

# 33RD EUROPEAN SYMPOSIUM ON COMPUTER AIDED PROCESS ENGINEERING

ESCAPE-33

Edited by  
ANTONIS C. KOKOSSIS  
MICHAEL C. GEORGIADIS  
STRATOS PISTIKOPOULOS



COMPUTER-AIDED CHEMICAL ENGINEERING, 52

33<sup>rd</sup> EUROPEAN SYMPOSIUM ON  
COMPUTER AIDED PROCESS  
ENGINEERING

VOLUME 1



**COMPUTER-AIDED CHEMICAL ENGINEERING, 52**

**33<sup>rd</sup> EUROPEAN SYMPOSIUM ON  
COMPUTER AIDED PROCESS  
ENGINEERING**

**VOLUME 1**

*Edited by*

**Prof. Antonios C. Kokossis**  
*National Technical University of Athens  
School of Chemical Engineering  
Athens, Greece*

**Prof. Michael C. Georgiadis**  
*Aristotle University of Thessaloniki  
School of Engineering  
Department of Chemical Engineering  
Greece*

**Prof. Efstratios Pistikopoulos**  
*Chemical Engineering  
Texas A&M Energy Institute  
TX, USA*



**ELSEVIER**

**Amsterdam – Boston – Heidelberg – London – New York – Oxford  
Paris – San Diego – San Francisco – Singapore – Sydney – Tokyo**

Elsevier

Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands  
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK  
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, USA

Copyright © 2023 Elsevier B.V. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: [www.elsevier.com/permissions](http://www.elsevier.com/permissions).

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

### Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

### British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

### Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

ISBN (Volume 1): 978-0-443-23553-5

ISBN (Set) : 978-0-443-15274-0

ISSN: 1570-7946

For information on all Elsevier publications visit our website at <https://www.elsevier.com/>



Working together  
to grow libraries in  
developing countries

[www.elsevier.com](http://www.elsevier.com) • [www.bookaid.org](http://www.bookaid.org)

*Publisher:* Candice Janco

*Acquisition Editor:* Anita Koch

*Editorial Project Manager:* Lena Sparks

*Production Project Manager:* Paul Prasad Chandramohan

*Designer:* Greg Harris

Typeset by STRAIVE

# Contents

<b>Preface</b>	<b>xv</b>
<b>LOCAL ORGANIZING COMMITTEE</b>	<b>xvii</b>
<b>T1: Modelling and optimization for multi-scale integration</b>	
<b>1. An exploratory model-based design of experiments technique to aid parameters identification and reduce prediction uncertainty</b> <i>Francesca Cenci, Arun Pankajakshan, Solomon Gajere Bawa, Asterios Gavriilidis, Pierantonio Facco, Federico Galvanin</i>	<b>1</b>
<b>2. GHG emission reduction via multiple fuel options in carbon-constrained industrial clusters</b> <i>Christelle Bechara, Sabla Y. Alnouri</i>	<b>7</b>
<b>3. An optimization-free Fisher information driven approach for online design of experiment</b> <i>Andrea Friso, Federico Galvanin</i>	<b>13</b>
<b>4. A Python-based approach for thermodynamic consistency tests of binary VLE data</b> <i>Pedro Cezareth, Martina Costa Reis</i>	<b>19</b>
<b>5. Modelling building's life-cycle decarbonization in China: A multi-level and multi-region optimization approach</b> <i>Chenxi Li, Pei Liu, Zheng Li</i>	<b>25</b>
<b>6. Automated Kinetic Model Discovery – A Methodological Framework</b> <i>Miguel Ángel de Carvalho Servia, Ilya Orson Sandoval, Dongda Zhang, Klaus Hellgardt, King Kuok (Mimi) Hii, Ehecatl Antonio del Rio Chanona</i>	<b>33</b>
<b>7. Assessing process systems models for pharmaceutical development</b> <i>Margherita Geremia, Samir Diab, Charalampos Christodoulou, Gabriele Bano, Massimiliano Barolo, Fabrizio Bezzo</i>	<b>39</b>
<b>8. Simulation and optimisation of a medium scale industrial reverse osmosis desalination system</b> <i>Mudhar A. Al-Obaidi, Alanood A. Alsarayreh, and Iqbal M. Mujtaba</i>	<b>45</b>

<b>9. Modeling strategies in multi-scale food-energy-water nexus system optimization</b>	
<i>Marcello Di Martino, Patrick Linke, Efstratios N. Pistikopoulos</i>	<b>51</b>
<b>10. MULTI-FIDELITY SURROGATE MODELING FOR CHEMICAL PROCESSES WITH PHYSICS-INFORMED NEURAL NETWORKS</b>	
<i>Yu-Ting Liu, Chuan-Yu Wu, Tao Chen, Yuan Yao</i>	<b>57</b>
<b>11. A CFD-model-based approach to continuous freezing process design for human induced pluripotent stem cells</b>	
<i>Yusuke Hayashi, Benedikt X. Scholz, Hirokazu Sugiyama</i>	<b>65</b>
<b>12. Multiscale Modeling of Spatial Area-Selective Thermal Atomic Layer Deposition</b>	
<i>Matthew Tom, Sungil Yun, Henrik Wang, Feiyang Ou, Gerassimos Orkoulas, Panagiotis D. Christofides</i>	<b>71</b>
<b>13. Modeling the shrinkage effect during the drying process of a green ceramic material</b>	
<i>Achilleas L. Arvanitidis, Margaritis Kostoglou, Michael C. Georgiadis</i>	<b>77</b>
<b>14. Investigating physics-informed neural networks for bioprocess hybrid model construction</b>	
<i>Alexander William Rogers, Ilya Orson Sandoval Cardenas, Ehecatl Antonio Del Rio-Chanona, Dongda Zhang</i>	<b>83</b>
<b>15. Decision Making Approaches to Improve Resilience in Food Supply Chains and Enhance Food Security Against Climate Change Risks</b>	
<i>Bashar Hassna, Sarah Namany, Mohammad Alherbawi, Adel Elomri, Tareq Al-Ansari</i>	<b>89</b>
<b>16. Lot-sizing and Production Scheduling of a Beverage Industry</b>	
<i>Maria E. Samouilidou, Eleonora Diakoumi, Georgios P. Georgiadis, Antonios Dikaiakos, Michael C. Georgiadis</i>	<b>95</b>
<b>17. Modeling and Analysis of Clinical &amp; Municipal Waste Incineration Process using Aspen Plus</b>	
<i>Mohamad Rizza Othman, Sivanesh Kumar Anpalagam, Nur Fitriyanni Jafary</i>	<b>101</b>
<b>18. On the minimization of cycle time in periodic production scheduling</b>	
<i>Alexandros Koulouris, Georgios P. Georgiadis</i>	<b>107</b>

- 19. Non-Pareto optimal solutions as enablers for versatile heat exchanger networks**  
*David Huber, Felix Birkelbach, René Hofmann* **113**
- 20. Enhanced Automatic Initialization Method for Solving Large Nonlinear Algebraic Systems**  
*Saskia Bublitz, Erik Esche, Jens-Uwe Repke* **119**
- 21. Electroplating scheduling: Closing a research gap from an automation vendor's perspective**  
*Sophie Reimschuessel, Uwe Fuchs, Guido Sand* **125**
- 22. Real time prediction of ozone pollution using data-enabled deep learning modeling**  
*Fugui Hong, Cheng Ji, Fangyuan Ma, Chang Chen, Wei Sun* **131**
- 23. Numerical Study of a Single-Spout and Z-shaped Fluidized Bed Biomass Gasifier**  
*Fatima Ez-zahra El Hamra, and Radouan Boukharfane* **139**
- 24. Intensification of ATJ Process Using Catalytic Distillation, Optimization Considering Control and Economic Issues**  
*Gabriel Contreras-Zarazúa<sup>1</sup>, Juan José Quiroz Ramírez, Eduardo Sánchez-Ramírez,<sup>b</sup> Juan Gabriel Segovia-Hernández* **145**
- 25. Efficient physical model building algorithm using equations extracted from documents**  
*Shota Kato and Manabu Kano* **151**
- 26. Developing a rigorous chemical reaction network reduction strategy for *n*-hexadecane hydroisomerisation**  
*Fernando Vega-Ramon, Wei Wang, Wei Wu, Dongda Zhang* **157**
- 27. Troubleshooting high-pressure issues in an industrial biorefinery process by feature-oriented modeling**  
*Elia Arnese-Feffin, Pierantonio Facco, Daniele Turati, Fabrizio Bezzo, Massimiliano Barolo* **163**
- 28. Optimal Determination of the Binary Interaction Parameters for the Vapor-Liquid-Liquid Equilibrium on the Vinyl Acetate/Acetic Acid/Water System**  
*Jose Alfredo Paredes-Ortiz, Úrsula Fabiola Rodríguez-Zúñiga, Fernando Israel Gómez-Castro* **169**
- 29. Assessment of subgrid-scale models for large-eddy simulation of a gas-liquid stirred reactor**  
*Safae Elmisaoui, Radouan Boukharfane, Lhachmi Khamar, and Jean-Michel Ghidaglia* **175**



- 30. Mechanistic model for inkjet printing and applications in perovskite solar cells engineering**  
*Naveen Bhati, Mohammad Khaja Nazeeruddin, and François Maréchal* **183**
- 31. Modeling and Optimization of Low-carbon Oriented Urban Energy Systems Spatial Planning Considering Energy Networks Topology**  
*Ruiyu Zhang, Pei Liu, Zheng Li* **189**
- 32. Multi-objective optimisation for microbrewery retrofitting**  
*Ruosi Zhang, Melis Duyar, Lesley Foulkes, James Berrow, Michael Short* **197**
- 33. Thermodynamic equilibrium modelling for the optimal performance of a wood biomass downdraft gasifier for hydrogen production**  
*Anamika Kushwah, Tomas Ramirez Reina, Michael Short* **203**
- 34. A model predictive control approach for recirculating aquaculture systems integrated with sustainable hybrid energy systems**  
*Ruosi Zhang, Tao Chen, Michael Short* **209**
- 35. Multi-stage optimization for marketing industrial flexibility**  
*Martin Fischer, Karl-Wilhelm Schenzel, René Hofmann* **215**
- 36. Intelligent Process Flowsheet Synthesis and Design using Extended SFILES Representation**  
*Vipul Mann, Rafiqul Gani, Venkat Venkatasubramanian* **221**
- 37. Physics-informed Neural Network based Modeling of an Industrial Wastewater Treatment Unit**  
*Tuse Asray, Ece Serenat Koksul, Elif Ecem Esenboga, Ahmet Cosgun, Gizem Kusoglu, Duygu Aydin, Erdal Aydin* **227**
- 38. Approaches to Reduce Optimization Time for Stochastic Optimization of Complex Chemical Processes**  
*Fanyi Duanmu, Dian Ning Chia, Eva Sorensen* **235**
- 39. Design of a CO<sub>2</sub> capture plant: A sustainable approach using deep eutectic solvents**  
*Adrián Martínez-Lomovskoi, Ana-Gabriela Romero-García, Eduardo Sánchez-Ramírez, Juan-Gabriel Segovia-Hernández, Cataldo De Blasio* **241**

- 40. Optimal Retrofitting of Conventional Oil Refinery into Sustainable Bio-refinery under Uncertainty**  
*Lifeng Zhang, Ana Inés Torres, Bingzhen Chen, Zhihong Yuan, Ignacio E. Grossmann* 247
- 41. A Combined D-optimal and Estimability Model-Based Design of Experiments of a Batch Cooling Crystallization Process**  
*Xuming Yuan, Brahim Benyahia* 255
- 42. A novel mathematical formulation for short-term scheduling of multipurpose batch plants in chemical manufacturing**  
*Dan Li, Taicheng Zheng, Jie Li* 261
- 43. Heat integration and heat exchanger network design for oxyfuel cement plants**  
*Leif E. Andersson, Avinash Subramanian, Mari Voldsund, Rahul Anantharaman, Kristina Fleiger, Francisco Carrasco, Mirko Weber* 267
- 44. Technological Portfolio Optimisation for de-risking within the energy, water and food nexus**  
*Maryam Haji, Mohammad Alherbawi, Sarah Namany, Tareq Al-Ansari* 273
- 45. Prediction of Melt Flow Rate (MFR) in polymer production, with auto-regressive Machine Learning algorithms, for multivariate time-series measured in irregular timestamps**  
*Symon Doe, Christoforos Kassianides, Symeon Kassianides, Christos Christodoulou, Nikos Bakas* 279
- 46. Achieving Operational Excellence by Combining Material Tracking and On-line Polymer Analysis and Classification data in an all-in-one Integrated Decision Support System**  
*Symeon Kassianides, Symon Doe, Sanja Micunovica, Artemis Theodoropoulou, Charis Komodromos, Sofia Louloudi* 285
- 47. Markov Chain Monte Carlo simulation-based optimization for a production scale milk drying process**  
*Adrián Ferrari, Soledad Gutiérrez, Gürkan Sin* 291
- 48. Polymer Grade Transition Control via Reinforcement Learning Trained with a Physically Consistent Memory Sequence-to-Sequence Digital Twin**  
*Zhen-Feng Jiang, David Shan-Hill Wong, Jia-Lin Kang, Yuan Yao, Yao-Chen Chuang* 297

- 49. A Rolling Horizon Approach for Ethylene Production Scheduling with Daily Inventory Constraints**  
*Haoran Li, Tong Qiu* **305**
- 50. Simulation and optimization of fixed bed and moving bed TSA processes for post-combustion CO<sub>2</sub> capture**  
*Shreenath Krishnamurthy, Richard Blom, Giorgia Mondino, Schalk Cloete, John Morud, Abdelghafour Zaabout, Philip Llewelyn, Cecile Pereira, and Veronique Pugnet* **313**
- 51. AI-enhanced BoM Optimization with MILP for the EMS Industry**  
*Hossein Mostafaei, Eero Hiltunen, Antti Liski* **319**
- 52. Reduced cleaning model for highly viscous non-Newtonian fluids in pipelines**  
*Vera Liebmann, Matti Heide, Hannes Köhler, Christian Golla, Jochen Fröhlich, Frank Rüdiger* **325**
- 53. Optimal Retrofit of Simple Distillation Sequences to Thermally Coupled Side-Stream Configurations**  
*Anna S. Horsch, Andres R. Acevedo, Mirko Skiborowski* **331**
- 54. Computer aided-design of energy efficiency tools on a microalgae biorefinery scheme**  
*Araceli Guadalupe Romero-Izquierdo, Claudia Gutiérrez-Antonio, Fernando Israel Gómez-Castro, Salvador Hernández, Juan Fernando García-Trejo* **337**
- 55. Optimization of Simulated Moving Bed Chromatographic Processes using Surrogate Models**  
*Rojiar Pishkari, Marcus Fechtner, Tobias Keßler, Achim Kienle* **343**
- 56. Multi-Objective Model-based Design of Experiments of Pharmaceutical Tableting Process**  
*Ilias Bouchkira, Brahim Benyahia* **349**
- 57. Water bodies restoration as an optimal control problem with integrated internal and external strategies. The inclusion of artificial floating islands**  
*Amira Siniscalchi, Laura J. Fritz, Vanina Estrada, Maria Soledad Diaz* **355**
- 58. Developing robust hybrid-models**  
*Peter Jul-Rasmussen, Xiaodong Liang, Xiangping Zhang, Jakob Kjøbsted Huusom* **361**

- 59. A gPC-based Global Sensitivity Analysis for Phosphate Slurry Flow in Pipelines**  
*Marwane Elkarii, Radouan Boukharfane, Saad Benjelloun, Chakib Bouallou, Nabil Elmocayd* **367**
- 60. Hybrid dynamic model of monoclonal antibody production using CHO cells**  
*Mariana Monteiro, Cleo Kontoravdi* **375**
- 61. Combined Bayesian optimization and global sensitivity analysis for the optimization of simulation-based pharmaceutical processes**  
*Niki Triantafyllou, Nilay Shah, Maria M. Papathanasiou, Cleo Kontoravdi* **381**
- 62. Competitive adsorption of copper, nickel, and chromium ions onto EDTA modified SBA-15**  
*Bawornpong Pornchuti, Yuttana Phoochahan, Prarana Padma, Suchada Ruengrit, Pravit Singtothong* **387**
- 63. Auto-extractive reactive distillation for the production of fatty esters**  
*Hernán D. Muñoz, Christian Hoffmann, Markus Illner, Álvaro Orjuela, Jens-Uwe Repke* **393**
- 64. Long-term Capacity Management of Pharmaceutical Manufacturing Networks**  
*Simon B. Lindahl, Deenesh K. Babi, Gürkan Sin* **399**
- 65. Modeling of a trickle bed reactor: Study of arabinose oxidation on gold catalyst extrudates**  
*Mouad Hachhach, Vincenzo Russo, Dmitry Yu. Murzin, Tapio Salmi* **405**
- 66. Numerical simulation of flow in porous media using the SPH method and Ergun force**  
*Carlos E. Alvarado-Rodriguez, Lamberto Diaz-Damacillo, Eric Plaza* **411**
- 67. Impact of structural change in electricity and hydrogen end use on low-carbon transition of an energy system: A case study of China**  
*Nuobei Zhang, Pei Liu, Zheng Li* **417**
- 68. Optimal planning for thermal power plants with CCS under source–sink matching constraints: A case study of China**  
*Binglin Du, a Pei Liu, Zheng Li* **425**

- 69. On discrete time chemical production scheduling MILP models containing record keeping variables**  
*Amin Samadi, Nathan Adelgren, Christos T. Maravelias* **433**
- 70. ESTAN – A toolbox for global sensitivity based estimability analysis**  
*Ilias Bouchkira, Abderrazak M. Latifi, Brahim Benyahia* **439**
- 71. Multi-technology separation network synthesis**  
*Garry S.P. Taifan, Christos T. Maravelias* **445**
- 72. A Reinforcement Learning Development for The Exact Guillotine with Flexibility on Cutting Stock Problem**  
*Jie-Ying Su, Chia-Hsiang Liu, Cian-Shan Syu, Jia-Lin Kang, Shi-Shang Jang* **451**
- 73. Real-time optimization of a chemical plant with continuous flow reactors via reinforcement learning**  
*Min Wu, Furkan Elmaz, Ulderico Di Caprio, Dries De Clercq, Siegfried Mercelis, Peter Hellinckx, Leen Braeken, Florence Vermeire and M. Enis Leblebici* **457**
- 74. From problem specification to converged flowsheet simulations: Integrated workflow in an industrial context**  
*Mirko Skiborowski, Norbert Asprion, Sergej Blagov, Michael Bortz, Dennis Manuel Heim, Kai Fabian Kruber, Thulasi Rolland, Tobias Seidel* **463**
- 75. A Hybrid Modelling Framework for Dynamic Modelling of Bioprocesses**  
*Haiting Wang, Cleo Kontoravdi, Ehecatl Antonio del Rio Chanona* **469**
- 76. Simultaneous selection and optimization of reaction pathways and downstream separation network**  
*Shuang Xu, Selen Cremaschi, Mario R. Eden and Anjan K. Tula* **475**
- 77. Workflow for adaptation, analysis and application of mechanistic models for experimental planning of protein refolding processes**  
*Jan Niklas Pauk, Chika Linda Igwe, Christoph Herwig* **481**
- 78. Mechanistic modeling of product formation in recombinant *Escherichia coli* cultivations**  
*Don Fabian Müller, Philipp Pably, Daniel Wibbing, Julian Kager and Christoph Herwig* **487**

- 79. Physics-informed neural networks for optimization of polymer reactor design**  
*Yubin Ryu, Sunkyu Shin, Jay J. Liu, Wonbo Lee, Jonggeol Na* **493**
- 80. Extending the SI Decomposition to Continuous-Time Two-Stage Scheduling Problems**  
*Daniel Montes, José Luis Pitarch, César de Prada* **499**
- 81. Modelling of water vapor adsorption/desorption on hopcalite in an industrial medical air dryer**  
*Cristian Cardenas, & David Farrusseng* **505**
- 82. Multi-fidelity Bayesian Optimisation of Reactor Simulations using Deep Gaussian Processes**  
*Tom Savage, Nausheen Basha, Omar K. Matar, Antonio del Rio Chanona* **511**
- 83. Design of a sustainable biodiesel production process by a multi-objective optimization**  
*Barcia-Quimi A.F., Risco-Bravo A., Alcivar-Espinoza K., Tinoco-Caicedo D.L* **519**
- 84. Modeling of an Industrial Delayed Coker Unit**  
*Anne Firstauthor, Tim B. Secondauthor, James Q. Thirdauthora* **525**
- 85. A sustainable energy supply system in smart cities: an integrated natural gas and electricity network operation**  
*Ali Ahmadian, Ali Almansoori, Ali Elkamel* **533**
- 86. A hybrid forecasting framework for the energy sector**  
*Stefanos G. Baratsas, Funda Iseri, Efstration N. Pistikopoulos* **539**
- 87. A Chance-Constrained Programming Approach to Optimal Planning of Low-carbon Transition of a Natural Gas Supply System: A Case Study of China**  
*Jiaqi Zhang, Pei Liu, Zheng Li* **547**
- 88. Efficient 1D modelling of hot melt extrusion process for pharmaceutical applications**  
*Michaela Vasilaki, Umair Zafar, Ioannis .S. Fragkopoulos, Ioannis K. Kookos* **553**
- 89. Towards Machine Learning of Power-2-Methanol Processes**  
*Carl Julius Martensen, Christoph Plate, Tobias Kefler, Christian Kunde, Lothar Kaps, Achim Kienle, Andreas Seidel-Morgenstern, Sebastian Sager* **561**

- 90. Water resource portfolio design for optimum water productivity in agriculture**  
*Fatima-Zahra Lahlou, Rajesh Govindan, Bilal Hazrat, Hamish R. Mackey, Tareq Al-Ansari* **569**
- 91. S-GNN: State-Dependent Graph Neural Networks for Functional Molecular Properties**  
*Adem R.N. Aouichaoui, Alessandro Cogliati, Jens Abildskov, Gürkan Sin* **575**
- 92. Switching Multi-Objective Dynamic Optimization (MODO) for the Production of Value-Added Products**  
*Juan C. Acosta-Pavas, Carlos. E Robles-Rodríguez, Jérôme Morchain, David Camilo Corrales, Claire Dumas, Arnaud Cockx, César A. Aceves-Lara.* **583**
- 93. Dynamic modeling of a waste incinerator furnace in view of supervision**  
*Lionel Sergent, François Lesage, Abderrazak M. Latifi, Jean-Pierre Corriou, Alexandre Grizeau* **589**
- 94. Reinforcement learning combined with model predictive control to optimally operate a flash separation unit**  
*Dean Brandner, Torben Talis, Erik Esche, Jens-Uwe Repke, Sergio Lucia* **595**
- 95. Architecture and Design of a Modern Commercial Process Simulator**  
*Ian Boys, Jochen Steimel* **601**
- 96. High and Robust Fault Detection via Polynomial Approximated Isomap Embeddings**  
*Burak Alakent* **607**
- 97. A comparative study of distributed feedback-optimizing control strategies**  
*Vegard Aas, Risvan Dirza, Dinesh Krishnamoorthy, Sigurd Skogestad* **613**
- 98. Superstructure configuration and optimization of the biofuel production perspectives in Austria**  
*Safdar Abbas, Ada Josefina Robinson, Stavros Papadokonstantakis* **619**
- 99. Optimal operation of an industrial low pressure carburizing furnace**  
*Fatima Matamoros, Pierre-Alexandre Glaude, Roda Bounaceur, Hubert Monnier, Abderrazak Latifi* **625**

<b>100. Development of a holistic Python package for optimal selection of experimental design criteria in kinetic model discrimination</b> <i>Maerthe Theresa Tillmann, Federico Galvanin</i>	<b>631</b>
<b>101. A Model-based Computational Approach to Optimal Design of Nose-to-Brain Drug Delivery Systems</b> <i>Athina Vasileiadou, Costas Kiparissides</i>	<b>637</b>
<b>102. Global estimability analysis and model-based design of experiments in surrogate modeling</b> <i>Ilias Bouchkira, Sanae Elmisaoui, Saad Benjelloun, Abderrazak M. Latifi</i>	<b>643</b>
<b>103. Tractable Data-driven Solutions to Hierarchical Planning-scheduling-control</b> <i>Damien van de Berg, Roberto Xavier Jimenez Jimbo, Nilay Shah, Ehecatl Antonio del Rio-Chanona</i>	<b>649</b>
<b>104. A graph classification approach to determine when to decompose optimization problems</b> <i>Ilias Mitrai, Prodromos Daoutidis</i>	<b>655</b>
<b>105. An Enhanced Group-Interaction Contribution Method for the Prediction of Freezing Point and Enthalpy of Fusion for Pure Organic Compounds</b> <i>Khaled Mansour, Mourad Korichi</i>	<b>661</b>
<b>106. Single and Multi-Objective Superstructure Optimization of an Integrated Continuous Multistage Reaction-Crystallization-Filtration Process with Recycles</b> <i>Jiaxu Liu, Brahim Benyahia</i>	<b>667</b>
<b>107. Prediction of product distribution using machine learning techniques</b> <i>Praveen Kumar Bommineni, Manohar Kakunuri, Sarat Babu Anne</i>	<b>673</b>
<b>108. Architectural Design of Chemical and Biological Pathways through Parameter Sensitivity Oriented Mixed Integer Formulations</b> <i>Emir Topac, Emrullah Erturk, Ozgun Deliismail, Sahin Sarrafi, Hasan Sildir</i>	<b>679</b>
<b>109. Data-driven fouling modeling in crude oil heat exchanger</b> <i>Pelin Dologlu, Alihan Tabak, A. Eren Vedin, Gizem Kayar</i>	<b>685</b>



- 110. Explainable formation energy prediction for uncovering the relationship between the electronic structure and stability of the heterogeneous catalyst**  
*Daeun Shin, Dong Hyeon Mok, Seoin Back, Jonggeol Na* **691**
- 111. Modular stochastic optimization for optimal rainwater harvesting system design**  
*Qiao Yan Soh, Edward O'Dwyer, Salvador Acha, Nilay Shah* **697**
- 112. Knowledge-based model and simulations to support decision making in wastewater treatment processes**  
*Claire Valentin, Frédéric Lagoutière, Jean-Marc Choubert, Françoise Couenne, Christian Jallut* **703**
- 113. Homotopic Method for binodal curves computation in ternary liquid-liquid separation**  
*Nataliya Shcherbakova, Kevin Roger, Joseph Gergaud, Olivier Cots* **709**
- 114. Modelling of iron oxide reduction with hydrogen**  
*Emiliano Salucci, Snigdha Ghosh, Vincenzo Russo, Henrik Grénman, Henrik Saxén* **715**
- 115. Parameter estimation approach applied to microalgae-bacteria photobioreactor**  
*Irina Bausa-Ortiz, Raul Muñoz, Smaranda P. Cristea, Cesar de Prada* **721**
- 116. An Event-Based Continuous-Time MILP for Short-Term Crude Oil Scheduling**  
*İrem Marttin, Sena Kurban, Gizem Kuşoğlu Kaya, Özgür Kabak, Y. İlker Topcu* **727**
- 117. Modeling aqueous multi-electrolyte systems for the simulation and optimization of concentrated brine processes**  
*Wajeha Tauqir, Adaeze C. Maduako, Xi Yang, George M. Bollas* **733**
- 118. Stable two-stage scenario generation via game-theoretic optimisation**  
*Georgios L. Bounitsis, Lazaros G. Papageorgiou, Vassilis M. Charitopoulos* **739**
- 119. Global One-dimensional Root-finding Algorithm for Derivable and Continuous Functions**  
*Andrea Galeazzi, Kristiano Prifti, Flavio Manenti* **745**

- 120. A Demethanizer column Digital twin with non-conventional LSTM neural networks arrangement**  
*Marta Mandis, Roberto Baratti, Jorge Chebeir, Stefania Tronci, José A. Romagnoli* **751**
- 121. Soft modelling of spruce conversion into bio-oil through pyrolysis – Note I: steam explosion and LPMO-activated enzymatic saccharification**  
*Matteo Gilardi, Filippo Bisotti, Olaf T. Berglihn, Roman Tschentscher, Vincent G.H. Eijssink, Anikó Várnai, Bernd Wittgens* **757**
- 122. Molecular Modelling and Optimization of Diesel Hydrotreating Processes**  
*Xiaolin Bi, Nan Zhang, Robin Smith* **763**
- 123. Soft modelling of spruce conversion into bio-oil through pyrolysis – Note II: pyrolysis**  
*Filippo Bisotti, Matteo Gilardi, Olaf T. Berglihn, Roman Tschentscher, Vincent G.H. Eijssink, Anikó Várnai, Bernd Wittgens* **769**
- 124. Numerical Modeling of Gas-liquid Hydrodynamics in a U-loop Bioreactor**  
*Johan Le Nepvou De Carfort, Tiago Pinto, Ulrich Krühne* **775**
- 125. Effect of vaporisation models on the FCC riser modelling**  
*Thabang W. Selalame, Raj Patel, Iqbal M. Mujtaba, Yakubu M. John* **783**
- 126. A joint model-based design of experiments approach for the identification of Kriging models in geological exploration**  
*Philipp Deussen, Federico Galvanin* **789**
- 127. Reinforcement Learning for inventory management in multi-echelon supply chains**  
*Guoquan Wu, Miguel Ángel de Carvalho Servia, Max Mowbray* **795**
- 128. Monte Carlo Simulation of Photo Induced Atom-Transfer Radical Polymerization for Microscopic Properties**  
*Rui Liu, Xiaowen Lin, Antonios Armaou, Xi Chen* **801**
- 129. Identification of a dynamic model of the N-vinylcaprolactam-co-glycidyl methacrylate microgel synthesis**  
*Luisse F. Kaven, Thomas Nevolianis, Nadja Wolter, Adel Mhamdi, Andrij Pich, Kai Leonhard, Alexander Mitsos* **807**

- 130. A Machine Learning-assisted Hybrid Model to Predict Ribbon Solid Fraction, Granule Size Distribution and Throughput in a Dry Granulation Process**  
*Yan-Shu Huang, David Sixon, Phoebe Bailey, Rexonni B. Lagare, Marcial Gonzalez, Zoltan K. Nagy, Gintaras V. Reklaitis* **813**
- 131. A Data-driven Modeling Approach for Water Flow Dynamics in Soil**  
*Zeyuan Song, Zheyu Jiang* **819**
- 132. Numerical Simulation of No-Newtonian Fluids in Stirred-Tank with proximity impellers**  
*Alvarado-Rodriguez C. E., Martinez-Herrera G., Cortez-Gonzalez J., Murrieta-Dueñas R* **825**
- 133. Hybrid Dynamic Surrogate Modelling for a Once-Through Steam Generator**  
*Sindre Stenen Blakseth, Leif Erik Andersson, Rubén Mocholí Montañés, Marit Jagtøyen Mazzetti* **831**
- 134. PINN-based Design of Experiment Concept for Process Model Parameter Identification**  
*Aike Aline Tappe, Subiksha Selvarajan, Caroline Heiduk, Stephan Scholl and René Schenkendorf* **837**
- 135. Performance analysis of virtual platforms focused on multi-objective process optimization**  
*Jaime David Ponce-Rocha, César Ramírez-Márquez, Ricardo Morales-Rodriguez* **843**
- 136. Mathematical Modeling of Anion Exchange Membrane Water Electrolyzer**  
*Donggyun Lee, Minsu Kim, Il Moon* **849**
- 137. An Aided Optimization Method for Molecular Reconstruction of Crude Oil Based on Pre-optimized Dataset**  
*Chen Zhang, Tong Qiu* **855**
- 138. A Global Optimization Algorithm for the Solution of Mixed-Integer Quadratic Adjustable Robust Optimization Problems under Endogenous Uncertainty**  
*Byungjun Lee, Styliani Avraamidou* **861**
- 139. Cross-domain Fault Diagnosis for Chemical Processes through Dynamic Adversarial Adaptation Network**  
*Ruoshi Qin, Jinsong Zhao* **867**

- 140. Design and performance analysis of energy efficient hydrogen liquefaction process**  
*Ahmad Naquash, Moonyong Lee* **875**
- 141. Optimal design of a microbubble cap to enhance mass transfer efficiency using image processing**  
*Yup Y., Jonghun L., Chonghyo J., Seongbin G., Sunghyun C., Jaewon L., Junghwan K., Hyungtae C.* **881**
- 142. Design of a heterogeneous azeotropic distillation processes using stochastic optimization**  
*Josué J. Herrera Velázquez, Julián Cabrera Ruiz, J. Rafael Alcántara Avila* **889**
- 143. Single and multi-objective optimization of heat integrated distillation columns using PYMOO**  
*Josué J. Herrera Velázquez, Julián Cabrera Ruiz, J. Rafael Alcántara Avila, Salvador Hernández Castro* **897**
- 144. Design of multi-component gradient SMBs**  
*Jesper Frandsen, Jakob Kjøbsted Huusom, Krist V. Gernaey, Jens Abildskov* **903**
- 145. A Method for Finding a Design Space as Linear Combinations of Parameter Ranges for Biopharmaceutical Development**  
*Thomas Oberleitner, Thomas Zahel, Christoph Herwig* **909**
- 146. Optimal operation of large gas networks: MILP model and decomposition algorithm**  
*Lavinia Marina Paola Ghilardi, Francesco Casella, Daniele Barbatì, Roberto Palazzo, Emanuele Martelli* **915**
- 147. An Integrated Framework for the Geometric Optimization of Bio-catalytic Microreactor**  
*Yueheng Han, Runze Liang, Zhihong Yuan* **921**
- 148. A genetic algorithm-active learning approach in energy systems optimization**  
*Rafael Castro Amoedo, Julia Granacher, Yi Zhao, and François Maréchal* **929**
- 149. Data-driven Distributionally Robust Joint Chance-Constrained Optimization for Industrial Utility systems under Uncertainty**  
*Hanxiu Li, Liang Zhao* **935**
- 150. Online real-time multi-parameter optimization solution based on parallel EGO algorithm**  
*Xuerong Gu, Siyu Yang* **943**

- 151. Optimal design of Water Treatment Networks: effluent and piping disaggregated modelling**  
*Francisco J. G. Patrocínio, Hugo M. D. Carabineiro, Henrique A. Matos and Nuno M. C. Oliveira* **949**
- 152. Evaluation of the potential of a deep eutectic solvent for liquid-liquid extraction of furfural using optimization-based process design**  
*Kai Fabian Kruber, Mariann Kroll, Christoph Held, Mirko Skiborowski* **955**
- 153. Profit Allocation in Industrial Symbiosis Networks: Utility Exchanges**  
*Fabian Lechtenberg, Antonio Espuña, Moisès Graells* **961**
- 154. Operational optimisation of steam power system under uncertainty using time series prediction optimization model with Markov chain**  
*Kenian Shi, Jingyuan Zheng, Siyu Yang* **967**
- 155. An approach for modelling simultaneous fluid-phase and chemical-reaction equilibria in multicomponent systems via Lagrangian duality: The Reactive HELD algorithm**  
*Felipe A. Perdomo, George Jackson, Amparo Galindo, Claire S. Adjiman* **973**
- 156. Optimization of the Biomass Supply Chain for Power Generation with the Steam Rankine Cycle**  
*Monika Dokl, Zdravko Kravanja, Lidija Čuček* **979**
- 157. Hybrid modelling and data-driven parameterization of monoclonal antibody cultivation processes: Shifts in cell metabolic behavior**  
*Sara Badr, Kota Oishi, Kozue Okamura, Sei Murakami, Hirokazu Sugiyama* **985**
- 158. Optimal Design of Anion-pillared Metal-organic Frameworks for Gas Separation**  
*Xiang Zhang, Kai Sundmacher* **991**
- 159. Probabilistic neural networks for mechanical properties and tensile behavior prediction of polymer composites considering uncertainty in tensile testing**  
*Jinkyung Son, Tim B. Jaewook Lee, Dongil Shin* **997**

<b>160 Data-Driven Indication of Flooding in an Industrial Debutanizer Column</b> <i>Martin Mojto, Karol Ľubušký, Miroslav Fikar, Radoslav Paulen</i>	<b>1001</b>
<b>161. Can we deploy carbon capture, utilization and storage at a scale?</b> <i>Ahmed Alhajaj</i>	<b>1007</b>
<b>162. Prediction of capacitance using artificial neural networks for carbon nanofiber-based supercapacitors</b> <i>Kiran Donthula, Naresh Thota, Sarath Babu Anne, Manohar Kakunuri</i>	<b>1013</b>
<b>163. Model size determination using data analysis tools in the context of reaction network identification capabilities</b> <i>Carolina S. Vertis, Belmiro P.M. Duarte, Nuno M.C. Oliveira</i>	<b>1019</b>
<b>164. Modeling of particle formation in pan granulators with sieve-mill recycle</b> <i>Robert Dürr, Mateusz Przywara, Eric Otto, Dorota Antos, Achim Kienle</i>	<b>1025</b>
<b>165. Synthesis and optimization of NGL separation as a complex energy-integrated distillation sequence</b> <i>Qing Li, Adrian J. Finn, Stephen J. Doyle, Robin Smith, Anton A. Kiss</i>	<b>1031</b>
<b>166. Multi-scale model of solid oxide fuel cell: enabling microscopic solvers to predict physical features over a macroscopic domain</b> <i>Hamid Reza Abbasi, Masoud Babaei, Arash Rabbani</i>	<b>1037</b>
<b>167. Parameter estimation combined with model reduction techniques for identifiability analysis of biological models</b> <i>Michael Binns, Alessandro Usai, Constantinos Theodoropoulos</i>	<b>1045</b>
<b>168. Improving Model Robustness with Transfer Learning for Product Property Models</b> <i>Per Julian BECKER, Loic IAPTEFF, Benoit CELSE</i>	<b>1051</b>
<b>169. Modeling and simulation approach for modularized hydrogen electrolyzer plants</b> <i>Kumar Rajan Gopa, Isabell Viedt, Lucien Beisswenger, Leon Urbas</i>	<b>1057</b>
<b>170. Design and optimization of a shared heat exchanger network for an integrated rSOC system</b> <i>Xinyi Wei, Shivom Sharma, Francois Marechal, Jan Van Herle</i>	<b>1063</b>

- 171. Multi-objective optimisation of the algae *Arthrospira platensis* under uncertainty: a Pareto ellipsoids approach**  
*Wannes Mores, Satyajeet S. Bhonsale, Filip Logist, Jan F.M. Van Impe* **1069**
- 172. Predicting the Evolution of Flammable Gases During Li-ion Battery Thermal Runaway Using Micro-Kinetic Modelling**  
*Dr. Peter Bugryniec, Dr. Sergio Vernuccio, Prof. Solomon Brown* **1075**
- 173. Modeling of carbon offset networks for process systems to achieve net zero emissions**  
*Soo Hyoung Choi* **1081**
- 174. Identifying first-principles models for bubble column aeration using machine learning**  
*Peter Jul-Rasmussen, Arijit Chakraborty, Venkat Venkatasubramanian, Xiaodong Liang, Jakob Kjøbsted Huusom* **1087**
- 175. Optimization framework based on a sensitivity analysis for the identification of the critical design variables**  
*Maria de los A. Villarreal-de-Aquino, Jaime D. Ponce-Rocha, Eduardo S. Perez-Cisneros, Verónica Rodríguez-López, Edgar I. Murillo-Andrade, Divanery Rodriguez-Gomez, Ricardo Morales-Rodriguez* **1093**
- 176. Effect of coalescence models on the prediction of the separation of dispersed oil-water pipe flows**  
*Nikola Evripidou, Federico Galvanin, Panagiota Angeli* **1099**
- 177. Learning interpretable multi-output models: Kaizen Programming based symbolic regression for estimating outlet concentrations of a splitter**  
*Jimena Ferreira, Martín Pedemonte, Ana I. Torres* **1105**
- 178. Decision regression for modelling of supply chain resilience in interdependent networks: LNG case**  
*Adnan Al-Banna, Brenno C. Menezes, Mohammed Yaqot, Jeffrey D. Kelly* **1111**
- 179. Multi objective decision-making methodologies applied to GREENSCOPE sustainability indicators**  
*Ricardo N. Dias, Rui M. Filipe, Henrique A. Matos* **1117**
- 180. Process as a battery: Robust dynamic optimal operation of zeolite crystallization in a COBR with respect to the carbon footprint of electric power**  
*Robin Semrau, Sebastian Engell* **1127**

- 181. Machine Learning models to optimize dairy cow culling: the case of the experimental farm UniLaSalle Beauvais**  
*Abdifatah OUSELEYE, José RODRIGUEZ, Hamilton ARAUJO* **1133**
- 182. Analyzing the Large-Scale Supply of Low-Carbon Hydrogen in Germany**  
**Paul Sizaire, Emre Gençer** **1139**
- 183. Equilibrium sensitivity analysis of Carbon dioxide hydrogenation to methanol reaction system and comparison with conversions achievable in a simplified packed bed reactor model**  
*Milton D. Cárdenas, Camilo Rengifo, Manuel Figueredo* **1147**
- 184. Improving Transferable Force-Fields for Describing Crystal Structures Containing Hydrogen-Bonds**  
*Benjamin I. Tan, David H. Bowskill, Adam Keates, Constantinos C. Pantelides and Claire S. Adjiman* **1153**
- 185. A data-driven uncertainty modelling and reduction approach for energy optimisation problems**  
*Julien Vaes, Vassilis M. Charitopoulos* **1159**
- 186. Moisture and Throughput Control in an Integrated Pharmaceutical Purification Platform using PharmaPy**  
*Inyoung Hur, Daniel Casas-Orozco, Gintaras Reklaitis, Zoltan K. Nagy* **1167**
- 187. Parameter estimation and dynamic optimization of an industrial fed-batch reactor**  
*Jan G. Rittig, Jan C. Schulze, Lars Henrichfreise, Sebastian Recker, Rolf Feller, Alexander Mitsos, Adel Mhamdi* **1173**





# Preface

The 3 volumes contain papers presented at the 33<sup>rd</sup> European Symposium of Computer Aided Process Engineering (ESCAPE-33) held in Athens, Greece from 18-21 June 2023. The ESCAPE Series started in 1992 and builds on a strong foundation of 32 previous events of the European Federation of Chemical Engineers (EFCE) Working Party on Computer Aided Process Engineering (CAPE). Hosting countries to the conference have been Austria (1993, 2018), Ireland (1994), Slovenia (1995, 2016), Greece (1996, 2011), Norway (1997), Belgium (1998), Hungary (1999, 2014), Italy (2000, 2010, 2020), Denmark (1992, 2001, 2015), The Netherlands (2002, 2019), Finland (2003, 2013), Portugal (2004), Spain (2005, 2017), Germany (2016), Romania (2007), France (1992, 2008, 2022), Poland (2009), United Kingdom (2012), and Turkey (2021).

ESCAPE-33 addresses emerging and significant challenges in **Process Systems Engineering as a driver to rebuild industry with a systemic and holistic approach**. Contributions relate to the sustainable development of chemical processes, the development and the systematic evaluation of processing technologies and process innovations, and research promoting **circular economy** paradigms including social aspects and social engagement networks to incentivize and involve citizens. The conference recognizes the exceptional importance of **industrial biotechnology** as a driver and enabler for new chemistries, alongside **Artificial Intelligence and data engineering** as a technological challenge with an apparent impact on many systems technologies. Overall, ESCAPE 33 covered up-to-date topics in Process Systems Engineering and attracted a significant and diversified number of people from academia, research institutions, and industrial organizations worldwide. The themes of the conference included:

## **Modelling and optimization for multi-scale integration**

Coordinators: Antonio Espuña and Fani Boukouvala

## **Control, scheduling, and operability at the process and enterprise-level**

Coordinators: Ana Barbosa-Povoa and Chrysanthos Gounaris

## **Safe and sustainable products by design**

Coordinators: André Bardow and Fengqi You

## **Green and sustainable processes for the circular economy**

Coordinators: Seyed Mansouri and Bhavik Bakshi

## **Systems methods in industrial biotechnology and biomedical applications**

Coordinators: Joern Viell and Christos Maravelias

## **Multi-scale energy systems engineering (organized by the EFCE energy section)**

Coordinators: Francois Maréchal and Fabrizio Bezzo

**Sustainable supply chains and ecosystems**

Coordinators: Franjo Cecelja and Edwin Zondervan

**Education and knowledge transfer**

Coordinators: Ludovic Montastruc and Iqbal Mujtaba

The conference has been oversubscribed and we had to face a difficult task to review an excess of 800 submitted abstracts. The review process included reviews of abstracts, reviews of manuscripts, and the final selection of the revised manuscripts. Reviews involved 185 reviewers. They were completed primarily thanks to the generous support by members of the ESCAPE international scientific committee who contributed with evaluations, comments, and recommendations. The review process converged to 619 contributions from 57 countries. A total of 561 contributions by means of 6-page papers are included in these 3 volumes.

We hope that the books will serve as valuable reference documents to the scientific and industrial community and that they will contribute to the effective and innovative use of process systems engineering in the design, operation, and development of sustainable processes.

Antonis C. Kokossis  
Michael C. Georgiadis  
Stratos Pistikopoulos

## **LOCAL ORGANIZING COMMITTEE**

### Chair board

**Antonis Kokossis** (Chair) – National Technical University of Athens

**Michael Georgiadis** (Co-chair) – Aristotle University of Thessaloniki

**Ioannis Kookos** (Co-chair) – University of Patras

**Konstantina Kosmidou** (Treasurer) – National Technical University of Athens

**Polyxeni Lazaropoulou** (Secretary) – National Technical University of Athens

### Members

**Kostas Pyrgakis** – National Technical University of Athens

**Melina Psychas** – National Technical University of Athens

**Theodoros Damartzis** – Aristotle University of Thessaloniki

**Nikolaos A. Dangelakis** – Technical University of Crete

**Athanasios Papadopoulos** – Centre for Research & Technology Hellas (CERTH)

**Effie Marcoulaki** – National Centre for Scientific Research ‘Demokritos’

**Chrysoula Kappatou** – Imperial College London



# An exploratory model-based design of experiments technique to aid parameters identification and reduce prediction uncertainty

Francesca Cenci,<sup>a</sup> Arun Pankajakshan,<sup>b</sup> Solomon Gajere Bawa,<sup>b</sup> Asterios Gavriilidis,<sup>b</sup> Pierantonio Facco,<sup>a</sup> Federico Galvanin<sup>b</sup>

<sup>a</sup>CAPE-Lab – Computer-Aided Process Engineering Laboratory, Department of Industrial Engineering, University of Padova, via Marzolo 9, 35131 Padova, Italy

<sup>b</sup>Department of Chemical Engineering, University College London, Torrington Place, London WC1E 7JE, United Kingdom

## Abstract

When developing mathematical models to describe reaction processes, model parameters require to be estimated from experimental data. Experiments are traditionally designed through techniques aiming at space exploration, like space-filling methods (e.g., Latin Hypercube sampling or LHS), or at information maximization, like model-based design of experiments (MBD<sub>oE</sub>). However, the former methods do not minimize parameters uncertainty, while the latter do not ensure a minimization of model prediction uncertainty in the entire experimental design space. In this work, we propose a novel exploratory MBD<sub>oE</sub> (*eMBD<sub>oE</sub>*) approach based on G-optimality calculation (*G-map eMBD<sub>oE</sub>*) to simultaneously enhance space exploration and minimize model prediction variance. The method is tested on a case study related to the identification of kinetic parameters of catalytic total methane oxidation in a flow microreactor. Results show that the method is more explorative than conventional MBD<sub>oE</sub> and more efficient than LHS and MBD<sub>oE</sub> in reducing model prediction uncertainty and parameters uncertainty.

**Keywords:** model-based design of experiments; model prediction uncertainty; parameters identification; design space exploration; total methane oxidation

## 1. Introduction

For a model to be representative of the system under study, its parameters must be estimated from experimental data. However, not all experiments are equally informative for estimating model parameters and scarce information may lead to unsatisfactory parameters precision even with a considerable amount of data. The information of each possible experiment can be estimated through the Fisher information matrix (FIM; Fisher, 1950), which depends on the sensitivity of the model response with respect to every model parameter. This is used by model-based design of experiments (MBD<sub>oE</sub>; Asprey and Macchietto, 2000), which is an optimization problem where the objective function is a scalar measure of the FIM and the result is a set of values for the manipulated variables that ensure a maximization of the information. Maximum information leads to the minimization of parameters uncertainty. However, optimal experimental design techniques are inherently exploitative and may suffer from scarce space exploration. Moreover, there is no guarantee that the minimization of parameters uncertainty leads to a minimization of model prediction uncertainty in the whole design space. In this work, we propose a modification of the MBD<sub>oE</sub> optimization framework in order to enhance

space exploration and to minimize model prediction variance in the entire domain of model utilization. The proposed procedure is based on the G-optimality (Kiefer and Wolfowitz, 1959), which estimates model prediction variance starting from the FIM, thus it does not require the acquisition of additional data. Therefore, a mapping of G-optimality values (*G-maps*) can be built for the whole design space and MBDoE optimization is performed in the regions that satisfy a user-defined requirement on the G-optimality value. This explorative MBDoE procedure based on G-map (*G-map eMBDoE*) is applied to a case study on the identification of kinetic parameters of Mars-van Krevelen reaction mechanism for total oxidation of methane. Herein, an automated flow micropacked bed catalytic reactor platform is simulated to generate in silico data. Thanks to simulated experimental campaigns G-map eMBDoE is compared against a purely information-based MBDoE and a purely exploration-based technique, namely Latin Hypercube sampling (LHS; McKay et al., 1979).

## 2. Methods

The estimation of the  $N_\theta$  model parameters  $\hat{\boldsymbol{\theta}}$  is carried out through Maximum Likelihood Estimation (MLE; Bard, 1974). The reference techniques for design of experiments are: model-based design of experiments (MBDoE), which aims at maximizing the experiments information, and Latin Hypercube sampling (LHS), which aims at an explorative design. The former can be represented by the optimization problem:

$$\boldsymbol{\varphi}_{opt} = \arg \min_{\boldsymbol{\varphi}} \psi(\mathbf{H}_{\hat{\boldsymbol{\theta}}}) \quad (1)$$

where  $\boldsymbol{\varphi}$  is the design vector, i.e. the set of control variables that determine the experimental conditions, and  $\psi(\mathbf{H}_{\hat{\boldsymbol{\theta}}})$  is a scalar measure of the FIM  $\mathbf{H}_{\hat{\boldsymbol{\theta}}}$ . Consider a general differential and algebraic model  $\mathbf{f}(\dot{\mathbf{x}}, \mathbf{x}, \mathbf{u}, t, \boldsymbol{\theta}) = \mathbf{0}$ ,  $\mathbf{y} = \mathbf{h}(\mathbf{x})$ , with  $N_y$  measurable responses ( $\hat{\mathbf{y}}$ ),  $N_\theta$  parameters ( $\hat{\boldsymbol{\theta}}$ ),  $N_u$  control variables ( $\mathbf{u}$ ),  $N_x$ -dimensional vectors of state variables and their first derivatives ( $\mathbf{x}$  and  $\dot{\mathbf{x}}$ , respectively) that change over time ( $t$ ). Then, the  $N_\theta \times N_\theta$  FIM is defined as:

$$\mathbf{H}_{\hat{\boldsymbol{\theta}}}(\hat{\boldsymbol{\theta}}, \boldsymbol{\varphi}) = [\mathbf{V}_{\hat{\boldsymbol{\theta}}}^0]^{-1} + \sum_{i=1}^{n_{sp}} \left( \frac{d\hat{\mathbf{y}}}{d\hat{\boldsymbol{\theta}}} \right)_i^T \boldsymbol{\Sigma}_y^{-1} \left( \frac{d\hat{\mathbf{y}}}{d\hat{\boldsymbol{\theta}}} \right)_i, \quad \text{with: } \mathbf{H}_{\hat{\boldsymbol{\theta}}}(\hat{\boldsymbol{\theta}}, \boldsymbol{\varphi}) \cong [\mathbf{V}_{\hat{\boldsymbol{\theta}}}]^{-1} \quad (2)$$

Thus, the FIM depends on the prior information  $[\mathbf{V}_{\hat{\boldsymbol{\theta}}}^0]^{-1}$  (related to parameters bounds; Bard, 1974), on the  $N_y \times N_y$  response variance covariance matrix  $\boldsymbol{\Sigma}_y$  and on the  $N_y \times N_\theta$  matrix of sensitivities  $\left( \frac{d\hat{\mathbf{y}}}{d\hat{\boldsymbol{\theta}}} \right)_i$  calculated at the  $i$ -th sampling point ( $n_{sp}$  is the total number of sampling points, given by the sum of the  $N_{sp}$  sampling points for every response). As shown in Eq. (2), the parameters variance covariance matrix  $\mathbf{V}_{\hat{\boldsymbol{\theta}}}$  can be approximated as the FIM inverse: therefore, the higher the experiments information content, the smaller the parameters uncertainty region. In this work, we minimize the direction of maximum variability of the uncertainty region by maximizing the minimum FIM eigenvalue  $\psi(\mathbf{H}_{\hat{\boldsymbol{\theta}}}) = -(\lambda_{\min}(\mathbf{H}_{\hat{\boldsymbol{\theta}}}))$  (E-optimal criterion).

After parameters identification, the reliability of model predictions can be assessed through the G-optimality  $V_y$ , i.e. a  $N_y N_{sp} \times N_y N_{sp}$  matrix where the  $ji$ -th diagonal element represents the prediction variance of the  $j$ -th response at the  $i$ -th time point. The  $ji$ -th element  $V_y(\hat{\boldsymbol{\theta}}, \boldsymbol{\varphi})|_{j,i}$  is calculated as:

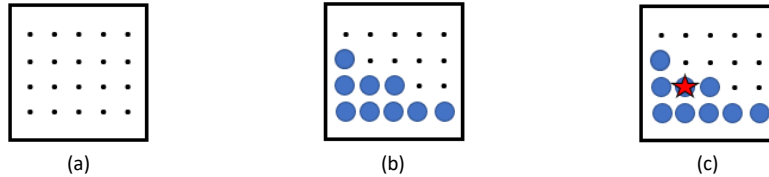
$$V_y(\hat{\boldsymbol{\theta}}, \boldsymbol{\varphi})|_{j,i} = \left( \frac{d\hat{y}_j}{d\hat{\boldsymbol{\theta}}} \right)_i^T [\mathbf{H}_{\hat{\boldsymbol{\theta}}}]^{-1} \left( \frac{d\hat{y}_j}{d\hat{\boldsymbol{\theta}}} \right)_i \quad \text{for } j = 1, \dots, N_y; i = 1, \dots, N_{sp} \quad (3)$$

where  $\left(\frac{d\hat{y}_j}{d\hat{\theta}}\right)_i$  is the  $N_{\theta} \times 1$  vector of sensitivities.  $V_y$  depends on the current parameter estimates  $\hat{\theta}$  and it is evaluated at each condition  $\boldsymbol{\varphi}$  in the design space. For the same  $\boldsymbol{\varphi}$ ,  $N_y N_{sp}$  values of  $V_y(\hat{\theta}, \boldsymbol{\varphi})|_{j,i}$  are calculated; a scalar measure  $J_G$  summarizing them all can be defined as:

$$J_G = \sum_{j=1}^{N_y} \sum_{i=1}^{N_{sp}} V_y(\hat{\theta}, \boldsymbol{\varphi})|_{j,i} \quad (4)$$

We define *G-map* as the mapping of  $J_G$  in the whole design space and we integrate it into the MBDoE optimization seeking enhancement of space exploration and minimization of model prediction uncertainty in the whole domain of model utilization. This is done through the steps illustrated in Fig.1:

- the design space is discretized into a grid of equally spaced points (black dots in (a)) representing experimental conditions  $\boldsymbol{\varphi}$ . In the figure, two inputs are shown for sake of simplicity, but it can be extended to any number  $N_u$ ;
- every point is characterized in terms of information content  $\psi$  and total variance  $J_G$ ;
- only a subset of points is retained for the optimization of Eq. (1). These are the  $\boldsymbol{\varphi}$  satisfying the condition:  $J_G \geq J_{G,thr} J_{G,max}$  (blue circles in (b)), where  $J_{G,max}$  is the maximum  $J_G$  calculated in the grid, while  $J_{G,thr} \in [0,1]$  is a user defined threshold that allows to balance space exploration and information maximization;
- finally, the  $\boldsymbol{\varphi}_{opt}$  satisfying Eq. (1) is selected among the candidates (red star in (c)). It is used as initial guess for a subsequent constrained optimization in Python 3.9 with the Scipy optimization package.



**Figure 1.** Graphical representation of the explorative MBDoE based on *G*-maps.

This procedure is iterated until the budget of experiments ( $N_e$ ) is reached; model calibration is performed at every iteration as soon as the new experiment is designed.

### 3. Results and discussion

To test the performance of the *G*-map eMBDoE, a case study related to the identification of reaction models for catalytic total oxidation of methane over a 5% Pd/Al<sub>2</sub>O<sub>3</sub>- catalyst is considered. The mechanism is described through Mars-van Krevelen reaction kinetics in a flow microreactor system, which proved to be the most suitable kinetic model in Bawa et al. (2022). The control variables and their ranges are: temperature ( $u_1$ ), 250-350 °C; total flow rate ( $u_2$ ), 20-30 Nml min<sup>-1</sup>; inlet methane concentration ( $u_3$ ), 0.005-0.025 mol mol<sup>-1</sup>; oxygen to methane mole ratio ( $u_4$ ), 2-4 mol mol<sup>-1</sup>. The three responses  $\mathbf{y}$  are: unreacted methane, unreacted oxygen and carbon dioxide mole fractions, with measurement errors standard deviations  $\boldsymbol{\sigma}_y = [\sigma_{CH_4}, \sigma_{O_2}, \sigma_{CO_2}] = [0.00043, 0.00202, 0.00051]$ . Consider the kinetic rate of methane:

$$r_{CH_4} = k_1 k_2 P_{CH_4} P_{O_2} / (k_1 P_{O_2} + 2k_2 P_{CH_4} + (k_1 k_2 / k_3) P_{O_2} P_{CH_4}) \quad (5)$$

To aid parameters identification, Arrhenius and Van't Hoff equations are reparametrized:



$$k(T) = \exp(\log k(T_{ref}) - (E_a/R)(1/T - 1/T_{ref})) \quad (6)$$

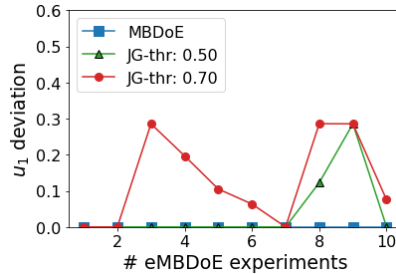
$$K(T) = \exp(\log K(T_{ref}) - (\Delta H/R)(1/T - 1/T_{ref})) \quad (7)$$

Thus, six parameters must be estimated:  $\hat{\theta} = [(-\log k_1(T_{ref})), (\frac{E_{a,1}}{10^4}), (-\log k_2(T_{ref})), (\frac{E_{a,2}}{10^4}), (-\log k_3(T_{ref})), (\frac{E_{a,3}}{10^4})]$ . Parameters estimation is carried out with initial guesses  $\hat{\theta}_{guess} = [3, 8, 3, 10, 10, 8]$  and  $[0,15]$  as parameter Bounds. Experiments are executed in silico by simulating the flow reactor model concentration responses assuming the following set of parameters to represent the system:  $\hat{\theta}_{true} = [5.31, 6.96, 4.88, 10.49, 10.44, 7.95]$  and by adding random gaussian noise with zero mean and standard deviation  $\sigma_y$ .

A G-map eMBDoe with thresholds  $J_{G,thr} = 0.50$  and  $J_{G,thr} = 0.70$  is compared against classical MBDoe and LHS. Since LHS can change randomly the allocation of design points in the design space, three different runs are considered. Finally, 8 LHS points are used as preliminary experiments to obtain initial parameters estimates as well as to gain minimum threshold information prior to applying MBDoe; they are the same in all simulations. The results are analyzed in terms of:

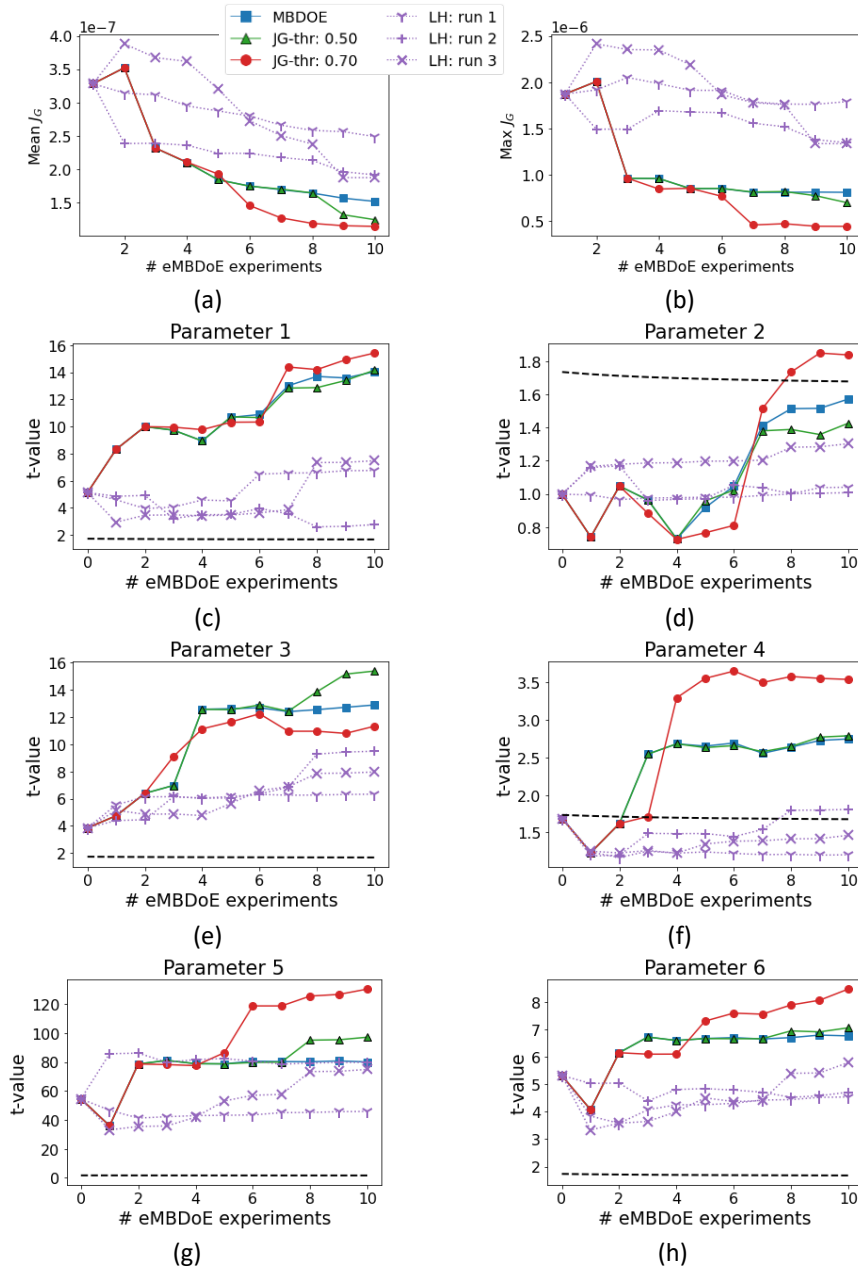
- space exploration, evaluated as deviation from the MBDoe design:  $d_i = \text{abs}(u_{i,l} - u_{i,ref})/u_{i,ref}$ , with  $i = 1,2,3,4$ ;  $l$  refers to  $J_G = \{0.50, 0.70\}$ , while the reference ref is MBDoe;
- mean and maximum  $J_G$  in the whole space; the smaller these indices, the smaller the model prediction uncertainty in the whole space of model utilization;
- parameters precision tests satisfied when  $t_{value,i} \geq t_{ref}$  (Aprey and Naka, 1999).

As seen in Fig. 2, the higher the G-optimality threshold, the greater the deviation of the optimal input values from the ones calculated through a standard MBDoe, especially at the first iterations of the experimental campaign. For instance, variable  $u_1$  with  $J_G = 0.70$  has  $d_i > 0$  starting from the 3<sup>rd</sup> iteration, while with  $J_G = 0.50$  it has  $d_i > 0$  after the 8<sup>th</sup> iteration. The results in terms of LHS are not shown since their departure from the MBDoe design is expected due to its explorative nature.



**Figure 2.** Deviation of the Gmap eMBDoe design with  $J_G = \{0.50, 0.70\}$  from the MBDoe optimal design. Only the deviations of variable 1 is shown for sake of conciseness. The x-axis shows the number of eMBDoe experiments after 8 LHS preliminary experiments.

The effect of this deviation from MBDoe conditions is assessed in terms of reduction of model prediction variance (Fig. 3 a-b) and increase of parameters precision (Fig. 3 c-h).



**Figure 3.** Results of eMBDOE compared to LHS and MBDOE: (a)-(b) mean and maximum  $J_G$  in the whole design space; (c)-(h) t-values for the full set of model parameters. In (c)-(h), the black dotted line represents the reference t-value, while the 0 in the x-axes represents the results of model calibration with the 8 LHS preliminary experiments only.

As shown in Fig.3 a-b, a purely explorative technique such as Latin Hypercube sampling has the worst performance in the reduction of model prediction uncertainty; this is true

for all the three runs considered. MBDoE is able to reduce model prediction uncertainty, but the best performance is obtained by the explorative MBDoE, particularly when a stricter threshold (namely,  $J_{G,thr} = 0.70$ ) is used.

Finally, assuming a maximum budget of 10 experiments, parameters identification is characterized by (Fig. 2 c-h):

- statistically satisfactory estimation of parameters  $\hat{\theta}_1, \hat{\theta}_3, \hat{\theta}_5$  and  $\hat{\theta}_6$  with the preliminary dataset in all scenarios;
- $\hat{\theta}_4$  estimated precisely from 3 optimal experiments for MBDoE and eMBDoE, while the estimate is more critical for LHS (estimated in one run over 3);
- the most critical parameter, e.g.  $\hat{\theta}_2$ , is sufficiently precise only when a Gmap eMBDoE is used with a strict threshold of 0.70.

This suggests that the proposed explorative MBDoE does not cause a loss of information content.

## Conclusion

A novel MBDoE technique is proposed in this paper in order to favor experimental design space exploration towards regions characterized by higher prediction uncertainty. Such regions are detected by using G-optimality maps, while the trade-off between information maximization and space exploration can be handled through a G-optimality threshold. The stricter the threshold, the higher the deviation of eMBDoE design from a conventional MBDoE. The new explorative and optimal design has benefits in terms of reduction of prediction variance as it reduces the mean and maximum G-optimality values in the entire design space with respect to MBDoE. Furthermore, G-map eMBDoE is more effective on estimating the full set of model parameters, since all t-tests are passed within 8 eMBDoE experiments. However, results show how a LHS space filling design, which is inherently explorative, can perform poorly both in terms of prediction variance minimization and increase of parameters precision. This suggests that the best experimental design performance is obtained only if a balance between space exploration and information maximization can be achieved.

## References

- S. P. Asprey, S. Macchietto (2000). Statistical Tools for Optimal Dynamic Model Building. *Comput. Chem. Eng.*, **24**, 1261-1267.
- S. P. Asprey, Y. Naka, (1999). Mathematical problems in fitting kinetic models-some new perspectives. *J. Chem. Eng. Jpn.*, **32(3)**, 328–337
- Y. Bard (1974). *Nonlinear Parameter Estimation*. New York, NY: Academic Press.
- S. G. Bawa, A.Pankajakshan, C. Waldron, E. Cao, F. Galvanin, A. Gavriilidis (2022). Rapid Screening of Kinetic Models for Methane Total Oxidation using an Automated Gas Phase Catalytic Microreactor Platform. *Chemistry—Methods*, e202200049.
- R. A., Fisher (1950). Contributions to Mathematical Statistics. Papers 10, 11,38, JWS.
- J. Kiefer, Wolfowitz (1960), The equivalence of two extremum problems, *Canad. J. Math.* **12**, 363-366.
- M.D. McKay, R.J. Beckman, W.J. Conover (1979). "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code". *Technometrics*. American Statistical Association. **21** (2), 239–245.

# GHG emission reduction via multiple fuel options in carbon-constrained industrial clusters

Christelle Bechara<sup>a</sup>, Sabla Y. Alnouri<sup>b\*</sup>

<sup>a</sup>*Baha and Walid Bassatne Department of Chemical and Petroleum Engineering, American University of Beirut, P.O. Box 11-0236, Riyad El-Solh, Beirut, Lebanon*

<sup>b</sup>*Gas Processing Center, College of Engineering, Qatar University, Doha, Qatar*  
*sabla@qu.edu.qa*

## Abstract

This paper primarily investigates the effect of incorporating multiple fuel options onto carbon constrained industrial clusters. The purpose of doing so was to maximize the profit through the optimal allocation of multiple types of fuel. The cluster consists of a number of different production plants and a power plant for the processing of value-added products as well as the generation of electricity. The power plant is connected to the electric grid and constitutes the main source of electricity for the cluster. It also has the ability to produce more power and export it to the grid. The integration of additional plants that could generate new products, convert, or store carbon dioxide through CCUS technologies, and produce cleaner power through the use of renewable and/or multiple and blended fuels is assessed. It was found that the incorporation of multiple types of fuel within a single cluster resulted in the lowest post-capture emissions with 9469 t CO<sub>2</sub>/d generating a high-end profit of \$907.16 million/y, when compared to cases which relied on a single fuel option only.

**Keywords:** GHG Emissions, Industrial Clusters, Biomass, Energy, Fuel

## 1. Introduction

Fossil fuels, mainly oil, coal, and gas, have been dominating the world's energy market. In fact, they are considered to be the most commercially attractive, (Wang and Economides, 2009). Moreover, the versatile functions that are associated with such fossil fuels, such as the production of value-added products to the supply of electricity and power, has all led to a drastic increase in industrial dependence on such energy sources. The industrial sector's energy consumption has been forecasted to grow by 32% by the year 2050 (Wang, 2019), which inevitably instigates major environmental concerns. Fossil fuel utilization is associated with high CO<sub>2</sub> emissions. Recent studies have shown that 18% of global emissions are a result from the combustion of natural gas, 40% from the combustion of coal, and 31% from the combustion of oil and (Olivier and Peters, 2018). This has pushed many governments to start enforcing carbon taxes on industrial clusters and set CO<sub>2</sub> emission limits in compliance with the targets imposed by the Paris Agreement. The need to satisfy the rising energy demands, while conforming to the set emission bounds, has pushed mitigation techniques that could be implemented in industrial clusters. Bechara and Alnouri (2022) have previously assessed and compared different energy scenarios in industrial clusters. In this work, the incorporation of multiple and simultaneous fuel options within an industrial cluster is studied, and compared against single type fuel clusters.

## 2. Methodology

The generic plant in this model can have several active sinks and sources, as shown in Figure 1. The sink options include: a fuel sink, a power sink, and a CO<sub>2</sub> sink, while the source options include: a product source, a CO<sub>2</sub> source, and a power source. The plant can be of two types: a production plant or a power plant with one source of fuel allocated to it. In terms of inputs, the plant requires a fuel input either to produce a value-added product or to generate power, a power input to satisfy its electricity requirements, and/or a carbon input, as in the case of a carbon utilization plant. In terms of outputs, a plant can produce products to be exported out of the industrial cluster, generate power to either satisfy the cluster's need or to be exported as revenue to the electricity grid, and emit carbon dioxide, which in turn is either released into the atmosphere or captured, stored, or utilized by the carbon capture and utilization technologies. Moreover, each production plant contributes to the total CO<sub>2</sub> emissions of the cluster based on the number of products it produces, the quantity of fuel, and the amount of electricity it requires to meet the set production. The power plant, in turn, contributes based on the amount of electricity the cluster requires and the quantity exported to the grid as an additional revenue source. The integration of the additional plants, the option of utilizing several types of fuels as well as renewable power within the cluster serve the following functions:

- The reduced dependence of the cluster on one type of fossil fuel.
- The reduction of the total CO<sub>2</sub> footprint of the cluster through the use of carbon utilization technology with the possibility of monetizing the captured carbon.
- The production of a clean source of power to satisfy the cluster's requirement and reduce the fossil fuel requirement for power production.

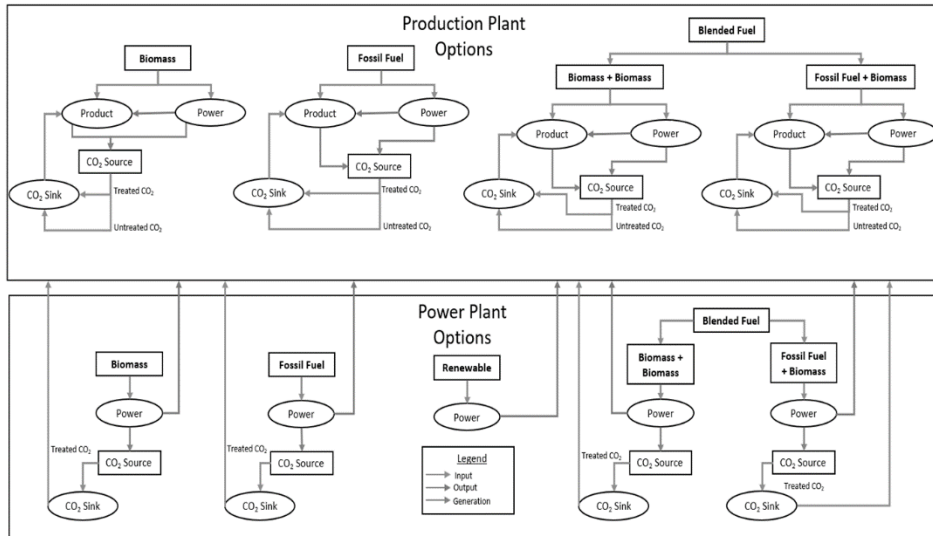


Figure 1: A generic illustration of sinks and sources within an industrial cluster

Because of the large number of variables involved in the optimization problem, from fuel allocation to CO<sub>2</sub> utilization, and power generation, a superstructure network approach is used to capture all the available configurations and determine the optimal selection. The cluster has an infrastructure suitable for the allocation and transportation of the different types of fuels that are introduced, from natural gas, to coal, to biomass, and blended fuel. All the product streams as well as the power exported to the grid are

considered as revenue streams that serve the end profit of the cluster. All active power sources are connected to the cluster's electricity grid which is supplied by the existing power plant. The CO<sub>2</sub> streams, treated and untreated, are connected to the CO<sub>2</sub> active sinks of all plants and the uncaptured CO<sub>2</sub> sources are emitted into the atmosphere. Taking all the aforesaid connections and networks into account, the optimization model was primarily designed to determine the optimal scenario that respects all constraints and results in the maximum profit for the cluster. To achieve the aforementioned, this work presents a superstructure optimization-based approach that aims at developing sustainable fuel and carbon utilization strategies to meet the environmental restrictions at a maximum end-profit. To summarise, the optimization model is able to determine:

- The selection of the production plants and the type of power plant.
- The integration of renewable energy as a replacement to the fuel-derived energy with regard to the economic and environmental aspects.
- The optimal allocation of the fuel options to the production and power plants.
- The integration of carbon capture technologies onto the cluster.
- The capturing, treating, and monetizing the captured carbon.

Moreover, the following factors are assumed to be known:

- The existing production plants, their capacities, and their locations in the cluster.
- The products that are produced in every production plant in the cluster.
- The characteristics and costs of the fuels to be allocated.
- The power requirements of the production plants in the cluster.
- The maximum renewable percentage that can replace the fuel-derived power.
- The product-to-fuel and power-to-fuel requirements.
- The parameter that describes the carbon emitted per ton of product, kWh of power produced, during the transport of solid fuels
- The revenue/cost parameter of every revenue/cost stream.
- The capital and operating costs of every plant.
- The cost of the carbon integration network (transportation, capture, treatment, compression, and pumping parameters) and the carbon capture limit.

### **3. Case Study**

The integration of different types of fuels was assessed based on both the economic and environmental feasibilities. The cluster investigated includes a set of existing and optional production plants. The existing plants were Methanol A, Aluminium, Cement and the natural gas power plant. Whereas the optional entities were Methanol B, Greenhouse, Saline storage and Enhanced Oil Recovery. Moreover, a renewable power plant with PV units has also been considered as an optional addition. Choosing between a natural gas, coal, biomass, and blended fuel power plants, and carbon utilization technologies were considered as part of the optimization problem. All process data has been obtained from Al-Mohannadi et al. (2017). The following additional assumptions were used:

- Natural gas, coal, biomass, and blended fuel are all integrated in the cluster.
- Only one type of fuel-derived power plant exists.
- Each production plant can have one type of fuel allocated to it.
- The net carbon capture rate is set to a minimum of 20%. The optimization model will select the optimal carbon capture percentage based on the set constraints.

The Mixed Integer Linear Program (MILP) involved a total of 634 variables, with a CPU

time of 3 mins and 20s using What'sBest (Lindo systems, 2006) Lindo Branch-and-Bound solver for MS-Excel 2013 via a laptop PC with Intel Core i7 Duo processor, 8 GB RAM and a 64-bit operating System. It was found that with a natural gas flow of 444.25 t natural gas/d, a coal flow of 888.35 t coal/d, a biomass flow of 585.39 t biomass/d, and a blended fuel flow of 756.14 t blended fuel/d and a carbon capture rate constraint set to a minimum of 20%, the cluster produced 5000 tMeOH(A), 10000 t Portland-cement, 1118.44 tAluminum, and 2600 tMeOH(B). The optimization selected the natural gas plant as the operating power plant, which satisfied the cluster's power requirements. The option of exporting power to the grid was not reported in the optimal solution, instead, the option of operating the renewable energy plant at the highest capacity to replace 5% (1010526.32 kWh/d) of the power supplied by the fossil fuel plant. The cluster's profit turned out to be \$907.16 million/y, slightly lower than that for the natural gas case study, but higher than the remaining case studies. However, the different fuel requirements are much lower, reducing the dependency on one type and showing the advantage of integrating more than one fuel into an industrial cluster. The activation of the renewable solar plant further shows the opportunity of gradually integrating more sources of green energy, such as green hydrogen, without being limited by capacity or by efficiency. To better understand the cluster's economics, with multiple fuels enabled, a cost distribution is presented in Figure 2. With the cluster's large product requirements, the highest costs resulted from the production sector, requiring \$1010.12 million/y, followed by the power sector, consisting of the natural gas power plant and the solar photovoltaic plant costs with \$151.53 million/y.

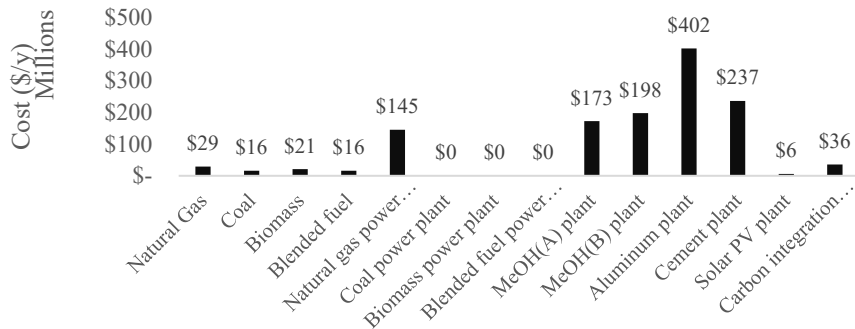


Figure 2: Optimal cost distribution attained

Analysing the carbon integration network's cost distribution, displayed in Figure 3, it can be seen that the highest cost is attributed to the treatment unit (Almohannadi and Alnouri, 2021), accounting for 65.28% of the total carbon integration costs. This high number is also associated with the fact that the optimization model selected to treat 99.91% of the carbon streams.

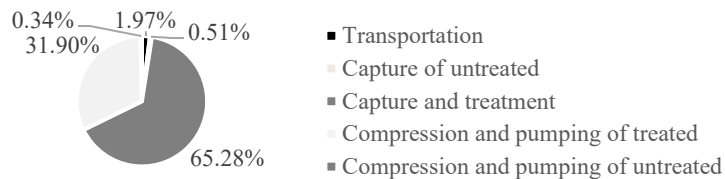


Figure 3: Carbon Network Cost Distribution

Regarding the carbon footprint, the total emissions of the cluster without capture was found to be 17640 t CO<sub>2</sub>/d, lower than those for the natural gas, biomass, and blended fuel cases, but slightly higher than that for the coal case study. The latter can be attributed to the higher production capacity of this cluster in comparison to the coal cluster. After applying carbon capture and utilization technologies and selecting the optimal source to sink mapping (refer to Figure 4) the carbon captured amounted to 13796 t CO<sub>2</sub>/d reducing the cluster's emissions by 47% resulting in a final footprint of 9345 t CO<sub>2</sub>/d, and a capture rate of 79%. Compared to the standalone and blended fuel case studies, the resulting emissions post capture of the multiple fuel case study turned out to be the lowest showing the great economic and environmental potentials of integrating multiple fuels into a cluster to mitigate its carbon footprint. Furthermore, the optimization chose to operate the solar photovoltaic plant at full capacity. Therefore, adding more capacity or integrating more than one type of renewable energy could result in a higher replacement of the fossil-fuel derived energy. Figure 4 summarizes the optimal source to sink mapping obtained. A treated carbon stream was allocated from aluminum to methanol (B), treated carbon from cement and a mixture of treated and untreated carbon from the natural gas power plant to the enhanced oil recovery, treated carbon from methanol (B) to the greenhouse, and treated carbon from methanol (A) and the natural gas power plant to the saline storage.

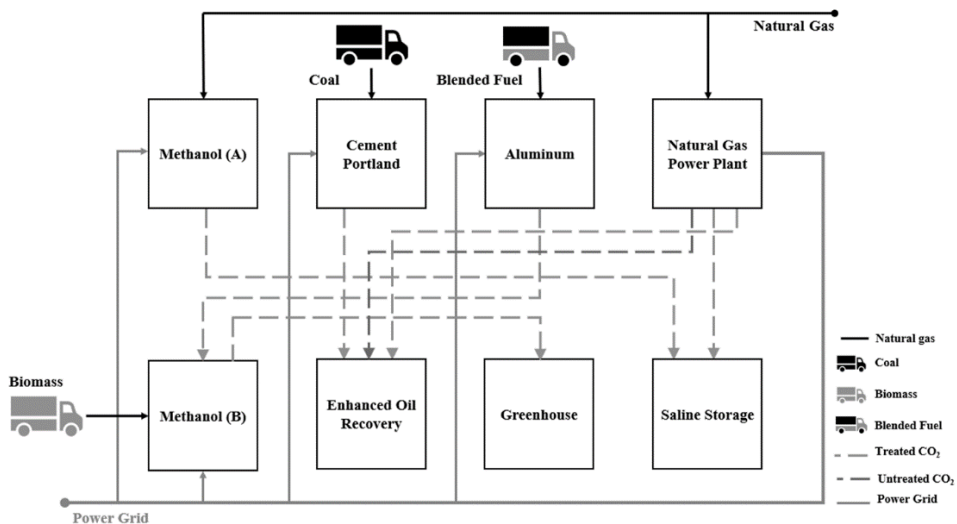


Figure 4: multiple fuel CO<sub>2</sub> network superstructure

Looking at the carbon emissions distribution in the cluster (illustrated in Figure 5) it can be seen that the production and power sector generated the highest emissions with 10780.58 t CO<sub>2</sub>/d and 6302.69 t CO<sub>2</sub>/d. Comparing the power emissions in every case study, natural gas resulted in 0.000691 tCO<sub>2</sub>/kWh, coal in 0.00077 tCO<sub>2</sub>/kWh, biomass in 0.000773 tCO<sub>2</sub>/kWh, blended fuel in 0.00077 tCO<sub>2</sub>/kWh, and finally multiple fuels in 0.000328 tCO<sub>2</sub>/kWh. The values were obtained by dividing the carbon emissions by the fuel derived power. Hence, the incorporation of multiple fuels yielded the lowest emissions per kWh, proving the advantage of adding renewable sources of energies into the power portfolio of a cluster. The flow rates of the different types of fuel are represented in Figure 6, showing the lower flow requirements needed in this analysis compared to those of the standalone and blended fuels. This further proves that integrating



different types of fuel results in a reduction in the reliance of the industry on one type. The total attained revenue for this case was found to be \$2170 million/y.

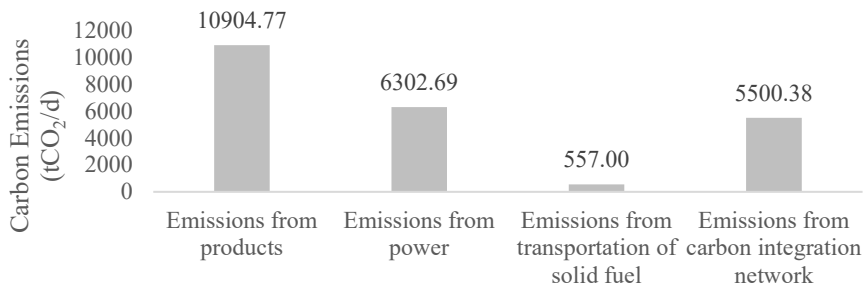


Figure 5: carbon emissions distribution in the cluster

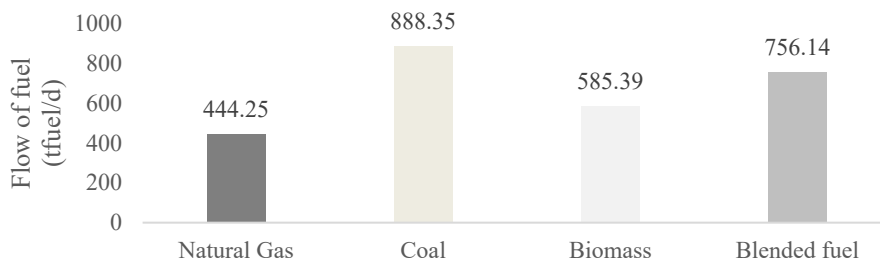


Figure 6: Total fuel flows fuel into the multiple fuel industrial cluster

#### 4. Conclusion

A mathematical model for a systematic approach that integrates multiple types of fuels into industrial clusters was developed, by integrating the following multiple fuel options into the cluster: natural gas, coal, biomass, and blended fuel. The results obtained pertaining to fuel use, and its distribution amongst product and power sinks have been discussed. Moreover, the carbon emission footprint has analysed, and potential reduction strategies via CCUS utilization have been proposed, so as to meet the set emission target.

#### References

- C. Bechara, S. Y. Alnouri, 2022, Energy assessment strategies in carbon-constrained industrial clusters. *Energy Conversion and Management*, 254, 38, 115204.
- D. Al-Mohannadi, K. Abdulaziz, S. Y. Alnouri, P. Linke, 2017, On the synthesis of carbon constrained natural gas monetization networks, *Journal of Cleaner Production*, 168, 735-745.
- D. Al-Mohannadi, S. Y. Alnouri, 2021, Studying the effect of solvent properties in treatment design within CO<sub>2</sub> Integration Networks, *Computer Aided Chemical Engineering*, 50, 295-300
- R. E. Ooi, D. C. Foo, R. R. Tan, D. K. Ng, R. Smith, 2013, Carbon Constrained Energy Planning (CCEP) for Sustainable Power Generation Sector with Automated Targeting Model. *Industrial & Engineering Chemistry Research*, 52, 29, 9889-9896.
- X. Wang, M. Economides, 2009, Natural Gas Supply, Alternative Energy Sources, and the Environment, *Advanced Natural Gas Engineering*, 303-331.
- T. Wang, 2019, Forecast for worldwide industrial energy consumption from 2018 to 2050, by fuel, In. *Statista* 2020.
- J. G. Olivier, J. A. Peters, 2018, Trends in Global CO<sub>2</sub> and Total Greenhouse Gas Emissions, The Hague: [www.pbl.nl/en](http://www.pbl.nl/en)

# An optimization-free Fisher information driven approach for online design of experiment

Andrea Friso,<sup>a</sup> Federico Galvanin,<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering University College London, Torrington Place,  
WC1E 7JE London, United Kingdom*

## Abstract

Developing mathematical models used to describe reaction kinetics is pivotal for the design, control and optimization of chemical processes. One of the most challenging tasks in the model development procedure is the identification of the unknown parameters within the model. This problem can be addressed using Model-Based Design of Experiment (MBD<sub>oE</sub>) techniques that allow experiments to be designed in such a way that parameters can be precisely estimated with the minimum number of runs and analytical resources. However, MBD<sub>oE</sub> techniques rely on an optimization procedure that is affected by the uncertainty related to the identified parameters and can be computationally expensive and prone to local optimality issues. MBD<sub>oE</sub> techniques are also applied to online procedures for faster identification of the kinetic model in autonomous platforms, and for this reason it is necessary to ensure a fast convergence and avoid numerical convergence issues during the operation. In this paper, a new optimization-free technology is proposed to tackle the above-mentioned problems.

**Keywords:** design of experiment, parameter estimation, Fisher Information matrix

## 1. Introduction

In chemical engineering, the role of kinetic models is fundamental to describe in a quantitative way the progress of chemical reactions occurring in a reaction system. These models, in the form of mathematical equations, play a key role in the design, control and optimization of chemical processes. When conventional model-building procedures are used reaction kinetics are modelled by proposing a set of candidate models based on preliminary experimental observations and hypotheses on potential reaction mechanisms. The key subsequent step is to determine the best model among the candidates to describe the system under analysis. The standard sequential methodology proposed by Franceschini and Macchietto (2000) consists of three different stages: 1) Preliminary analysis based on the identifiability and distinguishability of the model structure; 2) Model-based Design of Experiments to discriminate among the rival models that passed the first stage (MBD<sub>oE</sub>-MD); 3) Model-based Design of Experiments to improve the precision of the identified parameters for the best model selected in step 2 (MBD<sub>oE</sub>-PE). If no model passes the identifiability test of step 1 it is necessary to propose a new set of candidates. When a new experiment must be performed (either in steps 1 and 2) the new experimental conditions are optimally computed according to the expected amount of information for model discrimination (step 2) or parameter precision (step 3) the new experiments will eventually return to the experimenter. Automated model identification platforms were recently used to estimate the parameters of kinetic models online and speed up the identification procedure to reduce the costs of experimentation (Waldron et al, 2020, Pankajakshan et al, 2019). However, the parameter identification problem is often ill-conditioned and consequently the objective function of the MBD<sub>oE</sub> problem

may become difficult to compute (Quaglio et al. 2019). Moreover, optimization in these platforms can be difficult to handle in online model identification procedures for the following additional reasons: *i*) when the system under analysis is structurally very complex, and consequently requires model candidates containing a large set of variables and model parameters, numerical optimization algorithms may easily incur into infeasible solutions; *ii*) particularly for systems characterized by a high number of state variables and model parameters MBDDoE techniques can become computationally very expensive, thus reducing the benefits of online identification of kinetic models. A new MBDDoE approach to increase the precision of parameter estimation is proposed in this paper where a new optimization-free Fisher information-driven algorithm has been developed to overcome the above-mentioned problems. The method has been tested on a benchmark case study for the identification of the parameters of the Bakers yeast growth model (Asprey et al. 2000). Results show that the information-driven approach is comparable with MBDDoE methods in terms of experiments required for the precise identification of the parameters, but the information-driven approach is less computationally expensive.

## 2. Proposed framework and methodologies

The set of initial candidates used in the MBDDoE procedures are formulated as differential and algebraic equations in the form:

$$\begin{cases} f(\dot{\mathbf{x}}(\mathbf{t}), \mathbf{x}(\mathbf{t}), \mathbf{u}(\mathbf{t}), \boldsymbol{\theta}, \mathbf{t}) = 0 \\ \hat{\mathbf{y}} = \mathbf{g}(\mathbf{x}(\mathbf{t})) \end{cases} \quad (1)$$

In Eq. 1  $\mathbf{x}$  is the vector of  $N_x$  states variables,  $\dot{\mathbf{x}}$  is the vector of the derivative of the state variables,  $\mathbf{u}$  is the vector of the input of the system under analysis,  $\boldsymbol{\theta}$  is the vector of the model parameters of  $N_\theta$ -dimensions,  $t$  is the time variable,  $\hat{\mathbf{y}}$  is the vector of the measured output of the system. The objective of the proposed methodology is to reduce the number of experiments needed to estimate the parameters of the model with adequate precision, so the objective is to perform experiments that lead to estimate precisely the model parameters, i.e. with minimum uncertainty. The proposed procedure for the optimization-free Fisher Information Matrix (FIM)-driven MBDDoE algorithm is divided in 7 sequential steps, as reported in Figure 1:

1. Definition of experimental design bounds: in this step, all the control variables that can be manipulated and optimized during the process are collected in the design vector:  $\boldsymbol{\varphi} = \boldsymbol{\varphi}[\mathbf{u}(\mathbf{t}), \mathbf{y}^0, \boldsymbol{\tau}, \bar{\mathbf{w}}, \mathbf{t}_{sp}]$  where  $\mathbf{u}(t)$  is the vector of the time-varying input controls,  $\mathbf{y}^0$  is the vector of the initial conditions,  $\boldsymbol{\tau}$  is the experiment duration,  $\bar{\mathbf{w}}$  is the vector of the time-invariant input controls and  $\mathbf{t}_{sp}$  is the vector of the sampling times. In addition, the bounds on the experimental vector are defined to create the experimental space;
2. Definition of preliminary experiments using DoE techniques:  $N_{exp}$  sets of experimental conditions to test are computed using a space-filling DoE method, for example, Latin Hypercube Sampling (LHS).
3. Selection of the first experiment to run: since no information about the system is available at this stage the selection of the first experiment is random.
4. Estimation of the parameter: after collecting the experimental data, in step 4 the model parameters are estimated minimizing the loss function:

$$LOSS = \sum_{i=1}^{N_{exp}} \left( \frac{y_i - \hat{y}_i}{\sigma} \right)^2 \quad (2)$$

where  $y_i$  is the measurement of experiment  $i$ ,  $\hat{y}_i$  is the predicted output of experiment  $i$  obtained using the model and  $\sigma$  is the standard deviation. The initial guesses on model parameters are updated with the new estimation ( $\hat{\theta}$ ).

5. Statistics on model parameters and FIM evaluation: the  $t$ -test is used to evaluate the relative precision of the estimates. At this point the FIM (Zullo 1991) is evaluated using the following equation:

$$\mathbf{M}(\hat{\theta}, \boldsymbol{\varphi}) = \sum_{j=1}^{n_{\text{exp}}} \sum_{r=1}^{n_{\text{resp}}} \sum_{s=1}^{n_{\text{resp}}} \tilde{\sigma}_{rs,j} \mathbf{Q}_{r,j}^T \mathbf{Q}_{s,j} + \mathbf{M}^0 = [\mathbf{V}(\boldsymbol{\theta}, \boldsymbol{\varphi})]^{-1} \quad (3)$$

In (3)  $n_{\text{exp}}$ ,  $n_{\text{resp}}$  are respectively the number of experiments and the number of responses,  $\mathbf{Q}$  is the sensitivity matrix and  $\mathbf{M}^0$  is the prior information on the parametric system. By inverting the matrix  $\mathbf{M}$ , it is possible to obtain the variance-covariance matrix ( $\mathbf{V}$ ).

6. The evaluation of  $\mathbf{M}$  is done for each experimental design point generated in step 2 to compute the expected amount of information achievable from each experiment. Experiments are then ranked based on the computed information content using the concept of Relative Fisher information (RFI) (Galvanin et al. 2016).
7. At this point the experimental conditions with the highest expected information are selected (Eq. 4) and the corresponding experiment is performed.

$$\boldsymbol{\varphi}_{i+1} = \operatorname{argmax}(\mathbf{M}(\hat{\theta}, \boldsymbol{\varphi})) \quad (4)$$

Steps 4, 5, 6 and 7 are repeated iteratively for the fixed number of experiments set in step 2. Within this framework, it is possible to select at each iteration the most informative experiment to run in the candidate set without performing an optimization. The optimization-free approach is less prone to local optimization issues affecting the standard MBDofE approaches because an ‘exploratory part’ in the selection of the experiments is forced in step 2 as the initial selection of the experiments to perform using the LHS as preliminary DoE. In fact, since the optimal experiment is only selected based on the maximum evaluation of information from Eq. 4 there is no risk of incurring in local optima. The absence of optimization can, however, also be a limitation of the methodology, because the solution found through this procedure will always be sub-optimal. In addition, the fact that the first experiment is selected randomly (as would be done in practice if no preliminary knowledge of the system is available) can affect the computational performance and efficiency in subsequent iterations.

### 3. Case study description

The proposed framework has been tested using an in-silico case study related to a fed-batch reactor in which a fermentation reaction is carried out using baker’s yeast (Asprey et al. 2000). The model used to simulate the experiments is presented in Eq. 3.

$$\begin{cases} \frac{dx_1}{dt} = \left( \frac{\theta_1 x_2}{\theta_2 x_1 + x_2} - u_1 - \theta_4 \right) x_1 \\ \frac{dx_2}{dt} = - \frac{\theta_2 x_1 x_2}{(\theta_2 x_1 + x_2) \theta_3} + u_1 (u_2 - x_2) \end{cases} \quad (5)$$

This system has been chosen to test the method because the governing equations and their parameters are very well known. The experimenter aims to compute the dynamics of yeast concentrations  $x_1(t)$  [ $gL^{-1}$ ] and the substrate concentrations  $x_2(t)$  [ $gL^{-1}$ ] as a function

of two different inputs: the dilution factor,  $u_1 [h^{-1}]$ , and the substrate concentration in the feed  $u_2 [gL^{-1}]$ . In this model the vector of model parameters that need to be identified precisely is  $\theta = [\theta_1, \theta_2, \theta_3, \theta_4]^T$ . Parameters and constant input variables used to perform in-silico experiments are reported in Table 1.

**Table 1:** Parameter and constant input variables values used for the in-silico experiments.

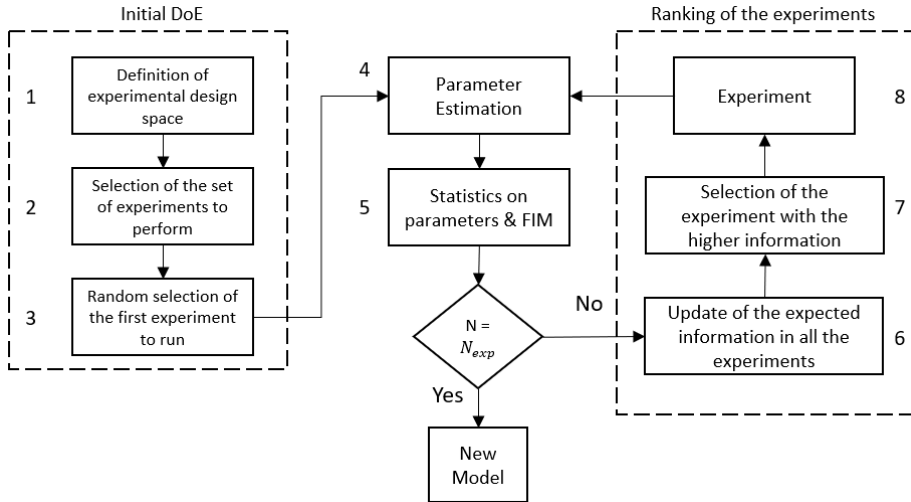
Parameter	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$x_2(0)$	$u_2$
Values	0.31	0.18	0.55	0.05	0.01	5
Initial guess	5	5	5	5	-	-

The design vector is  $\phi = [x_1, u_1]$ , where  $x_2$  and  $u_2$  are kept constant. The bounds used to define the experimental space are the same used by Asprey et al. (2000) and are reported in Table 2.

**Table 2:** Bounds on the experimental conditions.

Parameter	Lower Bound	Upper Bound
$x_1(0) [gL^{-1}]$	1	10
$u_1 [h^{-1}]$	0.05	0.2

The proposed FIM-driven method is compared with a classical MBDoE approach and with a DOE LHS approach. In the MBDoE framework, the experimental conditions are selected iteratively after a sequential optimization procedure as in Asprey et al (2000). In the DOE procedure the same experiments generated for the optimization-free method in step 2 are used but randomly selected so without operating a ranking of the experiments.



**Figure 1:** Block diagram representation of the proposed methodology for the design of experiment for parameter identification.

#### 4. Results

In this section, three experimental design methodologies are compared: DOE (LHS), the proposed FIM-driven approach and a standard MBDoE. The comparison of the profiles of parameter estimates with the number of experiments is reported in Figure 2 for the full set of model parameters. The profile of estimated values and the variance of parameters

are reported to quantify the precision of the estimates during the model identification procedure.

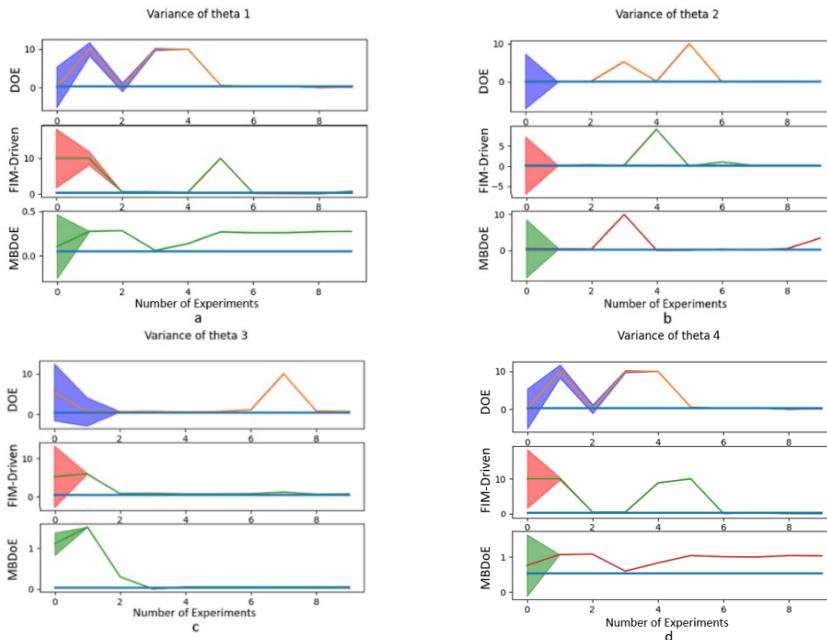


Figure 2 Profiles of  $\theta_1$  (a),  $\theta_2$  (b),  $\theta_3$  (c) and  $\theta_4$  (d) estimated values and relative variances.

From Figure 2 it is possible to observe that by using MBDoE it is possible to obtain the fastest reduction in parameter variance. Instead, the variance reduction obtained with the proposed FIM-driven approach is slower than the one obtained with MBDoE and faster than the one obtained with a conventional DOE methodology. As expected, in MBDoE the variance is generally lower and decreases faster than the variance obtained from the other two methods. This means that the precision of the estimations increases with every iteration faster than the other frameworks. The proposed methodology allows convergence to the assumed true value and negligible variance in 6 experiments. Instead, DoE requires more experiments (9) to obtain the same results. In Figure 3 the profiles of the  $t$ -values of the identified parameters are reported to evaluate and compare the relative precision of the estimates using the three methods. In Figure 2 the blue line represents the reference  $t$ -value, and the orange line represents the profile of the  $t$ -value of the parameter. If the computed  $t$ -values of the parameters are higher than the reference  $t$ -value the parameters are evaluated with good precision. From Figure 3 it is possible to notice that with a DoE method, the  $t$ -value for parameter  $\theta_3$  is lower than the reference until the sixth experiment and so the estimate is uncertain. Regarding the convergence speed of the three different procedures, the CPU time has been evaluated for all three methods. DoE is the fastest method (1.70 s) followed by FIM-driven (2.42 s) while MBDoE is the slowest method (10.47 s). These algorithms have been tested using an Intel® Core™ i9-10885H @ CPU 2.40 GHz with 16.0 GB RAM.

## 5. Conclusion

This study allows to define a new framework for the fast identification of parameters of kinetic models using an optimization-free Fisher Information Matrix-driven approach.

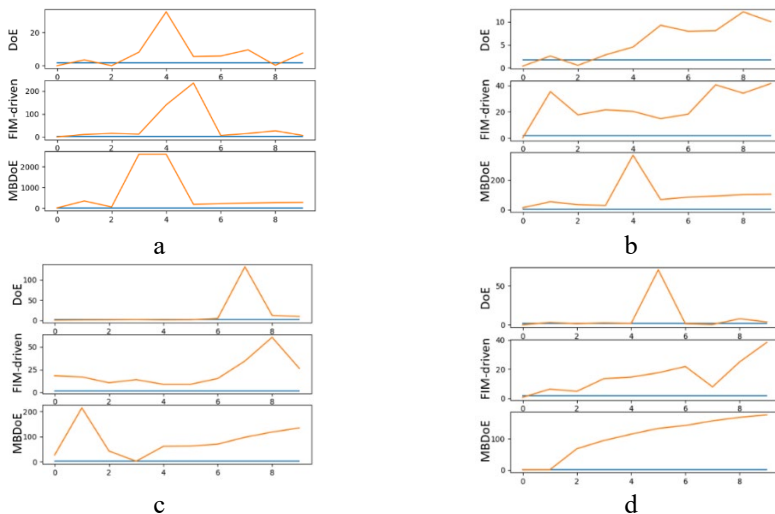


Figure 3 t-values of the parameters  $\theta_1$  (a),  $\theta_2$  (b),  $\theta_3$  (c) and  $\theta_4$  (d)

Preliminary results show how the proposed method can outperform standard space-filling design managing to identify the full set of model parameters precisely with a limited number of runs and providing results in parameter accuracy that are comparable to a standard sequential MBDoE but less computationally intensive. Future works will include re-sampling methods in the proposed procedure and the application to real chemical systems in automated platforms to test the robustness of the method and the comparison in terms of computational time required in practical lab settings.

## Acknowledgment

This project has received funding from EPSRC (EP/R032807/1). The support is gratefully acknowledged.

## References

- Asprey, S.P., Macchietto, S., 2000. Statistical tools for optimal dynamic model building. *Computers and Chemical Engineering* 24, 1261–1267
- Quaglio, M., Waldron, C., Pankajakshan, A., Cao, E., Gavriilidis, A., Fraga, E. A., Galvanin, F., 2019, An online reparametrisation approach for robust parameter estimation in automated model identification platforms, *Comp. Chem. Eng.*, 124, 270.
- Galvanin, F., Cao, E., Al-Rifai, N., Gavriilidis, A., Dua, V., 2016, A joint model-based experimental design approach for the identification of kinetic models in continuous flow laboratory reactors, *Comp. Chem. Eng.*, 95, 202.
- Waldron C., Pankajakshan A., Quaglio M., Cao E., Galvanin F., Gavriilidis A., Model-based design of transient flow experiments for the identification of kinetic parameters, *React. Chem. Eng.*, 2020, 5, 1, 112-123, The Royal Society of Chemistry.
- Pankajakshan A., Waldron C., Quaglio M., Gavriilidis A., Galvanin F., A Multi-Objective Optimal Experimental Design Framework for Enhancing the Efficiency of Online Model Identification Platforms, *Engineering*, Volume 5, Issue 6, 2019.
- Zullo, L., 1991. Computer aided design of experiments. An engineering approach. Ph.D. Thesis, University of London, UK.

# A Python-based approach for thermodynamic consistency tests of binary VLE data

Pedro Cezareth,<sup>a</sup> Martina Costa Reis,<sup>a</sup>

<sup>a</sup>*University of São Paulo, School of Engineering, São Paulo 05508-000, Brazil*

## Abstract

Although reliable experimental data are fundamental in process design and optimization, critical assessment of experimental vapor-liquid equilibrium data is usually time-consuming as it requires data mining and multiple thermodynamic and statistical analyses. Given this, any initiative towards the development of computational tools that helps to speed up the assessment of experimental vapor-liquid equilibrium data is essential. Therefore, in this work, a Python-based approach for thermodynamic consistency tests of vapor-liquid equilibrium data of binary mixtures is presented. While the consistency of experimental isobaric vapor-liquid equilibrium data was checked through the L-W and Redlich-Kister tests, the thermodynamic consistency of isothermal vapor-liquid equilibrium data was analyzed through the area, Redlich-Kister, and L-W tests. The objective of this work is to introduce an open-source, efficient computational tool implemented with Python that embeds different thermodynamic consistency tests which helps chemical engineers make rational use of experimental data in process simulation and model parameterization.

**Keywords:** Thermodynamic consistency tests, vapor-liquid equilibrium, binary mixtures.

## 1. Introduction

A fundamental requirement for the design and optimization of chemical processes is the use of a reliable set of experimental data of pure component and mixture properties. Since the design and optimization of chemical processes depend on thermodynamic models whose key adjustable parameters are determined by regression of experimental data, the use of a thermodynamically consistent data set is not only desirable but also imperative. To this end, it is fundamental that the chemical engineer is familiar with the different thermodynamic consistency tests available in the literature, as well as their pitfalls and limitations. In general, it is improbable that a particular data set succeeds in all thermodynamic consistency tests. As noted by Wisniak et al. (2017), there is no unique test capable of providing an absolute answer to the question “Are the experimental data thermodynamically consistent?”. This is exactly why it is important to use a data-informed approach, where the chemical engineer can submit all data sets to several consistency tests, analyze the output, and then decide whether the experimental data set must be rejected or not.

Therefore, in this work, a Python-based approach for thermodynamic consistency tests of VLE (vapor-liquid equilibrium) data of binary mixtures is developed. For the sake of convenience, all consistency tests embedded in the Python-based approach only depend on tabulated physicochemical properties, which are easily found in compendia of physical properties. Unlike some well-known computerized systems for the retrieval, correlation, and prediction of phase equilibria data, this work presents an open-source,



efficient computational tool that helps chemical engineers to make rational use of experimental data in process simulation and model parameterization.

## 2. Theoretical background of the thermodynamic consistency tests

Overall, the Gibbs-Duhem equation plays a very important role in the development of thermodynamic consistency tests. Since one usually neglects the effect of the pressure  $p$  on the activity coefficient  $\gamma_i$ , for a binary mixture at constant temperature  $T$  the Gibbs-Duhem equation becomes

$$x_1 d \ln \gamma_1 + x_2 d \ln \gamma_2 = 0, \quad (1)$$

whose mathematical structure implies that the activity coefficients  $\gamma_i$  are coupled, that is, if one knows the value of  $\gamma_1$  at composition  $x_1$ , then the value of  $\gamma_2$  at composition  $x_2$  is known too. In practical terms, if  $\gamma_1$  and  $\gamma_2$  are determined separately, Equation (1) may be used to check the consistency of the data. Hence, if both  $\gamma_1$  and  $\gamma_2$  satisfy Equation (1) over the range  $0 \leq x_1 \leq 1$ , the data are said to be thermodynamically consistent.

Equation (1) may be rewritten as

$$x_1 \left( \frac{d \ln \gamma_1}{dx_1} \right) + x_2 \left( \frac{d \ln \gamma_2}{dx_1} \right) = 0, \quad (2)$$

which is the main equation of the slope test. Likewise, Equation (1) may be integrated over the range  $0 \leq \ln \gamma_i \leq \ln \gamma_i^\infty$  to give

$$\int_{\ln \gamma_1^\infty}^0 x_1 d \ln \gamma_1 + \int_0^{\ln \gamma_2^\infty} x_2 d \ln \gamma_2 = \underbrace{\int_0^1 \ln \gamma_1 dx_1}_{A_1} + \underbrace{\int_1^0 \ln \gamma_2 dx_2}_{A_2} = 0, \quad (3)$$

where  $\gamma_i^\infty$  is the activity coefficient at infinite dilution. Equation (3) is the basis for the area test that asserts that in a plot of  $\ln \gamma_i$  versus  $x_i$  the quantities  $|A_1|$  and  $|A_2|$  must be equal. Of course, because the measurements of isothermal VLE data are associated to an experimental error the quantities  $|A_1|$  and  $|A_2|$  are in practice not equal. Thus, it is common to assume a more lenient criterion for the thermodynamic consistency: the difference between  $|A_1|$  and  $|A_2|$  should be smaller than 2.0% of the sum  $|A_1| + |A_2|$ . Here, note that the requirement that  $|A_1|$  and  $|A_2|$  be equal is only a necessary condition for the thermodynamic consistency of the experimental data. In fact, it might be the case that some experimental data satisfy Equation (3), but not Equation (2).

Another test that may be applied to isothermal VLE data is the Redlich-Kister test. After some algebra, the Gibbs-Duhem equation gives

$$\ln \gamma_1 dx_1 + \ln \gamma_2 dx_2 = 0, \quad (4)$$

which can be integrated over the range  $0 \leq x_1 \leq 1$  to give

$$\int_0^1 \ln \left( \frac{\gamma_1}{\gamma_2} \right) dx_1 = 0. \quad (5)$$

Equation (5) indicates that in a plot of  $\ln(Y_1/\gamma_2)$  versus  $x_1$  the areas above  $A_A$  and below  $A_B$  the axis  $x_1$  must be equal. However, because of the experimental errors, this criterion is relaxed to

$$D = 100 \frac{|A_A - A_B|}{A_A + A_B} < 10. \quad (6)$$

If, however, the Redlich-Kister test is applied to isobaric VLE data, then the contribution of the excess enthalpy on the excess Gibbs energy  $g^E$  is not negligible. In this case, the criterion of thermodynamic consistency becomes

$$D = \left| 100 \frac{|A_A - A_B|}{A_A + A_B} - 150 \frac{|\Delta T^{max}|}{T^{min}} \right| < 10, \quad (7)$$

where  $\Delta T^{max}$  is the largest difference of temperature observed and  $T^{min}$  is the lowest temperature value in the data set.

A more comprehensive consistency test for VLE data is the  $L$ - $W$  test proposed by Wisniak (1993). The starting point of the  $L$ - $W$  test is the equations  $g^E = RT \sum_i x_i \ln \gamma_i$  and  $\gamma_i = y_i p / x_i p_i^*$ , where  $R$  is the gas constant,  $p_i^*$  is the vapor pressure, and  $y_i$  and  $x_i$  are the mole fraction of the species  $i$  in the vapor and liquid phase, respectively. By restricting the analysis to mixtures composed by species whose boiling points are not too different, so that the Clausius-Clapeyron equation may be used, one obtains after some algebra the following expression

$$\frac{\sum_i x_i \Delta S_i^* T_i^*}{L_i} - T = \frac{g^E - RT \sum_i x_i \ln(y_i/x_i)}{\sum_i x_i \Delta S_i^*}, \quad (8)$$

where  $\Delta S_i^*$  and  $T_i^*$  are the entropy of vaporization and the boiling point of the pure constituents at the pressure  $p$  of the solution. By integrating Equation (8) over  $0 \leq x_1 \leq 1$ , one has:

$$\int_0^1 L_i dx_1 - \int_0^1 W_i dx_1 = 0. \quad (9)$$

Equations (8) and (9) comprise the two equations that the  $L$ - $W$  test relies on. While Equation (8) may be used to check the consistency of every single point of the experimental data set, Equation (9) may be employed to test the consistency of the full data set. According to the  $L$ - $W$  test, a single experimental point is thermodynamically consistent, if the condition  $0.92 < (L_i/W_i) < 1.08$  is fulfilled. In addition, an experimental data set is said to be consistent, if the difference between  $L$  and  $W$  is smaller than 3.0% of the sum  $L + W$ .

### 3. Implementation of the thermodynamic consistency tests

The thermodynamic consistency tests herein discussed were implemented in Python by using the libraries Numpy, Pandas, Scypy, Matplotlib, and Pyplot. Once experimental phase equilibrium data have been imported as a Python dataframe, the user must specify the type of data and provide some pure component parameters. Next, the thermodynamic consistency of the experimental data set is checked and the results are exported as a

worksheet, which should be analyzed by the user. A sketch of the algorithm architecture is shown in Figure 1.

For the sake of illustration, Figure 2 shows the results of thermodynamic consistency tests for experimental VLE data retrieved from the Dortmund Data Bank (2022). In addition to the numerical results exported in the worksheet (Table 1), the results of some thermodynamic consistency tests (e.g. area test, Redlich-Kister test, and  $L$ - $W$  test) may also be analyzed through plots. This approach is particularly interesting to identify possible outliers in the data set and help the user to decide if the data set should be totally or partially rejected. In addition, note that the results of the thermodynamic consistency tests for VLE data strongly depend on the quality of the pure component parameters, the applicable temperature range of the Antoine equation parameters, the amount of available experimental data, and how large the deviations from the ideality are. Regarding this last aspect, thermodynamic consistency tests based on the Gibbs-Duhem equation tend to face some issues when dealing with VLE data of very ideal mixtures. One of the reasons for this behavior is the fact that the measured pressure and temperature values are usually the same order of magnitude as the experimental error and, as a consequence, the output of the thermodynamic consistency tests is somehow imprecise with large values for the parameter  $D$ , as illustrated in Table 1. Thus, in such cases, the results of the thermodynamic consistency tests might be dubious and should be carefully analyzed.

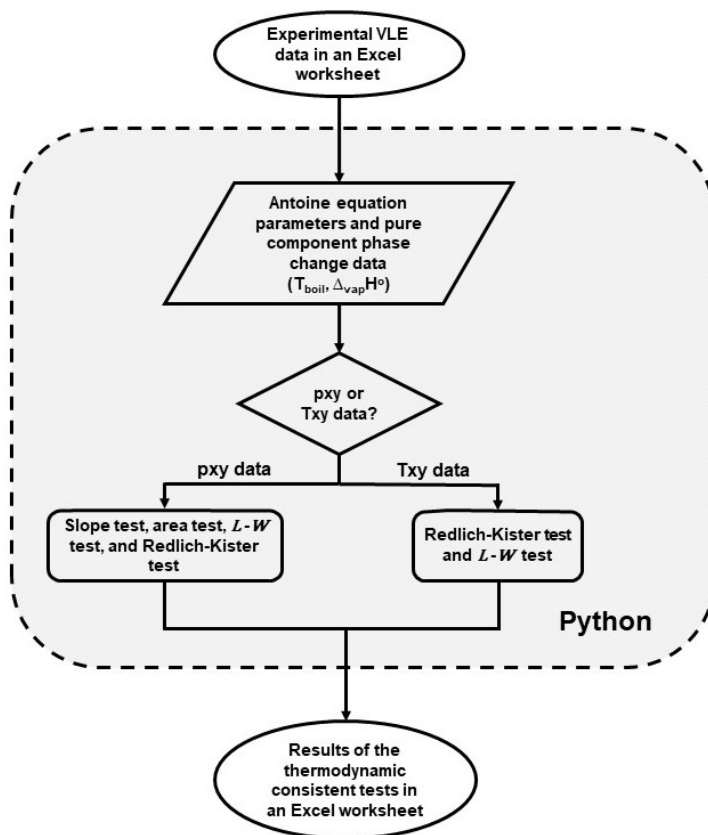


Figure 1. Structure of the proposed algorithm for thermodynamic consistency tests.

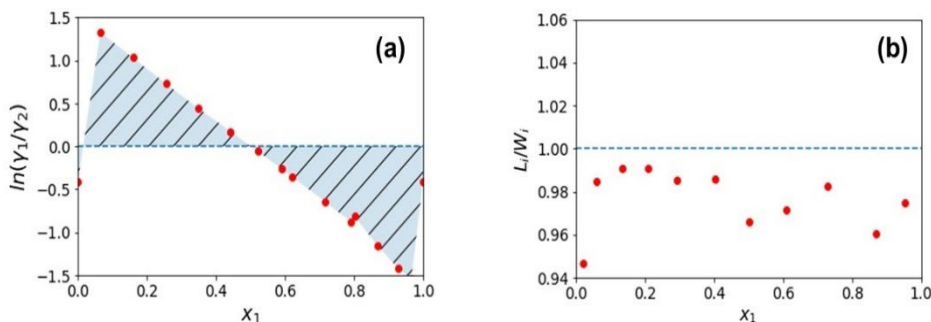


Figure 2. (a) Graphical visualization of the Redlich-Kister test ( $D= 1.48$ ) for the isothermal VLE data of the mixture acetone-hexane at 318.15 K and (b) results of the  $L$ - $W$  test for each point of the isobaric VLE data of the mixture acetone-ethanol at 101.33 kPa.

Table 1. Results of thermodynamic consistency tests for isobaric and isothermal VLE data.

			Thermodynamic Consistency Tests				
			Area	$L$ - $W$		Redlich-Kister	
			$D$	$L$	$W$	$D$	$D$
Butanol-Chlorobenzene	101.33 kPa	Txy	-	5.9	6.1	1.8	5.2
Methanol-Water	97.22 kPa	Txy	-	9.3	8.7	3.0	9.8
Acetone-Ethanol	101.33 kPa	Txy	-	4.4	4.7	3.2	7.1
Benzene-Toluene	333.15 K	pxy	1.6	35.9	2.4	87.7	14.9
Benzene-Toluene	353.15 K	pxy	13.1	16.0	2.3	75.0	11.1
1-Hexene-Hexane	328.15 K	pxy	77.9	11.2	0.3	95.2	11.3
1-Hexene-Hexane	333.15 K	pxy	38.3	4.7	0.2	91.1	5.8
Acetone-Hexane	318.15 K	pxy	10.6	15.4	10.1	21.1	1.5
Chloroform-Benzene	298.20 K	pxy	70.4	33.4	1.1	93.9	15.5
Diethyl ether-Acetonitrile	293.65 K	pxy	0.2	46.2	37.5	10.3	0.5

#### 4. Conclusion and future perspectives

Although checking the reliability of experimental data is a long-established problem, the development of new thermodynamic consistency tests and computational tools for the assessment of VLE experimental data is still the motive of intensive research in chemical engineering thermodynamics. There are some commercial tools (Frenkel et al., (2005)) that were designed to dynamically evaluate the consistency of experimental data, but unfortunately, open-source codes are still non-existent. Then, since the assessment of experimental VLE data is a time-consuming and laborious task, any initiative towards the development of computational tools that help in the critical data-evaluation process is not only welcome but also necessary.

Therefore, in this work, an open-source algorithm written in Python was used to check the thermodynamic consistency of VLE data of binary mixtures. To this end, initially, experimental data are imported as a Python dataframe and then submitted to a series of thermodynamic consistency tests, according to its type. Next, the obtained results are exported as a worksheet that should be carefully analyzed by the user since passing a thermodynamic consistency test does not imply that the data set is accurate. In

fact, as observed in this work, thermodynamic consistency tests based on the Gibbs-Duhem equation seem to be not so efficient when applied to very ideal solutions. One of the reasons for this trend might be related to the fact that the measured pressure and temperature values are usually the same order of magnitude as the experimental error, which leaves room for further research in chemical thermodynamics.

Even though the open-source tool presented in this work is still under development, it may be already used to test the thermodynamic consistency of VLE data. In the near term, other consistency tests for vapor-liquid and solid-liquid equilibria data should be implemented (Kojima test, Fredenslund test, McDermott-Ellis test, Kang test, and Cunico test), as well as tests of internal and external consistency.

## References

- J. Wisniak, J. Ortega, L. Fernández, 2017, A Fresh Look at the Thermodynamic Consistency of Vapour-Liquid Equilibria Data, *The Journal of Chemical Thermodynamics*, v. 105, p. 385-395.
- J. Wisniak, 1993, A New Test for the Thermodynamic Consistency of Vapor-Liquid Equilibrium, *Industrial & Engineering Chemistry Research*, v. 32, p. 1531-1533.
- Dortmund Data Bank, 2022, [www.ddbst.com](http://www.ddbst.com)
- NIST Chemistry WebBook, 2022, NIST Standard Reference Database Number 69, National Institute of Standards and Technology, <https://doi.org/10.18434/T4D303>
- M. Frenkel, R. D. Chirico, V. Diky, X. Yan, Q. Dong, C. Muzny, 2005, ThermoData Engine (TDE): Software Implementation of the Dynamic Data Evaluation Concept, *Journal of Chemical Information and Modeling*, v. 45, p. 816-83.

# Modelling building's life-cycle decarbonization in China: A multi-level and multi-region optimization approach

Chenxi Li,<sup>a</sup> Pei Liu,\* Zheng Li

*<sup>a</sup>State Key Lab of Power Systems, Department of Energy and Power Engineering, Tsinghua-BP Clean Energy Center, Tsinghua University, Beijing 100084, China  
liu\_pei@tsinghua.edu.cn*

## Abstract

The decarbonization of life-cycle emissions is a huge challenge to the building sector because it involves multiple levels. However, previous studies rarely evaluated the potential emission reduction contribution of various technologies in the life cycle, especially China, whose life-cycle CO<sub>2</sub> emissions of the building sector account for over 40 percent of the total emissions. A multi-level and multi-region optimization model is proposed to quantify the life-cycle CO<sub>2</sub> emission reduction path of China's building sector in this paper. Considering both indirect and direct emissions, this model can find the lowest-cost path of CO<sub>2</sub> emission reduction in China's building sector under specified emission scenarios. Results show that extending the building life from 30 years to 50 years has a great impact on emission reduction, leading to a CO<sub>2</sub> emission reduction of more than 50 percent if low-carbon building materials are also used. Moreover, electrification and emission reduction of the power system also have the potential of reducing the CO<sub>2</sub> emissions of buildings by nearly 40 percent, with an end-use electrification rate of more than 70 percent by 2060.

**Keywords:** building sector, decarbonization, optimization, life-cycle, China

## 1. Introduction

The building sector is an important sector which connects the upstream industries and the user terminal, leading to its high life-cycle CO<sub>2</sub> emissions. To achieve the goal of CO<sub>2</sub> mitigation in the whole society, the life-cycle emission reduction of the building sector which connects the upstream and downstream is very important, while the multi-level structure of the building sector brings great challenges to the planning of the life-cycle decarbonization.

Previous studies assessing the CO<sub>2</sub> mitigation of the building sector most focused on one level, such as the material of buildings, the operation of buildings or the reconstruction of a specific house. Li et al. (2017) and Ma et al. (2016) quantified the impacts of decarbonization in China's cement sector and steel sector by using China TIMES model separately. In order to achieve the goal of climate change, the energy demand for heating and cooling of German residential buildings was assessed comprehensively (Olonscheck et al., 2011). A multi-objective optimization model was utilized to find a building retrofit strategy to improve energy efficiency (Asadi et al., 2012). Although the above studies described the CO<sub>2</sub> mitigation pathway of the building sector or one specific house

detailedly, the research is still insufficient because of a lack of life-cycle analyzing. Each level of the building sector, material, construction, operation and demolition, affects each other, making it essential to carry out life-cycle CO<sub>2</sub> mitigation planning.

However, most life-cycle research on the building sector emphasized the accounting of historical CO<sub>2</sub> emissions instead of the mitigation potential of different low-carbon methods. Yang et al. (2018), Clabeaux et al. (2020), and Atmaca et al. (2022) counted the life-cycle emissions of a specific building or a building complex in China, USA and Turkey separately.

Overall, previous studies are rarely able to evaluate the emission reduction potential of buildings in the life cycle. Therefore, a multi-level and a multi-region optimization model is established to plan the life-cycle mitigation pathway of the building sector. The model is applied to the building sector of China, whose CO<sub>2</sub> emissions surpasses 4 billion tonnes, accounting for nearly 20 percent of the total CO<sub>2</sub> emissions all over the world.

## 2. Methodology

### 2.1. Model structure and Assumptions

The lifecycle of buildings includes material, construction, operation and demolition, which is shown in Figure 1. In this multi-level and multi-region optimization model, three main building materials and three main building structures are considered. Buildings are divided into three categories according to their use, and six energy demands are regarded. In this model, 40 technologies can be applied to the decarbonization of building sector, which are listed in Table 1.

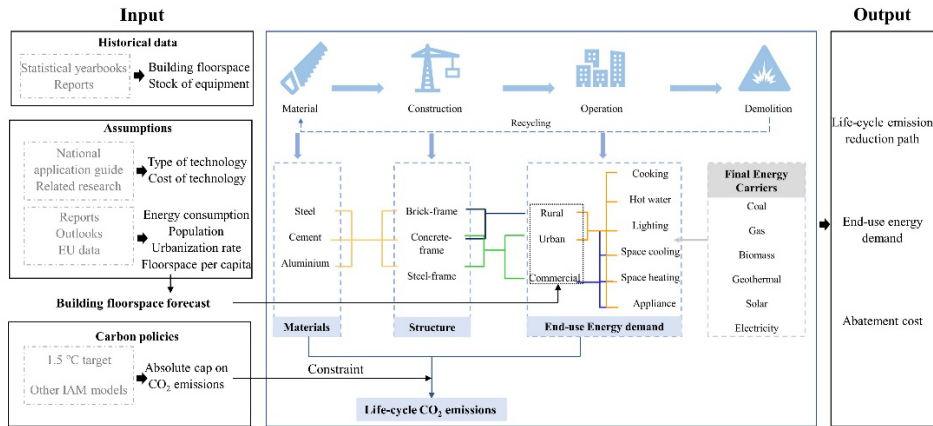


Figure 1. The structure of multi-level and multi-region optimization model of buildings

### 2.2. Mathematical formulation

Mathematical equations of the optimization model are presented in this model, including objective function and model constraints. Eight sets,  $t$ ,  $mat$ ,  $u$ ,  $str$ ,  $tech$ ,  $r$ ,  $op$ ,  $ref$  stand for time, material, type of utilization, structure, technology, region, energy demand and insulation of buildings.

#### 2.2.1. Objective function

The objective function of this model is to minimize the total cost of the building sector in its lifetime from 2020 to 2060. The total cost is composed of material costs, end-use device costs, central heating costs, insulation material costs and fuel costs, as Equations (1)-(2) shows.

$$objective = \sum_{t=2020}^{2060} \frac{sumcost_t}{(1+i)^{t-2020}} \quad (1)$$

$$sumcost_t = matc_t + devicec_t + heatc_t + refc_t + fuelc_t \quad (2)$$

Table 1. The technologies used in the model

Type of technology	Name of technology
Low carbonization of building materials	Far-infrared ceramic coating mixture, Electric arc furnace, Hydrogen steelmaking, New cement clinker, Solid regenerative fuel, Carbon capture and storage, Carbon capture and usage, 600kA potcell
Changing of building structure	Brick-frame, Concrete-frame, Steel-frame
Decarbonization of central heating	Coal-fired boiler, Gas-fired boiler, Coal-fired cogeneration, Biomass boiler, Geothermal, Solar heating, Heat pump
Insulation of buildings	Single glazing windows, 2bl glazing windows, Low-e windows, Stone wool insulation materials, Glass wool insulation materials, Expanded polystyrene, Sprayed polyurethane
	Cooking
	Electric water heater, Solar thermal water heater, Gas-fired water heater
End-use electrification	Hot water
	Lighting
	Incandescent lamp, Fluorescent lamp
	Space cooling
	AC-I, AC-II, AC-III
	Decentralized space heating
	Coal-burning fireplace, Solar heating, Heat pump

### 2.2.2. Material constrains

The demand of various materials depends on the newly-built area of different structures, which can be described by Equation (3). Although there may be demolition areas and new areas every year, it is still necessary to ensure that the total area is not less than the required area, listed as Equation (4). The CO<sub>2</sub> emissions caused by building materials are calculated by emission factors and the costs of materials are computed by unit price.

$$matd_{t,mat} = \sum_{str,u} narea_{t,u,str} * MATPRO_{str,mat} \quad (3)$$

$$\sum_{str}(area_{t-1,u,str} + narea_{t,u,str} - darea_{t,u,str}) \geq AREAD_{t,u} \quad (4)$$

$$matco2_t = \sum_{mat,tech} matp_{t,mat,tech} * TECHCO2_{t,mat,tech} \quad (5)$$

$$matc_t = \sum_{mat,tech} matp_{t,mat,tech} * MATPC_{t,mat,tech} \quad (6)$$

### 2.2.3. End-use demand constrains

The end-use demand can be divided into six categories: cooking, hot water, lighting, space cooling, space heating and electricity. Decentralized space heating and central heating constitute space heating. The decentralized equipment at the user end is supposed to meet the end-use demand, including cooking, hot water, lighting, space cooling and decentralized space heating, according to Equation (7). Due to the use of thermal insulation materials, the heating and cold demands of buildings will be reduced accordingly, as described in Equation (8). The coefficient  $COEFC_{r,ref}$  and  $COEFH_{r,ref}$  is related to the climate of the area where the building is located, and is estimated according to the average temperature. The heat provided by central heating should not be less than the demand, with the operation and heat supply of central heating boilers meeting Equation (9). The total power demand is the sum of all appliance electricity consumption, as shown in Equation (10). The consumption of fossil energy, biomass and other renewable energy is calculated according to the actual operation time of the equipment, as shown in Equation (11). The CO<sub>2</sub> emissions from fossil energy and electricity are computed by emission factors, which are described by Equation (12).



$$ENDUSED_{t,r,op} \leq \sum_{u,de,f} ndevice_{t,r,u,de} * OH_{t,r,u,de} * CONSUMPTION_{t,de,f} * CV_f \quad (7)$$

$$endused_{t,r,cooling} = ENDUSED_{t,r,op} - \sum_{ref} refarea_{t,ref} * COEFC_{r,ref} \quad (8a)$$

$$endused_{t,r,heating} = ENDUSED_{t,r,op} - \sum_{ref} refarea_{t,ref} * COEFH_{r,ref} \quad (8b)$$

$$\sum_{th} centralheat_{t,r,th} \leq CENTRALHEATD_{t,r} \quad (9a)$$

$$nbheat_{t,r,th} * OHMIN_{t,r,th} \leq heat_{t,r,th} \leq nbheat_{t,r,th} * OHMAX_{t,r,th} \quad (9b)$$

$$elec_{t,r} = ELECD_{appliance,t,r} + devicefuel_{t,elec} \quad (10)$$

$$devicefuel_{t,f} = \sum_{u,de,r} ndevice_{t,r,u,de} * OH_{t,r,u,de} * DCON_{t,de,f} \quad (11)$$

$$endco2_t = \sum_f devicef_{t,f} * FCO2_f \quad (12)$$

#### 2.2.4. CO<sub>2</sub> emissions constrains

The life-cycle CO<sub>2</sub> emissions of the building sector are composed of material emissions and end-use emissions, which can be described as Equation (13). The total emissions in the life cycle are supposed to meet the upper limit requirements of CO<sub>2</sub> emissions, as shown in Equation (14).

$$sumco2_t = matco2_t + endco2_t \quad (13)$$

$$sumco2_t \leq UPPERCO2_t \quad (14)$$

The model is implemented in General Algebraic Modelling System (GAMS) by using the CPLEX solver.

### 3. Case Study

The proposed model is implied in a case study of the building sector of China, which is divided into 9 regions according to the climate and data availability, as shown in Figure 2. The average life of buildings in China is about 30 years. It is assumed that the life of buildings are extended to 50 years to achieve the decarbonization goals. Several basic assumptions are listed in Table 2.

Table 2. Basic assumptions

Year	2030	2040	2050	2060
Floor area/100 million m <sup>2</sup> (ICCS, 2021)	806	844	861	828
Population/100 million	14.0	13.4	12.6	11.3
Emission factor of electricity /kg·kWh <sup>-1</sup> BAU	0.36	0.24	0.21	0.21
(Song et al., 2022) Decarbonization	0.34	0.13	0.06	0
Total energy demand/100 million tce (Chen et al., 2015)	7.1	8.0	8.8	8.5

#### 3.1. Decarbonization of building materials

Selecting low-carbon materials enables the building sector to lower its life cycle CO<sub>2</sub> emissions. According to the proposed model, average carbon emission factor of materials given in Figure 3 can be referred to elect low-carbon materials when new buildings are constructed in the future. In 2060, the average CO<sub>2</sub> emissions factor of steel reduces to half of that of 2020, and that of cement reduces to one-third of that of 2020. Due to the carbon neutralization of the power system, the average CO<sub>2</sub> emissions factor of aluminium is close to 0.



Figure 2 Geographic division of China

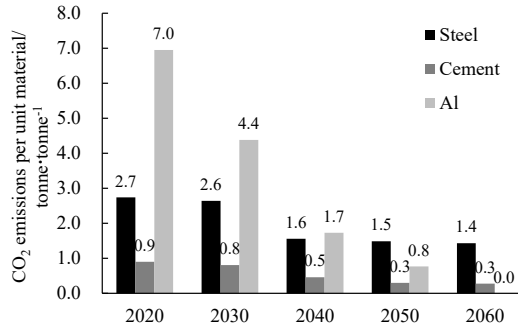


Figure 3 Emission factor of building materials

### 3.2. Decarbonization of building operation

Owing to the application of thermal insulation materials, heat loss is reduced in each region, especially after 2040, which is depicted in Figure 4. Because of the low temperature and long heating period, the application of thermal insulation materials in the northeast is significantly earlier than that in other regions. Despite rare heating demand in South China, thermal insulation materials are supposed to be used after 2050, which can reduce the demand for refrigeration, beneficial to the decarbonization.

As can be seen from Figure 5, which shows the consumption of primary energy, there is a peak both in coal and natural gas. The maximum coal consumption occurs before 2025, which is not much higher than that in 2020. The maximum gas consumption is supposed to appear between 2040 and 2045, with a peak of 250bcm. The electricity demand rises steadily until 2050. Although the electricity demand decreases after 2050 because of the shrinking population and total energy demand, the electrification rate gradually increases from 2020 to 2060, and exceeds 70% in 2060 according to Table 3.

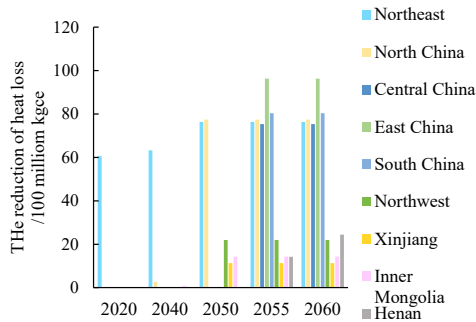


Figure 4 The reduction of heat loss

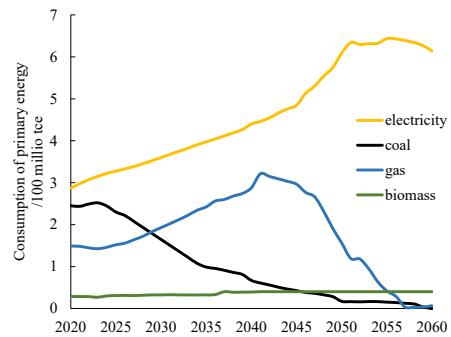


Figure 5 Fuel consumption during operation

Table 3. Proportion of clean energy

Year	2020	2030	2040	2050	2060
End-use Penetration of renewable energy	4.6%	5.0%	5.3%	12.4%	26.7%
End-use electrification rate	40.2%	27.7%	52.4%	68.2%	72.5%

### 3.3. Comparison of life cycle decarbonization potential in the building sector

Table 4 illustrates the decarbonization potential of various emission reduction methods. It is apparent from this table that extending the average lifespan of Chinese buildings from 30 years to 50 years can significantly reduce the life-cycle emissions, due to the demand for building materials being greatly reduced. This mitigation measure can be achieved by government order. Combined with low-carbon building materials, the

contribution rate of CO<sub>2</sub> reduction in non-operational processes of buildings is higher than 50%. What's more, although previous studies have implied the importance of building reconstruction (Motalebi et al., 2022), only the replacement of windows and the addition of building insulation layers contribute little to decarbonization pursuant to our study. In addition, even if there are no measures taken by the building sector, one fifth of the life-cycle CO<sub>2</sub> emissions of the building sector can be reduced by the power sector. The decarbonization of electricity and the electrification of the building contribute nearly 40% to the reduction of emissions in the building sector, meaning that support for electrification is very necessary for the Chinese government to the public to realize the low-carbon transition of the building sector at a lower cost.

#### 4. Conclusions

This paper proposes a multi-level and multi-region optimization model to analyze the life-cycle decarbonization pathway of the building sector and China is taken as a case study. It can be found that life extension plays a significant role in buildings' life-cycle CO<sub>2</sub> emissions reduction by lowering the demand for building materials. During the operation stage, natural gas becomes one of the main fuels in the next 20 years and the consumption of natural gas declines rapidly, which is consistent with the current policy of China. Although biomass-related technologies are relatively mature, the decarbonization transition of China's building sector cannot be entirely dependent on biomass due to lack of biomass resources. The improvement of end-use electrification rate is essential for the emission mitigation of buildings, which is supposed to be increased from 40% in 2020 to 70% in 2060. Therefore, to achieve life-cycle CO<sub>2</sub> reduction in the building sector in China, attention to building life extension and electrification must be paid, meaning that reasonable urban planning and electrification reform of buildings should be the major work of China's building sector in the future.

Table 4. The decarbonization potential of various emission reduction methods

Methods	CO <sub>2</sub> reduction in 2060 /100 million tonnes	Proportion	Total CO <sub>2</sub> reduction in 40 years /100 million tonnes	Proportion
Life extension	3.4	14.5%	268.7	41.3%
The use of low-carbon materials	1.7	7.2%	70.6	10.8%
Decarbonization of central heating	4.0	17.0%	62.9	9.7%
Insulation	0.2	0.9%	3.4	0.5%
Electrification of distributed heating, cooking and hot water	6.1	26.0%	109.4	16.8%
Decarbonization of electricity	8.1	34.4%	135.8	20.9%

#### References

- E. Asadi, M. G. da Silva, C. H. Antunes, L. Dias, 2012, Multi-objective optimization for building retrofit strategies: A model and an application, *Energy and Buildings*, 44, 81-87.
- A. Atmaca, N. Atmaca, 2022, Carbon footprint assessment of residential buildings, a review and a case study in Turkey, *Journal of Cleaner Production*, 340, 130691.
- P. Chen, Y. Da, S. Guo, S. Hu, J. Yi, 2015, Building energy use in China: Ceiling and scenario, *Energy and Buildings*, 102, 307-316.

R. Clabeaux, M. Carbajales-Dale, D. Ladner, T. Walker, 2020, Assessing the carbon footprint of a university campus using a life cycle assessment approach, *Journal of Cleaner Production*, 273, 122600.

ICCSA, 2021, *The Comprehension of Carbon Neutral: The Low Carbon Development Action Roadmap of China during 2020–2050*, China Citic Press, Beijing, China.

N. Li, D. Ma, W. Chen, 2017, Quantifying the impacts of decarbonisation in China's cement sector: A perspective from an integrated assessment approach, *Applied Energy*, 185, 1840-1848.

D. Ma, W. Chen, X. Yin, L. Wang, 2016, Quantifying the co-benefits of decarbonisation in China's steel sector: An integrated assessment approach, *Applied Energy*, 162, 1225-1237.

M. Motalebi, A. Rashidi, M. M. Nasiri, 2022, Optimization and BIM-based lifecycle assessment integration for energy efficiency retrofit of buildings, *Journal of Building Engineering*, 49, 104022.

M. Olonscheck, A. Holsten, J. P. Kropp, 2011, Heating and cooling energy demand and related emissions of the German residential building stock under climate change, *Energy Policy*, 39, 9, 4795-4806.

S. Song, T. Li, P. Liu, Z. Li, 2022, The transition pathway of energy supply systems towards carbon neutrality based on a multi-regional energy infrastructure planning approach: A case study of China, *Energy*, 238, 122037.

X. Yang, M. Hu, J. Wu, B. Zhao, 2018, Building-information-modeling enabled life cycle assessment, a case study on carbon footprint accounting for a residential building in China, *Journal of Cleaner Production*, 183, 729-743.



# Automated Kinetic Model Discovery – A Methodological Framework

Miguel Ángel de Carvalho Servia <sup>a</sup>, Ilya Orson Sandoval <sup>a</sup>, Dongda Zhang <sup>b,\*</sup>,  
Klaus Hellgardt <sup>a</sup>, King Kuok (Mimi) Hii <sup>c</sup>, Ehecatl Antonio del Rio Chanona <sup>a,\*</sup>

<sup>a</sup> *Department of Chemical Engineering, Imperial College London, South Kensington, London SW7 2AZ, UK*

<sup>b</sup> *Department of Chemical Engineering, The University of Manchester, M13 9PL, UK*

<sup>c</sup> *Department of Chemistry, Imperial College London, 82, Wood Lane, London W12 0BZ, UK*

\* *Corresponding authors: [dongda.zhang@manchester.ac.uk](mailto:dongda.zhang@manchester.ac.uk); [a.del-rio-chanona@imperial.ac.uk](mailto:a.del-rio-chanona@imperial.ac.uk)*

## Abstract

The industrialization of catalytic processes benefits strongly from kinetic models for optimization and control purposes. Nevertheless, mechanistic models are difficult to construct; data-driven and hybrid models lack interpretability and the flexibility to leverage physical knowledge. Thus, a different approach called automated knowledge discovery has been recently popularized. Existing methods in literature suffer from important drawbacks: necessitating assumptions about model structures, a lack of model selection automation, and sensitivity to noise. To overcome these challenges, the present work constructs a methodological framework for the automated generation of catalytic kinetic models. We leverage symbolic regression for model generation, a hybrid optimization algorithm for parameter estimation, and a robust criterion for model selection. The framework is tested with an illustrative isomerization case study, where it showcases the ability to retrieve the underlying kinetic model with a limited amount of noisy data from the catalytic system.

**Keywords:** catalysis, kinetic model generation, automated knowledge discovery, information criteria, machine learning.

## 1. Introduction

The industrialization of catalytic processes requires kinetic models, which are a mathematical representation of the dynamical trajectories of a chemical system, typically presented in the form of ordinary differential equations. Kinetic models can be constructed using any of the three classical modelling paradigms: mechanistic (white-box), data-driven (black-box) or hybrid (grey-box). Mechanistic models are constructed by using existing fundamental laws, such as mass, energy, and momentum balances, making them interpretable and extrapolatory (Baker, 2018). However, the construction of these models can be time-consuming and experimentally expensive. Data-driven models do not use any fundamental laws for their construction. Typically, these models are constructed by solely using collected data of a system, making their construction time-efficient but experimentally expensive. Furthermore, these models generally have poor extrapolatory abilities and are uninterpretable, which may classify their usage for certain industrial scenarios as unsafe. In comparison, hybrid models aim to exploit the advantages

of mechanistic and data-driven modelling. Hybrid models are composed of two parts: a mechanistic backbone and a data-driven block which aims to improve the backbone's fit. Generally, hybrid models retain the extrapolation capabilities of a mechanistic model and the flexibility and ease of construction of a data-driven model.

Hybrid modelling offers an elegant solution to the problems posed by the other classical modelling paradigms, but not an optimal one. A better solution would be to utilize existing data to automatically generate and select mechanistic models by exploiting state-of-the-art statistical and machine learning methods. In this way, the benefits of mechanistic models are maintained, whilst some of their drawbacks are eliminated. This modelling paradigm has been coined *automated knowledge discovery* and has recently gained popularity. In current literature, many automated knowledge discovery frameworks have been developed, such as: the ALAMO approach (Wilson, 2017), the SINDy algorithm (Brunton, 2016), and symbolic regression using genetic programming (Koza, 1994). These frameworks have three drawbacks that limit their ability to retrieve underlying ground-truth models, and consequently, their real-world applicability. Firstly, they necessitate substantial structural assumptions of the underlying data-generating model. This is particularly true of ALAMO and SINDy, as a design matrix needs to be constructed for their execution (Wilson, 2017; Brunton, 2016). Secondly, they lack a motivated and rigorous model selection routine (Wilson, 2017; Brunton, 2016; Koza, 1994). Lastly, they are sensitive to noisy data (Wilson, 2017; Brunton, 2016; Koza, 1994). Hence, this work aims to overcome these obstacles by constructing a generalizable and robust methodological framework that integrates a rigorous model selection routine for the automated kinetic model discovery of catalytic systems.

The rest of the paper will be structured as follows: in Section 2, the proposed automated knowledge discovery framework is motivated and described in detail; in Section 3, the illustrative case study used to showcase the capabilities of the proposed methodology is introduced; in Section 4, the results of the study are amply discussed; in Section 5, the key findings are presented, and the outlook of the research is explored.

## 2. Methodological Framework

Given the objective of automated knowledge discovery, the proposed methodological framework constitutes three main stages: model generation, model refinement, and model selection and discrimination.

Symbolic regression using genetic programming (SR-GP) is often considered the most generalizable and reliable model generation method found in literature, as this method does not require the construction of a design matrix. In other words, assumptions regarding an underlying ground-truth model are minimal. The basic concept of SR-GP is to specify a set of state variables and operators (e.g.: '+', '/') that may be present in the underlying mechanistic model. With this, an initial population of models can be constructed, and based on Darwin's theory of evolution, the best models – based on a specified performance metric – are evolved via genetic operations (e.g.: crossover and mutation), and the worst models are discarded. This process is iterated until convergence is achieved or a termination criterion is met. As such, for the model generation stage, we propose a weak reformulation of SR-GP, where models generated represent a mapping between input and output variables in the derivative hyperspace. In other words, this reformulation generates functions dependent on the state variables that can map the output via an integration step. In this way, weak SR-GP receives the dynamic trajectories of

concentration as inputs, proposes rate models ( $r = f(C_1, C_2, \dots)$  where  $C_i$  is the concentration of species  $i \in \mathbb{Z}^+$ ) integrates the rate models with respect to time at each given time-step where concentration data is available ( $\int r dt = \int \frac{1}{v_i} \frac{dC_i}{dt} dt = \frac{1}{v_i} C_i(t)$ ), and compares the results from the integration with the original dataset.

Once promising rate models are proposed by the weak reformulation of SR-GP, model refinement needs to take place. Specifically, the model refinement comes as a parameter estimation problem, where the error between the model's response and the data are minimized by finding the best set of kinetic parameters. This optimization problem is solved by deploying a hybrid optimization algorithm composed of an explorative phase, carried out by an artificial bee colony (ABC) algorithm, and an exploitative phase carried out by a limited memory Broyden Fletcher Goldfarb Shanno (LBFGS) algorithm. The former is a stochastic optimization algorithm shown to have excellent explorative characteristics (Cho, 2021), while the latter is a gradient-based optimization algorithm shown to have excellent exploitative characteristics (Malouf, 2002). The objective function used for the parameter estimation was the negative log-likelihood (NLL), presented below:

$$f(\boldsymbol{\theta}) = \sum_{i,j,k} \left[ \frac{(C_{i,j,k} - y_i(\boldsymbol{\theta}, \mathbf{x}_0)_{j,k})^2}{2\hat{\sigma}_i^2} - \log \left( \frac{1}{\sqrt{2\pi\hat{\sigma}_i^2}} \right) \right] \quad (1)$$

where  $C_{i,j,k}$  is the simulated concentration (i.e.: in-silico data) of species  $i \in S$  for dataset  $j \in D$  at time  $k \in T$ , where  $S$ ,  $D$ , and  $T$  represent the species set, data set and time set, respectively;  $y_i(\boldsymbol{\theta}, \mathbf{x}_0)_{j,k}$  is the concentration of species  $i \in S$  for dataset  $j \in D$  at time  $k \in T$  proposed by an arbitrary model which is dependent on its parameters  $\boldsymbol{\theta}$  and the initial conditions  $\mathbf{x}_0$ ;  $\hat{\sigma}_i^2$  is the variance of the noise that we assume the concentrations of species  $i \in S$  has. Once all the promising models have been explicitly optimized, the models are ranked based on their Akaike information criterion (AIC) values. The AIC was selected after a thorough analysis of the performance of different criteria under several conditions (e.g.: different amounts of additive noise, different amounts of data). This analysis concluded that AIC has a statistically higher probability of selecting the correct data-generating rate model than the other criteria tested.

After the ranking of the optimized models has been established based on AIC values, the Akaike weights of each model,  $w_i(\boldsymbol{\theta}_i, \mathbf{x}_0)$ , (i.e.: a statistical measure that quantifies the probability of a model structure being the correct one based on AIC values) are also evaluated. Provided a user-defined tolerance  $\epsilon \in \mathbb{R}^+$ , if  $|w_1(\boldsymbol{\theta}_1, \mathbf{x}_0) - w_2(\boldsymbol{\theta}_2, \mathbf{x}_0)| \leq \epsilon$ , then the modeler cannot be confident enough of the AIC selection, extra discriminatory experiments should be performed, and the process repeated. The discriminatory experiment can be determined by solving the Hunter-Reiner criterion, which aims at finding the experimental conditions that will maximize the difference between the response of two competing models. If  $|w_1(\boldsymbol{\theta}_1, \mathbf{x}_0) - w_2(\boldsymbol{\theta}_2, \mathbf{x}_0)| > \epsilon$ , then the modeler can be confident of the AIC selection and should integrate the proposed rate model and compare it with the dynamic trajectories of the concentrations. If the results are satisfactory, the methodological framework is terminated. Otherwise, further data should be collected by applying (model-based) design of experiments – a strategy aimed at computing maximally informative experiments. For clarity and simplicity, the flowchart of the proposed methodological framework is presented in Figure 1.



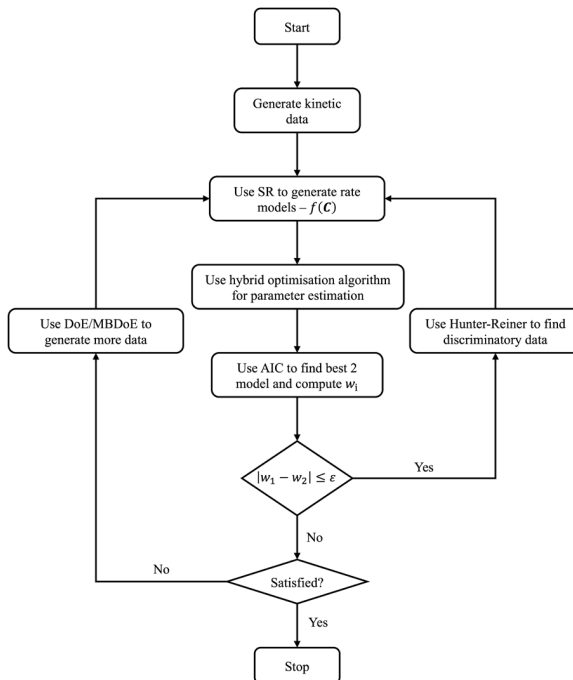


Figure 1: The flowchart of the proposed methodological framework featuring a weak reformulation of SR-GP

### 3. Kinetic Case Study

To showcase the performance of the proposed methodology, an illustrative case study of an isomerization reaction was chosen, where A is transformed to B reversibly over a catalytic active site. The reaction mechanism and the King-Altman graph of the case study are displayed in Figure 2(a).

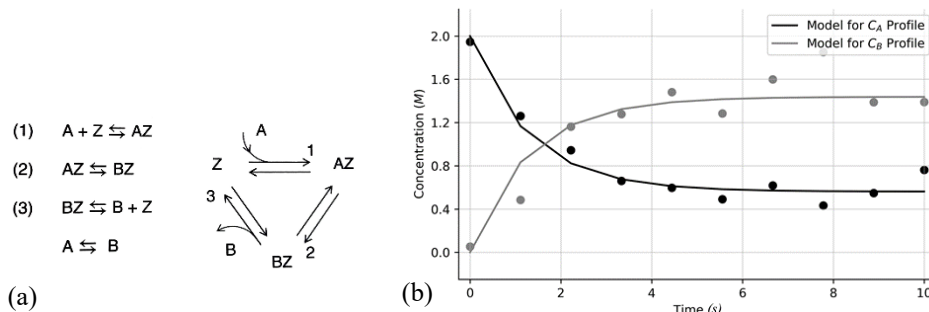


Figure 2: (a) Mechanism and King-Altman graph of illustrative isomerization reaction (Marin, 2019); (b) response of the chosen rate model for one of the three computational experiments

Given the presented mechanism, a rate model was derived and is shown in Equation 2. A detailed derivation of the rate model is found in (Marin, 2019).

$$r = -\frac{dC_A}{dt} = \frac{dC_B}{dt} = \frac{k_A C_A - k_B C_B}{k_C C_A + k_D C_B + k_E} \quad (2)$$

In Equation 2,  $C_A$  and  $C_B$  represent the concentration of reactant A and product B, respectively. The kinetic parameters of the rate model are represented by  $k_i$  where  $i \in$

$\{A, B, \dots, E\}$ . To generate the necessary data to test the proposed framework, three computational experiments were carried out, each with different initial conditions. The simulations were run with the following initial conditions (in molar units):  $(C_{A0}, C_{B0}) \in \{(2,0), (6,1), (10,2)\}$ . For each simulation, the concentration of the reactant and product are recorded at evenly spaced intervals between time  $t = 0$  s and  $t = 10$  s. For the simulations, the kinetic parameters were defined as:  $k_A = 7 \text{ M s}^{-2}$ ,  $k_B = 3 \text{ M s}^{-2}$ ,  $k_C = 4 \text{ s}^{-1}$ ,  $k_D = 2 \text{ s}^{-1}$  and  $k_E = 6 \text{ M s}^{-1}$ . To ensure that the generated data is as realistic as possible, Gaussian noise was added to the simulation results. The defined Gaussian noise had zero mean and a standard deviation of 0.13 for the concentrations of A, and 0.21 for the concentrations of B (the standard deviations represent 5% of the mean of the simulated concentrations of A and B). For the parameter estimation task, it would be futile to assume that as modellers the exact variance of the noise would be known. Thus, a conservative assumption was made by setting  $\hat{\sigma}_A = 0.26$  and  $\hat{\sigma}_B = 0.42$ , (i.e.: assuming a standard deviation of 10% of the mean of the simulated concentrations of A and B).

#### 4. Results and Discussion

As explained in Section 2, the proposed methodological framework implements a weak reformulation of SR-GP which automatically evolves rate models by integrating and comparing them with the concentration data available. Note that the law of conservation of mass is satisfied by construction under this integrating scheme. Departing from the kinetic data generated, the SR algorithm was used to generate rate models. The expression construction rules exclusively included the arithmetic operators ‘+’, ‘-’, ‘×’ and ‘/’, since rates including other operators are less common. The rate models generated were a function of the species whose concentrations were measured,  $r(C_A, C_B)$  in this case. The best expression proposed by the weak SR-GP, sorted by degree of complexity (i.e.: the number of operators and variables), are shown in Table 1.

The estimation of each kinetic parameter,  $k_i$  where  $i \in \{1, 2, \dots, 6\}$ , was carried out as explained previously: NLL was used as the objective function where the assumed variances for NLL are double of the real variances used to generate the Gaussian noise, and this is solved by using the hybrid algorithm comprised of the ABC and the LBFGS algorithms. The NLL, AIC and Akaike weight (AW) values for each of the weak SR-GP proposed and optimized models are also presented in Table 1 below. This shows that  $r_6$  and  $r_3$  are the best and second-best models for the given data, and by setting an Akaike weight tolerance of  $\epsilon = 0.05$ , we can be confident in the AIC’s choice that  $r_6$  is the best model to represent the dynamical catalytic reactive system under investigation, without requiring further discriminatory data. The response of the selected model is shown in Figure 2(b). For the sake of brevity, only one of the experiments is presented. The final rate model output by the proposed framework had a near identical structure to the data-generating one, differing only by a single parameter in the numerator. However, the hybrid optimization algorithm used for parameter estimation was able to determine that the extra parameter was practically zero, and therefore may be non-existent in the actual model. The discrepancy in other parameters may be attributed to the additive Gaussian noise in the data. Table 2 details the estimated kinetic parameter values and the original values of the true data-generating rate model.

Table 1: The NLL value, the AIC values and the Akaike weight (AW) values of all weak SR-GP proposed rate models

Model	NLL Value	AIC Value	AW Value
$r_1 = k_1$	540.410	1082.821	0.000

$r_2 = k_1 C_A$	53.857	109.715	0.000
$r_3 = \frac{k_1 C_A - k_2 C_B}{k_3 C_A}$	-0.831	4.337	0.244
$r_4 = \frac{k_1 C_A - k_2 C_B - k_3}{k_4 C_A}$	-0.988	6.025	0.105
$r_5 = \frac{k_1 C_A - k_2 C_B - k_3}{k_4 C_A + k_5}$	-2.724	4.551	0.219
$r_6 = \frac{k_1 C_A - k_2 C_B - k_3}{k_4 C_A + k_5 C_B + k_6}$	-4.275	3.450	0.380
$r_7 = \frac{k_1 C_A^2 C_B - k_2 C_A C_B^2 - k_3 C_A + k_4 C_B}{k_5 C_A^2 C_B - k_6 C_A}$	-2.289	7.422	0.052

Table 2: The estimated kinetic parameters from the chosen rate model proposed by proposed framework and the original kinetic parameter's values

Kinetic Parameter	Estimated Value	True Value
$k_1, k_A$	5.557 M s <sup>-2</sup>	7.000 M s <sup>-2</sup>
$k_2, k_B$	2.335 M s <sup>-2</sup>	3.000 M s <sup>-2</sup>
$k_3$	0.000 M <sup>2</sup> s <sup>-2</sup>	N/a
$k_4, k_C$	3.114 s <sup>-1</sup>	4.000 s <sup>-1</sup>
$k_5, k_D$	1.379 s <sup>-1</sup>	2.000 s <sup>-1</sup>
$k_6, k_E$	6.033 M s <sup>-1</sup>	6.000 M s <sup>-1</sup>

## 5. Conclusions and Outlook

In this work, we proposed an automated knowledge discovery framework for the generation of catalytic kinetic models from data. A combination of a reformulation of SR-GP, parameter estimation through hybrid optimization, and a rigorous model selection and discrimination routine alleviate drawbacks of alternative automated knowledge discovery methods in literature. The presented case study showcases the framework's ability to retrieve the underlying kinetic model of a catalytic system with a limited amount of noisy data. Future work will pursue strategies to incorporate physical constraints into the model generation, to increase its efficiency through physical guidance.

## References

- Ruth E. Baker, Jose-Maria Peña, Jayaratnam Jayamohan, and Antoine Jérusalem, may 2018, Mechanistic models versus machine learning, a fight worth fighting for the biological community? *Biology Letters*, 14(5):20170660, pages 2
- Zachary T. Wilson and Nikolaos V. Sahinidis, nov 2017, The ALAMO approach to machine learning. *Computers & Chemical Engineering*, 106:785–795, pages 3, 4, 7, 8, 9, 10
- Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz, mar 2016, Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937, pages 3, 4, 10, 12, 13
- JohnR. Koza, jun 1994, Genetic programming as a means for programming computers by natural selection. *Statistics and Computing*, 4(2), pages 14
- Bovinille Anye Cho, Miguel Ángel Carvalho Servia, Ehecatl Antonio del Río Chanona, Robin Smith, and Dongda Zhang, mar 2021, Synergising biomass growth kinetics and transport mechanisms to simulate light/dark cycle effects on photo-production systems. *Biotechnology and Bioengineering*, 118(5):1932–1942, pages 31
- Robert Malouf, 2002, A comparison of algorithms for maximum entropy parameter estimation. In *proceeding of the 6th conference on Natural language learning - COLING-02*. Association for Computational Linguistics, pages 31
- Guy B Marin, Gregory S Yablonsky, and Denis Constales, 2019, *Kinetics of chemical reactions: decoding complexity*. John Wiley & Sons, pages 36

# Assessing process systems models for pharmaceutical development

Margherita Geremia<sup>a</sup>, Samir Diab<sup>b</sup>, Charalampos Christodoulou<sup>c</sup>,  
Gabriele Bano<sup>d</sup>, Massimiliano Barolo<sup>a</sup>, Fabrizio Bezzo<sup>a,\*</sup>

<sup>a</sup> *CAPE-Lab – Computer-Aided Process Engineering Laboratory. Department of Industrial Engineering, University of Padova, via Marzolo 9, 35131 Padova (PD), Italy*

<sup>b</sup> *GlaxoSmithKline, Park Road, Ware SG12 0DP, United Kingdom*

<sup>c</sup> *GlaxoSmithKline, Gunnels Wood Road, Stevenage SG1 2NY, United Kingdom*

<sup>d</sup> *GlaxoSmithKline, 1250 S Collegeville Rd, Collegeville (PA) 19426, United States*

\* *Corresponding author e-mail: [fabrizio.bezzo@unipd.it](mailto:fabrizio.bezzo@unipd.it)*

## Abstract

Quantitative models are useful tools to accelerate the development of pharmaceutical processes. The assessment of the predictive capability of such models is fundamental to enhance their usage in a systematic way, particularly when the focus is on flowsheet models. In this study, we propose a systematic procedure that combines techniques that are typically used in different modeling contexts – namely, global sensitivity analysis, model-based design of experiments, and data reduction by means of multivariate statistical methods, with the advantage of enhancing readability and interpretation. The methodology is effective for the assessment of model fidelity and can support practitioners in the development of pharmaceutical processes. A direct compression systems model for manufacturing oral solid dosage products is used as a case study. Results show that just a subset of model parameters require precise estimation to meet the target critical quality attribute, and that calibrating flowsheet models on a unit-by-unit basis may be unnecessary when the focus is on one final quality attribute.

**Keywords:** systems model, pharmaceutical development, model identifiability, parameter estimation, model fidelity, pharmaceutical engineering.

## 1. Introduction

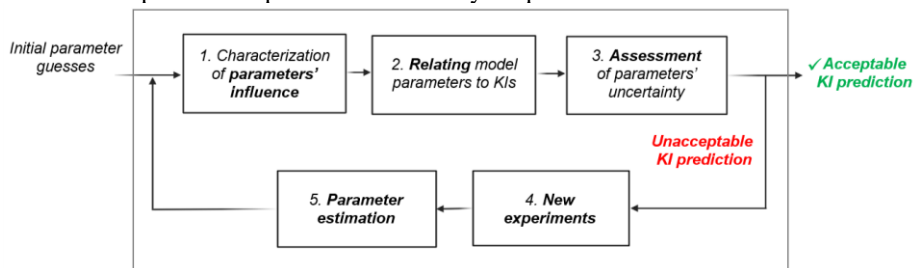
Quantitative models have been progressively adopted to accelerate the development of pharmaceutical manufacturing processes that traditionally requires time and resource-intensive experimental campaigns (Bano et al, 2022). Despite the potential benefits of modeling in pharmaceutical process development, the systematic use of these methods is not widespread, as stakeholders generally show a lack of confidence in the prediction capability of quantitative models with respect to key performance indicators (KPIs) or critical quality attributes (CQAs), which we will generically call key indicators (KIs) (Braakman et al, 2022). It is worth adding that pharmaceutical processes involve very complex phenomena, which may not be easily captured by first-principles models; therefore, high in-house expertise and resources are required (Boukouvala et al, 2011; Polak et al, 2023).

The prediction reliability of a quantitative model largely depends on the precision of model parameter estimates. A particular case is given by flowsheet models (or *systems models*, as they are often called in the pharmaceutical sector), where several unit operation models are connected, and the KIs are the outputs of some units (often the final one). It

is evident that, in general, the prediction fidelity of a KI depends not only on the parameters of the specific unit operation model, but also on the parameters of all unit operation models impacting the unit of interest. In this context, one may wonder whether *all* parameters should be estimated in a statistically satisfactory way or just a subset of them, and what parametric precision is sufficient to satisfy the prediction requirements for the KI. In this study, we propose a model evaluation framework to support the critical usage of quantitative models within pharmaceutical manufacturing process development. The methodology combines traditional tools typically adopted for parameter estimation purposes – namely global sensitivity analysis (GSA) and model-based design of experiments (MBDoE) – with techniques based on data analytics, that allow to graphically represent both model parameters and the KI onto a common space. A direct compression systems model for manufacturing oral solid dosage (OSD) products is used as a case study.

## 2. Methodology

In this section we describe the methodology for the assessment and quantification of parameters impact on the prediction reliability of quantitative models.



**Figure 1:** Schematic of the proposed methodology to quantify the model parameters' impact on the prediction reliability of quantitative models.

The workflow consists of multiple sequential steps (Fig. 1):

- **Step 1.** The objective is to characterize the parameter impacts on the prediction of the selected KI, so that the most influential parameters can be identified. We use Sobol's GSA to understand how the variability of the selected KI can be allocated to each model parameter.
- **Step 2.** The goal is quantifying the parameter impacts on the model prediction fidelity. In order to assist the interpretation of these results, here we suggest using a partial least-squares (PLS) regression model to relate the model parameters to the selected KI. Details are reported in § 2.1.
- **Step 3.** It is necessary to assess if the precision of the model parameters is sufficient. To this purpose, the estimated uncertainties of model parameters are projected onto the PLS model. If some projections fall outside the KI acceptability region, new experiments are needed to increase the parameters' precision. No improvement is required if all uncertainties fall inside the KI acceptability region (stop criterion).
- **Step 4.** If the parameters' precision requires improvement based on the output of Step 3, new experiments should be designed to collect new data. Experiments are designed through MBDoE techniques (Franceschini and Macchietto, 2008) to strengthen the precision of the most influential parameters identified at Step 1.
- **Step 5.** Once new data are available, all model parameters are estimated using a maximum likelihood estimator.

The procedure is iteratively repeated until satisfaction of the stop criterion.

### 2.1. Use of a PLS model

Let  $\mathbf{X}$  [ $N \times V$ ] be the input (regressor) matrix of  $N$  observations of  $V$  variables, and  $\mathbf{Y}$  [ $N \times L$ ] the response matrix of  $L$  variables. PLS (Geladi and Kowalski, 1986) is a multivariate regression technique that projects the regressor and response variables onto a common latent space, according to the model structure:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E}_X ; \quad \mathbf{Y} = \mathbf{T}\mathbf{Q}^T + \mathbf{E}_Y . \quad (1)$$

First, we generate a set of parameters combinations using Monte Carlo simulations, given their initial uncertainties. These combinations of model parameters make the set of regressors  $\mathbf{X}$ , while the predicted KI is the response variable  $\mathbf{y}$ .  $\mathbf{T}$  [ $N \times V$ ] is the score matrix,  $\mathbf{P}$  [ $V \times A$ ] and  $\mathbf{Q}$  [ $L \times A$ ] are arrays of loadings,  $A$  = number of latent variables (LVs), while  $\mathbf{E}_X$  and  $\mathbf{E}_Y$  are the residuals in the inputs and responses, respectively. We fix operating conditions so that the KI is equal to the target value, given the initial parameters guesses. Then, the PLS model is used in its inverse form (Jaeckle and MacGregor, 2000) to determine the set of input combinations that yield the desired tolerance on the target KI. The quality target is described by inequality constraints (bound values for the response prediction), which in turn defines the KI acceptability region. If the dimensionality of the latent space is greater than the  $y$ -space, multiple solutions of the model inversion problem exist, and are defined by the null space:

$$W = \{(\mathbf{t}_{new} + \mathbf{t}), \mathbf{t} \in \ker(\mathbf{Q})\}, \quad (2)$$

where  $\mathbf{t}_{new} = (\mathbf{Q}^T \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{y}^{*T}$  is the score vector for the target response variable  $\mathbf{y}^*$ . The use of a PLS model to relate parameters combinations and KIs gives us the possibility to monitor the evolution of parameter uncertainties along each iteration of the methodology in Fig. 1. We obtain a clear graphical representation of both regressor and response variables onto a common latent space that can be easily interpreted.

## 3. Case Study

A direct compression systems model for manufacturing OSD products is used as a case study. The systems model comprises a tablet press unit operation and a tablet disintegration unit.

### 3.1. Tablet press unit operation

First, the variation in tablet solid fraction caused by the compaction pressure,  $P$  [MPa], exerted by the press is expressed according to Kawakita and Lüdde (1971):

$$sf = \frac{a_{sf}(1 + b_{sf}P)}{1 + a_{sf}b_{sf}P} , \quad (3)$$

where  $sf$  [-] is the attained tablet solid fraction, while  $a_{sf}$  [-] and  $b_{sf}$  [MPa<sup>-1</sup>] are model parameters to be estimated. The Kushner (2012) equation is used to relate the effect of the extent of lubrication  $K$  [dm] attained in the upstream powder blending on the tablet tensile strength:

$$TS = TS_0((1 - \beta) + \beta \exp(-\gamma K)) , \quad (4)$$

where  $TS$  [MPa] is the tensile strength,  $TS_0$  [MPa] is the tensile strength at zero porosity,  $\gamma$  [dm<sup>-1</sup>] is the lubrication rate constant, and  $\beta$  [-] is the total fraction of tensile strength that can be lost due to lubrication. The empirical model by Nassar et al. (2021) accounts for the dependence of the Kushner parameters on the attained tablet solid fraction:

$$TS_0 = a_1 \exp(b_1(1 - sf)) , \quad (5)$$

$$\beta = a_2(1 - sf) + b_2 . \quad (6)$$

The model parameters  $a_1$  [MPa],  $b_1$  [-],  $a_2$  [-] and  $b_2$  [-] need to be estimated.

### 3.2. Tablet disintegration

Both erosion and swelling (Markl et al., 2017) mechanisms are considered.

Erosion is described as:

$$V_c = (H_{coat} - \dot{\epsilon}t)A_t, \quad (7)$$

where  $V_c$  [m<sup>3</sup>] is coating volume varying with time  $t$  [s],  $H_{coat}$  [m] is the thickness of the coating layer,  $A_t$  [m<sup>2</sup>] is the tablet surface area, and  $\dot{\epsilon}$  [m s<sup>-1</sup>] is the constant erosion rate. The dynamic evolution of the penetration depth due to swelling is modelled as:

$$\frac{dP_d}{dt} = \left(\frac{P}{F_L/A_t}\right)^{n(T_{t/2}-P_d)/T_{t/2}} \left[\frac{d_h^2 \epsilon}{S_p \mathcal{T}_{or}^2 \mu P_d}\right] p_c, \quad (8)$$

where  $P_d$  [m] is the water penetration depth,  $F_L$  [MN] the tablet hardness,  $d_h$  [m] the tablet hydraulic diameter,  $\mathcal{T}_{or}$  [-] the average tablet tortuosity,  $\mu$  [Pa s] the liquid viscosity,  $p_c$  [Pa] the capillary pressure.  $S_p$  [-] and  $n$  [-] are formulation-dependent model parameters to be estimated. Parameter  $n$  can be estimated if dynamic penetration depth data are available, e.g., via terahertz (THz) spectroscopy. However, this is a relatively new technology if compared to a standard end-point disintegration time test. If

we only have end-point disintegration data, the term  $\left(\frac{P}{F_L/A_t}\right)^{n(T_{t/2}-P_d)/T_{t/2}}$  can be replaced with a lumped parameter  $\Lambda$  [-] which can be fitted.  $T_{t/2}$  [m] is the time-dependent half tablet thickness, and  $\epsilon$  [-] represents the average porosity of the swollen product. The stress due to tablet expansion from swelling is defined according to Peppas and Colombo (1989):

$$\tau = -TS + C_2 w_l + C_3 \sqrt{w_l}, \quad (9)$$

where  $\tau$  [MPa] is the total stress,  $w_l$  [-] is the liquid content in the tablet, and  $C_2$  [MPa] and  $C_3$  [MPa] are model parameters to be estimated. From  $\tau$ , we can compute  $\epsilon$  to be included in Eq. (8):

$$\tau = \frac{G_0 \exp\left(-\frac{B\epsilon}{1-\epsilon}\right) \lambda t}{T_{t/2}}, \quad (10)$$

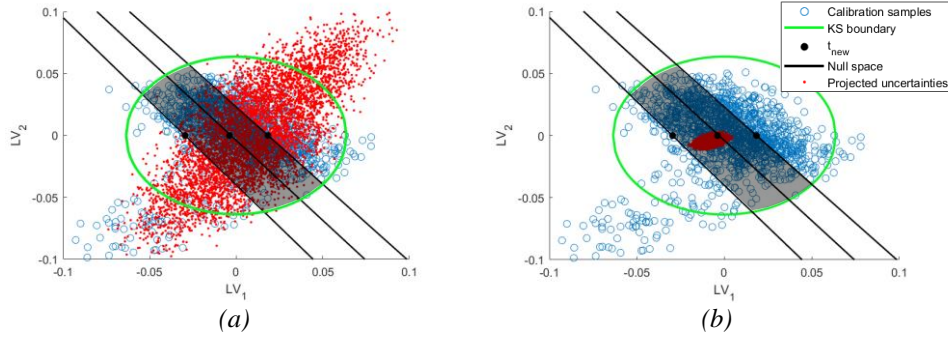
with  $G_0$  [MPa] and  $B$  [-] elastic constants from literature, and  $\lambda$  [s<sup>-1</sup>] the swelling rate (Kuentz and Leuenberger, 1998). We define the disintegration time as the time for which the tablet stops disintegrating, i.e.,  $dP_d/dt = 0$ . Five model parameters need to be estimated:  $C_2$ ,  $C_3$ ,  $\dot{\epsilon}$ ,  $n$ ,  $S_p$ .

### 3.3. Product quality assessment

We consider an immediate release product, with a target disintegration time of 4 min. According to USP <701> (2011) disintegration test specifications, the time limit for the formulation to completely disintegrate is 5 min; therefore, we set  $\pm 1$  min as the admissible tolerance with respect to the target value of the disintegration time. Results from Sobol's GSA (Step 1) are collected in two different metrics: the first-order sensitivity index  $S_i$ , which accounts for the direct effect of each parameter on the KI, and the total sensitivity index  $S_{i,TOT}$ , which also accounts for interactions with the other model parameters (Table 1). The use of both indices allows us to identify possible parameter interactions. Uniform distributions for the model parameters were assumed, and bounds were chosen on typical values that can be found in the literature (Peppas and Colombo, 1989; Nassar et al., 2021). The initial parameters' uncertainties were assumed equal to  $\pm 50\%$  of the nominal values, as reported in Table 2. The five most influential parameters, as characterized by the greatest total sensitivity indices, are:  $a_{sf}$ ,  $b_1$ ,  $b_{sf}$ ,  $n$  and  $S_p$ . Note that  $a_{sf}$ ,  $b_1$  and  $b_{sf}$  are parameters related to the tablet press model, suggesting that considering only the second unit may cause a loss of information compared to the consideration of both the tablet press and the disintegration together.

**Table 1.** Sobol's sensitivity indices for the parameters of the direct compression systems model with respect to the disintegration time. The most influential model parameters are in boldface.

Parameter	$S_i$	$S_{i,TOT}$
$a_1$	$4.767 \times 10^{-4}$	$1.925 \times 10^{-2}$
$a_2$	$3.762 \times 10^{-4}$	$5.478 \times 10^{-3}$
<b><math>a_{sf}</math></b>	<b><math>8.425 \times 10^{-2}</math></b>	<b>0.286</b>
<b><math>b_1</math></b>	<b><math>1.562 \times 10^{-3}</math></b>	<b><math>5.431 \times 10^{-2}</math></b>
$b_2$	$1.273 \times 10^{-4}$	$1.634 \times 10^{-2}$
<b><math>b_{sf}</math></b>	<b><math>2.245 \times 10^{-2}</math></b>	<b>0.139</b>
$\gamma$	$2.332 \times 10^{-4}$	$2.695 \times 10^{-4}$
$C_2$	$2.815 \times 10^{-4}$	$3.992 \times 10^{-3}$
$C_3$	$2.209 \times 10^{-4}$	$3.731 \times 10^{-3}$
$\dot{\epsilon}$	$3.603 \times 10^{-4}$	$5.042 \times 10^{-4}$
<b><math>n</math></b>	<b>0.641</b>	<b>0.881</b>
<b><math>S_p</math></b>	<b><math>1.312 \times 10^{-2}</math></b>	<b>0.101</b>



**Figure 2.** Application of the proposed methodology to the direct compression systems model for the assessment of tablet disintegration time: (a) initial iteration, (b) final iteration where all projected uncertainties expressed using new estimates CIs (red points) fall inside the KI acceptability region (black area).

The relationship between model parameters and the KI is assessed by building a PLS model (Step 2). We verified that two LVs are sufficient to capture enough variance of the data. With reference to the first iteration of the proposed workflow in Fig. 1, LV1 explains  $\sim 44\%$  the total variance, while LV2  $\sim 28\%$ . Confidence limits are considered in the latent space in the shape of an ellipse, whose semi-axis  $s_a$  along the  $a^{\text{th}}$  direction is defined as  $s_a = \sqrt{\lambda_a T_{lim}^2}$  ( $\lambda_a$  = eigenvalue of the matrix  $\mathbf{X}^T \mathbf{y} \mathbf{y}^T \mathbf{X}$ ;  $T_{lim}^2 = 95\%$  confidence limit); this region defines the knowledge space (KS) boundary. By assessing the effect of the initial model parameter uncertainty on the KI (Step 3), it can be observed that the prediction fidelity requirement for disintegration time cannot be satisfied, as many projections lie outside of the KI acceptability region (Fig. 2a). Thus, MBDoE techniques are applied to increase the precision of the five relevant parameters (Step 4); (in-silico) experimental data are used to re-estimate all parameters (Step 5). Seven iterations are required to reach the required model fidelity with respect to the KI prediction, i.e., seven experimental runs would need to be simultaneously performed for the two units. The results after the final iteration are shown in Fig. 2b: all projected uncertainties of parameter estimates fall inside the KI acceptability region. Finally, we report the estimated values of model parameters with their 95% confidence intervals (CIs) and  $t$ -values (Table 2). It can be observed that parameters  $C_2$ ,  $C_3$  and  $\dot{\epsilon}$  do not require precise estimation ( $t$ -values are lower than the reference) to meet the KI specification, and additional experimental effort can be saved.



**Table 2.** Estimated values of model parameters with their 95% CIs and *t*-values. † = 95% CI larger than  $\pm 50\%$  the parameter guess value, \* = precision is not statistically satisfactory.

Parameter	Units	Nominal	Estimated	95% CI	<i>t</i> -value
$a_1$	MPa	11.04	10.31	0.294	35.05
$a_2$	-	1.091	0.658	0.336	1.96
$a_{sf}$	-	0.463	0.439	$2.021 \times 10^{-3}$	$2.170 \times 10^2$
$b_1$	-	-8.202	-9.037	$6.638 \times 10^{-2}$	$1.361 \times 10^2$
$b_2$	-	0.326	0.150	$6.670 \times 10^{-2}$	2.26
$b_{sf}$	MPa <sup>-1</sup>	$2.460 \times 10^{-2}$	$2.984 \times 10^{-2}$	$3.863 \times 10^{-4}$	$1.036 \times 10^2$
$\gamma$	dm <sup>-1</sup>	$1.211 \times 10^{-3}$	$1.148 \times 10^{-3}$	$7.309 \times 10^{-5}$	15.71
					<i>t</i> <sub>ref</sub> = 1.690
$C_2$	MPa	$1.000 \times 10^2$	$1.032 \times 10^2$	$2.636 \times 10^4$ †	$3.942 \times 10^{-3}$ *
$C_3$	MPa	$1.000 \times 10^2$	82.46	$7.231 \times 10^2$ †	0.114*
$\dot{\epsilon}$	m s <sup>-1</sup>	$1.000 \times 10^{-3}$	$8.361 \times 10^{-4}$	$7.493 \times 10^{-2}$ †	0.112*
$n$	-	0.900	0.929	$3.591 \times 10^{-2}$	$1.037 \times 10^2$
$S_p$	-	0.524	0.488	$8.451 \times 10^{-2}$	5.77
					<i>t</i> <sub>ref</sub> = 1.646

## 4. Conclusions

In this study, a systematic workflow for the development and assessment of the predictive capability of quantitative models has been proposed and assessed. The methodology combines traditional tools with techniques based on data analytics, with the advantage of graphically represent both model parameters and the KI onto a common latent space. The methodology can be applied to any model, and be used to guide practitioners in model development and the subsequent usage for pharmaceutical process development.

## Acknowledgements

This study was funded by a Digital Design capability project at GlaxoSmithKline (GSK).

## References

- G. Bano, R.M. Dhenge, S. Diab, D.J. Goodwin, L. Gorringer, M. Ahmed, R. Elkes, S. Zomer, 2022. Streamlining the development of an industrial dry granulation process for an immediate release tablet with systems modelling. *Chem. Eng. Res. Des.*, 178, 421–37.
- F. Boukouvala, F.J. Muzzio, M.G. Ierapetritou, 2011. Dynamic data-driven modeling of pharmaceutical processes. *Ind. Eng. Chem. Res.*, 50, 6743–6754.
- S. Braakman, P. Pathamanathan, H. Moore, 2022. Evaluation framework for systems models. *CPT Pharmacometrics Syst. Pharmacol.*, 11, 264–289.
- G. Franceschini, S. Macchietto, 2008. Model-based design of experiments for parameters precision: state of the art. *Chem. Eng. Sci.*, 63, 4864–4872.
- P. Geladi, B.R. Kowalski, 1996. Partial least-squares regression: A tutorial. *Anal. Chim. Acta.* 185,1–17.
- C.M. Jaeckle, J.F. MacGregor, 2000. Industrial applications of product design through the inversion of latent variable models. *Chemom. Intell. Lab Syst.*, 50, 199–210.
- K. Kawakita, K.H. Lüdde, 1971. Some considerations on powder compression equations. *Powder Technol.*, 4, 61–68.
- M. Kuentz, H. Leuenberger, 1998. Modified Young's modulus of microcrystalline cellulose tablets and the directed continuum percolation model. *Pharm. Dev. Technol.*, 3, 13–19.
- J. Kushner, 2012. Incorporating Turbula mixers into a blending scale-up model for evaluating the effect of magnesium stearate on tablet tensile strength and bulk specific volume. *Int. J. Pharm.*, 399, 19–30.
- D. Markl, S. Yassin, D.I. Wilson, D.J. Goodwin, A. Anderson, J.A. Zeitler, 2017. Mathematical modelling of liquid transport in swelling immediate release tablets, *Int. J. Pharm.*, 526, 1–10.
- J. Nassar, B. Williams, C. Davies, K. Lief, R. Elkes, 2021. Lubrication empirical model to predict tensile strength of directly compressed powder blends. *Int. J. Pharm.*, 592, 119980.
- N.A. Peppas, P. Colombo, 1989, Development of disintegration forces during water penetration in porous pharmaceutical systems, *J. Control. Release*, 10, 245–250.
- J. Polak, M. Von Stosch, M. Sokolov, L. Piccioni, A. Streit, B. Schenkel, B. Guelat, 2023. Hybrid modeling supported development of an industrial small-molecule flow chemistry process. *Comput. Chem. Eng.*, 170, 108127.
- USP <701> Disintegration. The United States Pharmacopeial Convention. 2011.

# Simulation and optimisation of a medium scale industrial reverse osmosis desalination system

Mudhar A. Al-Obaidi <sup>a</sup>, Alanood A. Alsarayreh <sup>b</sup>, and Iqbal M. Mujtaba <sup>c</sup>

<sup>a</sup> *Middle Technical University, Technical Institute of Baquba, Baquba, Dayala – Iraq*

<sup>b</sup> *Chemical Engineering Department, Mu'tah University, Al Karak – Jordan*

<sup>c,\*</sup> *Chemical Engineering Department, University of Bradford, Bradford BD7 1DP, UK.*

*Email: [I.M.Mujtaba@bradford.ac.uk](mailto:I.M.Mujtaba@bradford.ac.uk)*

## Abstract

Desalination is one of the techniques used for meeting increasing water demand around the world. However, as the technology is energy intensive, the increasing energy price around the globe will force the practitioners to optimise the process from the point of minimising water production cost. The motivation of this study is to carry out simulation and optimisation of a reverse osmosis (RO) system to estimate and minimise the freshwater production cost. For this purpose, a detailed mathematical model of the process is developed using a combination of an earlier model of RO developed by the authors and a set of cost model equations collected from the open literature. The medium-scale industrial brackish water RO desalination system of Arab Potash Company in Jordan is selected as a case study. The model is first used in simulation mode which provides the detailed insight of the process and feasible operation envelope. The model is then embedded in a single objective non-linear optimisation framework to determine the best operating parameters in order to minimise the freshwater production cost while optimising several operating parameters and meeting the desired water quality specification in terms of salinity. Sensitivity of energy cost on the optimum operating conditions will also be presented in detail.

**Keywords:** Brackish water desalination; Reverse Osmosis system; Simulation; Optimisation; Water production cost.

## 1. Introduction

One of the methods being utilised to fulfil the rising water demand of the world is desalination (Mujtaba and Sowgath, 2022). However, such processes require huge amount of energy, and the rising cost of energy globally compels practitioners to optimise the process from the standpoint of reducing the freshwater production cost. Brackish Water Reverse Osmosis (BWRO) process is commonly used as a successful desalination method in many countries including Jordan and Saudi Arabia using groundwater with salinity ranging between 1,000 ppm to 10,000 ppm (Afonso et al., 2004).

According to the latest study of Pearson et al. (2021), the freshwater production cost varies between 0.39\$ to 0.66\$ per cubic meter of freshwater for water productivities ranging between 10,000 to 70,000 m<sup>3</sup>/d. However, there is still a scope of reducing the freshwater production cost further while enhancing the BWRO performance. Model based techniques can be applied to BWRO processes to identify such opportunity. Several researchers were successful in identifying optimal design and operating conditions of RO processes via simulation and optimisation leading to reduced freshwater production cost

of water. For instance, Arroyo and Shirazi (2012) evaluated the total freshwater production cost of six brackish water RO desalination facilities in Texas with productivity ranging between 5455.3 to 125017.47 m<sup>3</sup>/day and the cost ranging between 0.24 to 0.528\$ per cubic meter of freshwater. Ghaffour et al. (2013) carried out a techno-economic evaluation of the freshwater production cost considering the main influential parameters of different water desalination facilities resulting in 0.5 \$/m<sup>3</sup> of freshwater. Using simulation and optimisation, an economic study was achieved by Atab et al. (2016) to assess the specific energy consumption and freshwater production cost of brackish water RO desalination system which utilised a pressure exchanger as an energy recovery device. A specific model of RO process was used to carry out. The study resulted in 24000 m<sup>3</sup>/day of productivity at 0.11 \$/m<sup>3</sup> and freshwater salinity of 400 ppm from feed water salinity 15,000 ppm with specific energy consumption of 2.8 to 0.8 kWh/m<sup>3</sup>.

Recently, Pearson et al. (2021) attempted to obtain real cost information of seven BWRO plants in Florida, built between 2004 to 2013, for the period between 2017 and 2019. Considering the calculations of capital cost and operating cost of the plants with capacity ranging between 1364 to 56818 m<sup>3</sup>/day and brackish water salinity ranging between 2,000 to 6,000 ppm, the freshwater production cost varied between 0.23 to 0.63 \$/m<sup>3</sup> of freshwater. Using solar energy for different RO-PV designs, most recently Shalaby et al. (2022) reported the freshwater cost between 0.74-1.58 \$/m<sup>3</sup>.

The cost of desalination varies significantly depending on the location and is influenced by several factors, such as the feed water source, feed water quality, plant size, process type and design including energy recovery, intake type, pre- and post-treatment processes, concentrate disposal method, regulatory issues, land costs, and water conveyance to and from the plant. Because of these variations in water quality and quantity from site to site, and occasionally even within the same site, it is expected that the cost of RO brackish water desalination will be the same for all cases and condition. However, each case should be optimised to obtain the minimum freshwater production cost. To the best of our knowledge, there has not been any study to predict the freshwater production cost of a medium scale industrial BWRO system. This study estimates the freshwater production cost of the RO plant via simulation and optimisation using model-based technique. To carry out this task, an earlier RO model developed by the same authors and a set of cost model equations gathered from the open literature are combined to create a thorough mathematical model of the RO process comprising the estimation of freshwater production cost and associated specific energy consumption. The model is initially utilised in simulation mode, which offers an understanding of the actual freshwater production cost and a workable operational envelope. In order to reduce the cost of producing water while optimising a number of operating parameters and achieving the appropriate salinity specification, the model is then incorporated into a single objective non-linear optimisation framework.

## 2. Description of BWRO desalination system

Fig. 1 depicts the design of a brackish water RO for the APC plant with a 1200 m<sup>3</sup>/day capacity. Permeate and retentate reprocessing designs are used in two passes. Two stages of pressure vessels are arranged in the first pass in 4:2 order. Three pressure vessel stages are present in the second pass and arranged in 2:1:1 order. The permeate from the first pass is transferred to the second pass for further purification. The brine line from the first pass is drained. A high-quality water of 2 ppm is produced from the second pass. The retentate of the second pass is returned and mixed with the raw feed water to maintain a high productivity.

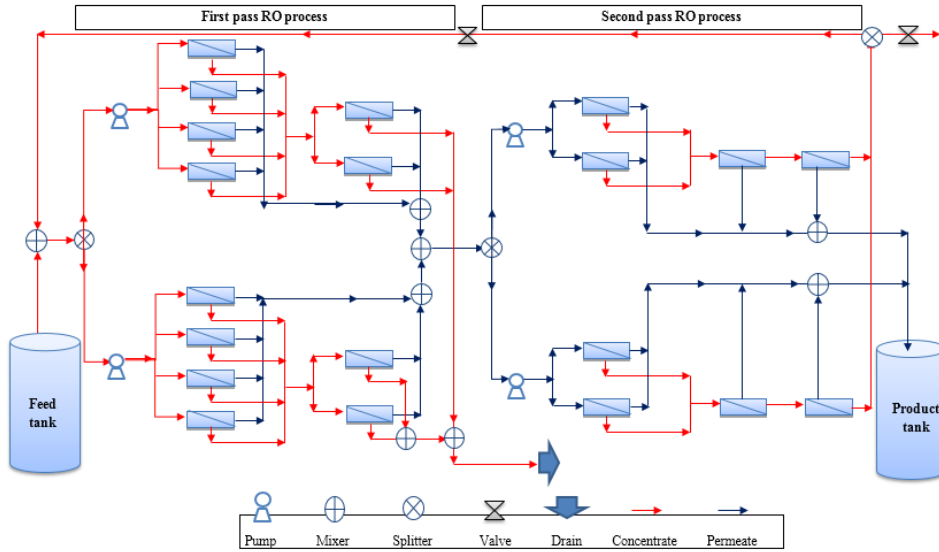


Fig. 1. A schematic representation of BWRO system of APC

### 3. Modelling of BWRO desalination system

A steady-state model was created by Al-Obaidi et al. (2018) for the APC medium-scale brackish water RO desalination plant. The model was constructed based on a number of assumptions. The most important ones are: 1) steady state operation, 2) fixed membrane features and channel geometries, 3) the film theory expresses the concentration polarisation phenomenon, 4) 1 atm in the permeate channel, 5) fixed temperature, and 6) The spacer properties are used to quantify the pressure drop in the feed channel. Before being utilised to examine the plant performance under various operating situations, the model was verified using APC plant data. In this work, this model has been improved by including a sub-model to calculate the cost of producing freshwater using data from several research (Filippini et al., 2018; Al-Obaidi et al., 2019; Malek et al., 1996). Table 1 presents a set of important equations of the RO model and the sub economic model used in this study. For further details of the model developed, parameters and variables and associated effects on the performance indicators can all be found in Al-Obaidi et al. (2018).

### 4. Simulation of BWRO desalination system

Simulation of the process is carried out to estimate the most important performance indicators including the freshwater production cost and specific energy consumption. Table 2 depicts the simulation results for the set of original inlet conditions of 988.93 ppm (salinity), 9.22 atm (pressure of 1<sup>st</sup> pass), 9.832 atm (pressure of 2<sup>nd</sup> pass), 2034.33 m<sup>3</sup>/day (feed flowrate) and 25 °C (temperature).

Furthermore, the estimation of electrical cost of pumping the brackishwater into the 1<sup>st</sup> and 2<sup>nd</sup> passes of RO system is important to be addressed. The power price (\$/day) is calculated using the following correlation:

$$\text{Power price per day} \left( \frac{\text{USD\$}}{\text{day}} \right) = \text{SEC} \left( \frac{\text{kWh}}{\text{m}^3} \right) \times \text{productivity} \left( \frac{\text{m}^3}{\text{day}} \right) \times \text{Factor} \left( \frac{\$}{\text{kWh}} \right)$$

Factor is the specific price of each kWh in Jordan which is 0.123\$/kWhr (Jordan electricity prices). Table 2 shows the energy price per day.

**Table 1.** RO model

Equation	Description	Unit
$Q_p = A_{w(r)} \left( P_f - \frac{\Delta P_{drop,E}}{2} - P_p - \pi_w - \pi_p \right) A_m$	Water flux	(m <sup>3</sup> /s)
$Q_s = B_{s(r)} (C_w - C_p)$	Solute flux	(kg/m <sup>2</sup> s)
$\Delta P_{drop,E} = \frac{9.8692 \times 10^{-6} A' \rho_b Q_b^2 L}{2 d_h R e_b^2 (W t_f \epsilon)^2}$	Pressure drop per element	(atm, -)
$\frac{(C_w - C_p)}{(C_b - C_n)} = \exp \left( \frac{Q_p / A_m}{k} \right)$	Membrane surface concentration	(-)
$k = 0.664 k_{dc} R e_b^{0.5} S c^{0.33} \left( \frac{D_b}{d_h} \right) \left( \frac{2 d_h}{L_f} \right)^{0.5}$	Mass transfer coefficient	(m/s, -)
$C_p = \frac{B_s C_f e^{\frac{J_w}{k}}}{J_w + B_s e^{\frac{J_w}{k}}}$	Permeate concentration	(kg/m <sup>3</sup> )
$Rej = \frac{C_f - C_p}{C_f} \quad Rec = \frac{Q_p}{Q_f}$	Solute rejection and water recovery	(-)
$SEC = \frac{(P_f(plant) \times 10^{1325}) Q_f(plant)}{Q_p(plant) \epsilon_{pump} \times 36 \times 10^5}$	Specific energy consumption	(kWh/m <sup>3</sup> )
$TAC = TCC + AOC$	Total annual cost	(\$/yr)
$TCC = [(C_{wip} + C_{Pump} + C_{me}) SD CC]$	Total capital cost	(\$/yr)
$AOC = OC_{Pu} + OC_{sc} + OC_{ch} + OC_{me} + OC_{lab} + OC_{maint} + OC_{bd}$	Total annual operating cost	(\$/yr)
$C_{wip} = 996 (86400 Q_f(plant))^{0.8}$	Water intake and pre-treatment cost	(\$)
$C_{Pump} = [52 (3600 Q_f(plant) (P_f(plant) 0.101325))^{0.96}]$	Capital cost of high-pressure pump	(\$)
$C_{me} = N_s N_{PV} (C_{ele} N_{ele} + C_{PV})$	Membrane module and pressure vessel capital cost	(\$)
$OC_{pu} = 365 \times 24 \left[ \left( \frac{(3600 (P_f(plant) 0.101325) Q_f(plant))}{3.6 \epsilon_{pump} \epsilon_{motor}} \right) \right] E_c L_f$	Pumping operating cost	(\$/yr)
$OC_{sc} = 3600 \times 24 \times 365 C_{cf} Q_p(plant) L_f$	Annual operating spares cost	(\$/yr)
$OC_{bd} = 3600 \times 24 \times 365 C_{bd} Q_p(plant) L_f$	Effluents disposal cost	(\$/yr)
$OC_{ch} = 3600 \times 24 \times 365 C_{ct} Q_f(plant) L_f$	Annual chemical treatment cost	(\$/yr)
$OC_{me} = 0.2 C_{me}$	Annual membrane replacement cost	(\$/yr)
$OC_{lab} = C_{lab} 3600 \times 24 \times 365 Q_p(plant)$	Annual labour cost	(\$/yr)
$FWC_{RO} = \frac{(TCC) + AOC}{3600 \times 24 \times 365 Q_p(plant)}$	Freshwater production cost	(\$/m <sup>3</sup> )
$CCRF = \frac{[(i+1)^n - 1]}{i(i+1)^n}$	Capital cost recovery factor	(yr)
$OC_{main} = 0.02 PUC 3600 \times 24 \times 365 Q_p(plant)$	Annual maintenance costs	(\$/yr)

**Table 2.** Simulation results

Water productivity (m <sup>3</sup> /day)	Product salinity (ppm)	Salt rejection (%)	Water recovery (%)	Specific energy consumption (kWh/m <sup>3</sup> )	Total capital cost (\$)	Total operation cost (\$)	Freshwater production cost (\$/m <sup>3</sup> )	Energy price per day (\$/day)
1164.51	2.0	99.79	57.24	0.834	64005.66	75753.3	0.192	119.45

### 5. Optimisation of BWRO desalination system

This section intends to optimise the operation of BWRO system of APC to guarantee the lowest freshwater production cost (objective function) via determining the optimal operating condition of feed pressure of the 1<sup>st</sup> and 2<sup>nd</sup> passes and feed flowrate. The other operating conditions of groundwater salinity and temperature will be taken as fixed values of 988.93 ppm and 25 °C, respectively. Furthermore, the optimisation framework has considered several inequality constraints including the upper and lower feed flowrate for each membrane. Also, the specific energy consumption has been constrained at its simulation value. The non-linear single-objective optimisation framework can be described as:

**Given:** Feed water pressure and flowrate, RO module specifications.

**Optimise:** The optimisation variables of feed pressure, and flow rate.

**Minimise:** the freshwater production cost (*FWPC*) of RO system.

**Subject to:** Equality and inequality constraints and limits of optimisation variables.

Thus, the optimisation problem can be mathematically represented as:

$$\text{Min} \quad \text{FWPC}$$

$$Q_{f(plant)}, P_{p(plant)}$$

Subject to: Equality constraints: RO process model and fixed specific energy consumption (simulation value) ( $SEC = 0.834 \text{ kWh/m}^3$ )

Inequality constraints:

a) lower and upper limits of feed flow rate of RO system

$$(696.96 \text{ m}^3/\text{day}) \quad Q_{f(plant)}^L \leq Q_{f(plant)} \leq Q_{f(plant)}^U \quad (3707.52 \text{ m}^3/\text{day})$$

a) lower and upper limits of feed pressure of the 1<sup>st</sup> and 2<sup>nd</sup> passes of RO system

$$(5 \text{ atm}) \quad P_{f(1st \text{ pass})}^L \leq P_{f(1st \text{ pass})} \leq P_{f(1st \text{ pass})}^U \quad (12 \text{ atm})$$

$$(5 \text{ atm}) \quad P_{f(2nd \text{ pass})}^L \leq P_{f(2nd \text{ pass})} \leq P_{f(2nd \text{ pass})}^U \quad (12 \text{ atm})$$

c) lower and upper limits of feed flow rate of each membrane module

$$(87.12 \text{ m}^3/\text{day}) \quad Q_{f(membrane)}^L \leq Q_{f(membrane)} \leq Q_{f(membrane)}^U \quad (463.44 \text{ m}^3/\text{day})$$

Table 3 shows the optimisation results of several performance indicators including the freshwater production cost and the optimal feed pressure and flowrate. Note that the freshwater production cost could be reduced by 6.7% by increasing the pressure of the 1<sup>st</sup> and 2<sup>nd</sup> passes from 9.22 and 9.832 atm to 9.93 and 10.59 atm, respectively. Also, it can be noticed that the feed flowrate has been decreased from 2034.33 to 1900.8 m<sup>3</sup>/day to assure increasing the rate of permeation through the membranes due to increased residence time of the water inside the module. The productivity increases by 8% while maintaining a high-quality freshwater of 2.1 ppm (as desired). However, the electrical cost is inevitably increased (by 8%) due to increasing the productivity in cubic meter per day.

**Table 3.** Optimisation results

Optimised inlet conditions			Optimised results		
Parameter	Value	Unit	Parameter	Value	Unit
Feed pressure of 1 <sup>st</sup> pass	9.93	atm	Water recovery	66.33	%
Feed pressure of 2 <sup>nd</sup> pass	10.59	atm	Salt rejection	99.78	%
Feed flow rate	1900.8	m <sup>3</sup> /day	Salinity of product water	2.1	ppm
			Specific energy consumption	0.834	kWh/m <sup>3</sup>
			Water productivity	1260.81	m <sup>3</sup> /day
			Freshwater production cost	0.179	\$/m <sup>3</sup>
			Electricity cost	129.33	\$/day

Reduction of freshwater production cost (in comparison to simulation) = 6.7%

## 6. Conclusions

This paper attempted to reduce the overall freshwater production cost of a medium scale brackish water RO desalination system of Arab Potash Company (Jordan) via process modelling, simulation and optimisation. The simulation of the studied RO system obtained the values of performance indicators while the optimisation has obtained the minimum freshwater production cost with optimal feed pressure of the 1<sup>st</sup> and 2<sup>nd</sup> passes and inlet feed flowrate for the same specific energy consumption (as in the real plant and simulation). The freshwater production cost is reduced by 6.7% in comparison to the original simulation value while achieving a high-quality of product freshwater. Note, the actual electrical cost increases by more than 8% due to increase in the water productivity of RO system (compared to actual plant productivity). In the context of model refinement, relaxing the model assumptions would certainly provide more accurate predictions of the performance metrics. Furthermore, it should be noted that the developed model did not thoroughly incorporate the influence of fouling on the performance metrics. The enhancement of the established model is expected to introduce actual estimation of the freshwater production cost, which should be addressed in future study.

## References

- A.A. Alsarayreh, M.A. Al-Obaidi, S.K. Farag, R. Patel, I.M. Mujtaba, 2021, Performance evaluation of a medium-scale industrial reverse osmosis brackish water desalination plant with different brands of membranes. A simulation study, *Desalination*, 503, 114927.
- A. Malek, M.N.A. Hawlader, J.C. Ho, 1996, Design and economics of RO seawater desalination. *Desalination*, 105, 245–261.
- G. Filippini, M.A. Al-Obaidi, F. Manenti, I.M. Mujtaba, 2018, Performance analysis of hybrid system of multi effect distillation and reverse osmosis for seawater desalination via modeling and simulation. *Desalination*, 448, 21–35.
- I.M. Mujtaba, M.T. Sowgath, 2022, *Desalination Technology: Design and Operation*, Elsevier.
- J. Arroyo, and S. Shirazi, 2012. Cost of brackish groundwater desalination in Texas. Texas Water Development Board, 1–7.
- J.L. Pearson, P.R. Michael, N. Ghaffour, T.M. Missimer, 2021. Economics and Energy Consumption of Brackish Water Reverse Osmosis Desalination: Innovations and Impacts of Feedwater Quality. *Membranes*, 11(8), 616.
- Jordan electricity prices. GlobalPrices.com. [https://www.globalpetrolprices.com/Jordan/electricity\\_prices/](https://www.globalpetrolprices.com/Jordan/electricity_prices/) (Accessed at 3-11-2022).
- M.A. Al-Obaidi, G. Filippini, F. Manenti, I.M. Mujtaba, 2019, Cost evaluation and optimisation of hybrid multi effect distillation and reverse osmosis system for seawater desalination. *Desalination*, 456, 136–149.
- M.D. Afonso, J.O. Jaber, M.S. Mohsen, 2004. Brackish groundwater treatment by reverse osmosis in Jordan. *Desalination*, 164, 157–171.
- M.S. Atab, A.J. Smallbone, A.P. Roskilly, 2016. An operational and economic study of a reverse osmosis desalination system for potable water and land irrigation. *Desalination*, 397, 174-184.
- N. Ghaffour, T.M. Missimer, G.L. Amy, 2013. Technical review and evaluation of the economics of water desalination: current and future challenges for better water supply sustainability. *Desalination*, 309, 197–207.
- S.M. Shalaby, M.K. Elfakharany, I.M. Mujtaba, B. M. Moharram, H. F. Abosheisha, 2022. Development of an efficient nano-fluid cooling/preheating system for PV-RO water desalination pilot plant, *Energy Conversion and Management*, 268, 115960

# Modeling strategies in multi-scale food-energy-water nexus system optimization

Marcello Di Martino<sup>a,b</sup>, Patrick Linke<sup>a,c</sup>, Efstratios N. Pistikopoulos<sup>a,b</sup>

<sup>a</sup>*Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX, USA*

<sup>b</sup>*Texas A&M Energy Institute, Texas A&M University, College Station, TX, USA*

<sup>c</sup>*Department of Chemical Engineering, Texas A&M University at Qatar, Education City, Doha, Qatar*

## Abstract

The modeling and optimization of multi-scale process systems is based on several interconnected process sub-systems. Due to the complexity of each individual sub-model, the resulting integrated process framework optimization formulations are computationally challenging to solve. While the richness of the multi-scale model employed is desired to maintain in order to obtain a solution with some degree of accuracy, simpler surrogate models are typically more attractive as a means to tame the underlying complexity, albeit often leading to an increase of the problem size. Here, the food-energy-water nexus (FEWN) is selected as a representative multi-scale process system, with focus on the reverse osmosis (RO) water supply sub-system. Based on a RO desalination model, two models are developed and compared in terms of accuracy, model complexity and size as well as computational efficiency, (i) a mixed-integer non-linear programming (MINLP) surrogate model, and (ii) a mixed-integer linear programming (MILP) surrogate model of reduced complexity but larger size. The results indicate that improved computational times can be obtained for a valid (lower bound) solution based on the MILP modeling strategy within the same level of accuracy, further underlying the importance of the selection of an appropriate surrogate model.

**Keywords:** optimization, surrogate modeling, food-energy-water nexus, reverse osmosis.

## 1. Introduction

The optimization of multi-scale process systems, which are composed of several interconnected sub-systems, remains challenging. Encountered challenges besides the multi-scale nature, encompass the complexity, as well as the size of the resulting optimization models. One approach to resolve these challenges is the usage of surrogate models to approximate sub-systems of the overall process systems (Kakodkar et al. 2022, Guillén-Gosálbez et al. 2019). This work evaluates in detail how changes in the employed surrogate model can have significant impact on the solution generation of the integrated process system. One of these interconnected multi-scale systems is the food-energy-water nexus (FEWN), which postulates that sustainable solutions regarding food, energy and water systems can only be obtained when all resources are considered holistically. Recently, more and more interest has been placed on the FEWN due to depleting natural resources and increasing global demands (Di Martino et al. 2019, Garcia and You 2016). FEWN models result in complex optimization programs, where surrogate models can



mitigate computational difficulties by reducing the complexity of the system. However, such systems usually increase in number of constraints and number of variables, which in turn can result in an intractable optimization model as well, especially if the surrogate is solved over a given time horizon multiple times integrated in an overarching process system (Allen et al. 2022). Thus, the scope of this work is to analyze and compare a mixed-integer linear programming (MILP) and a mixed-integer non-linear programming (MINLP) surrogate model which are representative of a sub-system in an interconnected multi-scale process model. As an example for such an interconnected system the FEWN is selected, with focus on a reverse osmosis (RO) desalination plant as the water supply system. The comparison is based on the accuracy of the obtained minimized solutions, the computational efficiency in the form of the necessary computational time to obtain solutions and the difference in the size and complexity of the optimization models. Next, the system under investigation is described. Then, the detailed modeling equations are discussed, before analyzing and discussing the results.

## 2. System Under Investigation

To analyze and compare the impact of MILP and MINLP surrogate models for the integration in a multi-scale process system the performance of RO desalination plant models is compared. As visualized in Figure 1, the goal is to integrate the RO model in a multi-scale FEWN model which in turn specifies restrictions regarding the design and operation of the RO plant. The goal is to solve this integrated model at an hourly time scale for a one-year time horizon, resulting in 8760 time points. Solutions specify the sustainable and interconnected water, energy and food supply system design and operation (Cook et al. 2022, Di Martino et al. 2022a). Thus, it is important to evaluate the model approximations in terms of accuracy and computational efficiency for the solution generation. All obtained surrogate models are based on one-and-a-half years of operational RO plant data, as presented in Di Martino et al. 2022b.

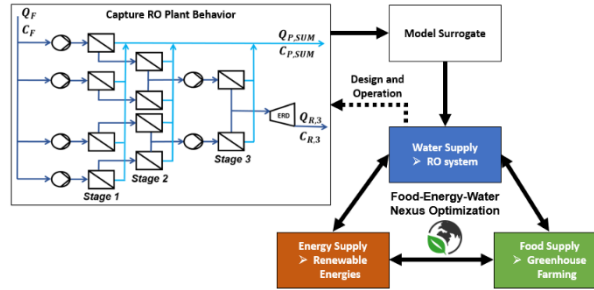


Figure 1: Schematic overview of the placement of the sub-system (RO) in the multi-scale process system (FEWN).

## 3. Modeling

The RO desalination plant is modeled and optimized according to Eq. 1 to 9 (Di Martino et al. 2022b).

$$\min SEC^* = \frac{1}{4 \cdot \eta_1} \sum_{j=1}^4 P_{f,1,j} + \frac{1}{10 \cdot \eta_1} \left( \sum_{j=1}^2 P_{f,3,j} - \frac{1}{2} \sum_{j=1}^4 P_{r,2,j} \right) - \frac{\eta_2}{20} \sum_{j=1}^2 \Delta P_{ERD,j} \quad (1)$$

s.t.

$$WR_{sys} \geq WR^{res} \leftrightarrow WR_i \geq WR_i^{res} \quad \forall i \in \{1,2,3\} \quad (2)$$

$$Q_f \geq Q_f^{res} \quad (3)$$

$$C_{p,t-1} = C_{p,t} \leq C_p^{res} \quad (4)$$

$$WR_i = WR_i^{max} \cdot \frac{\sum_j P_{i,j}}{\sum_j \max(P_{i,j})} \quad \forall i \in \{1,2,3\} \quad (5)$$

$$P_{r,i,j} = f_1(P_{f,i,j}) \quad \forall i = \{1,2\}, \forall j = \{1,2,3,4\} \quad (6)$$

$$P_{f,2,j} = P_{r,1,j} \quad \forall j \in \{1,2,3,4\} \quad (7)$$

$$\Delta P_{ERD,j} = f_2(P_{f,3,j}) \quad \forall j = \{1,2\} \quad (8)$$

$$C_{p,t} = f_3(Q_f, WR_1, WR_2, WR_3, C_{p,t-1}) \quad (9)$$

The presented objective function (Eq. 1) minimizes the linearized specific energy consumption of the system ( $SEC^*$ ) based on the feed ( $P_{f,i,j}$ ) and retentate pressures of each stage  $i$  and parallel flow  $j$  ( $P_{r,i,j}$ ).  $\eta_1$  and  $\eta_2$  denote the efficiency of the utilized pumps and energy recovery device, respectively. Eq. 2 to 9 summarize all necessary linear constraints of the system to specify the operation and design of the RO plant. The overall water recovery ( $WR_{sys} = Q_p/Q_f$ ), the feed flow ( $Q_f$ ) and the permeate concentration at the given and previous time point ( $C_{p,t}$ ,  $C_{p,t-1}$ ) are restricted to satisfy given water demand scenarios (Eq. 2 to 4). It is important to note that  $WR_{sys}$  is implicitly restricted by restricting the water recovery of each stage. The water recovery of each stage ( $WR_i \quad \forall i \in \{1,2,3\}$ ) is estimated based on the observed maximum water recovery ( $WR_i^{max} \quad \forall i \in \{1,2,3\}$ ) and pressures, according to Eq. 5. Eq. 6 and 7 summarize the pressure modeling of the system, with the linear regression ( $f_1(P_{f,i,j})$ ) used for calculating the retentate pressure of stages one and two for all parallel flows based on the respective feed pressure and the utilization of the retentate of stage one as the feed of stage 2. The pressure difference across the energy recovery device ( $\Delta P_{ERD,j} \quad \forall j = \{1,2\}$ ) as well as the overall permeate concentration ( $C_{p,t}$ ) are calculated based on the MILP reformulation of a feedforward artificial neural network (ANN) with rectified linear units (ReLU) as activation functions, as stated in Eq. 8 and 9 ( $f_2(P_{f,3,j}) \quad \forall j = \{1,2\}$ ,  $f_3(Q_f, WR_1, WR_2, WR_3, C_{p,t-1})$ ).

### 3.1. MINLP Surrogate Model

$SEC^*$  is derived by applying process specific, non-generic assumptions as described in Di Martino et al. 2022b. Therefore, obtained solutions actually specify lower bounds of the true solution. Accordingly, the non-linearized formula to calculate the exact specific energy consumption ( $SEC$ ) of the system is utilized in this study, as shown in Eq. 10. Furthermore, to obtain the energy consumption of the system ( $EC$ ),  $SEC$  has to be multiplied by  $WR_{sys}$  and  $Q_f$ , resulting in a trilinear term (Eq. 11). In addition,  $WR_{sys}$  can be derived based on  $WR_i \quad \forall i \in \{1,2,3\}$  with the nonlinear equation presented in Eq. 12. Thus, the MINLP surrogate model is given by minimizing Eq. 10 subject to the constraints summarized from Eq. 2 to 9, 11 and 12. With this modified model formulation no non-generic process specific assumptions are required.

$$SEC = \frac{1}{WR_{sys}} \cdot \left( \frac{1}{4 \cdot \eta_1} \sum_{j=1}^4 P_{f,1,j} + \frac{(1-WR_1) \cdot (1-WR_2)}{\eta_1} \left[ \frac{1}{2} \sum_{j=1}^2 P_{f,3,j} - \frac{1}{4} \sum_{j=1}^4 P_{r,2,j} \right] - \frac{(1-WR_1) \cdot (1-WR_2) \cdot (1-WR_3) \cdot \eta_2}{2} \sum_{j=1}^2 \Delta P_{ERD,j} \right) \quad (10)$$

$$EC = SEC \cdot WR_{sys} \cdot Q_f \quad (11)$$

$$WR_{sys} = WR_1 + (1 - WR_1) \cdot WR_2 + (1 - WR_1) \cdot (1 - WR_2) \cdot WR_3 \quad (12)$$

### 3.2. MILP Surrogate Model

Next, a linearized version of the nonlinear model is obtained by training ANNs with ReLUs as activation functions and reformulating them as MILPs for the calculation of  $SEC$  and  $EC$ , ultimately resulting in piecewise linear surrogate models. Here the results of the MILP approximations are specified as  $\widehat{SEC}$  (see Eq. 13 to 17) and  $\widehat{EC}$  (see Eq. 18 to 22). The obtained weights and biases together with the  $R^2$  value of both surrogate models are summarized in Table 1. The auxiliary positive variables  $(x_{SEC}^{(k_{SEC})}, s_{SEC}^{(k_{SEC})}, x_{EC}^{(k_{EC})}, s_{EC}^{(k_{EC})})$ , binary variables  $(z_{SEC}^{(k_{SEC})}, z_{EC}^{(k_{EC})})$ , together with lower and upper bounds  $(LB_{SEC}^{(k_{SEC})}, LB_{EC}^{(k_{EC})}, UB_{SEC}^{(k_{SEC})}, UB_{EC}^{(k_{EC})})$  are introduced for the MILP reformulation.

$$W_{SEC,1}^{(k_{SEC})} \cdot [WR_1 \ WR_2 \ WR_3 \ Q_f]^T + b_{SEC,1}^{(k_{SEC})} = x_{SEC}^{(k_{SEC})} - s_{SEC}^{(k_{SEC})}, \forall k_{SEC} \in \{1,2\} \quad (13)$$

$$x_{SEC}^{(k_{SEC})} - z_{SEC}^{(k_{SEC})} \cdot UB_{SEC}^{(k_{SEC})} \leq 0, \forall k_{SEC} \in \{1,2\} \quad (14)$$

$$s_{SEC}^{(k_{SEC})} - (1 - z_{SEC}^{(k_{SEC})}) \cdot LB_{SEC}^{(k_{SEC})} \leq 0, \forall k_{SEC} \in \{1,2\} \quad (15)$$

$$x_{SEC}^{(k_{SEC})}, s_{SEC}^{(k_{SEC})} \geq 0, z_{SEC}^{(k_{SEC})} \in \{0,1\}, \forall k_{SEC} \in \{1,2\} \quad (16)$$

$$\widehat{SEC} = \sum_{k_{SEC}} W_{SEC,2}^{(k_{SEC})} \cdot x_{SEC}^{(k_{SEC})} + b_{SEC,2} \quad (17)$$

$$W_{EC,1}^{(k_{EC})} \cdot [SEC \ WR_{sys} \ Q_f]^T + b_{EC,1}^{(k_{EC})} = x_{EC}^{(k_{EC})} - s_{EC}^{(k_{EC})}, \forall k_{EC} \in \{1,2,3\} \quad (18)$$

$$x_{EC}^{(k_{EC})} - z_{EC}^{(k_{EC})} \cdot UB_{EC}^{(k_{EC})} \leq 0, \forall k_{EC} \in \{1,2,3\} \quad (19)$$

$$s_{EC}^{(k_{EC})} - (1 - z_{EC}^{(k_{EC})}) \cdot LB_{EC}^{(k_{EC})} \leq 0, \forall k_{EC} \in \{1,2,3\} \quad (20)$$

$$x_{EC}^{(k_{EC})}, s_{EC}^{(k_{EC})} \geq 0, z_{EC}^{(k_{EC})} \in \{0,1\}, \forall k_{EC} \in \{1,2,3\} \quad (21)$$

$$\widehat{EC} = \sum_{k_{EC}} W_{EC,2}^{(k_{EC})} \cdot x_{EC}^{(k_{EC})} + b_{EC,2} \quad (22)$$

Table 1: Weights  $(W_{na,1}^{(k)}, W_{na,2}^{(k)})$ , biases  $(b_{na}^{(k)}, b_{na,2})$  and  $R^2$  of the surrogate models for the approximation of the specific energy consumption ( $SEC$ ) and energy consumption of the system ( $EC$ ).

$na$	$k_{na}$	$W_{na,1}^{(k)}$	$W_{na,2}^{(k)}$	$b_{na,1}^{(k)}$	$b_{na,2}$	$R^2$
$SEC$	1	[1.1556 - 0.5436 0.4116 0.4630]	0.4638	1.0465	-0.9950	0.99
	2	[0.2843 0.1137 0.1071 1.3843]	0.5371	-0.3829		
$EC$	1	[0.7492 - 0.3236 0.4624]	0.8180	0.4966	-0.9443	0.99
	2	[-1.0350 2.5363 1.3228]	0.2926	-0.3701		
	3	[-0.1111 - 0.1307 - 0.7936]	-0.2929	-1.1133		

Lastly, Eq. 12 is linearized with a first-order Taylor expansion around the nominal operating point of the RO plant, as shown in Eq. 23 ( $\overline{WR}_{sys} = 0.6039, \overline{WR}_1 = 0.3113, \overline{WR}_2 = 0.2935, \overline{WR}_3 = 0.1860$ ). For the relevant operating range of  $0.4 \leq WR_{sys} \leq 0.85$ , the approximation results in  $R^2 = 0.98$ . Overall, the MILP surrogate model is composed of minimizing  $\widehat{SEC}$  subject to Eq. 2 to 9 and Eq. 13 to 23.

$$\begin{aligned} \widehat{WR}_{sys} \approx & \overline{WR}_{sys} + (1 - WR_2 - WR_3 + WR_2 \cdot WR_3) \cdot (WR_1 - \overline{WR}_1) \\ & + (1 - WR_1 - WR_3 + WR_1 \cdot WR_3) \cdot (WR_2 - \overline{WR}_2) \\ & + (1 - WR_1 - WR_2 + WR_1 \cdot WR_2) \cdot (WR_3 - \overline{WR}_3) \end{aligned} \quad (23)$$

The described nonlinearities are overcome by introducing piecewise linear approximations resulting in a less complex optimization model. However, the resulting

system increases in size, with an increased number of constraints (60 vs. 80) and variables (continuous 47 vs. 57; binaries 9 vs. 14). The optimization models have been solved in GAMS with the CPLEX solver for the MILP surrogate model and the BARON solver for the MINLP surrogate model. All surrogate models have been derived in MATLAB.

#### 4. Results and Discussion

To analyze the trade-offs between system complexity and system size of the presented MILP and MINLP surrogates, the monthly energy consumption of the year 2017 is minimized based on the obtained operational RO data. The results of the optimization are visualized in Figure 2. The results of both models are of the same order of magnitude, where for the same set of restrictions in terms of  $WR_{sys}$  and  $Q_f$  the MINLP surrogate results in higher energy consumptions than the MILP surrogate, as expected.

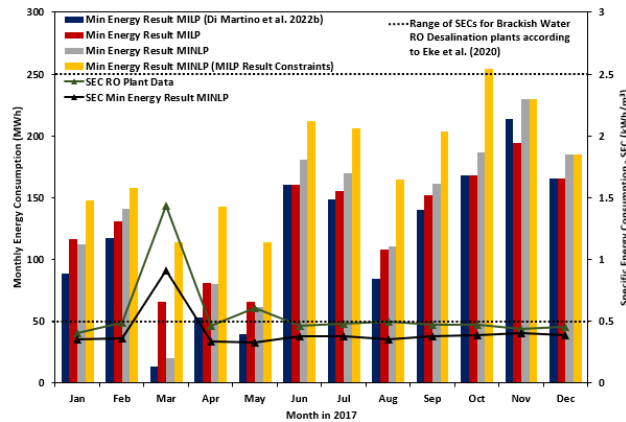


Figure 2: Comparison of the monthly energy consumption of the RO plant based on the optimization of the MILP, MINLP and MINLP with MILP result constraints.

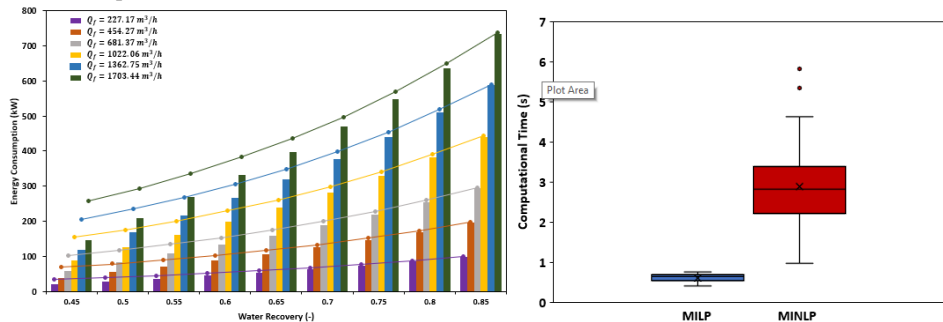


Figure 3: Left: Pareto front of the energy consumption of the RO system dependent on the system's water recovery and feed flow; bars – results of MILP, points connected with line – results of MINLP. Right: Comparison of the necessary computational time to solve the MILP and MINLP; boxplot: {min – 1st quartile – median – 3rd quartile – max}, x – mean, outliers – 1.5 times the interquartile range larger than the third quartile.

Next, the energy consumption of the RO plant is systematically minimized for different sets of  $WR_{sys}$  and  $Q_f$  resulting in the Pareto front shown in Figure 3. Again, for all observed results, the MILP denotes the lower bound of the solution. The difference between the MINLP and MILP solution decreases with increasing water recovery. This trend is observed for all feed flows under investigation. In addition, the computational efficiency of obtained solutions was recorded, as shown in Figure 3. In all observed cases,

the computational time for the MILP solution generation is less than the MINLP one, even in extreme cases. On average, the difference between the observed computational times is 2.28s. Thus, utilizing the MILP indeed results in significant reduction of computational time (for an hourly time scale over one year, 5.55 hours) while the obtained lower bounds are good approximations of the true optimum for high water recoveries.

## 5. Conclusion

This work analyzed and compared a MILP and MINLP surrogate model for the integration in a multi-scale process system. For illustration, the surrogate modeling of a RO process for the integration in a FEWN system has been studied due to its importance for sustainable solution generation. The derived MILP surrogate provides lower bounds to the true optimum which are of the same level of accuracy as the MINLP surrogate solutions, while resulting in significant increase of computational efficiency, viz. the reduction of computational time. This trade-off between increasing the size of the optimization model to linearize it, is particularly important to consider in applications for interconnected multi-scale process systems.

## 6. Acknowledgments

This publication was made possible by the National Priorities Research Program (NPRP) grant No. NPRP11S-0107–180216 from the Qatar National Research Fund (a member of Qatar Foundation). The findings herein reflect the work, and are solely the responsibility of the authors. The authors also gratefully acknowledge support from Texas A&M University, Texas A&M Energy Institute and Texas A&M University at Qatar.

## References

- R. C. Allen, S. G. Baratsas, R. Kakodkar, S. Avraamidou, C. D. Demirhan, C. F. Heuberger-Austin, M. Kolokkenburg, E. N. Pistikopoulos, 2022, A multi-period integrated planning and scheduling approach for developing energy systems, *Optim. Control. Appl. Methods*.
- J. Cook, M. Di Martino, R. C. Allen, E. N. Pistikopoulos, S. Avraamidou, 2022, A decision-making framework for the optimal design of renewable energy systems under energy-water-land nexus considerations, *Sci. Total Environ.*, 827, 154185.
- J. Eke, A. Yusuf, A. Giwa, A. Sodiq, 2020, The global status of desalination: An assessment of current desalination technologies, plants and capacity, *Desalination*, 495, 114633.
- M. Di Martino, S. Avraamidou, J. Cook, E. N. Pistikopoulos, 2021, An Optimization Framework for the Design of Reverse Osmosis Desalination Plants under Food-Energy-Water Nexus Considerations, *Desalination*, 503, 114937.
- M. Di Martino, S. Namany, S. Avraamidou, T. Al-Ansari, P. Linke, E. N. Pistikopoulos, 2022a, Energy-water Scheduling Decisions for Agricultural Scenario Planning, 32nd European Symposium on Computer Aided Process Engineering (ESCAPE-32), Elsevier, 1573-1578.
- M. Di Martino, S. Avraamidou, E. N. Pistikopoulos, 2022b, A Neural Network Based Superstructure Optimization Approach to Reverse Osmosis Desalination Plants, *Membranes*, 12 (2), 199.
- D. J. Garcia, F. You, 2016, The water-energy-food nexus and process systems engineering: a new focus, *Compu. Chem. Eng.*, 91, 49–67.
- G. Guillén-Gosálbez, F. You, A. Galan-Martin, C. Pozo, I. E. Grossmann, 2019, Process systems engineering thinking and tools applied to sustainability problems: Current landscape and future opportunities, *Current Opinion in Chemical Engineering* 26, 170–179.
- R. Kakodkar, G. He, C. D. Demirhan, M. Arbabzadeh, S. G. Baratsas, S. Avraamidou, D. Mallapragada, I. Miller, R. C. Allen, E. Gencer, E. N. Pistikopoulos, 2022, A review of analytical and optimization methodologies for transitions in multi-scale energy systems, *Renewable and Sustainable Energy Reviews*, 160, 112277.

# MULTI-FIDELITY SURROGATE MODELING FOR CHEMICAL PROCESSES WITH PHYSICS-INFORMED NEURAL NETWORKS

Yu-Ting Liu<sup>a</sup>, Chuan-Yu Wu<sup>b</sup>, Tao Chen<sup>b</sup>, Yuan Yao<sup>a\*</sup>

<sup>a</sup> *Department of Chemical Engineering, National Tsing Hua University, Hsinchu, 30013, Taiwan, ROC*

<sup>b</sup> *Department of Chemical and Process Engineering, University of Surrey, Guildford GU2 7XH, UK*

\*Corresponding Author's E-mail: [yao@mx.nthu.edu.tw](mailto:yao@mx.nthu.edu.tw)

## Abstract

In this study, a multi-fidelity surrogate modeling method based on physics-informed neural network (PINN) was proposed, which integrates high-fidelity simulation data and low-fidelity governing equations described by differential equations. By leveraging governing equations in the training of deep neural networks, the reliance on large amount of data has been relaxed. In the meantime, imposing physical laws ensures that the achieved surrogate models have clear physical meanings, which also improves the extrapolation performance of the models. Herein, the proposed multi-fidelity PINN surrogate modeling method was implemented to the simulation of the startup phase of a continuous stirred-tank reactor (CSTR) for illustrating its feasibility and advantages. From the computer experiment results, it is observed that the proposed method successfully reduced the sample size needed in model training and significantly improved the model extrapolation performance, facilitating its potential industrial applications.

**Keywords:** physics-informed neural network, multi-fidelity, surrogate modeling, deep learning.

## 1. Introduction

Numerical simulations, which use mathematical formulae to simulate complex physical systems, are becoming increasingly important in engineering tasks and decision-making processes (Moin and Mahesh, 1998), but the computational burden is a major practical problem limiting their applications. Surrogate models are a good solution to this problem. Nevertheless, data collection from high-fidelity simulations for training surrogate models is still computational expensive. Multi-fidelity surrogate modeling is an alternative approach, which utilizes both high-fidelity and low-fidelity data for model training (Fernández-Godino et al., 2019). A remaining problem is that, without a guidance of physical laws, the performance of surrogate models usually degrades significantly in the extrapolation applications.

In this study, we combined the strengths of data-driven and physics-based models to reduce the requirement on the amount of data for surrogate model training and improve the model extrapolation performance. In detail, we used the low-fidelity governing law described by a differential equation to constrain the training of the neural network model and obtain a multi-fidelity surrogate model based on physics-informed neuron network (PINN) (Raissi et al., 2019). With the aid of the physical law, the multi-fidelity PINN (MFPINN) surrogate model better approximates the simulated process especially in the

situation of small data or extrapolation. These advantages facilitate the potential applications of MFPINN to industries.

## 2. Numerical Simulation

In this study, a non-isothermal continuous stirred-tanked reactor (CSTR) was simulated using ANSYS Fluent to demonstrate the feasibility of proposed method. Figure 1 show a full-scale three-dimensional (3D) model of the system and Table 1 lists the system parameters and design space of input variables (Chuang et al., 2018). Here,  $C_{Ai}$  is the inlet concentrations of species A,  $T_i$  is the inlet temperature,  $V$  is the volume of the reactor,  $F_i$  is the inlet and outlet mass flow rate,  $S$  is the stirring speed of impeller, and  $\tau=V/F_i$  is the time constant of the CSTR. The physical characteristics, such as material density  $\rho$ , specific heat  $C_p$ , thermal conductivity  $\kappa$ , and dynamic viscosity  $\eta$ , are assumed to be constants with respect to temperature and compositions. This system consists of a first-order reaction,  $A \rightarrow B$ , where the activation energy is  $E_A$  and the pre-exponential Arrhenius constant is  $k_A$ .  $HR_A$  is the molar heats of the reaction. This 3D simulation provides information on product concentration and temperature as outputs, and we have selected the outlet concentration  $C_A$  as the predicted variable in presenting our results in this study due to its significance in determining product quality.

The mesh shape is tetrahedral with a minimum orthogonal quality of 0.238 and maximum skewness of 0.762 for accurate simulation results. It has 1867511 elements and 342112 nodes. The time consumption of each simulation run is about five to six hours. Consequently, this high-fidelity model is not suited for the engineering applications requiring multiple runs of simulations, such as optimization and sensitivity analysis. To deal with this problem, an MFPINN is proposed in this work, which can be implemented to build a surrogate model with a relatively small amount of simulation data.

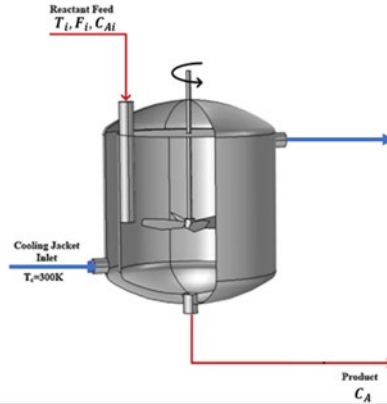


Figure 1 Schematic diagram of 3D model of CSTR process

Table 1. System parameters and their values used in the simulation

Parameters	Values	Units	Descriptions
$k_A$	$1.4 \times 10^4$	1/sec	Physical constant
$HR_A$	$-2.12 \times 10^7$	J/kgmol	Physical constant
$E_A$	$3.64 \times 10^7$	J/kgmol	Physical constant

$\rho$	1180	Kg/m <sup>3</sup>	Physical constant
$C_p$	$3.2 \times 10^3$	J/kg/K	Physical constant
$\kappa$	0.61	W/m/K	Physical constant
$\eta$	0.0008	Kg/m/s	Physical constant
V	18.764	dm <sup>3</sup>	Physical constant
S	120	rpm	Physical constant
R	8.314	J/mol/K	Physical constant
$C_{Ai}$	2-3	M	Design variable
$F_i$	0.0167-1.67	L/s	Design variable
$T_i$	300-370	K	Design variable
$C_A$		mol/dm <sup>3</sup>	System output

### 3. Multi-Fidelity Surrogate Model base on PINN

In the course of surrogate modeling of high-fidelity simulations that are computationally expensive, the cost of simulation data acquisition for surrogate model training is prohibitive. Therefore, we are inevitably confronted with the challenge of small data. Most conventional machine learning techniques, including most deep neural networks, are not robust in such a situation, whose performance cannot be guaranteed. The state-of-the-art PINN technique proposed by Raissi et al. (2019), which aims to solve supervised learning problems while respecting physical laws described by differential equations, can be adopted to address this issue. Inspired with PINN, a multi-fidelity surrogate model named MFPINN is developed by integrating both high-fidelity simulation data and low-fidelity physical law.

The low-fidelity physical law used in the MFPINN captures the essential physics of the process by assuming that the CSTR is perfectly mixed and isothermal at 300K. The following ordinary differential equation (ODE) can be obtained according to the conservation law to the mass.

$$\tau \frac{dC_A}{dt} = C_{Ai} - C_A + r_A \tau \quad (1)$$

With a first-order reaction,  $-r_A = k C_A$  and  $\tau = \frac{V}{v}$ , where  $k$  is defined by the Arrhenius equation.

$$k = k_A e^{\frac{-E_A}{RT}} \quad (2)$$

Consequently, eq. (1) is expressed as:

$$\frac{dC_A}{dt} + \frac{1 + \tau k}{\tau} C_A - \frac{C_{Ai}}{\tau} = 0 \quad (3)$$



MFPINN can be understood as a feedforward neural network (FFNN) trained with high-fidelity data generated with the numerical simulations described in Section 2 and the low-fidelity ODE described in eq.(3), whose loss function is composed of both two terms, i.e. a prediction loss and an ODE loss as depicted in eq.(4) ~ (6).

$$\text{Loss} = \text{Prediction loss} + \text{ODE loss} \quad (4)$$

$$\text{prediction loss} = \frac{1}{N_{\text{training}}} \sum_{i=0}^{N_{\text{training}}} (\widehat{C}_A(T_i, F_i, C_{Ai}, t_i) - C_A)^2 \quad (5)$$

$$\text{ODE loss} = \frac{1}{N_{\text{collocation}} + N_{\text{training}}} \sum_{i=0}^{N_{\text{collocation}} + N_{\text{training}}} \left( \frac{d\widehat{C}_A(T_i, F_i, C_{Ai}, t_i)}{dt} + \frac{1 + \tau k}{\tau} \widehat{C}_A - \frac{C_{Ai}}{\tau} \right)^2 \quad (6)$$

where  $T_i, F_i, C_{Ai}$  are three process parameters, which represent inlet temperature, inlet volume flow rate, and inlet concentration, respectively,  $t_i$  is the time value,  $\widehat{C}_A(T_i, F_i, C_{Ai}, t_i)$  is the prediction of the system output, i.e. the outlet concentration  $C_A$ , provided by the neural network,  $N_{\text{training}}$  denotes the number of training data points, and  $N_{\text{collocation}}$  denotes the number of collocation points which are chosen by Latin hypercube sampling (LHS) (Stein, 1987) within the design space of  $C_{Ai}, F_i$ , and  $T_i$  as listed in Table 1. The utilization of collocation points reduces the requirement on the amount of training data points. The automatic differentiation techniques (Baydin et al., 2018) was adopted for calculating the differentials. The training process of the MFPINN model is illustrated in Figure 2, where  $\theta^*$  denotes the model parameters including all weights and biases.

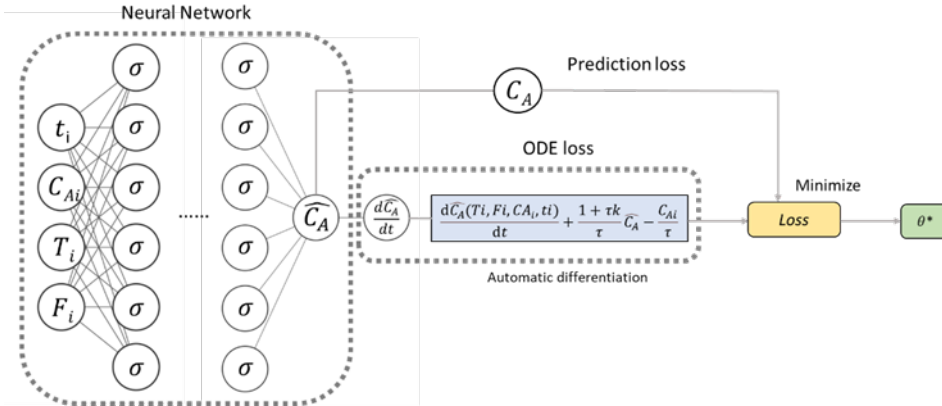


Figure 2. Training process of MFPINN

#### 4. Case Study

For the training and testing of both the proposed MFPINN model and the conventional FFNN model used for comparison, LHS was adopted to generate eight different operating conditions within the design space of  $C_{Ai}, F_i, T_i$  as listed in Table 1. The process was simulated under each operating condition for 150 seconds and the sampling time interval was set as 0.1 seconds for data generation. The hyperparameters

of the MFPINN model are given in Table 2. As shown in Table 2, the number of hidden layers and neurons per hidden layer of both the MFPINN and FFNN models are 8 and 100, respectively. Moreover, the number of collocation points collected for MFPINN is 100,000.

Table 2. Hyperparameter setup of both MFPINN and FFNN models

Number of hidden layers	8
Number of neurons per hidden layer	100
Number of collocation points	100,000

#### 4.1 Modeling with Small Data

In the first scenario, we used the simulation data collected from the 100<sup>th</sup> to 400<sup>th</sup> time intervals under each operating condition as the test set for both models. The training and validation sets with different sizes were randomly selected from the rest of the data. The ratio between the sizes of the training and validation sets was 4:1. The model test results show that, when the size of training set is relatively large, e.g. over 300, the performance of the two surrogate models are comparable. However, when the size of training set is reduced to 200, the MFPINN model is significantly better than the FFNN model, especially when applied to operating conditions 5~7. Tables 3 and 4 show the values of the coefficient of determination ( $R^2$ ) and mean squared error (MSE) of both models.

Table 3. Test results of MFPINN and FFNN when the size of training set is 300

	$R^2$ of MFPINN	$R^2$ of FFNN	MSE of MFPINN	MSE of FFNN
Condition 1	0.9964	<b>0.9996</b>	0.0002	<b><math>1.87 \times 10^{-5}</math></b>
Condition 2	<b>0.9987</b>	0.9952	<b><math>8.97 \times 10^{-5}</math></b>	0.0003
Condition 3	<b>0.9996</b>	<b>0.9996</b>	<b><math>2.57 \times 10^{-5}</math></b>	<b><math>2.57 \times 10^{-5}</math></b>
Condition 4	0.9962	<b>0.9990</b>	$6.49 \times 10^{-5}$	<b><math>1.79 \times 10^{-5}</math></b>
Condition 5	0.9777	<b>0.9817</b>	$6.68 \times 10^{-5}$	<b><math>5.46 \times 10^{-5}</math></b>
Condition 6	<b>0.9994</b>	0.9651	<b><math>3.89 \times 10^{-5}</math></b>	0.0023
Condition 7	0.9955	<b>0.9997</b>	0.0003	<b><math>1.90 \times 10^{-5}</math></b>
Condition 8	<b>0.9998</b>	0.9972	<b><math>7.40 \times 10^{-6}</math></b>	0.0001

Table 4. Test results of MFPINN and FFNN when the size of training set is 200

	$R^2$ of MFPINN	$R^2$ of FFNN	MSE of MFPINN	MSE of FFNN
Condition 1	<b>0.9898</b>	0.9534	<b>0.0005</b>	0.0021

Condition 2	0.9983	<b>0.9999</b>	0.0001	<b><math>8.09 \cdot 10^{-6}</math></b>
Condition 3	<b>0.9999</b>	0.9905	<b><math>7.10 \cdot 10^{-6}</math></b>	0.0006
Condition 4	0.9981	<b>0.9993</b>	$3.22 \cdot 10^{-5}$	<b><math>1.13 \cdot 10^{-5}</math></b>
Condition 5	<b>0.9183</b>	0.4434	<b>0.0002</b>	0.0017
Condition 6	<b>0.9862</b>	-0.9912	<b>0.0009</b>	0.1316
Condition 7	<b>0.9982</b>	0.3650	<b>0.0001</b>	0.0073
Condition 8	<b>0.9865</b>	0.9568	<b>0.0006</b>	0.0020

## 4.2 Extrapolation

In order to evaluate the extrapolation capability of the proposed model, the second scenario was designed, where all data collected under operating conditions 2 and 6 were only used for test, while the data of other operating conditions were divided to the training and validation sets under a ratio of 4:1. The test results are shown in Table 5, while Figure 3 compares the ground truth and the prediction values of different models. Obviously, the proposed MFPINN model outperforms its competitor. Because the MFPINN model is guided by the physical law during the model training process, it provides a predicted profile of  $C_A$  much closer to the high-fidelity simulation data, i.e. the ground truth. Its  $R^2$  values for operating condition 2 and 6 reach 0.967 and 0.977, respectively, although this is an extrapolation application. In contrast, although the training data is sufficient in this scenario, the FFNN model only achieves  $R^2$  values of 0.769 and 0.887 for operating condition 2 and 6, respectively. It is easy to understand that the extrapolation performance cannot be guaranteed unless the physics is taken into consideration.

Table 5 Regression analysis of MFPINN and FFNN models

	$R^2$ of MFPINN	$R^2$ of FFNN	MSE of MFPINN	MSE of FFNN
Condition 2	<b>0.9673</b>	0.7694	<b><math>7.10 \cdot 10^{-3}</math></b>	$5.03 \cdot 10^{-2}$
Condition 6	<b>0.9765</b>	0.8873	<b><math>4.80 \cdot 10^{-3}</math></b>	$2.28 \cdot 10^{-3}$

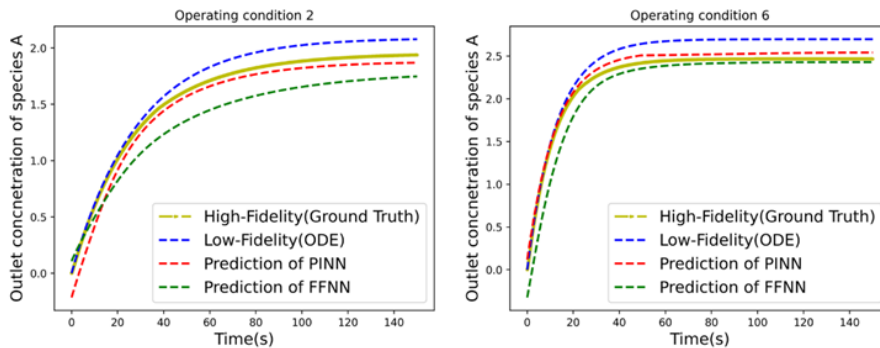


Figure 3. Predicted  $C_A$  of different models

## **5. Conclusions**

In this study, we integrate the high-fidelity simulation data and the low-fidelity physical law to build an MFPINN model for surrogate modeling. The experimental results show that, thanks to the guidance of the physical law, the proposed MFPINN outperforms the conventional FFNN when the training data is insufficient. Furthermore, the physical law comes in handy in the extrapolation. As a result, the MFPINN model predicts the outlet concentration of the CSTR system more accurately than the FFNN model. The application to the CSTR system illustrates the feasibility and potentials of the proposed MFPINN model.

## **Acknowledgment**

This work was supported in part by the National Science and Technology Council under project numbers NSTC 111-2622-8-007-011 and NSTC 111-2221-E-007-005.

## **References**

- A.G. Baydin, B.A. Pearlmutter, A.A. Radul and J.M. Siskind, (2018). Automatic Differentiation in Machine Learning: a Survey, *Journal of Machine Learning Research*, 18(1453), 1-43.
- S. Chakraborty, (2021). Transfer Learning Based Multi-Fidelity Physics Informed Deep Neural Network, *Journal of Computational Physics*, 426, 109942.
- M.G. Fernández-Godino, C. Park, N.H. Kim and R.T. Haftka (2019), Issues in Deciding Whether to Use Multifidelity Surrogates, *AIAA Journal*, 57(5), 2039-2054.
- S. Markidis, (2021), The Old and the New: Can Physics-Informed Deep-Learning Replace Traditional Linear Solvers?, [arXiv:2103.09655](https://arxiv.org/abs/2103.09655).
- P. Moin and K. Mahesh, (1998), DIRECT NUMERICAL SIMULATION: A Tool in Turbulence Research, *Annual Review of Fluid Mechanics*, 30(1), 539-578.
- M. Raissi, P. Perdikaris and G.E. Karniadakis, (2019), Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations, *Journal of Computational Physics*, 378, 686-707.
- M. Stein, (1987), Large Sampling Properties of Simulations Using Latin Hypercube Sampling, *Technometrics*, 29(2), 143-151.



# A CFD-model-based approach to continuous freezing process design for human induced pluripotent stem cells

Yusuke Hayashi<sup>a,\*</sup>, Benedikt X. Scholz<sup>a</sup>, Hirokazu Sugiyama<sup>a</sup>

<sup>a</sup>*Department of Chemical System Engineering, The University of Tokyo, Japan*

\**y-hayashi@pse.t.u-tokyo.ac.jp*

## Abstract

This work presents a computer fluid dynamics (CFD)-model-based approach to continuous freezing process design for human induced pluripotent stem (hiPS) cells. A process model was developed that combined a hybrid single-cell freezing model covering the cell level with a CFD model covering the cryovial and freezer levels. Given an inlet coolant temperature, an inlet coolant velocity, and a residence time, the developed hybrid model can calculate the cell survival rate as a quality indicator. We applied the developed process model to three operating conditions of continuous freezing processes for hiPS cells. As a result, the survival rate in one of the operating conditions could achieve more than 0.90. Thus, it was demonstrated that designing the continuous freezing process could be performed with appropriate value selection of the decision parameters.

**Keywords:** Regenerative medicine, Cryopreservation, Cell therapy, Hybrid model, Optimization.

## 1. Introduction

Human induced pluripotent stem (hiPS) cells (Takahashi et al., 2007) are regarded as one of the most promising sources for regenerative medical products because of various advantages over the conventional sources, e.g., human embryonic stem cells (Narsinh et al., 2011). Along with recent successful clinical studies, e.g., Parkinson's disease (Morizane, 2019) and retinitis pigmentosa (Tagawa et al., 2021), the implementation of medical treatments using hiPS cells is in progress.

In hiPS cell manufacturing, freezing processes are one of the most critical steps because the process is needed to the transportation and preservation of the cells. Generally, two methods can be adopted for the freezing process of hiPS cells: slow freezing and vitrification. Slow freezing has been applied to many cell types, where cryovials filled with cell suspension are cooled in a freezer at a predetermined cool rate. In vitrification, the vials are immediately cooled using liquid nitrogen. At a commercial scale, vitrification is rarely adopted because of vitrification is rarely adopted in the cell therapy industry because of a scale limitation and process complexity (Vajta and Nagy, 2006). Therefore, in this work, we focus on the slow freezing option.

In general, the slow freezing of hiPS cells has been performed by a direct contact freezer that can accommodate only a limited number of cryovials. On the other hand, with the expected future commercialization, there is a need to design continuous slow freezing processes for hiPS cells that can handle many cryovials. Most recently, a computer fluid dynamics (CFD) model-based approach to scale-up batch slow freezing processes for hiPS cells using forced convection-based freezers was published by the authors' research

group (Scholz et al., 2022). However, a process design of continuous slow cell freezing was yet to be performed.

This work presents a CFD-model-based approach to continuous freezing process design for hiPS cells. A process model was developed that combined a hybrid single-cell freezing model covering the cell level with a CFD model covering the cryovial and freezer levels. The cell level model was based on our previous publications (Hayashi et al., 2021, 2020), and the cryovial and freezer level models were newly constructed. Given an inlet coolant temperate, an inlet coolant velocity, and a residence time, the developed hybrid model can calculate the cell survival rate as a quality indicator. We applied the developed model to three operating conditions of continuous freezing processes for hiPS cells.

## 2. Continuous freezing processes for hiPS cells

The continuous freezing process assumed in this work is represented in Figure 1. The freezer adopted convection-based cooling and consisted of four stations according to the three-temperature zones proposed by Hayashi et al. (2021), i.e., dehydration (Station 1), nucleation (Stations 2 and 3), and further cooling (Station 4). The decision parameters in each station were defined as the inlet coolant temperate,  $T_i^{\text{st}}$  [K], the inlet coolant velocity,  $u_i^{\text{st}}$  [ $\text{m s}^{-1}$ ], and the residence time,  $t_i^{\text{st}}$  [s] ( $i = 1, 2, 3, 4$ ).

## 3. Process model

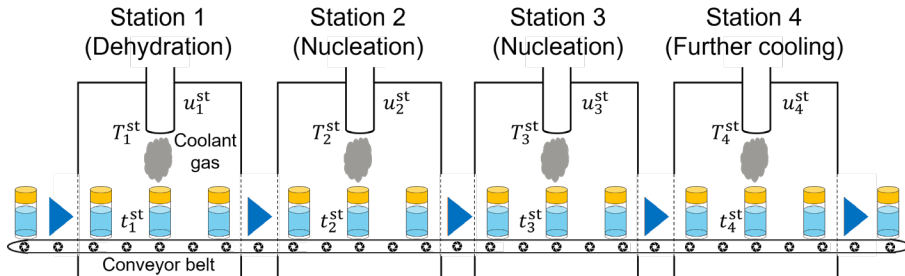
Figure 2 shows an overview of the developed process model consisting of the freezer, cryovial, and cell levels. The inputs were defined as the decision parameters, i.e.,  $T_i^{\text{st}}$ ,  $u_i^{\text{st}}$ , and  $t_i^{\text{st}}$ , and the output was defined as the cell survival rate,  $r_{\text{surv}}$  [-]. The freezer and cryovial level models constituted of the CFD model, and the cell model was based on Hayashi et al. (2021, 2020). The process model was defined, based on the following assumptions:

- The number of rows in the freezer is one.
- The values of the decision parameters are strictly controlled.
- The flow is laminar ( $Re_{\text{max}} = 2.0 \times 10^2$ ).

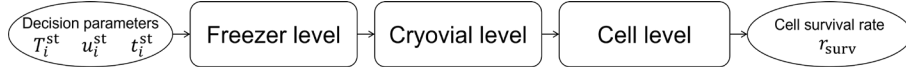
### 3.1. Freezer and cryovial level models

#### 3.1.1. Flow calculation

For the fluid flow calculation, the continuity equation and the Navier-Stokes equation were adopted as follows:



**Figure 1.** Continuous freezing processes for hiPS cells.



**Figure 2.** Overview of the developed process model.

$$\nabla \cdot v_f = 0 \quad (1)$$

$$\rho_f \frac{\partial v_f}{\partial t} + \rho_f (v_f \cdot \nabla) v_f = \nabla \cdot [-pI + \mu_f (\nabla v_f + (\nabla v_f)^T)] \quad (2)$$

where  $\rho_f$  [kg m<sup>-3</sup>] is the density,  $\mu_f$  [Pa s] is the dynamic viscosity,  $v_f$  [m s<sup>-1</sup>] is the velocity,  $p$  [Pa] is the pressure, and  $t$  [s] is the time.

### 3.1.2. Heat transfer

The energy balance of the fluid in the freezer was represented using the following equation:

$$\rho_f c_{p,f} \left[ \frac{\partial T}{\partial t} + (v_f \cdot \nabla) T \right] = -\nabla q_{\text{cond}} - \frac{T}{\rho_f} \frac{\partial \rho_f}{\partial T} \bigg|_p \left( \frac{\partial p}{\partial t} + (v_f \cdot \nabla) p \right) \quad (3)$$

where  $c_{p,f}$  [J kg<sup>-1</sup> K<sup>-1</sup>] is the specific heat capacity of the fluid,  $T$  [K] is the temperature,  $q_{\text{cond}}$  [J s<sup>-1</sup> m<sup>-3</sup>] is the conductive heat flux. The energy balance of the solid was also modeled as follows:

$$\rho_s c_{p,s} \frac{\partial T}{\partial t} = -(\nabla \cdot q_{\text{cond}}) \quad (4)$$

## 3.2. Cell level model

### 3.2.1. Mass transfer

The following equation was adopted to model the mass transport of water across the cell membrane:

$$\frac{dV_{\text{cell}}}{dt} = L_p A_{\text{cell}} \Delta \Pi \quad (5)$$

where  $V$  [m<sup>3</sup>] is the volume,  $L_p$  [m s<sup>-1</sup> Pa<sup>-1</sup>] is the water permeability,  $A$  [m<sup>2</sup>] is the surface area,  $\Delta \Pi$  [Pa] is the pressure difference. The normalized maximum cell volume change,  $\overline{\Delta V}_{\text{cell}}^{\text{max}}$  [-], was calculated as follows:

$$\overline{\Delta V}_{\text{cell}}^{\text{max}} = \max \left\{ \frac{|V_{\text{cell}}^{\text{fin}} - V_{\text{cell}}^{\text{init}}|}{V_{\text{cell}}^{\text{init}}} \right\} \quad (6)$$

where the superscripts initial and final represent the initial and final state of freezing, respectively.

### 3.2.2. Crystallization

The radius at time  $t$  of an intracellular crystal that nucleated at time,  $\tau$  [s], was estimated as follows:



$$r_{ice,i} = \begin{cases} 0 & (0 \leq N_{ice} < 1) \\ \sqrt[3]{\int_{\tau_i}^t \gamma_{ice}^2 \bar{D} dt} & (N_{ice} \geq 1) \end{cases} \quad (7)$$

where  $r_{ice}$  [m] is the radius of an ice crystal,  $N_{ice}$  [-] is the number of intracellular ice crystals,  $\gamma_{ice}$  [-] is the nondimensional ice crystal growth parameter, and  $\bar{D}$  [ $m^2 s^{-1}$ ] is the average water diffusion coefficient. The number of intracellular ice crystals,  $N_{ice}$ , was calculated using the following equations:

$$N_{ice} = \text{int}[\bar{N}_{ice}] \quad (8)$$

$$\frac{d\bar{N}_{ice}}{dt} = J_{ice} \quad (9)$$

where  $\bar{N}_{ice}$  [-] is the ensemble average of the number of intracellular ice crystals and  $J_{ice}$  [ $s^{-1}$ ] is the ice nucleation rate. The total volume of intracellular ice crystals,  $V_{ice}$  [ $m^3$ ], was modeled as follows:

$$V_{ice} = \sum_{i=1}^{N_{ice}} \frac{4}{3} \pi r_{ice,i}^3 \quad (10)$$

The normalized maximum ice crystal volume,  $\bar{V}_{ice}^{max}$  [-], was calculated using the following equation:

$$\bar{V}_{ice}^{max} = \max \left[ \frac{V_{ice}^{fin}}{V_{cell}^{fin}} \right] \quad (11)$$

### 3.2.3. Cell survival rate

The cell survival rate was estimated as follows (Hayashi et al., 2021):

$$r_{surv} = \omega_1 + \omega_2 \bar{\Delta V}_{cell}^{max} + \omega_3 \bar{V}_{ice}^{max} + \omega_4 \bar{\Delta V}_{cell}^{max} \cdot \bar{V}_{ice}^{max} \quad (12)$$

where  $\omega_1$  [-],  $\omega_2$  [-],  $\omega_3$  [-], and  $\omega_4$  [-] are the fitting coefficients, for which the same values published in Hayashi et al. (2021) were used in the calculation.

## 4. Results and discussion

Three operating conditions of continuous freezing processes for hiPS cells were applied as defined in Table 1. The conditions were selected based on the results presented in Hayashi et al. (2021). Figure 3 shows (a) the temperature profiles of the vial center and (b) the cell survival rates for the three operating conditions. The cell survival rate in Condition 1 was highest, followed by Condition 2 and Condition 3. This result can be explained using the three-temperature zones proposed by Hayashi et al. (2021). Above 233 K, fast cooling would be preferred because cell dehydration mainly occurred in the temperature range. In contrast, the cooling rate between 233 K and 213 K should be slow because intracellular ice nucleation mainly happened in the temperature range. In addition, below 213 K, any cooling rates could be applied because the cooling rate would

**Table 1.** Detailed values of the inlet coolant temperature,  $T_i^{st}$ , the inlet coolant velocity,  $u_i^{st}$ , and the residence time,  $t_i^{st}$ , applied in this work

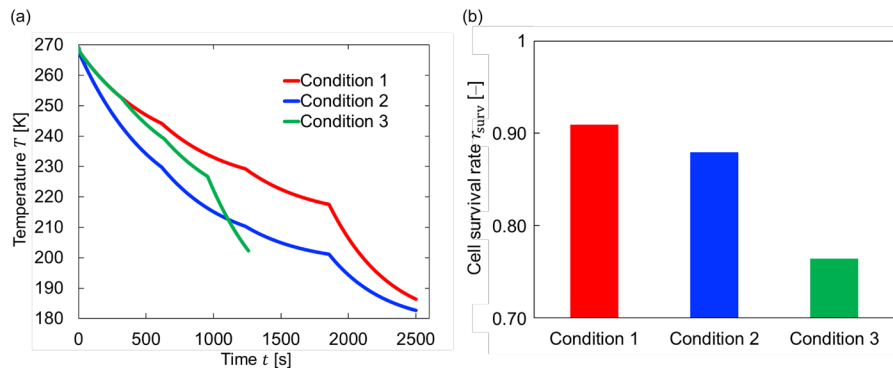
Station	Condition 1			Condition 2			Condition 3		
	$T_i^{st}$ [K]	$u_i^{st}$ [m s <sup>-1</sup> ]	$t_i^{st}$ [s]	$T_i^{st}$ [K]	$u_i^{st}$ [m s <sup>-1</sup> ]	$t_i^{st}$ [s]	$T_i^{st}$ [K]	$u_i^{st}$ [m s <sup>-1</sup> ]	$t_i^{st}$ [s]
1	233	0.100	625	213	0.100	625	233	0.100	315
2	223	0.100	625	203	0.100	625	223	0.100	315
3	213	0.100	625	198	0.100	625	213	0.100	315
4	178	0.100	625	178	0.100	625	178	0.100	315

not affect the cell survival rate. In the range of above 233 K, Condition 2 would be the best for the cell survival rate because the cooling rate was the fastest. On the other hand, in the range of between 233 K and 213 K, Condition 1 would be the most desirable for the cell survival rate because the cooling rate was the slowest. According to Mazur et al. (1972), cell damage due to intracellular ice nucleation was more critical than due to cell dehydration. Therefore, the highest survival rate (more than 0.90) was observed in Condition 1. Thus, it was demonstrated that designing continuous freezing processes for hiPS cells could be performed with appropriate value selection of the decision parameters.

The flow and heat transfer calculations were performed using COMSOL Multiphysics® version 6.1, and the calculations of mass transfer, crystallization, and cell survival rate were conducted by Python 3.9. The total CPU time that the results shown in Figure 3 were obtained was about 60 hours, using Intel® Xeon® Gold 6230 CPU @ 2.10 GHz with 512 GB RAM memory.

## 5. Conclusions and outlook

This work presented a CFD-model-based approach to continuous freezing process design for hiPS cells. A process model was developed that combined a hybrid single-cell freezing model covering the cell level with a CFD model covering the cryovial and freezer levels. The developed process model enabled calculating the cell survival rate of hiPS cells, given an inlet coolant temperature, an inlet coolant velocity, and a residence time. We



**Figure 3.** (a) Temperature profile of the vial center and (b) cell survival rate for the three operating conditions.

applied the developed process model to three operating conditions of continuous freezing processes for hiPS cells. As a result, the cell survival rate in one of the three operating conditions could achieve more than 0.90. Thus, it was demonstrated that designing the continuous freezing process could be performed with appropriate value selection of the decision parameters. In the field of computer-aided process engineering, cell therapy related studies are becoming relevant, e.g., Triantafyllou et al. (2022) and Hirono et al. (2022). Further model-based studies in this area are encouraged.

### Acknowledgements

H. S. is thankful for financial support by a Grant-in-Aid for Challenging Research (Exploratory) No. 20K21102 from the Japan Society for the Promotion of Science.

### References

- Y. Hayashi, I. Horiguchi, M. Kino-oka, H. Sugiyama, 2021, Model-based assessment of temperature profiles in slow freezing for human induced pluripotent stem cells, *Comput. Chem. Eng.*, 144, 107150.
- Y. Hayashi, I. Horiguchi, M. Kino-oka, H. Sugiyama, 2020, Slow freezing process design for human induced pluripotent stem cells by modeling intracontainer variation, *Comput. Chem. Eng.*, 132, 106597.
- K. Hirono, I. Udugama, Y. Hayashi, M. Kino-oka, H. Sugiyama, 2022, Design space determination of mesenchymal stem cell cultivation by dynamic modeling under uncertainty, *Comput. Aided Chem. Eng.*, 51, 721.
- P. Mazur, S. Leibo, E. Chu, 1972, A two-factor hypothesis of freezing injury: Evidence from Chinese hamster tissue-culture cells, *Exp. Cell Res.*, 71, 345.
- A. Morizane, 2019, Cell therapy for Parkinson's disease with induced pluripotent stem cells, *Rinsho Shinkeigaku*, 59, 119.
- K. Narsinh, J. Plews, J. Wu, 2011, Comparison of human induced pluripotent and embryonic stem cells: fraternal or identical twins?, *Mol. Ther.*, 19, 635.
- B. Scholz, Y. Hayashi, I. Udugama, M. Kino-oka, H. Sugiyama, 2022, A multilayered approach to scale-up forced convection-based freezing of human induced pluripotent stem cells, *Comput. Chem. Eng.*, 163, 107851.
- M. Tagawa, H. Ikeda, Y. Inoue, S. Iwai, Y. Iida, M. Hata, I. Asaka, A. Tsujikawa, 2021, Deterioration of phagocytosis in induced pluripotent stem cell-derived retinal pigment epithelial cells established from patients with retinitis pigmentosa carrying Mer tyrosine kinase mutations, *Exp. Eye Res.*, 205, 108503.
- K. Takahashi, K. Tanabe, M. Ohnuki, M. Narita, T. Ichisaka, K. Tomoda, S. Yamanaka, 2007, Induction of pluripotent stem cells from adult human fibroblasts by defined factors, *Cell*, 131, 861.
- N, Triantafyllou, A. Bernardi, M. Lakelin, N. Shah, M. Papathanasiou, 2022, A bi-level decomposition approach for CAR-T cell therapies supply chain optimization, *Comput. Aided Chem. Eng.*, 49, 2197.
- G. Vajta, Z. Nagy, 2006, Are programmable freezers still needed in the embryo laboratory? Review on vitrification, *Reprod. Biomed. Online*, 12, 779.

# Multiscale Modeling of Spatial Area-Selective Thermal Atomic Layer Deposition

Matthew Tom<sup>a</sup>, Sungil Yun<sup>a</sup>, Henrik Wang<sup>a</sup>, Feiyang Ou<sup>a</sup>,  
Gerassimos Orkoulas<sup>b</sup>, Panagiotis D. Christofides<sup>a,c,\*</sup>

<sup>a</sup>*Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA 90095-1592, USA*

<sup>b</sup>*Department of Chemical Engineering, Widener University, Chester, PA 19013, USA*

<sup>c</sup>*Department of Electrical and Computer Engineering, University of California, Los Angeles, CA 90095-1592, USA*

*pd@seas.ucla.edu*

## Abstract

Area-selective atomic layer deposition (ASALD) is an advancement of conventional thin-layer deposition processes that utilizes an additional chemoselective inhibition reaction to improve nanopatterning. However, ASALD is a challenging procedure that is difficult to characterize in the surface and fluid phases. A three-dimensional (3D) multiscale model that employs an atomistic-mesoscopic density functional theory (DFT) and kinetic Monte Carlo (kMC) simulation and macroscopic computational fluid dynamics (CFD) modeling are performed for a spatial rotary reactor configuration for the ASALD of a SiO<sub>2</sub>/Al<sub>2</sub>O<sub>3</sub> substrate *in silico*. Due to the complexity of the multiscale simulation, various reactor configurations are devised to minimize the computational requirements of the dynamic mesh of the reactor while reproducing results that are analogous to experimental data.

**Keywords:** Area-selective atomic layer deposition; multiscale modeling; computational fluid dynamics; kinetic Monte Carlo simulation; density functional theory.

## 1. Introduction

Area-selective atomic layer deposition is an attractive thin-layer deposition process that does not require subsequent processing steps (e.g., etching and lithography) to achieve effective nanopatterning and self-limiting deposition on growth areas of the substrate. In particular, semiconducting materials require thin-layer deposition processes to promote the self-aligned stacking of transistors, which is essential for producing high-performance wafers. In accordance with the predictions of Moore's Law, the stacking of the transistors is needed to improve the computational performance and efficiency of electronics; however, the lack of self-alignment introduces a fabrication challenge for wafers. Although conventional thin-layer deposition reactors are widely used, these reactors require extensive purging steps to maintain the self-limiting nature of the process. Spatial reactor configurations, including the rotary type, introduce reagent in physically isolated regions to prevent intermixing while maintaining the self-limiting tendency of the reactions (Poody et al., 2012). The stringent dimensions of these transistors require much experimentation through a tedious procedure of optimizing the reactor design such that reagent intermixing is minimized and substrate exposure to the reagents is uniformly distributed. However, multiscale modeling has emerged as a cost-effective and timely solution for relating multiphase deposition processes for various time and length scales while reproducing experimental results with high fidelity. This work will employ an atomistic-mesoscopic density

functional theory (DFT) and kinetic Monte Carlo model in conjunction with a three-dimensional (3D) macroscopic computational fluid dynamics (CFD) to simultaneously simulate surface-scale reaction kinetics comprising steric hindrance and repulsion effects and macroscopic fluid transport phenomena for a silicon wafer ( $\text{SiO}_2/\text{Al}_2\text{O}_3$ ) substrate in accordance to the ABC-type reaction kinetics proposed by Merckx et al. (2020) and Mameli et al. (2017) composed of (A) inhibition, (B) adsorption, and (C) oxidation steps. Additionally, the relation of various constraints in the design of the rotary reactor (e.g., gap distance, geometries of reaction zones, and orientations of inlets) to the surface uniformity and deposition growth rate are studied for process optimization. Such constraints provide significant insight for application to current fabrication procedures.

## 2. Multiscale Simulation

### 2.1. Atomistic-mesoscopic modeling

Area-selective atomic layer deposition (ASALD) processes integrate a variety of synergistic, chemoselective, and self-limiting reactions that deposit monolayers of substrate material. Mameli et al. (2017) proposed an ABC-type mechanism composed of three steps (Steps A, B, and C). First, an inhibition preprocessing step (Step A) to deactivate subsequent reactions in the non-growth area using a small molecule inhibitor, acetylacetone (Hacac). Next, a precursor, bis(diethylamino)silane (BDEAS), adsorbs onto the growth area to produce a modified surface layer (Step B) composed of hydrogen-terminated ligands. Finally, an oxidation step (Step C) using ozone ( $\text{O}_3$ ) oxidizes the hydrogen terminals to produce hydroxyl-ligands, thereby concluding a single monolayer of the deposition cycle. To ensure the self-limiting behavior of these reactions, each step is followed by a purging action by employing an inert gas such as nitrogen,  $\text{N}_2$ .

The abundance of reaction mechanisms and the lack of reference and experimental data would make the characterization of the kinetics difficult to quantify in the atomistic and mesoscopic phases. Thus, *ab initio* quantum mechanics that uses first principles density functional theory (DFT) is performed through the open-source software, Quantum ESPRESSO (QE) to compute the thermophysical and kinetics parameters of species and reactions, respectively. QE performs molecular structure and electronic optimization through pseudopotential data and nudged elastic band (NEB) methods to compute the minimal energy path to compute the activation energy of the reaction, to be substituted into the Arrhenius Equation. Additionally, a typical elementary reaction mechanism for rate-determining reaction steps is proposed for each step (A, B, and C) in the ASALD process to integrate transition state theory (TST) for non-adsorption reactions and collision theory (CT) for adsorption reactions.

With the kinetics mechanisms and reaction parameters defined, reaction parameters are defined into the kinetic Monte Carlo (kMC) algorithm that adopts a stochastic procedure to exemplify the behavior of realistic reactions. The kMC method, performed through the Python programming language, adopts a sequential procedure that sums probable reaction rates at a surface site to preferentially weight reactions of greater magnitude while devising a random number to arbitrarily select a reaction for the active site. Such procedure enables the computation of species generation and consumption terms from surface pressure and temperature data (from the CFD simulation), which will be integrated into the CFD simulation. Subsequently, the kMC method simulates the progression of the reaction path by calculating process time through a secondary random number described by Yun et al. (2022a). Additionally, steric hindrance effects are simulated by randomly selecting the angle of rotation for adsorption reactions encountered in Step A, and portraying the deflection of adsorption sites that are hindered by bulky molecules (Roh et al., 2022).

## 2.2. Macroscopic modeling

Computational fluid dynamics (CFD) is performed to simulate the transport phenomena in the fluid regime through the commercial multiphysics software, ANSYS Fluent. Prior to conducting the CFD simulation, a three-dimensional (3D) rotary reactor model, illustrated in Fig. 4 is constructed through ANSYS DesignModeler, which is then meshed through ANSYS Workbench's meshing software using tetrahedral cells, until an ideal mesh quality (per ANSYS standards) and minimal number of cells is obtained. An ideal mesh is composed of a minimal number of cells to lessen the dependence of robust computational power while maintaining accurate computations that are reflective of experimental results recorded in academia and ensuring convergence is obtainable. Additionally, the movement of the wafers between reaction zones are achieved through a dynamic mesh method that performs remeshing and diffusion-based smoothing methods to retain mesh quality and preserve solution convergence. In Fig. 4, the reactor design separates each wafer into 10 sections to define source generation and consumption flux terms (evaluated by the kMC simulation) on the wafer surface through user-defined functions (UDFs) in a "C++ like" script. The CFD simulation is performed using a fixed time step of 0.001 s, employing a pressure-based solver method, and adopting a first-order implicit numerical solver method that is defined through a journal file and executed through a Scheme script. Lastly, various thermophysical parameters are defined by integrating results from the atomistic *ab initio* first principles quantum mechanics simulations and through the ANSYS Fluent database of gaseous molecules.

In addition to the optimization of the reactor mesh, the chamber design of the rotary reactor and the operating conditions have a significant role in the quality and conformance of the substrate (De la Huerta et al., 2018). The reactor design must ensure that effective reagent separation is made to prevent species intermixing, which disrupts the perceived self-limiting behavior of the ASALD process. Reagent intermixing is prevented by specifying an optimal gap distance, the distance between the substrate surface and the ceiling of the reactor that separates the reaction zones, introducing output streams at precise locations to remove byproducts generated, and defining a large purge flow rate. Following the procedures to minimize reagent intermixing, the delivery of gases to the surface of the wafer and the exposure time are of paramount importance to ensure that a total surface deposition is achieved. Constraints including the reagent (Hacac, BDEAS, O<sub>3</sub>) flow rates and rotation speed have a profound impact on the deposition rate onto the substrate surface (Pan et al., 2017). Although large reagent flow rates and low rotation speeds ensure complete deposition is obtained, such quantities are also nonideal from an economics and production perspective. Lastly, the uniformity of the deposition is a qualitative parameter that discusses the effectiveness of the nanopatterning performance. This modification provides better control of the film uniformity to ensure that all regions in the wafer have the same exposure time. The aforementioned quantities provide a meaningful study for optimizing the reactor design and process operation for practical applications.

## 2.3. Multiscale modeling

The multiscale simulation allows the calculation of various time and length scale in various domains including the atomistic, mesoscopic, and macroscopic regimes (Raimondeau and Vlachos, 2002) as described in Fig. 1. Such multiscale modeling allows the relating of various phases that cannot be discerned through *ex silico* methods. The multiscale simulation is performed autonomously through a Linux cluster system through the UCLA Hoffman2 Cluster and allows the interconnection of various programming languages and application software. This work will utilize 36-core compute nodes consisting of 192 GB

of dynamic random-access memory (DRAM) that are standard for ANSYS Fluent to conduct the multiscale simulation by adopting both a parallel computation strategy in the macroscopic simulation and serial method for the microscopic simulation taking an average of 1 to 3 days, which depends on the rotation speed of the reactor. Such coding logic allows ANSYS Fluent to evaluate surface pressure ( $P$ ) and temperature ( $T$ ) data on the wafer surface to be read by the kMC coding script to evaluate the source generation and consumption terms,  $S_h$  and  $S_{m,i}$ , for the heat and mass of species  $i$  balances, which will then be defined through the UDF by ANSYS Fluent as summarized in the process diagram in Fig. 2. The simulation will conclude until the termination condition (full coverage) is obtained or if the wafer exits the reaction zone completely.

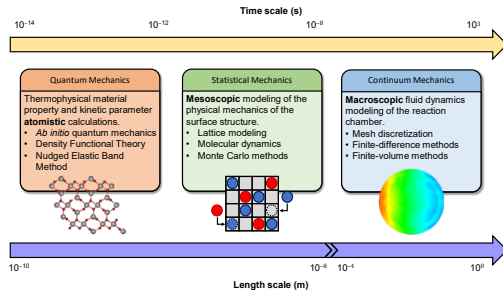


Figure 1: Multiscale model application for various phases and time and length domains.

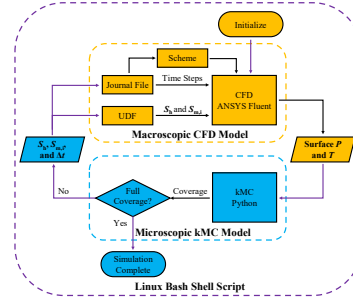


Figure 2: Multiscale model simulation process diagram.

### 3. Results

#### 3.1. Atomistic-mesoscopic simulation results

The atomistic-mesoscopic simulation is conducted for a pressure range of 10 to 500 Pa in 10 Pa intervals and a temperature range of 423 to 573 K in 10 K intervals for a total of 800 data points. The collection of a diverse data set is employed to determine the effects of the operating conditions on the total process time to obtain full coverage on the wafer surface for Steps A, B, and C on a  $100 \times 100$  lattice. Graphical results for the process times as a function of the temperature and pressure are provided in Fig. 3. Results from

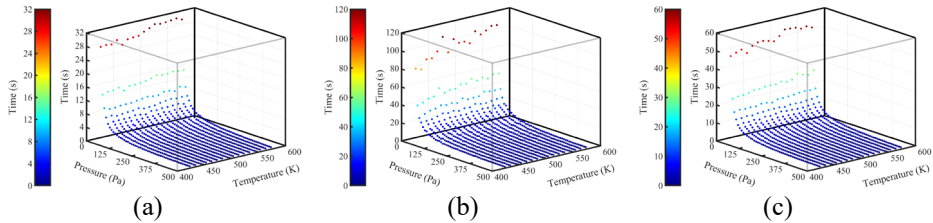


Figure 3: Kinetic Monte Carlo process times for (a) Hacac inhibition, (b) BDEAS adsorption, and (c)  $O_3$  oxidation at various pressure and temperature conditions.

Step A in Fig. 3(a) demonstrates that the Hacac pulse time of 3.0 s or more above was observed for pressures below the 100 Pa, which was analogous to reported findings by Mameli et al. (2017) who reported a saturation dosage time of 5.0 s for such operating conditions. Additionally, the BDEAS adsorption in Step B displayed in Fig. 3(b) has a BDEAS dosage time of 2.3 s for a pressure of 400 Pa to reach full coverage, which is comparable to the reported findings of 2.0 s of BDEAS saturation times by Merckx et al. (2020). The aforementioned results indicate that the combined atomistic and mesoscopic simulations are reflective of experimental results; thus, the kMC simulation will have a minimal contribution to deviation of results in the multiscale simulation.

3.2. Multiscale computational fluid dynamics simulation results

The reactor mesh is fully optimized using less than 1.3 million tetrahedral cells and a gap distance of 5 mm by adopting a reactor geometry with reaction barrier zones conforming to the radial direction of the reactor as illustrated in Fig. 4. In addition to the radial separation of the vacuum ports, the wafers were also separated into 10 regions in the radial direction for the UDF specifications of the source terms. Results from the multiscale CFD simulation in Fig. 5 reveal that the dividers provide substantial reagent exposure in the radial direction when rotating with an angular velocity of 0.40 rad/s. The

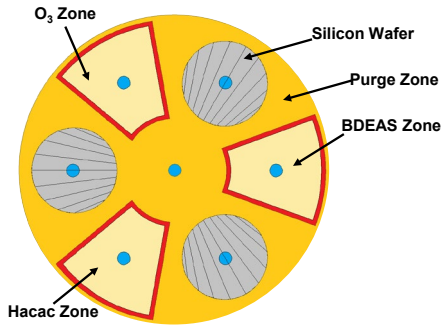


Figure 4: Top view of the rotary reactor design. Inlet and outlet streams are in blue and red, respectively.

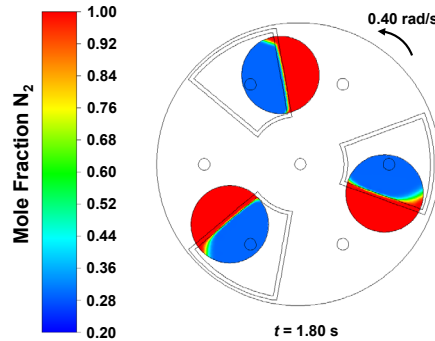


Figure 5: Multiscale CFD simulation results illustrated in N<sub>2</sub> mole fraction wafer surface contours.

contours of mole fraction of N<sub>2</sub> also illustrate that reagent intermixing is effectively minimized, thereby maintaining the self-limiting nature of the ASALD process. However, it is notable that intermixing is visible in reaction zone corners; thus, further study is needed to minimize the effects of the purge and reagents intermixing within the reaction zone by

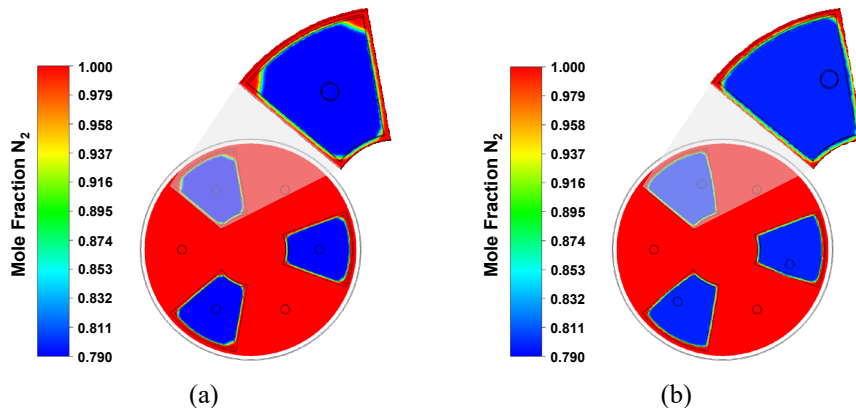


Figure 6: Contours of N<sub>2</sub> mole fraction illustrating intermixing for a conventional rotary reactor design (a) and the mitigation of intermixing by the asymmetric design (b).

adjusting the orientation of the inlets, modifying the diameters of the outlets, and adjusting the reaction zone geometry. The optimization of the reactor configuration is illustrated in Fig. 6 to demonstrate the effects of adjusting the inlet location to an asymmetric model in Fig. 6b, which moderates the effects of purge transfer into the reaction zones for the conventional reactor configuration in Fig. 6a. Also, the process operation parameters can be considered for further study such as balancing the reagent flow rates with a proper



composition of reagent. The collection of a diverse data set is still needed to further characterize various constraints (e.g., exposure uniformity and deposition growth rate).

#### 4. Conclusion

A novel multiscale simulation of an area-selective ASALD process was performed to study the relation between various operating conditions and reactor geometry on the uniformity and deposition rate of a thin film on a  $\text{SiO}_2/\text{Al}_2\text{O}_3$  substrate in various phases. First, an atomistic-mesoscopic model was developed in conjunction with *ab initio* first principles quantum mechanics simulation to calculate kinetic parameters of rate-determining elementary reactions and thermophysical parameters of species and with a kinetic Monte Carlo (kMC) algorithm to simulate the stochastic nature of surface reactions on the wafer. The resulting mesoscopic simulation was performed with computational fluid dynamics (CFD) to study the transport phenomena of the gases. It was determined that the atomistic-mesoscopic model, when compared to experimental results found in literature, demonstrated accurate process time results in exact operating conditions. The multiscale CFD simulation illustrated that the reactor design successfully minimized the effects of reagent intermixing while improving exposure uniformity to enhance nanopatterning. Although further study is needed to characterize the relation of various reactor design and operation constraints to fully understand their effects on the deposition growth rate and film uniformity, initial results reveal that the reactor design and operation conditions determined by the model lead to improved deposition performance.

#### References

- De la Huerta, C.M., Nguyen, V.H., Dedulle, J.M., Bellet, D., Jiménez, C., Muñoz-Rojas, D., 2018. Influence of the geometric parameters on the deposition mode in spatial atomic layer deposition: a novel approach to area-selective deposition. *Coatings* 9, 5.
- Mameli, A., Merckx, M.J.M., Karasulu, B., Roozeboom, F., Kessels, W.E.M.M., Mackus, A.J.M., 2017. Area-selective atomic layer deposition of  $\text{SiO}_2$  using acetylacetone as a chemoselective inhibitor in an ABC-type cycle. *ACS Nano* 11, 9303–9311.
- Merckx, M.J.M., Sandoval, T.E., Hausmann, D.M., Kessels, W.M.M., Mackus, A.J.M., 2020. Mechanism of precursor blocking by acetylacetone inhibitor molecules during areaselective atomic layer deposition of  $\text{SiO}_2$ . *Chemistry of Materials* 32, 3335–3345.
- Pan, D., Jen, T.C., Yuan, C., 2016. Effects of gap size, temperature and pumping pressure on the fluid dynamics and chemical kinetics of in-line spatial atomic layer deposition of  $\text{Al}_2\text{O}_3$ . *International Journal of Heat and Mass Transfer* 96, 189–198.
- Poodt, P., Cameron, D.C., Dickey, E., George, S.M., Kuznetsov, V., Parsons, G.N., Roozeboom, F., Sundaram, G., Vermeer, A., 2012. Spatial atomic layer deposition: A route towards further industrialization of atomic layer deposition. *Journal of Vacuum Science & Technology A* 30, 010802.
- Raimondeau, S., Vlachos, D.G., 2002. Recent developments on multiscale, hierarchical modeling of chemical reactors. *Chemical Engineering Journal* 90, 3–23.
- Roh, H., Kim, H.L., Khumaini, K., Son, H., Shin, D., Lee, W.J., 2022. Effect of deposition temperature and surface reactions in atomic layer deposition of silicon oxide using bis(diethylamino)silane and ozone. *Applied Surface Science* 571, 151231.
- Yun, S., Tom, M., Luo, J., Orkoulas, G., Christofides, P.D., 2022a. Microscopic and data-driven modeling and operation of thermal atomic layer etching of aluminum oxide thin films. *Chemical Engineering Research and Design* 177, 96–107.
- Yun, S., Tom, M., Orkoulas, G., Christofides, P.D., 2022b. Multiscale computational fluid dynamics modeling of spatial thermal atomic layer etching. *Computers & Chemical Engineering* 163, 107861.
- Yun, S., F. Ou, H. Wang, M. Tom, G. Orkoulas and P. D. Christofides, 2022c. Atomistic-mesoscopic modeling of area-selective thermal atomic layer deposition. *Chemical Engineering Research and Design* 188, 271–286.

# Modeling the shrinkage effect during the drying process of a green ceramic material

*Achilleas L. Arvanitidis,<sup>a</sup> Margaritis Kostoglou,<sup>b</sup> Michael C. Georgiadis<sup>a,\*</sup>*

*<sup>a</sup>Department of Chemical Engineering, Aristotle University of Thessaloniki, University Campus, Thessaloniki, 54124, Greece*

*<sup>b</sup>Department of Chemistry, Aristotle University of Thessaloniki, University Campus, Thessaloniki, 54124, Greece*

*Corresponding author, e-mail: [mgeorg@auth.gr](mailto:mgeorg@auth.gr)*

## Abstract

In this work, a 1-D diffusion model is developed to describe the drying process of a ceramic slab with respect to shrinkage. The proposed mathematical framework consists of a set of partial differential and algebraic equations along with moving boundary conditions, making it suitable for predicting the drying behavior and the physical phenomena that accompany the drying process of a green ceramic slab, such as shrinkage and porosity increase. Initially, the developed model is validated against experimental data from the open literature. Then, the drying behavior of a clay material, produced at a local factory, is investigated. This requires the characterization of the material with respect to its desorption kinetics. The parameter values of a semi-empirical desorption isotherm equation are estimated based on drying tests conducted in the lab. Finally, the drying kinetics of the porous material is studied and the shrinkage effect is discussed.

**Keywords:** Ceramic drying, shrinkage, porosity, moving boundary problem

## 1. Introduction

The drying process constitutes an essential step in the ceramic sector, which aims at removing the water from the green ceramic material before firing takes place. This is a critical production step because if the material is inserted wet into the extreme environment conditions of the firing chamber, then the probability of moisture evaporation inside the pores of the material would be high. As a consequence, this could lead to the development of intense stresses inside the porous body that can potentially lead to cracks and fracture. Cracks could also occur during the drying process itself, because of the non-uniform moisture distribution along the body and hence, non-uniform shrinkage (Kowalski and Pawłowski, 2010).

The understanding of the shrinkage mechanism during the drying procedure of a ceramic green body plays a key role for the product quality assurance and the fabrication of new drying strategies. Mancuhan et al. (2016) studied the shrinkage behavior of ceramic green bricks through Bigot curves. They proposed that during the initial stages of drying, the volume variation of the material is equal to the volume of the removed water. However, as drying proceeds, a critical moisture content is reached. This indicates that most clay particles of the solid matrix are already in direct contact and cannot move any further. At that point, shrinkage is almost complete and drying proceeds without any volume variations taking place.

It is a common practice in the open literature, to calculate the drying-induced shrinkage using mechanical interaction models which use strain – stress relationships (Hammouda and Mihoubi, 2014). Other approaches use more simplistic methods to account for the

volume variation of a porous material during drying. Barbosa de Lima et al. (2016) used polynomial expressions to correlate the moisture content of a clay material to its volumetric strain. However, this method of approaching the shrinkage calculation does not provide any physical insight for the drying process.

In this work, the drying-induced shrinkage calculation is approached by using a shrinkage velocity, which reflects the velocity at which a specific nodal point of the clay matrix recedes to the interior of the bulk body. The proposed modeling framework leads to high quality predictive results with low computational complexity and facilitates the numerical calculation of key physical quantities, such as the solid concentration and the body porosity, which are closely associated to qualitative characteristics of the green material.

## 2. Mathematical model

In this section, a mathematical model is developed to capture the drying behavior of a shrinking ceramic roof tile. Since the material aspect ratio is very small, its geometry is simplified to that of a semi-infinite tile. Consequently, the model considers that heat and mass transfer phenomena occur only on the thickness dimension of the tile.

The evolution of the water mass concentration,  $C_w$  and temperature,  $T$ , with time along the green material is given by equations. (1) and (2), respectively. Both equations also consider the shrinkage effect via a local shrinkage velocity term,  $v_{shr}$ , which determines the velocity at which a nodal point of the clay matrix moves towards the bulk body.

$$\frac{\partial C_w}{\partial t} = \frac{\partial}{\partial z} \left[ D_{eff} \frac{\partial C_w}{\partial z} - v_{shr} C_w \right] \quad , \quad z \in (0, L) \quad (1)$$

$$\rho c_p \frac{\partial T}{\partial t} = \frac{\partial}{\partial z} \left[ k_{eff} \frac{\partial T}{\partial z} - v_{shr} \rho c_p T \right] \quad , \quad z \in (0, L) \quad (2)$$

Where  $D_{eff}$  is the effective diffusivity of the moisture and  $k_{eff}$  is the effective thermal conductivity of the material. Both properties are constant for the present study.

Preheated air flows over the wet tile and drying is achieved through forced convection. This model assumes that moisture diffuses from the bulk body towards the air-solid interface, where it instantaneously evaporates and desorbs to the blowing air. As such the mass balance boundary equation is given by equation (3).

$$-D_{eff} \frac{\partial C_w}{\partial z} \Big|_{z=L} = h_m \frac{MW_w}{R_g} \left( \frac{P_{sat}(T_{z=L})}{T_{z=L}} RH_{eq} - \frac{P_{sat}(T_{air})}{T_{air}} RH \right) \quad (3)$$

Where,  $h_m$ , is the convective mass transfer coefficient,  $MW_w$  is the water molecular weight and  $R_g$  is the global gas constant. Also  $P_{sat}$  is the saturation pressure of water and  $RH$  is the relative humidity of the air, which is defined as  $P_w/P_{w,sat}$ .

It is assumed that at the air - solid interface, the moisture content is at chemical equilibrium with the air humidity. This is given by a desorption isotherm correlation.

$$RH_{eq} = f(X|_{z=L}, T|_{z=L}) \quad (4)$$

Where  $X$  is the moisture mass fraction (dry base) of the material and is defined as  $X = C_w/C_s$ .

As the preheated air flows over the colder wet tile surface, heat is transferred through convection. Additionally, the latent heat of vaporization,  $\Delta H_{vap}$ , is abducted from the drying surface of the material to achieve the water evaporation.

$$-k_{eff} \frac{\partial T}{\partial z} \Big|_{z=L} = h_T(T|_{z=L} - T_{air}) - \Delta H_{vap} D_{eff} \frac{\partial C_w}{\partial z} \Big|_{z=L} \quad (5)$$

The bottom surface of the green body is not in direct contact with the air.

$$\frac{\partial C_w}{\partial z} \Big|_{z=0} = 0 \quad (6)$$

$$\frac{\partial T}{\partial z} \Big|_{z=0} = 0 \quad (7)$$

The only mechanism that affects the solid concentration is the shrinkage-induced convection. Hence, the temporal evolution of the solid concentration  $C_s$ , is given by (8).

$$\frac{\partial C_s}{\partial t} = - \frac{\partial [v_{shr} C_s]}{\partial z} \quad , \quad z \in (0, L) \quad (8)$$

The local porosity,  $\varepsilon$ , is calculated from the phase volume balance.

$$1 = \varepsilon + \frac{C_s}{\rho_{intrinsic}} + \frac{C_w}{\rho_w} \quad (9)$$

Where,  $\rho_{intrinsic}$  and  $\rho_w$  are the intrinsic solid density and water density respectively. Since all heat and mass transfer phenomena are considered exclusively on the thickness direction, the tile thickness,  $L$ , evolution should be described by equation (10).

$$\frac{dL}{dt} = v_{shr} \Big|_{z=L} \quad (10)$$

The shrinkage velocity is a local term, which shows the velocity at which a nodal point recedes to the interior of the body. More specifically, the shrinkage velocity is the velocity at which a volume of water diffuses towards the air – solid interface.

$$v_{shr} = \int_0^z \alpha \frac{\partial}{\partial z} \left[ \frac{D_{eff}}{\rho_w} \frac{\partial C_w}{\partial z} \right] dz \quad , \quad z \in [0, L] \quad (11)$$

Where  $\alpha$  is a proportionality factor. Its physical meaning is that when  $\alpha$  equals to 1, the volume variation of the porous body is respective to the volume variation of the removed water (ideal shrinkage case). When  $\alpha = 0$ , no volume variation takes place as water leaves the body (no shrinkage case).

As explained earlier, during the initial stages of drying, the shrinkage is ideal (e.g.  $\alpha = 1$ ). However, when the moisture content,  $X$ , reaches a critical value,  $X_{cr}$ , almost all shrinkage phenomena are complete because most of the clay particles are already in direct contact with each other (Mançuhan et al., 2016). At that point, shrinkage stops (e.g.  $\alpha = 0$ ) and the porosity of the material starts to increase since air penetrates into the body to substitute the volume of the removed water.

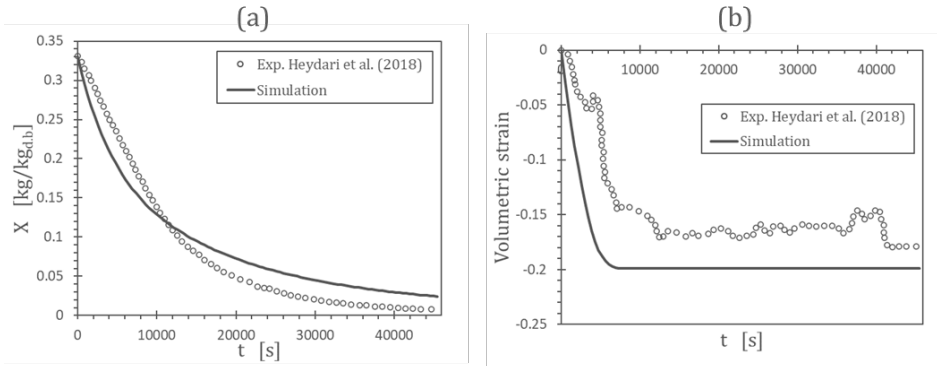
$$\alpha = \begin{cases} 1, & X \geq X_{cr} \\ 0, & X < X_{cr} \end{cases} \quad , \quad z \in [0, L] \quad (12)$$

### 3. Solution Strategy

The integration of the above system of equations on the  $(z, t)$  plane is not trivial because the boundary conditions are located on a moving boundary. The position of the  $z = L$  node must be determined as an integral part of the solution at any time. That for, a simple Lagrangean transformation is used, by introducing the spatial variable  $y = z/L$  to immobilize the moving boundary. The  $y$  spatial variable lies between 0 and 1, and enables the normal discretization of the thickness dimension. The newly transformed model is implemented into gPROMS<sup>TM</sup> modeling environment. The partial differential equations are approximated using 2<sup>nd</sup> order Central Finite Differences and the differential – algebraic solver DASOLV is used for the integration.

### 4. Model Validation

For the validation of the model, a suitable set of experimental data from the open literature was used. The experimental measurements that are presented in the work of Heydari et al. (2018) facilitate the corroboration of the model predicted shrinkage behavior of the green ceramic body. In Fig. 1, the model predicted moisture content and volumetric strain is plotted against the experimental measurements.



**Fig. 1:** Comparison of experimental data (Heydari et al., 2018) and model predictions for (a) moisture content and (b) volumetric strain.

The results indicate that a sufficiently good agreement is achieved. The deviations of the model predictions from the experimental points are ought to the fact that the studied material of Heydari et al. (2018) has much larger aspect ratio than the semi-infinite plate that is of interest in the present work. It is expected that the use of a higher order model could result to a better match between the model predictions and the experimental data.

### 5. Parameter Estimation

The material of interest in this work is a clay tile that is produced in a large-scale production plant in Greece. The simulation of the drying process of the material requires the establishment of its desorption kinetics. That for, equation (4) has to be substituted by an equation that correlates the relative humidity to the moisture content of the material and its temperature. For this reason, the GAB model is selected.

$$X_{eq} = \frac{Ckm_0RH_{eq}}{(1 - kRH_{eq})(1 - kRH_{eq} + CkRH_{eq})} \quad (13)$$

Where  $C, k, m_0$  are temperature dependent parameters and are described by Arrhenius type equations :  $C = C' e^{-\frac{\Delta H_C}{R_g T}}$ ,  $k = k' e^{-\frac{\Delta H_k}{R_g T}}$  and  $m_0 = m_0' e^{-\frac{\Delta H_{m_0}}{R_g T}}$ . The parameters  $C', k', m_0'$  and  $\Delta H_C, \Delta H_k, \Delta H_{m_0}$  are estimated based on drying tests that were conducted in the lab and their values are illustrated in Table 1.

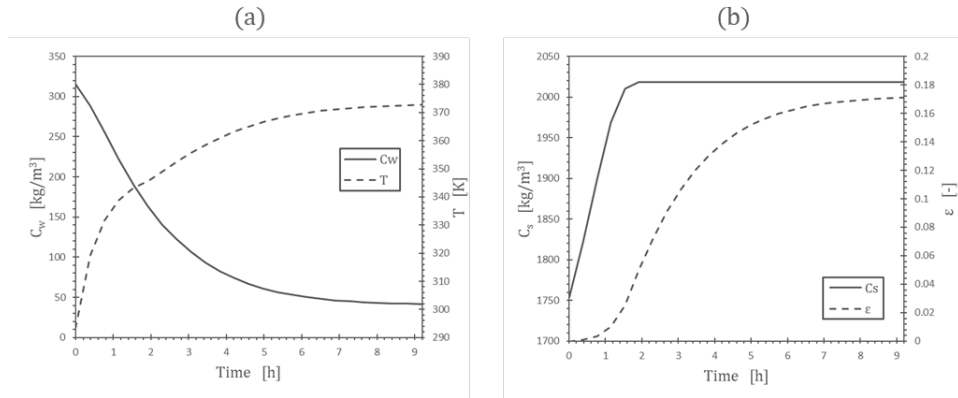
**Table 1:** GAB model parameters

GAB parameters					
$C''$	$k''$	$m_0''$	$\Delta H_{C'}$	$\Delta H_{k'}$	$\Delta H_{m_0'}$
0.195	0.0254	0.175	$-1.62 \times 10^7$	$-5.05 \times 10^6$	$2.42 \times 10^5$

## 6. Results and Discussion

Simulation results of the drying process for the studied material are summarized here, to gain insight on how the shrinkage phenomenon affects the dynamics of various quantities of the process. The model inputs required for the simulation are summarized as follows. The blowing air temperature,  $T_{air}$ , is  $100^\circ\text{C}$  and the air relative humidity,  $RH$ , is 3%. The initial water concentration and initial solid concentration are  $316 \text{ kg/m}^3$  and  $1754 \text{ kg/m}^3$  respectively. The initial tile thickness is  $1.87 \text{ cm}$  and the critical moisture content, which determines when shrinkage stops, is  $X_{cr} = 0.1 \text{ kg}_w/\text{kg}_{d.s.}$ . As for the heat and mass transfer coefficients, their values are chosen arbitrarily as  $28 \text{ W/m}^2\text{K}$  and  $0.028 \text{ m/s}$ , respectively.

Fig. 2 illustrates the water concentration and body temperature time profiles. It is observed that the water concentration inside the body decreases with time as its temperature rises. Finally, a steady state is reached and the drying procedure is complete.

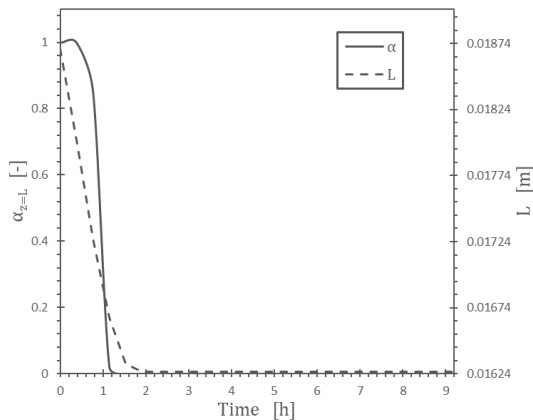


**Fig. 2:** (a) Mean water concentration and temperature vs time and (b) Mean solid concentration and body porosity vs time

The solid concentration and porosity temporal evolutions are illustrated in Fig. 2(b). At the beginning of the process, the body porosity remains unchanged and the solid concentration rises abruptly. This is justified by the fact that during these times shrinkage is ideal and  $\alpha = 1$ . When the critical moisture content,  $X_{cr}$ , is achieved, a transition to the no-shrinkage state occurs and shrinkage stops. Consequently, the solid concentration stops to vary and air starts to replace the removed water volume, which leads to the porosity increase.

In Fig. 3, the shrinkage parameter on the moving boundary and the tile thickness are plotted against time. It is observed that the body shrinks for some time despite that the shrinkage parameter  $\alpha|_{z=L}$  obtains the zero value. This is because even if  $X < X_{cr}$  at the

boundary,  $z = L$ , shrinkage continues to take place to the interior of the body, as long as  $X > X_{cr}$ .



**Fig. 3:**  $\alpha$  parameter on the boundary  $z = L$  and slab thickness vs time

## 7. Conclusions

A 1-dimensional model is developed to predict the drying behavior of a green ceramic tile with respect to shrinkage. The proposed model facilitates the calculation of key physical quantities that have not been studied extensively in the literature. Such quantities are the local solid concentration and porosity. The prediction of the transient evolution of these quantities is crucial because they are severely associated to most qualitative characteristics of the green product.

## Acknowledgments

This research has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the specific call “Aquacultures”- “Industrial Materials”- “Open Innovation in Civilization” (project code: T6YBII-00251). Project title: Development of Computer-Aided Tools for Optimal Energy Consumption in Industrial Ceramics (CATOPEC-IC)

## References

- I. Hammouda, D. Mihoubi, 2014, Modeling of Thermo-Hydro-Viscoelastic Behavior of a Partially Saturated Ceramic Material During Drying. *Drying Technol.*,32,1219-1230.
- M. Heydari M, K. Khalili,S.Y. Ahmadi-Brooghani, 2018, More comprehensive 3D modeling of clay-like material drying. *AIChE J.*,64,1469-1478.
- E. Mancuhan, S. Özen, P. Sayan, S.T. Sargut, 2016, Experimental investigation of green brick shrinkage behavior with Bigot’s curves. *Drying Technol.*,34,1535-1545.
- S.J. Kowalski, A. Pawłowski, 2010, Modeling of kinetics in stationary and intermittent drying. *Drying Technol.*,28,1023-1031.
- A.G. Barbosa de Lima, J. Barbosa da Silva, G.S. Almeida, J.J.S. Nascimento, F.V.S. Tavares, V.S. Silva, 2016, Clay products convective drying: Foundations, modeling and applications. *Adv Struct Mater.*,63,43-70.

# Investigating physics-informed neural networks for bioprocess hybrid model construction

Alexander William Rogers,<sup>a</sup> Ilya Orson Sandoval Cardenas,<sup>b</sup> Ehecatl Antonio Del Rio-Chanona,<sup>b</sup> Dongda Zhang,<sup>a\*</sup>

<sup>a</sup> *Department of Chemical Engineering, University of Manchester, Oxford Road, Manchester, M1 3AL, UK.*

<sup>b</sup> *Department of Chemical Engineering, Imperial College London, South Kensington Campus, London, SW7 2AZ, UK.*

\* *Corresponding author email: [dongda.zhang@manchester.ac.uk](mailto:dongda.zhang@manchester.ac.uk)*

## Abstract

Integrating physical knowledge and machine learning is a cost-efficient solution to modelling complex biochemical processes when the underlying mechanisms are not fully understood. However, hybrid model structure identification is still time-consuming for new processes, requiring iteration over different hypotheses to explain the observed process dynamics while minimizing over-parameterization. Unfortunately, conventional approaches to automatic model structure identification do not always converge for highly nonlinear models and cannot estimate time-varying model parameters. To address this and accelerate the design of new biochemical processes, a Reinforcement Learning (RL) based framework recently reformulated synchronous hybrid model structure-parameter identification into a process optimal control problem. To further investigate other possible solutions, in this study, a novel Physics Informed Neural Network (PINN) based framework was proposed for the first time to infer time-varying kinetic parameters. This framework first combines possible kinetic structures from phenomenological knowledge, then simultaneously identifies the most likely hybrid model structure and time-varying parameter trajectories. To demonstrate the performance of the PINN based framework, several in-silico case studies were conducted using a known ground truth bioprocess. We thoroughly examined the advantages and limitations of the framework, elucidating its potential for high-fidelity hybrid model construction in biochemical engineering research.

**Keywords:** automatic model structure identification, time-varying parameter estimation, physics-informed neural network, hybrid modelling, machine learning.

## 1. Introduction

Mathematical modelling is pivotal to understanding and designing biochemical processes. A thoroughly validated dynamic model can predict biomass growth and product synthesis under different operating conditions, reducing the number of experiments required to characterize and optimize novel biochemical processes. Kinetic and data-driven models have been proposed to describe biological processes. However, identifying a suitable kinetic model structure is time-consuming, and they are often overparameterized (i.e., multiple structures and many parameters) in an effort to capture the complex dynamics, leading to high parameter and propagated state uncertainties. On the other hand, data-driven models risk overfitting without enough experimental data – which is time-consuming to generate – and poor generalization to new operating conditions.



Integrating physical knowledge and machine learning is a cost-efficient solution to modelling complex biochemical processes when the underlying mechanisms are not fully understood. Hybrid models either use a data-driven model to correct the discrepancy between a kinetic model and the observed process dynamics or update selected time-varying kinetic model parameters. A properly validated hybrid model can effectively resolve the issue of incomplete physical knowledge and low-quality-quantity data (Zhang et al., 2020) while improving prediction accuracy and confidence compared with pure kinetic models (Vega-Ramon et al., 2021). However, hybrid models inherit the risk of over-parameterizing the kinetics and overfitting the data-driven model, which can lead to high uncertainty and poor generalization. Consequently, it is essential to identify the kinetic model structure that best represents the underlying mechanisms to reduce the burden on the data-driven model to compensate and risk overfitting. This is no trivial task since the current bioprocess kinetics depend on the present culture conditions and microenvironmental history due to stochastic effects on the controlling mechanisms or metabolic stores and systematic intracellular metabolic regulation mechanisms. As a result, the kinetic model parameters and structure will evolve with time and history.

Despite this challenge, there are few attempts outside case-specific studies to resolve automatic model structure identification for complex combinatorial, history-dependent or time-varying kinetics. Conventional approaches such as mixed integer nonlinear programming (MINLP) do not always converge for highly nonlinear models and cannot estimate time-varying model parameters. To address this, we recently proposed a novel Reinforcement Learning (RL) based framework to reformulate synchronous hybrid model structure-parameter identification into a process optimal control problem (Wu et al., 2022). The RL-based framework proved promising for recovering the correct kinetic model structure or time-varying kinetic parameters, but the combined case with more combinatorial options and time-varying kinetic parameters remains unexplored. In recent years, Physics-Informed Neural Networks (PINNs) (Raissi et al., 2019) have emerged as a novel approach to discovering underlying governing equations; however, they have not been applied before to infer time-varying physical parameters. Therefore, in this work we aim to propose a novel PINN-based framework for hybrid bioprocess model construction. Several in-silico case studies will thoroughly examine the advantages and limitations of this technique for synchronous hybrid model structure-parameter identification.

## 2. Problem Statement

The PINN-based framework first combines several possible kinetic structures from phenomenological knowledge, then simultaneously identifies the most likely hybrid model structure and time-varying parameter trajectories. To demonstrate the performance of this framework, several in-silico case studies were conducted using a known ground truth model, shown in Equation 1, that we developed in our previous work (Rogers et al., 2022). This high-fidelity hybrid model can predict the temperature-dependent biomass growth, glucose consumption and  $\gamma$ -linolenic acid (GLA) accumulation rates during fermentation of the fungus *Cunninghamella echinulata*. In Equation 1,  $X$ ,  $F$ ,  $S$  and  $P$  are the total biomass, fat-free biomass, glucose and GLA concentration, respectively, while  $T$  is temperature. The half-saturation  $K_S$  and maintenance  $k_S$  coefficients, the total  $k_X$  and fat-free  $k_F$  biomass decay rates and specific GLA decay rate  $k_P$  are all constants, while the total  $Y_X$  and fat-free  $Y_F$  biomass growth rates, the glucose-to-biomass yield coefficient  $Y_S$  and growth-dependent GLA yield coefficient  $Y_P$  are all time-varying parameters. In

our previous work, we estimated the constants and built the Gaussian Process (GP) model mapping the states and temperature to the time-varying parameters (Rogers et al., 2022).

$$\frac{dX}{dt} = Y_X(X, F, S, P, T) \cdot \frac{S}{S + K_S \cdot X} \cdot X - k_X \cdot X \quad (1a)$$

$$\frac{dF}{dt} = Y_F(X, F, S, P, T) \cdot \frac{S}{S + K_S \cdot X} \cdot X - k_F \cdot X \quad (1b)$$

$$\frac{dS}{dt} = Y_S(X, F, S, P, T) \cdot \frac{S}{S + K_S \cdot X} \cdot X - k_S \cdot X \cdot \frac{S}{S + 0.1} \quad (1c)$$

$$\frac{dP}{dt} = Y_P(X, F, S, P, T) \cdot \frac{S}{S + K_S \cdot X} \cdot X - k_P \cdot X \quad (1d)$$

Now the ground-truth model was used to generate six in-silico batches from low ( $S_0 = 60 \text{ g L}^{-1}$ ) and high ( $S_0 = 100 \text{ g L}^{-1}$ ) initial glucose concentrations under three different temperatures ( $T = 14^\circ\text{C}$ ,  $T = 28^\circ\text{C}$ ,  $T = 37^\circ\text{C}$ ). The four GP models updated the four time-varying parameters every 24 hours while the constants were fixed. Each batch was ‘fermented’ for 408 hours, and the four state variables ‘measured’ every 24 hours.

### 3. Methodology

#### 3.1. Physics-Informed Neural Network Structure

The PINN aims to learn the solution to the system of ordinary differential equations (ODEs) presented in Equation 1 by reformulating the parameter estimation and numerical integration problem into a single nonlinear optimization problem. The ANN aims to learn the concentration profile and kinetic parameters that satisfy the measured states and the system of ODEs of the form:  $d\mathbf{x}/dt = \mathcal{N}(\mathbf{x}, \boldsymbol{\phi}, \boldsymbol{\theta})$ , where  $\mathbf{x} = [X, F, S, P]^T$ , while  $\boldsymbol{\phi}$  and  $\boldsymbol{\theta}$  are vectors of time-varying and time-constant parameters, respectively. The unknown solution to the concentration profile  $\mathbf{u}_x(\boldsymbol{\lambda}, t, \mathbf{x}_0)$  and kinetic parameters  $\mathbf{u}_\phi(\boldsymbol{\lambda}, t, \mathbf{x}_0)$  was represented by a single ANN as a function of batch time  $t \in [0, 408 \text{ hours}]$  and the initial state  $\mathbf{x}_0 = [X_0, F_0, S_0, P_0]^T$ , where  $\boldsymbol{\lambda}$  denotes the tunable weights and biases.

$$\boldsymbol{\lambda}^* = \min_{\boldsymbol{\lambda}} \mathcal{L}(\boldsymbol{\lambda}) = \min_{\boldsymbol{\lambda}} [\omega_0 \mathcal{L}_0(\boldsymbol{\lambda}) + \omega_d \mathcal{L}_d(\boldsymbol{\lambda}) + \omega_c \mathcal{L}_c(\boldsymbol{\lambda})] \quad (2a)$$

$$\mathcal{L}_0(\boldsymbol{\lambda}) = \frac{1}{N_0} \sum_{i=1}^{N_0} |\hat{\mathbf{x}}_0^i - \mathbf{x}_0^i|^2 \quad (2b)$$

$$\mathcal{L}_d(\boldsymbol{\lambda}) = \frac{1}{N_d} \sum_{i=1}^{N_d} |\hat{\mathbf{x}}_d^i - \mathbf{x}_d^i|^2 \quad (2c)$$

$$\mathcal{L}_c(\boldsymbol{\lambda}) = \frac{1}{N_c} \sum_{i=1}^{N_c} \left| \frac{\partial \hat{\mathbf{x}}_c^i}{\partial t} - \mathcal{N}(\hat{\mathbf{x}}_c^i, \hat{\boldsymbol{\phi}}_d^i, \hat{\boldsymbol{\theta}}) \right|^2 \quad (2d)$$

The PINN was trained by minimizing the composite loss function in Equation 2, where  $\omega_0$ ,  $\omega_d$  and  $\omega_c$  balance the interplay between the different loss terms during training. Here  $\{\mathbf{x}_0^i\}_{i=1}^{N_0}$ ,  $\{t_d^i, \mathbf{x}_d^i\}_{i=1}^{N_d}$  and  $\{t_c^i, \mathbf{x}_c^i\}_{i=1}^{N_c}$  are the time-state value pairs for the initial conditions,

measured data points and collocation points, respectively. Each batch is divided into  $N$  intervals of  $\Delta t = 24$  hours, where  $t_d \in \{0, \Delta t, 2\Delta t, 3\Delta t \dots N\Delta t\}$ , while each interval is subdivided into  $\Delta t/\delta t = 5$  collocation points, such that  $t_c \in \{0, \delta t, 2\delta t, 3\delta t \dots n\delta t\}$  is on the same timescale. The terms  $\hat{\mathbf{x}}_d = \mathbf{u}_x(\boldsymbol{\lambda}, t_d, \mathbf{x}_0)$  and  $\mathbf{x}_d$  denote the predicted and measured states, respectively, at the known data points, while  $\hat{\mathbf{x}}_c = \mathbf{u}_x(\boldsymbol{\lambda}, t_c, \mathbf{x}_0)$  denotes the predicted state at the collocation points for which the states were not measured. Once  $\hat{\mathbf{x}}_c$  is evaluated, the gradients with respect to time  $\partial \hat{\mathbf{x}}_c / dt$  were computed by reverse automatic differentiation. The time-constant  $\hat{\boldsymbol{\theta}} = \mathbf{u}_\phi(\boldsymbol{\lambda}, 0, 0)$  and time-varying  $\hat{\boldsymbol{\phi}}_d = \mathbf{u}_\phi(\boldsymbol{\lambda}, t_d, \mathbf{x}_0)$  kinetic parameters were both predicted by the ANN, implicitly coupling them to the same network parameters for more stable training. However, unlike the continuous state variables,  $\hat{\boldsymbol{\theta}}$  is a fixed constant, while  $\hat{\boldsymbol{\phi}}$  is a piecewise constant. To represent this behaviour,  $\hat{\boldsymbol{\theta}}$  was evaluated at a fixed input while  $\hat{\boldsymbol{\phi}}_d$  was evaluated at each data sampling point  $t_d$  and the result used over  $t_d \leq t_c < t_d + \Delta t$  when computing  $\mathcal{N}(\hat{\mathbf{x}}_c, \hat{\boldsymbol{\phi}}_d, \hat{\boldsymbol{\theta}})$ . The network was built from a single hidden layer of 40 neurons with a hyperbolic tangent activation function in the hidden layer and a linear output layer. This combination mirrors the typical bioprocesses sigmoid-like growth and parameter profiles.

### 3.2. Physics-Informed Neural Network Training and Simulation

The PINN was fitted simultaneously to all six in-silico batches. The PINN was trained in two stages: first (i) the network was trained for 6000 epochs with  $\{\omega_0 = 5, \omega_d = 1, \omega_c = 0\}$  and a learning rate of  $5 \times 10^{-2}$  until  $\mathcal{L}_d(\boldsymbol{\lambda})$  converged, then (ii) the network was trained for 10000 epochs with  $\{\omega_0 = 5, \omega_d = 1, \omega_c = 5 \times 10^3\}$  and a learning rate of  $5 \times 10^{-4}$  until  $\mathcal{L}_c(\boldsymbol{\lambda})$  converged. Fitting the state profile before the kinetic parameters approximates the initial gradients  $\partial \hat{\mathbf{x}}_c / dt$  and was found empirically to be more robust to becoming trapped within low-quality local optima than single-stage training.

Once the PINN was constructed, the multistep-ahead state trajectory was re-simulated using Equation 3 from  $\mathbf{x}_0 = [X_0, F_0, S_0, P_0]^T$  using the predicted time-constant parameters  $\hat{\boldsymbol{\theta}}^* = \mathbf{u}_\phi(\boldsymbol{\lambda}^*, 0, 0)$  and the predicted time-varying parameters  $\hat{\boldsymbol{\phi}}_d^* = \mathbf{u}_\phi(\boldsymbol{\lambda}^*, t_d, \mathbf{x}_0)$  updated once every 24 hours.

$$\hat{\mathbf{x}}'_{d+1} = \hat{\mathbf{x}}'_d + \int_{t_d}^{t_{d+1}} \mathcal{N}(\mathbf{x}', \hat{\boldsymbol{\phi}}_d^*, \hat{\boldsymbol{\theta}}^*) dt \quad (3)$$

## 4. Results and Discussion

Once the PINN was fitted to the six in-silico batches, the fitted state  $\hat{\mathbf{x}}_d^* = \mathbf{u}_x(\boldsymbol{\lambda}^*, t_d, \mathbf{x}_0)$  trajectory, the time-constant parameters  $\hat{\boldsymbol{\theta}}^* = \mathbf{u}_\phi(\boldsymbol{\lambda}^*, 0, 0)$  and the time-varying kinetic parameter  $\hat{\boldsymbol{\phi}}_d^* = \mathbf{u}_\phi(\boldsymbol{\lambda}^*, t_d, \mathbf{x}_0)$  trajectory was retrieved. In addition, the fitted PINN was used to re-simulate the multistep-ahead state trajectory  $\hat{\mathbf{x}}'_d$  for the six in-silico batches from their different initial states and operating temperatures, as described in Section 3.2. Figure 1 compares the fitted (i.e.,  $\hat{\mathbf{x}}_d^*$ ) and the re-simulated (i.e.,  $\hat{\mathbf{x}}'_d$ ) state trajectories for one of the six in-silico batches, showing that the PINN could fit and re-simulate the state trajectory with very high accuracy, with a mean absolute percentage error (MAPE) of 1% and 5%, respectively. Since  $\mathbf{x}_0$  is satisfied (MAPE of 1%) by strong regularization (i.e.,  $\omega_0 > \omega_d$ ) the mismatch between the fitted and re-simulated trajectories is due to the PINN not being strictly required to satisfy the system of ODEs during PINN training.

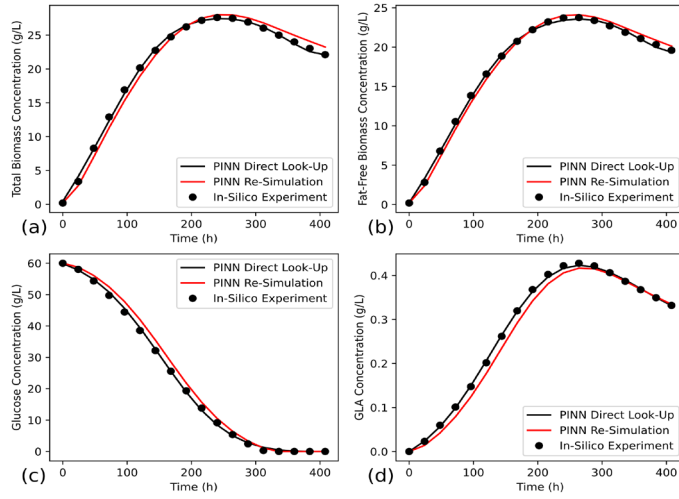


Figure 1: Fitted and re-simulated state trajectory for one of the six in-silico training batches for total biomass (a), fat-free biomass (b), glucose (c) and GLA (d) concentration.

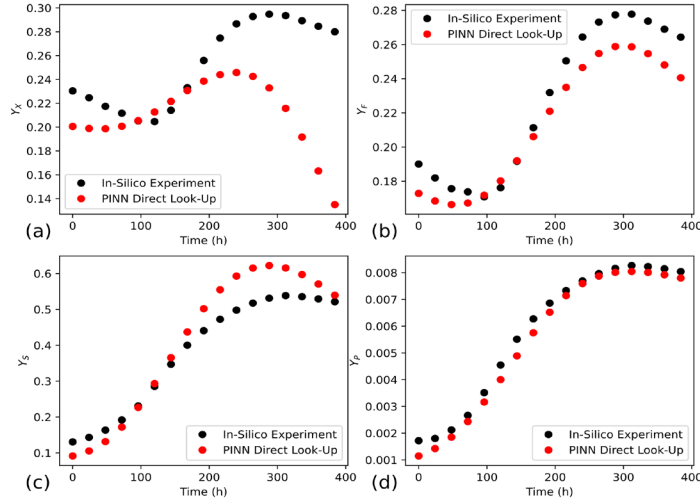


Figure 2: Ground truth and estimated time-varying parameters for one of the six in-silico batches for the specific growth rate  $Y_X$  (a) and  $Y_F$  (b),  $Y_S$  (c) and  $Y_P$  (d) yield coefficients.

Table 1 compares the PINN estimated and ground-truth constants, demonstrating that the PINN can accurately recover (MAPE of 8%) the correct constants when there is no measurement uncertainty. The residual between the ground truth and estimated values can likely be attributed to the practical identifiability of the parameters given only six in-silico batches. Figure 2 compares the PINN estimated and ground-truth time-varying parameter trajectories, demonstrating that the PINN can recover  $Y_F$ ,  $Y_S$  and  $Y_P$  well (MAPE of 15%), particularly considering the time-varying parameters (i.e.,  $Y_X$ ,  $Y_F$ ,  $Y_S$  and  $Y_P$ ) become non-identifiable as  $S \rightarrow 0$  at the end of the batch, as their influence becomes negligible compared to the specific decay constants (i.e.,  $k_X$ ,  $k_F$ ,  $k_S$  and  $k_P$ ). However, the PINN could not recover well  $Y_X$  (MAPE of 150%). Given the excellent fit in Figure 1,

there could be two reasons for this: either (i) the PINN cannot converge to the globally optimum ground-truth parameters, and the model is insensitive to the exact parameter values, or (ii)  $K_S$  is partially non-identifiable, and extra regularization is required.

Table 1: Comparison of ground-truth and estimated time-constant parameters.

Parameter	$K_S$	$k_X$	$k_F$	$k_S$	$k_P$
Ground-Truth	48.64	1.70E-03	1.35E-03	3.10E-03	3.23E-05
Estimated	49.60	1.50E-03	1.32E-03	2.57E-03	2.96E-05
Absolute Error (%)	2.0	11.8	2.2	17.1	8.4

## 5. Conclusion

The PINN-based approach has potential for kinetic parameter estimation and for providing a way to understand real bioprocess underlying system dynamics. At present, embedded hybrid model construction follows a two-step procedure: (i) simultaneous time-constant and time-varying parameter estimation, then (ii) data-driven model construction to correlate the time-varying parameters with the state and operating conditions for predictive simulation. This novel PINN-based framework has the potential to accelerate hybrid model construction by directly building the predictive data-driven model. However, there remains the challenge of identifying a high-quality solution to the PINN network parameters. Therefore, future work will explore more advanced techniques, such as adaptive time-sampling, self-attenuation, time-marching and causal training (Wang et al., 2022). In addition, the PINN-based framework is currently only employed to estimate the kinetic parameters for a pre-defined kinetic model structure. Therefore, future work will also explore synchronous parameter-structure identification and compare the performance of the PINN-based framework against an extended RL-based framework (Wu et al., 2022) when multiple possible structures are possible.

## References

- Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707.
- Rogers, A. W., Song, Z., Vega-Ramon, F., Jing, K., & Zhang, D. (2022). Investigating ‘greyness’ of hybrid model for bioprocess predictive modelling. *Biochemical Engineering Journal*.
- Vega-Ramon, F., Zhu, X., Savage, T. R., Petsagkourakis, P., Jing, K., & Zhang, D. (2021). Kinetic and hybrid modeling for yeast astaxanthin production under uncertainty. *Biotechnology and Bioengineering*, 118(12), 4854–4866.
- Wang, S., Sankaran, S., & Perdikaris, P. (2022). *Respecting causality is all you need for training physics-informed neural networks*.
- Wu, C., Mowbray, M. R., Rogers, A. W., Rio-Chanona, E. A. Del, & Zhang, D. (2022). *A reinforcement learning based hybrid modelling framework for bioprocess kinetics identification*.
- Zhang, D., Savage, T. R., & Cho, B. A. (2020). Combining model structure identification and hybrid modelling for photo-production process predictive simulation and optimisation. *Biotechnology and Bioengineering*, 117(11), 3356–3367.

## **Decision Making Approaches to Improve Resilience in Food Supply Chains and Enhance Food Security Against Climate Change Risks**

Bashar Hassna, Sarah Namany, Mohammad Alherbawi, Adel Elomri, Tareq Al-Ansari\*

*College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar*  
*\*talansari@hbku.edu.qa*

### **Abstract**

As far as food security is concerned, the world is challenged by growing resource scarcity, changing diets, environmental challenges, volatile prices, and supply chain disruptions. Various technological, political, economic, and social factors drive food system dynamics that fluctuate in terms of their impact on food security. Many studies have attempted to alleviate risks governing food systems as means to achieve food security by suggesting numerous modelling approaches and decision-making tools. The purpose of this paper is to design a composite indicator analysing the risk status of a set of trade partners supplying Qatar with a predetermined food basket. The latter involves perishable fruits and vegetables. To design risk indicators, a set of individual factors reflecting the impact of climate change, inclusive of seasonality and sudden risks, are considered. The evaluation of these risks is performed based on their impact on Qatar's demand satisfaction. Whereas the assessment is conducted using the Analytical Hierarchy Process (AHP). The second stage of this work formulates a multi-objective optimization model in MATLAB that determines the optimal network of suppliers to satisfy the local need of Qatar, considering their climate risk profiles and their production capacities. The purpose of the optimization framework is to identify the contribution percentage of each trading partner to the total demand for the predetermined food basket while minimizing economic costs and risk composite factors. Results of the study assert that diversifying trade partners is the most resilient option to satisfy local demands yet with a high economic cost.

**Keywords:** food imports, climate change, risks, AHP, multi-objective optimisation.

### **1. Introduction**

One of the most pressing issues in today's world is climate change, which has significantly impacted the earth's ecosystems. Although the world has always experienced some degree of climate change in the last 100 years, the rate of this variation has multiplied in recent decades. Since the nineteenth century, anthropogenic activities have caused an increase in average temperature of 0.9 °C, primarily as a result of greenhouse gas (GHG) releases into the atmosphere. This rise is anticipated to reach 1.5 °C by 2050 due to continuous deforestation and the contamination of air, soil, and water bodies (Roe et al., 2019).

Long-term variations in temperature and precipitation patterns contribute to changing the comparative advantage of food production around the world and may even slow the expansion of agricultural productivity in areas with high rates of hunger. As a result, global agriculture and food systems may be altered, causing an inter-annual volatility in food supplies. If the current climate change patterns continue to evolve, crop losses may

rise in the very near future at an unprecedented rate; significantly reducing production, driving up food costs, and making it difficult to meet the rising demands of a growing population. Therefore, future food security must be strengthened by considering the inter-annual and inter-decadal effects of climate change.

Considering food systems are susceptible to climate change and natural disasters. The effect of these factors either can be sudden or can be slow, resulting in a positive or negative change (Gomez-Zavaglia, 2020), noting that it mainly affects the production stages, which have been found to only contribute to around 20% of the value chain (Cucagna & Goldsmith, 2018). Whereas the complexity of the global agricultural markets has increased due to more focus on food standards that are related to food safety, food quality, and technological advances, which creates changes in the industry along the value chain. These considerations make the task of creating a robust and resilient food supply system even more challenging.

Therefore, it is imperative to adopt a holistic value chain to enhance resilience rather than considering the production stage alone (Ringsberg, 2014). Consequently, there is a need to increase the efforts related to identifying and classifying risks related to food systems considering data from different parts of the value chain.

Hence, the main aim of this study is to develop a clearer understanding of the food system, influencing factors, and respective risk management tools, with a focus on climate change as an external and unpredictable risk to food systems.

In this study, food system's associated risks are evaluated considering their impact on the demand satisfaction for the State of Qatar. While the Analytical Hierarchy Process (AHP) tool is utilised to evaluate the relative weight of these risk factors based on quantitative and qualitative measures.

A multi-objective optimization model is then developed and solved using MATLAB's Genetic Algorithm. The goal of the optimization framework is to select optimal suppliers and define their contribution percentage towards the food basket of Qatar, while ensuring the local demand is satisfied within the suppliers' capacities, with reduced costs and minimal associated risks. The model in hand may provide insights on possible enhancement to global food systems considering economic, social, and environmental aspects.

## **2. Methodology**

### **2.1. The Analytical Hierarchy Process to design a climate risk indicator**

The Analytical Hierarchy Process (AHP) was used to develop a composite risk index which involves five different climate change risk related indicators counting, surface temperature change, water stress, arable land, sea level rise and natural disasters management in addition to agriculture policy adoption. Four countries were selected which are USA, India, China, and Iran. The AHP is used to quantify the level of risk of each trade partner through determining its priority rank and following the decision tree described in figure 1. Priorities were selected based on the situation of the country in each of the indicators and was based on historical data and grey literature (Dahlman, (n.d); EEA(2021) Level of Water Stress (2023); The Economist Group (2018)).

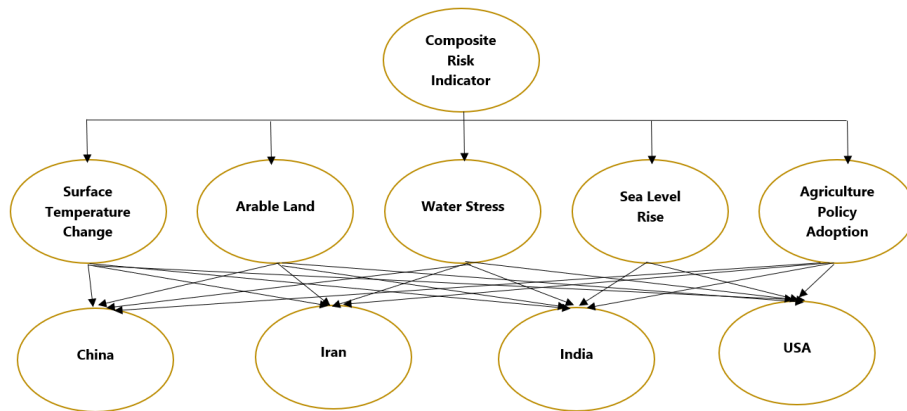


Figure 1. The hierarchical decision tree

## 2.2. Optimization model

The purpose of this part of the methodology is to design a multi-objective optimisation model that determines the optimal network of suppliers that can satisfy the local need of Qatar, considering their climate risk profiles and their production capacities. The purpose of the optimization framework is to identify the per cent contribution of each trading partner to the total demand for the predetermined food basket while minimizing economic costs, and risk composite indicators. The following table 1 presents the mathematical formulation of the suggested optimisation model. Considering the variable used,  $x_i$  presents the decision variable of the model which is the percent contribution of each country to the food basket.  $Q$  is the yearly local demand for food in kg. As for  $c_i$ , it represents the food basket average overall unit cost in \$/kg.  $r_i$  is the overall risk indicator generated from the AHP model. As for the  $cap_i$ , it represents the allowable exportable quantities in kg that can be supplied by each trade partner.

Table 1. The mathematical formulation of the optimisation model

	Equation	Connotation
<b>Objective function 1</b>	$\min \sum_{i=1}^5 Qc_i x_i$	Identifies the optimal set of trade partners that supply Qatar’s food basket while minimising the overall cost.
<b>Objective Function 2</b>	$\min \sum_{i=1}^5 r_i x_i$	Identifies the optimal set of trade partners that supply Qatar’s food basket while minimising the overall composite risk indicators.
<b>Constraint 1</b>	$\sum_{i=1}^5 x_i = 100\%$	Logical constraint implies that the sum of all contributions per cent of water technologies must be 100%.
<b>Constraint 2</b>	$Qx_i < cap_i$	The imported quantity from each trading partner should not exceed their individual exportable capacity.



### 3. Results and discussion

After conducting the AHP, 4 composite risks factors were generated for each country describing its overall risk performance (table 4). Those risks factors were derived from the composite indicators involving the set of climate risks sub-indicators as highlighted in table 3.

Composite risk factors were later used in the optimisation framework as parameters to select the optimal mix of exporting countries that can supply the market with food products.

**Table 3.** Results of the AHP priority score for each exporting country

	China	USA	India	Iran
<b>Disaster Risk Management</b>	0.0611	0.0614	0.0656	0.0738
<b>Surface Temperature Change</b>	0.14065	0.1484	0.1322	0.1264
<b>Arable land</b>	0.1612	0.1738	0.3004	0.2546
<b>Water Stress</b>	0.1388	0.1921	0.2432	0.2726
<b>Sea level rise</b>	0.4522	0.3755	0.2116	0.2255
<b>Agricultural Policy Adoption</b>	0.0459	0.0482	0.0467	0.0466

**Table 4.** Composite risk factors for each exporting country ((Namany et al., 2019)

	China	USA	India	Iran
<b>Unit cost (\$/kg)</b>	1.22	7.08	1.45	0.81
<b>Composite Risk Factor</b>	0.2222	0.2155	0.2114	0.2088

Results of the optimisation model are summarised in Figure 2 and 3. The optimal solution suggests a slightly diversified network of exporting countries with a large dominance of Iran with around 90%, while the remaining 10% is distributed between China and India. As for USA, its contribution is below the 1%. This can be explained by the large unit costs of products originating from the USA and the relatively low risk factors and costs of the other trade partners. Taking the graphical average from the Pareto front, the average optimal solution, generates an average total cost of \$672M.

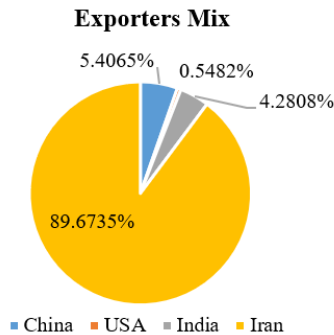


Figure 2. The average optimal trade partners' network

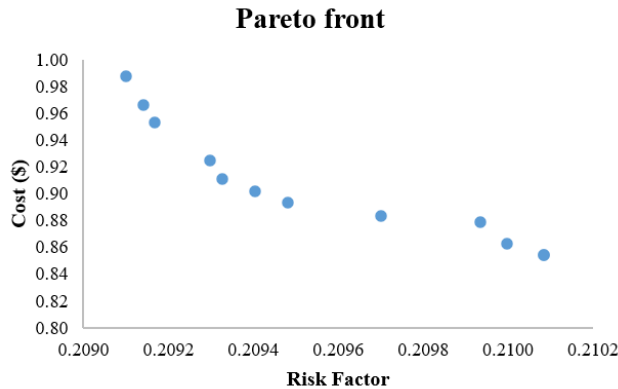


Figure 3. The Pareto front.

#### 4. Conclusion

In this paper a multi-objective optimization model was developed to identify the ideal network of suppliers that can meet Qatar's local demand while taking into account their production capacities and risk profiles related to the climate. The optimization framework's goal was to calculate each trade partner's percentage contribution to the overall demand for the predefined food basket while minimizing composite risk factors and economic costs four different countries. According to study findings, diversifying trade partners is the most durable way to meet local wants, however it can come at a high financial cost.

#### 5. Acknowledgments

This research was made possible by an Award (GSRA7-1-0521-20080) and supported by proposal number NPRP11S-0107-180216 from Qatar National Research Fund (a member of Qatar Foundation). The contents herein are solely the responsibility of the authors[s].

## 6. References

- Cucagna, M. E., & Goldsmith, P. D. (2018). Value adding in the agri-food value chain. *International Food and Agribusiness Management Review*, 21(3), 293-316.
- Dahlman, R. L. A. N. D. L. A. (n.d.). Climate change: Global temperature. NOAA Climate.gov. Retrieved February 5, 2023, from <https://www.climate.gov/news-features/understanding-climate/climate-change-global-temperature>
- Global and European sea-level rise. European Environment Agency. (2021, May 11). Retrieved February 5, 2023, from <https://www.eea.europa.eu/data-and-maps/indicators/sea-level-rise-6/assessment>
- Gomez-Zavaglia, A., Mejuto, J., & Simal-Gandara, J. (2020). Mitigation of emerging implications of climate change on food production systems. *Food Research International*, 134, 109256. <https://doi.org/10.1016/j.foodres.2020.109256>
- Namany, S., Al-Ansari, T., & Govindan, R. (2019). Optimisation of the energy, water, and food nexus for food security scenarios. *Computers & Chemical Engineering*, 129, 106513. <https://doi.org/10.1016/J.COMPCEMENG.2019.106513>
- Level of Water Stress: Freshwater withdrawal as a proportion of available freshwater resources - sustainable development goals - united nations economic commission for Europe. (n.d.). Retrieved February 5, 2023, from <https://w3.unece.org/SDG/en/Indicator?id=140>
- Ringsberg, H. (2014). Perspectives on food traceability: a systematic literature review. *Supply Chain Management: An International Journal*.
- Roe, S., Streck, C., Obersteiner, M., Frank, S., Griscom, B., Drouet, L., ... & Lawrence, D. (2019). Contribution of the land sector to a 1.5 C world. *Nature Climate Change*, 9(11), 817-828.
- The Economist Group. (2018). *Global Food Security Index 2018: Build resilience in the face of rising food-security risks*. The Economist Group.

# Lot-sizing and Production Scheduling of a Beverage Industry

Maria E. Samouilidou<sup>a</sup>, Eleonora Diakoumi<sup>a</sup>, Georgios P. Georgiadis<sup>a</sup>, Antonios Dikaiakos<sup>b</sup>, Michael C. Georgiadis<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece*

<sup>b</sup>*Green Cola Hellas, National Road Athens – Lamia & Kalamatas 2, New Kifisia Athens, 14564, Greece*  
[mgeorg@auth.gr](mailto:mgeorg@auth.gr)

## Abstract

In this work, the lot-sizing and production scheduling problem of a real-life beverage industry is addressed. In particular, a Big bucket-Small bucket Mixed-Integer Linear Programming (MILP) framework is proposed for the optimal scheduling of a Greek soft-drink production facility. The production process consists of two main stages, a batch syrup preparation stage and a continuous bottling stage. The synchronization between the syrup tanks and filling lines is of key importance for the feasibility and efficiency of the production schedules. The main goal is the generation of an optimal weekly production schedule which satisfies a given demand. A rigorous scheduling formulation for the lot-sizing and production scheduling of an industrial facility including both batch and continuous processes is presented. The derived optimal schedules lead to an increased plant productivity while reducing the utilization and cost of labor resources.

**Keywords:** production scheduling, optimization, beverage industries, MILP

## 1. Introduction

Nowadays, process industries operate in a challenging economic environment with low profit margins, where the demands they have to meet significantly increase and competition is rapidly intensifying. The ever-increasing number and diversity of products, as well as the complexity of production processes, especially in the food and beverage industry, render production scheduling an essential procedure for acquiring efficient production plans and ensuring the plant's viability and profitability. The main challenge regarding these industries is the integrated modelling of all the production stages (Toledo et al., 2015). In the beverage industry, a few scheduling techniques have been proposed in the literature, with even fewer addressing complex real-life industrial applications (Georgiadis et al., 2021). Baldo et al. (2014) proposed an MILP model, combined with MIP-based heuristics for the production lot-sizing and scheduling problem of a brewery industry with long lead times and unfixed bottleneck. Mediouni et al. (2022) studied the lot-sizing and scheduling problem in a dairy soft-drink production process, considering the limited shelf life of intermediate products and utilizing a relax-and-fix heuristic. Ferreira et al. (2009) presented an MILP model that integrates production lot-sizing and scheduling decisions of beverage plants with sequence-dependent setup costs and times and applied it to a real life instance generating improved results. The efficient lot-sizing and scheduling problem of a real-life soft-drink production facility is examined in this

work, including the integrated modelling of all production stages. The proposed mathematical framework tackles the existing production restrictions and provides optimal results regarding the productivity and the production costs of the plant.

## 2. Problem Statement

The plant includes both batch and continuous processes. More than 60 final products are produced over a 5-day time horizon. A daily 8-hour shift is used, while overnight and weekend shifts can take place if needed. It is noted that more than one product can be produced from the same syrup. A brief layout of the production process is shown in Figure 1.

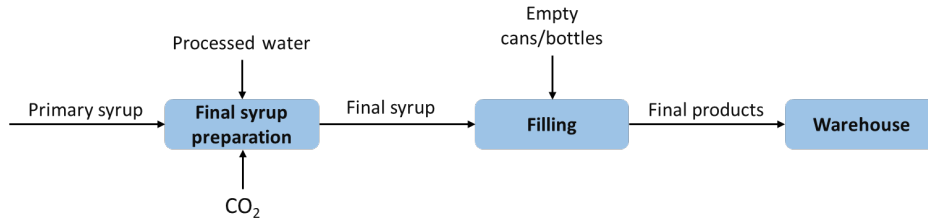


Figure 1. Production process layout

According to the production recipe, initially, a primary flavored syrup is prepared. There are two tanks available for the preparation of syrups. At the same time, water is treated and deaerated so as to be suitable for mixing with the primary syrup. Inline mixing takes place between the primary syrup, the treated water and  $\text{CO}_2$  and the final mixture reaches the filling lines. In the meanwhile, empty cans/bottles are also transferred to the appropriate packing lines, where they are rinsed. There is a total of three packing lines in the facility, one for cans, one for PET bottles and one for glass bottles, but according to the company's policy only two of them can operate per day. Then, the cans/bottles are filled with the final mixture and before they finally form a multipack, labels and other messages are printed on them. Additionally, a brief quality check happens at this point. Finally, the final products are transferred to the warehouse, where they are stored until they are later distributed according to customer demand. It can be considered that the production process consists of two main stages, a batch syrup preparation stage and a continuous packing stage. The production of soft-drinks requires the synchronization between these two stages, in order to secure feasible schedules and to minimize overtime costs. Furthermore, time consuming changeovers take place in both stages depending on factors such as flavor, packing size and label. Taking into account all of these limitations while satisfying a given demand is quite challenging in real life industrial environments. Therefore, the production scheduling of a facility presenting the aforementioned characteristics is examined in this work.

## 3. Mathematical framework

This model constitutes an extension of the mathematical frameworks previously developed by Ferreira et al. (2009) and Mediouni et al. (2022). It is a Big bucket-Small bucket Mixed-Integer Linear Programming (MILP) framework extended in such a way that up to three products can be produced in the same Small Bucket (micro-period) from a single syrup and additionally the option of overtime production is considered. The planning horizon is divided in  $T$  macro-periods with the duration of a day and each macro-period is further divided into  $S$  micro-periods with flexible lengths. In each micro-period

only one syrup can be prepared in each tank. A brief description of the mathematical model follows below.

**1<sup>st</sup> stage constraints**

$$\sum_{l \in LU_{l,u}} Yl_{u,l,s} \leq 1 \quad \forall u, s \quad (1)$$

$$\sum_m Yum_{u,m,l,s} = Yl_{u,l,s} \quad \forall u, s, l \in LU_{l,u} \quad (2)$$

$$\sum_{l, u \in UL_{u,l}} Yum_{u,m,l,s} \leq 1 \quad \forall m, s \quad (3)$$

$$Zl_{u,k,l,s} \geq Yl_{u,k,s-1} + Yl_{u,l,s} - 1 \quad \forall u, t, s \in St_{s,t}, k \in LU_{k,u}, l \in LU_{l,u}, k \neq l \quad (4)$$

$$\sum_{u \in UL_{u,l}} q_{l,u} \cdot Yum_{u,m,l,s} \leq \sum_{j \in JML_{j,l} \cap JM_{j,m}} r_{j,l} \cdot x_{m,j,s} \leq \sum_{u \in UL_{u,l}} Ku_u \cdot Yum_{u,m,l,s} \quad \forall m, s, l \quad (5)$$

**2<sup>nd</sup> stage constraints**

$$In_{j,t-1} + \sum_{m \in JM_{j,m}, s \in St_{s,t}} x_{m,j,s} = In_{j,t} + d_{j,t} \quad \forall j, t \quad (6)$$

$$\sum_{s \in St_{s,t}} TSm_{m,s} \leq w_{m,t} + Ov_t \quad \forall m, t \quad (7)$$

$$TSm_s \geq \sum_{j \in JM_{j,m}} a_{m,j} \cdot x_{m,j,s} + \sum_{i \in JM_{i,m}, j \in JM_{j,m}, i \neq j} bb_{i,j} \cdot Zij_{m,i,j,s} + \sum_{u,k \in LU_{k,u}, l \in LU_{l,u}, k \neq l} b_{k,l} \cdot Zum_{u,m,k,l,s} \quad (8)$$

$$+ tr \cdot \sum_{u,l \in UL_{u,l}} Yum_{u,m,l,s} \quad \forall m, s$$

$$TSm_s \geq \sum_{j \in JM_{j,m}} a_{m,j} \cdot x_{m,j,s} + \sum_{i \in JM_{i,m}, j \in JM_{j,m}, i \neq j} bb_{i,j} \cdot Zij_{m,i,j,s} + \sum_{i \in JM_{i,m}, j \in JM_{j,m}, i \neq j} bb_{i,j} \cdot Zjm_{i,j,s} \quad \forall m, s \quad (9)$$

$$\sum_{i \in JM_{i,m}, i \neq j} Zij_{m,i,j,s} + \sum_{ii \in JM_{ii,m}, ii \neq j} Zij_{m,j,ii,s} + 1 \geq Yjm_{j,s} + Yjm_{j,j,s}, \quad (10)$$

$\forall m, j \in JM_{j,m}, jj \in JM_{jj,m}, j \neq jj, s$

$$\sum_{i \in JM_{i,m}, j \in JM_{j,m}} Zij_{m,i,j,s} + V_{m,s} = \sum_{j \in JM_{j,m}} Yjm_{j,s} \quad \forall m, s \quad (11)$$

$$G \leq Zjm_{i,j,s} \leq 0.5 + \frac{G}{2} \quad \forall m, i \in JM_{i,m}, j \in JM_{j,m}, i \neq j, t, s, \text{ where} \quad (12)$$

$$G = Yjm_{i,s-1} + Yjm_{j,s} - 1 - \sum_{jj \in JM_{jj,m}, i \neq jj} Zij_{m,i,jj,s-1} - \sum_{ii \in JM_{ii,m}, j \neq ii} Zij_{m,ii,j,s} \leq Zjm_{i,j,s}$$

**Objective function**

$$\min \sum_{m,t,s \in St_{s,t}} TSm_{m,s} \quad (13)$$

Constraint set (1) imposes that in each micro-period  $s$ , each tank  $u$  can prepare at most one syrup  $l$ , using the binary allocation variable  $Yl_{u,l,s}$ . Constraints (2) and (3) express that if a syrup  $l$  is prepared in tank  $u$  in micro-period  $s$ , then it must be transferred to a single packing line  $m$  and this packing line cannot receive another syrup from another tank. The binary variable  $Yum_{u,m,l,s}$  indicates that the tank  $u$  prepares syrup  $l$  in micro-period  $s$  for the line  $m$ . Constraints (4) define the changeover of syrups in a tank using the immediate precedence binary variable  $Zl_{u,k,l,s}$ . With constraints (5) it is ensured that the amount of final products produced from the same syrup in a micro-period  $s$  (where  $r_{j,l}$  is the required amount of syrup  $l$  to produce one unit of product  $j$  and  $x_{m,j,s}$  represents the produced units of product  $j$  on packing line  $m$  in micro-period  $s$ ), is bounded by the minimum quantity of filling  $q_{l,u}$  and the maximum capacity  $Ku_u$  of the tank. Constraints (6) demonstrate that the inventory of product  $j$  in macro-period  $t$ ,  $In_{j,t}$ , equals the inventory of the previous macro-period plus the produced amount in the current macro-period, minus the customer demand  $d_{j,t}$ . Constraint set (7) guarantees that the total production time in macro-period  $t$ ,  $\sum_{s \in St,s,t} Ts_{m,s}$ , does not exceed the daily production time horizon  $w_{m,t}$  including any possible overtime shifts  $Ov_t$ . Constraints (8) and (9) establish that the functioning time of line  $m$  in micro-period  $s$ ,  $Ts_{m,s}$ , must be equal to or greater than the packing time (with  $a_{m,j}$  being the required production time of one unit of product  $j$  on line  $m$ ), plus the changeover time between products from the same syrup  $bb_{i,j}$ , plus the line's waiting time. The line's waiting time in a micro-period is either equal to the time needed for the tank's changeover  $b_{k,l}$  and preparation of syrup  $tr$  or to the line's changeover  $bb_{i,j}$  regarding the previous micro-period, depending on which one lasts longer. Constraints (10) and (11) monitor the changeover between products in the same micro-period  $Zij_{m,i,j,s}$  while constraints (12) define the changeover between products in consecutive micro-periods  $Zj_{m,i,j,s}$ . The binary variable  $Yj_{m,j,s}$  specifies whether product  $j$  is being packed in line  $m$  in micro-period  $s$  and the binary variable  $V_{m,s}$  is activated if line  $m$  functions at micro-period  $s$ . The objective of the model, described by constraint (13), is the minimization of the total production time.

#### 4. Results & Discussion

An industrial case study concerning the production plant of Green Cola Hellas is demonstrated. The MILP model was implemented in GAMS and solved using CPLEX 12.0. Optimality was reached in acceptable computational times by the company in all cases, as shown in Table 1. Figures 2 and 3 illustrate the generated schedules for two given weekly demands. According to the company's policy two of the three available packing lines operate per week. Figure 4 depicts the plant's weekly schedule if all three packing lines were to operate in a week. Moreover, Figure 5 shows the percentage of functioning time each machine devotes to different kinds of operations in this case. It is noted that in the following Gantt charts only the syrup preparation and changeover time are illustrated with colors for the tanks, whereas the time dedicated to feeding the packing lines is not. The different syrup flavors prepared are annotated as RX while the final products are annotated as PX.

Table 1. Solution report for the satisfaction of three weekly product demands

	CPU (s)	GAP (%)	Overtime (hrs)	Weekly functioning time (hrs)				
				Tank 1	Tank 2	Can line	PET line	Glass line
Demand 1	28	0	16.3	40.3	38.1	30.6	-	23.8
Demand 2	900	0.4	12.6	36.1	47.0	25.9	21.3	-
Demand 3	900	0.3	16.3	46.4	38.0	31.7	12.8	23.8

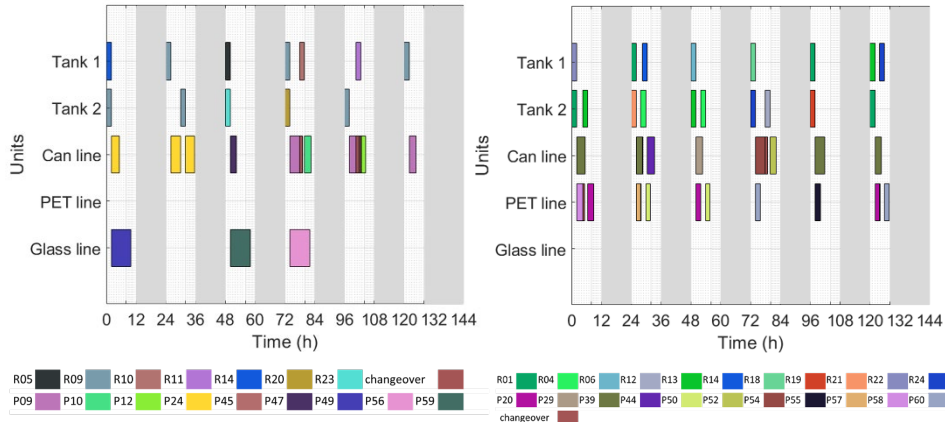


Figure 2. Weekly schedule for Demand 1

Figure 3. Weekly schedule for Demand 2

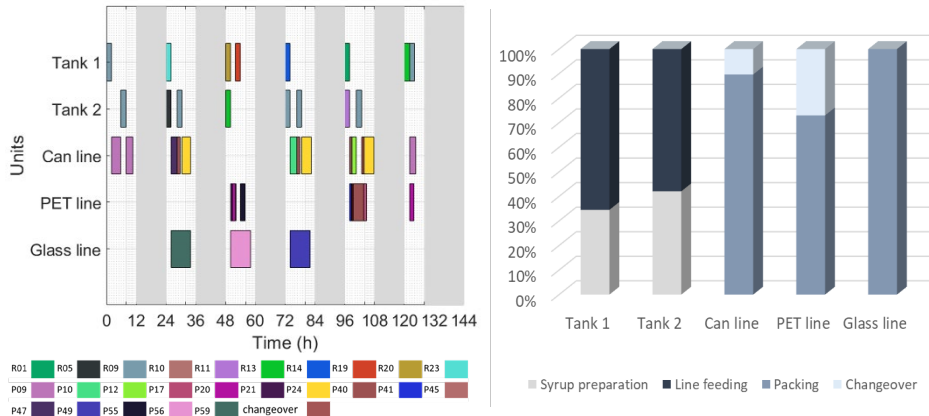


Figure 4. Weekly operation of all three lines

Figure 5. Functioning time allocation for each machine (Demand 3)

In all cases, overtime is required to satisfy the given demand. In Figures 2 and 4, it is observed that multiple products are produced from a single syrup, a flexibility given by the extended mathematical model presented in this work. Furthermore, the correct synchronization of the various units is achieved. The syrup tanks are identified as the production’s bottleneck since they are utilized during the whole production horizon, either preparing syrups or feeding the packing lines. Syrup preparation requires only a little time, nevertheless a tank cannot start preparing a new syrup unless it is empty. In



fact, significantly more time is dedicated to line feeding compared to the time dedicated to syrup preparation, as shown in Figure 5. The operation of all three available packing lines in a week gives flexibility to the production while larger product demands can be satisfied. However, this leads to an increased inventory.

## 5. Conclusions

The lot-sizing and production scheduling problem of a real-life beverage industry is presented in this work. The production facility under consideration can be identified as multiproduct and multistage, with each stage including both batch and continuous processes, making the problem particularly complex. Furthermore, the synchronization between these stages is important for the feasibility and efficiency of the production schedules, an issue often met in food and beverage industries. For the optimal scheduling of the facility, an MILP framework is proposed. To the best of our knowledge this study is the first to allow multiple products to be produced in the same Small Bucket from a single syrup, thus significantly improving the facility's productivity. This work's novelty is further extended by the inclusion of overtime production if necessary, for the timely satisfaction of the given weekly demand. Optimal production schedules are generated in short computational times leading to an increased plant productivity and reduced cost of labor resources. Moreover, the syrup tanks are identified as the production's bottleneck. Future work focuses on the short-term storage of the intermediate syrups in the preparation tanks while taking into account their shelf life and also on the introduction of buffer tanks between the preparation stage and the packing lines so that the syrup tanks can function at their full capacity.

## Acknowledgements

This research is co-financed by the European Union and Greek National Funds through the Region of Central Macedonia, under the operational program "Region of Central Macedonia 2014-2020" and the specific action/call "Investment Plans on Innovation". (Project code: KMP6-0077560). Project title: Development of a software tool for the optimization of production scheduling in the manufacturing industries.

## References

- D. Ferreira, R. Morabito & S. Rangel, 2009, Solution approaches for the soft drink integrated production lot sizing and scheduling problem, *European Journal of Operational Research*, 196(2), 697-706
- G. P. Georgiadis, A. P. Elekidis & M. C. Georgiadis, 2021, Optimal production planning and scheduling in breweries, *Food and Bioproducts Processing*, 125, 204-221
- A. Mediouni, N. Zufferey, M. Rached & N. Cheikhrouhou, 2022, The multi-period multi-level capacitated lot-sizing and scheduling problem in the dairy soft-drink industry, *Supply Chain Forum: An International Journal*, 23(3), 272 - 284
- C. F. Toledo, A. Kimms, P. M. França & R. Morabito, 2015, The synchronized and integrated two-level lot sizing and scheduling problem: evaluating the generalized mathematical model, *Mathematical Problems in Engineering*, 2015
- T. A. Baldo, M. O. Santos, B. Almada-Lobo & R. Morabito, 2014, An optimization approach for the lot sizing and scheduling problem in the brewery industry, *Computers & Industrial Engineering*, 72, 58-71

# Modeling and Analysis of Clinical & Municipal Waste Incineration Process using Aspen Plus

Mohamad Rizza Othman,<sup>a</sup> Sivanesh Kumar Anpalagam,<sup>a</sup> Nur Fitriyanni Jafary,<sup>a</sup>

<sup>a</sup>*Dep. of Chem. Eng. Tech., Fac. of Chem. & Process Eng. Tech., Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Gambang, Pahang, Malaysia*

## Abstract

Modeling and analyzing the solid waste incineration process is a salient study for future advanced waste disposal technology and optimizing process efficiency. However, the development of a mechanistic model that describes the process deal with several challenges and has yet to be extensively explored. Thus, this paper aims to develop a mechanistic model for clinical and municipal solid waste incineration in Aspen Plus. The waste was modeled using customized non-conventional (NC) solid properties, and a general coal enthalpy model with heat correlations to calculate the heating value. The model results accurately depict the waste heating values found in the literature. Further analysis reveals the ability of the model to describe the effect of waste feed quality on incinerator performance.

**Keywords:** Clinical and municipal waste, incineration, Aspen Plus.

## 1. Introduction

The generation of municipal solid waste (MSW) has been increasing globally owing, primarily due to an increase in global population, urbanization, and economic growth, as well as changes in production and consumption behavior (Lohri et al., 2017). On the other hand, the amount of clinical waste (CW) has increased over the past two years due to a high number of COVID-19 infections. Studies found that the weight of CW has increased by 27% due to higher cases of COVID-19 being recorded. In 2013 about 50 metric tons of CW was generated; however, for now, it is estimated that about 90 metric tons of CW are generated each day, and about 25 metric tons of it is due to COVID-19 (Agamuthu & Barasarathi, 2021). Various solid waste management treatments exist, including conventional sanitary landfill treatment, biological treatment, and thermal treatment. Predominantly, landfill is the primary method of waste disposal in most countries around the world. However, such a common method is ineffective in managing the increasing volume of solid waste apart from its drawback towards land, air, ground, and surface water pollution. Hence, this combustible waste may be incinerated or combusted as an alternative to landfill.

The advantages of waste incineration over landfilling are the reduction in waste volume, waste stabilization, waste-to-energy (WtE) conversion, and sanitization of waste, especially clinical waste. Besides, it also runs stably for a long period. In the incinerator, waste is burned at 850 to 1000°C, where CO<sub>2</sub>, SO<sub>2</sub>, NO, NO<sub>2</sub>, HCl, HF, CO, dioxins, and furans are released during incineration. The government law has mandated that waste incineration facilities use gas purification technologies to reduce volume and hazardous characteristics and capture or destroy potentially harmful substances that may be released

during incineration. As a result, atmospheric gas emission limits are typically set to minimize health and environmental impacts. According to Williams (2005), the combustion properties and emissions produced by the combustion system will be influenced by the waste's composition and characteristics. The solid waste's fuel properties and the proximate and ultimate analyses play an important role in how the waste will burn in the incinerator and the emissions that are likely to result. The composition and heating value were the most important parameters governing thermal processing (Zhou et al., 2014). Heating values measure the chemical energy bound in a feedstock where this energy is released during combustion (Erol et al., 2010) and categorized into two, namely higher heating value (HHV) and lower heating value (LHV). The former involves the heat of condensation of water vapor produced in the heat of combustion, while the latter is related to the HHV via the heat of vaporization of water. In some instances, the HHV and LHV are associated with the heat of vaporization of both HCl and water if the waste contains chlorine and produces HCl as one of the products.

There are two methods for experimentally determining the heating values: a full-scale boiler as a calorimeter and a laboratory bomb calorimeter. Heating values can also be determined by calculation using empirical models, and several models have been developed by researchers (Ogwueleka, 2010). To this day, the most popular method for predicting heating value is linear regression. Using regression analysis to develop an empirical model to predict the heat value is thought to be easier and more cost-effective (Khuriati et al., 2017). A model that describes the oxidation, pyrolysis, and incineration process is useful to provide insights into the incineration performance and operations. However, the development of such a model deal with several challenges and has yet to be extensively explored. Therefore, the focus of this study was to develop a working model by employing Aspen Plus simulation software. The establishment of such a model is useful to describe the solid waste incineration process, determine the factors impacting its performance, and predict incineration performance under various operating conditions.

## 2. Methodology

### 2.1. Calculation of heating value

The solid waste was modeled in Aspen Plus as a non-conventional (NC) solid. The NC component models for enthalpy and density properties calculations were empirical correlations that require solid material characterization information derived from constituent ultimate and proximate analyses. The estimates were based on the general coal enthalpy model, including several correlations to compute the heat of combustion, heat of formation, and heat capacity. In this work, the heat of combustion in dry, mineral-matter-free basis,  $\Delta_c h_i^{dm}$ , was calculated using the Boie correlation according to the following equation,

$$\Delta_c h_i^{dm} = (a_{1i}w_{C,i}^{dm} + a_{2i}w_{H,i}^{dm} + a_{3i}w_{St,i}^{dm} + a_{4i}w_{O,i}^{dm} + a_{5i}w_{N,i}^{dm})10^2 + a_{6i} \quad (1)$$

Where  $w$  is the weight fraction and the constant terms,  $a_{1i}$  to  $a_{6i}$  were obtained from the Institute of Gas (IGT) study. Products were in the form of ash, liquid water, and gaseous CO<sub>2</sub>, SO<sub>2</sub>, and NO<sub>2</sub>. The heat of formation,  $\Delta_f h_i^d$  was based on the standard heat of combustion-based correlation. The numerical coefficients were combinations of stoichiometric coefficients and heat of formation for CO<sub>2</sub>, H<sub>2</sub>O, HCl, and NO<sub>2</sub> at 298.15K represented by,

$$\Delta_f h_i^d = \Delta_c h_i^d - (1.418 \times 10^6 w_{H,i}^d + 3.278 \times 10^5 w_{C,i}^d + 9.264 \times 10^4 w_{S,i}^d - 2.418 \times 10^4 w_{N,i}^d - 1.426 \times 10^4 w_{Cl,i}^d) 10^2 \quad (2)$$

Whereas the heat capacity calculations were based on Kirov Correlations (Kirov, 1965), which treats the heat capacity as a weighted sum of the heat capacities of the constituents, which include moisture, fixed carbon, volatile matter, and ash.

### 2.2. Aspen Plus flowsheet for CW incineration

Modeling and simulation for solid waste incineration were performed in Aspen Plus V12.1. The solid waste was fed constantly, and the combustion chamber's process was adiabatic. Gaseous equilibrium is achieved between the primary and secondary chambers. Heat transfer occurs within the chamber via the mixing of hot gases as well as radiation. The flowsheet comprised two reactor blocks, RYIELD and RSTOIC. The former was used to convert non-conventional components into conventional components and the heat of formation required, while the latter was used to simulate the combustion process. In addition, two calculator blocks were added to convert the ultimate analysis element into a wet basis component for the RYIELD reactor block, while the second calculator block was used to calculate the heating value. The IDEAL base method was used as the property method. The thermophysical properties (enthalpy, density, and specific heat) of solid waste were calculated using the built-in coal property models HCOALGEN and DCOALIGT based on their proximate and ultimate analysis. For calculating heating values, all unit operations use 25°C and 1 bar as the reference state. Excess air with 78% nitrogen and 21% oxygen was used as the oxidizing agent.

### 2.3. Case study

Table 1 shows the proximate and ultimate analysis and its heating values of several municipal and clinical samples found in the literature. Each sample's proximate and ultimate values were inserted into the model to calculate the heating values. The calculated results were then compared. For all cases, the waste fed rate was assumed to be 500 kg/h. To ensure excess air between a stoichiometric ratio of 1.5 to 2.0, the combustion air fed rate used was 7000 kg/h.

Table 1: Proximate and ultimate analysis of municipal and clinical waste samples

	CW (Erdogan & Yilmazoglu, 2021)			MSW (Gebreslassie et al., 2020)		
	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
<u>Proximate</u>						
Moisture (%)	0.32	0.29	7.04	0.24	5.76	2.99
Ash (%)	0	2.3	1.89	0.12	6.23	13.73
Volatiles (%)	99.13	78.52	82.37	94.44	77.81	75.23
Carbon (%)	0.55	18.89	8.7	5.2	10.2	8.05
<u>Ultimate</u>						
C (%)	81.81	65.92	48.98	63.91	39.32	36.51
H (%)	12.17	10.03	7.2	5.331	7.2	5.31
O (%)	5.76	23.09	43.52	30.714	53.48	58.18
N (%)	0.15	0.74	0.22	0	0	0
S (%)	0.11	0.22	0.08	0.045	0	0
HV (kJ/kg)	42650 <sup>a</sup>	29120 <sup>a</sup>	15570 <sup>a</sup>	23193 <sup>b</sup>	16217 <sup>b</sup>	12174 <sup>b</sup>

<sup>a</sup> LHV, <sup>b</sup> HHV

### 2.4. Statistical analysis

To determine the accuracy of the proposed model, mean absolute percentage error (MAPE) was used to denote the closeness of the calculated heating value to the actual

value. MAPE is mainly used to determine model accuracy and is calculated using the following equation,

$$MAPE = \left( \frac{1}{n} \sum_{i=1}^n \frac{|E_i - P_i|}{E_i} \right) \times 100\% \quad (3)$$

where  $n$  is the sample number,  $E_i$  is the actual value, and  $P_i$  is the calculated value. Based on Lewis (1982), a MAPE value of less than 10% gives a highly accurate prediction.

### 2.5. Model analysis

Analysis of the model was performed to illustrate the effect of process parameters on the incineration performance of CW and MSW. Two parameters were studied for each type: moisture and VM content. Moisture content, in CW, for example, could consist of up to 38% (Li & Jeng, 1993), and for some biomedical compositions, VM values were from 30% (Olanrewaju, 2019). These two values were varied, and the corresponding normalized proximate and ultimate analyses were then used in the model to determine the heating value and combustion temperature. Furthermore, some waste, i.e., CW, must be incinerated to around 1000 C due to regulatory requirements. Therefore, additional heat duty was calculated to determine the combustion heat duty required to achieve the desired decomposition temperature. Typically, the extra heat duty was supplied by auxiliary fuel i.e., diesel and natural gas.

## 3. Result & Discussion

### 3.1. Model validation

In Table 2, the model calculated Sample 1 to have the highest heating values, whereas Sample 6 had the lowest. Moreover, each sample's descending pattern of calculated heating values was consistent with the experimental heating values. In general, carbon elements have a direct impact on heating value and the results were coherent with the carbon composition in Table 1. Table 2 also shows the percentage difference between the calculated value and the experimental data. The percent MAPE range for all samples were between 3-12%. On average, the percentage error was 7%. This implies that the simulation model estimated the heating value of the CW and MSW samples with an excellent range of accuracy.

Table 2: Heating values comparison and simulated combustion temperature results.

Category	Sample #	Experiment, kJ/kg	Simulation, kJ/kg	Simulated comb. temp, C	MAPE, %
Clinical waste	1	42650 <sup>a</sup>	39042	1075	8
	2	29120 <sup>a</sup>	29868	857	3
	3	15570 <sup>a</sup>	17471	543	12
Municipal solid waste	4	23193 <sup>b</sup>	24824	719	7
	5	16217 <sup>b</sup>	16533	444	2
	6	12174 <sup>b</sup>	13693	398	12
<b>Average MAPE, %</b>					<b>7</b>

<sup>a</sup> LHV, <sup>b</sup> HHV

### 3.2. Model analysis

The results for variation in moisture and VM content of CW and MSW feed to heating value, and combustion temperature is shown in Figure 1. Figure 1a shows that when the moisture content is increased, the heating value decreases. These findings are intuitive

and agree with Liang et al. (2008). Higher moisture content typically leads to lower calorific value as various moisture content levels in mass burning affect combustion performance. Furthermore, the higher moisture content also reduces reaction temperature and results in poor ignition (Johari et al., 2012). This, in return, lowers the combustion temperature, as shown by the decreasing combustion temperature. On the other hand, the increase in VM had an inverse effect on heating values. As shown in Figure 1b, the heating values increase upon increasing VM content. High VM content contributes to easy fuel ignition, where a lower minimum ignition temperature is needed. The impact, however was less significant compared to moisture content. Nevertheless, the result ties well with theoretical data wherein VM significantly impacts thermal decomposition.

Specifically for CW, an additional heat duty is required to increase the temperature to around 1000 C for complete disintegration. Figure 2 shows the additional heat duty required with variation in moisture content. This is typically achieved through burning of fuel i.e., diesel and natural gas. Obviously, a larger temperature difference needs more heat or fuel. Therefore, proper handling, segregation, sorting, waste volume, and size and feed time of the waste are detrimental to reducing the interference with reaction product combustion, increasing combustion quality, and consequently decreasing the fuel cost.

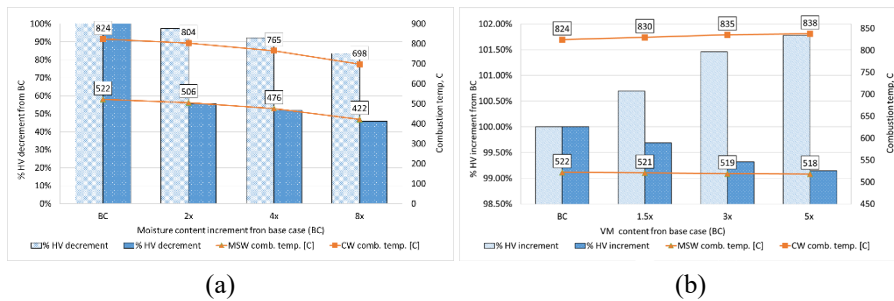


Figure 1: Effect on LHV and combustion temperature of CW and MSW for variation in (a) moisture and (b) VM content.

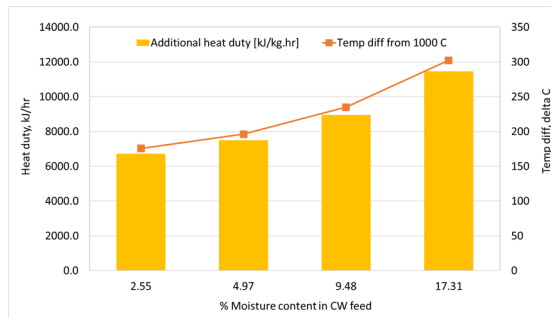


Figure 2: Additional heat required and temperature difference to achieve the CW decomposition temperature of 1000 C.

#### 4. Conclusion

The objective of this study to develop a mechanistic model of the CW and MSW incineration process was achieved. Using customized NC solid properties, flowsheet, and a general coal enthalpy model with heat correlations in Aspen Plus, the model demonstrates excellent agreement with experimental data with a mean absolute

percentage error of 7%. Model analysis of waste feed content agrees well with results found in the literature. The model shows that lower moisture and higher VM content contribute to higher heating values. However, the effect of moisture content is much more significant than VM content. The model was also extended to predict combustion temperature and the additional heat required to achieve decomposition temperature. The model reveals that waste feed quality is detrimental to the incinerator's performance. In future work, we will expand the model to other applications, such as optimization, air pollution control, and waste-to-energy systems.

## 5. Acknowledgements

The authors wish to acknowledge assistance and support from Universiti Malaysia Pahang through grant number PDU213003-1.

## References

- A. Johari, H. Hashim, R. Mat, H. Alias, M.H. Hassim, M. Rozainee, 2012, Generalization, formulation and heat contents of simulated MSW with high moisture content, *Journal of Engineering Science and Technology*, 7(6), pp. 701–710
- A.A. Erdogan, M.Z. Yilmazoglu, 2021, Plasma gasification of the medical waste, *International Journal of Hydrogen Energy*, 46(57), pp. 29108-25
- C.D. Lewis, 1982, *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*, London; Boston: Butterworth Scientific
- C.R. Lohri, S. Diener, I. Zabaleta, A. Mertenat, C. Zurbrügg, 2017, Treatment technologies for urban solid biowaste to create value products: a review with focus on low-and middle-income settings, *Reviews in Environmental Science and Biotechnology*, 16(1), pp.81-130
- C.S. Li, F.T. Jeng, 1993. Physical and chemical composition of hospital waste. *Infection Control & Hospital Epidemiology*, 14(3), pp.145-150.
- H. Zhou, A.H. Meng, Y.Q. Long, Q.H. Li, Y.G. Zhang, 2014, An Overview of Characteristics of Municipal Solid Waste Fuel in China. *Physical, Chemical Composition and Heating Value. Renewable And Sustainable Energy Reviews*, 36, pp. 107-122
- M.G. Gebreslassie, H.B. Gebreyesus, M.T. Gebretsadik, S.T. Bahta, S.E. Birkie, 2020, Characterization of Municipal Solid waste's Potential for Power Generation at Mekelle City as a Waste Minimisation strategy, *Int. Journal of Sustainable Engineering*, 13(1), pp. 68-75
- N.Y. Kirov, 1965, Specific Heats and Total Heat Contents of Coals and Related Materials are Elevated Temperatures, *BCURA Monthly Bulletin*, pp. 29-33
- O. Olanrewaju, 2019. Quantification and characterization of medical waste in public health care facilities within Akure Metropolis, Ondo State, Nigeria. *EPH-International Journal of Agriculture and Environmental Research*, 55, pp.15-30.
- P. Agamuthu, J. Barasarathi, 2021, Clinical waste management under COVID-19 scenario in Malaysia', *Waste Management & Research*, 39(1\_suppl), pp. 18-26
- P.T. Williams, 2005, *Waste treatment and disposal*, John Wiley & Sons
- R.S. Liang, Jun Fei, Shaohua Wu, Xiang Liu, Kui Dai, Na Yao, 2008, Experimental study on effects of moisture content on combustion characteristics of simulated municipal solid wastes in a fixed bed. , 99(15), pp. 7238–7246
- T.Ch. Ogwueleka, F.N. Ogwueleka, 2010, Modelling Energy Content Of Municipal Solid Waste Using Artificial Neural Network, *Iran. J. Environ. Health. Sci. Eng.*, 7(3), pp. 259-266

# On the minimization of cycle time in periodic production scheduling

Alexandros Koulouris,<sup>a</sup> Georgios P. Georgiadis,<sup>b</sup>

<sup>a</sup>*International Hellenic University, P.O. Box 141, Sindos 57400, Greece*

<sup>b</sup>*Intelligen Europe, S. Kazantzidi 47, Thermi 57001, Greece*

## Abstract

Periodic scheduling (commonly implemented in batch processing) is characterized by the process cycle time. Minimizing the cycle time maximizes throughput. This paper proposes two approaches for the calculation of the minimum cycle time for batch processes with resource sharing among tasks: an algorithmic and an MILP-based. The first approach implements an exact, exhaustive algorithm which also allows the calculation of all feasible cycle time ranges. The second approach is optimization-based, yields only the minimum cycle time but is more flexible than the algorithmic in the type of problems it can handle. Both approaches are demonstrated with illustrative examples.

**Keywords:** periodic scheduling, cycle time minimization, MILP.

## 1. Introduction

Cyclic, and more specifically, periodic scheduling is preferred by many batch industries because of the simplified shop floor control, the easily implementable schedules and easier shift decisions for the staff. Furthermore, studies show that cyclic schedules tend to be more robust compared to non-cyclic ones. Despite those facts, cyclic scheduling has not attracted much attention from the PSE research community.

Periodic scheduling is characterized by the process cycle time, i.e. the constant time interval between the start (or end) of two consecutive batches. The cycle time is directly related to throughput; minimizing the cycle time is essential in the operation of a plant and in capacity studies. Many researchers proposed exact algorithmic polynomial solvable methods for cyclic scheduling problems, where tasks do not share resources (Hanan, 1994). However, the methods available in the literature cannot optimize cases with sharable resources. Few researchers developed novel mathematical optimization approaches for cyclic scheduling (Pinto and Grossmann, 1994; Wu and Maravelias, 2019). In most of these formulations, each task can only be processed in exactly one unit throughout the cyclic schedule. Moreover, they do not address the cycle time minimization problem.

The objective of this paper is to present solutions to the minimization of cycle time in the presence of tasks sharing multiple resources. Two approaches are presented: an algorithmic and an MILP-based. In the algorithmic case, the non-overlapping constraint between tasks on the same resource yields bounds on the feasible cycle time between any pair of tasks under the assumption that all available resources are used periodically by all tasks. The developed MILP model is based on the general precedence framework (Kopanos et al., 2011) and takes into account the cyclic nature of batch scheduling through the introduction of a new set of constraints. In comparison to the solutions



found in literature, the model relaxes the requirement of using a single resource for each task throughout the schedule allowing for a better utilization of the available resources. The proposed model is relatively general and can be easily extended to consider any specific operational constraint, thus providing a basis for tackling realistic large-scale problems. Two case studies are presented, demonstrating the applicability of the proposed solutions in real-life industrial scenarios and highlighting their differences. The case studies have been modeled and scheduled within the scheduling software SchedulePro (by Intelligen, Inc., Scotch Plains, NJ, USA, [www.intelligen.com](http://www.intelligen.com)) where the algorithms have also been implemented. The MILP models have been solved using the HiGHS open-source solver (Huangfu and Hall, 2018).

## 2. Periodic Scheduling Set-up

Let  $T = \{1, 2, \dots, n\}$  be a set of  $n$  tasks corresponding to a batch that is to be executed indefinitely often. Let  $p_i > 0$  be the duration of task  $i$ . The batches must be executed periodically with a fixed cycle time,  $a$ . Let  $t(i, k)$  denote the start of the execution of task  $i$  in batch  $k$  and  $t_i$  denote the start of task  $i$  for the (randomly selected) reference batch 0. Then,

$$t(i, k) = t_i + ak \quad (1)$$

Within each iteration, the tasks are to be executed at a specific order with fixed time intervals between their starts. Let  $D_{i,j} \forall i, j \in T$  be the fixed time distance between the start of task  $j$  and the start of task  $i$  within the same batch. Then,

$$\begin{aligned} D_{ij} &= t(j, k) - t(i, k) \Rightarrow D_{ij} = t_j - t_i \\ D_{ji} &= -D_{ij}, \quad D_{ii} = 0 \end{aligned} \quad (2)$$

It is assumed that all tasks within  $T$  are ordered in ascending start times. Each task uses a pool of available resources to be used for its execution. Different tasks may use the same resources, therefore, the resource pools may overlap (partially or fully) between tasks. Feasible cycle times are those that do not create overlaps of tasks (in the same or different batches) on the same resource. Under this setting, a unique schedule can be generated with the assignment of the cycle time,  $a$ , and the allocation of resources for each task in every batch.

## 3. Algorithmic Approach

To obtain the minimum and feasible cycle time ranges in an algorithmic way, it is assumed that each task uses exclusively its own resource pool, or, if tasks are to share resources, the resource pools are identical. It is further assumed that if common resources are available, they are used in a rotation mode. As a result, all tasks will eventually "meet" on each resource but it is not known a priori to what batch these tasks will belong. Under these assumptions, a periodic schedule is completely determined by setting the value of the cycle time and the resource used for every task at the first batch. To find the minimum cycle time, it is possible to split the task set into subsets composed of either a single task (using its own resources) or of a group of tasks sharing a common resource pool. If a task uses exclusively one or multiple resources, then its feasible cycle time range is simply  $[p_i/m_i, \infty)$  where  $m_i$  is the number of resources available for task  $i$ . The analysis below will deal with tasks sharing a pool of multiple resources.

Due to the periodicity and the cyclic rotation between all resources, each task  $i$  is executed on the same resource every  $m$  batches (where  $m$  is the number of common

resources). The indices of repetitions of the same task  $i$  on the same resource can therefore be represented by  $mk_i + s_i$ , where  $k_i, s_i \in Z$  and  $s_i$  ( $0 \leq s_i < m$ ) is an integer representing the index of the batch (with respect to reference batch 0) for the first occurrence of task  $i$  in a specific resource. Note that  $s_i$  is not the same for all resources. The non-overlap constraints between two tasks  $i, j$  on a given resource can be written as:

$$\begin{aligned}
 t(i, mk_i + s_i) + p_i &\leq t(j, mk_j + s_j) \quad \text{or,} \quad t(j, mk_j + s_j) + p_j \leq t(i, mk_i + s_i) \\
 t_i + amk_i + as_i + p_i &\leq t_j + amk_j + as_j \quad \text{or,} \quad t_j + amk_j + as_j + p_j \leq t_i + amk_i + as_i \\
 t_j - t_i &\geq p_i - \alpha(mK_{ij} + S_{ij}) \quad \text{or,} \quad t_j - t_i \leq -p_j - \alpha(mK_{ij} + S_{ij}) \\
 D_{ij} &\geq p_i - \alpha(mK_{ij} + S_{ij}) \quad \text{or,} \quad D_{ij} \leq -p_j - \alpha(mK_{ij} + S_{ij})
 \end{aligned} \tag{3}$$

where  $K_{ij} = k_j - k_i$ , and  $S_{ij} = s_j - s_i$ ,  $K_{ij}, S_{ij} \in Z, -m < S_{ij} < m$ . Note that because of the cyclical rotation over all resources, and despite the fact that  $s_i$  and  $s_j$  are resource-dependent, their difference  $S_{ij}$  is not.

For a feasible schedule to exist, for every pair of tasks  $i, j$  executed on the same resource with given value of  $S_{ij}$ , there must exist an integer value of  $K_{ij}$  so that  $D_{ij}$  belongs to one of the allowed intervals, namely  $[p_i - \alpha(mK_{ij} + S_{ij}), -p_j - \alpha(m(K_{ij} - 1) + S_{ij})]$ . In other words, the non-overlapping constraints can be re-written as:

$$p_i - \alpha(mK_{ij} + S_{ij}) \leq D_{ij} \leq -p_j - \alpha(m(K_{ij} - 1) + S_{ij}) \tag{4}$$

or, if the two inequalities are separated,

$$\begin{aligned}
 D_{ij} &\geq p_i - \alpha(mK_{ij} + S_{ij}) \\
 D_{ij} \leq -p_j - \alpha(m(K_{ij} - 1) + S_{ij}) &\Rightarrow D_{ij} \geq p_j - \alpha(m(1 - K_{ij}) - S_{ij})
 \end{aligned} \tag{5}$$

These two constraints are, in fact, symmetrical provided that in the second inequality we use:  $K_{ji} = 1 - K_{ij}$  and  $S_{ji} = -S_{ij}$ . The feasible values of the cycle time,  $\alpha$ , should therefore satisfy the following constraints:

$$\begin{aligned}
 D_{ij} &\geq p_i - \alpha(mK_{ij} + S_{ij}) \quad \forall i, j \in T \\
 K_{ij} + K_{ji} &= 1 \\
 S_{ji} &= -S_{ij}
 \end{aligned} \tag{6}$$

Depending on the sign of the term  $p_i - D_{ij}$ , Eq. (6) provides upper or lower bounds on  $\alpha$  for different values of  $K_{ij}, S_{ij}$ . After a few algebraic steps (outlined in Koulouris and Georgiadis, 2022), these bounds can be calculated as follows:

$$\begin{aligned}
 a &\geq \frac{p_i + p_j}{m} \\
 \text{if } p_i - D_{ij} > 0 \text{ then,} \\
 a &\geq \max \left\{ \frac{p_j + D_{ij}}{m - S_{ij}}, \frac{p_i - D_{ij}}{S_{ij}} \right\} \text{ for } S_{ij} \neq 0 \\
 \text{if } p_i - D_{ij} = 0 \text{ then,}
 \end{aligned} \tag{7}$$

$$a \geq \frac{p_j + D_{ij}}{m - S_{ij}}$$

if  $p_i - D_{ij} < 0$  then,

$$\frac{p_j + D_{ij}}{m(1 - K_{ij}) - S_{ij}} \leq a \leq \frac{p_i - D_{ij}}{mK_{ij} + S_{ij}} \text{ for } \frac{p_i - D_{ij}}{p_i + p_j} - \frac{S_{ij}}{m} \leq K_{ij} < -\frac{S_{ij}}{m} \text{ and } K_{ij} \neq 0$$

$$a \geq \frac{p_j + D_{ij}}{m - S_{ij}} \text{ for } K_{ij} = 0$$

As seen in Eq. (7), the values of  $K_{ij}$  are either fixed or bounded on both sides and that means that the search of feasible ranges is also bounded. Using different values of  $K_{ij}, S_{ij}$  within their admissible ranges, bounds on cycle time are calculated. The union of all these feasible ranges provides the values of the cycle time that, for a given pair of tasks, guarantee repetitive execution of these tasks with no overlaps. This calculation process must be repeated for all task pairs; the intersection of all pair-wise feasible ranges will result in the feasible cycle time ranges for all tasks sharing common resources. The lower bound of these ranges is the minimum process cycle time.

#### 4. MILP Approach

While computationally efficient, the previously described exact method is based on a heuristic of utilizing the resources in a cyclical manner that could affect the output resulting to very good but potentially suboptimal solutions. Therefore, an MILP-based approach, that can ensure the optimality of the generated solutions and can tackle problems with non-identical resource pools, is also developed. The model is based on the known general precedence framework, which has been extended to address periodic scheduling optimization problems. Given a specific number of batches  $B$  of tasks  $T$  to be processed in the available resources  $R$ , the optimization model generates a complete optimal schedule with minimum cycle time. A set of binary allocation variables ( $Y_{i,b,r}$ ) and binary sequencing variables ( $X_{i,b,i',b'}$ ) are employed. The first are enabled when a batch  $b$  of task  $i$  is processed by resource  $r$ , while the latter when a batch  $b$  of task  $i$  is processed prior to batch  $b'$  of task  $i'$  using the same resource. The proposed model consists of constraints (8)-(15) and the objective function (16), which are described below.

$$\sum_{r \in IR_{i,r}} Y_{i,b,r} = 1 \quad \forall i \in T, b \in IN_b \quad (8)$$

$$L_{i,b} + \sum_{r \in IR_{i,r}} (Y_{i,b,r} \cdot p_i) = C_{i,b} \quad \forall i \in T, b \in B \quad (9)$$

$$L_{i,b} = t_i + (b - 1) \cdot \alpha \quad \forall i \in T, b \in B \quad (10)$$

$$L_{i',b'} \geq C_{i,b} - M \cdot (1 - X_{i,b,i',b'}) - M \cdot (2 - Y_{i,b,r} - Y_{i',b',r}) \quad \forall i, i' \in T, i < i', b, b' \in B, r \in (IR_{i,r} \cap IR_{i',r}) \quad (11)$$

$$L_{i,b} \geq C_{i',b'} - M \cdot X_{i,b,i',b'} - M \cdot (2 - Y_{i,b,r} - Y_{i',b',r}) \quad \forall i, i' \in T, i < i', b, b' \in B, r \in (IR_{i,r} \cap IR_{i',r}) \quad (12)$$

$$L_{i,b'} \geq C_{i,b} - M \cdot (2 - Y_{i,b,r} - Y_{i,b',r}) \quad \forall i \in T, b, b' \in IN_b, b < b', r \in IR_{i,r} \quad (13)$$

$$Y_{i,b,r} = Y_{i,b+NB,r} \quad \forall i \in R, b \leq |B| - NB, r \in IR_{i,r} \quad (14)$$

$$L_{i,b+NB} \geq C_{i,b} \quad \forall i \in T, b \leq |B| - NB \quad (15)$$

$$\text{Min } \alpha \quad (16)$$

Constraint set (8) ensures that each batch  $b$  of a task  $i$  will be processed by exactly one resource  $r$ . Notice that in order to reduce the model's size, we only consider the batches in the set  $IN_b$ . This set denotes the batches for which we optimize the allocation decisions. Let us assume that we want to create an optimal schedule for a set of batches  $B=\{b0, b1, \dots, b10\}$  and decide that  $IN_b=\{b0, b1\}$ , then the allocation of the tasks in the available resources is repeated every two batches. For example, if task  $i1$  is allocated to resource  $r2$  and  $r3$  in batches  $b0$  and  $b1$ , then the allocation for the upcoming batches will be  $b2 \rightarrow r2, b3 \rightarrow r3$  etc. The size of set  $IN_b$  is a user-defined parameter depending on the needs and wants of the production engineer using the model. A larger size means potentially lower cycle times but requires more CPU time. Next, constraint set (9) guarantees the proper connection between the starting and completion times of every batch of each task. In particular, the completion time  $C_{i,b}$  must equal the starting time  $L_{i,b}$  plus the required processing time  $p_i$ . Constraints (10) are essential since they define the problem as a cyclic scheduling problem. More specifically, we define that the starting time of a batch  $b$  of task  $i$  must be equal to the start time of the first batch  $t_i$ , plus the cycle time times the batch number minus 1. For example, let us assume a task with starting time 5 hours, if the optimal cycle time is calculated to be 2 hours, then the starting time of the first batch will be 5 hours, of the second 7 hours, the third 9 hours and so on. Constraints (11) and (12) are the precedence constraints that are necessary in order to respect the disjunctive constraints (any resource can only process one task at a time). Constraints (11) and (12) are complementary. If a batch  $b$  of a task  $i$  is processed prior to a batch  $b'$  of task  $i'$  ( $X_{i,b,i',b'}=1$ ) and both of them are processed on the same resource ( $Y_{i,b,r}=Y_{i',b',r}=1$ ), then the batch  $b'$  of task  $i'$  must start after the completion of batch  $b$  of task  $i$ .  $M$  is a large number (big-M parameter), which is necessary in order to relax the constraints, in case the previously mentioned conditions are not valid. Constraint set (13) states that a batch  $b'$  of task  $i$  must start after the completion of a batch  $b < b'$  of the same task in case both are processed by the same resource. Again, to reduce the model's size we only consider this constraint for the subset  $IN_b$ . The next two constraints ensure that the generated schedule will be a cyclic one.  $NB$  is the number of batches in the subset  $IN_b$ . Constraints (14) guarantee that a batch  $b+NB$  of a task  $i$  will be processed in the same resource with batch  $b$ , since the schedule must be cyclic. Similarly, for the same couple of batches, batch  $b+NB$  must start after the completion of batch  $b$ . Goal is the minimization of the cycle time  $\alpha$ .

## 5. Examples

To demonstrate the use of the above approaches, two examples are presented here. The first example refers to a process producing a small molecule API. It consists of 11 tasks: 4 tasks sharing 3 reactors, 5 tasks sharing 2 tanks and two tasks using exclusive equipment (1 filter and 3 dryers, respectively). Details about the start time and duration of these tasks can be extracted from Intelligen, Inc. (2022) where this process is presented as a SchedulePro case study. As an extension to this example, it was assumed that 2 transfer panels are used as auxiliary equipment for all material transfers (11 in total). Their duration and start times in hours are  $\{4, 1, 1, 1, 1.5, 1, 1, 1, 1, 8, 2\}$  and  $\{0, 11, 15, 16, 17, 19.5, 20.5, 24, 25, 43, 51\}$ , respectively. Both the algorithmic and the MILP approaches yield instantaneously the same minimum cycle time of 17.67 hours. The bottleneck resources are the transfer panels. A Gantt chart showing the bottleneck equipment in a sequence of 15 batches scheduled at the minimum cycle time is shown in Figure 1. The algorithmic approach yields also all feasible cycle time ranges which are:  $\{[17.67, 18.5], [19.5, 20], [27, 27.75], 29.25, [31.5, 32.5], [33.5, \infty)\}$ .

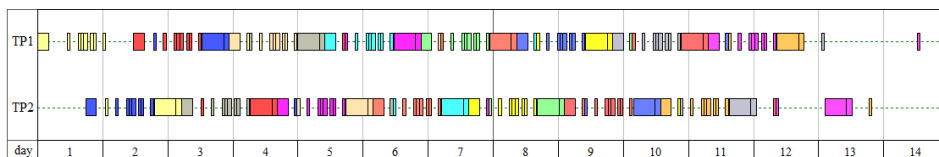


Figure 1. Gantt chart of 15 batches for the small molecule API case study.

Knowledge of the feasible ranges is useful if the plant wants to operate at specific cycle time which is not necessarily the minimum.

The second case is an example process where the algorithmic approach yields a suboptimal solution as a result of the resource rotation assumption. The process has 3 tasks with durations and start times  $\{6, 2, 11\}$  and  $\{0, 2, 5\}$  hours, respectively. It is assumed that 3 equipment resources (Eq1-3) are available. In this case, the algorithmic approach yields a minimum cycle time of 8 hours while the optimal value calculated by the MILP approach ( $IN_b=3$ ) is 7 hours. As shown in Figure 2, unlike the algorithmic approach, the optimal policy assigns an exclusive resource (Eq1) to the first task.

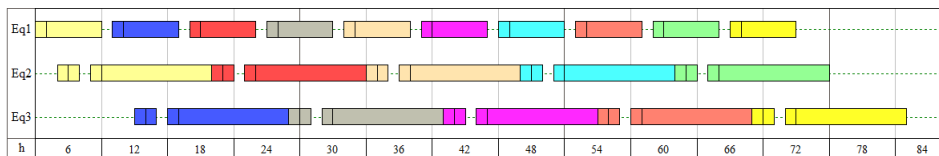


Figure 2. Gantt chart of 15 batches for the example process.

## 6. Conclusions

Two approaches are presented in this paper for the calculation of the minimum cycle time in periodic batch process scheduling with resource sharing. The algorithmic approach is an exact, exhaustive method that yields the minimum cycle time and all feasible cycle time ranges for an infinite scheduling horizon under the assumption that all common resources are used in rotation. The MILP approach is more general since it makes no a priori assumption on resource assignments and generates optimal solutions for predefined scheduling horizons in short computational times.

## References

- Hanen, C., 1994. Study of a NP-hard cyclic scheduling problem: The recurrent job-shop. *European journal of operational research*, 72(1), pp.82-101.
- Huangfu, Q. and Hall, J.J., 2018. Parallelizing the dual revised simplex method. *Mathematical Programming Computation*, 10(1), pp.119-142.
- Intelligen Inc. (2022) <https://www.intelligen.com/products/schedulepro-examples/> (access November 2022)
- Kopanos G., Puigjaner L., Georgiadias M.C., 2011. Production Scheduling in Multiproduct Multistage Semicontinuous Food Processes. *Industrial & Engineering Chemistry Research*, 50 (10), pp. 6316-6324.
- Koulouris, A., Georgiadias, G., 2022. An exact algorithm for calculating the minimum and feasible ranges of cycle time in periodic scheduling with shared resources, *submitted to Computers and Chemical Engineering*.
- Pinto, J.M. and Grossmann, I.E., 1994. Optimal cyclic scheduling of multistage continuous multiproduct plants. *Computers & Chemical Engineering*, 18(9), pp.797-816.
- Wu, Y. and Maravelias, C.T., 2019. A general model for periodic chemical production scheduling. *Industrial & Engineering Chemistry Research*, 59(6), pp.2505-2515.

# Non-Pareto optimal solutions as enablers for versatile heat exchanger networks

David Huber<sup>a</sup>, Felix Birkelbach<sup>a</sup>, René Hofmann<sup>a</sup>

<sup>a</sup>*Institute for Energy Systems and Thermodynamics, TU Wien, Getreidemarkt 9/BA,  
Wien, 1060, Austria  
[david.huber@tuwien.ac.at](mailto:david.huber@tuwien.ac.at)*

## Abstract

Assessing a heat exchanger network (HEN) based on total annual costs (TAC) alone is no longer appropriate. Climate-damaging emissions must also be considered. Thus, we use CO<sub>2</sub> emissions and TAC as objectives in the fully linearized heat exchanger network synthesis (HENS). Due to uncertainties concerning emissions and equipment costs, a refurbishment of the HEN may become necessary even before its life cycle has been reached. Since a significant share of the system costs is related to the piping, it is economical to change only the heat exchanger (HEX) areas. To assess the potential for the refurbishment of a given HEN in the future, we constrain the Pareto optimal HEN configurations and recompute sub-Pareto fronts. Our results show that there are obviously superior configurations regarding distance and coverage of the sub-Pareto front concerning the initial Pareto front. The hypervolume indicator (HVI) and the coverage were used as quantitative indicators to evaluate different configurations. However, these characteristic parameters do not correlate with empirically determined high-quality configurations. This contribution has laid an essential foundation towards developing versatile and optimal HEN.

**Keywords:** MILP, multi-objective optimization, non-dominated solutions, locked HEN configuration

## 1. Introduction

To save energy and costs, industrial energy systems can connect hot and cold process streams using a HEN. The optimal interconnection of the HEN can be calculated using HENS. If we look at various literature on HENS, the aim is almost entirely to achieve a cost minimum. The single-objective approach cannot consider environmental impacts. To produce a holistic set of choices for decision-makers, multi-objective optimization (MOO) with CO<sub>2</sub> emissions and costs is the way to go.

Many optimal HEN can provide a comprehensive basis for decision-making. However, determining which HEN to choose is a significant challenge for decision-makers. The selected HEN should meet the current conditions and be reasonable in case of unforeseen developments. For example, political or economic changes such as increased CO<sub>2</sub> certificate costs or energy costs can adversely influence the selected HEN's performance. Since such changes can never be fully considered, it is more effective to present decision-makers with fewer but more robust solutions.

According to Peters et al. (2003), the piping costs in fluid processing plants can be up to 80 % of equipment costs. This corresponds to about 25 % of the total investment costs. The costs associated with adjusting the operating point of the HEN can be reduced

significantly when keeping the piping configuration as it is. With an unchanged HEN, the operating point, respectively the load on the utilities can be influenced by varying the HEX areas.

In this paper, we use HENS to propose a method to assess the versatility of Pareto-optimal configurations and to identify the ones that can be refurbished economically in the future if necessary. To determine the versatility of a configuration, we compute sub-Pareto fronts from the Pareto optimal configurations. Investigation of the Pareto front has shown that only a few unique configurations of HEN exist as non-dominated solutions. Sub-Pareto fronts can be generated using those unique configurations, which differ only in the utility load and HEX areas. The results show that the sub-Pareto fronts cover large areas of the initial Pareto front while containing a variety of only weakly dominated solutions.

## 2. Methods

The HENS formulation used in this paper is based on the stage-wise superstructure formulation according to Yee and Grossmann (1990). With this formulation, heat exchange can occur in counterflow between streams in  $k$  stages with stream splits. The hot and cold utilities are located at the ends of the streams, respectively.

### 2.1. Objectives

The first of the two objectives describes the TAC and was adopted from Yee and Grossmann (1990). The second objective represents the CO<sub>2</sub> emissions due to the operation and manufacturing of the HEN, according to Pintarič et al. (2019). This method evaluates CO<sub>2</sub> emissions based on hot utility consumption and economic activity level according to the economic input-output life cycle assessment (EIO-LCA). In contrast to Pintarič et al. (2019), we also consider emissions from the circulation pumps power consumption.

### 2.2. Linearization

To apply highly efficient mixed integer linear programming (MILP) solvers, all non-linear terms (reduced stream HEX area, reduced utility HEX area and logarithmic mean temperature difference (LMTD)) are linearized. The non-linear reduced stream HEX area is tightened to the feasible solution space according to Beck and Hofmann (2018) and approximated with super positioned planes. The convex approximation allows the efficient transfer to MILP without using binary variables. Analogously, the reduced utility HEX area is tightened. In contrast to Beck and Hofmann (2018), the non-linear relation is approximated with several straight lines. The function of LMTD is approximated using plane triangles on a regular grid. The transformation to MILP is done using a highly efficient logarithmic coding of the binary variables introduced by Vielma and Nemhauser (2011). With this method,  $T$  Simplices require only  $\lceil \log_2(T) \rceil$ , instead of  $T$  binary variables.

### 2.3. Optimization Approach

The developed method can be divided into the following main steps:

1. Calculate Pareto optimal solutions
2. Constrain incidence matrix in Pareto optimal HEN configuration
3. Calculate Pareto optimal solutions of subproblems

The Pareto front is calculated using the adaptive weighted sum method from Kim and de Weck (2005). Compared to the weighted sum method, the Pareto front can also be solved in non-convex regions and with well-distributed solutions. In the remainder of this paper, we will refer to this Pareto front as an initial Pareto front.

In the second step, the unique configurations of the initial Pareto front are identified. From the resulting set of unique configurations, new subproblems are created. The incidence matrix representing the HEN configuration is blocked with constraints. Therefore, the degree of freedom and thus, the complexity can be reduced significantly, leading to short computation time.

In the final step, a new Pareto front is calculated for each subproblem with the adaptive weighted sum method. Note that the Pareto optimal solutions of the subproblems are not optimal regarding the initial problem. The sub-Pareto fronts always contain at least one non-dominated Pareto optimal solution concerning the initial problem.

### 3. Use Case

The method presented in the previous section was applied to a representative use case with four hot and five cold process streams, according to Linnhoff and Flower (1978). The stream data, HEX, and utility costs are summarized in Table 1. The CO<sub>2</sub> emissions factor for hot utilities  $\omega_{HU}$  and investment costs  $\omega_{INV}$  were taken from Pintarič et al. (2019). The emission factor for the cold utilities is calculated assuming that 10 % of the cooling load must be supplied as electrical energy. With 8000 h full load hours and an average electricity mix,  $\omega_{CU}$  results in 0.0016 t/kWy. The minimum allowed approach temperature was set to  $\Delta T_{\min} = 20^\circ\text{C}$  and heat transfer is possible in two stages.

Table 1: Case-study: stream and cost data according to Linnhoff and Flower (1978).

Stream	$T_{\text{in}} / ^\circ\text{C}$	$T_{\text{out}} / ^\circ\text{C}$	$F / \text{kW/K}$	$H / \text{kW/m}^2\text{K}$
H1	327	40	100	0.50
H2	220	160	160	0.40
H3	220	60	60	0.14
H4	160	45	400	0.30
C1	100	300	100	0.35
C2	35	164	70	0.70
C3	85	138	350	0.50
C4	60	170	60	0.14
C5	140	300	200	0.60
HU	330	250	-	0.50
CU	15	30	-	0.50

HEX data:  $c_f = 2000 \text{ \$/y}$ ,  $c_v = 70 \text{ \$/m}^2\beta\text{y}$ ,  $\beta = 1$ ,  $\omega_{INV} = 399 \cdot 10^{-6} \text{ t/\$y}$

Utility data:  $c_{HU} = 60 \text{ \$/kW}_y$ ,  $c_{CU} = 6 \text{ \$/kW}_y$

CO<sub>2</sub> emissions:  $\omega_{HU} = 1.436 \text{ t/kW}_y$ ,  $\omega_{CU} = 0.0016 \text{ t/kW}_y$

#### 3.1. Optimization Framework

The optimization problem has been modeled using *Yalmip* R20210331 Lofberg (2004) in *MATLAB* R2022a. The optimization problem was solved with the state-of-the-art solver *Gurobi* 9.5.2 on a 128-core system (AMD EPYC 7702P) with 128 GB RAM.

#### 3.2. Linearization Settings

The reduced stream HEX areas were approximated with as many planes as necessary to obtain a root mean square error (RMSE) of less than 1.0 %. In most cases, this was achieved with four planes. The LMTD was approximated with 32 plane triangles resulting in an RMSE of 1.0 %. The reduced HEX areas of the utilities were approximated with two straight lines resulting in an RMSE below 2.0 %.



### 3.3. Pareto Front Settings

The problem was normalized between the nadir and utopia point to calculate the Pareto front. The solutions were estimated to have a minimum relative distance of 2.5 %. Below that, solutions are considered overlapping and excluded. Within the adapted weighted sum method, iteration was performed until a maximum relative point distance of 5.0 % was met. If larger distances occur in the Pareto front, no solution exists.

The normalization of the problem with locked HEN configurations is done based on the respective nadir and utopia points. Each subproblem was thus normalized with respect to its nadir and utopia point, different from the initial solution.

## 4. Results

Figure 1 shows the calculated Pareto fronts of the use case with 27 different solutions. Out of  $2^{49}$  possible HEN configurations, only 13 are identified as a solution to the problem. Identical configurations are shown in the same color.

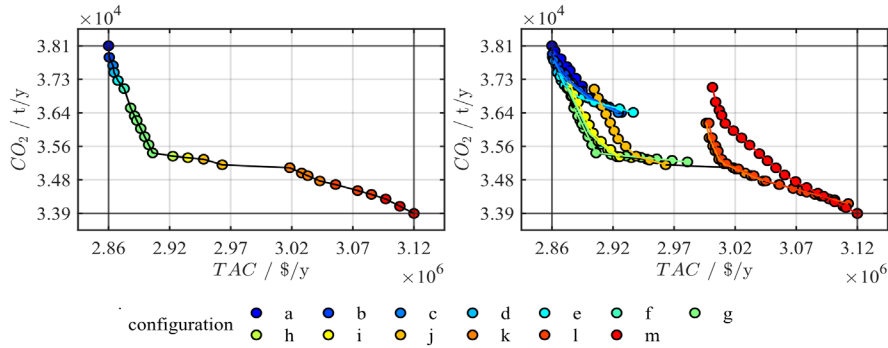


Figure 1: Left: initial Pareto front with 27 non-dominant solutions and 13 different HEN configurations. Right: initial Pareto front with 27 non-dominant, sub-Pareto fronts with 125 dominated solutions and 13 different HEN configurations.

The resulting 13 sub-Pareto fronts with locked HEN configuration are shown in Figure 1 (right). Obviously, some solutions are closer to the initial Pareto front and cover a broader range than others. The hypervolume indicator (HVI) according to Guerreiro et al. (2021) was calculated for each Pareto front to evaluate the quality of the different configurations. The reference point is always the nadir point of the initial Pareto front. The ratio of the HVI between a sub-Pareto front and the initial Pareto front can be interpreted as a dimensionless area ratio. The closer the HVI ratio is to 100 %, the closer the sub-Pareto front is to the initial Pareto front. The results are summarized in Table 2. Since the single-criterion evaluation of the different configurations considers versatility only to a limited extent, the coverage is also evaluated and listed in Table 2. The coverage quantifies the overlap of the sub-Pareto front concerning the initial Pareto front.

Since there is no direct correlation between coverage and HVI ratio, it is up to the decision-maker to find a trade-off between these two indicators.

Table 2: Characteristic values of subproblems with locked HEN configuration.

config.	# of solutions	HVI ratio		coverage	
		/ %		annual CO <sub>2</sub> emissions	TAC
/ -	/ -			/ %	/ %
a	7	32.42		9.90	23.95
b	9	52.23		22.49	34.91
c	8	52.64		21.12	31.14

d	2	33.46	3.67	8.16
e	7	52.51	23.90	21.39
f	8	85.60	16.92	40.99
g	12	88.57	37.13	32.21
h	9	85.56	15.94	41.01
i	9	84.74	16.20	43.33
j	11	54.16	23.39	45.11
k	9	55.93	18.67	34.49
l	15	55.93	45.67	48.00
m	19	48.46	47.44	75.22

Configurations f-i (light green to orange in Figure 1) have a high HVI ratio of over 80 % but cover only small areas of the Pareto front. The peripheral areas are not covered at all. Configuration g offers the best trade-off between HVI ratio and coverage. However, configurations l and m (red and dark red in Figure 1) cover at least the peripheral area with low CO<sub>2</sub> emissions well. Due to the higher HVI ratio, configuration l is preferable. The second configuration (blue in Figure 1) covers a substantial area in the low TAC region. Based on this empirical analysis, the obviously relevant solutions can be reduced to the three configurations b, g and l. The stream plots of the three configurations can be found in the Appendix. As can be seen in Figure 2, these solutions cover almost the entire Pareto front. Remarkably, all solutions have large areas with only weakly dominated solutions.

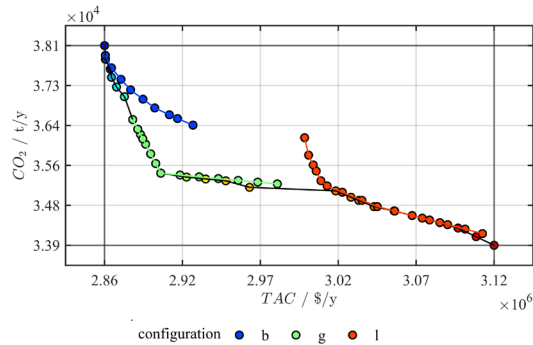


Figure 2: Pareto front with three non-dominant solutions and 33 weakly-dominated solutions of configuration b (blue), g (light green) and l (red).

## 5. Conclusion

In this paper, we used MOO to incorporate TAC and CO<sub>2</sub> emissions into the design optimization process of a HEN. Our results show that some sub-Pareto fronts cover the initial Pareto front to a large extent. Although each sub-Pareto front contains at least one Pareto optimal solution, some HEN configurations are obviously superior to others. We argue that a desirable configuration is one whose solutions are as close as possible to the initial Pareto front and at the same time provide high coverage and therefore versatility. In this paper, we attempted to quantify the trade-off between versatility and optimality of a given configuration. HVI and coverage were considered, but results were not satisfactory. The analysis of Pareto-optimal configurations, which we proposed in this paper, can be a useful tool to identify the most versatile HEN. More research on quantitative measures for the optimality of these configurations is still required.

## 6. Acknowledgment

This work was funded by the Austrian research promotion agency (FFG) under grant number 884340. The authors want to express their great acknowledgment.

## References

- M. S. Peters, K. D. Timmerhaus, R. E. West, 2003, *Plant Design and Economics for Chemical Engineers*, 5th ed. McGraw-Hill Chemical Engineering Series, New York: McGraw-Hill.
- T. F. Yee, I. E. Grossmann, 1990, Simultaneous Optimization Models for Heat Integration - II. Heat Exchanger Network Synthesis, *Computers & Chemical Engineering*, 14 (10): 1165-84, [https://doi.org/10.1016/0098-1354\(90\)85010-8](https://doi.org/10.1016/0098-1354(90)85010-8).
- A. Beck, R. Hofmann, 2018, A Novel Approach for Linearization of an MINLP Stage-Wise Superstructure Formulation, *Computers & Chemical Engineering*, 112 (April): 17-26, <https://doi.org/10.1016/j.compchemeng.2018.01.010>.
- J. P. Vielma, G. L. Nemhauser, 2011, Modeling Disjunctive Constraints with a Logarithmic Number of Binary Variables and Constraints, *Mathematical Programming*, 128 (1): 49-72, <https://doi.org/10.1007/s10107-009-0295-4>.
- I. Y. Kim, O. L. de Weck, 2005, Adaptive Weighted-Sum Method for Bi-Objective Optimization: Pareto Front Generation, *Structural and Multidisciplinary Optimization*, 29 (2): 149-58, <https://doi.org/10.1007/s00158-004-0465-1>.
- B. Linnhoff, J. R. Flower, 1978, Synthesis of Heat Exchanger Networks: I. Systematic Generation of Energy Optimal Networks, *AIChE Journal*, 24 (4): 633-42, <https://doi.org/10.1002/aic.690240411>.
- Z. N. Pintarič, P. S. Varbanov, J. J. Klemeš, Z. Kravanja, 2019, Multi-Objective Multi-Period Synthesis of Energy Efficient Processes under Variable Environmental Taxes, *Energy*, 189 (December): 116182, <https://doi.org/10.1016/j.energy.2019.116182>.
- J. Lofberg, 2004, YALMIP: A toolbox for modeling and optimization in MATLAB, *IEEE International Conference on Robotics and Automation (IEEE Cat. No.04CH37508)*, 284-289, <https://doi.org/10.1109/CACSD.2004.1393890>.
- A. P. Guerreiro, C. M. Fonseca, L. Paquete, 2021, The Hypervolume Indicator: Problems and Algorithms, *ACM Computing Surveys*, 54 (6): 1-42, <https://doi.org/10.1145/3453474>.

## Appendix

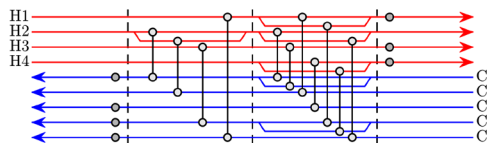


Figure 3: Stream plot for configuration b: 11 stream HEX, 4 hot and 3 cold utilities.

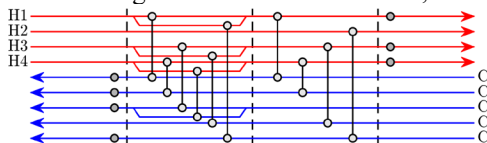


Figure 4: Stream plot for configuration g: 10 stream HEX, 4 hot and 3 cold utilities.

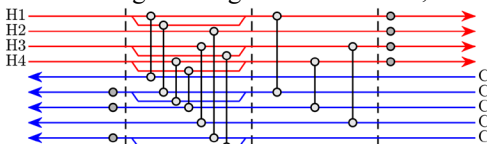


Figure 5: Stream plot for configuration l: 10 stream HEX, 3 hot and 4 cold utilities.

# Enhanced Automatic Initialization Method for Solving Large Nonlinear Algebraic Systems

Saskia Bublitz<sup>a</sup>, Erik Esche<sup>\*a</sup>, Jens-Uwe Repke<sup>a</sup>

<sup>a</sup>*Technische Universität Berlin, Chair of Process Dynamics and Operation, Str. des 17. Juni 135, 10623 Berlin, Germany*

<sup>\*</sup>[erik.esche@tu-berlin.de](mailto:erik.esche@tu-berlin.de)

## Abstract

New aspects of an interval-arithmetic (IA) based automatic initialization scheme for root-finding algorithms to efficiently solve nonlinear algebraic equation systems (NLEs) are presented. Using the model of a partial condenser, it is demonstrated how additional constraints may eliminate non-physical solutions by the help of interval arithmetic for initialization. Finally, the number of stages of a multicomponent distillation column is varied to investigate the required time of the approach to obtain initial values and solutions by root-finding as a function of the system size. The overall method finds a physically feasible solution for the distillation column with 20 stages without any initial values required.

**Keywords:** NLE, initialization, interval arithmetic, numerical iteration

## 1. Motivation

Chemical processes can become quite complex so that corresponding mathematical models are frequently hard to solve. Newton-based methods guarantee convergence in Lipschitz continuous areas, but fail, in case functions or derivatives are not defined at individual iteration points. In our study, we present further developments of an already introduced hybrid approach (Bublitz et al., 2021) that can automatically initialize nonlinear algebraic systems for root-finding algorithms using IA-based iteration methods. Its aim is to ease the challenging task of finding a converging initial point. The approach is designed to solve square problems of the form

$$f(\mathbf{x}) = \mathbf{0} \quad . \quad (1)$$

## 2. Notation

Throughout this work, multidimensional quantities are printed in bold. An interval  $\bar{x}$  is a convex set of real numerical values delimited by the lower bound  $\underline{x}$  and the upper bound  $\bar{x}$ . Intervals of multiple variables span a convex variable space, which we term a box  $\bar{x}$ . The midpoint  $m(\bar{x})$ , width  $w(\bar{x})$  and magnitude  $|\bar{x}|$  of an interval are defined as

$$m(\bar{x}) := \frac{\underline{x} + \bar{x}}{2}; \quad w(\bar{x}) := \bar{x} - \underline{x}; \quad |\bar{x}| = \max\{|\underline{x}|; |\bar{x}|\} \quad (2)$$

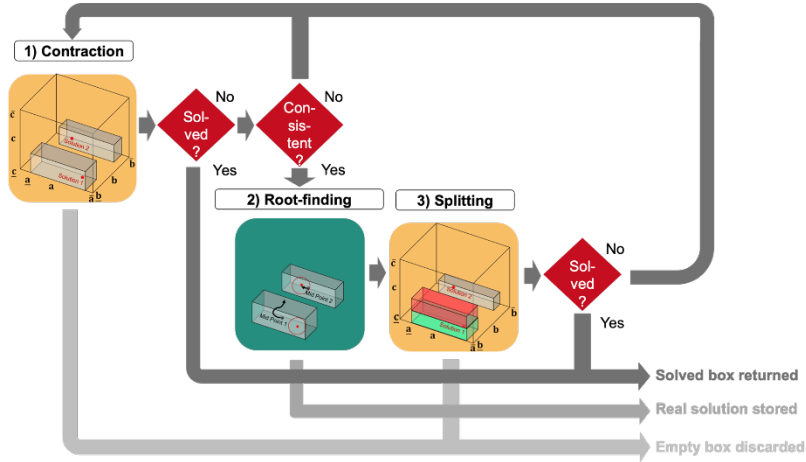


Figure 1. Three steps of the hybrid approach: 1) Contractions; 2) Root-finding; 3) Splitting

### 3. Hybrid Approach

The hybrid approach applied, consists of three steps, which are shown in Figure 1. As input, the user must provide initial intervals for all iteration variables. They span the initial box, which is reduced in the first step via IA-based contraction methods to one or more boxes (see section 3.1). A box  $\bar{x}$  is termed solved when all its intervals  $\bar{x}$  satisfy the condition

$$\Delta_{Abs} + \Delta_{Rel} \cdot |\bar{x}| \geq w(\bar{x}) . \quad (3)$$

The absolute and relative tolerances,  $\Delta_{Abs}$  and  $\Delta_{Rel}$ , must be specified. Unsolved boxes are reduced until they are consistent, i.e., their intervals no longer change from reduction step  $k - 1$  to  $k$  and both conditions,

$$\Delta_{Abs} + \Delta_{Rel} \cdot \max\{|\underline{x}^{(k-1)}|; |\underline{x}^{(k)}|\} \geq w(|\underline{x}^{(k)} - \underline{x}^{(k-1)}|), \quad (4)$$

$$\Delta_{Abs} + \Delta_{Rel} \cdot \max\{|\bar{x}^{(k-1)}|; |\bar{x}^{(k)}|\} \geq w(|\bar{x}^{(k)} - \bar{x}^{(k-1)}|) \quad (5)$$

hold for all of them. In the second step, a consistent, unresolved box is passed on to a Newton-based solver that starts from the center of the box to potentially find a root in its interior. A solution  $x^*$  to Eq. (1) in a set tolerance  $\Delta_{FTOL}$  is found when

$$\|f(x^*)\|_2 \leq \Delta_{FTOL} \quad (6)$$

holds, and is immediately returned to the user. If only one numerical solution is of interest, the procedure can be configured so that it terminates after a successful numerical iteration. Otherwise, the consistent box is split in step 3 (see section 3.3), and then further processed in contraction.

#### 3.1. Contraction

Contraction methods can reduce an initial variable space  $\bar{x}$  based on the given constraints from Eq. (1). This can be generalized by the introduction of the contracting operator  $\Gamma(f(x), \bar{x})$  (Granvilliers, 2020):

$$\Gamma(\mathbf{f}(\mathbf{x}), \bar{\mathbf{x}}) \subseteq \bar{\mathbf{x}} \quad . \quad (7)$$

While removing inconsistent variable domains, contraction methods never lose any solution within the initial box. A general contraction step is defined as

$$\{\bar{\mathbf{x}}^{(c+1)}\} := \bar{\mathbf{x}}^{(c)} \cap \Gamma(\mathbf{f}(\mathbf{x}), \bar{\mathbf{x}}^{(c)}) \quad . \quad (8)$$

Eq. (8) either breaks a box into one or more sub-boxes, verifies that the box cannot contain any solution, or proves that the box has a unique solution (Moore et al., 2009). For contraction, we use a combination of our own method Bnormal, presented in Bublitz et al. (2021), as well as the method HC4revise from Benhamou & Puget (1998). Both contracting operators have different requirements and capabilities. Details of which can be found in the respective publications. The latter method is particularly fast in reducing wide intervals. Our method on the other hand can find tighter enclosures for variable domains in equations with many instances of the same variable. For example, the equation

$$0 = x^3 - x^2 + x + y \quad (9)$$

can be used to reduce the initial interval  $[-10; 10]$  of  $x$  to  $[-1.37; 1.57]$  by HC4revise in 17 ms<sup>1</sup> through the given variable interval  $[0; 10]$  of  $y$ . Bnormal needs 29 ms but the reduced interval of  $x$  is  $[-0.55; 0]$ . Besides, Bnormal can filter infeasible regions within variable intervals to pave the way for the upcoming root-finding step. However, to use the advantages of both methods, a combination of them is also tested, in which HC4revise is applied first due to its speed. It continues until a box becomes consistent and forwards the box subsequently to Bnormal. When Bnormal could further contract the box, it returns it to HC4revise, and the alternation continues until the box does not change anymore according to Eq. (4) and (5). This alternation is denoted as a box reduction step. One box reduction step usually encompasses many contraction steps to reach consistency.

### 3.2. Root-finding

The root-finding problem is reformulated to a minimization problem of the sum of the quadratic function residuals of Eq. (1)

$$\min_x \frac{0.5}{\Delta_{FTOL}} \cdot \sum f(\mathbf{x})^2 \quad (10)$$

and can be solved by any nonlinear optimization solver, e.g., ipopt (Wächter & Biegler, 2006). The reduced variable bounds from the contraction are directly considered here. To avoid an early termination of ipopt at a point with a low curvature of the objective, Eq. (10) is scaled by the hyperparameter  $\Delta_{FTOL}$ . In addition, a Newton method is started on the root-finding problem at the local minimum in case it does not fulfill Eq. (6), to find the root of Eq (1) with the required tolerance.

### 3.3. Splitting

The applied splitting scheme bisects the variable interval, which has changed the least since the last split. If no split has occurred yet, it refers to the initial box. If there are several variable intervals that were reduced equally little, the interval is taken that causes

Table 1. Initial bounding of 36 iteration variables of the partial condenser.

Iteration Variables	Symbol	Initial Interval	Unit
Mole fractions	$x_i; y_i$	[0; 1]	-
Outlet flow rates	$F^L; F^V$	[0; 75]	kmol/h
Compressibility factors	$Z^L; Z^V$	[0; $10^2$ ]	-
Cohesion pressure	$a^L; a^V$	[0; $10^4$ ]	Pa
First derivative (EoS)	$\delta_{d1}^L; \delta_{d1}^V$	[0; $10^9$ ]	-
Second derivative (EoS, Liquid)	$\delta_{d2}^L$	$[-10^9; 0]$	-
Second derivative (EoS, Vapor)	$\delta_{d2}^V$	[0; $10^9$ ]	-
Other iteration variables		$[-10^9; 10^9]$	various

the greatest box reduction in a subsequent contraction. Ideally, one of the bisected boxes can be identified as empty and is directly discarded.

#### 4. Computer Experiments

We apply the hybrid approach on two NLEs: a partial condenser and a distillation column. The hybrid procedure stops as soon as ipopt finds a solution. The tolerances of the box reduction  $\Delta_{Abs}$  and  $\Delta_{Rel}$  are set to  $10^{-12}$  and  $10^{-10}$ . These values ensure that no solution near an infeasibility is filtered accidentally. This could otherwise be the case due to low mole fractions in the order of  $10^{-6}$ . For root-finding,  $\Delta_{FTOL}$  is set to  $10^{-8}$  to ensure that all state variables are sufficiently converged. A maximum of 10 box reduction steps and 50,000 iteration steps during root-finding are applied. Each test run is repeated three times and the averaged CPU time is presented in the results<sup>i</sup>. The success of the box reduction at reduction step  $k$  is measured by parameter  $\varepsilon$ :

$$\varepsilon^{(k)} := \sum_{b=1}^{n_{box}} \left( \sum_{v=1}^{n_{var}} \frac{w(\bar{x}_v^{(k)})}{w(\bar{x}_v^{(0)})} \cdot \frac{1}{n_{var}} \right) \cdot \frac{1}{n_{box}} . \quad (11)$$

After reduction step  $k$ , the contraction methods might have split the initial variable space into  $n_{box}$  reduced boxes. A value of  $\varepsilon = 1$  corresponds to no reduction of the initial box, while a value of  $\varepsilon = 0$  equals tight enclosures of all real valued solution(s), which fulfill the required tolerances, i.e., all solutions have been found.

##### 4.1. Partial Condenser

In this unit, a gas mixture of Toluene, Biphenyl, Benzene, Methane and Hydrogen is partially condensed to separate the low-boiling components Methane and Hydrogen at a given pressure and temperature of 35 bar and 298.15 K. The inlet flow rate is 75 kmol/h. Thermodynamic equilibrium between the ideally mixed liquid and vapor phases prevails in the unit. The chemical equilibrium according to the  $\phi$ - $\phi$  approach as well as the enthalpies rates are calculated using the cubic equation of state of Soave-Redlich-Kwong (SRK) (Soave, 1972). The full model consists of 36 equations and can be found in MOSAICmodeling<sup>ii</sup>. Table 1 summarizes its coarse initialization.

For each phase, a dimensionless equation of state is implemented (Rao, 1997) to determine the compressibility factor  $Z$

$$0 = Z^3 - Z^2 + (A - B - B^2) \cdot Z - A \cdot B . \quad (12)$$

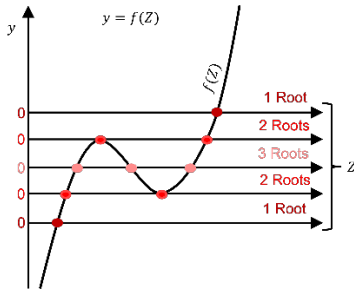


Figure 2. Number of real valued roots of a cubic equation  $y = f(Z)$  depending on its distance to the origin.

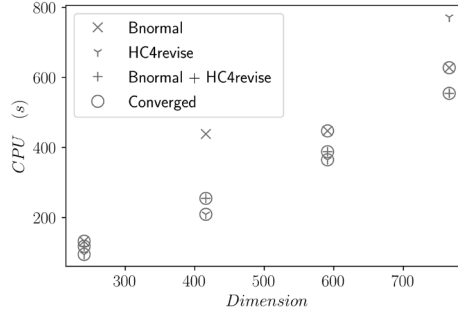


Figure 3. CPU time and convergence results after one box reduction step for different number of trays of the distillation column<sup>1</sup>.

According to Figure 2, such a cubic equation has one up to three real roots for  $Z$ . Physically interesting, however, are only the smallest root for the liquid phase and the largest one for the vapor phase. A potential third solution in between can be filtered by requiring the first order derivative of Eq. (12) with respect to  $Z$  to be greater than or equal to zero

$$\delta_{d1} = 3 \cdot Z^2 - 2 \cdot Z + (A - B - B^2), \quad \delta_{d1} \geq 0 \quad (13)$$

The compressibility factors of the liquid and the vapor phases differ in their second order derivatives. To obtain the compressibility factor of the liquid phase  $Z^L$  the condition

$$\delta_{d2}^L = 6 \cdot Z^L - 2, \quad \delta_{d2} \leq 0 \quad (14)$$

must be fulfilled, while for  $Z^V$  the condition

$$\delta_{d2}^V = 6 \cdot Z^V - 2, \quad \delta_{d2} \geq 0 \quad (15)$$

holds. To avoid interval dependency during contraction, Eq. (12) und Eq. (13) are reformulated to become

$$0 = Z^3 - Z^2 + (A - (B - 0.5)^2 + 0.25) \cdot Z - A \cdot B, \quad (16)$$

$$\delta_{d1} = \left( \sqrt{3} \cdot Z - \frac{1}{\sqrt{3}} \right)^2 - \frac{1}{3} + (A - (B - 0.5)^2 + 0.25) \quad (17)$$

This ensures that, e.g.,  $\overline{\delta_{d1}}$  can be further tightened. Let us assume  $\overline{A} = [0; 5]$ ,  $\overline{B} = [-1; 1]$ , and  $\overline{Z} = [0; 1]$ : Evaluating the righthand side of Eq. (13) by IA can only reduce  $\overline{\delta_{d1}}$  to  $[-4; 9]$ , while its reformulation according to Eq. (17) achieves  $[-2. \overline{3}; 6.25]$ . Eq. (14) to (17) have been implemented to limit the variable space to the desired solution without having to make any case differentiation.

Table 2 shows the results of the tested contraction schemes. While Bnormal finds a physical solution after one box reduction step, HC4revise has not solved the system numerically even after 10 reduction steps. The combination of both finds a solution. It is slower than Bnormal but has a better box reduction performance in terms of  $\epsilon$ .



Table 2. Results after one box reduction step applied on the partial condenser<sup>i</sup>.

Method	Solved	CPU (s)	$\epsilon$ (-)
Bnormal	yes	4.63	0.60
HC4revise	no	2.13	0.59
HC4revise + Bnormal	yes	4.82	0.57

#### 4.2. Distillation Column

A mixed liquid stream of the same components as in the previous example is thermally separated in a distillation column. The unit removes the byproduct Biphenyl from the hydrodealkylation process. Its model was originally implemented by Rajes (2020) and includes equilibrium trays, a total condenser and a total reboiler. The number of trays is varied from 5, 10, 15 to 20 which corresponds to 241, 416, 591, and 766 state variables. The processing time in dependence of the problem size is of interest in this investigation. The vapor liquid equilibrium and the enthalpies are also calculated via SRK. The initialized NLEs can be found in MOSAICmodeling<sup>iii</sup>. According to Figure 3 the combination of Bnormal and HC4revise outperforms the single methods in most cases regarding its run time. Unlike the other two methods, it always finds a solution for all NLEs after one box reduction step.

### 5. Conclusion

Without a laborious search for suitable initial values, just one box reduction step of our IA-based initialization approach suffices to reduce the initial variable space to such an extent that the applied root-finder determines feasible solutions for both test examples. Nevertheless, it should be noted that overly wide variable intervals, e.g., of  $[-10^9; 10^9]$ , may slow down the method due to extreme box splitting. Hence, process knowledge remains indispensable to narrow the operating range. The more precise this estimation is, the faster also the hybrid approach becomes.

### References

- A. Wächter, L. T. Biegler, 2006, On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, *Math. Programming*, 106, 1, p. 25 – 57
- F. Benhamou, J.-F. Puget, 1998, Revising hull and box consistency, 16-th Conference Logic Programming, p. 230 – 244
- G. Soave, 1972, Equilibrium constants from a modified Redlich-Kwong equation of state, *Chem. Eng. Science*, 27, 6, p. 1197 – 1203
- L. Granvilliers, 2020, Filtering Domains of Factorable Functions Using Interval Contractors, *Optimization of Complex Systems*, p. 99 – 108
- R. E. Moore, R. Baker Kearfott, M. J. Cloud, 2009, *Introduction to interval analysis*, SIAM
- S. Bublitz, E. Esche, J.-U. Repke, 2021, Automatic initial value generation with interval arithmetic for nonlinear process models, *Comput. Chem. Eng.*, 151, p. 107342
- S. Rajes, 2020, *Modellierung und Modellvalidierung des Hydrodealkylierungsprozesses*, Master Thesis
- Y. V. C. Rao, 1997, *Chemical Engineering Thermodynamics*, Sangam Books

<sup>i</sup> Tested on a Notebook Intel i7 Processor (8x 1.8 GHz, 8. Generation)

<sup>ii</sup> Internal ID (Partial condenser): 170665, free account from <https://mosaic-modeling.de/>

<sup>iii</sup> Internal IDs (Distillation column), 5 to 20 trays: 170907, 171127 171022 171020

# Electroplating scheduling: Closing a research gap from an automation vendor's perspective

Sophie Reimschüssel,<sup>a</sup> Uwe Fuchs,<sup>b</sup> Guido Sand,<sup>a</sup>

<sup>a</sup>*Pforzheim University of Applied Science, Tiefenbronner Straße 65, 75175 Pforzheim, Germany*

<sup>b</sup>*UF-automation GmbH, Am hohen Markstein 13, 75177 Pforzheim, Germany*

## Abstract

The production scheduling of batch operated electroplating plants gives rise to hoist scheduling problems. A literature review and real-world case studies reveal that a broad spectrum of electroplating scheduling problems can be modeled as mixed integer programs. Since problems of real-world size cannot be solved by state-of-the-art standard solvers in reasonable times, problem-specific simplification heuristics can be manually applied. An automation vendor strives for an automated solution for a wide spectrum of hoist scheduling problems to reduce the engineering effort. Therefore, the selection and parameterization of models and heuristics need to be automated. This paper argues for an approach which combines mixed integer programming with hyper-heuristics to close this research gap.

**Keywords:** Hoist Scheduling, Electroplating Plants, Mixed Integer Programming, Hyper-Heuristics

## 1. Electroplating scheduling

In batch-operated electroplating plants, piece or bulk material is coated with thin layers of metal using electrolysis. Such a plant, schematically shown in Figure 1, consists of several tanks arranged in one or more coupled rows, which are filled with liquids for coating as well as for pre- and post-treatment of the products.

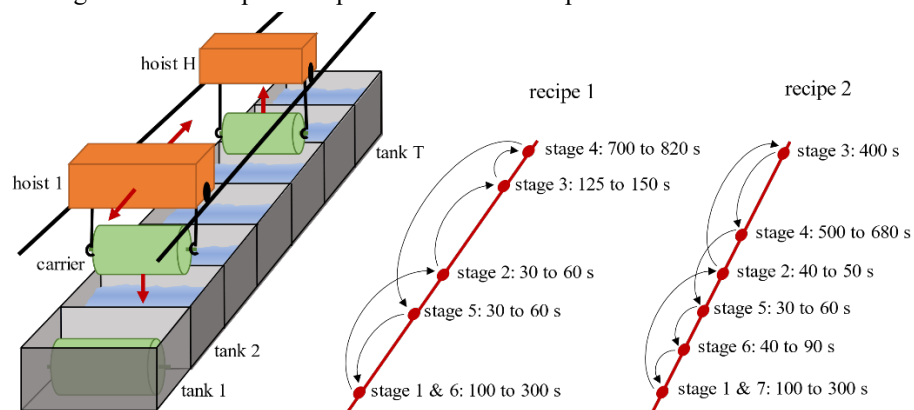


Figure 1 Schematic representation of a batch operated electroplating plant (left) and different recipes with their production sequence (right)

Items to be coated are mounted on racks or filled into barrels (in general carriers) and dipped into specific tanks in a sequence given by recipes. The processing times may vary for some stages within a given range (see Figure 1). The carriers are automatically picked up, transported and lowered by hoists, moving above the tanks on a rail. Multiple hoists operate on one rail, so they cannot pass each other and must not collide. The plant operation constitutes Hoist Scheduling Problems (HSPs), which can be thought of as Flexible Job Shop Problems (FJSP) with additional consideration of hoists. The dominating degrees of freedom are the movements of the hoists, which also determine the occupancy of the tanks. The carriers usually move back and forth and must be returned to the start of the line, as shown in Figure 1. The most common scheduling objectives are the minimization of the makespan or cycle time. In the HSPs of the electroplating processes at hand, no-wait conditions as well as time intervals for process durations must be satisfied, so that the hoists arrive at the tanks in time. The reachability combined with the collision avoidance of the hoists increase the complexity of the HSPs compared to FJSPs and simplification strategies become necessary to solve real-world problems. Multiple aspects like the combination of products, their geometrical shape, base and coating material, layer-thickness or the required space for the plant determines the plant structure and the recipes of the products. Virtually each electroplating plant is designed individually for a specific application according to these requirements. Consequently, a broad spectrum of scheduling problems exists, which differ from application to application and require different simplification strategies.

## 2. State of research and technology

### 2.1. MIP-based methods

For the HSP, different Mixed Integer Programming (MIP) models for different plant and process structures exist. The HSP belongs to the class of NP-hard problems (Lei et al., 1989) and, depending on the size of the problem instance, cannot be solved in reasonable time with a standard solver like CPLEX ([ibm.com/de-de/analytics/cplex-optimizer](http://ibm.com/de-de/analytics/cplex-optimizer)). Therefore, often simplification strategies are applied to compute a feasible schedule for one or a few days. A first distinction of simplification strategies is made between cyclic and acyclic scheduling. Cyclic scheduling is suitable for plants processing few product types in large quantities. Each cyclic schedule covers one or a few products and is executed several times in succession (Feng et al., 2018). There also exist MIP-models to switch from one cyclic schedule to another in an optimal way (Steneberg, 2013). Acyclic scheduling is suitable for plants processing many product types in small quantities. To compute an acyclic schedule in reasonable times, a rolling horizon approach can be applied, where the entire scheduling horizon is decomposed into smaller time segments (Ramin et al., 2022). If cyclic or acyclic scheduling with rolling horizons does not reduce the solution time sufficiently additional simplification heuristics become necessary.

The simplification heuristics can be classified as follows: (1) decomposition, (2) fixing variables, (3) adding constraints, (4) aggregation, (5) relaxation of constraints and (6) algorithm tuning. Cyclic scheduling (decomposition by products) and rolling horizon (decomposition by time) can be assigned to class (1). Furthermore, there are different zone partitioning approaches, in which the electroplating line is decomposed into zones which are assigned to the hoists (Li et al., 2015). A further decomposition can be done by the choice of the cycle degree (decomposition by batches). A higher degree leads to an equal or shorter cycle time but to a longer scheduling horizon and thus to a longer solution time (Feng et al., 2018). Simplification heuristics from class (2) are the assignment of certain transport operations to hoists (Fröhlich et al., 2009) or the fixed assignment of

stages to identical tanks (Reimschüssel et al., 2022). Simplification strategies from class (3) include fixed processing times (Che et al., 2010), restriction of the transport direction (Chtourou et al., 2013), and prevention of a cycle transition during a transport operation (Leung et al., 2004). The aggregation of picking up, transporting, and lowering a carrier is a simplification from class (4) (Mao et al., 2018). Different approaches exist for relaxing constraints (5) and dealing with infeasible schedules caused by this simplification. For instance, Yan et al. (2018) check whether the solution violates relaxed constraints and reinserts them into the model. Algorithm tuning (6) can be done by tighter model formulations (Leung et al., 2004) or passing a warmstart solution (Ramin et al., 2022) or lower/upper bounds on variables to the solver (Che et al., 2007).

### *2.2. Metaheuristics, constructive heuristics and hyper-heuristics*

In addition to MIP, some metaheuristics are applied to solve the HSP, such as evolutionary algorithms (Amraoui et al., 2013), variable neighborhood search (Laajili et al., 2021) or tabu search (Wang et al., 2019). There are also various constructive heuristics like variations of the earliest starting time heuristic (Amraoui et al., 2016) or the adaptive time window heuristic (Paul et al., 2007). Hyper-heuristics are used to automate the selection and parameterization of metaheuristics or constructive heuristics. For a given problem instance, a hyper-heuristic (often a metaheuristic itself or a neural network) selects a suitable metaheuristic/constructive heuristic from a pool of low-level heuristics and determines an appropriate parameterization (Burke et al., 2019).

### *2.3. Commercial solutions*

Generic production planning systems are usually not applicable for the electroplating scheduling since they neglect the hoists. Scheduling without hoists and adding them to an existing schedule later is usually infeasible due to collisions of hoists or hoists not reaching transport operations in time. Specific software solutions for scheduling batch-operated electroplating plants exist, such as GALVATEK ([galvatek.eu](http://galvatek.eu)), DiTEC ([ditec-gmbh.com](http://ditec-gmbh.com)) or ICOM Automation ([icom-automation.de](http://icom-automation.de)), but their market penetration is limited due to complex manual adjustments required for the specific problem instances.

## **3. Automation vendor's perspective**

In practice the scheduling is typically done individually for each plant through the manual application of different types of heuristics by experienced experts. Sometimes the scheduling procedure is split between the automation vendor and the plant operator. However, this manual heuristic procedure suffers from long response times, a lack of experienced experts and a considerable loss of optimization potential – in particular for plants processing many product types in small quantities.

To overcome the mentioned drawbacks of the manual heuristic procedure, automated MIP-based scheduling services are proposed. In order to exploit economy of scale effects, they should be offered by automation vendors. However, a key requirement for an automated scheduling procedure for a broad spectrum of problem instances is an efficient and widely automated engineering process. The engineering process is meant to adapt the generic scheduling procedure to individual problems.

## **4. Research gap**

This chapter outlines the gap between an automation vendor's requirements for an automated solution to schedule individual electroplating plants as described in chapter 3 and existing approaches.

#### *4.1. MIP-based solution*

Mathematical programming methods offer a high modeling flexibility and standard solvers that are independent of the model. The applicability of MIP methods to real-world electroplating scheduling problems depends on two prerequisites: The models need to represent the real-world problems sufficiently well and the computational limits of the standard solvers need to be overcome.

##### *4.1.1. Case studies*

Three real-world use cases were examined: (a) a medium-scale manually operated plant, (b) a small- and (c) a large-scale automatically operated plant. Their key properties can be adequately reflected by MIP models. To solve them by CPLEX, (a) was simplified by an aggregation approach and (b) and (c) by a product-decomposition approach. For (b), the optimized schedule was computed within 2 min CPU-time and with a 14,9 % shorter cycle time than the manually created schedule. Further it was shown that varying problem instances require different models and simplification heuristics. The model for (a) cannot be used for (b) and (c) because the hoists are not covered and depending on the problem size, certain combinations of simplifications are no longer sufficient to compute a schedule in reasonable time. For (c) for example, a schedule could be computed within 48 hours only for small recipes. The selection of a suitable model and suitable heuristics for a given problem instance is based on an educated guess.

##### *4.1.2. Literature analysis*

A complementary literature analysis revealed, that according to the classification scheme by Manier (2003), nearly the full spectrum of HSPs can be modelled as MIPs. However, an analysis of the papers also showed that different problems require different models and heuristics, as well as different parameterizations of the heuristics. A typical approach in the literature for selecting models and heuristics and parameterizing heuristics for MIP is also an educated guess.

#### *4.2. Automated model and heuristic selection*

The models existing in literature collectively represent almost the entire spectrum of the HSPs. Considered individually, the models as well as the heuristics are suitable for a small subset of problem instances and are usually applied to individual selected or benchmark instances. However, an automation vendor needs a solution that can be applied to a spectrum of real-world problem instances automatically. Which model and heuristics should be applied for a given problem instance to compute a solution in a short solution time with high quality is still not considered in the state of research and not yet automated. Manually selecting the models and the simplification heuristics for each instance takes much time and requires a lot of knowledge. From an automation vendor's perspective an automated engineering process i.e. an automated selection, parameterization and solution of models and heuristics for specific instances is crucial.

## **5. Approach**

Studies on real-world instances show that MIP can be used in practice for hoist scheduling in electroplating plants if simplification heuristics are applied. As described in section 4.2, an automated selection and parameterization is needed. According to literature the application of hyper-heuristics is limited to metaheuristics and constructive heuristics. Methods to select and parameterize simplification heuristics used to solve MIP models are not known to the authors. The authors propose to combine the principles of hyper-heuristics with MIP as follows. The real-world problem instances as well as the models are classified according to the scheme of Manier. Based on the classification, a preselection of models suitable for the problem is possible. The suitable models are

parameterized for the problem instance and a pool of suitable heuristics is chosen. The preselection of the models already decides which low-level heuristics are to be added to the pool, since only a subset of heuristics is applicable to each model. In the research phase, a parameterized model with a suitable heuristic and its parameterizations is to be determined with an evolutionary algorithm. A possible criterion to evaluate the fitness of individuals could be the objective value after a given solution time. This process is performed for multiple problem instances, resulting in pairs of scheduling problems with model-heuristic combinations. These will be used to train a neural network to avoid selecting a model and heuristic each time again with the evolutionary algorithm during the application phase.

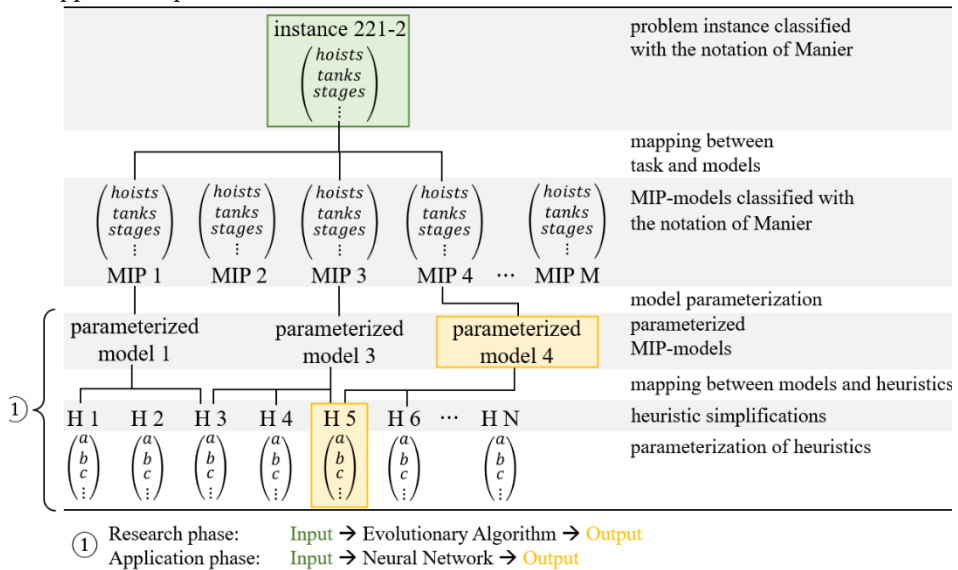


Figure 2 Procedure to select and parameterize a model and a heuristic for a scheduling problem

## 6. Outlook

Young research areas like hyper-heuristics deal with automated algorithm selection and parameterization but have not yet been applied to simplification heuristics for solving MIP models. The approach presented in this paper combines hyper-heuristics with MIP to compute good schedules in reasonable time for new, unknown problem instances. The study of this approach shall provide a deeper insight into the problem as well as knowledge about the relationships between the scheduling problem and suitable models-heuristic combinations. It is expected that the research will reveal the potentials of an intelligent, automated model and heuristic selection, regarding solution time and scheduling quality, compared to the educated guesses applied in literature and at the automation vendor. Furthermore, information on suitable structures for hyper-heuristics for selecting MIP models and heuristics as well as their essential differences compared to classical hyper-heuristics shall be obtained. The results can not only be transferred to further applications in the field of HSPs such as robotic cells or cluster tools, but also to optimization problems for which many models and heuristics exist which are differently suited for different problem instances. The expected overall goal is to bring the theoretical results in the form of MIP models and simplification heuristics into practice through the approach presented here.

## References

- A. E. Amraoui et al., 2013, A genetic algorithm approach for a single hoist scheduling problem with time windows constraints, *Engineering Applications of Artificial Intelligence*, Vol. 26, pp. 1761-1771, DOI: 10.1016/j.engappai.2013.02.004
- A. E. Amraoui et al., 2016, An efficient new heuristic for the hoist scheduling problem, *Computers & Operations Research*, Vol. 67, pp. 184-192, DOI: 10.1016/j.cor.2015.10.006
- E.K. Burke et al., 2019, A Classification of Hyper-Heuristic Approaches: Revisited, *International Series in Operations Research & Management Science*, Vol 272, pp. 453-477, DOI: 10.1007/978-3-319-91086-4\_14
- A. Che et al., 2007, Cyclic hoist scheduling in large real-life electroplating lines, *OR Spectrum*, Vol. 29, pp. 445-470, DOI: 10.1007/s00291-006-0040-9
- A. Che et al., 2010, Optimal cyclic scheduling of a hoist and multi-type parts with fixed processing times, *International Journal of Production Research*, Vol. 48, pp. 1225-1243, DOI: 10.1080/00207540802552659
- S. Chtourou et al., 2013, A hybrid algorithm for the cyclic hoist scheduling problem with two transportation resources, *Computers & Industrial Engineering*, Vol. 65, pp. 426-437, DOI: 10.1016/j.cie.2013.03.013
- J. Feng et al., 2018, Cyclic jobshop hoist scheduling with multi-capacity reentrant tanks and time-window constraints, *Computers & Industrial Engineering*, Vol. 120, pp. 382-391, DOI: 10.1016/j.cie.2018.04.046
- R. Fröhlich and S. C. Steneberg, 2009, Optimal Cyclic Multiple Hoist Scheduling for Processes with Loops and Parallel Resources, *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 293-298, DOI: 10.1109/ICSMC.2009.5346151
- E. Laajili et al., 2021, An Adapted Variable Neighborhood Search based algorithm for the cyclic multi-hoist design and scheduling problem, *Computers & Industrial Engineering*, Vol. 157, DOI: 10.1016/j.cie.2021.107225
- L. Lei and T. Wang, 1989, A Proof: The Cyclic Hoist Scheduling Problem is NP-Complete, Technical Report 89-0016, Graduate School of Management, Rutgers University
- J. M. Leung et al., 2004, Optimal Cyclic Multi-Hoist Scheduling: A Mixed Integer Programming Approach, *Operations Research*, Vol. 52, pp. 965-976, DOI: 10.1287/opre.1040.0144
- X. Li et al., 2015, Optimal multi-degree cyclic scheduling of multiple robots without overlapping in robotic flowshops with parallel machines, *Journal of Manufacturing Systems*, Vol. 36, pp. 62-75, DOI: 10.1016/j.jmsy.2015.03.003
- M. Manier, 2003, A Classification for Hoist Scheduling Problems, *International Journal of Flexible Manufacturing Systems*, Vol. 15, pp. 37-55, DOI: 10.1023/A:1023952906934
- Y. Mao et al., 2018, Mixed-integer linear programming method for multi-degree and multi-hoist cyclic scheduling with time windows, *Engineering Optimization*, Vol. 50, pp. 1978-1995, DOI: 10.1080/0305215X.2017.1418865
- H. J. Paul et al., 2007, A heuristic scheduling procedure for multi-item hoist production lines, *International Journal of Production Economics*, Vol. 105, pp. 54-69, DOI: 10.1016/j.ijpe.2005.11.008
- R. Reimschüssel et al., 2022, AALE Conference 2022, DOI: 10.33968/2022.06
- D. Ramin et al., 2022, Dynamic Hoist Scheduling for Multi-Recipe and Multi-Stage Production Lines: A Logical Framework, *SSRN Electronic Journal*, DOI: 10.2139/ssrn.4105670
- S.C. Steneberg, 2013, MILP model for multi-product, multi-part and multi-hoist cycle shops, 2013 IEEE International Conference on Industrial Technology (ICIT), pp. 1339-1346, DOI: 10.1109/ICIT.2013.6505867
- H. Wang et al., 2019, The Printed-Circuit-Board Electroplating Parallel-Tank Scheduling With Hoist and Group Constraints Using a Hybrid Guided Tabu Search Algorithm, *IEEE Access*, Vol. 7, DOI: 10.1109/ACCESS.2019.2915587
- P. Yan et al., 2018, A dynamic scheduling approach for optimizing the material handling operations in a robotic cell, *Computers & Operations Research*, Vol. 99, pp. 166-177, DOI: 10.1016/j.cor.2018.05.009

# Real time prediction of ozone pollution using data-enabled deep learning modeling

Fugui Hong, Cheng Ji, Fangyuan Ma, Chang Chen, Wei Sun\*

*College of Chemical Engineering, Beijing University of Chemical Technology, North Third Ring Road 15, Chaoyang District, Beijing, 100029, China*

*\*Corresponding Author's E-mail: [sunwei@mail.buct.edu.cn](mailto:sunwei@mail.buct.edu.cn)*

## Abstract

Tropospheric ozone pollution is one of the most harmful secondary air pollutants that could have brought challenges to air pollution prevention. Therefore, early prediction and warning of high-level tropospheric ozone concentration plays an important role in the management of ambient air pollution. Deep learning methods have been widely applied for their powerful data fitting ability, but nonstationary characteristic is exhibited in ozone concentration data owing to the instability of the atmosphere, which pose a challenge to the prediction of data with high-level ozone concentration. Considering the effectiveness of difference in handling nonstationary sequence, a nonstationary modeling method called single difference-embedded long short-term neural network (SDELSTM) is applied to realize the prediction of high-level tropospheric ozone concentration. Case study on an ozone dataset between 2013 and 2017 in Los Angeles is used to verify the effectiveness of the SDELSTM.

**Keywords:** SDELSTM, high-level tropospheric ozone concentration, nonstationary characteristics

## 1. Introduction

Tropospheric ozone pollution is a secondary pollutant produced by nitrogen oxides (NO<sub>x</sub>) and volatile organic compounds (VOCs) under solar radiation (Jenkin and Clemitshaw, 2000). The long-term exposure to tropospheric ozone contributes to the risk of respiratory and circulatory mortality, and is potentially associated with cardiovascular disease risk and premature death (Jerrett et al., 2009, Turner et al., 2016). The real time prediction of tropospheric ozone pollution can realize the early alarm and prevention of ozone pollution and reduce the risk of human health. Generally, data-driven methods are widely used to achieve this task. Among them, LSTM is widely used in ozone prediction since it could model long-term time dependence (Cabaneros et al., 2019). However, the air quality data show short-term nonstationary characteristics due to the unstable atmospheric composition and the complexity in the formation mechanism of pollutants. Under this circumstance, traditional data-driven models show poor performance in the field of ozone prediction. Zakaria et al. (Zakaria et al., 2021) introduced stationary and nonstationary model to forecast high ground-level concentrations and found that nonstationary models showed better performance in ozone prediction. Therefore, the consideration of nonstationary characteristics of data can enhance the performance of ozone prediction model. Difference is widely used in statistics to deal with nonstationary time series, which could eliminate the impact of random fluctuations in data, so as to improve the performance of the model. It is reported that several difference methods have been combined with deep learning methods for time series prediction. Zhou et al. (Zhou et al., 2022) proposed difference-LSTM (DLSTM) in which the difference of input variables was seen as a dynamic input, and the method was verified by the prediction of key variables in complex industrial process. Wang et al. (Wang et al., 2019) proposed the memory in memory (MIM) networks. Then the MIM module was utilized to replace the



forgotten gate and retain more nonstationary characteristics. In this study, a model called SDELSTM is introduced to perform ozone prediction from the perspective of nonstationary characteristics of data. In this model, an additional network layer including a differential module is embedded into LSTM neural network to capture the nonstationary characteristics of data under the premise of avoiding the loss of original data information (Ji et al., 2023). The effectiveness of the SDELSTM model is validated by performing ozone prediction using the historical tropospheric data from Los Angeles.

## 2. Description of the model

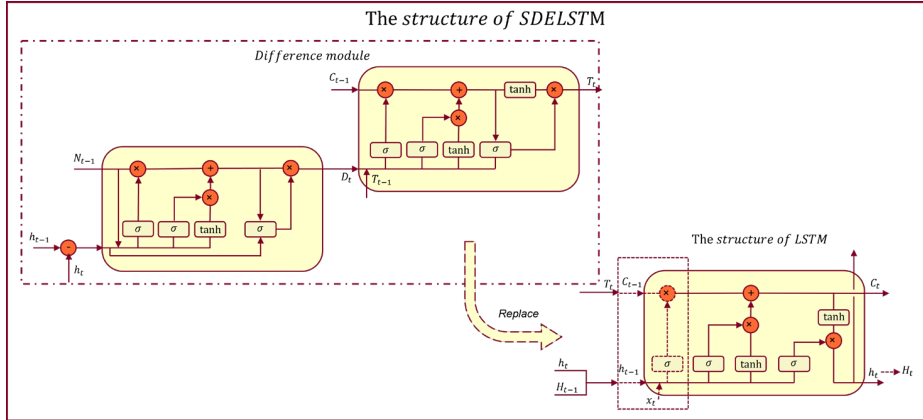


Figure 1 The structure of LSTM and SDELSTM

### 2.1. long short-term memory neural network

The LSTM neural network is a special recurrent neural network which contains many neurons in the hidden layer. The structure of LSTM is shown in Figure 1. The initial input  $h_{t-1}$  and  $x_t$  represent the state of the previous hidden layer and input of current time respectively.  $C_t$  and  $C_{t-1}$  record the cell state which is the key value to LSTM.  $h_t$  represents the output of the current layer. The neurons in hidden layer contains three gates including a forget gate, an input gate and an output gate which could control the transmission of information. The unique structure helps LSTM learn the long-term characteristics of the data. The forget gate control which part of information should be retained. Then the input gate determines the information to be added to cell state and the output gate control the output of the cell state. However, the short-term nonstationary characteristics cannot be obtained by a single forget gate and some measures have been adopted to make LSTM become nonstationary model.

### 2.2. The structure of single difference-embedded long short-term neural network

To address this problem, the SDELSTM model is improved from LSTM where an additional layer including a differential module is embedded into the hidden layer. The dashed box indicates the replacement of the forget gate by the differential module. The input of difference module is states of hidden layers of the current and previous moment ( $h_t$  and  $h_{t-1}$ ) which is regarded as difference information.  $N_{t-1}$  represents cell state of the previous difference layer. The output of difference layer is a new cell state  $D_t$  which contains difference information. The next layer combines  $D_t$  and  $C_{t-1}$  whose purpose is selective retain difference characteristics and long-term characteristics of the original sequence.  $T_t$  represents cell state which records both short-term and long-term information. Eventually, in order to avoid the loss of original information,  $h_t$  and  $H_{t-1}$  which represents the state of SDELSTM are utilized as the input of LSTM neural network

and  $H_t$  represents the output of SDELSTM. Then, the short-term nonstationary characteristics can be obtained by the unique forget gate module. In summary, the proposed SDELSTM model could learn short-term nonstationary characteristics and long-term dependence at the same time by a difference module, and therefore the performance of ozone concentration prediction can be improved.

### 2.3. Implementation procedure of SDELSTM

The implementation procedure of SDELSTM is shown in Figure 2 which contains three parts: data processing, stationary analysis and modelling. The procedures can be summarized as follows,

Step 1: The missing values are filled by K-Nearest Neighbor (KNN).

Step 2: The datasets are divided into training data and test data in a ratio of 4:1 and 20 percent of training data are grouped into validation data.

Step 3: The stationarity of time series is tested by the ADF test.

Step 4: The SDELSTM model is set up by training data and the hyperparameters of the model are determined by validation data.

Step 5: The performance of the model is verified by test data.

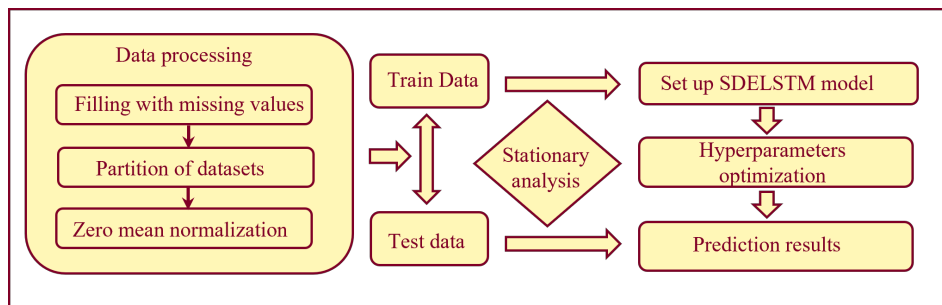


Figure 2. The training process of SDELSTM model

## 3. Results and discussions

### 3.1. Data preparation

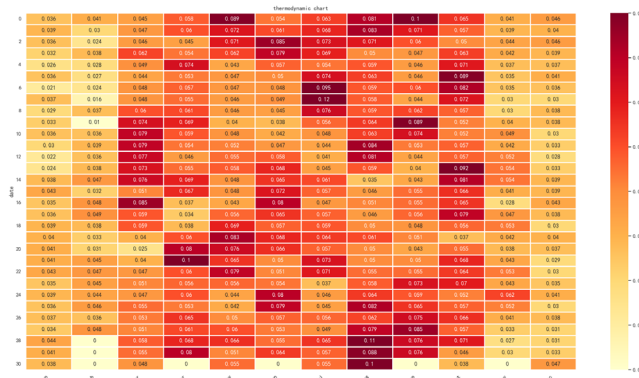


Figure 3 Daily Maximum ozone distribution thermal map of Los Angeles in 2017  
 Due to the special geographical location of Los Angeles and the year-round high temperature and little rain characteristics, ozone pollution is serious. In 2017, more than 100 million people in the United States were exposed to tropospheric ozone pollution,

especially in Los Angeles (Zhang et al., 2019). Figure 3 shows that the thermal map of daily maximum ozone of Los Angeles in 2017. The national ambient air quality standards from EPA for ozone concentration is 0.070 parts per million (ppm). In addition, according to world health organization (WHO) air quality guidelines, the standard for ozone concentration is 0.05ppm which suggests more people are exposed to tropospheric ozone pollution. It is found that the darker the color, the more ozone pollution there is and most ozone pollution is concentrated during the warmer months. Besides the dataset mentioned above, the Los Angeles air quality datasets between 2013 and 2017 are obtained from the air quality system of environment protection agency (EPA) which were collected from the monitoring site at N main St in Los Angeles. Datasets from 2013 to 2016 are used as training sets and the data of 2017 is selected as the test set. Considering that there are missing values in the data, KNN is utilized to fill missing data. After data preprocessing, twelve air monitoring variables are selected, including SO<sub>2</sub>, NO<sub>2</sub>, PM<sub>2.5</sub>, PM<sub>10</sub>, RH\_DP, TEMP, CO<sub>0</sub>, CO<sub>1</sub>, NO<sub>x</sub>, NO<sub>y</sub>, NO<sub>0</sub>, and NO<sub>1</sub>.

### 3.2. Case study

#### 3.2.1. Stationary analysis

Since the complexity of atmospheric environment, the air quality monitoring data obtained from monitoring stations would fluctuate randomly within a certain range, which is easy to be ignored from a long-term perspective, and the nonstationary characteristics generated by random fluctuations will affect the performance of ozone prediction model. The Augmented Dickey-Fuller (ADF) test was widely used to test the stationary of the time series and conducted for monthly tropospheric ozone concentration to demonstrate the non-stationary characteristics of the data. It is shown in Table 1 that 8 months ozone concentration data are nonstationary at 1% confidence. Therefore, the nonstationary characteristics of the data cannot be ignored while traditional LSTM models do not take these characteristics into account. To address this issue, the SDELSTM model is introduced to learn the nonstationary characteristics within the air quality data, which could enhance the performance of ozone prediction.

Table 1 Los Angeles monthly data ADF statistics in 2017

Months	Test statistics	1%	5%	10%
January	<b>-2.97</b>	-3.71	-2.98	-2.63
February	<b>-2.44</b>	-3.66	-2.96	-2.62
March	-4.79	-3.67	-2.96	-2.62
April	-4.97	-3.78	-3.01	-2.64
May	<b>-3.34</b>	-3.66	-2.96	-2.62
June	<b>-2.53</b>	-3.66	-2.96	-2.62
July	-4.07	-3.75	-2.99	-2.63
August	-4.92	-3.67	-2.96	-2.62
September	<b>-3.27</b>	-3.66	-2.96	-2.62
October	<b>-3.43</b>	-3.66	-2.96	-2.62
November	<b>-1.75</b>	-3.71	-2.98	-2.63
December	<b>-2.99</b>	-3.66	-2.96	-2.62

#### 3.2.2. Performance of SDELSTM on ozone concentration

With the current air quality data fed in, the model is able to predict ozone concentration for the next hour. It is shown in Figure 4 that the prediction results of the model for different months. The model makes accurate predictions of the concentration data where there is a health risk while the model performed poorly in the prediction of low concentration data. It is found that ozone data show successive near-zero values at regular

intervals after the analysis of actual hourly ozone data. Then we speculate that it is caused by artificial manipulation which leads to the deviation in the predicted result graph. However, only high concentrations of ozone (>0.05ppm) pose a huge threaten to human health. Therefore, when evaluating the predictive performance of the model, the prediction results of low concentration values will not greatly affect the predictive performance of the model.

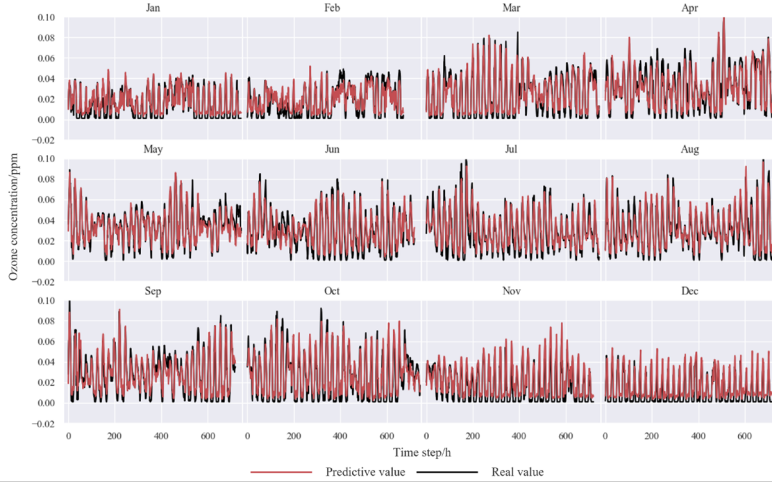


Figure 4 Comparison of true and predicted values for different months

In order to evaluate the prediction effect of the model more scientifically and objectively, it is necessary to use quantitative indicators to further evaluate the model, coefficient ( $R^2$ ), root mean square error (RMSE) are widely used to evaluate the performance of prediction models. The indicators are defined as follows,

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_j - \hat{y}_j)^2} \quad (1)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_j - \hat{y}_j)^2}{\sum_{i=1}^N (y_j - \bar{y}_j)^2} \quad (2)$$

where  $N$  is the number of test samples.  $y_j$  and  $\hat{y}_j$  indicate the actual and predicted values at time  $j$ .  $\bar{y}_j$  indicates the mean of the predicted values over all times. The specific calculation results are presented in the next section along with the comparison of other models.

### 3.2.3. Model comparison

In order to demonstrated the performance of the SDELSTM model in ozone prediction, the related models including artificial neural network (ANN), LSTM, gate recurrent unit

(GRU) are utilized to compare the results with the SDELSTM model. The specific model comparison results are shown in Table 3. It is shown in Table 2 that the SDELSTM model performs better than other models in all indicators. The  $R^2$  of the model is 13.4% higher than that of the traditional LSTM and RMSE is 34.6% lower than other models. The SDELSTM model shows more accuracy on the tropospheric ozone concentration because nonstationary characteristics of the ozone data cannot be considered by traditional predictive models.

Table 2 Comparison of different models

	RMSE	$R^2$
ANN	0.0120	0.6119
GRU	0.0110	0.6869
LSTM	0.0104	0.7265
SDELSTM	0.0068	0.8237

#### 4. Conclusions

In this work, a model called SDELSTM is introduced to perform ozone prediction from the perspective of nonstationary characteristics of data. An additional network layer including a differential module is embedded into LSTM neural network to capture the nonstationary characteristics of data under the premise of avoiding the loss of original data information. Based on literature research, the ADF stability test was conducted on the ozone data of Los Angeles in 2017, which confirmed the nonstationary characteristics of the ozone data. Then the effectiveness of SDELSTM model is validated by performing ozone prediction using the historical tropospheric data from Los Angeles. The results demonstrates that the SDELSTM model shows better performance in hourly troposphere ozone prediction than other related models, indicating its application prospect in preventing ozone pollution.

#### References

- Cabaneros, S. M., Calautit, J. K. and Hughes, B. R., 2019. A review of artificial neural network models for ambient air pollution prediction. *Environmental Modelling & Software*, 119 285-304.
- Jenkin, M. E. and Clemitshaw, K. C., 2000. Ozone and other secondary photochemical pollutants: chemical processes governing their formation in the planetary boundary layer. *Atmospheric Environment*, 34 (16), 2499-2527.
- Jerrett, M., Burnett, R. T., Pope III, C. A., Ito, K., Thurston, G., Krewski, D., Shi, Y., Calle, E. and Thun, M., 2009. Long-term ozone exposure and mortality. *New England Journal of Medicine*, 360 (11), 1085-1095.
- Ji, C., Ma, F., Wang, J. and Sun, W., 2023. Profitability related industrial-scale batch processes monitoring via deep learning based soft sensor development. *Computers & Chemical Engineering*, 170 108125.
- Turner, M. C., Jerrett, M., Pope, C. A., 3rd, Krewski, D., Gapstur, S. M., Diver, W. R., Beckerman, B. S., Marshall, J. D., Su, J., Crouse, D. L. and Burnett, R. T., 2016. Long-Term Ozone Exposure and Mortality in a Large Prospective Study. *Am J Respir Crit Care Med*, 193 (10), 1134-1142. 10.1164/rccm.201508-1633OC.
- Wang, Y., Zhang, J., Zhu, H., Long, M., Wang, J. and Yu, P. S., 2019. Memory in memory: A predictive neural network for learning higher-order non-stationarity from spatiotemporal

dynamics. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,

Zakaria, S. A., Amin, N. A. M., Radi, N. F. A. and Hamidin, N., 2021. Stationary and Non-Stationary Models of Extreme Ground-Level Ozone in Peninsular Malaysia. *Mathematics and Statistics*, 9 (3), 357-370. 10.13189/ms.2021.090318.

Zhang, J., Wei, Y. and Fang, Z., 2019. Ozone pollution: a major health hazard worldwide. *Frontiers in immunology*, 10 2518.

Zhou, J., Wang, X., Yang, C. and Xiong, W., 2022. A Novel Soft Sensor Modeling Approach Based on Difference-LSTM for Complex Industrial Process. *IEEE Transactions on Industrial Informatics*, 18 (5), 2955-2964. 10.1109/tii.2021.3110507.



# Numerical Study of a Single-Spout and Z-shaped Fluidized Bed Biomass Gasifier

Fatima Ez-zahra El Hamra,<sup>a</sup> and Radouan Boukharfane<sup>a,\*</sup>

<sup>a</sup>*Mohammed VI Polytechnic University (UM6P), MSDA Group, Benguerir, Morocco  
radouan.boukharfane@um6p.ma*

## Abstract

Gas-solid multiphase systems are ubiquitous in several applications such as pharmaceutical, mining, and petrochemical industry (Geldart, 1986). The vigorous interaction between gas and solid particles involves complex dynamics in the determination of mass and heat transfer. Most simulations and experiments use fluidized beds with simple geometries, but industrial multiphase reactors generally have complex configurations (Zhao et al., 2020). The aim of this work is to study the gas–solid flow characteristics of a Z-shaped gas-solid fluidized bed by means of an in-house solver DEM-LES solver. The fluid phase is computed through solving the volume-averaged four-way coupling Navier–Stokes equations in which the Smagorinsky sub-grid scale tensor model is used. Results show that particles and fluid phases are not uniformly distributed, and a core-annulus structure is observed. These findings provide valuable insight into the development and optimization of particle purification appliances in multiphase reactors.

**Keywords:** Fluidized bed, Computational fluid dynamics, DEM-LES, Biomass particles, Particle sorting.

## 1. Introduction

Finding sustainable, clean, and low-carbon alternatives to traditional fossil fuels is becoming crucial given the world's rising energy needs and the severity of global warming. The many inherent benefits of biomass for bioenergy, such as its global availability, renewability, and carbon neutrality, make it one of the most attractive possibilities. Torrefaction, pyrolysis, gasification, and catalytic conversion are among the thermochemical conversion processes that may turn raw biomass feedstocks into high-quality energy products. Because it can transform biomass into three different products-biogas, bio-oil, and biochar, pyrolysis has drawn the most attention among these processes. The fluidized-bed pyrolyzer is one of the most popular reactors for biomass pyrolysis because of its excellent particle mixing capability, high heat transfer efficiency, and operational flexibility. The fluidized-bed reactor, however, typically presents a complicated system in which the hydrodynamics of dense gas-solid flow, heat/mass transfer, chemical reactions, and turbulence are intimately inter-coupled, making it challenging to fully comprehend the mechanics of biomass pyrolysis.

In order to develop and scale-up fluidized systems, it is necessary to understand the flow characteristics between particles and gas but doing so through experiments is challenging. The usage of numerical simulation as an alternate tool for engineering analysis has increased with the advancement of numerical methods and processing power.



Particle-resolved direct numerical simulation (PR-DNS), two-fluid model (TFM), and CFD-DEM are the three categories into which numerical methods of particle fluidization can be categorized. PR-DNS is only useful for flow systems with no more than 1000 particles due to its high computing requirements. The TFM technique treats fluid and solid as two distinct continuous phases. The CFD-DEM model uses Newton's second law to monitor particles, which is why a good match was obtained between simulations and experiments.

Numerous studies have concentrated on how operating parameters affect the particle flow characteristics. Ejim et al., (2017) investigated how the inner geometry of mesoscale tubes with baffles affected the axial dispersion and suspension of materials. Hu et al. (2019) assessed how collisional characteristics affected the hydrodynamics and heat transport in a spouted bed. In a full-loop circulating fluidized bed at varied surface gas velocities, Xu et al. (2018) computationally and empirically examined the particle movements, proving that solid behaviors (such as residence duration, dispersion, and mixing) are connected to the fluidization regions. Particle size and gas velocity effects on tar yield in rapid pyrolysis reactors were discussed by Xue et al. (2012). Wang et al. (2018) also thoroughly investigated the particle behaviors, including the particle force and velocity. The effectiveness of chemical reactions is impacted by the fact that biomass is made up of heterogeneous particles of various densities, sizes, and forms. Consequently, a number of separation and purification techniques were suggested.

Konrath et al. (2014) investigated the circumstances of fine particle separation in centrifugal classification. The influence of particle organization and spigot widths on the enhancement of separation of the hydrocyclone was numerically modeled by Vega-Garcia et al. (2020). Further discussion was made regarding the impact of dipleg shape on the functionality of gas-solid cyclones.

Lv et al. (2018) investigated the motion and separation behavior of coal in a fluidized bed for gas-solid separation.

Since most previous research places emphasis on the effects of operating conditions but gives the reactor structure relatively little attention, the current work aims to investigate the influence of the reactor structure on biomass fluidized-bed pyrolysis behavior. The separation of two-component suspension has typically been shown to be improved by the fluidized bed structure's inclination (Masliyah et al., 1989). In the present study, multiple inclination cross-sections that make up a Z--shaped channel, which has considerable potential for the fluidization and separation of mixed biomass particles, is investigated.

## 2. Mathematical modeling

The single-spout fluidized bed is simulated in the current study using a coupled Computational Fluid Dynamics and Discrete Element Method (CFD-DEM). The Newtonian laws of motion control the dynamics of the particles suspended in the bed. To describe the fluid phase, the locally volume-filtered Navier-Stokes equations are solved. The whole momentum exchange between the solid and fluid phases is then included in these sets of equations (two-way coupling).

$$\left\{ \begin{aligned} \partial_t(\theta_f \bar{\rho}_f) + \partial_i(\theta_f \bar{\rho}_f \tilde{u}_{f,i}) &= 0, \end{aligned} \right. \quad (1a)$$

$$\left\{ \begin{aligned} \partial_t(\theta_f \bar{\rho}_f \tilde{u}_{f,i}) + \partial_j(\theta_f \bar{\rho}_f \tilde{u}_{f,i} \tilde{u}_{f,j}) &= -\partial_j \bar{p} + \partial_j(\bar{\tau}_{ij} + \bar{\tau}_{ij}^{SGS}) + \theta_f \bar{\rho}_f g_i + \mathcal{F}_i^{inter}, \end{aligned} \right. \quad (1b)$$

Where  $\theta_f$ ,  $\rho_f$ ,  $p$ , and  $u_f$  are the fluid-phase volume fraction, density, dynamic pressure, and velocity, respectively. In Eq.(1b),  $g$  stands for the acceleration due to gravity and  $F^{inter}$  for the inter-phase exchange term that arises from filtering the divergence of the stress tensor. The volume-filtered stress tensor,  $\tau$ , is expressed as  $\overline{\tau_{ij}} = \mu \left[ \partial_i \overline{u_{f,j}} + \partial_j \overline{u_{f,i}} - \frac{2}{3} \partial_i \overline{u_{f,i}} \delta_{ij} \right] + \mathcal{R}_{\mu,ij}$ , where  $\mu$  is the dynamic viscosity and  $\delta_{ij}$  is the Kronecker function. The SGS stress tensor  $\overline{\tau_{ij}}^{SGS} = \overline{\rho_f u_{f,i} u_{f,j}} - \overline{\rho_f} \widetilde{u_{f,i}} \widetilde{u_{f,j}}$  is modeled within the Boussinesq's framework evaluated using the classical Smagorinsky (Smagorinsky, 1963) model.

A Lagrangian particle-tracking approach is used for the solid phase. The displacement of an individual solid particle indicated by the subscript  $p$  is calculated using Newton-Euler equations for rigid body motion as follows:

$$\left\{ m_p \frac{du_{p,i}}{dt} = f_{p,i}^{inter} + f_{p,i}^{col} + m_p g_i, \text{ with } \frac{dx_{p,i}}{dt} = u_{p,i} \right\} \quad (2a)$$

$$\left\{ \mathcal{J}_p \frac{d\omega_{p,i}}{dt} = \mathcal{M}_{p,i}^{drag} + \mathcal{M}_{p,i}^{col} \right\} \quad (2b)$$

where the particle mass is defined by  $m_p = \pi \rho_p d_p^3 / 6$  and  $u_p$  is the centroid particle velocity. The force  $f_p^{col}$  accounts for the particle-particle and particle-wall repulsion modeled using a soft-sphere model (Cundall & Strack, 1979). The force  $f_p^{inter}$  exerted on a single particle  $p$  by the surrounding fluid is related to the inter-phase exchange term in Eq.(1b) by  $\mathcal{F}^{inter} = \sum_{p=1}^{N_p} \xi(|x - x_p|) f_p^{inter}$ , where  $N_p$  is the total number of particles,  $\xi$  is the filtering kernel used to volume filter the Navier-Stokes equations,  $x_p$  is the position of the  $p^{th}$  particle, and  $f_p^{inter}$  is approximated by  $f_{p,i}^{inter} \approx \mathcal{V}_p \partial_j \tau_{ij} + f_{p,i}^{drag} \approx -\mathcal{V}_p \partial_i p^\circ + f_{p,i}^{drag}$ , where  $\mathcal{V}_p$  is the volume of the  $p^{th}$  particle given by  $\pi d_p^3 / 6$  and  $\partial_i p^\circ$  is the local pressure gradient interpolated at the center of the particle. The drag force is computed using the drag force coefficient of Tennesi et al. (2011). Particle-particle and particle-wall collisions are modeled using the adaptive collision time model ACTM proposed by Kempe & Fröhlich (2012), which is a variation of the classical soft-sphere model. In this work, a fourth-order central scheme is used for the spatial integration, and a third-order accurate semi-implicit Crank-Nicolson scheme is employed for time integration. The reader is referred to (el Hamra et al., 2022) for further information on the current CFD-DEM solver as well as to review the numerical validation.

### 3. Single-spout and Z-shaped fluidized bed system

In our study, three spouted-fluidized beds with a single orifice with the same cross-section are considered as schematically depicted in Fig. 1. The difference between the second and third configuration is the angle between the two sections, which is equal to  $155^\circ$  in the Case 2, while it is equal to  $132^\circ$  in the Case 3. In all base case simulations carried out for this work, the spouted-fluidized bed height (BH) and depth are 400 mm and 14 mm, respectively. There are two portions

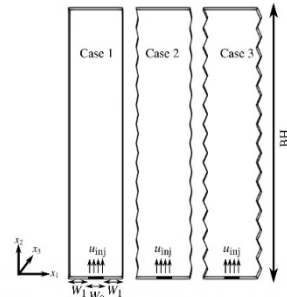


Figure 1: Different geometric models of Z-shaped particle fluidized system.

across the bed's width (shown as  $W_0 = 0.007$  mm and  $W_1 = 0.007$  mm in Fig. 1).

The computational domain is divided into  $82 \times 12 \times 370$  quad grids. In the spouted region, 12 grids are considered. Gas flows upwards through  $W_0$ , which induces three distinct regions: a dilute central core, where the solid phase is entrained in the concurrent jetting gas flow, a fountain region in the uppermost part, and a dense annular region with a counter-current flow between both phases. Simulation settings of the validation cases are summarized in Table 1.

Fluid phase	
Fluid density $\rho_f$ [kg/m <sup>3</sup> ]	1.165
Fluid viscosity $\mu$ [Pa · s]	$2.0 \times 10^{-5}$
Superficial mass flow rate at the injection [kg/s]	$2.0 \times 10^{-3}$
Solid phase	
Number of particles	93808
Bed mass $m_{bed}$ [kg]	0.122
Particle diameter $d_p$ [mm]	1.0
Particle density $\rho_p$ [kg/m <sup>3</sup> ]	2500
Inter-particle & Particle-wall restitution coefficient	0.97
Inter-particle & Particle-wall friction coefficient	0.1

Table 1: Simulation settings of the validation case for the three spouted bubbling fluidized configurations.

A typical illustration of particle velocity distribution for the three cases at  $t=25$  s can be seen in Fig. 2, where the mesoscopic bubble formed in the mixture with different sizes is observed. During this process, the particles are impulsed into the bubbles and scattered to the freeboard and fall back to bed due to the gravitational force. As the angle between the two sections in the Z-shaped fluidized bed, the bed height is higher. A higher bed expansion rate and an intense particle motion is therefore caused by applying the change of lateral wall's shape compared to the flat wall case.

The time-averaged fluid volume fraction and gas velocity at the midplane of the fluidized bed is portrayed in Fig. 3 for three configurations.

A large velocity gradient is observed around the irregular wall. The Z-shaped fluidized bed breaks the flow symmetry, thus allowing the mixed particles to achieve effective screening in each section. Furthermore, in the corners, heaps of unfluidized particles accumulate in much larger quantities in the case of flat walls. While less unfluidized zones are obtained when the angle between the two sections of the Z-shaped fluidized bed gets smaller. Note that due to these corners actually act as dead zones and create more resistance to bubble movement.

Because gas and solids transfer mass at slower rates in dead zones than they do in a fully fluidized bed, mixing rates may be lowered as a result. Dead zones in corners thereby reduce the cross section available for bubble passage, which in turn reduces bubble size and/or speed.

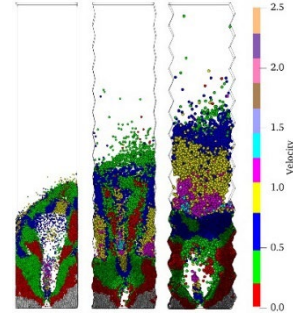


Figure 2: Instantaneous distribution characteristic of particles colored by particle velocity in the three cases.

The mixing rate in case 1 bed is lower than in case 3 bed because dead zone particles move into the active zone of the bed extremely slowly. In other words, Z-shaped fluidized beds remove dead zones and enhance the movement, homogeneous mixing, and subsequent fluidization of the solids.

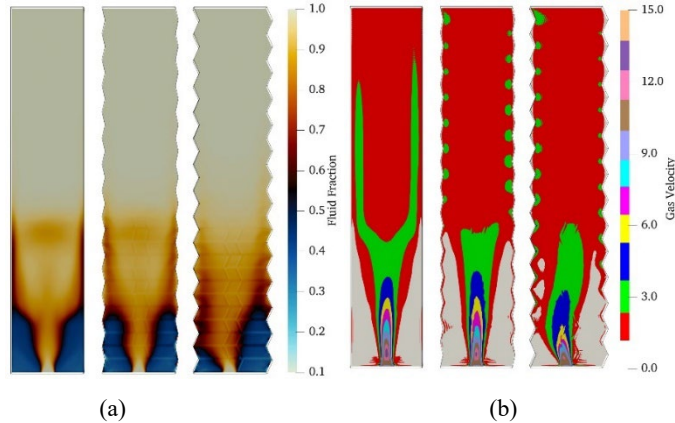


Figure 3: Time-averaged (a) fluid volume fraction and (b) gas velocity at the midplane for the three studied cases.

Solid flow rate is a key parameter for the operation and optimization of fluidized beds. To quantify this parameter, the  $x_2$ -direction of the bed, and mainly the interval  $x_2 \in [0, 0.35]$  m, is subdivided into seven equal sections, and the mean number of particles within each section is computed for  $t \in [5, 25]$  s as can be seen in Fig. 4. The section #1 is the lowest one (at the injection) and the section #7 is the highest one. It is confirmed that the Z-shaped beds tend to provide sufficient space compared to flat walls for the separation and purification of particles as they occupy more space in the  $x_2$ -direction.

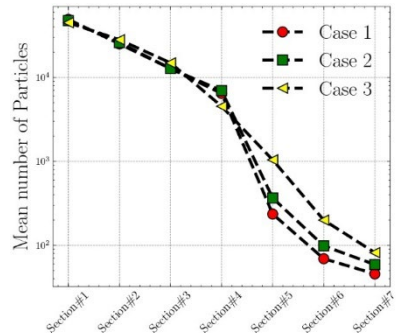


Figure 4: Mean number of particles in different sections for three studied cases.

Due to collisions and energy loss, many particles congregate near the corners of flat walls. However, in cases 2 and 3, where zigzag patterns are present, larger flocules form, break, and then slide down the wall.

#### 4. Conclusion

This work presented a preliminary attempt to use single-spout and Z-shaped fluidized beds as fluidized bed gasifiers of mixed biomass particles. A numerical study based on a large-eddy Simulation (LES) solver coupled with Discrete Element Method (DEM) was employed to study the effect of zigzagging the lateral walls with two angles. The Z-shaped fluidized bed was found to modify the uniformity of the flow domain and to allow an enhanced particle's mixing. These results offer an important direction for

designing clean particle products and can be expanded to include research into environmental chemical and process engineering.

**Acknowledgement:** This work was supported by OCP Group (Morocco). The authors gratefully acknowledge the support and computing resources from the African Supercomputing Center (ASCC) and SIMLAB HPC center at UM6P (Morocco).

## References

- Cundall, P. A., & Strack, O. D. L. (1979). A discrete numerical model for granular assemblies. *Geotechnique*, 29(1), 47–65.
- Ejimi, L. N., Yerdelen, S., McGlone, T., Onyemelukwe, I., Johnston, B., Florence, A. J., & Reis, N. M. (2017). A factorial approach to understanding the effect of inner geometry of baffled meso-scale tubes on solids suspension and axial dispersion in continuous, oscillatory liquid–solid plug flows. *Chemical Engineering Journal*, 308, 669–682.
- el Hamra, F., Boukharfane, R., Er-raiy, A., & Chakraborty, N. (2022). Development and assessment of algorithms for DEM-LES simulations of fluidized bed. *Submitted*.
- Geldart, D. (1986). *Gas fluidization technology*.
- Hu, C., Luo, K., Wang, S., Junjie, L., & Fan, J. (2019). The effects of collisional parameters on the hydrodynamics and heat transfer in spouted bed: A CFD-DEM study. *Powder Technology*, 353, 132–144.
- Kempe, T., & Fröhlich, J. (2012). Collision modelling for the interface-resolved simulation of spherical particles in viscous fluids. *Journal of Fluid Mechanics*, 709, 445–489.
- Konrath, M., Hackbarth, M., & Nirschl, H. (2014). Process monitoring and control for constant separation conditions in centrifugal classification of fine particles. *Advanced Powder Technology*, 25(3), 991–998.
- Lv, B., Luo, Z., Zhang, B., Qin, X., & Zhu, C. (2018). Particle motion and separation behavior of coal in gas–solid separation fluidized bed. *Powder Technology*, 339, 344–353.
- Masliyah, J. H., Nasr-El-Din, H., & Nandakumar, K. (1989). Continuous separation of bidisperse suspensions in inclined channels. *International Journal of Multiphase Flow*, 15(5), 815–829.
- Smagorinsky, J. (1963). General circulation experiments with the primitive equations: I. The basic experiment. *Monthly Weather Review*, 91(3), 99–164.
- Tenneti, S., Garg, R., & Subramaniam, S. (2011). Drag law for monodisperse gas–solid systems using particle-resolved direct numerical simulation of flow past fixed assemblies of spheres. *International Journal of Multiphase Flow*, 37(9), 1072–1092.
- Vega-Garcia, D., Cilliers, J. J., & Brito-Parada, P. R. (2020). CFD modelling of particle classification in mini-hydrocyclones. *Separation and Purification Technology*, 251, 117253.
- Wang, S., Luo, K., Hu, C., Sun, L., & Fan, J. (2018). Effect of superficial gas velocity on solid behaviors in a full-loop CFB. *Powder Technology*, 333, 91–105.
- Xu, J., Lu, X., Zhang, W., Chen, J., Wang, Q., Chen, Y., & Guo, Q. (2018). Effects of superficial gas velocity and static bed height on gas-solid flow characteristics in a 60-meter-high transparent CFB riser. *Chemical Engineering Journal*, 334, 545–557.
- Xue, Q., Dalluge, D., Heindel, T. J., Fox, R. O., & Brown, R. C. (2012). Experimental validation and CFD modeling study of biomass fast pyrolysis in fluidized-bed reactors. *Fuel*, 97, 757–769.
- Zhao, P., Xu, J., Ge, W., & Wang, J. (2020). A CFD-DEM-IBM method for Cartesian grid simulation of gas-solid flow in complex geometries. *Chemical Engineering Journal*, 389, 124343.

# Intensification of ATJ Process Using Catalytic Distillation, Optimization Considering Control and Economic Issues

Gabriel Contreras-Zarazúa<sup>1, a, b</sup>, Juan José Quiroz Ramírez,<sup>a</sup> Eduardo Sánchez-Ramírez,<sup>b</sup> Juan Gabriel Segovia-Hernández<sup>b</sup>

<sup>a</sup>*CONACyT-CIATEC A.C Center of Applied Innovation in Competitive Technologies, Omega 201, Industrial Delta, 37545 León, Gto., Mexico.*

<sup>b</sup>*Department of Chemical Engineering University of Guanajuato, Noria Alta S/N, Noria Alta, 36000, Guanajuato, Gto., Mexico.*

## Abstract

In this work the design and simultaneous optimization of an intensified process for jet fuel production by the ATJ process was developed. The intensified process consists of a catalytic reactive distillation column. The process integrates the oligomerization, hydrogenation and purification stages in a single unit. The design and simultaneous optimization process were performed using the differential evolution with tabu list (DETL) algorithm. Total annual cost and condition number were considered as objective functions to evaluate cost issues and control properties, respectively. The optimized process was contrasted against the original intensified process. The results indicate that the optimized process reduces energy consumption by 30% and total annual cost by up to 80%, due to reductions in energy consumption and equipment size. Regarding the control properties, it is observed that the optimized process has a condition number of 82.52 against 915,000 of the original process. This means that the optimized process withstands disturbances better, which facilitates control and ensures the quality of the bioturbosine in a more reliable way.

**Keywords:** Biojet-fuel, reactive distillation, process intensification.

## 1. Introduction

The world is currently facing increasing energy demands. This growth in energy is associated with an increase in the production of goods and services, increase in the quality of life, among others. Unfortunately, most of the energy produced by humans derives from the burning of fossil fuels such as coal, natural gas, gasoline, etc., which generates severe environmental problems by the emission of greenhouse gases such as CO<sub>2</sub>. Carbon dioxide is the main greenhouse gas, recent studies reveal that by 2018 its concentration in the atmosphere had gone from 280 to 408ppm, taking the year 1900 as a reference. This increase means a 0.8°C increase in the earth's global temperature and it is expected that by the year 2100 the temperature increase will reach 5°C (Bains et al., 2017).

In order to reduce their emissions, various industries such as the aviation sector have started to develop technologies to reduce their CO<sub>2</sub> emissions. The aircraft industry is responsible for 2.6% of CO<sub>2</sub> emissions and it is expected that by mid-century this industry will emit 20.2% of CO<sub>2</sub> emissions. An interesting alternative to reduce polluting emissions is to replace jet fuel (also called biojet-fuel), which comes from fossil sources,

with biojet-fuel, which is a renewable fuel derived from biomass residues, vegetable oils, etc. Several processes and biomasses have been proposed to produce biojet-fuel, however, a little explored alternative is the ATJ (Alcohol to Jet fuel) process, which consists of converting lignocellulosic residues into sugars, which are transformed into hydrocarbons (Wang and Tao, 2016). As with other alternatives to produce biojet-fuel or even other biofuels, economic viability remains the main challenge for this type of technology. In the case of the ATJ process, a critical and energy-intensive stage is the oligomerization and hydrogenation stage, since it is at this stage that hydrocarbons are produced and separated. In this sense, process intensification coupled with rigorous optimization techniques can help to reduce energy consumption and process operating costs, in order to increase the profitability and adoption of these processes. In this work, an intensified process based on catalytic reactive distillation is proposed to replace the current oligomerization and hydrogenation steps in the ATJ process. This column is simultaneously designed and optimized using the differential evolution stochastic optimization method with tabu list (DETL). The minimization of the total annual cost is proposed as the economic objective function. On the other hand, because the catalytic column is notably more complex and in order to ensure a viable operation of this equipment, the minimization of the condition number is proposed as the second objective function in order to evaluate the control properties.

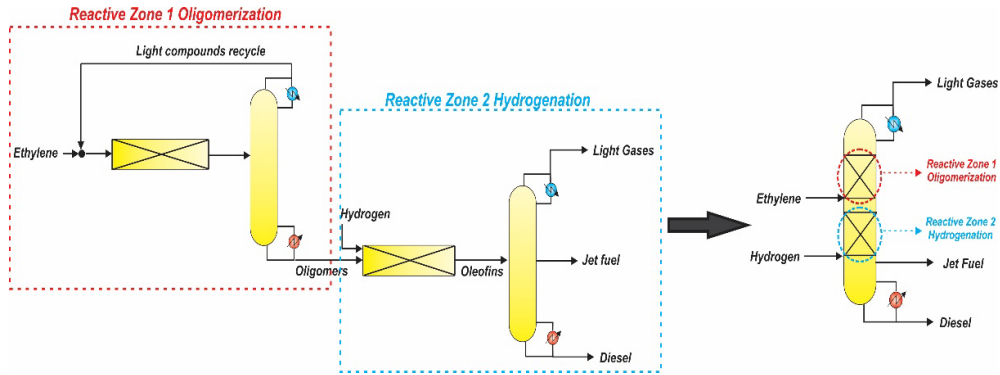
## 2. Methodology

First, the intensified catalytic reactive distillation process is generated from the conventional process. As shown in Figure 1, the conventional process consists of 2 reactive zones, the first one is the oligomerization stage, where ethylene is fed to a fixed bed reactor, then the reactor products which are light compounds (C2- C5) and oligomers (C6-C17) are fed to a distillation column, where the light compounds are recirculated, while the heavy oligomers are sent to the hydrogenation zone. In this second reaction zone, the oligomers are transformed to alkanes by hydrogenation reactions, which are carried out in a second fixed-bed reactor. Finally, the products of the hydrogenation reactor are separated into three main fractions, light, jet-fuel and diesel (Wang and Tao, 2016).

The catalytic column integrates these separation and reaction stages in a single equipment, the oligomerization zone is located at the top of the column, it is in this zone where the proper distribution of hydrocarbons is obtained. The compounds that descend from this zone are transformed into olefins (alkanes), which at the same time as they are produced are separated into the 3 main fractions. The mathematical modeling of this intensified process was performed in Aspen Plus software, considering an ethylene input flow of 2100kg/hr. The kinetic models were taken from (Goortani et al., 2015). NRTL-RK (Interan, 2021) was selected as thermodynamic model. It is important to mention that the original design of the intensified column was obtained through a sensitivity analysis of some important design parameters, such as hold ups, number of hold ups, feed stages, etc.

The preliminary design of the catalytic column will be used as the starting point for the optimization algorithm. As mentioned earlier, the total annual cost (TAC) and the condition number were the objective functions to be minimized. The total annual cost consists of the sum of the annualized capital costs, plus the operating costs. As shown below:

$$TAC = \frac{\text{Capital cost}}{\text{Payback period}} + \text{Operating cost} \quad (1)$$



**Figure 1.** Intensification of ATJ process using catalytic distillation

The capital cost is associated to the construction cost of process equipment, such as condensers, reboilers, distillation columns, trays, process vessels and compressors; whereas the operating cost corresponds to the cost of use of electricity, cooling water, steam and solvents. The Guthrie method is used to calculate the total annual cost. The parameters and equations required to estimate the cost of equipment and utilities were taken from Turton et al., 2018. Carbon steel is the construction material considered, and a payback period of ten years is used, which corresponds to an average payback time of a chemical plant according to (Turton et al., 2018). Trays type Sieve with 0.61m spacing were considered for distillation columns. The utilities costs are calculated considering 8,500 hours of operation per year. Cooling water (0.355USD/GJ), electricity (16.8 USD/kwh) and Fired heat (20.92 USD/GJ) were the utilities considered (Turton et al., 2018).

The condition number is a common index used to determine in a qualitative way the controllability of a specific process. It has proven be a powerful tool as it analyzes the control properties of a process in order to detect potential operational problems. Mathematically, the condition number is calculated through a singular value decomposition of the gain matrix according to equation 2:

$$K = W \cdot S \cdot V^T \quad (2)$$

Where K is the relative gains matrix, W and V are unitary matrices and S is the singular value matrix. From the matrix S we take the maximum singular value ( $\sigma^*$ ) and the minimum singular value, to calculate the condition number as shown below:

$$\gamma = \frac{\sigma^*}{\sigma_*} \quad (3)$$



The condition number represents the sensitivity of a system to assimilate a disturbance without process destabilization. indicate that the processes could be susceptible to destabilization by small perturbations. Furthermore, low condition numbers represent a robust process that is resistant to perturbations

Based on the above, the optimization problem can be stated mathematically as equation 4.

$$\begin{aligned} \min Z &= [TAC, \gamma] \\ \text{Subject to } & y_i \geq TE_i \\ & w_i \geq u_i \end{aligned} \quad (4)$$

The objective functions are constrained to meet the minimum boiling point specifications ( $TE_i$ ) for each of the products which are 220°C for jet fuel and 300°C for diesel. On the other hand, the optimization problem is also subject to meet minimum production quantities ( $u_i$ ) which were 1000kg/hr for jet fuel and 100kg for diesel. As decision variables, the number of stages, feed stages, number of reactive stages, hold up values, reflux ratio and thermal load were taken.

Differential evolution with tabu list (DETL), the mathematical method used to solve the optimization problem. DETL is a direct population-based search method, specially designed for non-continuous and highly non-linear functions. This method consists of the 4 stages: generation of the initial population, mutation, crossover and selection. Its main feature is its ability to avoid or revisit areas with bad values of objective functions, by tracking previous searches using the tabu list, improving its computational efficiency (Srinivas and Rangaiah, 2007). The parameters used for the algorithm are: population size: 120, number of generations: 1250, Tabu list size 60, crossover factor: 0.8, mutation factor: 0.3 and tabu radius of 0.01. The algorithm parameters were taken from (Contreras-Zarazúa et al., 2017). Finally, the implementation of the optimization strategy involves a hybrid platform, where, the optimization algorithm is programmed in Excel and the mathematical model of the process and equation solving is in Aspen Plus.

### 3. Results

This section presents the results obtained during the design and simultaneous optimization stage of the catalytic column. In order to determine the solution with the best balance between both objective functions, the utopia point methodology was used to determine the solution, which requires the Pareto front analysis. The Pareto front of the solution is shown in Figure 2.

As can be seen in Figure 2, the range of condition numbers for the different solutions is quite wide, ranging from 11.34 to 2,500 000, indicating that the design of the column design parameters has a strong impact on the process accounting. Note the presence of a red dot in the Pareto, this dot corresponds to the solution with the best trade-off between the two objectives. Table 1 shows the comparison of the original intensified process versus the process obtained by the optimization technique.

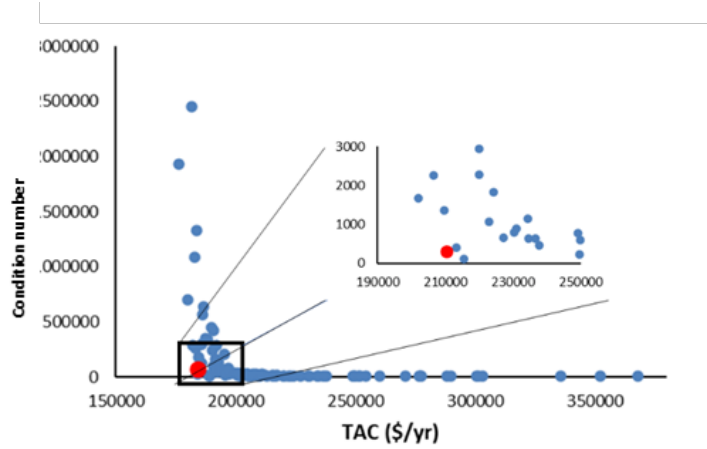


Figure 2. Pareto front for catalytic distillation

Table 1. Design parameters for the original process and intensified process

Process variable	Original process	Optimized process
Number of stages	80	29
Ethylene feed stage	15	21
Hydrogen feed stage	59	26
Reactive stages (oligomerization)	2-29	12-16
Reactive stages (Hydrogenation)	30-59	17-25
Reflux ratio	100	51.55
Reboiler duty (kW)	170	117.642
Hydrogen mass flowrate (kg/hr)	22	20.6417
Diameter(m)	4	1.39
Condition number	915, 000	82.58
Total annual cost (USD/yr)	1,028,914	215,553

As can be seen in Table 1, the intensified process shows remarkable improvements over its original counterpart. It can be seen that the optimized process presents energy savings of 30% while the cost savings are almost 80%. As for the control properties, the optimized process has a condition number of 82.58 against 915,000 of the original process. This means that the optimized process is less sensitive to changes and disturbances. This is due

to the reactive stages and column diameters. The oligomerization and hydrogenation reactions are very sensitive to temperature changes, so having many reactive stages for these areas means that it is easier for the process to have a temperature perturbation, which dramatically affects hydrocarbon conversion. In contrast, the optimized process has determined that it is not necessary to have many reactive stages for these zones, and the optimization process has identified the optimal location of the column temperature reactive zones, thus avoiding the process to be sensitive to disturbances.

#### 4. Conclusions

In this work, the design and simultaneous optimization of a reactive distillation column for jet fuel production by the ATJ process was developed. The objective functions, total annual cost and condition number were used to evaluate the cost and control properties. The results indicate that the optimization algorithm is able to generate a design that has 30% energy savings and 80% cost savings over the original process. Also, the optimized process has a condition number of 82.58 which is a notably lower value than the original process. The results indicate that the design of a process using robust optimization techniques generates a significantly cheaper and more efficient process. Future work proposes the optimization of the conventional process and a rigorous control analysis in order to have a broader picture of the improvements of the optimized process.

#### References

- Combust. Sci. 63, 146–172. <https://doi.org/10.1016/j.pecs.2017.07.001>
- Contreras-Zarazúa, G., Vázquez-Castillo, J.A., Ramírez-Márquez, C., Segovia-Hernández, J.G., Alcántara-Ávila, J.R., 2017. Multi-objective optimization involving cost and control properties in reactive distillation processes to produce diphenyl carbonate. *Comput. Chem. Eng.* 105. <https://doi.org/10.1016/j.compchemeng.2016.11.022>
- Goortani, B.M., Gaurav, A., Deshpande, A., Ng, F.T.T., Rempel, G.L., 2015. Production of isooctane from isobutene: Energy integration and carbon dioxide abatement via catalytic distillation. *Ind. Eng. Chem. Res.* 54, 3570–3581. <https://doi.org/10.1021/ie5032056>
- Interan, R.R., 2021. Optimización De Sistemas Intensificados Para La Producción De Bioturbosina A Través De Alcoholes De Forma Sustentable. Universidad de Guanajuato.
- Srinivas, M., Rangaiah, G.P., 2007. Differential evolution with tabu list for solving nonlinear and mixed-integer nonlinear programming problems. *Ind. Eng. Chem. Res.* 46, 7126–7135. <https://doi.org/10.1021/ie070007q>
- Turton, R., Bailie, R.C., Whiting, W.B., Shaeiwitz, J.A., 2018. Analysis, synthesis and design of chemical processes. Pearson Education., 1-1200.
- Wang, W.C., Tao, L., 2016. Bio-jet fuel conversion technologies. *Renew. Sustain. Energy Rev.* 53, 801–822. <https://doi.org/10.1016/j.rser.2015.09.016>

# Efficient physical model building algorithm using equations extracted from documents

Shota Kato and Manabu Kano

*Department of Systems Science, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan*

## Abstract

To facilitate arduous physical model building tasks, we aim to develop automated physical model builder, AutoPMoB, which automatically builds physical models from multiple documents. This study focuses on a method for building desired models from equations, a fundamental technology for realizing AutoPMoB. We defined two requirements desired models must fulfill and proposed a gradual method. The gradual method obtains combinations of input equations satisfying one requirement and then modifies the combinations to fulfill the other. For comparison, we used an exhaustive method, which obtains all possible combinations of the equations and checks whether each combination fulfills the requirements. We compared the models built by the two methods and their computational time in four cases. In all cases, both methods built models including all correct models. The gradual method and exhaustive method took  $1.0 \times 10^{-4}$  s and 30 s, respectively, to build models from 23 equations.

**Keywords:** Physical model building, Modeling algorithm, Process systems engineering.

## 1. Introduction

In the process industry, a digital twin, which is a model that mimics the behavior of a real process like a twin, is expected to yield significant benefits. Digital twins should be based on physical models derived from scientific principles rather than statistical models derived only from data since data collection through experiments under various conditions, especially abnormal ones, in the manufacturing process is difficult from cost and safety perspectives. However, the physical model building requires a deep understanding of the process and trial-and-error improvements in the model's accuracy. To facilitate the arduous task, we aim to develop automated physical model builder, AutoPMoB, which automatically builds physical models from multiple documents [Kato and Kano (2022)]. AutoPMoB 1) retrieves documents related to the target process from literature databases, 2) unifies the documents' formats, 3) extracts information vital for model building from the documents, 4) unifies the expression of the extracted information, and 5) combines the information to build physical models and presents the models with their rankings. Users of AutoPMoB decide which model to use based on the rankings.

This study focuses on how to build models by combining equations extracted from documents to perform task 5. We assume that users give input variables (IVs) and output variables (OVs) of the models. For example, when a user builds a model to predict a variable, the variable is an OV, and the other measurable variables are IVs.

The present study proposes a gradual method and uses an exhaustive method as a baseline. We compare the models built by the two methods and the required computational time.

## 2. Methods

The number of degrees of freedom (DoF)  $N_{DF}$  of a model calculating the OVs from the IVs must be equal to the number of the IVs  $N_{IV}$ . We propose a gradual method for building such desired models from a set of  $N_e$  equations  $S_e = \{e_1, \dots, e_{N_e}\}$ . Each desired model is an equation group (EG) that meets the following requirements: 1) the EG includes the IVs and OVs, and 2)  $N_{DF}$  of the EG is equal to  $N_{IV}$ . Here, the given set of IVs and that of OVs are denoted by  $S_{IV} = \{v_1, \dots, v_{N_{IV}}\}$  and  $S_{OV} = \{v'_1, \dots, v'_{N_{OV}}\}$ , respectively. Since there has not existed any method for comparison, we use an exhaustive method, which obtains all combinations of  $N_e$  equations and checks whether each combination fulfills the requirements, as a baseline.

### 2.1. Exhaustive method

The exhaustive method is shown in Algorithm 1. The method obtains all possible combinations of  $N_e$  equations (ll. 3–6) and checks whether each combination fulfills the requirements above (ll. 7–13). The total number of the combinations is

$$\sum_{n=1}^{N_e} C_n = 2^{N_e} - 1. \quad (1)$$

This method checks whether each of the  $2^{N_e} - 1$  EGs satisfies the requirements and outputs EGs that meet the requirements. The time complexity of this exhaustive method is  $O(2^{N_e})$ .

---

#### Algorithm 1: Exhaustive method for building model

---

**Input:** set of equations  $S_e = \{e_1, \dots, e_{N_e}\}$   
 set of input variables  $S_{IV} = \{v_1, \dots, v_{N_{IV}}\}$   
 set of output variables  $S_{OV} = \{v'_1, \dots, v'_{N_{OV}}\}$

**Output:** set of equation groups  $S_{EG}$

- 1:  $S_{EG} \leftarrow \emptyset$
- 2:  $S'_{EG} \leftarrow \emptyset$
- 3: **for**  $n_e = 1$  to  $N_e$  **do**
- 4:      $S''_{EG} \leftarrow n_e$  combinations of  $S_e$
- 5:     add  $S''_{EG}$  to  $S'_{EG}$
- 6: **end for**
- 7: **for**  $G'$  in  $S'_{EG}$  **do**
- 8:      $S_{v,G'} \leftarrow$  set of variables in  $G'$
- 9:      $N_{DF,G'} \leftarrow$  the number of degrees of freedom of  $G'$
- 10:     **if**  $(S_{IV} \cup S_{OV}) \subseteq S_{v,G'}$  **and**  $N_{DF,G'}$  is equal to  $N_{IV}$  **then**
- 11:         add  $G'$  to  $S_{EG}$
- 12:     **end if**
- 13: **end for**
- 14: **return**  $S_{EG}$

---

## 2.2. Gradual Method

Algorithm 2 shows the pseudocode of the gradual method. The method first obtains sets of variables contained in each formula (ll. 2–4), sets of equations containing each of the IVs and OVs (ll. 5–8), and sets of EGs that include both IVs and OVs (l. 9). Then, the method checks each of the obtained EGs,  $G'$ . If the number of DoF of  $G'$ ,  $N_{DF,G'}$ , is equal to  $N_{IV}$ ,  $G'$  is judged desired (ll. 10–13). If not, to change  $N_{DF,G'}$ , the method searches the remaining equations for the equations containing the variables that appear in  $G'$  (ll. 14–17). If such equations exist, one or more of the equations are combined with  $G'$ , and the combined EGs whose number of DoF is equal to  $N_{IV}$  are decided desired (ll. 18–24). Finally, the method outputs all the desired EGs.

---

### Algorithm 2: Gradual method for building model

---

**Input:** set of equations  $S_e = \{e_1, \dots, e_{N_e}\}$   
 set of input variables  $S_{IV} = \{v_1, \dots, v_{N_{IV}}\}$   
 set of output variables  $S_{OV} = \{v'_1, \dots, v'_{N_{OV}}\}$

**Output:** set of equation groups  $S_{EG}$

- 1:  $S_{EG} \leftarrow \emptyset$
- 2: **for**  $n = 1$  to  $N_e$  **do**
- 3:      $S_{v,e_n} \leftarrow$  set of variables in  $e_n$
- 4: **end for**
- 5:  $N_{IOV} = N_{IV} + N_{OV}$
- 6: **for**  $n = 1$  to  $N_{IOV}$  **do**
- 7:      $S_{e,v''_n} \leftarrow$  set of equations including  $v''_n$  in  $S_e$ , where  $v''_n \in (S_{IV} \cup S_{OV})$
- 8: **end for**
- 9:  $S'_{EG} \leftarrow \{\{e'_1, \dots, e'_{N_{IOV}}\} \mid e'_n \in S_{e,v''_n} \text{ for every } n \in \{1, \dots, N_{IOV}\}\}$
- 10: **for**  $G'$  in  $S'_{EG}$  **do**
- 11:      $N_{DF,G'} \leftarrow$  the number of degrees of freedom of  $G'$
- 12:     **if**  $N_{DF,G'}$  is equal to  $N_{IV}$  **then**
- 13:         add  $G'$  to  $S_{EG}$
- 14:     **else**
- 15:          $S_{v,G'} \leftarrow \bigcup_{e \in G'} S_{v,e}$
- 16:          $\tilde{S}_{v,G'} \leftarrow S_{v,G'} \setminus (S_{IV} \cup S_{OV})$
- 17:          $S''_{EG} \leftarrow \{\{e''_1, \dots, e''_{|\tilde{S}_{v,G'}|}\} \mid e''_n \in S_{e,\tilde{v}_n} \text{ for every } \tilde{v}_n \in \tilde{S}_{v,G'}\}$
- 18:         **For**  $G''$  in  $S''_{EG}$  **do**
- 19:              $\tilde{G} = G'' \cup G'$
- 20:              $N_{DF,\tilde{G}} \leftarrow$  the number of degrees of freedom of  $\tilde{G}$
- 21:             **if**  $N_{DF,\tilde{G}}$  is equal to  $N_{IV}$  **then**
- 22:                 add  $\tilde{G}$  to  $S_{EG}$
- 23:             **end if**
- 24:         **end for**
- 25:     **end if**
- 26: **end for**
- 27: **return**  $S_{EG}$

---

The time complexity of the gradual method is

$$O(N_e N_{IOV}) + O\left(\prod_{n=1}^{N_{IOV}} |S_{e,v'_n}|\right) + \sum_{i=1}^{|S'_{EG}|} O\left(\prod_{n=1}^{|S_{v,G'}(G'_i)|} |S_{e,\bar{v}_n}(G'_i)| + |S''_{EG}(G'_i)|\right), \quad (2)$$

where  $G'_i$  denotes the  $i$ th element of  $S'_{EG}$  and  $S(G'_i)$  indicates that  $S$  depends on  $G'_i$ . The first, second, and third terms indicate the time complexities at lines 1–8, 9, and 10–26, respectively. The gradual method obtains combinations focusing on IVs and OV and judges whether each combination satisfies the requirements; hence, the time complexity is smaller than the exhaustive method.

### 3. Experiments

#### 3.1. Datasets

We generated datasets consisting of  $S_e$ ,  $S_{IV}$ , and  $S_{OV}$  shown in Table 1 based on the textbook about process control [Seborg et al. (2010)] and defined correct models, which can calculate the values of the OVs by substituting values into IVs. Then, we used these  $S_e$ ,  $S_{IV}$ , and  $S_{OV}$  as the input of the model building methods and checked whether the correct models could be obtained.

#### 3.2. Results and Discussion

The model building results of the exhausted method and the gradual method are summarized in Table 2. In Case 1, the two methods obtained the same correct models. In Case 2, where equations with only variables not included in IVs and OVs exist, the gradual method built the same models as in Case 1, while the exhaustive method built 2,097,152 models, most of which were incorrect. Since adding unnecessary equations does not change the number of DoF, the exhaustive method outputs much more models than the gradual method. In Case 3, similar to Case 1, the same models were obtained by the two methods. In Case 4, similar to Case 2, the models obtained by the two methods comprised all the correct models, and the number of models built by the gradual method is less than that by the exhaustive method.

All the built models in Table 2 meet the predefined requirements, but some models, such as those built by only the exhaustive method in Case 2 and the second model in Case 3, are impractical because the values of the OVs cannot be obtained. This is because the two methods judged only the necessary conditions. To remove such impractical models and obtain only practical ones, we need to judge whether the values of OVs can be obtained. If all the equations in one model are linear, the additional check can be done using the coefficient matrix. However, built models usually include nonlinear equations, as shown in Case 3. We must develop a method for selecting practical models.

Equations (1) and (2) indicate that the gradual method can build models within a shorter time than the exhaustive method, even when the number of equations is large. In Case 2, which took the longest computational time among the four cases, the computational time of the exhaustive method was 30 s, and that of the gradual method was  $1.0 \times 10^{-4}$  s. We used MacBook Pro with an Apple M1 Max processor, 10 Cores, 64GB RAM, and macOS 13.0.1.

Table 1 Equation set, input variable (IV) set, and output variable (OV) set used in experiments.

Case	Equation set $S_e$	IV set $S_{IV}$	OV set $S_{OV}$
1	$y = x_1 + x_2$ $x_2 = 0.1$ $x_2 = 1$	$x_1$	$y$
2	$y = x_1 + x_2$ $x_2 = 0.1$ $x_2 = 1$ $a_i = i (i = 1, \dots, 20)$	$x_1$	$y$
3	$V \frac{dC_A}{dt} = v_0 C_{A0} - v_0 C_A + r_A V$ $-r_A = k C_A^n (n = 0, 1, 2)$ $k = k_0 \exp(a/T)$ $k_0 = 0.1$ $a = 10$	$v_0$ $C_{A0}$ $T$ $V$ $C_A$	$\frac{dC_A}{dt}$
4	$V \frac{dC_A}{dt} = v_0 C_{A0} - v_0 C_A + r_A V$ $-r_A = k C_A^n (n = 0, 1, 2)$ $k = k_0 \exp(a/T)$ $k_0 = 0.1$ $a = 10$ $Q = UA(T_c - T)$ $V\rho \frac{dT}{dt} = wC(T_i - T) + H_r V r_A + Q$	$v_0$ $C_{A0}$ $T$ $V$ $C_A$ $T_c$ $U$ $A$ $\rho$ $w$ $C$ $T_i$ $H_r$	$\frac{dC_A}{dt}$     $\frac{dT}{dt}$

#### 4. Conclusion

We proposed a gradual method to build desired models from equations extracted from documents. We used an exhaustive method for comparison and compared the model building results of the two methods in four cases where the equations, IVs, and OVs were generated based on the textbook about process control [Seborg et al. (2010)]. Both methods built models including all correct models in all cases. In terms of the computational time, the gradual method took  $1.0 \times 10^{-4}$  s while the exhaustive method took 30 s when the number of equations was 23. However, the models built by the two methods sometimes had unnecessary ones. We will develop a method for filtering the obtained models to only correct models in our future work.



Table 2 Model building results of exhaustive method and gradual method.

Case	Method	#built models	Examples of built models
1	Exhaustive	2	$\begin{cases} y = x_1 + x_2 \\ x_2 = 1 \end{cases} \quad \begin{cases} y = x_1 + x_2 \\ x_2 = 0.1 \end{cases}$
	Gradual	2	
2	Exhaustive	2,097,152	$\begin{cases} y = x_1 + x_2 \\ x_2 = 1 \\ a_1 = 1 \\ a_2 = 2 \end{cases}$
	Gradual	2	
3	Exhaustive	10	$\begin{cases} V \frac{dC_A}{dt} = v_0 C_{A0} - v_0 C_A + r_A V \\ -r_A = k \\ k = k_0 \exp(a/T) \\ k_0 = 0.1 \\ a = 10 \end{cases}$
	Gradual	10	
4	Exhaustive	22	$\begin{cases} V \frac{dC_A}{dt} = v_0 C_{A0} - v_0 C_A + r_A V \\ -r_A = k \\ k = k_0 \exp(a/T) \\ k_0 = 0.1 \\ a = 10 \\ Q = UA(T_c - T) \\ V\rho \frac{dT}{dt} = wC(T_i - T) + H_r V r_A + Q \end{cases}$
	Gradual	13	

## References

- S. Kato and M. Kano, 2022, Towards An Automated Physical Model Builder: CSTR Case Study, Computer Aided Chemical Engineering, Volume 49, Pages 1669–1674.  
D. E. Seborg, *et al.*, 2010, Process dynamics and control, John Wiley & Sons.

## Acknowledgments

This work was supported by JSPS KAKENHI Grant Number JP21K18849.

# Developing a rigorous chemical reaction network reduction strategy for *n*-hexadecane hydroisomerisation

Fernando Vega-Ramon,<sup>a</sup> Wei Wang,<sup>b</sup> Wei Wu,<sup>b</sup> Dongda Zhang,<sup>\*a</sup>

<sup>a</sup> *Department of Chemical Engineering and Analytical Science, the University of Manchester, The Mill, Sackville Street, Manchester, M1 3AL, UK.*

<sup>b</sup> *School of Chemistry and Material Sciences, Heilongjiang University, Harbin, 150080, China*

<sup>\*</sup> *Corresponding author. Email: [dongda.zhang@manchester.ac.uk](mailto:dongda.zhang@manchester.ac.uk)*

## Abstract

The hydroisomerisation of long-chain alkanes on bifunctional zeolite catalysts is a promising synthetic route for the sustainable production of diesel fuel derived from vegetable oils. Significant efforts have been made in literature to propose potential reaction mechanisms and materials with improved catalytic activity, but the development of a kinetic model remains a challenge due to the complexity of the reaction network. In this study a microkinetic model was initially constructed to simulate both observable and non-observable states in the reaction mechanism. Upon fitting the model against packed bed reactor experimental data, two rigorous model simplification strategies were proposed and adopted to guide reaction network reduction. In the physics-based approach the ratios of the fitted microkinetic rate constants are evaluated against a threshold to identify irreversible reactions, whereas in the mathematics-based approach an additional penalty term is introduced during microkinetic parameter estimation to penalise the number of active chemical pathways. The reaction network reduction results were similar under both strategies and provided valuable insight on the irreversibility of isomerisation, cracking and desorption reactions, as also supported by observations in literature. Finally, simplified kinetic models that are only capable of simulating observable states based on the steady-state assumption were also constructed using the two reduced reaction networks. Through comparison, it is concluded that both simplified kinetic models can yield satisfactory fitting result of the process; however, their model structures and parameter values are highly sensitive to the mechanism reduction strategy adopted. This strongly indicates the impact of different reaction network reduction strategies on kinetic model construction, and that directly building a steady-state kinetic model may not be the best approach to investigate intrinsic reaction mechanisms. The kinetic modelling and reaction network reduction frameworks proposed in this work, therefore, provide a new avenue to infer mechanistic knowledge from kinetic data in a more efficient manner than traditional steady-state approaches.

**Keywords:** Hydroisomerisation, kinetic model, reaction network reduction, catalytic reaction, intrinsic mechanisms.

## 1. Introduction

### 1.1. Background and motivation

The hydroisomerisation of *n*-alkanes found in waxy feedstock is considered a promising sustainable route for diesel synthesis. Recent research in this area has mainly focused on

novel catalyst development in an attempt to minimize the yield toward cracking by-products and improve isomerization selectivities. This sets the motivation behind microkinetic modelling approaches, which can help elucidate potential reaction mechanisms, streamline the design of catalysts with improved performance, and provide guidelines for model-based design of experiments and process optimisation (Motagamwala & Dumesic, 2021). In this work, a microkinetic modelling and reaction network simplification framework was proposed to analyse experimental data of *n*-hexadecane hydrosiomerisation and cracking over a Pd/ZSM-12 catalyst.

## 2. Methodology

### 2.1. Kinetic model construction

#### 2.1.1. Base-case reaction mechanism

To construct a microkinetic model comprising equations for the rate of change of reactants, products and intermediate species, the reaction network presented in Fig. 1 was proposed by considering the plausible elementary steps occurring at the catalyst surface.

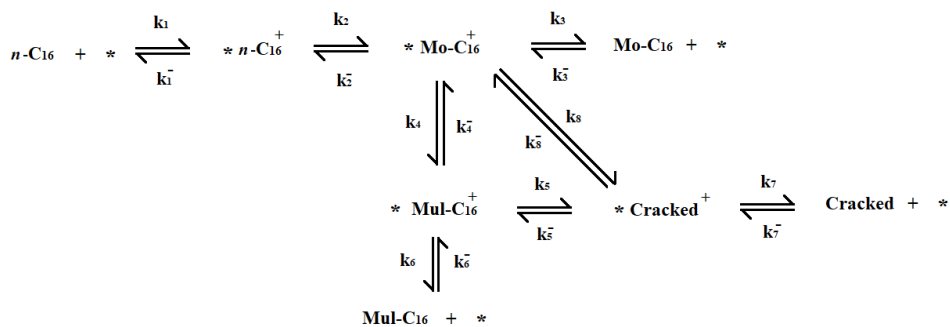


Figure 1: Proposed reaction mechanism. *n*-C<sub>16</sub>, Mo-C<sub>16</sub> and Mul-C<sub>16</sub> refer to unbranched, monobranched and multibranched hexadecane isomers, Cracked refers to cracking by-products. Their respective carbenium ion intermediates are denoted by the subscript <sup>+</sup>, while \* represents an active site. *k<sub>i</sub>* and *k<sub>i</sub><sup>-</sup>* are the forward and reverse reaction rates of elementary step *i*, respectively.

In the classical mechanism of *n*-alkane hydrosiomerisation described by Weitkamp (2012), the unbranched alkane reactant is first adsorbed at the metallic sites of the bifunctional catalyst, where it is dehydrogenated into its corresponding *n*-alkene. The *n*-alkene intermediate is then readily protonated at the acid catalytic sites, forming a primary alkylcarbenium ion intermediate. This intermediate then undergoes successive skeletal rearrangement reactions into more stable monobranched, dibranched and tribranched isomers. The carbenium isomers desorb from the acid sites as their corresponding branched alkenes, which are then hydrogenated at the metallic sites to yield branched isomers of the starting *n*-alkane. The branched isomers may also crack due to β-scission reactions, resulting in shorter-chain alkenes and carbenium ions which can also participate in further cracking reactions.

In this study, the packed-bed reactor (PBR) experimental data lumped dibranched and tribranched isomer products into a single multibranched species. Similarly, the by-products of the cracking side-reactions were lumped into a single species for convenience. The mechanism proposed in this study was therefore modified accordingly such that it was in line with the measured states, as seen in Fig. 1. Furthermore, the fractional site coverage of surface species is not observable during practical applications, and so including all the reaction intermediates and active site balances in the kinetic analysis

would lead to non-identifiable model structures. Coonradt and Garwood (1964) suggested that if there are sufficiently high concentrations of metallic sites for hydrogenation/dehydrogenation, the alkene intermediates assume their equilibrium concentrations. This, in turn, means that the acid site-catalysed isomerisation and cracking reactions become selectivity- and rate-determining for the overall reaction (Vandegehuchte et al., 2014). Thus, the metal-catalysed surface reactions were neglected in the present study, and it was assumed that the gas-phase species directly adsorb to (or desorb from) their corresponding alkylcarbenium ion intermediates at the acidic sites. For this reason, the alkene intermediates and the metallic active site balance were not considered during kinetic model construction.

### 2.1.2. Microkinetic model construction

The microkinetic model developed in this work consists of balance equations for both the gas-phase species as well as the non-observable intermediate states in the proposed reaction network (Fig. 1). These are presented in Eq. 1(a) – Eq 1(i), where  $x_i$  refers to the mass fraction of gas-phase species  $i$ ,  $\tau$  is the weight-based spacetime (h),  $\theta_i$  is the fractional coverage of the alkylcarbenium intermediate  $i$ , the subscripts “r”, “mo”, “mul”, and “cr” refer to unbranched, monobranched, multibranched, and cracked species, respectively. Forward and backward reaction constants ( $\text{g}_c \text{ mol}^{-1} \text{ h}^{-1}$ ) of the  $j^{\text{th}}$  elementary step are denoted by  $k_j$  and  $k_j^-$ , respectively.  $c_{\theta_T}$  is the total active site concentration of the catalyst ( $\text{mol g}_c^{-1}$ ),  $\rho_b$  is the catalyst bulk density ( $\text{g}_c \text{ m}^{-3}$ ) and  $c_{r,0}$  is the initial concentration of the reactant ( $\text{mol m}^{-3}$ ). In addition, the model assumes 1:1 overall reaction stoichiometry (this is exact for the isomerisation reactions but clearly an approximation for the cracking reactions) and isothermal reactor operation, such that the total pressure of the reacting mixture can be assumed constant.

$$\frac{dx_r}{d\tau} = c_{\theta_T} \cdot (-k_1 \cdot x_r \cdot \theta_V + k_1^- \cdot \theta_r) \quad \text{Eq. 1(a)}$$

$$\frac{d\theta_r}{d\tau} = \frac{c_{r_0}}{\rho_b} \cdot [k_1 \cdot x_r \cdot \theta_V + k_2^- \cdot \theta_{Mo} - (k_1^- + k_2) \cdot \theta_r] \quad \text{Eq. 1(b)}$$

$$\frac{dx_{Mo}}{d\tau} = c_{\theta_T} \cdot (k_3 \cdot \theta_{Mo} - k_3^- \cdot x_{Mo} \cdot \theta_V) \quad \text{Eq. 1(c)}$$

$$\frac{d\theta_{Mo}}{d\tau} = \frac{c_{r_0}}{\rho_b} \cdot [k_3^- \cdot x_{Mo} \cdot \theta_V + k_2 \cdot \theta_r + k_4^- \cdot \theta_{Mul} + k_8^- \cdot \theta_{Cr} - (k_2^- + k_3 + k_4 + k_8) \cdot \theta_{Mo}] \quad \text{Eq. 1(d)}$$

$$\frac{dx_{Mul}}{d\tau} = c_{\theta_T} \cdot (k_6 \cdot \theta_{Mul} - k_6^- \cdot x_{Mul} \cdot \theta_V) \quad \text{Eq. 1(e)}$$

$$\frac{d\theta_{Mul}}{d\tau} = \frac{c_{r_0}}{\rho_b} \cdot [k_6^- \cdot x_{Mul} \cdot \theta_V + k_4 \cdot \theta_{Mo} + k_5^- \cdot \theta_{Cr} - (k_4^- + k_5 + k_6) \cdot \theta_{Mul}] \quad \text{Eq. 1(f)}$$

$$\frac{dx_{Cr}}{d\tau} = c_{\theta_T} \cdot (k_7 \cdot \theta_{Cr} - k_7^- \cdot x_{Cr} \cdot \theta_V) \quad \text{Eq. 1(g)}$$

$$\frac{d\theta_{Cr}}{d\tau} = \frac{c_{r_0}}{\rho_b} \cdot [k_7^- \cdot x_{Cr} \cdot \theta_V + k_5 \cdot \theta_{Mul} + k_8 \cdot \theta_{Mo} - (k_5^- + k_7 + k_8^-) \cdot \theta_{Cr}] \quad \text{Eq. 1(h)}$$

$$\theta_V = 1 - \theta_r - \theta_{Mo} - \theta_{Mul} - \theta_{Cr} \quad \text{Eq. 1(i)}$$

### 2.1.3. Kinetic parameter estimation

The kinetic models were fitted against the PBR data; the parameter estimation problem was formulated as a weighted least squares objective function (Eq. 4(a)), subject to nonlinear process constraints and parameter bound constraints (Eq. 4(b)):

$$\min \sum_j (x_{j,E} - x_{j,M})^T \Lambda (x_{j,E} - x_{j,M}) \quad \text{Eq. 4(a)}$$

$$\text{st. } \frac{dx}{d\tau} = f(x, \theta, k), \quad \frac{d\theta}{d\tau} = g(x, \theta, k), \quad \mathbf{k}_{lb} < \mathbf{k} < \mathbf{k}_{ub} \quad \text{Eq. 4(b)}$$

where  $\mathbf{x}_{j,E}$  and  $\mathbf{x}_{j,M}$  are the experimental and model estimates of the state variables, respectively.  $f(x, \theta, k)$  and  $g(x, \theta, k)$  refer to the kinetic model differential equations for the gas-phase and surface species, respectively.  $\mathbf{k}_{lb}$  and  $\mathbf{k}_{ub}$  are the lower and upper bounds of the optimization variables  $\mathbf{k}$ . The weighting matrix  $\Lambda$  is used to normalise the residuals. The differential process constraints were fully discretised into algebraic profiles by applying orthogonal collocation on finite elements. The arising nonlinear programming problem was then solved via interior-point optimization (IPOPT) algorithms. Both the numerical discretization and IPOPT procedures were implemented via the *Pyomo* modelling environment in *Python* programming language.

## 2.2. Reaction network simplification

### 2.2.1. Physics-based network reduction

The physics-based reduction approach aims to identify kinetically negligible reactions by comparison of the forward and backward reaction rates of each elementary step. Similar rate-based procedures have been applied in literature to automatically discriminate elementary steps in complex reaction mechanisms (Goldsmith & West, 2017). In this study, the estimated equilibrium constants were compared against an upper and lower threshold:

- If  $\frac{k_i}{k_i^-} > 10$ , then the backward reaction is negligible ( $k_i^- \approx 0$ )
- If  $\frac{k_i}{k_i} < 0.1$ , then the forward reaction is negligible ( $k_i \approx 0$ )

These thresholds have been chosen on the basis that most of the estimated rate constants in the proposed mechanism were 1 order of magnitude apart ( $10^5$ -  $10^6$ ), and that both observable and non-observable states in the kinetic model were expressed in normalised units (mass fractions and fractional surface coverages, respectively). Upon identifying negligible steps, the microkinetic model was reconstructed in accordance with the physics-reduced mechanism and subsequently fitted against experimental data.

### 2.2.2. Mathematics-based network reduction

In the mathematics-based network reduction strategy, an additional penalty term is included in the parameter estimation objective function to penalise the number of non-zero reaction rate constants:

$$\min \sum_j (\mathbf{x}_{j,E} - \mathbf{x}_{j,M})^T \Lambda (\mathbf{x}_{j,E} - \mathbf{x}_{j,M}) + \omega \sum_i \left( \frac{k_i}{k_i + 1} + \frac{k_i^-}{k_i^- + 1} \right) \quad \text{Eq. 5}$$

For non-negligible microkinetic parameters, we have  $k_i \gg 1$  and the corresponding penalty term  $\frac{k_i}{k_i + 1} \approx 1$ , whereas for negligible reactions we obtain  $\frac{k_i}{k_i + 1} \approx 0$ . This, in turn, allows for the identification of kinetically redundant reactions without the introduction of binary variables and mixed-integer programming. The penalty weight  $\omega$  must be manually tuned so the sparsity of the model does not compromise its fitting accuracy.

## 3. Results

### 3.1. Microkinetic model results

The microkinetic model for the proposed mechanism exhibited good fitting performance, with an overall mean percentage error of 6.94%; the simulated process trajectories for the Pd/ZSM-12 catalyst with 0.5 Pd wt% composition are presented in Fig. 2 below. The

estimated microkinetic constants not only allow for the simulation of the process, but also provide mechanistic insight on the underlying reaction network. For example, comparison of the estimated rate constants  $k_3 = 8.02 \times 10^5$ ,  $k_6 = 5.99 \times 10^5$  and  $k_7 = 1.45 \times 10^6$   $\text{g}_c \text{ mol}^{-1} \text{ h}^{-1}$  indicate that the monobranched and cracking intermediates are more readily desorbed to the gas-phase than their multibranched counterpart. This trend can be justified with reference to the higher diffusional energy barriers of multibranched paraffins in zeolite catalysts (Oenema et al., 2020), recalling that the desorption step involves the diffusion of an alkene intermediate from the acid sites to the metallic sites. Similarly, the estimated multibranched cracking rate constant ( $k_5 = 3.44 \times 10^6$   $\text{g}_c \text{ mol}^{-1} \text{ h}^{-1}$ ) is much larger than its desorption rate constant ( $k_6 = 5.99 \times 10^5$   $\text{g}_c \text{ mol}^{-1} \text{ h}^{-1}$ ), thus providing a mechanistic justification for the high yield towards cracking by-products and the comparatively low yield towards the multibranched hexadecane isomers. Previous microkinetic studies of *n*-alkane hydroconversion have reached similar conclusions regarding the cracking affinity of multibranched alkanes and its effect on the product distribution (Vandegehuchte et al., 2014).

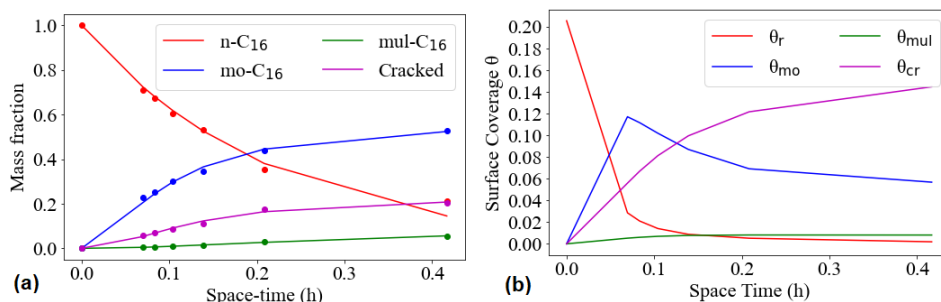


Figure 2: (a) Microkinetic model fitting results for gas-phase species profiles. Dots indicate experimental data while solid lines are model predictions; (b) Simulated coverages of surface species.

### 3.2. Reaction network simplification results

Physics- and mathematics-based network reduction analysis were conducted upon fitting the microkinetic model for the proposed mechanism; in either case five elementary steps were deemed to be irreversible, as seen in the simplified reaction networks presented in Fig. 3(a) and 3(b). The microkinetic model was also reconstructed in accordance with the simplified mechanisms, yielding fitting results (6.22% and 6.98% overall MAPE) of similar accuracy as the original model (6.15%) despite having 5 less parameters. This indicates that the neglected reactions were indeed kinetically insignificant.

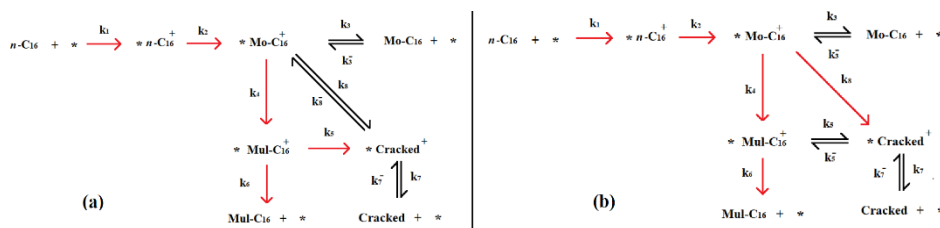


Figure 3: (a) Physics-reduced reaction network, (b) mathematics-reduced network. Red arrows indicate elementary steps that have been identified as irreversible.

Both simplification approaches identified the isomerization reactions (steps 2 and 4 in Fig. 3(a) and Fig. 3(b)) as irreversible. The irreversibility of these skeletal rearrangements has been previously analysed in literature and may be attributed to operating conditions and reactant conversions far away from thermodynamic equilibrium, such that the rates of the reverse isomerization reaction are negligible (Steijns & Froment, 1981). The main dissimilarity between the two reduced networks concerns the reversibility of the two cracking reactions in the mechanism (steps 5 and 8). The relative stabilities of the alkylcarbenium intermediates suggest that the cracking products would re-alkylate more readily into the multibranched species (i.e.  $k_5^- > k_8^-$ ). This is indeed what is observed in the mathematics-reduced network ( $k_8^- = 0$ ), but not in the physics-reduced network ( $k_5^- = 0$ ). This suggests that the reduced microkinetic mechanisms identified by different model reduction strategies should be further evaluated before finally applying them.

#### 4. Conclusion

In this work, a microkinetic modelling approach was adopted to simulate observable gas-phase species and non-observable intermediate states in the proposed *n*-hexadecane hydroisomerisation network. Upon parameter estimation, the model provided accurate fitting of the packed bed reactor data and mechanistic insight was gained through evaluation of the estimated microkinetic constants. Mathematics and physics-based network reduction strategies were then applied to identify kinetically insignificant reactions in the mechanism. The microkinetic model was modified in accordance with the network reduction results, also yielding satisfactory fitting of the experimental data. This study suggests that the two proposed reaction network reduction strategies have great potentials to be applied to generic processes for reaction mechanism investigation.

#### References

- Coonradt, H. L., & Garwood, W. E. (1964). Mechanism of Hydrocracking. Reactions of Paraffins and Olefins. *Industrial & Engineering Chemistry Process Design and Development*, 3(1), 38–45. <https://doi.org/10.1021/i260009a010>
- Goldsmith, C. F., & West, R. H. (2017). Automatic Generation of Microkinetic Mechanisms for Heterogeneous Catalysis. *The Journal of Physical Chemistry C*, 121(18), 9970–9981. <https://doi.org/10.1021/acs.jpcc.7b02133>
- Motagamwala, A. H., & Dumesic, J. A. (2021). Microkinetic Modeling: A Tool for Rational Catalyst Design. *Chemical Reviews*, 121(2), 1049–1076. <https://doi.org/10.1021/acs.chemrev.0c00394>
- Oenema, J., Harmel, J., Vélez, R. P., Meijerink, M. J., Eijsvogel, W., Poursaeidesfahani, A., Vlugt, T. J. H., Zečević, J., & de Jong, K. P. (2020). Influence of Nanoscale Intimacy and Zeolite Micropore Size on the Performance of Bifunctional Catalysts for *n*-Heptane Hydroisomerization. *ACS Catalysis*, 10(23), 14245–14257. <https://doi.org/10.1021/acscatal.0c03138>
- Steijns, M., & Froment, G. F. (1981). Hydroisomerization and hydrocracking. 3. Kinetic analysis of rate data for *n*-decane and *n*-dodecane. *Industrial & Engineering Chemistry Product Research and Development*, 20(4), 660–668. <https://doi.org/10.1021/i300004a014>
- Vandegehuchte, B. D., Thybaut, J. W., & Marin, G. B. (2014). Unraveling Diffusion and Other Shape Selectivity Effects in ZSM5 Using *n*-Hexane Hydroconversion Single-Event Microkinetics. *Industrial & Engineering Chemistry Research*, 53(40), 15333–15347. <https://doi.org/10.1021/ie500164q>
- Weitkamp, J. (2012). Catalytic Hydrocracking—Mechanisms and Versatility of the Process. *ChemCatChem*, 4(3), 292–306. <https://doi.org/10.1002/cctc.201100315>

# Troubleshooting high-pressure issues in an industrial biorefinery process by feature-oriented modeling

Elia Arnese-Feffin<sup>a</sup>, Pierantonio Facco<sup>a</sup>, Daniele Turati<sup>b</sup>,  
Fabrizio Bezzo<sup>a</sup>, Massimiliano Barolo<sup>a,\*</sup>

<sup>a</sup>*CAPE-Lab – Computer-Aided Process Engineering Laboratory, Department of Industrial Engineering, University of Padova, via Marzolo 9, 35131 Padova, Italy*

<sup>b</sup>*Novamont S.p.A., via G. Fauser 8, 28100 Novara, Italy*

\*Corresponding author: [max.barolo@unipd.it](mailto:max.barolo@unipd.it)

## Abstract

Biorefinery plants manufacture chemicals by conversion of sustainable raw materials. Membrane separation processes are commonly used to separate cells from fermentation broths, and account for most of the operating costs of a biorefinery process. Membrane fouling can disrupt normal operation, increasing cleaning costs and process downtime (hence, overall manufacturing costs). Whereas investigating the causes of fouling through mechanistic models can be challenging in an industrial environment, information in process data historians can be leveraged through data-driven modeling. In this study, principal component analysis and feature-oriented modeling are combined to identify potential causes of fouling in a semi-continuous membrane separation process of an industrial biorefinery. This approach can effectively address batch duration variability issues, while exploiting process knowledge to enhance information on effects of fouling. Membrane age and operating temperature were found to be the major variables related to fouling.

**Keywords:** biorefinery; membrane separation processes; process understanding; feature-oriented models; data analytics

## 1. Introduction

Biorefineries are facilities that integrate biomass conversion processes and equipment to sustainably produce fuels, power, and chemicals from biomass (Martín *et al.*, 2013). Process operations typically include: media preparation, production of microorganisms, and large-scale fermentation in the upstream section; broth sterilization, cell separation, and product recovery and purification in the downstream section (Bähner *et al.*, 2021). Operating costs of downstream in biorefineries usually range between 40% and 60% of the total processing cost. Energy is the main cost, as typical downstream operations are evaporation, distillation, and, recently, membrane separation processes (Bähner *et al.*, 2021). Membranes are becoming common in biorefineries to separate cells and/or large molecules from the fermentation broth (Ables *et al.*, 2013), for examples by ultrafiltration and nanofiltration operated (semi-)continuously. Such membrane separation processes can determine the major operating cost of a biorefinery (Satam *et al.*, 2019). Pressure-driven membranes separation processes (i.e., membrane filtration) are the most



widely used, especially to remove cells from the fermentation broth. However, this operation can greatly suffer from membrane fouling, as highlighted in many studies. Ables *et al.* (2013) highlighted a permeate flux decrease in constant pressure filtration, while Klimkiewicz *et al.* (2016) recorded remarkable pressure build-up in constant flow-rate separation. The latter case is particularly relevant as, besides obvious disruption of operation and increase in cleaning costs, it also implies an increase in energy expenditure.

Mathematical modelling is a valuable tool to analyze this problem. Membrane fouling mechanisms are well known for some membrane separation processes, and models are available (Meindersma *et al.*, 1997) to help diagnosing the root-causes of fouling and support decision making to mitigate its effect. However, reliable fouling models often feature remarkable complexity and require high quality data for parameter estimation (Bolton *et al.*, 2016). Although modern biorefineries are heavily sensorized and provide a wealth of data recorded online, data usually regard the history of process operation, and might not be adequate for estimation of parameters of fouling models.

Data historians contain valuable information on process operation and can be leveraged to enhance performance (Cuellar *et al.*, 2020). To this end, data-driven models such as principal component analysis (PCA; Wold *et al.*, 1987), proved valuable when applied to membrane separation processes, as shown by Klimkiewicz *et al.* (2016) and Naessens *et al.* (2017). Applying such methods to membrane separation processes requires some extra care, mostly because of the semi-continuous nature of the process. This implies frequent process downtimes for membrane cleaning and strong variability in the duration of operating periods (runs), in the process variables within runs, and in the profiles of process variables between runs. Variable batch duration entails a lack-of-synchronization issue when process datasets are explored by data analytics.

Feature-oriented modeling (Yoon *et al.*, 2001; Rendall, 2019) offers an elegant way to address the lack-of-synchronization issue, while also emphasizing the phenomena one wants to model by properly defining features (Rendall *et al.*, 2017). In this study, knowledge-driven feature-oriented PCA (Wold *et al.*, 2009) is used to investigate the observable effects of fouling, i.e., high-pressure issues, in the membrane separation section of an industrial biorefinery process. The analysis of features summarizing time-profiles of process variables can highlight potential causes of fouling, thus paving the way for targeted experimental investigations.

## 2. Materials and methods

### 2.1. The membrane separation process

The focus of this study is membrane filtration of a fermentation broth, aimed at separating cells and large molecules from the stream containing the product. The operation is performed as a part of the downstream processing of an industrial biorefinery process where a biopolymer is manufactured. Sensible details on process and data are not disclosed for confidentiality.

The sterilized fermentation broth is accumulated in two parallel tanks, and, during operation, is fed to the first membrane module of the sequence at a constant flow-rate. Part of the retentate of each module gets discharged to the retentate manifold, while the remainder is fed to the following module. The first modules perform a standard filtration, while last ones operate in diafiltration: part of the permeate of each module is sent to the permeate manifold, and the remainder is fed back to the preceding module; water is fed to the very last one. A simplified scheme of the process and sensors is shown in Figure 1. The process runs in semi-continuous mode. The membrane modules are initially filled with water, which is displaced by the fermentation broth during the startup phase. In the

steady-state phase, the pressure on the feed/retentate side of the membranes is manipulated to keep flow-rates constant. When the feed tanks are almost empty, the flushing phase begins, and the modules are filled again with water before cleaning-in-place takes place. However, premature interruption of the process occurs frequently due to excessive pressure build-up following membrane fouling. Investigating the potential causes of this issue is the objective of this study.

Production data over a seven-month timespan are available for this investigation. Namely, online measurements from all sensors are available; offline measurements from the upstream process are available as well, to characterize the broth being processed.

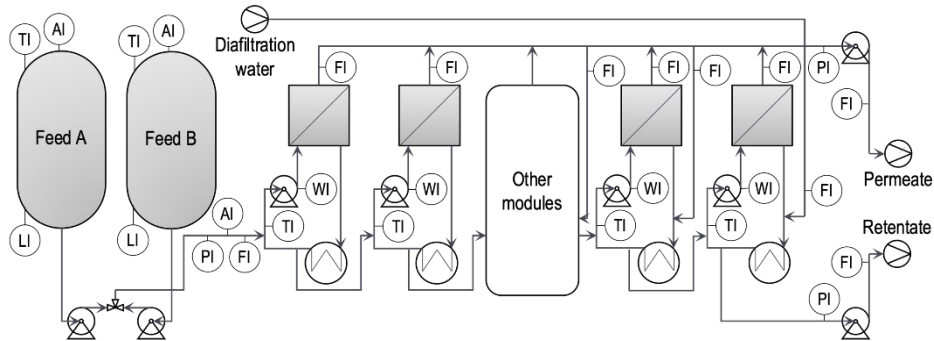


Figure 1. Simplified diagram of the membrane separation process under analysis. Analyzers on feed tanks measure pH, while the one on feed manifold measures conductivity.

## 2.2. Principal component analysis

PCA (Wold *et al.*, 1987) is a multivariate statistical method that allows extracting a sequence of orthogonal variables, called principal components (PCs), from a data matrix  $\mathbf{X} \in \mathbb{R}^N \times \mathbb{R}^V$  collecting  $N$  observations of  $V$  variables. Assuming that  $\mathbf{X}$  is mean-centered (and possibly scaled to unit variance), the PCA model is provided as a matrix decomposition in the form  $\mathbf{X} = \mathbf{T} \cdot \mathbf{P}^T + \mathbf{E}$ , where  $\mathbf{T} \in \mathbb{R}^N \times \mathbb{R}^A$  is the score matrix,  $\mathbf{P} \in \mathbb{R}^V \times \mathbb{R}^A$  is the loading matrix,  $A$  is the number of PCs of the model (to be set prior to model calibration), and  $\mathbf{E} \in \mathbb{R}^N \times \mathbb{R}^V$  is the residual matrix. The scores are projections of observations in  $\mathbf{X}$  onto the PC space and describe the relation among observations. The loadings describe the correlation among variables in  $\mathbf{X}$ , and the relation between the original space and the PC space. The columns of  $\mathbf{P}$  are set as to maximize the variance of data explained by the PCs (while respecting orthogonality constraints) and can be obtained by truncated singular value decomposition of the sample covariance matrix of data.

## 2.3. Feature-oriented modelling of membrane separation processes

PCA requires operating onto two-dimensional data arrays. However, the available data on the membrane separation process under investigation are a sequence  $\{\mathbf{X}_1, \dots, \mathbf{X}_B\}$ , where  $\mathbf{X}_b \in \mathbb{R}^{N_b} \times \mathbb{R}^V$  is a matrix collecting  $N_b$  observations of the  $V$  process variables during the  $b$ -th run of the process, with  $b = 1, 2, \dots, B$ , and  $B$  is the total number of runs (i.e., batches). To apply PCA to this process, one would need to concatenate all  $\mathbf{X}_b$  matrices as a single matrix. However, if the aim is to investigate on runs that suffered from high-pressure issues, this approach may be inadequate due to observations representing single time instants rather than entire runs. In this study, we tackle the problem by means of feature-oriented modeling (Yoon *et al.*, 2001; Rendall *et al.*, 2019). The rationale is to summarize profiles of variables into numerical indexes called features. This approach offers an elegant solution to the lack-of-synchronization issue, while also allowing to emphasize the phenomena one is interested in by properly defining features. To this end,

knowledge-driven features (Wold *et al.*, 2009; Rendall *et al.*, 2017) can be derived from time profiles, such as integrals, averages, slopes, minima, and maxima of measured variables in a run (note that all these operations are independent of  $N_b$ ). Furthermore, the available process knowledge can be used to improve such “standard” features. An example is splitting time-profiles into phases to better characterize the process (e.g., startup, steady operation, and washing). Process knowledge can also be used to design special features.

Feature synthesis can be interpreted as an operator  $\mathcal{F}: \mathbb{R}^{N_b} \times \mathbb{R}^V \rightarrow \mathbb{R}^F \mid \mathbf{X}_b \mapsto \mathbf{h}_b$ , where  $\mathbf{h}_b \in \mathbb{R}^F$  is a vector containing  $F$  features. As one vector of features is obtained from each run, a matrix  $\mathbf{H} \in \mathbb{R}^B \times \mathbb{R}^F$  is obtained, whose rows characterize single runs. Such a matrix can be analyzed by PCA for process understanding and troubleshooting purposes.

### 3. Results

After preliminary screening, data for  $B = 176$  runs were selected for analysis. Concerning data from process sensors, only the steady-state phase was used, as the startup and flushing phases usually feature excessive/unstructured variability and significant nonlinearities. Sensor measurements were first augmented with additional variables (e.g., pressure profile slope) to detect occurrence of fouling. Some engineering variables were added as performance indexes, for instance average energy consumption and volume conversion ratio (VCR, ratio of retentate flow-rate and feed flow-rate). A total of  $V = 61$  online variables was obtained. These variables, together with data from upstream process (characteristics of the broth being processed), were used to compute  $F = 179$  features for PCA analysis. Some of the most important features are collected in Table 1. Note that some features were also computed on sub-phases of the steady-state phase, whenever that seemed meaningful.

*Table 1. Examples of features extracted from profiles of online variables. Features marked with a star are encoded as binary variables.*

Average flow-rates in manifolds	Average VCR
Average feed conductivity	Energy consumed
Slope of feed conductivity	Processed volume
Max/min of pressures	Run duration
Average slope of pressures	Feed tanks average temperature
Max/min of pressure slopes	Feed tanks average pH
Average trans-membrane pressure	Concentration of cells in the feed
Average flow-rates of module permeates	Concentration of organic acids in the feed
Average temperatures of modules	Concentration of ions in the feed
Durations of pump steps	Concentration of product in the feed
Average pump power over steps	Contamination of upstream fermenter*
Pressure features computed over pump steps	Chemicals used in cleaning*

A preliminary PCA model including all variables featured limited interpretability. Therefore, a stepwise approach was adopted. Features were first grouped in subsets according to their “origin”, such as upstream, cleaning, single modules. The most important subset is the one of global features characterizing the membrane filtration, which are basically the ones regarding manifolds: flow-rates, pressures, and performance indexes. A first PCA model is developed on global features alone. Pressure-related features were identified as the main drivers of variability (first PC), as expected; the second source of variability (second PC) regards average flow-rates and some of the performance indexes, among which VCR. The explained variances of the other PCs were far lower than the

ones of the first two, and loadings appeared to model random variability. Therefore, the features modelled by the first two PCs were selected as the reference set of features; cell concentration in the feed was added to this set, as it was expected to be an important factor for fouling. A PCA model developed on these selected features proved to properly identify runs that suffered from high pressure issues, which are separated from the bulk of the runs in the score plot reported in Figure 2 (triangles).

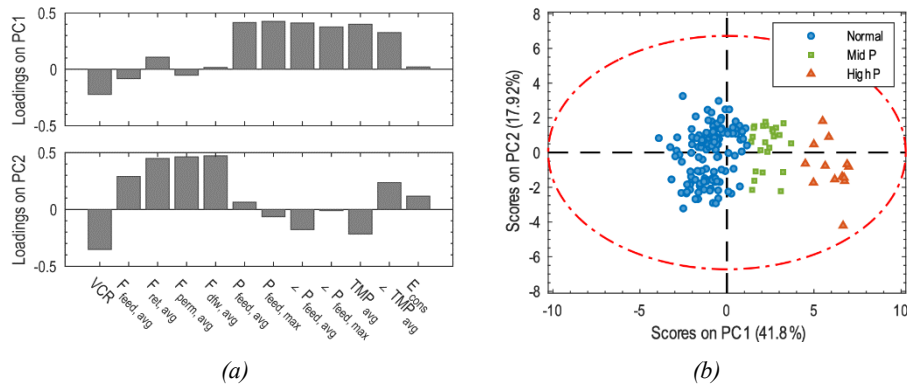


Figure 2. (a) Loadings and (b) scores of the PCA model developed on the set of global features. Runs suffering from high pressure clearly group along PC1.

More features were then added to the data matrix for PCA modelling, in a group-by-group fashion, to assess the occurrence of correlation with the pressure-related features. When such a correlation was found and was physically meaningful, newly added features correlating with pressure were regarded as potential causes of fouling.

The most important factor for high pressure issues was found to be membrane age. This is highlighted by features regarding permeates of single modules, which are proxies for membrane age (i.e., permeate flux decreases run after run). The second most important factor was found to be temperature: more precisely, runs featuring a high temperature of flows entering membrane modules suffered from high-pressures more frequently than runs with low inlet temperature. Such conclusions were verified by targeted analysis of raw data (see Figure 3) and make engineering sense. Therefore, they were deemed as potential causes of fouling, and are currently being verified by designed experiments.

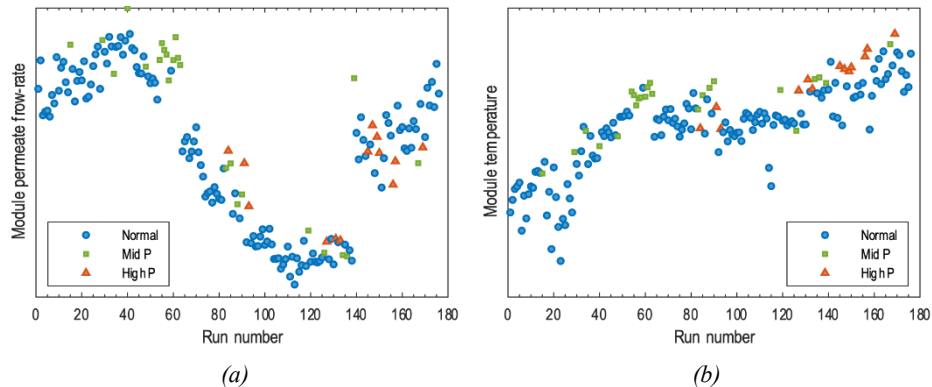


Figure 3. (a) Average permeate flow-rate and (b) temperature of a selected module. Pressure issues are more frequent for old membranes (flow-rate is a proxy for membrane age) and when modules operate at high temperature.

#### 4. Conclusions

Interpretable data-driven models are valuable tools to investigate issues in membrane separation processes, such as membrane fouling, relying solely on data collected during process operation. This study proved how principal component analysis can identify potential causes of fouling by analysis of data concerning the observable effect of this complex phenomenon, namely, pressure increase in membrane filtration. A feature-oriented approach was adopted: instead of using process data directly, numerical values characterizing each operating period were obtained so as to summarize time profiles into time-independent numerical features. This allowed addressing issues (e.g. lack of synchronization and uneven batch duration) that can complicate the use of other approaches. Process knowledge was exploited to design features that enhanced the phenomenon under investigation, thus maximizing the dataset information content and enhancing interpretability with respect to other methods. Incorporation of process knowledge into the data analysis workflow proved essential to identify potential causes of fouling, which are currently being verified by targeted experimental studies.

#### References

- C. Ables, F. Carstensen, M. Wessling, 2013, Membrane processes in biorefinery applications, *Journal of Membrane Science*, 444, 285–317
- F. Böhner, O. A. Prado-Rubio, J. K. Huusom, 2021, Challenges in Optimization and Control of Biobased Process Systems: An Industrial-Academic Perspective, *Industrial and Engineering Chemistry Research*, 60, 42, 14985–15003
- G. Bolton, D. LaCasse, R. Kuriyel, 2006, Combined models of membrane fouling: Development and application to microfiltration and ultrafiltration of biological fluids, *Journal of Membrane Science*, 277, 1–2, 75–84
- M. Cuellar, A. JJ Straathof, 2020, Downstream of the bioreactor: advancements in recovering fuels and commodity chemicals, *Current Opinion in Biotechnology*, 62, 189–195
- A. Klimkiewicz, A. E. Cervera-Padrell, F. W. J. van den Berg, 2016, Multilevel Modeling for Data Mining of Downstream Bio-Industrial Processes, *Chemometrics and Intelligent Laboratory Systems*, 154, 62–71
- M. Martín, I. E. Grossmann, 2013, On the systematic synthesis of sustainable biorefineries, *Industrial and Engineering Chemistry Research*, 52, 9, 3044–3064
- G. Meindersma, J. Augeraud, F. H. P. Vergossen, 1997, Separation of a biocatalyst with ultrafiltration or filtration after bioconversion, *Journal of Membrane Science*, 125, 2, 333–349
- W. Naessens, T. Maere, G. Gilibert-Oriol, V. Garcia-Molina, I. Nopens, 2017, PCA as tool for intelligent ultrafiltration for reverse osmosis seawater desalination pretreatment, *Desalination*, 419, June, 188–196
- R. Rendall, B. Lu, I. Castillo, S. T. Chin, L. H. Chiang, M. S. Reis, 2017, A Unifying and Integrated Framework for Feature Oriented Analysis of Batch Processes, *Industrial & Engineering Chemistry Research*, 56, 30, 8590–8605
- R. Rendall, L. H. Chiang, M. S. Reis, 2019, Data-driven methods for batch data analysis – A critical overview and mapping on the complexity scale, *Computers & Chemical Engineering*, 124, 1–13
- C. Satam, M. Daub, M. J. Realf, 2019, Techno-economic analysis of 1,4-butanediol production by a single-step bioconversion process, *Biofuels, Bioproducts and Biorefining*, 13, 5, 1261–1273
- S. Wold, K. Esbensen, P. Geladi, 1987, Principal Component Analysis, *Chemometrics and Intelligent Laboratory Systems*, 8, 1–3, 37–52
- S. Wold, N. Kettaneh-Wold, J. F. MacGregor, K. G. Dunn, 2009, Batch Process Modeling and MSPC, In: *Comprehensive Chemometrics*, 2, 163–197
- S. Yoon, J. F. MacGregor, 2001, Incorporation of External Information into Multivariate PCA/PLS Models, *IFAC Proceedings Volumes*, 34, 27, 105–110

# Optimal Determination of the Binary Interaction Parameters for the Vapor-Liquid-Liquid Equilibrium on the Vinyl Acetate/Acetic Acid/Water System

Jose Alfredo Paredes-Ortiz <sup>a</sup>, Úrsula Fabiola Rodríguez-Zúñiga <sup>b</sup>, Fernando Israel Gómez-Castro <sup>a\*</sup>.

<sup>a</sup> *Departamento de Ingeniería Química, División de Ciencias Naturales y Exactas, Campus Guanajuato, Universidad de Guanajuato, Noria Alta S/N, Guanajuato, Gto. 36050, México. [fgomez@ugto.mx](mailto:fgomez@ugto.mx)*

<sup>b</sup> *Universidad de Ingeniería y Tecnología, Departamento de Ingeniería Química, Jr. Medrano Silva 165, Barranco, Lima, Perú, 15063.*

## Abstract

The design of the vinyl acetate purification process requires modelling the vapor-liquid-liquid equilibrium (VLLE) of the mixture vinyl acetate/acetic acid/water. In this work, the equilibrium of the system is represented through the gamma-phi formulation using the Hayden O'Connell (HOC) model for the vapor phase and the universal quasi-chemical (UNIQUAC) model for the liquid phases. The parameters for the model are estimated through optimization tools, using the BARON global optimization algorithm. The objective function to be minimized was the least-squares function. The results show a good correlation for the LLE and VLE of vinyl acetate with water with errors less than 2%. In the case of the VLE equilibrium of water with acetic acid the error is close to 5% with a larger deviation at acid mole fractions above 0.4. The parameters determined allow a proper modeling of the VLLE for the analyzed mixture.

**Keywords:** Binary interaction parameters, vapor-liquid-liquid equilibrium, deterministic optimization

## 1. Introduction

Vinyl acetate is one of the most widely produced chemicals in the world. It is used in the production of polyvinyl acetate and polyvinyl alcohol. At present, the main route for vinyl acetate synthesis is via the acetoxylation reaction of ethylene in the presence of oxygen and palladium catalysts. During the vinyl acetate synthesis process, the mixture vinyl acetate/acetic acid/water is generated; this mixture is commonly separated by distillation. Over the last few years, several authors have conducted research related to the improvement of the vinyl acetate process focusing on both the design and control areas, with particular emphasis on the purification of the mixture. Nevertheless, the mixture presents two liquid phases, which may difficult the separation. On the last years, authors such as Li et al. (2022) and Xie et al. (2020) point to the use of intensified separation operations such as extractive distillation or azeotropic distillation to achieve the purification task.

The vinyl acetate purification by intensified processes involves modelling the VLLE of the mixture. However, this implies that the binary interaction parameters of the equilibrium models must properly predict both vapor-liquid equilibrium (VLE) and liquid-liquid equilibrium (LLE). Unfortunately, the currently available binary interaction parameters accurately describing the VLE show inaccurate predictions for the LLE, and vice versa. The data fitting of these models for the representation of vapor-liquid-liquid equilibrium is a complex task with a high probability for the solution to get stuck in a local optimum. Some authors propose global optimization strategies based on minimization of squares to solve this problem (e.g., Wyczesany, 2014). Therefore, in the present work, a least-squares optimization was performed using the BARON global optimization algorithm to obtain the binary interaction parameters that properly represent VLE and LLE for the mixture vinyl acetate/acetic acid/water.

## 2. Methodology

For mixtures with carboxylic acids, as is the case due to the presence of acetic acid, the HOC model describes the behavior quite well since the carboxylic acid interacts with itself or with other compounds in vapor phase due to dimerization. In addition to this model, the NRTL or UNIQUAC models proposed by Hsieh et al. (2008) are often used for this type of mixtures due to the presence of non-ideal solutions. Thus, in this work, the UNIQUAC-HOC combination has been used.

The Gamma-Phi approach is used to model the vapor-liquid equilibrium, as represented in equation 1.

$$\phi_i^V y_i P = x_i \gamma_i \phi_i^S P_i^S \quad (1)$$

Where the left-hand terms are the vapor phase fugacity of the component, the mole fraction of the component in vapor phase, and the pressure. On the right-hand side the terms appearing are the liquid phase mole fraction, the activity coefficient, the saturation fugacity, and the saturation pressure, all for the respective component  $i$ .

For the calculations of the vapor phase, the relationship between the apparent fugacity coefficients and the real fugacity coefficients is used, as well as the true mole fraction of the vapor phase, based on the chemical theory that also complements and uses the equation of state HOC, equation 2.

$$\phi_i^V y_i = z_i \phi_i^\# \quad (2)$$

Equation 2 corrects for the deviation generated due to dimer formation. In this case it is used because acetic acid has a predominance in dimer formation.

The Gamma-Gamma representation has been used for the liquid-liquid equilibrium, this is shown in equation 3, where the upper indices refer to each of the liquid phases.

$$x_i^I \gamma_i^I = x_i^{II} \gamma_i^{II} \quad (3)$$

The binary interaction parameters of the UNIQUAC model are represented with an independent and a temperature-dependent parameter (equation 4).

$$\tau_{ij} = \exp\left(a_{ij} + \frac{b_{ij}}{T}\right) \quad (4)$$

For the correction of the binary interaction parameters, the experimental data reported by Shanghai College Chemical Engineering (1976) and Zhang et al. (2011) has been used for the liquid-vapor equilibrium modeling. Complementarily, the data reported by Gao et al. (2017) has been employed for the liquid-liquid equilibrium correlation.

The least squares ratio for the experimental and the estimated data is established as the objective function, equation 5.

$$F = \sum_{i=1}^{N_{LLE}} \sum_{j=1}^{N_C} \sum_{k=1}^2 (x_{e_{k,j,i}} - x_{t_{k,j,i}})^2 + \sum_{i=1}^{N_{VLE}} \sum_{j=1}^{N_C} (y_{e_{j,i}} - x_{t_{j,i}})^2 + \sum_{i=1}^{N_{VLE}} (T_{e_i} - x_{t_i})^2 \quad (5)$$

The objective function refers to the number of experimental points reported ( $i$ ), the number of components ( $j$ ) and the number of phases ( $k$ ), organic and aqueous.

The model was run through the GAMS interface using the Global BARON optimization algorithm created by Sahinidis (1996) through the NEOS server, Hosted by the Wisconsin Institute for Discovery at the University of Wisconsin in Madison.

### 3. Results and Discussion

Based on the experimental data reported by Shanghai College Chemical Engineering (1976), Zhang et al. (2011) and Gao et al. (2017), and the adjustment of the interaction parameters, the following results have been obtained.

#### 3.1. LLE Vinyl Acetate/Acetic Acid/Water

As shown in Figure 1, favorable results have been obtained with respect to the liquid-liquid equilibrium of the mixture. The curve fit with the model is very close to the data reported by Gao et al. (2017), as observed from the deviation data shown in Table 1.

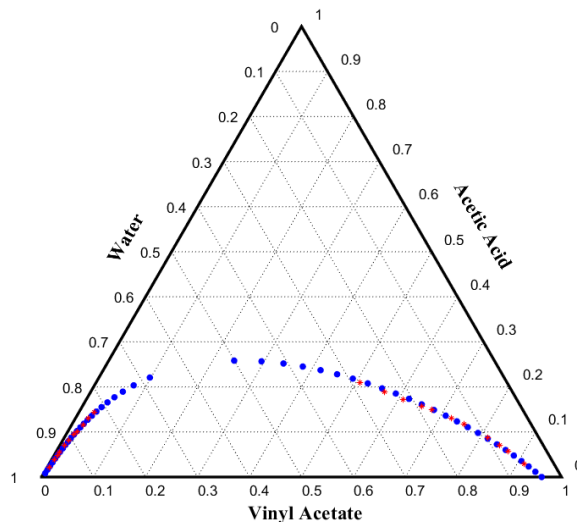


Fig 1. LLE for the ternary system vinyl acetate/ acetic acid / water: (red) experimental data at 298.15 K, (blue) estimated values.

Table 1. Deviation of the liquid-liquid equilibrium of the Vinyl Acetate (1) / Acetic Acid (2)/ Water (3) mixture at 298 K y 1 atm.

$\Delta x_i$						
$\Delta T$	(1) <sup>I</sup>	(2) <sup>I</sup>	(3) <sup>I</sup>	(1) <sup>II</sup>	(2) <sup>II</sup>	(3) <sup>II</sup>
5.29E-01	-3.16E-04	-2.26E-04	5.42E-04	-6.85E-04	1.58E-03	-8.95E-04
-2.34E-01	2.23E-04	-6.12E-04	3.89E-04	3.83E-04	-1.35E-03	9.72E-04
1.69E-01	-5.22E-05	-9.17E-04	9.69E-04	-2.87E-03	4.46E-05	2.83E-03
-1.27E-01	1.20E-04	5.66E-04	-6.86E-04	1.92E-03	-2.63E-03	7.04E-04



-4.17E-01	4.82E-04	1.08E-03	-1.56E-03	8.62E-03	-6.67E-03	-1.95E-03
-1.52E-01	2.26E-04	3.35E-04	-5.61E-04	3.77E-03	-2.87E-03	-9.05E-04
1.01E-01	-2.27E-04	-2.84E-03	3.06E-03	-6.30E-03	6.76E-03	-4.61E-04
-1.94E-02	-4.08E-07	2.89E-04	-2.88E-04	4.33E-04	1.32E-03	-1.75E-03
9.04E-02	-3.10E-04	1.16E-03	-8.49E-04	-2.88E-03	3.86E-03	-9.80E-04
9.03E-02	-2.33E-04	-1.30E-03	1.53E-03	-5.96E-03	7.06E-03	-1.10E-03
-3.15E-02	3.84E-04	-4.17E-03	3.78E-03	2.69E-03	2.39E-03	-5.08E-03

### 3.2. VLE Vinyl Acetate/Acetic Acid

A good correlation has been obtained between the model and the experimental data for the mixture of vinyl acetate with acetic acid (Figure 2), obtaining an adequate prediction, as shown in Table 2 in terms of the deviations.

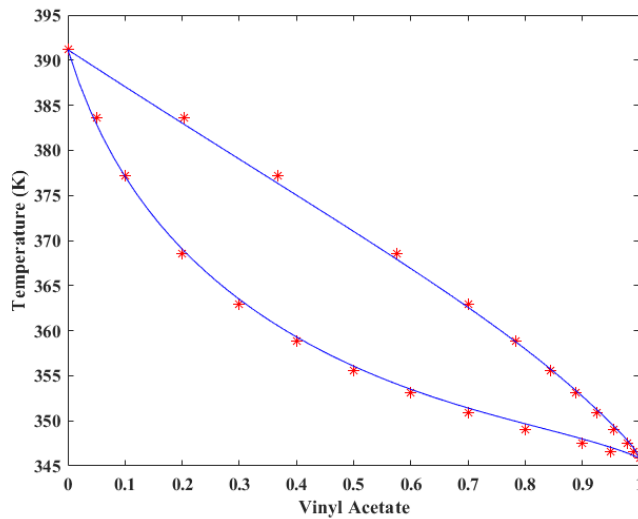


Fig 2. VLE of vinyl acetate/ acetic acid at 1 atm; (red) experimental data, (blue) estimated values.

Table 2. Deviation of the vapor-liquid equilibrium of the Vinyl Acetate (1) / Acetic Acid (2) at 1 atm.

$\Delta T$	$\Delta x_{(1)}$	$\Delta x_{(2)}$	$\Delta y_{(1)}$	$\Delta y_{(2)}$
-5.78E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00
-6.60E-01	-2.02E-05	2.02E-05	-1.03E-03	1.03E-03
-1.13E-01	2.69E-05	-2.69E-05	-1.99E-02	1.99E-02
4.80E-01	1.78E-04	-1.78E-04	-2.80E-02	2.80E-02
4.96E-01	4.17E-04	-4.17E-04	-1.97E-02	1.97E-02
4.26E-01	5.37E-04	-5.37E-04	-1.15E-02	1.15E-02
3.85E-01	6.37E-04	-6.37E-04	-6.53E-03	6.53E-03

3.33E-01	2.61E-04	-2.61E-04	-2.67E-03	2.67E-03
4.51E-01	2.05E-04	-2.05E-04	-2.67E-03	2.67E-03
5.79E-01	1.75E-04	-1.75E-04	-2.98E-03	2.98E-03
3.91E-01	1.23E-04	-1.23E-04	-2.38E-03	2.38E-03
3.52E-01	5.14E-05	-5.14E-05	-9.35E-04	9.35E-04
7.39E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00

### 3.3. VLE Acetic Acid/Water

The fit between the model and the experimental data in Figure 3 is good. However, the results presented in Figures 1 and 2 present a better fit to the data. The higher the acetic acid composition, the larger the deviation. The error with respect to the experimental points is detailed in Table 3. Finally, the computed parameters are presented in Table 4.

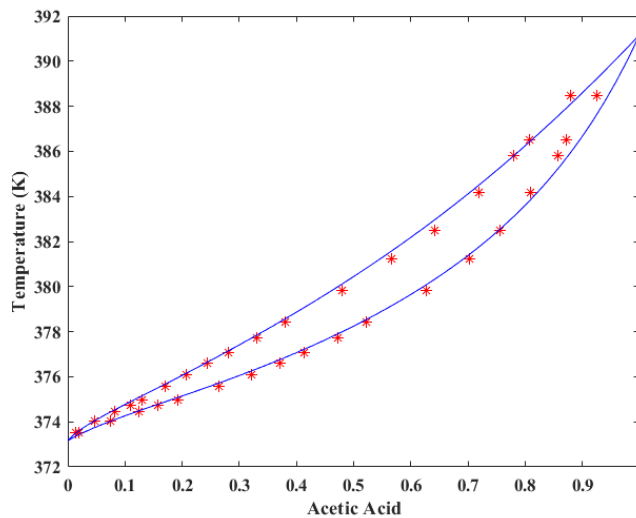


Fig 3. VLE of acetic acid/water at 1 atm; (red) experimental data, (blue) estimated values.

Table 3. Deviation of the vapor-liquid equilibrium of the Acetic Acid (1) / Water (2) at 1 atm.

$\Delta T$	$\Delta x_{(1)}$	$\Delta x_{(2)}$	$\Delta y_{(1)}$	$\Delta y_{(2)}$
-5.00E-02	6.85E-05	-6.85E-05	-4.29E-03	4.29E-03
5.43E-03	3.32E-05	-3.32E-05	-1.92E-03	1.92E-03
5.93E-02	4.98E-05	-4.98E-05	-3.09E-03	3.09E-03
4.25E-02	1.37E-04	-1.37E-04	-8.86E-03	8.86E-03
9.35E-02	9.54E-05	-9.54E-05	-6.27E-03	6.27E-03
1.13E-01	-4.11E-05	4.11E-05	2.07E-03	-2.07E-03
1.54E-01	-1.38E-04	1.38E-04	7.22E-03	-7.22E-03
1.43E-01	-1.54E-04	1.54E-04	7.80E-03	-7.80E-03
1.39E-01	-1.14E-04	1.14E-04	5.15E-03	-5.15E-03
1.75E-01	-1.04E-04	1.04E-04	3.26E-03	-3.26E-03

1.04E-01	1.53E-05	-1.53E-05	-3.13E-03	3.13E-03
2.20E-01	-3.31E-05	3.31E-05	-2.83E-03	2.83E-03
1.91E-01	-4.26E-06	4.26E-06	-9.44E-03	9.44E-03
5.49E-02	1.18E-04	-1.18E-04	-2.30E-02	2.30E-02
-2.79E-01	2.01E-04	-2.01E-04	-3.13E-02	3.13E-02
-5.33E-01	1.71E-04	-1.71E-04	-2.53E-02	2.53E-02
-7.34E-01	2.12E-04	-2.12E-04	-3.11E-02	3.11E-02
-7.49E-01	1.12E-04	-1.12E-04	-1.78E-02	1.78E-02

Table 4. UNIQUAC binary interaction parameters of the Vinyl Acetate (1) / Acetic Acid (2)/ Water (3) mixture.

		$A_{IJ}$	$B_{IJ}$	$A_{JI}$	$B_{JI}$
(1)	(2)	-4.118	936.916	-0.554	416.761
(1)	(3)	-10.000	2469.161	7.513	-2500.000
(2)	(3)	5.724	-2500.000	-3.855	1675.908

#### 4. Conclusions

The results of the methodology employed show that the UNIQUAC equation, combined with the HOC equation of state, can properly describe the LLE and VLE of the ternary mixture vinyl acetate/acetic acid/water with adequate accuracy if the parameters are simultaneously fitted for both equilibria. The VLE of acetic acid with water has a higher deviation, up to  $3.13 \times 10^{-2}$  for the vapor composition of acetic acid. Nevertheless, the parameters obtained allow to adequately represent the ELLV of the evaluated mixture. This data will be useful for the simulation and analysis of intensified processes for the separation of the mixture vinyl acetate/acetic acid/water in the vinyl acetate production process.

#### References

- J. Gao, D. Guan, D. Xu, L. Zhang, Z. Zhang, 2017, Measurement and modeling of liquid-liquid equilibrium for the systems vinyl Acetate + acetic acid/ethanol + water at 298.15 and 308.15 K, *Journal of Chemical and Engineering Data*, 62, 4, 1240–1246.
- C.-T. Hsieh, W.-Y. Ji, H.-m Lin, M.- J. Lee, 2008, Multiphase equilibria for mixtures containing water, acetic acid, propionic acid, methyl acetate and methyl propionate, *Fluid Phase Equilibria*, 271, 1–2, 69–75.
- Z. Li, K. Wang, X. Luo, X. Xu, 2022. Optimization and simulation of vinyl acetate process based on Aspen Plus, *Proceedings of the International Conference on Optoelectronic Materials and Devices*, 121641U.
- Shanghai College Chemical Engineering, 1976, Study on the vapor liquid equilibrium data for acetic acid-water-vinyl acetate ternary system- I.The mutually miscible liquid phase region, *Acta Chimica Sinica*, 34, 79–93.
- Wyczęsany, A. 2014. Calculation of vapor-liquid-liquid equilibria at atmospheric and high pressures. *Industrial and Engineering Chemistry Research*, 53(6), 2509–2519.
- Y.N. Xie, Z. Meng, F. Li, 2020, Purification process design of vinyl acetate based on Aspen Plus, *IOP Conference Series: Earth and Environmental Science*, 545.
- C. Zhang, H. Wan, L. Xue, G. Guan, 2011, Investigation on isobaric vapor liquid equilibrium for acetic acid + water + (n-propyl acetate or iso-butyl acetate), *Fluid Phase Equilibria*, 305, 1, 68–75.

# Assessment of subgrid-scale models for large-eddy simulation of a gas-liquid stirred reactor

Safae Elmisaoui,<sup>a,b</sup> Radouan Boukharfane,<sup>a</sup> Lhachmi Khamar,<sup>a,c</sup> and Jean-Michel Ghidaglia<sup>a,d</sup>

<sup>a</sup> Mohammed VI Polytechnic University, MSDA Program, Benguerir 43150, Maroc

<sup>b</sup> Mohamed V University, Laboratoire du génie civil et environnement, 11060, Maroc

<sup>c</sup> Sultan Moulay Slimane University, LIPIM, ENSA Khouribga, 25000, Maroc

<sup>d</sup> Centre Borrelli, ENS Paris-Saclay, CNRS, Université Paris-Saclay, 94235 Cachan, France

Radouan.boukharfane@um6p.ma

## Abstract

The LES-SGS simulations of a three-dimensional multiphase flow in an industrial preneutralizer reactor is performed, and the Multiple Reference Frame (MRF) approach is used. The effectiveness of the three SGS turbulent closures (Smagorinsky–Lilly model, Wall Adapting Local Eddy Viscosity (WALE) model, and the Algebraic Wall Modeled LES (WMLES) model) to predict the flow hydrodynamics and to capture the eddies near to high turbulent zones (in liquid injection point and in the turbine rotation zone) is assessed. The results show that the SGS model has a significant impact not only on the instantaneous flow field, but also on the time-average velocity magnitude, therefore the hydrodynamics of unsteady behaviour in turbulent chemical reactors is found to be significantly influenced by the choice of the SGS closure model.

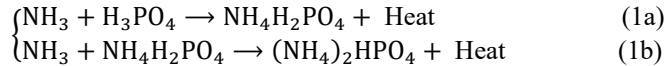
**Keywords:** CFD, LES- SGS, turbulence model, hydrodynamics, multiphase flow

## 1. Introduction

In the phosphate industry, Di–Ammonium Phosphate (DAP) is considered to be a very efficient chemical fertilizers that is used to provide the required nutrients to the plants and soil. The DAP manufacturing processes begin with a preneutralizer reactor that consists of the production of homogeneous slurry produced by mixing the Ammonia gas and Phosphoric acid (ACP), under specific conditions. The preneutralizer reactor is a non-standard stirred tank reactor, equipped with a pitched blades turbine agitator that (i) promotes mass and heat transfer, (ii) reduces foaming, and (iii) improves ammonia dispersion in the liquid phase. This unit operation represents the masterpiece that conducts the physicochemical phenomenon called "preneutralization" chemical reaction (Elmisaoui et al., 2020).

The reaction between ammonia and ACP in the reactor leads to MAP or DAP depending on the ratio of N (Nitrogen) and P (Phosphorus) ratio N/P. Other species are formed but in small quantities in such a way that they do not significantly affect the total nitrogen or P<sub>2</sub>O<sub>5</sub> in the final product (Campbell et al., 2006). When the ratio N/P is greater than unity,

the produced MAP in the first reaction is converted to DAP in the second one (Campbell et al., 2006). The overall reactions are listed as follows



These reactions occur immediately after Ammonia gas  $\text{NH}_3$  and ACP( $\text{H}_3\text{PO}_4$ ) are mixed. Produced slurry feeds the granulator, to be sprayed onto the granular fertilizer bed existing in the granulator, and granule growth occurs. Granules leaving the granulator are first dried and then screened to separate out the product size.

For the CFD modeling studies of multiphase flows in stirred reactors, different turbulence models have been developed to predict the hydrodynamics of the flow, each one of them demonstrating its effectiveness and robustness in each particular application (Derksen, 2001). The age-old CFD modeling method based on the Reynolds Averaged Navier-Stokes (RANS) approach is still considered the most employed turbulence modeling approach due to its acceptable accuracy and affordable computational cost for predicting the hydrodynamics involving complex geometries, but with the development of the HPC capabilities, the employment of the hybrid Detached Eddy Simulation (DES) model, which blends the RANS approach with LES finds its applications. The DES methodology shows its increasing potential and accuracy in flow prediction more than the RANS models. Especially for its informative ability to predict the flow regimes in the moving zone characterizing the impeller rotation. This work aims at evaluating the performance of three main Large-Eddy Simulation (LES) subgrid-scale (SGS) turbulence models in a stirred reactor.

## 2. Physical model of the preneutralizer

In this study, numerical simulations have been carried out using the workbench of the CFD commercial software ANSYS Fluent 2021-R2. The physical model tailored in this work is based on the actual size of a preneutralization chemical reactor (PN) with a working volume of  $47 \text{ m}^3$  to ensure throughput of around  $120 \text{ m}^3/\text{hr}$ . The ACP is injected from the top, and eight uniform Ammonia inlet nozzles are arranged in the bottom, which are 3m away from it. The rotational agitator is equipped with a Pitch Blade Turbine (PBT), consisting of four impellers slanted with an angle of  $45^\circ$  from the horizontal plane. Detailed characteristics of the CAD file (cf. Fig. 1) are based on the original preneutralizer geometry and are previously described in Elmisaoui et al. (2022).

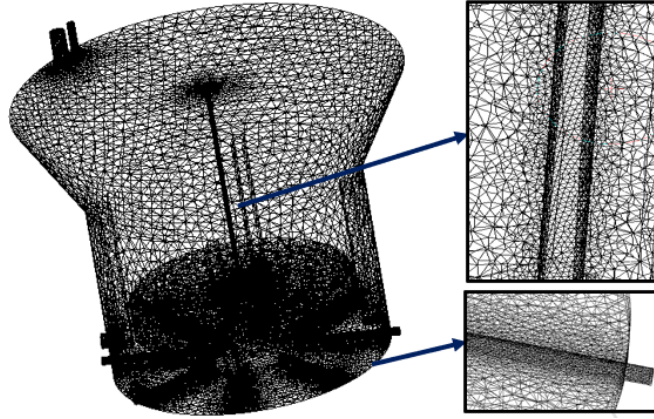


Figure 1: Mesh grid of the simulation domain, with zoom in shaft axis and gas spargers

### 3. Mathematical modeling and numerical schemes

LES comes in between the Direct Numerical Simulation (DNS) and the Reynolds Averaged Navier-Stokes (RANS) approaches in terms of turbulent scale resolution. Large eddies are directly resolved in LES, whereas small eddies are modeled. The time-dependent Navier-Stokes equations are filtered. The space-filtered equations for the conservation of mass and momentum of an incompressible Newtonian fluid can be written as:

$$\begin{cases} \frac{\partial \bar{u}_i}{\partial x_i} = 0 & (2a) \\ \frac{\partial \bar{u}_i}{\partial t} + \bar{u}_i \frac{\partial \bar{u}_j}{\partial x_i} = -\frac{\partial \bar{P}}{\partial x_i} + \nu \frac{\partial^2 \bar{u}_i}{\partial x_j^2} - \frac{\partial \tau_{ij}}{\partial x_j} & (2b) \\ \tau_{ij} = \bar{u}_i \bar{u}_j - \bar{u}_i \bar{u}_j & (2c) \end{cases}$$

where  $\tau_{ij}$  is the sub-grid scale stress tensor, which reflects the effect of the unresolved scales on the resolved scales. Subgrid-scale stresses resulting from the filtering operation are unknown, and require modeling. The subgrid-scale turbulence models employ the Boussinesq hypothesis as in the RANS models, computing subgrid-scale turbulent stresses from:

$$\tau_{ij} = \frac{1}{3} \tau_{kk} \delta_{ij} - 2\mu_t \bar{S}_{ij} \quad (3)$$

Three SGS turbulence models are considered in this work:

- Smagorinsky-Lilly model
 
$$\mu_t = \rho L_s^2 |\bar{S}| \quad (4)$$
- Wall Adapting Local Eddy Viscosity (WALE) model

$$\mu_t = \rho L_s^2 \frac{(s_{ij}^d s_{ij}^d)^{3/2}}{(s_{ij}^d s_{ij}^d)^{5/2} + (s_{ij}^d s_{ij}^d)^{5/4}} \quad (5)$$

- Algebraic Wall-Modeled LES Model

$$\mu_t = \rho m [(\kappa d_{\text{wall}})^2, (C_k \Delta)^2] \cdot S \cdot \{1 - \exp(-(y^+ / 25))\} \quad (6)$$

where  $d_{\text{wall}}$  is the wall distance,  $S$  is the strain rate,  $\kappa = 0.41$ , and  $C_k = 0.2$  are constants.

To consider chemical species transport, the previous system of equations is coupled to the species transport equation, as described in Elmisaoui et al. (2022).

### 3.1 Numerical simulation conditions

As previously stated in many CFD simulations of stirred reactors, a variety approaches may be used to cope with the movement of the impeller blades. The Multiple Reference Frame (MRF) approach was employed in this investigation Elmisaoui et al. (2022), and the SIMPLE algorithm was employed for pressure-velocity coupling.

### 3.1 Computational domain

The computational domain grid is composed of two parts: an inner spinning cylindrical volume holding the turbine and an outer, stationary volume containing the remainder of the tank. In both portions, structured grids of non-uniformly distributed hexahedral cells are utilized. To provide a more precise description of the impeller, the grid employed in the impeller region is refined. In the tank, there are a total of 970997 grid nodes.

It is worth noting that a grid-dependency study is carried out using ANSYS Mesher to evaluate mesh suitability for the present configuration. A three-dimensional mesh grid of the computational domain is discretized into unstructured cells with specific refined zones (cf. Fig. 1). Three mesh featuring grids are generated and defined as coarse, medium, and fine meshes, respectively. Refinements are carried out in the shaft region, near to the nozzles, at the inlets, and at the outlet of the domain, allowing for a minimum and maximum mesh size of 25 mm and 15 mm for the medium mesh grid. The last refinement level is retained based on a mesh grid sensitivity, as the grid allows for an accurate numerical solution at acceptable time costs.

## 4. Results and analysis

To evaluate the flow hydrodynamics in a three-dimensional multiphase turbulent flow, the identification of coherent structures is crucial, which allows a better understanding of the real space dynamics of turbulent flows. One of the most commonly used methods for categorizing three-dimensional flow structures is the Q-criterion proposed by Hunt et al. (1988). It is defined in terms of the instantaneous velocity gradient tensor as:

$$\mathbf{Q} = \mathbf{1}/2((\nabla \cdot \mathbf{u})^2 - \nabla \mathbf{u} : \nabla \mathbf{u}^T) = \mathbf{1}/2((\nabla \cdot \mathbf{u})^2 + \|\boldsymbol{\Omega}\|_2^2 - \|\mathbf{S}\|_2^2) \quad (7)$$

Here,  $S$  is the strain rate tensor and  $\Omega$  is the rotation rate tensor. The turbulence associated with the flow through the inlet and across the stirred reactor is visualized at two different instants  $t = 100$  s and  $t = 190$  s in Figs. 2 and 3. As can be noted for the three SGS-

models and as the flow is introduced from the nozzles into the tank, it is clear to have higher velocities near the inlet pipes. In addition, there is a velocity difference between the nozzle stream and the rotating flow in the tank, which creates local shear zones. As a consequence, inlet nozzles become the inherent turbulence generators. The intensity of the generated turbulence is less pronounced in the classical Smagorinsky model, while it seems to be more pronounced in the WALE and WMLES models.

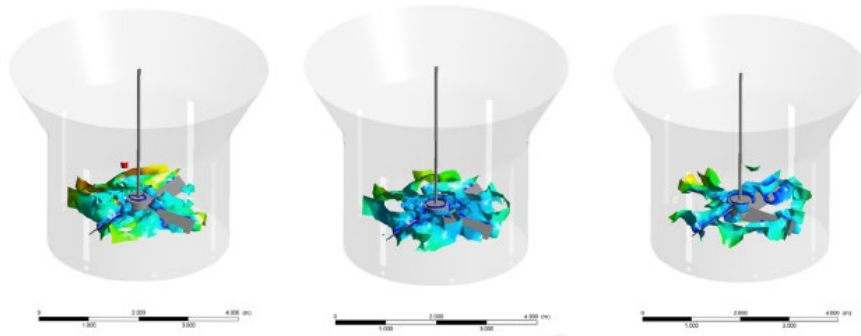


Figure 2: Turbulence structures for (left to right) Smagorinsky–Lilly model, WALE model, and WMLES model, respectively at  $t = 100$  s

Furthermore, for the latter models, the instantaneous flow seems to be characterized by nearly concentric streamlines in a circular fluid motion, which is due to tangential inflow. The resulting vortex column around the central outlet structure is wrapped by the external turbulent filaments, leading to intensive momentum transport.

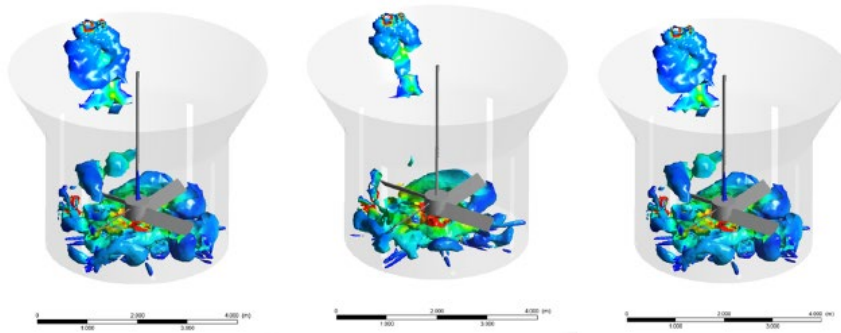


Figure 3: Turbulence structures for (left to right) Smagorinsky–Lilly model, WALE model and WMLES model, respectively at  $t = 190$  s

Figure 5 presents the profiles of instantaneous magnitude velocities in two locations of the domain (denoted by P1 and P2 in Fig. 4).



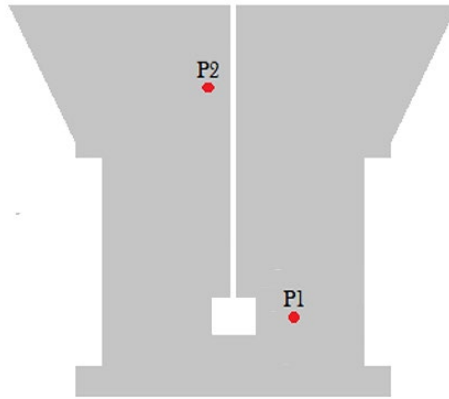


Figure 4: Axial projection plan in the middle of the domain

As a considerable throughput of liquid is injected near to position P2, a high velocity peak is observed in a short time. Then, as the mixing gets enhanced, the velocity decreases gradually until  $t = 150$  s and a considerable flow amount of Ammonia gas achieves the top of the domain. This explains the emergence of new turbulent eddies in the P2 location.

LES models were able to capture the generated eddies coming from the liquid injection and from the mixing zone. WALE model gives higher values of the velocity magnitude than WMLES and Smagorinsky-Lilly models.

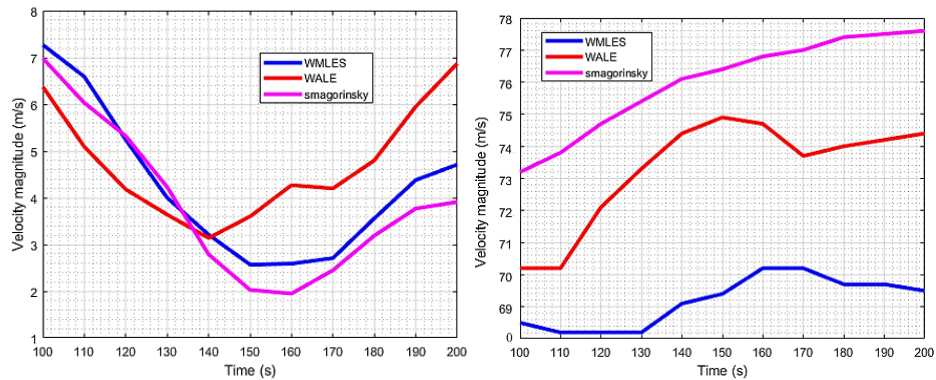


Figure 5: (a) & (b) Temporal evolution of velocity magnitude profiles in probes P1 and P2, respectively

However, based on the P2 probe position near to the shaft wall, the WALE is more appreciated to be used for near to the wall turbulence eddies prediction. As shown in Fig. 5b, the predicted velocity magnitude profiles fastly increase in the rotation zone (P2 probe). Strong eddies are developed by mixing, and interesting redistribution of the reactants is ensured by the circulation loop, which amplifies the magnitude of velocity. The occurrence of the gas flow from the nozzles with a considerable injection velocity considerably enhances the eddies and amplifies the circulation loops. The Smagorinsky-

Lilly model seems to capture the flow hydrodynamics with a good precision compared to WMLES and WALE models.

## **5. Conclusion**

In this paper, the large eddy simulation method based on an Eulerian–Eulerian CFD model is used to simulate hydrodynamics of the gas–liquid reacting flow in the preneutralizer stirred reactor. The liquid and gas phase flow fields in the stirred tank simulated by the LES approach are more asymmetric in the case of WMLES and WALE models. Overall, each of the three models has its own specific ability for the flow prediction. However, the choice of an LES Turbulence model should be done carefully, for a good physical flow prediction.

## **References**

- G. R. Campbell, Y.-K. Leong, C. C. Berndt, J. L. Liow, 2006. Ammonium phosphate slurry rheology and particle properties—the influence of Fe (III) and Al (III) impurities, solid concentration and degree of neutralization. *Chemical Engineering Science* 61 (17), 5856–5866.
- J. Derksen, 2001. Assessment of large eddy simulations for agitated flows. *Chemical Engineering Research and Design* 79 (8), 824–830.
- S. Elmisaoui, S. Elmisaoui, S. Benjelloun, L. Khamar, M. Khamar, 2022. Cfd investigation of industrial gasliquid preneutralizer based on a bioreactor benchmark for spargers optimization. *Chemical Engineering Transactions* 93, 73–78.
- S. Elmisaoui, L. Khamar, S. Benjelloun, M. Khamar, J. Ghidaglia, 2020. Modeling and study of hydrodynamic flow within the preneutralizer reactor using cfd approach. In: 30th European Symposium on Computer Aided Chemical Engineering. Elsevier, pp. 103–108.
- J. Hunt, S. Leibovich, K. Richards, 1988. Turbulent shear flows over low hills. *Quarterly Journal of the Royal Meteorological Society* 114 (484), 1435–1470.



# Mechanistic model for inkjet printing and applications in perovskite solar cells engineering

Naveen Bhati<sup>a,\*</sup>, Mohammad Khaja Nazeeruddin<sup>b</sup>, and François Maréchal<sup>a</sup>

<sup>a</sup>*Industrial Process and Energy Systems Engineering, Ecole Polytechnique Fédérale de Lausanne, Sion, 1951, Switzerland*

<sup>b</sup>*Group for Molecular Engineering of Functional Materials, Institute of Chemical Sciences and Engineering, Ecole Polytechnique Fédérale de Lausanne, Sion, 1951, Switzerland*

\**naveen.bhati@epfl.ch*

## Abstract

Perovskite solar cells (PSCs) are the most exciting third-generation PV technology, which has shown an unprecedented increase in power conversion efficiencies from < 4% to > 25% in just over a decade. However, most of the research results indicating efficiencies of more than 20% are based on spin coating techniques with small cells with an active area < 20mm<sup>2</sup>. With efficiency already been on par with established silicon solar cells, there is a need to upscale the technology to production scale and parallelly work on other facets of the technology like stability and toxicity. Among such techniques for mass production of PSCs is well-established inkjet printing, which has been explored to achieve PSC printing and can be used with various substrates and solvents. However, there is not much effort in modelling the inkjet printing process and understanding the impacts of different decision variables related to the process. Here, this article presents a mechanistic model to achieve perovskite printing (or entire stack of layers) using this technique and analyzing the influence of solvent selection and process parameters on film characteristic.

**Keywords:** Perovskite solar cells, inkjet printing, fabrication techniques, modelling.

## 1. Introduction

Inkjet printing is an attractive technology not only for research level but also for scaling the perovskite solar cells fabrication because of low cost, high production throughput and high material efficiency. Further, the non-contact nature of printing helps in depositing solutions on variety of materials. The technique has been discussed in literature with focus on droplet formation mechanism, travel to substrate and droplet interaction with substrate (Derby, 2010; Kang et al., 2020; Soltman and Subramanian, 2008). However, most of the studies in simulation regime were based on computational fluid dynamics which is time-consuming and resource intensive (Tofan et al., 2021; van der Bos et al., 2014) with less focus on process dynamics of the process. With the ever-increasing applications of inkjet printing in printable energy technologies like organic solar cells, polymer solar cells, fuel cells, energy storage devices, etc., it is important to understand the effect of different variables involved in the process on the final film quality.

Perovskite solar cells have shown unparalleled improvement in their efficiency over the last decade and have reached the levels of the established Si cells. Solution-based deposition methods have shown better results compared to their vacuum-based counterparts not only in efficiency but also provides cost effective, easily controllable, and scalable solutions. Among these solution-based deposition methods, inkjet printing

has gained significant interest in the community where drop-by-drop printing leads to complete formation of film with control over droplet spacing and morphology via solvent engineering (Yang et al., 2022). The power conversion efficiency (PCE) of inkjet-printed devices has risen from 11.6% (Wei et al., 2014) to 21.6% (Abzieher et al., 2019) since its first application. Moreover, many researchers have also shown the complete fabrication of PSCs except the transparent conducting oxide layer using inkjet printing technique which shows the potential of complete fabrication of perovskite solar cells based on inkjet printing (Gao et al., 2021). However, the dynamics of inkjet printing is more complex as compared to other solution-based techniques like spin-coating (0-D) or slot-die coating (1-D) processes. Therefore, it is crucial to understand the influence of different parameters on the inkjet printing process to obtain homogeneous liquid films and to control the morphology of the final dry film (Huckaba et al. 2019, Mathies et al., 2016).

## 2. Model Description

The entire process of inkjet printing can be divided into four stages:

### 2.1. Droplet generation

Inkjet printing technique can be used only with certain solvents or solutions depending on the solution properties and is governed by coating window which can be represented in terms on non-dimensional numbers Eq. 1-3 (Derby, 2010).

$$0.1 \leq Oh (= 1/Z) \leq 1; \quad Oh\text{-Ohnesorge number} \dots \dots \dots \text{Eq. 1}$$

$$We > 4; \quad We\text{-Weber number} \dots \dots \dots \text{Eq. 2}$$

$$We^{1/2} Re^{1/4} \leq 50; \quad Re\text{-Reynolds number} \dots \dots \dots \text{Eq. 3}$$

### 2.2. Droplet travel to substrate

Once the drop is generated and leaves the nozzle, the set of differential equations describing the system in second stage include Eq. 4-7 where  $C_D$  is based on Kunii and Levenspiel(1991), velocity,  $v$  is based on Sloth(2007) and  $k_m$  is mass transfer coefficient.

$$\frac{dv}{dt} = \left(1 - \frac{\rho_a}{\rho_d}\right) g - \frac{C_D \rho_a v^2 \pi r_d^2}{2m_d} + \frac{F_t}{m_d}; \quad C_D\text{-Drag coefficient, } m_d\text{:drop mass} \dots \dots \dots \text{Eq. 4}$$

$$\frac{dm_{\text{solvent}}}{dt} = \left(-k_m 4\pi r_d^2 \frac{M_{\text{solvent}} P_{\text{amb}}}{RT_{\text{avg}}}\right) (y_{v,\text{surface}} - y_{v,\text{amb}}); y_{v,\text{mole fraction in vapor}} \text{Eq. 5}$$

$$\frac{dr_d}{dt} = \frac{1}{4\pi r_d^2 \rho_{\text{solvent}}} \frac{dm_{\text{solvent}}}{dt}; \quad r_d\text{-droplet radius, } t\text{-time, } \rho\text{-density} \dots \dots \dots \text{Eq. 6}$$

$$\frac{dT_d}{dt} = \frac{h_c [4\pi r_d^2 (T_{\text{amb}} - T_d)] + \Delta H_v \frac{dm_{\text{solvent}}}{dt}}{m_{\text{solute}} c_{p,\text{solute}} + m_{\text{solvent}} c_{p,\text{solvent}}}; \quad T\text{-temperature, } h_c\text{-heat transfer coefficient} \dots \dots \text{Eq. 7}$$

The thermophoretic force component,  $F_t$  is a dominant force especially close to the surface when the temperature of substrate is very high (Filipovic et al., 2013).

### 2.3. Droplet impingement and positioning

Based on Stringer and Derby (2010), once the droplet reaches the surface, the equilibrium diameter of the droplet ( $D_{\text{eq,d}}$ ) can be expressed in terms of diameter of droplet at impingement and the equilibrium contact angle ( $\theta_{\text{eqb}}$ ). For stable line formation without any defects like waviness or bulging, criteria for maximum and minimum droplet spacing is used. Further, width of the bead is deduced based on Stringer and Derby (2010). Afterwards, the line formed is assumed to be a flat rectangle with height depending on the width and total volume of the bead from initial state of flat cylinder for single droplet.

### 2.4. Film formation

The system of differential equations for tracking the solution parameters during line/film stages can be expressed as given in Eq. 8-10 where  $\Delta H_v$  is enthalpy of vaporization.

$$\frac{dm_{\text{solvent}}}{dt} = \left[ -k_m(A_{\text{eff}}) \frac{M_{\text{solvent}} P_{\text{amb}}}{RT_{\text{avg}}} \right] (y_{v,\text{surface}} - y_{v,\text{amb}}); A_{\text{eff}} \text{ -mass transfer area...Eq. 8}$$

$$\frac{dH_{\text{bead}}}{dt} = \frac{1}{A_{\text{eff}} \rho_{\text{solvent}}} \frac{dm_{\text{solvent}}}{dt}; H_{\text{bead}} \text{ -bead height, } A_{\text{eff}} \text{ -heat transfer area from substrate.Eq. 9}$$

$$\frac{dT_{\text{bead}}}{dt} = \frac{h_c[A_{\text{eff}}(T_{\text{amb}} - T_{\text{bead}})] + \Delta H_v \frac{dm_{\text{solvent}}}{dt} + \frac{A_{\text{eff}}(T_{\text{subs}} - T_{\text{bead}})(\kappa_{\text{sol}})}{H_{\text{bead}}/2}}{m_{\text{solute}}c_{p,\text{solute}} + m_{\text{solvent}}c_{p,\text{solvent}}}; \kappa \text{ -thermal cond.. Eq. 10}$$

The total time of the deposition process,  $t_{\text{dp}}$  can be calculated based on Eq. 11-12.

$$t_{\text{dp}} = \frac{\text{cell}_{x\text{-dim}} \text{cell}_{y\text{-dim}} \text{dpi}_{\text{target-y}}}{v_{\text{ph}} \text{dpi}_{\text{native}} \text{width}_{\text{ph}}}; \text{cell}_{x\text{-dim}}, \text{cell}_{y\text{-dim}} \text{ -x and y-dimension of cell.....Eq. 11}$$

$$v_{\text{ph}} = \frac{25.4 f_{\text{DOD}}}{1000 \text{dpi}_{\text{target-x}}}; v_{\text{ph}} \text{ -printhead velocity, } f_{\text{DOD}} \text{ -drop ejection frequency .....Eq. 12}$$

During film formation stage, line width ( $W_{\text{line}}$ ) during multiple repetition ( $N_{\text{rep-y}}$ ) is based on the lateral spacing (based on  $1/\text{dpi}_{\text{target-y}}$ ), equilibrium diameter of droplet and initial line width ( $W_{\text{ini-line}}$ ) (Eq. 13). When merging of line happens, the width ( $W_{\text{rem-bead}}$ ) and length ( $L_{\text{rem-bead}}$ ) of the new bead is expressed based on Eq. 14-15.

$$W_{\text{line}} = \frac{W_{\text{ini-line}}}{2} + \max\left(\frac{W_{\text{ini-line}}}{2}, (N_{\text{rep-y}} - 1) * \frac{0.0254}{\text{dpi}_{\text{target-y}}} + \frac{D_{\text{eq,d}}}{2}\right) \dots \dots \dots \text{Eq. 13}$$

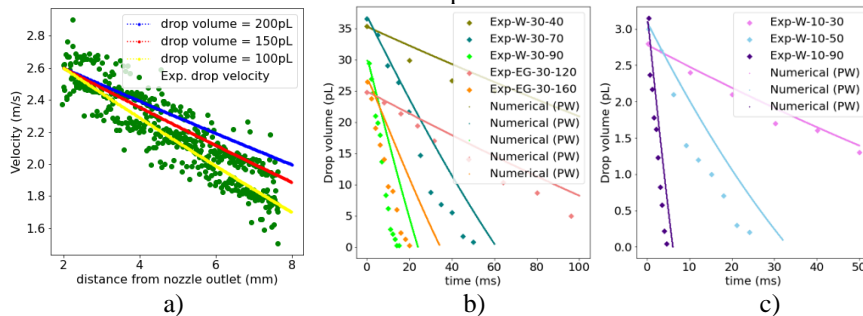
$$L_{\text{rem-bead}} = N_d * D_{\text{eq,d}} - (N_d - 1) * \frac{0.0254}{\text{dpi}_{\text{target-x}}}; N_d \text{ -number of drops deposited....Eq. 14}$$

$$W_{\text{rem-bead}} = \frac{0.0254}{\text{dpi}_{\text{native}}} - W_{\text{line}}; \text{dpi}_{\text{native}} \text{ - nozzles per inch.....Eq. 15}$$

### 3. Results

#### 3.1. Validation

In this section, model is validated based on existing experimental data before extending it to elucidate film formation dynamics. Figure 1 a) shows the droplet velocity as a function of distance from nozzle which lies in the range of data collected by Kang et al. (2020). Further, Figure 1 b) and c) shows the evolution of droplet volume based on present model and comparison with experimental results obtained by Lim et al. (2009) which shows very good agreement in all range of initial droplet volume and substrate temperature with water (W) and ethylene glycol (EG) as solvents. Since, the analytical expression for equilibrium droplet diameter and first bead/line width were directly used, they are not validated here. With these validation results, the model can be used for further understanding of the process dynamics involved in the inkjet printing process including line formation and film formation and the impact of different decision variables.

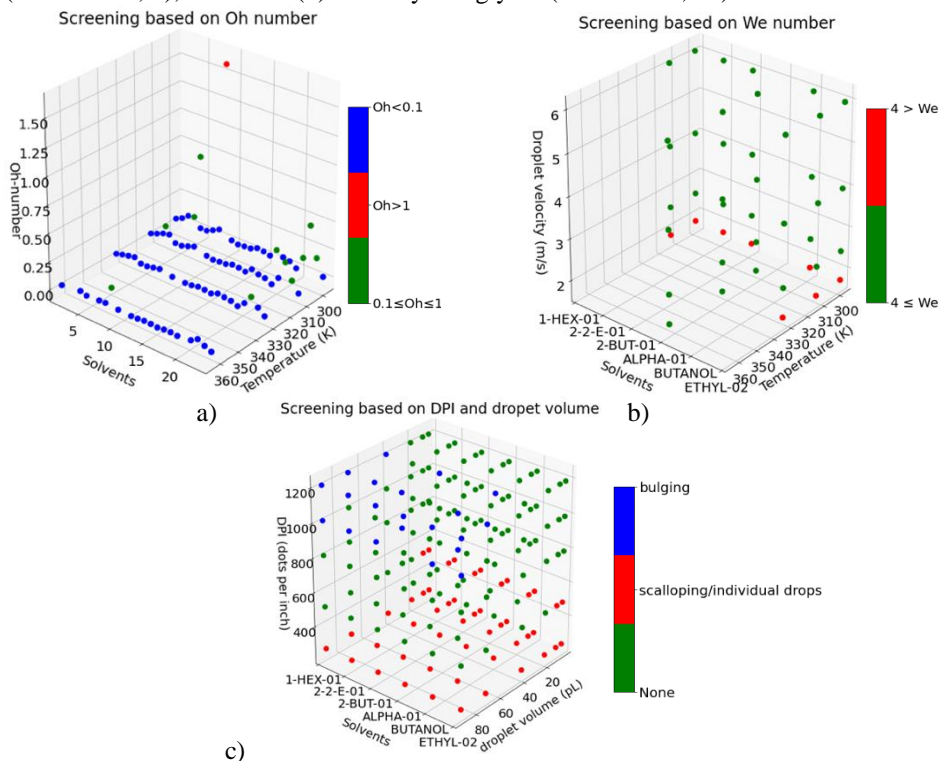


**Figure 1:** Validation results for a) droplet velocity, and b) and c) droplet volume during evaporation on substrate (PW-present work)

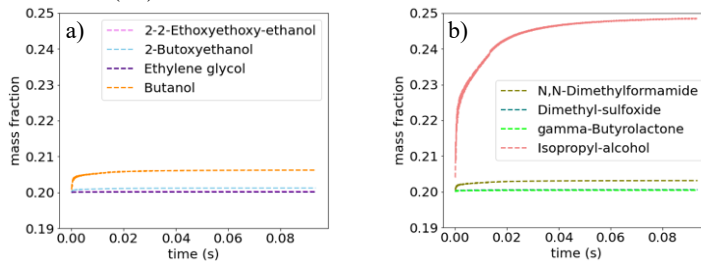
#### 3.2. Screening decision variables space

In the first part, the screening of parameter space based on coating window available for the process can be investigated. The result shows that out of 84000 cases only 12600 fall

within the coating window. Further, with constraints on the size of droplets with respect to nozzle and boiling point of solvents only 6480 cases remain feasible. Finally, with constraints on minimum spacing and maximum spacing between droplets only 4568 cases remain feasible for the inkjet printing process. Out of the 24 solvents commonly used in perovskite layer deposition, only 6 solvents have required properties with the given parameters selected for the analysis which include hexanol (1-HEX-01, 23), 2-2-ethoxyethoxy-ethanol (2-2-E-01, 21), 2-butoxyethanol (2-BUT-01, 17), alpha terpineol (ALPHA-01, 9), n-butanol (4) and ethylene glycol (ETHYL-02, 22).



**Figure 2:** Screening results based on a) Oh-number b) We number and c) line defects. Ethanol (2) and IPA (3) are on the edges of the applicability window. Further DMSO (14), DMF (13), gamma-butyrolactone (15), o/m-Dichlorobenzene (24, 10) and NMP (19) are also lying close to the edges of the application window at low temperature. Other solvents used include water (1), chlorobenzene (5), methanol (6), 2-methoxyethanol (7), propanol (8), acetone (11), toluene (12), Methyl-phenyl-ether (16), tetrahydronaphthalene (18) and acetonitrile (20). Ambient conditions are 25°C and 40% RH.



**Figure 3:** Influence of solvent properties a) within and b) outside coating window

3.3. Effect of solvent ( $T_{subs}=55^{\circ}C$ ,  $DPI=847$ ,  $LPI=635$ ,  $\theta_{eqb} = 1^{\circ}$ ,  $V_d=10pL$ ,  $f_{DOD}=5000$ )

Figure 3 shows the influence of different solvents on the wet film characteristics. IPA shows maximum change (>20%) in mass fraction over printing time (<100 ms). However, the solvents lying within the coating window don't show more than 5% change in the mass fraction over time with maximum change shown by Butanol. Therefore, mixing solvents with different properties can be used to influence the final film characteristics.

3.4. Effect of substrate temperature ( $T_{subs}$ )

Figure 4 shows the effect of substrate temperature on film properties with Dimethyl-sulfoxide (DMSO) as solvent (for all remaining cases). The effect of temperature becomes more dominant at high temperature resulting in >10% difference in film height with progressive increase in mass fraction of solute. Here, the maximum temperature cannot be more than boiling point of solvent to avoid pin-hole formation in the deposited film.

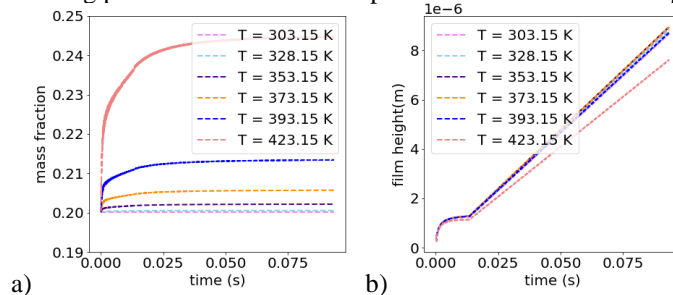


Figure 4: Influence of substrate temperature on a) mass fraction and b) film height

3.5. Effect of Target DPI ( $dpi_{target-x}$ ) ( $LPI=51$ )

Figure 5 shows the effect of DPI (dots per inch, relates to spacing in direction of printing) for DMSO at substrate temperature of  $55^{\circ}C$ . Depending on the frequency the total time for printing changes with dpi. With increase in DPI, the total time increases, however, the change in mass fraction is very less. Further, with increase in DPI, film height increase as more droplets are deposited on the same area of substrate. With increase in dpi, the morphology of wet film changes from individual lines (from different nozzles) to merged film as closer droplets leads to broader lines which become greater than nozzle spacing.

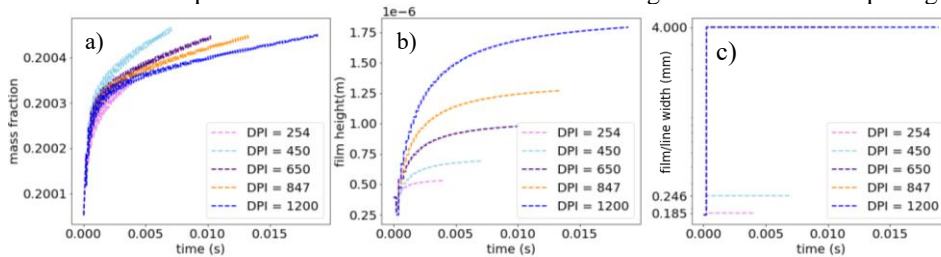


Figure 5: Influence of target DPI on a) mass fraction b) film height and c) film width

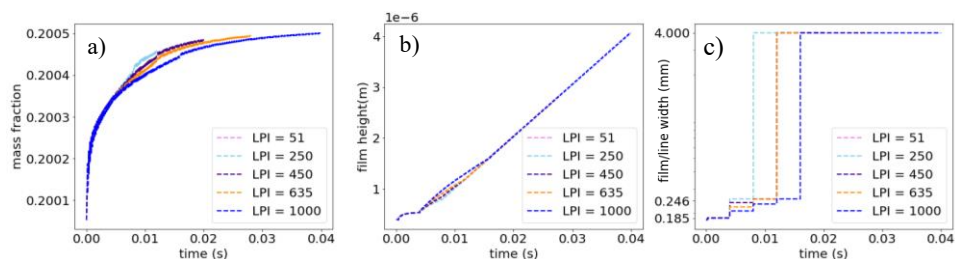
3.6. Effect of Target LPI ( $DPI=254$ )

In this subsection, influence of droplet spacing in lateral direction (lines per inch, lpi or  $dpi_{target-y}$ ) is shown. With LPI increase, total time of process increases resulting in more solution deposited on the same area thus leading to thicker wet films. Further, change in mass fraction decreases and time for film formation increase with increase in LPI(Fig. 6).

4. Conclusions

Here a mechanistic model for inkjet printing is presented which can be used to understand the impact of different decision variables involved in the process of inkjet printing.





**Figure 6:** Influence of target LPI on a) mass fraction b) film height and c) film width

Screening solvents based on existing literature should be done with more lenient conditions. Most of the solvents show similar behavior during the printing process except a few which have high vapor pressure. Substrate temperature can have significant impact of final mass fraction of film. DPI and LPI have direct influence on the film thickness, processing time and film morphology with less impact on change in mass fraction.

## 5. Acknowledgements

This work is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 945363.

## References

- Li, X., Bi, D., Yi, C., Décoppet, J.D., Luo, J., Zakeeruddin, S.M., Hagfeldt, A. and Grätzel, M., 2016. *Science*, 353(6294), pp.58-62
- Derby, B., 2010. *Annual Review of Materials Research*, 40, pp.395-414
- Kang, S.H., Kim, S., Lim, J.W., Sohn, D.K. and Ko, H.S., 2020. *Journal of Mechanical Science and Technology*, 34(8), pp.3311-3315.
- Soltman, D. and Subramanian, V., 2008. *Langmuir*, 24(5), pp.2224-2231
- Stringer, J. and Derby, B., 2010. *Langmuir*, 26(12), pp.10365-10372
- Tofan, T., Kruggel-Emden, H., Turla, V. and Jasevičius, R., 2021. *Applied Sciences*, 11(2), p.527
- van der Bos, A., van der Meulen, M.J., Driessen, T., van den Berg, M., Reinten, H., Wijshoff, H., Versluis, M. and Lohse, D., 2014. *Physical review applied*, 1(1), p.014004
- Yang, H., Wang, J., Yu, X., Feng, Y., Chen, X., Long, F., Ku, Z., Huang, F., Cheng, Y. and Peng, Y., 2022. *Chemical Physics Letters*, 807, p.140084
- Wei, Z., Chen, H., Yan, K. and Yang, S., 2014. *Angewandte. Chemie International. Edition*, 53(48), pp.13239-13243
- Abzieher, T., Moghadamzadeh, S., Schackmar, F., Eggers, H., Sutterlüti, F., Farooq, A., Kojda, D., Habicht, K., Schmager, R., Mertens, A. and Azmi, R., 2019. *Advanced Energy Materials*, 9(12), p.1802995.
- Gao, B. and Meng, J., 2021. *Solar Energy*, 230, pp.598-604
- Huckaba, A.J., Lee, Y., Xia, R., Paek, S., Bassetto, V.C., Oveisi, E., Lesch, A., Kinge, S., Dyson, P.J., Girault, H. and Nazeeruddin, M.K., 2019. *Energy Technology*, 7(2), pp.317-324
- Mathies, F., Abzieher, T., Hochstuhl, A., Glaser, K., Colsmann, A., Paetzold, U.W., Hernandez-Sosa, G., Lemmer, U. and Quintilla, A., 2016. *J. of Materials Chemistry A*, 4(48), pp.19207-19213
- Kunii, D. and Levenspiel, O., 1991. *Fluidization engineering*. Butterworth-Heinemann
- Sloth, J., 2007. *Formation of enzyme containing particles by spray drying* (Doctoral dissertation, PhD Thesis. Department of Chemical Engineering, Technical University of Denmark)
- Filipovic, L., Selberherr, S., Mutinati, G.C., Brunet, E., Steinhauer, S., Köck, A. and Schrank, F., 2013, July. In *Proceedings of the world congress on engineering* (Vol. 2, pp. 987-992).
- Lim, T., Han, S., Chung, J., Chung, J.T., Ko, S. and Grigoropoulos, C.P., 2009. *International Journal of Heat and Mass Transfer*, 52(1-2), pp.431-441

# Modeling and Optimization of Low-carbon Oriented Urban Energy Systems Spatial Planning Considering Energy Networks Topology

Ruiyu Zhang<sup>a</sup>, Pei Liu<sup>a,\*</sup>, Zheng Li<sup>a</sup>

<sup>a</sup>*State Key Lab of Power Systems, Department of Energy and Power Engineering,  
Tsinghua University, Beijing 100084, China  
[liu\\_pei@tsinghua.edu.cn](mailto:liu_pei@tsinghua.edu.cn)*

## Abstract

Spatial dimension planning of an energy system is critical at an urban scale due to the strong spatial heterogeneity of load levels and categories at different spots of a city and the high complexity of pipeline networks extension and connection topology. External requirements of low-carbon transition of a city may make its urban energy systems planning further complicated. In this paper we propose a novel urban energy systems spatial planning modeling and optimization approach considering energy systems form and topology. Coupling with integrated energy systems modeling and multiple energy pipeline networks topology modeling, the processes of energy generation, conversion, storage, and transmission are depicted synthetically. A case area is employed to illustrate feasibility of proposed method and the impacts of both low-carbon transition and user side participation in energy production on urban energy systems optimal layout. Results show that main energy supply technology shifts and corresponding networks topology changes significantly under the requirement of decarbonization and energy supply mode revolution. Decarbonization and participation of user side in power generation reduce the average network capacity by 19% and 26% respectively, indicating a more interconnected energy network with low capacity to balance energy between areas is a preferred form for a low-carbon city.

**Keywords:** urban energy systems; systems modeling and optimization; network topology; low carbon transition.

## 1. Introduction

An urban energy system consists of a variety of energy generating units and interconnected pipeline networks of multiple energy categories, including electricity, heat, or cold, thus has great potential to integrate multiple types of energy. However, the high complexity of such a system also brings great challenges to its planning. Moreover, urban areas are the main places of various kinds of social activities, in which might lead to massive carbon emissions and various environmental problems. It is expected more than 70% of the population will live in cities in 2050 (C40 Cites, 2017), and a considerable part of emissions may come from energy production and utilization. Therefore, it is of great importance to carry out low-carbon oriented urban energy systems planning, especially under the new trend of decarbonization and renewable energy utilization (Stoeglehner et al, 2016).

Currently, a series of urban energy related researches exist. One mainstream of concern is urban emissions measurement and analysis based on emissions inventory (Kennedy et al, 2010) and index decomposition method such as LMDI (Gu et al, 2019), IPAT and improved STRIPAT (Wang et al, 2017), to quantify the impact of various social and economic indicators on carbon emissions. Another group of scholars apply commercial software, for example LEAP (Huang et al, 2019), MARKAL (Isi et al, 2021) and Energy PLAN (Thellufsen et al, 2020), to implement simulation and designing for a low-carbon energy system. However, most of these studies are similar to decarbonization researches at a national scale (Sun et al, 2019), which generally apply lumped parameter method and regard objective into a whole, mainly discuss the impacts of relevant indicators change on carbon emissions. Spatial dimension, the most unique and vital aspect at a city scale, is neglected and heterogeneities and correlations within a city are thus not considered enough. Conversely, existing studies on urban space mainly focus on land use planning (Penazzi et al, 2019) and various location problems (Bélanger et al, 2019), where concerns of low-carbon oriented revolution of urban energy systems are absent. As a result, the purpose of this study is to investigate the impacts of decarbonization as well as energy supply revolution represented by user side participation in energy supply on the optimal urban energy systems form and layout, aiming at facilitating spatial planning of low-carbon oriented urban energy systems.

## 2. Methodology

### 2.1. Overview of model structure

The aim of this study is to explore the impacts of decarbonization and low-carbon revolution on urban energy systems form and layout. Therefore, an urban energy systems spatial planning model is proposed, as illustrated in Figure 1.

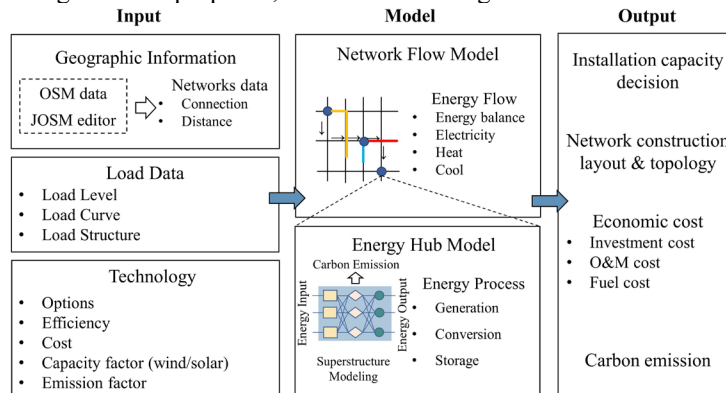


Figure 1 Model framework

Geographic information is derived from OpenStreetMap (OSM, 2022), a widely-used open-source map platform, and edited by Java OpenStreetMap (JOSM, 2022). Load data are used for depicting various energy demand and techno-economics parameters are applied for energy facilities modeling.

The core of the proposed model consists of two sub-models, namely the Network Flow Model and the Energy Hub Model. The purpose of the Network Flow Model is to characterize energy transmission in the energy network and energy balance at each node. The Energy Hub Model aims to establish a mathematical model for real energy facilities installed at energy hub nodes based on superstructure modeling method.

The planning goals are (a) which energy hub nodes should be selected, (b) which technologies and how much capacity to install in hubs, and (c) how to lay the energy pipe networks properly to satisfied all load nodes demand with minimum cost. By solving this problem, optimal installation decision, energy networks arrangement, detailed cost and carbon emissions can be obtained.

## 2.2. Mathematical formulation

### 2.2.1. Objective function

The optimization objective is to minimize the overall energy systems costs, including investment costs of technologies  $cost_{tech}$  and networks  $cost_{net}$ , operation and maintenance costs  $cost_{om}$  and fuel costs  $cost_{fuel}$ . The proportion coefficient of  $cost_{om}$  refers to literature (Xu et al., 2020).

$$cost_{total} = cost_{tech} + cost_{net} + cost_{om} + cost_{fuel} \quad (1)$$

$$cost_{tech} = \sum_{tech,node} CAPEX_{tech} \cdot ic_{node} \cdot \frac{r(1+r)^T}{(1+r)^{T-1}} \quad (2)$$

$$cost_{net} = \sum_{tdn,road_{id}} DISTANCE_{road_{id}} \cdot (C_0 + C_{1tdn}) \cdot icn_{tdn,road_{id}} \cdot \frac{r(1+r)^T}{(1+r)^{T-1}} \quad (3)$$

$$cost_{fuel} = \sum_{fuel,node_{hub}} FP_{fuel} \cdot \sum eimp_{node_{hub},s,dh,f} \quad (4)$$

$$cost_{om} = coefficient \cdot cost_{tech} \quad (5)$$

### 2.2.2. Network flow modeling

Expression (6) and (7) describe the upper limit of flow and energy balance of the energy flow in network.

$$-icn_{tdn,road_{id}} \leq flow_{road_{id},s,dh,e,tdn} \leq icn_{tdn,road_{id}} \quad (6)$$

$$\sum flow_{in,node} - \sum flow_{out,node} = \begin{cases} LOAD_{node}, & \text{if node} \in \text{load nodes} \\ 0 & \text{if node} \in \text{road nodes} \\ -output_{node_{hub},s,dh,e} & \text{if node} \in \text{hub nodes} \end{cases} \quad (7)$$

### 2.2.3. Energy hub modeling

Nine types of electricity generation technologies are considered, including Super-critical Pulverized Coal (SPC), Ultra-Supercritical Pulverized Coal (UPC), Combined-cycle Gas Turbine (CCGT), Biomass power plant (BE) and options of the above technologies with CCS. Solar Photovoltaic (PV) is selected as the option of renewable technologies. Besides, six types of heat and cool generation/conversion technologies, namely Biomass Power Plant Combined Heat and Power (BE\_CHP), Combined-cycle Gas Turbine Combined Heat and Power (CCGT\_CHP), Gas Boiler (GB) and Air-source Heat Pump (ASHP), Air-conditioner (AC), Absorption chiller (ABS\_CHILLER), and three kinds of energy storage technologies including Battery (BATTERY), Sensible Thermal Energy Storage (STES) and Ice Storage (ICES) are involved. Key expressions of energy generation, conversion and storage processes are presented in (8) to (11).

$$eimp_{node_{hub},s,dh,e} = \sum_{gnt} ei_{gnt,node_{hub},gnt,s,dh,e} \quad (8)$$

$$\sum_e ei_{gnt,node_{hub},gnt,s,dh,e} \times eff_{gnt,ee} = eo_{gnt,node_{hub},gnt,s,dh,e} \quad (9)$$

$$est_{node_{hub},stt,s,dh+1,e} = est_{node_{hub},stt,s,dh,e} \times (1 - LOSS_{STT}) - \frac{eo_{stt,node_{hub},stt,s,dh,e}}{eff_{stt,ee}} + ei_{stt,node_{hub},stt,s,dh,e} \quad (10)$$

$$\begin{aligned} output_{node_{hub},s,dh,e} = & \sum_{gnt} e_{gnt} e_{node_{hub},gnt,s,dh,e} + \\ & \sum_{cvt} e_{cvt} e_{node_{hub},cvt,s,dh,e} + \sum_{stt} e_{stt} e_{node_{hub},stt,s,dh,e} - \sum_{cvt} e_{cvt} e_{node_{hub},cvt,s,dh,e} - \sum_{stt} e_{stt} e_{node_{hub},stt,s,dh,e} \end{aligned} \quad (11)$$

### 2.2.4. Carbon emissions constrain

Emissions constrain is shown in (12), namely net emissions should be limited by quota. Specially, when quota is set as zero, it means carbon neutralization scenario.

$$\sum (e_{imp} e_{node_{hub},s,dh,f} - e_{ccs,node_{hub},gnt,s,dh,f} \cdot 0.9) \cdot CEI_f \cdot DAY_s \leq QUOTA \quad (12)$$

Eventually, the model proposed above is built in General Algebraic Modelling System (GAMS, 2022) platform and solved via applying built-in CPLEX solver.

## 3. Case study

### 3.1. Area information and scenarios setting

A case area commonly used in several literatures (Chen et al., 2018; Xu et al., 2020) is applied in this study. As shown in Figure 2, case area is consisted of 205 nodes, including 20 load nodes colored in orange triangle, representing points where energy demands exist, and 8 candidate energy hub nodes in red square meaning where energy facilities can be constructed, and 345 roads consist of existing roads in solid line and candidate in dotted line. All energy pipes are limited to build along roads. Basic load data and road distance are derived from literature (Xu et al., 2020). In this study the load data are further split into hourly scale via load profiles from a widely-used energy system planning software HOMER (NREL, 2022) to reflect the characteristics and changes of different categories of loads including residential, industry, and commerce, as shown in Figure 3.

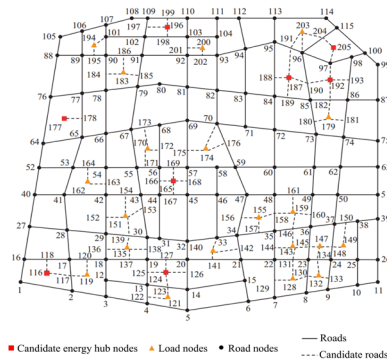


Figure 2 Networks of case area

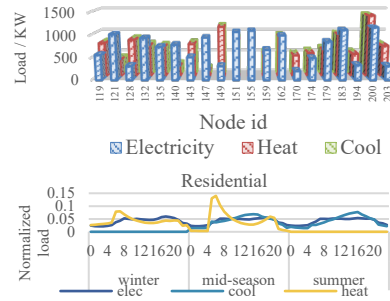


Figure 3 Load level and hourly-profile

To investigate the impacts of decarbonization on urban energy systems planning, three scenarios are set according to emissions constraint and energy supply mode, as shown in Table 1. **NC** is a reference scenario with no emissions constrain and follows the current mainstream idea of centralized energy supply, which means energy supply equipment except AC can only be installed in the energy hub. **DC-1** is a low-carbon scenario with the requirement of carbon neutralization but stays in traditional centralized energy supply mode, to illustrate the impacts of carbon neutralization. **DC-2** further indicates that distributed PVs are approved to be constructed at user side, namely load nodes and users play the role of both energy consumers and producers, which is an expected form of energy systems revolution in the process of decarbonization.

Table 1 Scenario setting and explanations

Scenario	Explanation
NC	No carbon emissions constrain; Centralized energy supply
DC-1	Decarbonization with carbon neutralization; Centralized energy supply
DC-2	Decarbonization with carbon neutralization; User side distributed PV available

3.2. Impacts of decarbonization on urban energy systems planning

3.2.1. Major energy technologies shifting

The choices of energy hubs in three scenarios are consistent, almost all candidate hub nodes are selected except node 205. However, the technology installing decisions are different as shown in figure 4 and figure 5. In DC-1, the capacity of PVs grows while UPC and ASHP descends markedly, and more types of technologies such as BE\_CHP and ABS\_CHILLER are used. The installed technologies at user side in these two scenarios are limited in AC only since scenarios setting. In DC-2 scenario, however, it can be found a considerable number of PVs are installed at the user side, and both batteries and thermal storage equipment are installed.

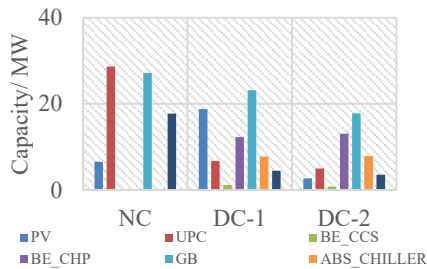


Figure 4 Installed capacity of energy hub

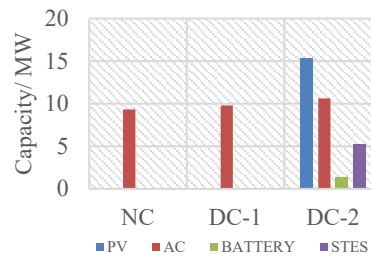
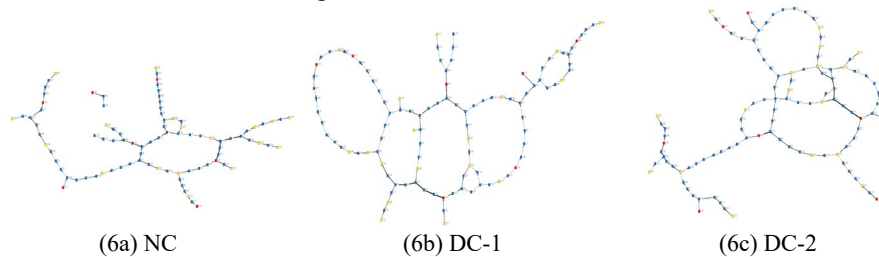


Figure 5 Installed capacity of user side

3.2.2. Energy networks topology

Figure 6 displays the electricity network and heat network topology between scenarios. Red, yellow, and blue nodes represent hub nodes, load nodes and road nodes respectively. Linkages between nodes mean where energy networks are constructed. It should be noticed that these topology diagrams only indicate the topological relationships, namely the connections between nodes, instead of real spatial direction and distance. Nodes unselected are not shown in diagrams.



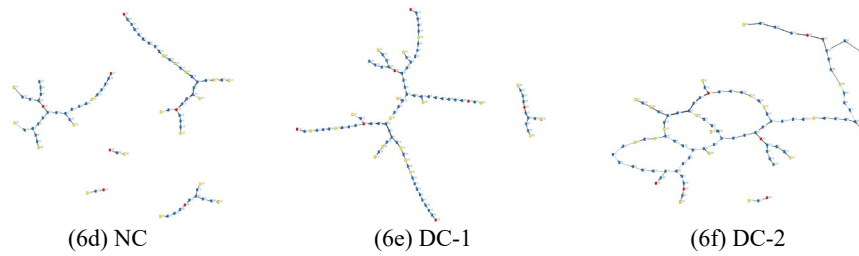


Figure 6 Electricity network (6a-6c) and heat network (6d-6f) topology

In detail, electricity network in NC scenario consists of two independent subnets. By comparison, the network exists as a whole and is more complex with six rings in DC-1. This is because abundant PVs are constructed under the constraints of carbon neutralization, which implies the mismatch between load and energy supply will be more severe due to the inherent volatility of renewable energy. Therefore, a more interconnected supply network is needed to balance the energy between nodes. And in DC-2 scenario, with the participation of user side distributed PVs, the number of latent energy supply nodes and the requirement for load-supply matching increase, explaining further complications on the network topology. Table 2 summarizes specific results of networks. It can be found that the results of total capacity in NC and DC-1 are almost the same, while the number of pipe lines increase 23%. Comparing DC-1 and DC-2, the number of lines is almost identical while total capacity decreases by 24%. A similar trend can be observed in heat network. However, the structure of heat network is more concise and its average capacity is approximately twice that of the electricity network.

Table 2 Energy networks capacity

Unit: MW	EN capacity	Number of lines	Average capacity	HN capacity	Number of lines	Average capacity
NC	128.3139	90	1.42571	235.1022	75	3.134695
DC-1	128.4098	111	1.156845	235.6052	89	2.64725
DC-2	97.06535	113	0.858985	188.0379	95	1.979346

#### 4. Conclusions

Under the targets of decarbonization and energy systems revolution, the decision making and spatial layout of urban energy systems will change significantly. Results indicate that a dispatching and balancing oriented network with more complex topology and lower average line capacity is needed for the requirements of decarbonization and user side participation in energy supply, instead of a transmission-oriented net that carries energy from centralized plants to load nodes.

#### References

- V. Bélanger et al, 2019, Recent optimization models and trends in location, relocation, and dispatching of emergency medical vehicles, *European Journal of Operational Research*, 272, 1, 1-23.
- C40 Cities, 2017, Focused acceleration: A strategic approach for climate action in cities, [www.mckinsey.com](http://www.mckinsey.com) <accessed in 2022.11.13>.
- J. Chen et al, 2018, Research on “Stations-Pipelines” Layout and Optimization of Regional Energy Internet, *Proceedings of the CSEE*, 38(03), 675-684. (In Chinese).
- GAMS, [www.gams.com](http://www.gams.com) <accessed in 2022.11.11>.
- S. Gu et al, 2019, Coupled LMDI and system dynamics model for estimating urban CO2 emission mitigation potential in Shanghai, China, *Journal of Cleaner Production*, 240, 118034.

- Y. Huang et al, 2019, Exploring potential pathways towards urban greenhouse gas peaks: A case study of Guangzhou, China, *Applied Energy*, 251, 113369.
- M. Isik et al, 2021, Transportation emissions scenarios for New York City under different carbon intensities of electricity and electric vehicle adoption rates. *Nature Energy* 6, 92–104.
- JOSM, [josm.openstreetmap.de](http://josm.openstreetmap.de) <accessed in 2022.11.11>.
- C. Kennedy et al, 2010, Methodology for inventorying greenhouse gas emissions from global cities, *Energy Policy*, 38, Issue 9, 4828-4837.
- NREL, [www.nrel.gov/HOMER](http://www.nrel.gov/HOMER) <accessed in 2022.11.11>.
- OSM, [www.openstreetmap.org](http://www.openstreetmap.org) <accessed in 2022.11.11>.
- S.Penazzi et al, 2019, Planning low carbon urban-rural ecosystems: An integrated transport land-use model, *Journal of Cleaner Production*, 235, 96-111.
- G. Stoeglehner et al, 2016, Chapter 3: Spatial Archetypes in the Energy Turn, *Integrated Spatial and Energy Planning Supporting Climate Protection and the Energy Turn with Means of Spatial Planning*, 35-51.
- Z. Sun et al, 2019, China's carbon emission peak pre-2030: Exploring multi-scenario optimal low-carbon behaviors for China's regions, *Journal of Cleaner Production*, 231, 963-979.
- J.Z. Thellufsen et al, 2020, Smart energy cities in a 100% renewable energy context, *Renewable and Sustainable Energy Reviews*, 129, 109922.
- S. Wang et al, 2017, Examining the impacts of socioeconomic factors, urban form, and transportation networks on CO<sub>2</sub> emissions in China's megacities, *Applied Energy*, 185, Part 1, 189-200.
- C. Xu et al, 2020, Planning of Energy Station and Pipeline Considering Topological Characteristics of Regional Integrated Energy System , *Automation of Electric Power Systems*, 44(03), 74-82. (In Chinese).





# Multi-objective optimisation for microbrewery retrofitting

Ruosi Zhang<sup>a,b</sup>, Melis Duyar<sup>a</sup>, Lesley Foulkes<sup>c</sup>, James Berrow<sup>c</sup>, Michael Short<sup>a\*</sup>

<sup>a</sup>*University of Surrey, Stag Hill, Guildford GU2 7XH, UK*

<sup>b</sup>*National Innovation Center for Digital Fishery, Beijing 100083, China*

<sup>c</sup>*Langham Brewery, The Granary, Lodsworth, West Sussex GU28 9BU, UK*

[m.short@surrey.ac.uk](mailto:m.short@surrey.ac.uk)

## Abstract

In a highly competitive retail market, many microbreweries have attempted to maximise profits and decrease energy consumption through retrofitting their operations with renewable energy. This paper develops an optimisation model to minimise investment and operation costs of a microbrewery meeting the dynamic energy demands, via an integrated photovoltaic (PV) system with energy storage and different boiler choices to lower carbon emissions. A microbrewery in UK has been used for the case study to demonstrate the approach on real data, with challenges in implementation and real-world constraints and considerations discussed. A set of rigorous multi-objective optimisation and sensitivity analyses are performed to analyse the resulting system. For the particular brewery, a modern electric boiler combined with photovoltaic system is an economic and sustainable choice, due to the cooling and other electric requirements in the brewery, leading to a 33 percent reduction in operational costs with a payback time of 2.6 years.

**Keywords:** microbrewery operation, optimisation model, PV system, sustainable.

## 1. Introduction

Beer is a popular alcoholic drink throughout the world. Global beer consumption reached approximately 177.50 million kiloliters in 2020, equivalent to 537.9 billion 33 cl bottles (Kirin Holdings, 2022). The UK is ranked 8<sup>th</sup> with 60.2 liters beer consumption per-capita. In the past two decades, environmental concerns and increasing energy cost have impacted companies and beer consumers. The new challenges for the brewing industry related to energy efficiency, water consumption, emission management and waste generation, have attracted great attention from breweries. Beer production is an energy intensive process, and the energy demand is approximately 262 MJ/hL (59% thermal, 41% electrical) (Fadare et al., 2010). Compared with large breweries, microbreweries consume up to 1.5-2.0 times more for thermal and electric energy (Kubule et al., 2016). In 2020, the number of active microbreweries in the UK was 1,852, with a steady increasing trend (Statista, 2022). Approximately 8% of the production cost is used in thermal processes like boiling, cooling, and fermentation (Kubule et al., 2016). Local microbreweries are trying to find a way to reduce energy consumption. Breweries have considered sustainability and lowering carbon footprints in their processes, seeking the strategies to minimise emissions and impact on the environment. Shifting load to a lower tariffs period is potential saving in other industries, but not suitable for microbrewery because the load demand is non-deferrable. Renewable energy sources and energy storage system can offer a solution for reducing the operation costs. Muster-Slawitsch et al.(2011) have analysed the use of different renewable sources, and identified solar energy as the best option to support the brewery industry.

The studies of solar sources for the brewery are focused on large size of breweries (Eiholzer *et al.*, 2017). Pino *et al.*(2019) used two case studies in Spain to model the integration of a hybrid PV system into thermal and electricity generation system. The simulation results showed that solar sources have potential benefits in microbreweries with payback period of 10.7 years. Another energy-related case study is based on a microbrewery in South Africa(Kusakana, 2020). A grid-connected tracking PV system and a battery were integrated into the microbrewery, and the payback period is 13.8 years based on simulation results. Very few published studies focused on the onsite small-scale PV panels electricity generation, and they have not looked at the application in the UK. Given the popularity of microbreweries and needs to reduce energy costs, this paper develops an optimisation model to minimise investment and operation costs of a microbrewery meeting the dynamic energy demands, via an integrated PV system with energy storage and different boiler choices to lower carbon emissions. A microbrewery in UK has been used for the case study to demonstrate the approach on real data, with challenges in implementation and real-world constraints and considerations discussed. The thermal and electricity load profile of the brewing process has been collected in summertime and, based on the collected data, the dynamic energy demand profiles for brewing days has been characterised. After an initial analysis, the number of PV panels, installation and type of the battery, and type of boiler are used as decision variables in model. A set of rigorous multi-objective optimisation and sensitivity analyses are performed to analyse the resulting system. The proposed optimisation model and hybrid system can be used in industrial, residential, or other demand sectors to reduce their operational energy costs through a detailed multi-objective analysis and sensitivity study.

## 2. Case study description

The case study for the model is a microbrewery called Langham Brewery located at 50°59'N,0°41'W, which regularly produces six types of craft beers. Langham is a typical, small-scale British microbrewery. Electricity is provided from Scottish Power for the cooling process, and thermal processes use a steam boiler fueled by Furnace Flame. The brewery has a capacity of producing one batch of 1,650 L of beer four times per week. Therefore, the yearly demand is assumed at an average of 208 brewing days. Additionally, the cold storage and the fermentation chamber run as the baseload throughout the year.

The load demand has been recorded in a summer day in July 2022. The first brewing day is the most energy intensive one, with the different processes shown in Figure 1 where the brewery's load demand is the highest. During the night, from 18h30 to 6h30 next day, the baseload is assumed at 4.4 kW including boiler (turned on, but not producing steam and on standby), cold storage and fermentation chamber. Starting at 6h30 to 7h15, the 2.9kW boiler and 0.65 kW pump begin to work and prepare for the mashing process. The temperature of hot water in the hot liquor tank (HLT) is 77°C and 700 L of hot water is discharged into an insulated mash tank (mixer). Malted barley and specialty grains are conveyed into the mash tank. The mashing stage starts from 7h15 to 8h30, boiler and pump are kept on during this stage reheating water in HLT for the next process. After mashing, the temperature of mixture is 66.3°C.

At 8h30, the sparging process begins. In this process, 1,950 L hot water from the HLT is added to the mash tank over a period of about 100 min. Simultaneously, using a 0.65 kW pump, the dissolved sugars in the water (wort) extracted from the grain husks is slowly drained from the bottom and pumped into a boiler where the hops are added to the mixture (wort). The mixture is heated for an hour. At the end of the sparging process, the temperature of the mixture is 90°C. From 11h00 to 12h00, the mixture is heated from

90°C to 100°C, preparing for boiling. The boiling process starts from 12h00 to 13h00. After 15 min rest, the mixture is cooled by cold water (5°C) via heat exchanger to reach 20°C, desirable for the required fermentation. The whole preparation process ends at 18h30 after pumping the mixture into fermentation tanks and finishing the cleaning process. The beer wort is moved to the fermentation which constitutes the 4.4 kW base load together with the cold storage.

Based on our data collection, the daily operation cost including electricity and fuel costs on a brewing day is approx. £158. This value excludes maintenance costs. The daily CO<sub>2</sub> emission is 0.35 tonnes per brewing day (ScottishPower, 2022).

### 3. Methodology

After analysing Figure 1, it can be noticed that most of the peak power demand is occurring during the day, matching the solar irradiance profile. The implementation of a PV system can be of great advantage in maximising the electrical energy that can be generated from the panels while minimising the cost and amount of energy acquired from the grid. A battery storage system can be added to the supply system to store some energy during excess generation times and can also be used in the event the PV system cannot supply the demand. This will also increase the availability of energy supply as well as provide an opportunity of applying demand response strategies through optimal power dispatch. This section will describe the proposed system's operation modeling; the selected size and the basis of the economic analysis to be performed.

#### 3.1. Optimal energy management model of the grid-connected PV system

The overall objective of this model is to minimize the total cost in the whole lifetime:

$$f = \min (C_{Operation} + C_{Investment}) = (\sum_{S=1,2,3,4} C_S^{grid} - I_S) * \text{Lifetime} + \sum_{Unit} AC_{Unit} \quad (1)$$

where  $AC_{Unit}$  is the capital cost of the equipment including PV panels, battery, and boilers;  $C_S^{grid}$  is the seasonal cost of buying electricity from the grid;  $I_S$  is the seasonal income from selling electricity to the grid. There are two choices of the boiler, one is a gas boiler, and the other is an electric boiler. The satisfaction of electricity demand  $P_{t,c}^{load}$  is described as below:

$$\sum_c P_{t,c}^{load} = P_t^{grid} + P_t^{PV} + P_t^{BAT} \quad (2)$$

where  $P_t^{grid}$  is the power got from the grid to satisfy the loads. If it is positive, electricity is bought from grid to satisfy the load; if it is negative, electricity is sold back to main distribution.  $P_t^{PV}$  is the amount of power generated by the PV that is used to satisfy the demand, and  $P_t^{BAT}$  is the amount of power used by the batteries. If it is positive, the battery is discharging; if it is negative, the battery charges. Note that indices  $t \in T$ ,  $c \in C$  refer to the timestep and boiler type. Similarly, the heat demand is satisfied by the heat generated (note that heat storage or heating networks are not included in the model).

The total power generated by the PVs, is described below:

$$P_t^{PV} \leq N^{Panel} * P_{PV} \quad (3)$$

where  $N^{Panel}$  is the number of panels installed (an integer variable) and needs to be decided,  $P_{PV}$  is the power generated each panel.

The model also ensures that electricity is not bought from or sold to the grid at the same time by the addition of the big-M constraints, which are similar to the constraints on the battery's charging or discharging status. The type of the boiler installed is limited by a binary variable  $K_c$ , constrained as follows:

$$\sum_c K_c = 1 \quad (4)$$

For proper operation, each power source must not be operated above its rated or maximum limit according to the manufacturer's specifications. The state variable also has boundaries where the minimum depends on the type of battery used.

#### 4. Optimisation results

All the constraints in the optimisation model are linear with integer variables. The optimisation problem are solved using cbc (Forrest & Lougee-Heimer, 2005) solver from Julia JuMP (Dunning et al., 2017) optimisation toolbox. Simulations are performed to assess the effectiveness of the developed grid-connected system optimal operation model to minimise the daily energy cost of power purchased from the grid. The simulations reported in the sections below explain the system's behavior for a brewing day in four seasons with different boundaries on investment, which means disposable investment is gradually decreasing. The maximum number of PV panels is 1000. The upper boundary of investment is £500,000.

##### 4.1. Different payback periods

We did not consider the maintenance costs for equipment, and there is no limit on the investment cost. In Figure 2, the relationships among daily operation cost, capacity of PV panels and payback period is shown. Under sufficient investment, the model chooses to install an electric boiler under all the payback periods. If the payback period is lower than 5 years, the brewery should only install an electric boiler, with no battery or PV panels. When the payback period is more than 5 years, with the installation of PV panels, operation costs decrease. To invest in an electric boiler, the operation cost will drop from approximately £158 per day to £102.6 per day, with an investment assumed to be £27,255, yielding a payback period of 2.6 years. This will also yield a CO<sub>2</sub> emissions reduction from the boiler system of 62 % and an overall business emissions reduction of 21 %.

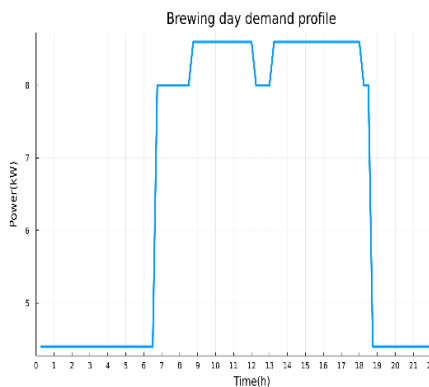


Figure 1 Brewing day electricity demand profile.

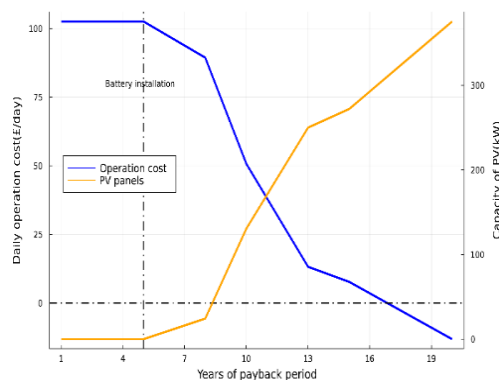


Figure 2 The relationships among daily operation cost, payback period, and capacity of PV panels.

The installation of PV panels has a great Impact on the daily operation cost of brewing days, especially in spring and summer, because the solar irradiance value is relatively high in these two seasons. The daily costs will decrease with longer payback periods especially 8-13 years if we have enough investment budget to install more PV panels. When the payback period is higher than 17 years, the operation costs are negative.

#### 4.2. Different gas price and investment cost

To analyse whether gas boilers are ever considered, the payback period was set to 10 years gas prices from 4 to 40 p/kWh were simulated. When the price of natural gas is lower than 18 p/kWh, the model chooses to install a gas boiler not an electric one. Compared with the gas boiler, the CO<sub>2</sub> emissions can be reduced more than 50% through electrification. If the Smart Export Guarantee can increase the export price 20% (from 5p to 6p), carbon emissions can be reduced by 28%, though the overall cost is only reduced by 2%.

We also analyse the simulation results under different investment cost constraints, and the results are shown in Figure 3. Even under different investment conditions, the model always chooses an electric boiler due to high gas prices. Daily operation cost drops as number of solar panels increases, and a significant positive linear correlation between operation cost and installation of PV panels are shown in Figure 3. When the value of investment is lower than £42,000, the optimisation model chooses not to install a battery, the turning point is presented by a vertical dotted line with 18 kW PV capacity. On the right side of the line, the model chooses to install batteries. As the investment cost decreases, the operation cost gradually increases, with the increase in CO<sub>2</sub> emissions displayed by the orange line.

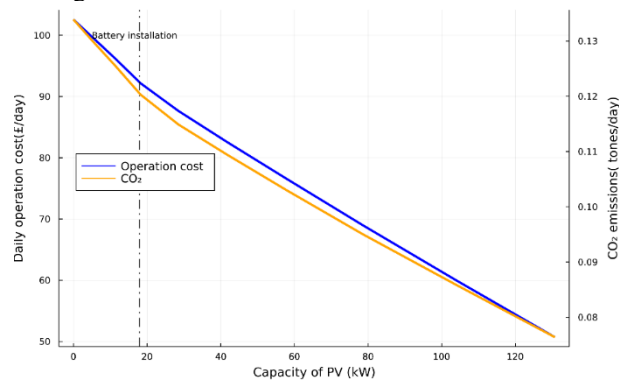


Figure 3 The relationships among daily operation cost, CO<sub>2</sub> emissions, and capacity of PV panels.

## 5. Conclusion

To be more competitive and sustainable, microbreweries need to reduce their energy cost without concession on the quality of final product. The aim of this work is to analyse the techno-economic benefits of using renewable sources by maximising profits and decreasing energy consumption through retrofitting their operations for a microbrewery. This paper develops an optimisation model to meet the dynamic energy demand based on data collection from a real microbrewery in UK, by minimising investment and operating costs and integrating PV systems (with energy storage and different boiler options to reduce carbon emissions). Along with the energy demand profiles, the solar resources, system components, and energy costs have been used as inputs to develop the model. Under the operational conditions and constraints, the model analyses the economic, technical, and sustainable performance of the system, together with company- and region-specific constraints related to budgets and incentives. The simulation results showed that the model can effectively optimise the capacity of PV system, installation of the battery, and type of boiler.

Compared to the scenario in which the grid is the only option for energy supply, the proposed system can significantly reduce energy costs. A potential energy cost saving of 33.3%, is possible when the optimal energy management model is applied to the proposed system, under the applicable operating condition. The length of the equipment life cycle has a great impact on the optimal investment strategy of the brewery, with a 7-year payback period important should the microbrewery wish to move towards electrification. Due to the installation of PV panels, there is a positive correlation between daily operation costs and carbon emissions. The carbon emissions of electric boilers are significantly lower than those of gas boilers, and the brewery prefers to install an electric boiler, unless the price of natural gas is lower than 18p/kWh. More PV panels installed brings lower daily operation cost and lower carbon emissions, but higher investment costs, and therefore the brewery needs to find a balance according to their own needs.

The satisfactory results obtained in this study demonstrate the potential techno-economic benefits of solar PV integration in the small-scale microbreweries. The proposed optimisation model and hybrid system can be used in industrial, residential, or other demand sectors to reduce their operational energy costs through a detailed multi-objective analysis and sensitivity study.

## References

- Dunning, I., Huchette, J., & Lubin, M. (2017). JuMP: A Modeling Language for Mathematical Optimization. *SIAM Review*, 59(2), 295–320. <https://doi.org/10.1137/15M1020575>
- Eiholzer, T., Olsen, D., Hoffmann, S., Sturm, B., & Wellig, B. (2017). Integration of a solar thermal system in a medium-sized brewery using pinch analysis: Methodology and case study. *Applied Thermal Engineering*, 113, 1558–1568. <https://doi.org/10.1016/j.applthermaleng.2016.09.124>
- Fadare, D., Nkpubre, D., Oni, A., Falana, A., Waheed, M., & Bamiro, O. (2010). Energy and exergy analyses of malt drink production in Nigeria. *Energy*, 35(12), 5336–5346.
- Forrest, J., & Lougee-Heimer, R. (2005). CBC User Guide. In *Emerging Theory, Methods, and Applications* (pp. 257–277). INFORMS. <https://doi.org/10.1287/educ.1053.0020>
- Kirin Holdings. (2022). *Global Beer Consumption by Country in 2020 | 2022 | Kirin Holdings*. Kirin Holdings Company, Limited. [https://www.kirinholdings.com/en/newsroom/release/2022/0127\\_04.html](https://www.kirinholdings.com/en/newsroom/release/2022/0127_04.html)
- Kubule, A., Zogla, L., Ikaunieks, J., & Rosa, M. (2016). Highlights on energy efficiency improvements: A case of a small brewery. *Journal of Cleaner Production*, 138, 275–286. <https://doi.org/10.1016/j.jclepro.2016.02.131>
- Kusakana, K. (2020). Optimal energy management of a grid-connected dual-tracking photovoltaic system with battery storage: Case of a microbrewery under demand response. *Energy*, 212, 118782. <https://doi.org/10.1016/j.energy.2020.118782>
- Muster-Slawitsch, B., Weiss, W., Schmitzer, H., & Brunner, C. (2011). The green brewery concept – Energy efficiency and the use of renewable energy sources in breweries. *Applied Thermal Engineering*, 31(13), 2123–2134. <https://doi.org/10.1016/j.applthermaleng.2011.03.033>
- Pino, A., Lucena, F. J. P., & Macho, J. G. (2019). Economic analysis for solar energy integration in a microbrewery. *2019 International Conference on Smart Energy Systems and Technologies (SEST)*, 1–6.
- ScottishPower. (2022). *Where does our energy come from? | Gas and Electricity Company | ScottishPower*. <https://www.scottishpower.co.uk/about-us/performance/fuel-mix>
- Statista. (2022). *Topic: Craft beer in the UK*. Statista. <https://www.statista.com/topics/6456/craft-beer-in-the-uk/>

# Thermodynamic equilibrium modelling for the optimal performance of a wood biomass downdraft gasifier for hydrogen production

Anamika Kushwah,<sup>a</sup> Tomas Ramirez Reina,<sup>a,b</sup> Michael Short,<sup>a\*</sup>

<sup>a</sup>*University of Surrey, Guildford, GU27XH, UK*

<sup>b</sup>*Institute of Materials Science of Seville, Seville, 41092, Spain*

## Abstract

The flexibility of gasification, which incorporates both characteristics of the kind of biomass and options for producing heat or fuel, is what stimulates interest in biomass gasification research and application possibilities. One of the challenges with gasification is the accurate prediction of gasification products from the wide array of feedstocks and operational conditions, making the design and optimisation challenging. There are several studies on equilibrium models that have been published, but most of the models examined the impact of moisture and heating value. Examining the other parameter effects within the same model can enhance the accuracy of a gasification model. This work aims to create stoichiometric equilibrium models that enable the analysis of parameter effects on a specific feedstock's gasification products for use in optimisation and system modelling. A mathematical model for fixed bed gasifiers that use downdraft gasification for wood as a feedstock is presented. When the algebraic model equations are solved, the conversion behaviour of wood biomass throughout a gasification process is predicted in terms of the composition of the end products. The model has been verified with published experimental data from the literature demonstrating good agreement. The created model may be regarded as a helpful tool to simulate the impact of numerous different biomass feedstocks and operating conditions on gas characteristics and the simple algebraic formulation allows for the application of the model for optimisation purposes.

**Keywords:** Biomass gasification, modelling, equilibrium models

## 1. Introduction

Global warming is regarded as one of the top global challenges for contemporary civilisation, notwithstanding COVID-19 crisis global challenges. More waste is produced because of growing urbanisation and economic growth. The primary problem is reducing environmental pollution and worldwide warming brought on by the extensive use of fossil fuels including coal, petroleum, and natural gas. These fossil fuels serve as the main source of energy generation in many nations, which results in continuous greenhouse gas emissions. Because of its widespread availability and practical processing, biomass is predicted to grow significantly among the many renewable energy sources in the near future (Yamany et al. 2016). There is ongoing research on methods for enhancing the production of energy from biomass, and several organisations are working on its development (Janajreh et al., 2021). One of the most adaptable industrial processes is biomass gasification, which converts biomass fuels into consumable gaseous fuels that may be used for heating, power production, and hydrogen fuel cells in a range of residential and commercial applications (Sidek et al., 2020). Due to its versatility, gasification - a thermochemical reaction that transforms solid organic resources into gaseous fuels at temperatures ranging from 700 to 900 °C is highly popular (Okolie et al.,



2021). It yields valuable products like CO, trace amounts of CH<sub>4</sub>, H<sub>2</sub>, and undesirable gases like N<sub>2</sub>, CO<sub>2</sub>, and other hydrocarbons. Because experimental conditions can affect the behaviour of reactions like water gas, steam reforming reactions, Boudouard and water gas shift reactions, and these all have an effect on syngas formation, stability, and thermal efficiency, restricting its future uses, it is crucial to use caution when simulating experimental conditions (Baruah & Baruah, 2014b).

Multiple studies have revealed that modelling and theoretical predictions based on equilibrium modelling have a great deal of similarities (Kushwah et al, 2022; Chaurasia, 2018). For instance, Pradhan et al. (2019) used varied equilibrium reaction rates to test five different biomass gasification models. When compared to other models, they discovered that the model that contained basic methanation and the water gas shift reaction generated syngas with the most accurate composition. The correlation output predictions resembled those of the equilibrium model, with the exception of CH<sub>4</sub>. On the other hand, reliable correlations will require the integration of more empirical observations and the effect of other parameters.

Experiments generally consume significantly more time, money, and effort than modeling. Modeling has been shown to be a relatively quick and cost-effective option in situations when doing so is not always practicable or economically sensible (Galvanin et al., 2009). Thermodynamic equilibrium models forecast the highest possible yield of desired products without requiring any understanding of the conversion processes. Like this, downdraft gasifiers are widely employed in various energy platforms due to their straightforward construction, inexpensive maintenance requirements, and ability to produce gas with a low tar content for use in small-scale electrical generation. This paper proposes a thermodynamic stoichiometric equilibrium model for a downdraft gasifier is used for the gasification of wood at 800 °C with moisture content from 0-40 % in order to predict the composition of the fuel gas produced and its associated heating values. The gasifier system is modelled using an optimisation modelling package which includes mass conservation equations, energy conservation equations and the equilibrium equations. The developed model, based on thermodynamic calculations, has been validated with experimental published data of other authors and provides opportunity to evaluate downdraft gasification processes in a fixed bed reactor, as well as effects of variations in biomass properties and operating conditions.

## 2. Methodology

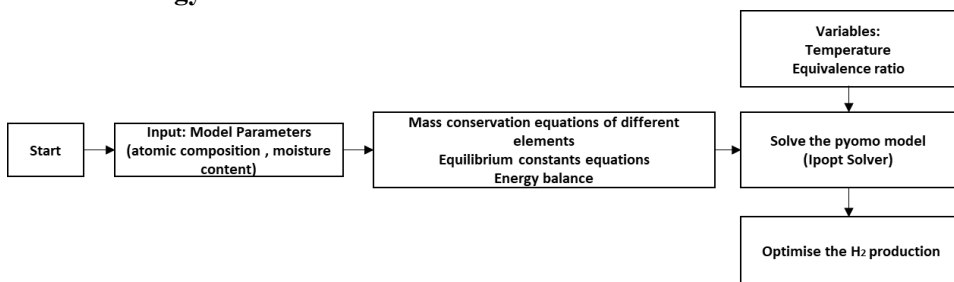


Figure 2.1: Structure of the simulation equilibrium model

In figure 2.1, the structure of the equilibrium model has been presented. In this, a mathematical model based on a single feedstock is created and developed in the algebraic modelling language, Pyomo (Hart et al. 2017) in order to determine the molar values of products. It includes the mass balance of carbon, hydrogen, oxygen, nitrogen in the woody

biomass, relative mass of the moisture and the molar water content. In next step, the composition of the producer gas is estimated, minimising the Gibbs free energy for calculation of the equilibrium constants. A Pyomo model has been implemented that relates equations between inlet biomass and equilibrium constants with temperature to solve the model. The equivalency ratio and temperature are used as optimization variables when the model is optimised using the IPOPT nonlinear programming solver to get the maximum hydrogen production rate.

### 2.1 Biomass composition

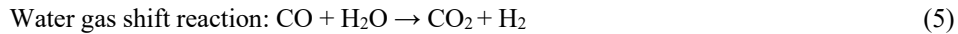
To develop the model, the chemical formula of feedstock is defined as  $CH_xO_yN_z$ . The following is an equation for the global gasification reaction:



Where x and y are the numbers of atoms of hydrogen and oxygen per atom of carbon in the feedstock, w is the amount of moisture per kmol of feedstock and m is the amount of oxygen per kmol of feedstock. On the right-hand side  $n_i$  are the numbers of moles of the species i that are unknown.

### 2.2 Chemical Equilibrium

Thermodynamic equilibrium is explained by minimisation of Gibbs free energy or by using an equilibrium constant. The two main equations (methane formation and water gas shift reaction) are considered in this case shown below:



Equation (5) is obtained by subtracting equation (2) from (3). In this model, thermodynamic equilibrium was assumed for all chemical reactions takes place in gasification zone. All the gases were assumed to be ideal and all reaction form at 1 atm. Therefore, the equilibrium constants which are functions of temperature for the water-gas shift reaction and methane formation reactions are:

$$K1 = \frac{(nCH_4)}{(nH_2^2)} \quad (\text{Equilibrium constant for methane reaction}) \quad (6)$$

$$K2 = \frac{(nCO_2)(nH_2)}{(nCO)(nH_2O)} \quad (\text{equilibrium constant for water gas shift reaction}) \quad (7)$$

The value of K1 and K2 are calculated from Gibbs free energy

$$K = \exp(-\Delta G^\circ_T / RT) \quad (8)$$

Where R is the universal constant, 8.314 KJ/ (kmol. K)

$-\Delta G^\circ_T$  represents the Gibbs function of formation at given temperature which can be expressed by empirical equation (Green, 1984):

$$\frac{d(\Delta G^\circ_T / RT)}{dT} = -\frac{\Delta H}{RT^2} \quad (9)$$

Where  $\Delta H$  is the heat of formation of biomass in KJ/kmol

$$\text{Therefore: } \frac{d \ln K}{dT} = -\frac{\Delta H}{RT^2} \quad (10)$$

Where, equation 9 shows the effect of temperature on the equilibrium constant;  $\Delta H$  represents the enthalpy difference between any given state and at reference state. It can be calculated by

$$\Delta H = \int_{298}^T Cp(T)dt \quad (11)$$

Where  $Cp(T)$  is specific heat at constant pressure in kJ/kmol K and is a function of temperature.

$$Cp(T) = a + bT + cT^2 + dT^3 \text{ where } a, b, c, d \text{ are specific gas coefficients} \quad (12)$$

Equations (8)-(12) is used to find the equilibrium constant for any reaction temperature T. Zainal et al. (2001), suggested the relationship for finding the equilibrium constant values for reaction temperature and calculating the lower heating value (LHV) of product gas that are:

$$K1 = \frac{7082.848}{T} - 6.567 * \ln(T) + 3.733 * 10^{-3} - 0.36066 * 10^{-6} + 0.3505 * \frac{10^{-5}}{T^2} + 32.54 \quad (13)$$

$$K2 = \frac{5870.53}{T} + 1.86 * \ln(T) + 2.7 * 10^{-4} - \frac{58200}{T^2} + 18.007 \quad (14)$$

$$LHV = n_1 LHV(H_2) + n_2 LHV(CO) + n_5 LHV(CH_4) \text{ MJ/Nm}^3 \quad (15)$$

However, the reference model is limited to moisture content and heating value but using an optimisation-based algebraic formulation can be used to study other parameter effects like equivalence ratio and gasification agent against temperature. For the desired product, this can also find the optimum parameters for gasification reaction. For different ranges of temperature, the values of equilibrium constant, numbers of moles and heating value have been calculated in a model, which is discussed in the next section.

### 3. Results and Discussion

The model developed in this study was used to optimise the production of hydrogen by varying temperature and equivalence ratio at 20% moisture content. Also, it was tested by comparing the calculation results of model with data from another researcher. To study the effect of moisture content of the wood biowaste, the amount of air was fixed at 0.3 of the stoichiometric requirement. Figure 3.1 shows the effect of moisture content in woody biomass on the product gas when the gasifier is operating at 800°C. It can be observed from the figure that the composition of methane produced is almost constant at a low percentage. The percentage of H<sub>2</sub> in the produced gas increases continuously with the moisture content from 0% to 40 %. A similar trend is also observed for the CO<sub>2</sub> however, the influence of the addition of more moisture will result in decrease in its concentration that might occur due to char, tar and soot formation, and water-gas shift reaction. Also, the concentration of CO reduces as moisture content increases. The results obtained from the model are in close agreement with experimental data from Zainal et al. (2001) which performed for wood biomass under similar operating conditions. Figure 3.2 shows the calorific value for wood biomass against moisture content. It is seen that the calorific values decrease with increases in moisture content. This is due to the greater reduction in carbon monoxide content compared to the increase in hydrogen content as moisture content increases.

Figure 3.3 explains the variation of calorific value against gasification temperature for the raw material wood with an initial moisture content of 20%. It can be observed that the predicted calorific values reduce with increase in temperature. Figure 3.4 shows the

comparison between model and experimental results for wood at 18.5 % moisture content and gasification temperature of 800°C. It can be observed from graph that the predicted value percentage of hydrogen and carbon monoxide content is predicted as 16.87% & 22.92% against an experimental value of 12.23% & 21.00%. The hydrogen percentage predicted is more than the experimental value, while the carbon monoxide percentage predicted is lower than the experimental result. Also, higher equivalence ratio (ER) result in higher operating temperatures, which in turn increases the degree of oxidation of available carbon.

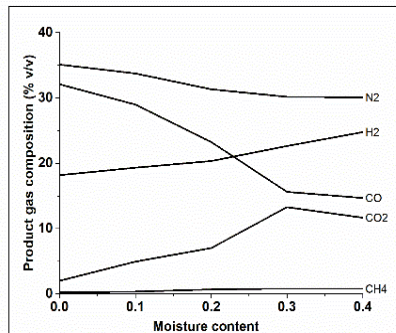


Figure 3.1: Effect of moisture content in wood biomass on product gas at 800°C

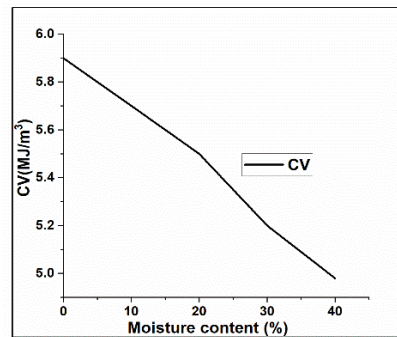


Figure 3.2: Effect of calorific value against moisture content at 800°C

The model is optimised using nonlinear programming solver IPOPT to find the maximum hydrogen production rate with the equivalence ratio and temperature as optimisation variables. Higher ER results in more CO<sub>2</sub> production because of oxidation of carbon; hence, the optimum ER, calculated by model is 0.1957 at 853.23°C and 20 % moisture, which yields 21.214% H<sub>2</sub>. So, by combining the two above-mentioned gasification variables, temperature, and equivalence ratio, it is possible to optimise the operating conditions to magnify the H<sub>2</sub> production.

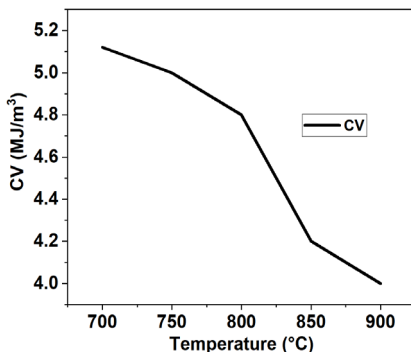


Figure 3.3: Calorific value against

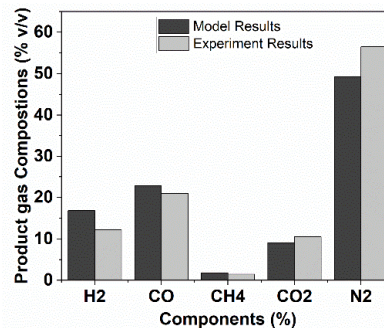


Figure 3.4: Comparison between model and experimental results at 800°C

#### 4. Conclusion:

Stoichiometric thermodynamic equilibrium models for the biomass gasification process were developed in this work, with the entire conversion of carbon considered to estimate syngas composition.

For a wide range of temperatures, the equilibrium constants of the water gas shift process and the methanation reaction have been computed. The models are formulated in an algebraic modelling language for optimisation of the gasification process. This model can predict the composition of syngas from biomass sources and can also predict biomass gasification performance under a variety of operating parameters, such as temperature, moisture content and equivalence ratio. The optimum ER, calculated by model is 0.1957 at 853.23°C and 20 % moisture, which yields 21.214% H<sub>2</sub>. The study leads to the conclusions that the amount of hydrogen in the product gas nearly rises linearly with the amount of moisture for biomass. However, carbon monoxide content in the product gas decreases with moisture content almost in a linear variation. The proportion of methane in the product gas is relatively low and increases linearly with moisture content. The calorific value of the producer gas decreases as the moisture content increase. In short, with a known ultimate analysis of any biomass and its calorific value, the model can predict the composition and the calorific value of the producer gas under different operating conditions.

## References

- Baruah, D., & Baruah, D. C. (2014). Modeling of biomass gasification: A review. *Renewable and Sustainable Energy Reviews*, 39, 806-815.
- Chaurasia, A. (2018). Modeling of downdraft gasification process: Studies on particle geometries in thermally thick regime. *Energy*, 142, 991–1009.
- Green, D. W. (1984). *Perry's Chemical Engineers' Handbook*, Eighth Edition by Don W. Green, McGraw-Hill Education. McGraw-Hill Education.
- Galvanin, F., Barolo, M., & Bezzo, F. (2009). Online Model-Based Redesign of Experiments for Parameter Estimation in Dynamic Systems. *Industrial & Engineering Chemistry Research*, 48(9), 4415–4427.
- Hart, W.E., Laird, C.D., Watson, J.P., Woodruff, D.L., Hackedbeil, G.A., Nicholson, B.L. and Sirola, J.D., 2017. Pyomo-optimization modeling in python, 67, 277.
- Janajreh, I., Adeyemi, I., Raza, S. S., & Ghenai, C. (2021). A review of recent developments and future prospects in gasification systems and their modeling. *Renewable and Sustainable Energy Reviews*, 138, 110505.
- Kushwah, A., Reina, T. R., & Short, M. (2022). Modeling approaches for biomass gasifiers: A comprehensive overview. *Science of The Total Environment*, 155243.
- Dogru, M. (2000) Fixed Bed Gasification of Biomass, Ph.D. thesis, University of Newcastle, UK.
- Pradhan, P., Arora, A., & Mahajani, S. M. (2019), A semi-empirical approach towards predicting producer gas composition in biomass gasification. *Bioresource Technology*, 272, 535-544.
- Sidek, F. N., Abdul Samad, N. A. F., & Saleh, S. (2020), Review on effects of gasifying agents, temperature, and equivalence ratio in biomass gasification process. *IOP Conference Series: Materials Science and Engineering*, 863(1).
- Yamany, W., Fawzy, M., Tharwat, A., & Hassanien, A. E. (2015, December). Moth-flame optimization for training multi-layer perceptrons. In 2015 11th International computer engineering Conference (ICENCO), 267-272.
- Zainal, Z. A., Rifau, A., Quadir, G. A., & Seetharamu, K. N. (2002). Experimental investigation of a downdraft biomass gasifier. *Biomass and Bioenergy*, 23(4), 283–289

# A model predictive control approach for recirculating aquaculture systems integrated with sustainable hybrid energy systems

Ruosi Zhang<sup>a,b</sup>, Tao Chen<sup>a</sup>, Michael Short<sup>a\*</sup>

<sup>a</sup>*University of Surrey, Stag Hill, Guildford GU2 7XH, UK*

<sup>b</sup>*National Innovation Center for Digital Fishery, Beijing 100083, China*  
*m.short@surrey.ac.uk*

## Abstract

Aquaculture is a growing industry that provides high-quality protein, however, the growth of aquaculture production has a range of environmental concerns. Recirculating aquaculture systems (RAS) are a popular solution to accomplish intensifying fish production, and integrating renewable technologies is an attractive option that may provide low-carbon energy and promote sustainability. This paper presents an optimisation model for the control and operation of a grid-connected distributed energy system (DES) integrated with an RAS. A linear optimal power flow (OPF) combined with model predictive control (MPC) strategy is developed which simulates the thermal and electricity balances present throughout the renewable energy integrated electricity network and RAS, to make real-time optimal operation scheduling plans on different time horizons. The optimal results show that the MPC strategy improves the control performance of system operation and improves the process economics. With renewable energy sources (RES), operational costs can be reduced by up to 27%. Compared with a conventional scheduling plan, the rolling horizon approach can provide 3% energy cost saving, while maintaining fish well-being and system safety.

**Keywords:** recirculating aquaculture system, distributed energy systems, renewable energy, optimal power flow, model predictive control.

## 1. Introduction

Aquaculture is a main source of protein and income in many countries globally (FAO, 2020). Sustainable aquaculture is directly related to targets of the Sustainable Development Goals. In this context, the implementation of RAS are gaining more attention globally (Chen et al., 2021). RAS has the potential to provide a sustainable fish-rearing system since it is easily expandable, potentially environmentally friendly, and improves production by utilizing a small space. RAS is competitive since it is able to recycle 90-99% water for reuse. However, RAS has received little attention in terms of understanding the synergies between energy consumption and fish production. The application of RAS is difficult to commercialise. The reasons are that the high initial investments are hard to recover (Schneider et al., 2006) and the high energy consumption for operation (Ayer & Tyedmers, 2009). RAS integrating renewable technologies is an attractive option providing low-carbon energy and promoting sustainability.

The use of solar energy in RAS decreases the energy cost on fossil fuel, especially for fish farms with higher water temperatures. The PV system combined with a wind turbine (WT) can enhance the reliability of the system and smooth the power generation curve caused by the intermediate characteristics of RESs (Ruggles & Caldeira, 2022). Jamroen

(2022) developed design models to determine the appropriate size of PV and battery storage for fish farms. These models have common disadvantages. The first is that uncertainty of renewable energy is not considered, and another is the timescale considered. In real-time, the scheduling of the power system could be in minutes, and in these studies the scale of days is used, wherein it is possible that the system may not respond to changes quickly. Silvente et al. (2015) presented a linear operational model considering the prediction uncertainties from RESs, which is solved by an MPC rolling horizon strategy to correct the deviations from the initial conditions. MPC rolling horizon strategy is suitable for dealing with system control problems that are easy to model and have high stability. MPC includes four parts, a prediction model, rolling optimisation, feedback mechanism and reference trajectory. The feedback mechanism corrects the prediction model, and the error is compared with the reference trajectory. The final control decision applied to the system is calculated by the rolling optimisation.

This paper proposes a linear MPC approach for sustainable hybrid energy systems for RAS. The model simulates the thermal and electricity balances present throughout the renewable energy integrated electricity network and RAS, to make real-time optimal operation scheduling plans on different time horizons. First, a multi-time scale optimal scheduling problem is divided based on three scales: day-ahead long-term scale scheduling in 1 h, intra-day predictive control in 15 min and real-time adjustment in 5 min. Then, we establish an optimisation model based on the best economic operation of the system and the minimum penalty cost of start-up and shutdown. The inputs include the relevant climate parameters for renewable energy generation, the energy storage schedules, operation of facilities including climate and oxygen control equipment, as well as the fish biomass production quantity.

## 2. Model formulation

The components of the system are divided into load demand of RAS and power resources including generator, WT, PV and national grid as backup power in on-grid operation mode. The main load demand of the system is the thermal process to keep water temperature as required in RAS, and other demands such as lights, water pumps, water treatment system and aeration are treated as constant load demand in this model. The overall objective of this model is to minimise the daily operation cost:

$$OPERATION = COST_{gen} + COST_{grid} \quad (1)$$

$$COST_{grid} = \sum_t (P_t^{grid, buy} * C^{grid} - P_t^{grid, sold} * C^{tariff}) * \Delta t \quad (2)$$

$$COST_{gen} = \sum_t ((P_t^{gen} * F_1 + Y_{gen} * F_0) * C^{fuel} + S_{gen} * u^t) * \Delta t \quad (3)$$

The cost of buying electricity from the grid is  $COST_{grid}$ , with  $C^{grid}$  and  $C^{tariff}$  the tariffs of electricity grid.  $P_t^{grid, buy}$  and  $P_t^{grid, sold}$  are the power of buying or selling electricity to the grid respectively,  $\Delta t$  is the time interval. The cost of generator's operation is  $COST_{gen}$ ,  $P_t^{gen}$  is the output power of generator,  $Y_{gen}$  is the capacity of the generator,  $F_1$  is generator fuel curve slope,  $F_0$  is generator fuel curve intercept coefficient,  $C^{fuel}$  is the price of generator fuel,  $S_{gen}$  is the cost of start-up generator, and  $u^t$  is the start-up variable of generator.

$$\sum_t P_t^{gen} + W_t + P_{ess,t} + P_t^{grid} + PV_t - P_t^{HP} - P_t^{eb} = P_t^{load} \quad (4)$$

$W_t$ ,  $P_{ess,t}$ ,  $P_t^{grid}$ ,  $PV_t$ ,  $P_t^{HP}$ ,  $P_t^{eb}$ ,  $P_t^{load}$  are the output power of WT, battery, main distribution, PV, heat pump, electric boiler, and load demand of RAS respectively. The

output power of wind turbine and PV should be lower than the predicted output value according to the weather. Unit commitment:

$$\begin{aligned} x^{t-1} - x^t + u^t &\geq 0 \\ x^t - x^{t-1} + v^t &\geq 0 \end{aligned} \quad (5)$$

where  $x, u, v$  are binary variables.  $x$  is the unit commitment of generators (1 if the generator is committed at time  $t$ , and 0 otherwise).  $u^t$  is the start-up variable of generator (1 if generator starts at time  $t$ , and 0 otherwise).  $v^t$  is the start-up variable of generator (1 if the generator starts at time  $t$  and 0 otherwise).

The model also ensures that electricity is not bought from or sold to the grid at the same time through big-M constraints, which are similar to the constraints on the battery's charging or discharging status. For proper operation, each power source must not be operated above its rated or maximum limit according to the manufacturer's specifications. The state variable also has boundaries where the minimum depends on the type of battery used.

The water heat balance is used to calculate the energy flows of the water volume each hour. In other words, the water heating or cooling demand to maintain its temperature. The formulated simplified heat balance is:

$$Q_{conv}(t) + Q_{exchange}(t) = Q_{demand}(t) \quad (6)$$

The convective heat transfer,  $Q_{conv}$  represents the heat exchange on the surface of the tank between water and air, which is determined by (Holman, 2009):

$$Q_{conv} = h_{conv,w} \cdot A_{surface} \cdot (T_{air}(t) - T_w(t-1)) \quad (7)$$

where  $h_{conv,w}$  ( $kW/(m^2 \cdot ^\circ C)$ ) is the convective heat transfer coefficient decided by indoor air velocity,  $T_w$  is the water temperature,  $T_{air}$  is the air temperature, and  $A_{surface}$  ( $m^2$ ) is the surface area of the tanks. The water exchange that occurs in the system implies a heat gain or loss of  $\dot{Q}_{exchange}$ . The rate of water exchange in the system is assumed to be constant, with the same amount of water entering and leaving the system. The heat gain through water exchange can be calculated by:

$$\dot{Q}_{exchange}(t) = \dot{m}_{exchange} \cdot c_{p,w} \cdot (T_{w,in} - T_w(t-1)) \quad (8)$$

where  $\dot{m}_{exchange}$  is the mass flow of the water exchange,  $c_{p,w}$  is the specific heat of the water and  $T_{w,in}$  is the temperature of water entering the system.

The heating system will be active until a certain water temperature has been reached. When the temperature set point has been reached the heating system will be turned off, the water temperature will then drop to a lower limit where the heating system is activated again.

The core idea of MPC is rolling optimization and feedback correction, which solve the optimization problems with uncertainties, and can also conveniently include a variety of constraints. The rolling optimization process of MPC is shown in Figure 1. If the system operates at time  $t$ , based on the prediction information of the system disturbance variables in a future period, obtain the optimal outputs in the prediction domain with the objective of minimizing the operating cost. However, the optimization result is used only in one control domain, and the optimization process is repeated at time  $t + 1$ . The proposed scheduling is divided into three timescales: day-ahead, intraday, and real-time. The main decisions of the day-ahead dispatching are the unit start-up and shutdown plan in the next 24 hours, the charging and discharging state of energy storage, and the power purchase and sale to the main distribution network. The time interval in this model is 1 h. The intraday scheduling mainly corrects the output of the controllable units, the charging and discharging power of the energy storage and the purchasing and selling power of the main distribution grid in the intraday dispatching according to the latest renewable energy power output and thermal load forecast data. The time interval of this model is 15min and



the dispatching period is 2h. The time interval of real-time dispatching is 5min, and the dispatching period is 15min. It is mainly based on the real-time forecast data of wind power, PV output and thermal load to finetune the intraday scheduling plan. In the intraday and real-time dispatching time, the unit start-up and shutdown plan, energy storage charging and discharging status and power purchase and sale status of the main distribution grid in the day-ahead scheduling results are kept unchanged. The level-by-level coordination of the three-time scales can better deal with uncertainties and make scheduling results more consistent.

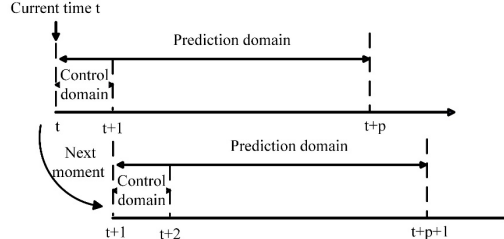


Figure 1 The rolling optimization process of MPC.

When optimising intraday and real-time scheduling, there is often a large error between the renewable energy output obtained by the short-term forecasting method and the daily forecast data, so the daily scheduling plan needs to be properly modified to ensure the effectiveness of the scheduling strategy. The objective function of intraday optimal scheduling model is:

$$\min \text{Operation}_{intra} = \min \sum_{t=1}^{N_T} [C_{G,t}^{intra} + C_{ESS,t}^{intra} + C_{M,t}^{intra}] \quad (9)$$

where  $N_T$  is intraday scheduling period,  $N_T = 2h$ ;  $C_{G,t}^{intra}$  is the generator operation cost at time  $t$ ;  $C_{ESS,t}^{intra}$  is the storage operation cost,  $C_{M,t}^{intra}$  is the operation cost of intraday purchasing or selling electricity from main distribution network.

During the intraday scheduling, the status of the generator is consistent with that of the day before, so the intraday scheduling cost of the controllable generator unit only includes the operation cost. The charging and discharging power of energy storage should follow the day-ahead scheduling results as much as possible. Therefore, intraday scheduling cost of energy storage is expressed as:

$$C_{ESS,t}^{intra} = \left\{ K_{ESS,i} \left[ \frac{1}{\eta_{E,dis}} P_{E,t}^{sum,dis} + \eta_{E,ch} P_{E,t}^{sum,ch} \right] + \lambda_{ESS}^{intra} |SOC_t - SOC_t^{ahead}| \right\} \Delta t \quad (10)$$

where  $\lambda_{ESS}^{intra}$  is the weight coefficient of energy storage adjustment during the intraday plan. Equation (10) has absolute value and auxiliary variables  $S_{E,t}^{au}$  are introduced and linearized:

$$\begin{aligned} C_{ESS,t}^{intra} &= \left\{ K_{ESS,i} \left[ \frac{1}{\eta_{E,dis}} P_{E,t}^{sum,dis} + \eta_{E,ch} P_{E,t}^{sum,ch} \right] + \lambda_{ESS}^{intra} S_{E,t}^{au} \right\} \Delta t \\ \text{s.t. } S_{E,t}^{au} &\geq SOC_t - SOC_t^{ahead} \\ S_{E,t}^{au} &\geq -[SOC_t - SOC_t^{ahead}] \end{aligned} \quad (11)$$

Similarly, the cost of intraday power purchase and sale of the main distribution network is the same as that of energy storage and also follows the day-ahead scheduling results as much as possible. During intraday scheduling, to track the day-ahead charging and discharging power of storage energy, the following constraint should be satisfied:

$$SOC_{4t}^{intra} = SOC_t^{ahead} \quad t = 1, 2, \dots, 24 \quad (12)$$

where  $E_{4t}^{Intra}$ ,  $E_t^{ahead}$  are the state of charge at the end of each hour for intraday and day-ahead scheduling respectively. Real-time optimal scheduling is similar to the intraday model, tracking the intraday dispatching plan as much as possible.

The process is summarized in Algorithm 1.

**Algorithm 1 Model predictive control**

1. Predict values for day ahead including renewable energy output, load demand. Input other parameters.
2. Solve the day-ahead model, with objective function (1), constraints are (2)-(8). Decide the binary variables of generator and battery.
3. Predict values for short-term scheduling, keep binary decision variables from day-ahead model, use rolling optimization, and objective function (9), solve intra-day model. Get the scheduling plan for intra-day.
4. Predict values for real-time, keep binary decision variables from day-ahead model, use rolling optimization, solve real-time model. Based on intra-day plan, adjust the output of each unit.

**3. Case study**

The test case considers a fictional RAS located in Lochinver, Scotland, where fish farming is a vital industry, as they are the largest producer of farmed salmon in the European Union. The system uses a heat pump (HP) and electric boiler to keep the water temperature in the feed tanks of RAS ranging 16°C to 17°C. The capacities of PVs, batteries, generator, and WT installed are decided by the solution of a design optimisation model, shown in Table 1. The results presented in Table 1 are used as inputs in this model, and a typical autumn day has been chosen specifically to test the model as renewable energy power generation is highest in this season.

Table 1 Model results summary.

PVs (kW)	Battery(kW)	Wind Turbines(kW)	Generator(kW)
13	7.7	19	35

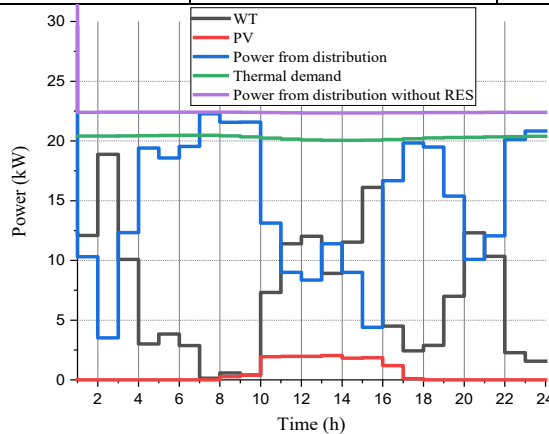


Figure 2 The scheduling plan of each unit in the system for day-ahead.

First, the day-ahead scheduling plan is shown in Figure 2. Due to the influence of geographical environment and weather, the output of the WT is far greater than that of PV. When the output of renewable energy is insufficient, the demand is met by purchasing power from the main distribution instead of starting up the generator, and the SOC of the battery is basically unchanged. If the parameters are changed, the electricity price is increased and the fuel price is reduced, the generator power generation will make up for

the shortage of renewable energy, and the SOC of the battery will also change. The thermal demand is maintained at a relatively stable value, because of the high specific heat capacity. Compared with the original system without RESs, the daily operation cost can be reduced by about 27%, from £144.5 to £105.

Considering the disturbance of renewable energy output, rolling optimisation of MPC is introduced into the model. In intraday and real-time dispatching, to keep the energy balance of the network, the system prefers to purchase/sell the electricity from the main network rather than batteries, which is due to the lower electricity price compared with the use price of energy storage. Faced with renewable energy disturbance, the scheduling change of the MPC rolling algorithms is more stable and has less variation, which enhances the system's ability to cope with uncertainties. By using the rolling-horizon MPC model, the daily running expenses can be reduced by 3%.

#### 4. Conclusion

RAS as a popular fish production system requires high energy demand due to thermal processes, water circulation, and aeration. This paper presents an optimisation model for the control and operation of a grid-connected distributed energy system integrated with an RAS, that is developed to optimise the performance of the system in economic, sustainable, and technical aspects in real time. An MPC strategy for a sustainable hybrid energy system for the RAS is proposed in this work. To analyse the performance of the model, the optimal results of the proposed system were compared with the traditional system in economic and environmental aspects. The optimal results show that the MPC strategy improves the control performance of system operation and improves the economy of system operation. Compared with a conventional RAS, the integration of renewable energy can save up to 27% operation cost in a typical day. With the MPC rolling-horizon algorithm, the system's ability to cope with uncertainties has been improved.

#### References

- Ayer, N. W., & Tyedmers, P. H. (2009). Assessing alternative aquaculture technologies: Life cycle assessment of salmonid culture systems in Canada. *Journal of Cleaner Production*, 17(3), 362–373.
- Chen, Z., Ye, Z., Ji, M., Zhou, F., Ding, X., Zhu, S., & Zhao, J. (2021). Effects of flow velocity on growth and physiology of juvenile largemouth bass (*Micropterus salmoides*) in recirculating aquaculture systems. *Aquaculture Research*, 52(7), 3093–3100.
- FAO. (2020). *GLOBEFISH Highlights 2019 Statistics: A quarterly update on world seafood markets*. FAO. <https://www.fao.org/documents/card/en/c/ca7459en/>
- Holman, J. (2009). *Heat Transfer* (10th edition). McGraw-Hill Education.
- Jamroen, C. (2022). Optimal techno-economic sizing of a standalone floating photovoltaic/battery energy storage system to power an aquaculture aeration and monitoring system. *Sustainable Energy Technologies and Assessments*, 50, 101862.
- Ruggles, T. H., & Caldeira, K. (2022). Wind and solar generation may reduce the inter-annual variability of peak residual load in certain electricity systems. *Applied Energy*, 305, 117773.
- Schneider, O., Blancheton, J., Varadi, L., Eding, E., & Verreth, J. (2006). Cost price and production strategies in European recirculation systems. *Conference: Aqua 2006: Linking Tradition & Technology Highest Quality for the Consumer, Firenze (Florence), Italy, 9-13 May 2006*, 855.
- Silvente, J., Kopanos, G. M., Pistikopoulos, E. N., & Espuña, A. (2015). A rolling horizon optimization framework for the simultaneous energy supply and demand planning in microgrids. *Applied Energy*, 155, 485–501.

# Multi-stage optimization for marketing industrial flexibility

Martin Fischer<sup>a</sup>, Karl-Wilhelm Schenzel<sup>a</sup>, René Hofmann<sup>a</sup>

<sup>a</sup> *Institute of Energy Systems and Thermodynamics, TU Wien, Getreidemarkt 9/302,  
1060 Vienna, Austria*

*[martin.fischer@tuwien.ac.at](mailto:martin.fischer@tuwien.ac.at)*

## Abstract

This work presents a method based on Mixed Integer Linear Programming (MILP) to consider wholesale energy markets in the operational optimization of industrial energy systems. A two-stage optimization framework is presented to consider price differences between the day-ahead and the intraday spot markets to optimize the electricity procurement in both markets. A representative use case of an industrial process is used to optimize the energy supply from the power generators and storage units and minimize energy costs as well as to benefit from participating in the wholesale market. The results show potential to save energy costs and make additional profits on the intraday market.

**Keywords:** Flexibility, Optimization, MILP, Intraday Market

## 1. Introduction

A rising share of renewables and an increasing number of volatile power generators and consumers are causing imbalances and congestions in the power grid. With decarbonization and the demise of conventional power plants, our energy system needs new flexibility measures and market participants. In most countries, the industrial sector is one of the largest electricity consumers and can therefore play a significant role in supplying flexibility to balance the power grid and provide security of supply. Previous work (Schenzel, et al., 2022) presented a method to identify existing flexibilities based on a MILP approach (Halmschlager et al., 2020), which is integrated into this work to consider the participation in wholesale markets, specifically focusing on the day-ahead and intraday markets.

## 2. Leveraging Energy Markets

Conventional optimization methods focus on high efficiency and low operational costs, but neglect opportunities by not considering energy markets. Most industrial customers procure their energy through energy suppliers, with fixed or floating electricity tariffs that reduce both the actual effort needed to operate on the wholesale market and the exposure to price volatilities (Schnorr 2019). At the same time, there are opportunities for industrial customers to make profits by marketing their flexibility - either by directly acting on the wholesale market or partnering with aggregators to have access to the wholesale markets. The advancement in energy storage technologies and the staggering increase in price spreads of wholesale markets support the case for industrial plant operators to act as a Flexibility Service Provider (FSP).

### *2.1. Flexibility*

Villar, Bessa, and Matos (2018) define flexibility as the possibility of modifying generation and/or consumption patterns in reaction to an external signal (price or activation signals). Generally, industrial energy systems have three ways of providing flexibility. Firstly, they can adjust their production schedule. This is highly dependent on the processes involved and often needs more lead time and computational effort. Secondly, energy carriers can be substituted, meaning that in order to meet a certain electricity demand, it can either be consumed from the power grid or produced by generators such as gas turbines. Thirdly, energy storages can be used to shift energy demands in the process. While conventional power plants usually store electric energy by converting it to other forms of energy and backwards, with significant losses each conversion, industries can directly utilize the converted energy - like heat - to cover the demand of their production processes. Also, thermal heat storages are cheap in comparison to batteries and can be used to make industrial processes more efficient by reducing waste heat. Once identified, the electric industrial flexibility can be offered as bids on the intraday market at an electricity exchange like the European Power Exchange (EPEX) or as an ancillary service at the balancing markets of the Transmission System Operator (TSO).

### *2.2. Balancing Markets*

These markets are used for the operating reserve to restore the frequency of the electricity grid in case of imbalances of the system. There are three different products available, Frequency Containment Reserve (FCR), automatic Frequency Restoration Reserve (aFRR), or manual Frequency Restoration Reserve (mFRR). Each product has their own requirements, pricing, and activation times. The remuneration model for aFRR and mFRR consists of two parts: one part is for keeping the capacity available (capacity price), and one is for the actual activation (energy price). Depending on the product, the activation must follow very shortly after the activation signal. Since these activations are very hard to predict they can interfere with production planning and safety.

### *2.3. Intraday Markets*

To compensate for changes in energy production, often caused by insufficient weather forecasts on the day-ahead, energy companies can buy or sell energy to level their position. This can be done during the day and up to several minutes prior to delivery. With the growing number of volatile renewables, these markets have gained much importance in the last decade with more energy being traded at these markets each year. FSPs can provide their flexibility to the market, to increase liquidity, reduce the amount of balancing energy and to make additional profits.

## **3. Modelling Approach**

Mathematical optimization represents a powerful method to maximize profits for power plant operators, often referred to as Unit Commitment (UC) problem. UC problems - a subclass of optimization problems - have been widely used in the energy market for many years. Industrial energy systems often face similar problems, when deciding when and how to utilize their power, heat, or steam generators to cover their demand in the most efficient way, especially since decentralized energy systems tend to get more diverse. In contrast to conventional UC problems of power plant operators, industrial plant operators are also constrained by their own operational targets, e.g., minimum

production output, minimum up-times for batch processing, working hours, etc. For our purposes a fixed set of demand time series is sufficient, meaning that there is a predetermined production plan, the demand for different energy carriers such as steam on different pressure levels, gas and electricity can be derived. Scheduling is often impractical, since it means a bigger intervention in the operation and is therefore not considered. With the increase of digitization and the development of energy control systems, a focus on optimizing the energy management system is more applicable and reduces interference to operation to a minimum. In the following we present a method to include the intraday market in the optimization problem.

### *3.1. Optimization Method*

The problem is formulated as a UC problem on the basis of MILP formulations, which are implemented in Matlab<sup>®</sup> and solved with a state-of-the-art solver (e.g. GUROBI). The model uses a combination of linear, continuous, semi-continuous and binary decision variables. They represent all operation trajectories of the energy supply system of the industrial plant, such as the power generated by the gas turbine, the charging and discharging of storages etc. Constraints are used to define the optimization problem, to depict operational constraints such as the satisfaction of energy demands, operational limits, power outputs or ramp-up/down constraints. The objective is to minimize operational costs, by optimizing the trajectories of the individual generating units of the industrial plant. This is done by utilizing the storages to meet the demands in the most efficient way and considering price curves of power spot markets for a cost-efficient operation plan. To better represent the market mechanisms of spot markets, the method is split in two optimization runs or “stages”.

### *3.2. Two-stage Optimization*

Figure 1 shows the two-stage optimization method to consider price signals between the day-ahead and intraday spot markets. Based on historical data, we developed a regression model to predict day-ahead market prices, based on weather forecasts and the previous day’s market coupling results. In Stage 1 (D-1), the predictions for the day-ahead prices are used to calculate an optimal operation plan using MILP. Based on the results, we derive the electricity demand procured directly via placing price inelastic bids on the day-ahead market. The industrial plant acts as a price taker, meaning that the orders are executed independent the market price and all bids are accepted. In Stage 2 (D-0), the results from the day-ahead auction and the actual weighted average intraday market prices are used as input to solve a second optimization problem. We assume that price signals from the continuous intraday market are available at a single point in time and the weighted average result from each quarter hour can be used. Considering these price signals the plant operator can buy or sell additional electricity at a 15-minute resolution to make additional savings and reduce energy costs.

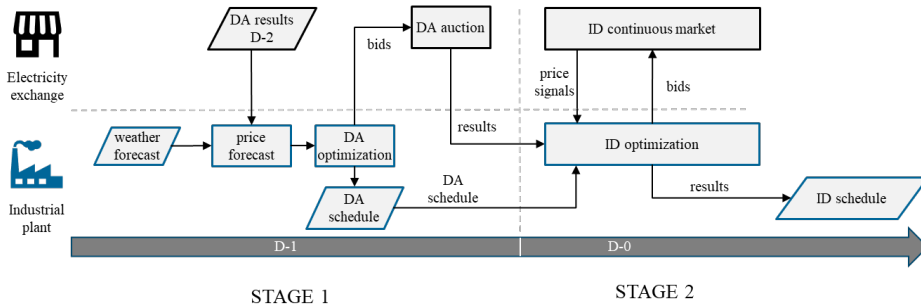


Figure 1: The two-stage optimization framework for day-ahead (DA) and intraday (ID) optimization.

## 4. Use Case

To demonstrate the proposed two-stage method, a simplified industrial energy system is modeled. The model consists of two demand curves, representing the power demand and steam demand, which result from a production plan. The power demand and the steam demand are fluctuating during the day but are set to a mean value of 1 MW. Both demands can be covered to a certain degree by a combined heat and power generator (CHP) with a maximal output of  $0.5 \text{ MW}_{\text{el}}$  and  $1 \text{ MW}_{\text{therm}}$ . A gas boiler can produce additional steam to cover the full steam demand. Additionally, a thermal and an electrical energy storage increase the efficiency and provide flexibility to the system.

### 4.1. Data

We used historical price data from the European Power Exchange (EPEX) spot market and the European Network of Transmission System Operators for Electricity (ENTSOE). We used hourly day-ahead market coupling results and weighted average results from the continuous intraday market for the Austrian bidding zone. Data from the year 2021 was used to train the linear regression model while data from the first half of 2022 was used to validate the regression model and to predict the daily day-ahead prices for the optimization framework.

### 4.2. Optimization Results

Figure 2 shows the optimization results of stage 2 for a single day in February 2022. The upper part of the plot shows the optimal trajectories to meet the electric power and steam demand of the production processes in the most cost-efficient way. Since the CHP has a limited electric power output, the industrial plant is dependent on power supplied by the public grid. The battery is charged and discharged to shift power demands to times of lower energy prices. Part B of Figure 2 shows the bids that were placed on the day-ahead and intraday markets to procure the power supply according to the optimal trajectories.

Figure 3 shows the potential energy cost savings for each day in correlation with the daily variance of the price spreads between the day-ahead and the intraday market. The savings are calculated by the difference of the cost function between stage 1 and stage 2. The mean value of the relative cost savings is 1.96 %. Generally, a higher price spread between these markets has shown a higher potential of saving energy costs.

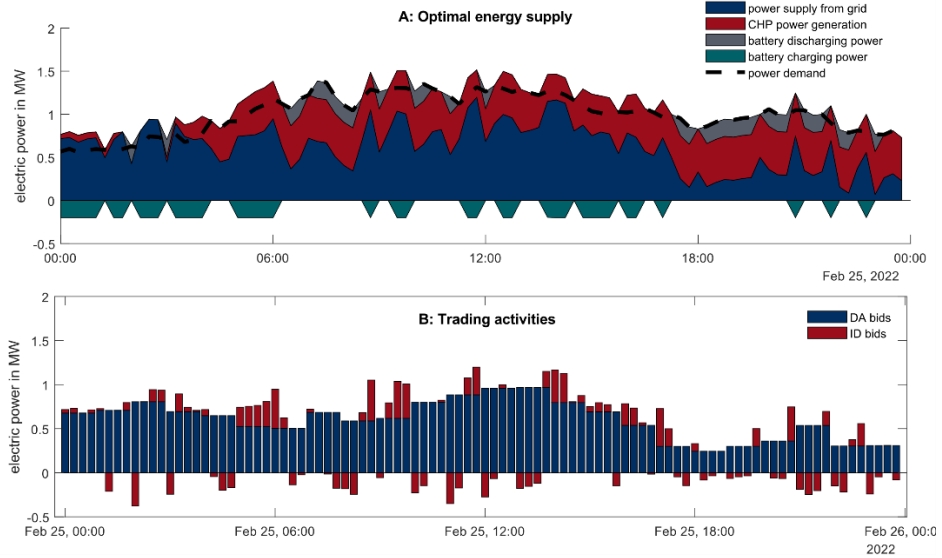


Figure 2: Optimization results for a single day. Part A shows the optimal operation trajectories of the energy supply system. Part B shows the corresponding bids on the day-ahead (DA) and intraday (ID) market.

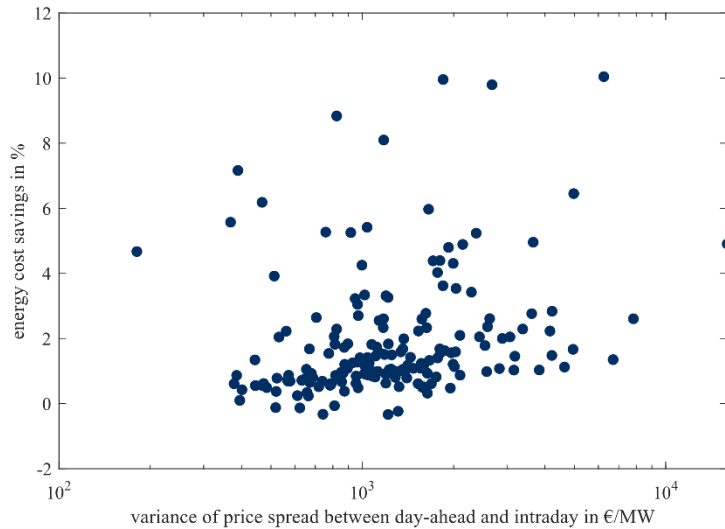


Figure 3: Relative energy cost savings per price spread



### 4.3. Discussions

The intraday bids in Figure 2 B are the corresponding reaction to price differences between the day-ahead and intraday markets. These can be interpreted as price signals from the current market situation and as reactions to deviations of the predicted day-ahead price curves that are used for optimization in stage 1.

By using the average result of the intraday market, we assume perfect foresight of the intraday market prices in the second stage. These prices are hard to predict though and are not traded at a single point in time, but continuously during the day. Thus, it may be possible to additionally exploit the volatility in this market by using the flexibilities of industrial energy systems.

## 5. Conclusion

The proposed two-stage optimization method enables a deeper integration of industrial energy systems in the short-term energy market. Hereby, industrial energy systems can react to fluctuations in the system in the form of price signals in the intraday market. Using real price data from the first half of 2022, our method also shows significant potential for reducing energy costs. Assuming direct access to wholesale market, the intraday market offers a supplementary opportunity for marketing flexibility, as the entry barriers are lower than in other markets such as balancing or redispatch markets. Further research will show the possible exploitation of volatility in the intraday market. In the prospect of reduced entry barriers for balancing markets and new opportunities for industrial energy systems to act as FSPs, mathematical optimization methods based on MILP may play an important role when participating in flexibility markets.

## Acknowledgement

This work was supported by the project “Industry4Redispatch”, which is part of the energy model region NEFI - New Energy for Industry and is funded by the Austrian Climate and Energy Fund [FFG, No.887780].

## References

- Halmschlager, Verena, and René Hofmann. 2020. „Flexibility identification of an industrial production“. In *Computer Aided Chemical Engineering*, 48:1375–80. Elsevier. <https://doi.org/10.1016/B978-0-12-823377-1.50230-5>.
- Schenzel, Karl, Fischer, Martin, Zlabinger, Erwin und René Hofmann. 2022. „Holistic Approach for the Optimization of Industrial Hybrid Energy Hubs with MILP“. In *Proceedings of the 2nd NEFI Conference 2022*, 21:30. NEFI.
- Schnorr, Stephan. 2019. *Energiebeschaffung in Industrieunternehmen: erfolgreiches Agieren am Energiemarkt*. Essentials. Wiesbaden [Heidelberg]: Springer Gabler. <https://doi.org/10.1007/>.
- Villar, José, Ricardo Bessa, und Manuel Matos. 2018. „Flexibility Products and Markets: Literature Review“. *Electric Power Systems Research* 154 (Januar): 329–40. <https://doi.org/10.1016/j.epr.2017.09.005>.

# Intelligent Process Flowsheet Synthesis and Design using Extended SFILES Representation

Vipul Mann,<sup>a</sup> Rafiqul Gani,<sup>b,c</sup> Venkat Venkatasubramanian,<sup>a</sup>

<sup>a</sup>*Columbia University, New York, NY 10027, USA*

<sup>b</sup>*PSE for SPEED Company, Charlottenlund, DK-2920, Denmark*

<sup>c</sup>*Sustainable Energy and Environment Thrust, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China*

## Abstract

A central problem in chemical engineering (and several related areas) is evaluating the correct sequence of unit operations, their design aspects, and the continuous optimization of their operations to efficiently convert input materials to final products. The numerous decisions to be made at each problem-solving stage renders this problem combinatorically complex. In this work, we propose a hybrid, artificial intelligence-based multi-level framework to perform fast, efficient, and reliable flowsheet design and optimization. We build upon the previously proposed SFILES-based text representation of flowsheets to incorporate additional contextual details in the extended SFILES framework using hypergraph representations. We discuss our eSFILES framework using the well-known hydridealkylation (HDA) process.

**Keywords:** process design; flowsheet prediction; artificial intelligence; computer-aided flowsheet synthesis

## 1. Introduction

Process flowsheet synthesis and design is a challenging task that involves identifying the correct set of unit operations and their optimal sequence that enables the conversion of input materials to desired output products, while considering energy consumption, environmental impact, safety, operability, and many more. The goal of synthesis is to identify the tasks, the operations that will perform these tasks, and their sequence to determine the flowsheet. The goal of design is to add operational and equipment details to the flowsheet so that the process can be verified, optimized, built, and operated. The size of the synthesis problem varies with the number of tasks and the number of alternative operations for each task. The complexity of the design problem is related to matching the tasks and their associated operations with design parameters so that the designed process matches the desired process specifications. The design parameters are further optimized to obtain a sustainable process design.

Model-based methods are usually used for the synthesis and design stages. Although, in principle, the synthesis, design and optimization steps could be performed simultaneously, the current practice is to solve them separately due to the size and complexity of the resulting mathematical problem. The challenge here is how does one incorporate issues, such as, economics, environmental impacts, operability, safety, and sustainability in the early stages of process synthesis and/or design (Tula et al., 2017).

Computer-aided approaches using AI-inspired methods combined with fundamental concepts offer several advantages over purely data-driven methods, which may lead to infeasible flowsheets, or purely model-based methods, which may run into difficulties in the numerical solution step due to the complexity of the models used.

To facilitate the development of such hybrid AI-based approaches combining process knowledge with computational algorithms, an appropriate flowsheet representation is needed that is concise, complete, and accurate. The SFILES representations developed by (D'Anterrosches, 2005), (Bommareddy *et al.*, 2011), and (Tula *et al.*, 2015) was the first step in this direction, and has been shown to have various applications, such as, flowsheet autocompletion (Vogel *et al.*, 2022), piping and instrumentation diagram generation (Hirtreiter *et al.*, 2022), and flowsheet pattern mining (Zhang *et al.*, 2019). The SFILES strings represent correctly and consistently a wide range of process flowsheets, involving typical operations found in chemical and biochemical processes. Note that, like the SMILES strings for a molecule, a parser is needed to convert SFILES strings to the actual process flowsheet diagrams. As originally developed, a process flowsheet is first represented by a set of process groups, which are similar to the functional groups that represent a molecule. The process group representation is then converted to a SFILES representation. Unlike the molecular representation with SMILES, several additional details need to be considered. For example, information about the number of chemicals and their effect on the system behavior; the direction of flow-paths for reactants, products, inerts, solvents, etc. need to be tracked; and, start and end of the process need to be clearly marked. These additional issues give opportunities for a symbolic AI-based intelligent system (Venkatasubramanian & Mann, 2022) need to be incorporated to the current process synthesis-design methods such that the application of SFILES to represent process flowsheets as well as its use in computer-aided process synthesis and design can be extended.

The objective of this paper is to present a hybrid multi-level AI framework for fast, efficient, and reliable flowsheet synthesis and design taking into account concepts and theory on chemical process development, together with the knowledge and data for hundreds of process flowsheets that are known already. We present an extended SFILES representation (eSFILES) based on SFILES of process flowsheets ((D'Anterrosches, 2005); (Bommareddy *et al.*, 2011); (Tula *et al.*, 2015)) and the annotated hypergraph representation, developed originally to study networks in organic chemistry (Mann & Venkatasubramanian, 2023). We highlight selected developments where we illustrate the hierarchical eSFILES representation framework for any chemical process flowsheet, also suitable for process synthesis, design, and innovation (Tula *et al.*, 2015). Namely, we propose a three-level, extended SFILES framework where the lowest level contains flowsheet connectivity information which could be used to generate text-based flowsheet representations like SFILES but with additional information; the middle level contains additional details on process groups and streams providing information necessary information for process synthesis (superstructure-based optimization or process group based enumeration and test); and the top level contains process operational data in terms of design parameters for process simulation and innovation through optimization and intensification. At the core of our approach lies a hypergraph representation of flowsheet (Mann & Venkatasubramanian, 2023) that represents process groups as hyperedges and streams as nodes with annotations indicating contextual information.

## 2. Current state of flowsheet representation: SFILES

A text-based representation for chemical flowsheets reported by (Tula et al., 2015) is called simplified flowsheet input line entry scheme (or SFILES) analogous to the SMILES representation for molecules. Such a representation offers a concise way of representing flowsheets and could also facilitate not only storage of flowsheets, but also, enumeration, analysis, and deployment for interactive text-based algorithms process synthesis, design, and simulation of a given flowsheet. (Tula et al., 2015) developed separate rule-based algorithms for process synthesis and design, whose solutions were not added to SFILES strings. The underlying idea behind SFILES is to identify analogous process components similar to a SMILES string as follows –

- represent processes as *process groups* akin to functional groups in molecules.
- represent process flow streams akin to chemical bonds in molecules.
- represent various connections, recycle streams, and branching of streams akin to branching in molecules.
- unlike SMILES, SFILES need special unidirectional input nodes and output nodes that indicate the start and end of a process flowsheet, through special characters ‘i’ and ‘o’, respectively.

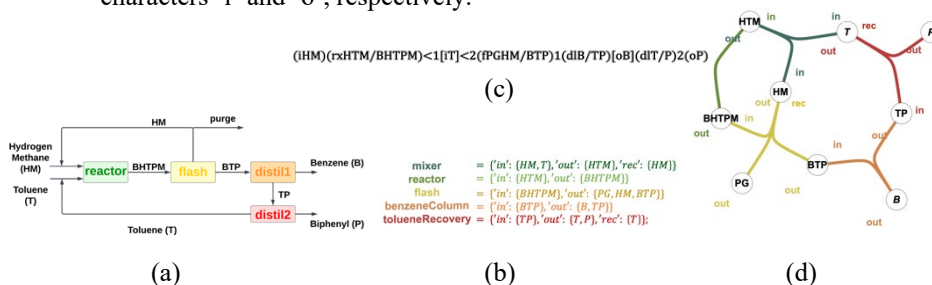


Figure 1: (a) HDA process flowsheet, (b) the various process groups, (c) the corresponding SFILES string, (d) HDA process represented by process groups. In the above subplots, the abbreviations refer to – mixture of hydrogen, methane (HM); mixture of benzene, hydrogen, toluene, biphenyl, methane (BHTPM); mixture of benzene, toluene, biphenyl (BTP); benzene (B); biphenyl (P); toluene (T); purge gas (purge). The hypergraph was generated using Wolfram Mathematica and algorithmic implementation would be done in Python.

We use the well-known hydrodealkylation (HDA) process to illustrate the basic concepts and the new developments. First the process flowsheet is represented by a set of process groups retrieved from the database (the process group representation is shown in Fig 1b) and the parser is used to generate the SFILES string (shown in Fig 1c). The reader is referred to (Tula et al., 2015) for detailed examples on SFILES generation.

## 3. Extended SFILES (eSFILES)

**Hypergraph:** The hypergraph representation of the HDA process is shown in Figure 1(d). While this new SFILES representation offer numerous benefits, there is scope for further improvement and incorporation of additional details that would aid in intelligent process flowsheet synthesis and design. The new information that, when incorporated, would generate additional, more complete information in the flowsheet representation such as:

- information on incoming/outgoing streams for each process group explicitly, especially when multiple streams enter or leave a process group.

- the minimum data needed to completely define a material and/or energy stream.
- information on driving force for the process groups indicating the ease/difficulty of conversion/separation.
- ability to systematically generate superstructure of flowsheet and enumerate valid alternatives aiding in the design problem.

Towards this end, we propose the extended SFILES (or eSFILES) representation that provides a framework for incorporation of this information in a hierarchical manner and further details on this are presented in the subsequent sections.

**Hierarchical framework:** A hierarchy is needed to efficiently organize process knowledge needed for different applications and to guide users to select, retrieve, and/or collect the data needed for a specific application. Namely, the eSFILES hierarchy has three levels of information -- first at the lower level, information needed to only represent the flowsheet (connectivity); next, at the middle-upper level, introduce additional details such as driving force and mixture composition that helps to establish the material and energy flows; and finally, the top level that contains information on composition and material/energy balance that would be needed to solve a flowsheet among other applications. The eSFILES hierarchy is shown in Figure 2.

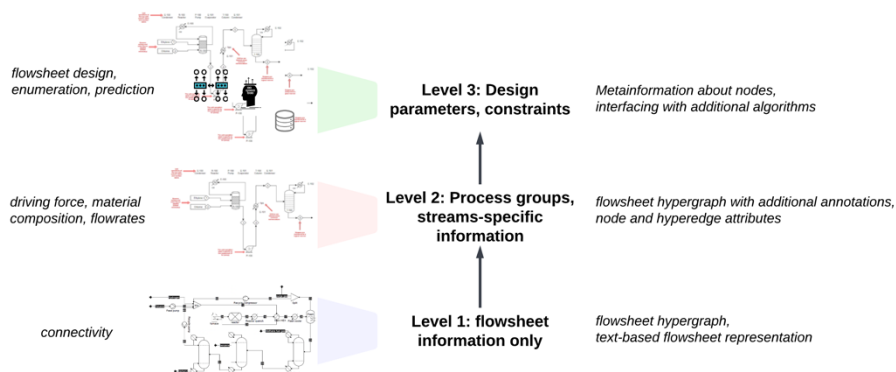


Figure 2: eSFILES framework with three levels of hierarchy

**Implementation of hierarchical framework:** The concept of hypergraphs is introduced to incorporate multi-level hierarchy of information with eSFILES. A hypergraph is a generalization of a graph where each edge is not limited to connecting only two vertices but could connect any number of vertices. Mathematically, a hypergraph consists of a set of hyperedges and vertices where each hyperedge consists of a non-empty subset of vertices. In addition, each vertex has a hyperedge-specific annotation indicating a predefined contextual information. The flexibility of hyperedges to connect more than one node and ability to have hyperedge-specific node annotations offers several advantages for flowsheet representation. Further details on the annotated hypergraph framework and its various applications in reaction engineering are presented in (Mann & Venkatasubramanian, 2023).

For instance, for the HDA process, the equivalent flowsheet hypergraph representation is shown in Figure 1(d) where each process group is represented as a colored hyperedge, process streams are represented as vertices, and the vertices have hyperedge-specific 'in'/'out' annotations indicating the incoming/outgoing nature of the process streams for

a corresponding process groups. Moreover, recycle streams are indicated with an additional annotation 'rec' wherever applicable for the process streams. Figure 1(b) shows each process group as a hyperedge which essentially is a set of named vertices where the names (or annotations) indicate the streams' nature.

The three levels of information in the eSFILES framework is shown for the 'toluene recovery' process group in Figure 3. In Figure 3(a), only the hyperedge representing the process group along with participating streams as vertices and role-specifying annotations are shown. In Figure 3(b), additional information for the entire process group (or hyperedge) such as driving force and stream (or vertex)-specific information such as mass fraction and composition is shown. Finally, information on design aspects and necessary information required for additional analyses such as mixture analysis, reaction analysis, and pure component properties are presented in Figure 3(c). Though the three levels in eSFILES are only shown for a process group brevity, the same could be done for an entire process using the same approach.

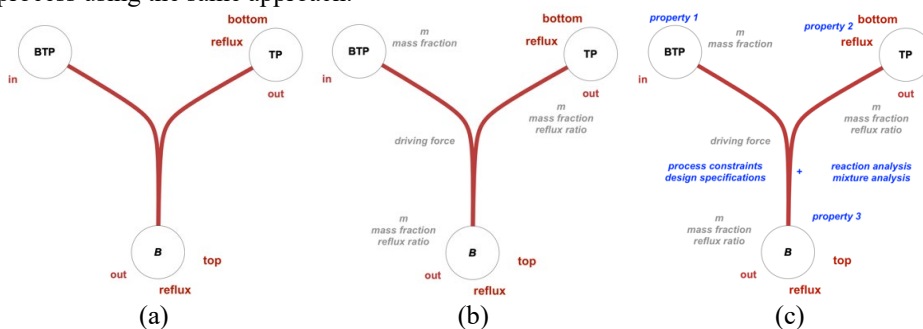


Figure 3: (a) Distillation column hyperedge with additional annotations. The corresponding text-based representation for this process groups is  $BTP(in) \rightarrow dl \rightarrow B(out, top, reflux) TP(out, bottom, reflux)$ , (b) the driving force, material compositions, reflux rate, and conversion indicated as hyperedge attributes (for process groups) and hyperedge-specific node attributes (for streams) (c) the design parameters, properties, and constraints are represented as additional meta information.

**Application of the eSFILES in sustainable process design:** (Tula et al., 2017) proposed a 3-stages sustainable process design method consisting of 12 hierarchical steps. Stage-1 consists of steps 1-4 where the process synthesis problem is solved. Stage-2 consists of steps 5-9 where the base case process design problem is solved. Stage-3 consists of steps 10-12 where the base case design is further improved to find innovative and more sustainable alternatives. The original multi-stage method of (Tula et al., 2017) now retrieves the necessary data from the stored hypergraph. Also, different alternatives for stage-3 can easily be generated through different alternatives of the hypergraph, allowing the user to generate and analyze multiple alternatives as well as store the information for future use without having to repeat any of the previously solved steps.

#### 4. Conclusions

The concept of hypergraph has been used to extend the original SFILES concept to multiple levels to store as well as retrieve information related to process synthesis, design, and innovation. In this way, an existing interactive multi-stage method has been made more intelligent and efficient through the adoption of an available AI method for

sustainable process design. Therefore, the eSFILES-based framework for process synthesis and design incorporates a combination of artificial intelligence-based methods and well-known chemical engineering knowledge incorporated through an intelligent system facilitating fast, correct, and consistent decision-making related to process synthesis and design.

Current and future work involves the development of an extended manuscript providing more details of the eSFILES, its adoption in the multi-stage sustainable process design method, and its application in several case studies. In particular, we plan to provide rigorous examples to demonstrate the usability of eSFILES for new process flowsheets and also use them to address real challenges such as provide flowsheet alternatives, perform process intensification, and flowsheet property prediction based on the text-based flowsheet representation using a natural language framework similar to molecular property prediction (Mann et al., 2022) and chemical product design (Mann et al., 2023).

## References

- Bommareddy, S., Eden, M. R., & Gani, R. (2011). *Computer Aided Flowsheet Design using Group Contribution Methods* (pp. 321–325). <https://doi.org/10.1016/B978-0-444-53711-9.50065-1>
- D'Anterrosches, L. (2005). *Process flowsheet generation & design through a group contribution approach*. [CAPEC], Department of Chemical Engineering, Technical University of Denmark.
- Hirtreiter, E., Balhorn, L. S., & Schweidtmann, A. M. (2022). *Towards automatic generation of Piping and Instrumentation Diagrams (P&IDs) with Artificial Intelligence*. <http://arxiv.org/abs/2211.05583>
- Mann, V., Brito, K., Gani, R., & Venkatasubramanian, V. (2022). Hybrid, Interpretable Machine Learning for Thermodynamic Property Estimation using Grammar2vec for Molecular Representation. *Fluid Phase Equilibria*, 561, 113531. <https://doi.org/10.1016/j.fluid.2022.113531>
- Mann, V., Gani, R., & Venkatasubramanian, V. (2023). Group contribution-based property modeling for chemical product design: A perspective in the AI era. *Fluid Phase Equilibria*, 568, 113734. <https://doi.org/10.1016/J.FLUID.2023.113734>
- Mann, V., & Venkatasubramanian, V. (2023). AI-driven hypergraph network of organic chemistry: network statistics and applications in reaction classification. *Reaction Chemistry & Engineering*. <https://doi.org/10.1039/D2RE00309K>
- Tula, A. K., Babi, D. K., Bottlaender, J., Eden, M. R., & Gani, R. (2017). A computer-aided software-tool for sustainable process synthesis-intensification. *Computers & Chemical Engineering*, 105, 74–95. <https://doi.org/10.1016/j.compchemeng.2017.01.001>
- Tula, A. K., Eden, M. R., & Gani, R. (2015). Process synthesis, design and analysis using a process-group contribution method. *Computers & Chemical Engineering*, 81, 245–259. <https://doi.org/10.1016/j.compchemeng.2015.04.019>
- Venkatasubramanian, V., & Mann, V. (2022). Artificial intelligence in reaction prediction and chemical synthesis. *Current Opinion in Chemical Engineering*, 36, 100749. <https://doi.org/10.1016/j.coche.2021.100749>
- Vogel, G., Balhorn, L. S., & Schweidtmann, A. M. (2022). *Learning from flowsheets: A generative transformer model for autocompletion of flowsheets*. <http://arxiv.org/abs/2208.00859>
- Zhang, T., Sahinidis, N. v., & Siirola, J. J. (2019). Pattern recognition in chemical process flowsheets. *AIChE Journal*, 65(2), 592–603. <https://doi.org/10.1002/aic.16443>

# Physics-informed Neural Network based Modeling of an Industrial Wastewater Treatment Unit

Tuse Asrav<sup>a,b</sup>, Ece Serenat Koksal<sup>a,b</sup>, Elif Ecem Esenboga<sup>c</sup>, Ahmet Cosgun<sup>c</sup>, Gizem Kusoglu<sup>c</sup>, Duygu Aydin<sup>c</sup>, Erdal Aydin<sup>a,b,\*</sup>

<sup>a</sup>*Department of Chemical and Biological Engineering, Koç University, Istanbul 34450, Turkey*

<sup>b</sup>*Koç University TUPRAS Energy Center (KUTEM), Koç University, Istanbul, 34450, Turkey*

<sup>c</sup>*Turkish Petroleum Refineries Corporation, Körfez, Kocaeli 41790, Turkey*  
*eydin@ku.edu.tr*

## Abstract

Wastewater treatment units consist of biological treatment with activated sludge and are subject to many disturbances such as influent flowrate, pollutant load and weather conditions bringing about many challenges for the modeling of such plants. Data-driven models may respond to these challenges at the cost of issues such as overfitting or poor fitting due to the lack of high-quality data. To benefit from the available physics-based knowledge and to eliminate the drawbacks of suboptimal and poor training, physics informed neural networks might be quite promising. In this work, artificial, recurrent and physics-informed neural network models are utilized for the wastewater plant in Tüpraş İzmit Refinery. For recurrent models with selected features based on correlation technique, test mean squared error is up to 82% smaller compared to the standard artificial neural network models. Physics-informed trained neural network models with selected features improved the test performance by decreasing mean squared error up to 87% with acceptable decreases in training performance which addresses its strength compared to fully data-driven models.

**Keywords:** wastewater treatment; physics-informed neural networks; recurrent neural networks; process optimization; wastewater control

## 1. Introduction

Wastewater treatment is the removal of contaminants from the wastewater before being discharged into the sea to minimize the damage to the environment. Pre-treatment of



wastewater is a physical treatment that involves grid sieves, sand traps, equalization basins, and corrugated plate interceptor. Coagulants, flocculants, sodium carbonate, and hydrogen peroxide addition to the wastewater corresponds to the chemical treatment where disinfection, sanitization and purification is achieved. Then, the stream is further treated by dissolved air flotation (DAF) which reduces the amount of total suspended solids and oil and grease. Biological treatment involves the removal of biodegradable organic matter by making use of the bacteria which is called the activated sludge process occurring in aeration tanks. Here, bacteria use oxygen to break down the biodegradable organic matter. Activated sludge process is followed by a sedimentation tank also called clarifier. Suspended particles settle out of wastewater as it flows slowly through the tank therefore further purification is achieved. Part of the exiting flow from the clarifier with high mixed liquor suspended solids (MLSS) content is recycled to the aeration tanks (Benyahia et al., 2006.)

Raising concerns around the world about environment have started to shift the regulations for several process systems areas towards tighter windows. Wastewater treatment is among these areas where optimal operation will become inevitable. Since wastewater treatment is a complex, non-linear system including physical, chemical, and biological processes as expressed above and subject to many disturbances, mechanistic modeling of a wastewater treatment unit is challenging (Guo et al., 2015). Alternatively, data-driven models, in particular artificial neural networks (ANN), are widely used at the cost of outcomes of black-box models such as overfitting, underfitting and lack of high-quality data. Similarly, recurrent neural networks (RNN) are able to handle time-dependent responses. They are suited better than the ANNs for dynamic systems. On the other hand, RNNs are also subject to issues such as vanishing gradient problem. Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) introduce more complex recurrent units with gates to retain long-term dependencies by controlling which information is passed through. The presence of dynamic data in wastewater treatment units makes it appropriate to construct RNN, LSTM and GRU models (Quaghebeur et al., 2022).

In this work, ANN, simple RNN, LSTM, GRU and physics informed ANN models are constructed to predict dissolved oxygen concentration (mg/L) at the aeration tanks, sludge volume index (mL/g) and waste activated sludge (m<sup>3</sup>/day) in Tüpraş İzmit Refinery. Prediction performance is further increased by removing some of the features based on a correlation technique. Section 2 involves the methodology. In Section 3, results are shown, and Section 4 concludes the results.

## 2. Methodology

### 2.1. Artificial and Recurrent Neural Networks

A feed forward, fully connected ANN can be expressed as follows:

$$\hat{y} = f_1(w_1 f_2(w_2 u + b_2) + b_1) \quad (1)$$

where  $f_1$  and  $f_2$  are activation functions,  $w_1$  and  $w_2$  are weights,  $b_1$  and  $b_2$  are bias terms at output and hidden layers, respectively. Moreover,  $u$  and  $\hat{y}$  represents input and output terms, respectively. RNNs extend the feed forward ANNs to handle the sequential data by repeatedly transferring information to the cell states throughout the time series. The prediction at the current time step depends on the input at the current time step and the memory of the cell state. LSTM and GRU architectures can track longer dependencies better than the traditional RNNs due to their complex gated mechanisms controlling the flow of information.

## 2.2. Physics-Informed Neural Networks

The training of the neural networks can be expressed as an optimization problem where the objective function is the loss function given in Eq. 2. Typically, the purpose of training the neural networks is to find the appropriate values of weights and biases that minimize the loss function which can be achieved using the backward propagation algorithm. In order to improve the performance via physics-informed knowledge, neural networks can be trained using a bi-objective loss function which includes user-defined physics term. A bi-objective loss function can be proposed as follows:

$$L = \frac{1}{N} \sum_i^N (y_i - \hat{y}_i)^2 + W_p P \quad (2)$$

where  $W_p$  is the scalar weight and  $P$  is the physics term.

## 3. Results

### 3.1. Comparison of Neural Networks for Prediction of Dissolved Oxygen Concentration at Aeration Tanks

Total suspended solids (TSS), chemical oxygen demand (COD), pH, oil and grease, ammonium ( $\text{NH}_4\text{-N}$ ), phenol, sulfide, TSS at return activated sludge which are measured after dissolved air flotation (DAF) are taken as inputs, and dissolved oxygen concentration in two parallel aeration tanks (Aeration tank A and B) is taken as output. The data is normalized between -1 and 1. The first 90% of the data is used as training data and the remaining is used as test data. An ANN, a simple RNN, an LSTM, and a GRU networks are developed using 3 hidden layers with 50 neurons with the rectified linear unit (ReLU) activation function. The mean squared error loss function with the Adam optimization algorithm in Python is used to train the models. Results based on mean squared error (MSE) and mean absolute percentage error (MAPE) given in Table 1 show that RNNs predicted the test data better than the feed-forward artificial neural network. Among the RNNs, LSTM can be selected as the network delivering the best prediction over test data since simple RNN and GRU converges to more oscillatory predictions as

shown in Figure 1, and it may be more useful to predict the direction. In addition, a correlation technique based on F-value for regression tasks in Sklearn library is used to select features for dissolved oxygen prediction in the aeration tanks based on the LSTM model. (Pedregosa et al., 2011). By this technique, COD, pH,  $\text{NH}_4\text{-N}$ , phenol, sulfide, and TSS at return activated sludge are chosen as inputs and TSS at DAF exit and oil are removed from features since the correlation between these features is low. The results are shown in Table 1 and Figure 1.

**Table 1.** Performance of the Data-driven Neural Network Models for Prediction of Dissolved Oxygen Concentration at Aeration Tanks.

	ANN	Simple RNN	GRU	LSTM	LSTM after feature selection
Train MSE	0.0194	0.0024	0.0028	0.0146	0.0089
Test MSE	0.1266	0.0374	0.0416	0.0327	0.0229
Train MAPE	1.7978	0.3580	0.3616	1.4553	1.224
Test MAPE	0.7006	0.3502	0.4316	0.4150	0.3293

Model

Training

Test

ANN

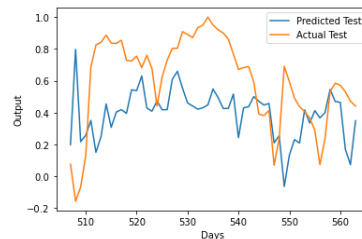
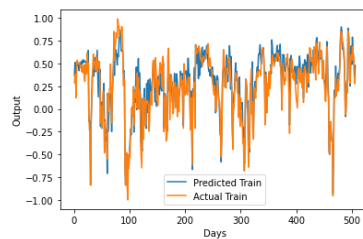


Figure 1.a.

Figure 1.b.

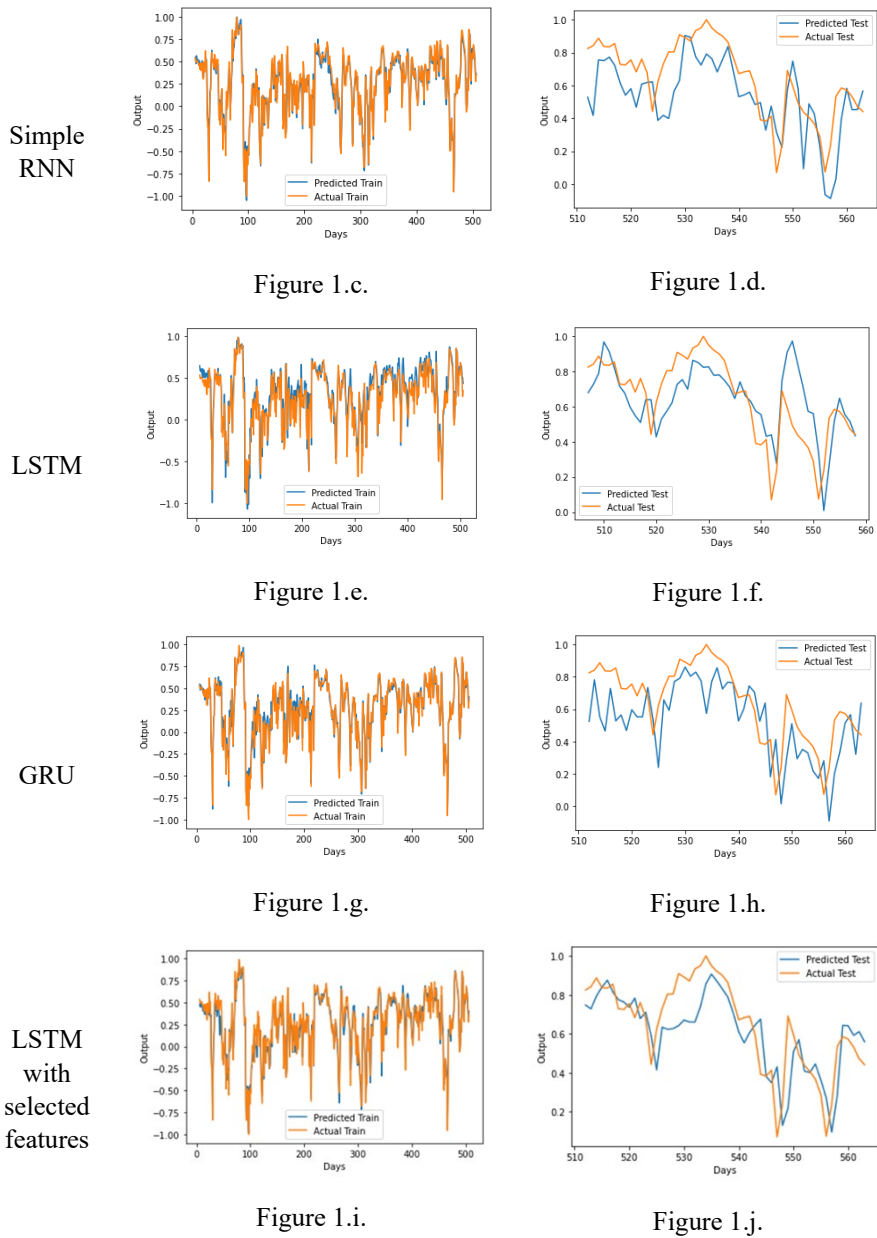


Figure 1. Training and test trends for the prediction of dissolved oxygen concentration at the aeration tanks.

### 3.2. The effect of Physics-Informed Loss Function

Feed-forward artificial neural networks trained with physics-informed (PI+) and physics-uninformed (PI-) loss functions are developed and the results are compared to observe the effect of bi-objective training.

#### 3.2.1. Sludge Volume Index (SVI) Prediction

Dissolved oxygen concentration, pH, temperature, TSS, results of imhoff cone experiments where settleable suspended solid is measured (mL/L), waste activated sludge, and food to microorganism ratio (F/M) in the aeration tank A are taken as inputs, and sludge volume index (SVI) in aeration tank A is taken as output. First 60% of the data is selected as training. A single hidden layer with 5 neurons with the ReLU activation function is used. Physics-term related to the SVI is constructed by a rule-of-thumb used in the operation in the refinery. In addition, based on the results of the feature selection, a physics-informed ANN is developed taking dissolved oxygen concentration, pH, TSS, and settleable suspended solids in the aeration tank A as inputs since the correlation is only observed between these parameters and SVI. Results are shown in Table 2.

**Table 2.** Performance of the ANN Models for SVI Prediction.

	PI-	PI+	PI+ with feature selection
Train MSE	0.0053	0.0067	0.0072
Test MSE	0.0650	0.0181	0.0088
Train MAPE	0.1187	0.1661	0.0969
Test MAPE	0.1520	0.1144	0.0623

Even though the bi-objective physics-informed model delivers slightly higher training errors, the test performance of the model is improved when the physics term is introduced to the loss function. Moreover, the performance of the physics-informed trained ANN is further increased with selected features. Overall, the MSE is reduced 86.5%.

*3.2.2. Prediction of Waste Activated Sludge*

Oily water sewer feed, DAF exit properties (pH, TSS, COD, oil and grease, NH<sub>4</sub>-N, phenol, sulfide), carbon-nitrogen ratio (C/N), COD load, aeration tanks' conditions (pH, dissolved oxygen concentration, temperature, TSS, F/M, settleable suspended solids, SVI), and TSS at return activated sludge are taken as inputs and waste activated sludge from both in aeration tank A and B are taken as outputs. First 80% of the data is selected as training data. There are 4 hidden layers with 50 neurons with the hyperbolic tangent (tanh) activation function. For the physics-term, as in the case for SVI, a rule of thumb related to the waste activated sludge is used. Furthermore, DAF exit properties are excluded from the inputs as a result of feature selection for physics-informed ANN model since they do not have a significant impact to predict the outputs. Training and test performances of the ANN models are evaluated and reported in Table 3. MSE values reduced 40.2% and 47.7% for tank A and B, respectively.

**Table 3.** Performance of the ANN Models for Prediction of Waste Activated Sludge.

	Waste Activated Sludge from tank A			Waste Activated Sludge from tank B		
	PI-	PI+	PI+ with selected features	PI-	PI+	PI+ with selected features
Train MSE	0.0005	0.0020	0.0004	0.0025	0.0022	0.0036
Test MSE	0.0169	0.0132	0.0101	0.0044	0.0029	0.0023
Train MAPE	0.0400	0.0882	0.0209	0.0338	0.0361	0.0171
Test MAPE	0.0307	0.0444	0.0308	0.0399	0.0259	0.0234

**4. Conclusion**

In this work, ANN, simple RNN, LSTM and GRU models are constructed to predict dissolved oxygen concentration at the aeration tanks in Tüpraş İzmit Refinery wastewater treatment unit. RNN-based models improved the prediction capability due to the dynamic nature of the data. LSTM is chosen as the model to be constructed since it predicts the direction better than the other models. Then, the prediction capability is further increased by selecting features based on correlation technique. As a result, the test MSE of LSTM with selected features is approximately 82 % better compared to standard ANN with a closer trend to the actual values. Accordingly, physics-informed knowledge is integrated into the training instances of the neural networks to improve test performance and reduce

the possibility of suboptimal training for the prediction of SVI and waste activated sludge. Test MSE of physics-informed feature selected ANN models for the prediction of SVI, waste activated sludge in aeration tanks A and B are 86.5%, 40.2% and 47.7% smaller, respectively compared to the standard ANN. Finally, this study shows that RNNs may result in more accurate models especially when the process nature is dynamic. In addition, physics-informed knowledge can also be included to the standard ANN models to increase the prediction accuracy.

## References

- Benyahia, F., Abdulkarim, M., & Embaby, A. (n.d.). *Refinery wastewater treatment: a true technological challenge*.
- Guo, H., Jeong, K., Lim, J., Jo, J., Kim, Y. M., Park, J. pyo, Kim, J. H., & Cho, K. H. (2015). Prediction of effluent concentration in a wastewater treatment plant using machine learning models. *Journal of Environmental Sciences (China)*, 32, 90–101. <https://doi.org/10.1016/j.jes.2015.01.007>
- Pedregosa, F., Varoquaux, Ga"el, Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... others. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(Oct), 2825–2830.
- Quaghebeur, W., Torfs, E., de Baets, B., & Nopens, I. (2022). Hybrid differential equations: Integrating mechanistic and data-driven techniques for modelling of water systems. *Water Research*, 213. <https://doi.org/10.1016/j.watres.2022.118166>

# Approaches to Reduce Optimization Time for Stochastic Optimization of Complex Chemical Processes

Fanyi Duanmu, Dian Ning Chia, Eva Sorensen\*

*Department of Chemical Engineering, University College London (UCL),  
Torrington Place, London WC1E 7JE, United Kingdom*

*\*e.sorensen@ucl.ac.uk*

## Abstract

Optimization is a crucial step to obtain an energy efficient and economically viable design in the chemical industry. The optimization of a chemical process is, however, often challenging due to the inherent complexity of chemical process designs. In this work, different combinations of speed up approaches – parallel computing (ParC), timeout function (TO), escape repeated simulations (ERS), dynamic bound (DB), and machine learning assisted optimization (ML) – are considered and their performances are compared. ParC is always activated as it has previously been found very effective to reduce optimization time. The results show that including DB as part of the speed up approach can ensure successful optimization for complex chemical processes. The TO and ML approaches can help reduce the optimization time significantly, but the ERS approach has no clear effect on the optimization performance. The approach with all speed up approaches activated together yields the best design in the shortest time.

**Keywords:** Stochastic optimization, genetic algorithm, parallel computing, machine learning, distillation

## 1. Introduction

Optimization is an important step in designing a chemical process, as an optimal design should not only have low capital and operating costs (CAPEX and OPEX) but also be energy efficient and sustainable. However, due to the complex nature of chemical process designs, their optimization is challenging, partly because of the difficulty in converging into a truly optimal design, but also due to the long optimization time. Different stochastic optimization methods (e.g., Genetic Algorithm (GA) and Particle Swarm Optimization (PSO)) and their corresponding settings (e.g., for GA, various types of parent selection, crossover, and mutation methods) are available that suit different designs and can help to improve convergence. However, the long optimization time is difficult to address as it is often caused by the high chance of failure of the initialization and re-initialization of the design (i.e., infeasible designs), which makes the solution time long. The optimization time can be significantly reduced if these infeasible designs could be avoided with the use of, for instance, dynamic bounds (Chia et al., 2021), removed by machine learning assisted optimization (Ibrahim et al., 2017), or timed-out using a timeout function. Moreover, escaping (i.e., skipping) repeated simulations is also a direct and efficient way to reduce optimization time. Finally, parallel computing (e.g., master-slave structure) is also widely used to speed up optimization.



## 2. Methodology

It has been reported that stochastic optimization, e.g., Genetic Algorithm (GA), is more robust but also more time-demanding compared with gradient-based deterministic optimization (Chia et al., 2021; Cavazzuti, 2013), and the stochastic optimization time of a chemical process may take up to a few hours. Therefore, in order to tackle long optimization times, different speed up approaches will be introduced in this work.

### 2.1. Parallel Computing (ParC)

The parallel computing approach (ParC) is one of the most successful and commonly applied methods. There are a few structures available to perform ParC, such as the master-slave model, island model, and cellular model, which are classified by different ways of communication within the model. Details of each type of parallel model can be found in Gong et al. (2015). In this work, the master-slave model is used. For this model, in each iteration, the individuals (e.g., chromosomes in GA) are assigned to available slaves (i.e., cores/processors) for evaluating the fitness and constraints. Once the results are obtained, the slaves return the results to the master to handle constraints and to generate the individuals for the next iteration (e.g., parent selection, crossover, and mutation in GA). Another important concept of ParC is the ways of assigning the tasks by either static or dynamic assignment (Mathew et al., 2014). In the case of static assignment, the tasks in a parallel job are assigned to a core/processor beforehand, while in the case of dynamic assignment, a task is assigned to whichever core/processor that is "free" at the moment. For the optimization of chemical processes, the dynamic assignment is preferred as the simulation time of each individual may vary considerably.

### 2.2. Timeout Function (TO)

From the authors' experience, some infeasible simulations (i.e., designs that fail to converge) may take a very long time before they fail to converge due to the multiple re-initialization performed to "confirm" that there is no numerical solution to that simulation. This is illustrated in Figure 1, which is one of the optimization results taken from the case study (10,000 simulations with 4,737 feasible simulations). Almost all the feasible simulations converged within 9 seconds but there is a proportion of infeasible simulations which take more than 9 s and may take up to about 38 s before they can be confirmed as "infeasible". A timeout function can therefore be used to rule out the simulations that take an unreasonably long time but have a high chance of being infeasible.

### 2.3. Escape Repeated Simulation (ERS)

Perhaps one of the most direct ways to reduce the optimization time is to escape any repeated simulations (i.e., skip the simulations that have the exact same input values). Repeated simulations can be present due to, for example, the elitism strategy adopted in GA. ERS can be achieved by saving and comparing the inputs to the simulation. In this work, a "short-term memory" strategy is adapted, where the inputs of the current iteration is saved and retained to the next iteration where the saved inputs and the "new" inputs generated are compared to identify any repeated simulations. The "old" inputs from the previous iteration are then discarded and replaced with the "new" inputs generated in this iteration. This strategy can reduce the need for memory allocation and avoid the slow down effect that will otherwise occur if a "long-term memory" strategy is used (i.e., saving and retaining all inputs that have occurred in all iterations so far).

### 2.4. Dynamic Bound (DB)

Dynamic bound (DB) is a function that changes (internally in the code) the upper and/or lower bound of a variable depending on the value of another variable. The benefit of DB may not be immediate in optimizing a purely mathematical problem, but the benefit is

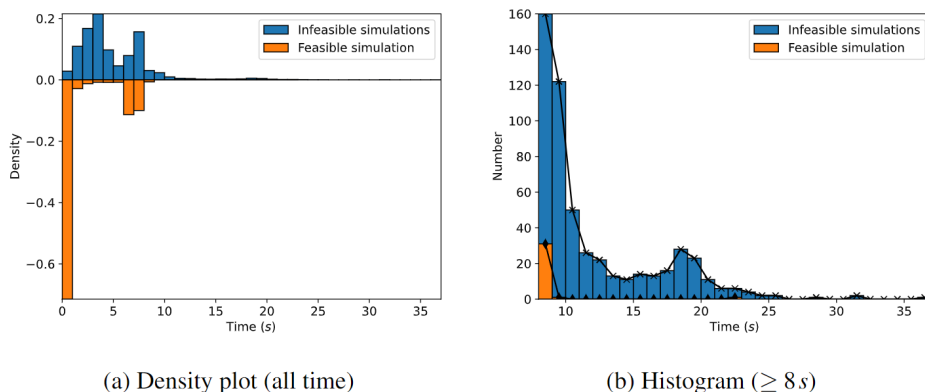


Figure 1. Distribution of infeasible and feasible simulations depicted as:  
 (a) density plot for the whole range of simulations times and  
 (b) histogram as a function of simulation time from 8 s onwards.

tremendous when it comes to optimizing a chemical process. For example, it is obvious that the feed stage ( $N_f$ ) of a distillation column should not be larger than the total number of stages ( $N_t$ ). However, if the bounds of  $N_t$  have a significant overlap with the bounds of  $N_f$  (e.g.,  $N_t$ : 10-40 and  $N_f$ : 5-35) due to lack of information about the process, there will be a very high chance that a design generated from the stochastic optimization has a higher value of  $N_f$  than  $N_t$ , which clearly leads to an infeasible simulation. The chance becomes even more significant if more feed locations and/or more distillation columns are involved in the optimization task. Therefore, if the rule (e.g.,  $N_f$  should not be larger than  $N_t$ ) is violated after generating new individuals (e.g., crossover in GA), then the DB function can be used to re-assign the feed location as a random integer within its lower bound and the total number of stages in the current design, thus reducing infeasibility.

### 2.5. Machine Learning Assisted Optimization (ML)

Due to the nature of the simulations of chemical processes, there are at least two sets of labels of a simulation, which are feasible/infeasible and on-/off-spec. The feasible/infeasible label describes if a simulation converges successfully or not, while the on-/off-spec label describes if a simulation stays within the constraint(s) or violates the constraint(s). Machine Learning (ML; e.g., Support Vector Machine (SVM)) is well suited for classification, which can be used to predict the status of a design before evaluating it. From the authors' experience, using ML to predict the on-/off-spec is more challenging as the training data requires a good balance between the number of on-spec and off-spec simulations, and for a complex chemical process, the proportion of on-spec simulation may be relatively low and even worse in the early iterations. Therefore, in this work, ML is applied to predict only the feasible/infeasible labels. ML will only collect the inputs and results of the simulations until the iteration reaches a user-defined iteration. After that, ML will use the information collected to train the model. Then, for the subsequent iterations, no training is required and only prediction with the trained model is carried out. For designs that are predicted to be infeasible, no evaluations will be carried out, thus the time required to process the "infeasible" designs can be saved.

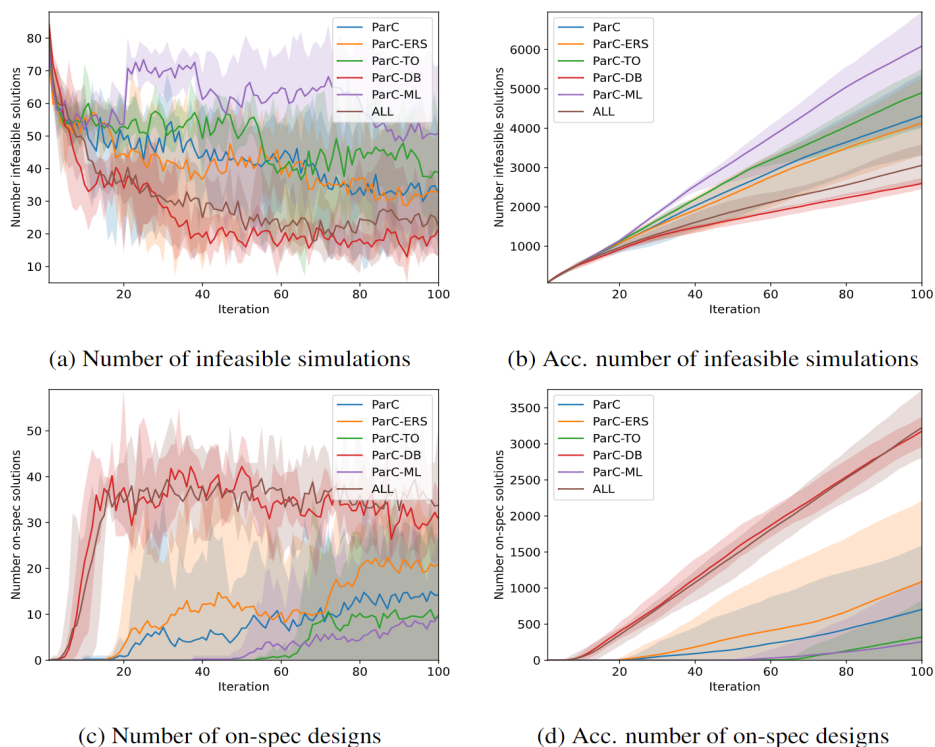


Figure 2. Performances of various combinations of speed up approaches on (a,b) reducing the number of infeasible simulations and (c,d) increasing the number of on-spec designs.

### 3. Case Studies

The case studies in this work were performed on a dual Intel Xeon Gold 6226R CPU with 16 Cores 2.90GHz (total 64 logic processors) for a highly integrated and complex hybrid dividing wall column (H-DWC) design to separate a mixture of ethyl acetate and ethanol, and the H-DWC details can be found in Chia et al. (2022). The simulation of a H-DWC involves two distillation columns (prefractionator and main column in the Petlyuk structure) and a membrane network, and the optimization variables are the design and operating variables of the distillation columns and membrane network. To compare the performance of different approaches, six optimization tasks are carried out: optimization with parallel computing (ParC); ParC with timeout function (ParC-TO); ParC with escaped repeated simulation (ParC-ERS); ParC with dynamic bound (ParC-DB); ParC with machine learning (ParC-ML); and all approaches activated together (ALL). ParC was always activated with 40 processors with the dynamic task assignment structure, as it had been found in previous work that ParC is very efficient in reducing optimization time (Chia et al., 2021). For TO, the timeout value was set as 10 s, decided through statistics shown in Figure 1. For ERS, the "short-term memory" strategy was adopted. For DB, a few rules related to column stages were defined, such as the feed locations should not be larger than its total number of stages (stage number is counted starting from the condenser), and the number of stages on both sides of the wall in the dividing wall column should be the same (although can of course be different). For ML, a support vector machine (SVM) library developed by Chang and Lin (2011) was used with a seventh-

degree polynomial,  $C = 5$ , and  $\epsilon = 0.1$ . The settings for SVM were decided through trial-and-error (not shown). The first 20 iterations were used for training the SVM. In this work, the H-DWC model was developed in gPROMS Process (Process Systems Enterprise, 2021). The Genetic Algorithm (GA) was used as external optimization (black-box optimization) and was coded in C# by the authors. GA was constructed with 100 iterations (generations) and 100 individuals (chromosomes), tournament with four players for the parent selection, flat crossover, and discrete random mutation.

Each optimization task was repeated five times and the results are plotted in Figure 2, where the solid lines are the average values. The bands bounded by the minimum and maximum values are plotted together. Figures 2a and 2b show the number of infeasible simulations and the number of accumulated infeasible simulations in each iteration, respectively. It can be seen that ParC-ML has the most infeasible simulations, and this is because the training of SVM is biased due to the fact that most of the simulations in the first 20 iterations are infeasible. ParC, ParC-ERS, and ParC-TO show similar trends, especially ParC and ParC-ERS. The finding is valid as for ERS, this function does not affect the operation of the GA, and for TO, a slightly larger deviation from ParC is also expected because there is still a chance that feasible simulations may be timed out and be considered as infeasible simulations, which may slightly reduce the performance of GA (considering the number of feasible and on-spec simulations). Considering ParC-DB and ALL, these have similar trends and the lowest number of infeasible simulations in each iteration. Compared with ParC, the total number of infeasible simulations (shown in Table 1) of ParC-DB and ALL drops from 4313 to 2596 (40 % reduction) and 3053 (29 % reduction), respectively, which indicates that DB is the key function to reduce the number of infeasible simulations. Considering the number of on-spec simulations shown in Figures 2c and 2d, it is not surprising that ParC-DB and ALL have higher number of on-spec simulations as many unreasonable (thus off-spec) designs are avoided by DB, so a higher chance of on-spec designs would be expected. Moreover, the use of DB balances the number of feasible and infeasible simulations in the first 20 iterations, thus the bias of ML in ALL is reduced compared to in ParC-ML.

To reflect the speed of the optimizations, three indicators are tabulated in Table 1, which are the convergence percentage (the number of successfully converged optimizations, i.e., the final optimal on-spec design divided by the total number of optimizations performed); quality of the fitness (e.g., the total annualized cost in the case studies); and optimization time. Starting with the convergence percentage, Table 1 shows that only ParC-DB and ALL could achieve 100 % convergence, which indicates that to obtain the same number of valid optimizations (i.e., converged optimizations), other approaches require more repeated optimizations leading to a longer total optimization time. Both ParC-DB and ALL yield designs with the lowest TACs ( $2.54 \text{ M } \$ \text{ y}^{-1}$ ), indicating better optimization efficiency. Note that all optimizations stop after 100 iterations, which is the stopping criterion. Moving to the most direct indicator, optimization time, starting with ParC, the averaged optimization time is 1886 s. The time for each individual is recorded and the summation of all individual times is considered as the total time taken without parallel. The optimization time without parallel computing is 30,963 s, which indicates the optimization with parallel computing only is about 15 times faster. The table shows that ERS does not save CPU time as the elitism percentage is low (10 %), and repeated designs rarely appeared due to the applied Flat crossover method. Moreover, ParC-DB has minor improvement due to the lower infeasible rate. For the other three approaches, ParC-TO,

Table 1. Key performance indicators of the different speed up approaches. The values reported are the mean values obtained from five optimizations unless stated otherwise.

	ParC	ParC-ERS	ParC-TO	ParC-DB	ParC-ML	ALL
No. infeasible solutions	4314	4134	4902	2596	6086	3053
No. on-spec solutions	704	1090	321	3075	259	3225
Optimization time (s)	1885	1939	1397	1756	1452	1033
Conv. percentage (%)	60	80	60	100	40	100
TAC* ( $M \$ y^{-1}$ )	2.65	2.63	2.80	2.54	2.67	2.54

\* Considering only converged optimizations

ParC-ML, and ALL, they all show good time savings with about 26 %, 23 %, and 45 % reduction, respectively.

#### 4. Conclusion

In conclusion, this work has compared five different approaches to speed up stochastic optimization. By considering the overall performance, the dynamic bound function is clearly the most effective function to ensure a successful optimization for a complex chemical process. Using a timeout function or machine learning has no clear improvement on the quality of the optimization results, but these functions show remarkable reduction in the optimization time. The escaping repeated simulation function shows no clear effect on the optimization speed. By applying all approaches together, the optimization achieves the best design, the same as for parallel computing with dynamic bound, but with the least time. In future work, a more comprehensive comparison will be performed with other stochastic optimization methods such as Particle Swarm Optimization and Simulated Annealing.

#### References

- M. Cavazzuti, 2013. Optimization Methods. Vol. 53. Springer Berlin Heidelberg, Berlin, Heidelberg.
- C.-C. Chang, C.-J. Lin, 2011. LIBSVM: A Library for Support Vector Machines. ACM Transactions on Intelligent Systems and Technology 2 (3), 1–27.
- D. N. Chia, F. Duanmu, E. Sorensen, 2021. Optimal Design of Distillation Columns Using a Combined Optimisation Approach. In: M. Turkay, R. Gani (Eds.), 31st European Symposium on Computer Aided Process Engineering. Elsevier B.V., pp. 153–158.
- D. N. Chia, F. Duanmu, E. Sorensen, 2022. Optimal Design of Hybrid Dividing Wall Columns for Azeotropic Separations. In: The 12th International Conference Distillation & Absorption, Toulouse, September 2022.
- Y. J. Gong, W. N. Chen, Z. H. Zhan, J. Zhang, Y. Li, Q. Zhang, J. J. Li, 2015. Distributed evolutionary algorithms and their models: A survey of the state-of-the-art. Applied Soft Computing Journal 34 (2013), 286–300.
- D. Ibrahim, M. Jobson, J. Li, G. Guillén-Gosálbez, 2017. Surrogate Models combined with a Support Vector Machine for the Optimized Design of a Crude Oil Distillation Unit using Genetic Algorithms. In: Computer Aided Chemical Engineering. Vol. 40. pp. 481–486.
- T. Mathew, K. C. Sekaran, J. Jose, 2014. Study and analysis of various task scheduling algorithms in the cloud computing environment. In: 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI). IEEE, pp. 658–664.
- Process Systems Enterprise, 2021. gPROMS Process version 2.2.

# Design of a CO<sub>2</sub> capture plant: A sustainable approach using deep eutectic solvents

Adrián Martínez-Lomovskoi,<sup>a</sup> Ana-Gabriela Romero-García,<sup>a,b</sup> Eduardo Sánchez-Ramírez,<sup>a</sup> Juan-Gabriel Segovia-Hernández,<sup>a</sup> Cataldo De Blasio<sup>b</sup>

<sup>a</sup>*Departamento de Ingeniería Química, Universidad de Guanajuato, Noria Alta s/n, Guanajuato, Gto., 36050, México.*

<sup>b</sup>*Åbo Akademi University, Faculty of Science and Engineering, Laboratory of Energy Technology, Rantakatu 2, Vaasa, 65100, Finland.*

## Abstract

In this work, a new design for post-combustion capture (PCC) of CO<sub>2</sub> is introduced, using green deep eutectic solvent Choline chloride/Urea (1:2). A multi-objective optimization using meta-heuristic algorithm differential evolution and tabu list (MODE-TL) considering environmental and economic **objectives** is performed to obtain a sustainable process design. The proposed process was analyzed for treating flue gases from combustion of the most common fuels used in power plants. The process performs with a lower cost and environmental impact when treating flue gases from the combustion of coal. Through a simulation in Aspen Plus, it is shown that using ChCl/Urea (1:2) it's possible to obtain a process capable of recovering more than 95% of the CO<sub>2</sub> contained in the combustion gases with a purity of 95% molar. A sustainable capture processes according to the objectives of green chemical engineering and circular economy is developed.

**Keywords:** CO<sub>2</sub> capture, MEA, deep eutectic solvent

## 1. Introduction

The United Nation's 2030 Agenda for Sustainable Development has called for the taking urgent action to combat climate change and its impacts, mainly through Goals 7, 9, 12 and 13. Citing examples of the efforts companies can make to contribute by decarbonizing their operations and supply chains. Seeking to reduce the carbon footprint of their products, services and processes, and setting ambitious emissions reductions targets in line with climate science. As well as scaling up investment in the development of innovative low-carbon products and services. Post-combustion carbon capture (PCC) process is one of the most promising technologies due to its ability to treat high CO<sub>2</sub> concentration streams from emission points from fossil fuels or biomass in industrial facilities. Moreover it can be adapted as a complementary process in existing thermal power plants. Monoethanolamine (MEA) is conventionally used as solvent because of its high reactivity with CO<sub>2</sub> and relatively low cost. This technology has severe disadvantages including a high rate of corrosion of process equipment and degradation of amines due to the presence of SO<sub>2</sub>, NO<sub>2</sub> and O<sub>2</sub> in the combustion gases (Luis, 2015). In addition, MEA is considered toxic, so its implementation entails a high environmental impact.

In recent years, research has focus in a new category of green solvents called Deep Eutectic Solvents (DES) with selective absorption capacity towards CO<sub>2</sub>. DESs have the characteristics of low cost, low to no toxicity, and good biodegradation. In recent years

the properties of several DES have been reported in literature. Among them, the eutectic combination of choline chloride and urea in a 1:2 (ChCl/Urea (1:2)) molar proportion has shown the highest absorption rate for CO<sub>2</sub> (García et al., 2015). This property, along with its low vapor pressure and high thermal stability, has made it a candidate for replacing MEA in CO<sub>2</sub> removal processes.

DESs industrial applications have been developed in recent years. Ma et al. modeled, simulated and evaluated the upgrading of biogas by removing CO<sub>2</sub> with aqueous ChCl/Urea (1:2), demonstrating that this solvent was suitable for this process. Using choline chloride/urea (1:2) as the absorbing agent, Luo et al. (2021) designed and simulated an industrial-scale PCC process, to treat flue gases from coal combustion. Evaluation of life cycle environmental sustainability showed that the CO<sub>2</sub> capture with DES has significantly lower values compared to MEA, highlighting the advantage of lower environmental impact. However, the application of DES in an optimization framework considering sustainable targets, is not yet developed. In other words, no study has been conducted to analyze the impact of post-combustion absorption process design variables and their effect on the sustainability of the process, using ChCl/Urea (1:2) as solvent. Furthermore, the capacity of the DES to treat flue gasses with different CO<sub>2</sub> concentrations needs to be studied, as traditional power plants require a wide operational range in terms of the fuel used for energy generation, and the associated CO<sub>2</sub> concentration.

In this work, a novel PCC process using green solvent ChCl/Urea (1:2) is designed and optimized for the first time under a sustainability scheme, using Aspen Plus. The impact of the design variables on environmental and economic metrics is analyzed. Furthermore, the plant's capacity for treating flue gases with different CO<sub>2</sub> concentrations is studied through four case studies, considering the use of Natural Gas (NG), Associated Gas (AG), Biogas (BG) and Coal.

## 2. Study Case

The study case presented in this work evaluates the scenario of a power plant coupled to a carbon capture plant using DES as solvent (Fig. 1). The power plant consists in a traditional steam-generating boiler applied in a combined-cycle configuration. A fuel feed flow of 1000 kmol/h is considered, for natural gas, biogas, and associated gas. For coal, a flue gas flow of the same order of magnitude as those obtained for the combustion of the gasses is considered. The combustion process simulation is based in the simplified model proposed by Luyben (2013). The Peng-Robinson (PR) thermodynamic model was used to calculate physical properties, and an Aspen RGibbs chemical-equilibrium reactor model is used for the combustor. The PCC process consists of an absorber/desorber system with two intermediate flash tanks. The RADFRAC module is used for the absorber and desorber columns, considering an equilibrium-based physical absorption. ChCl/Urea was modeled in Aspen Plus using a group contribution method, and critical and temperature dependent properties, as well as Henry constants for the gaseous components in the solvent were taken from experimental data. The NRTL thermodynamic model was used to calculate the activity coefficient in the liquid phase, for ChCl/Urea + water. PR thermodynamic model was used to calculate fugacity coefficient for the gaseous components, with parameters taken from Aspen databank.

The multi-objective optimization of the process modeled in Aspen Plus represents a non-linear and multivariable problem, where the objective function used as optimization criterion is generally non-convex with several local optimums. Stochastic global optimization methods, such as Genetic Algorithms, have been proven to be robust

and can handle both MESH equations and phase equilibrium calculations with complete models. A multi-objective meta-heuristic optimization algorithm differential evolution and tabu-list (MODE-TL) is used to obtain an optimal design of the PCC process. This algorithm allows the comparison of multiple solutions in terms of the objective functions, described after the economic and environmental metrics considered. For the economic indices, total annual costs (TAC) were calculated with the module costing technique outlined by Turton *et al.* (2003). Return on Investment (ROI) was calculated according to Jiménez-Gutiérrez. Eco-Indicator 99, as a measure of environmental impact, was calculated according to the methodology developed by Goedkoop *et al.* (2001). Total Solvent Recovery Energy (TSRE) and Green House Gas Emissions (GHGE) environmental indices were also considered, according to the definitions proposed by Jiménez-González *et al.* (2012). A general mathematical expression for the objective function and its respective decision variables involved in the optimization procedure are shown in Equation (1).

$$\min[TAC, -ROI, EI99, TSRE, GHGE] = f(F_{air}, x_{H_2O}, F_{DES}, T_{DES}, N_{abs}, N_{PS1}, N_{PS2}, P_{abs}, P_{F1}, \Psi_{F1}, P_{F2}, \Psi_{F2}, RR, D, N_{des}, N_{feed}, P_{des}) \quad (1)$$

$$s.t. \ y_{i,f} \geq x_{i,f}, \text{ and } w_{i,f} \geq u_{i,f}$$

Where the solutions are restricted to satisfy the recovery of 95% of the CO<sub>2</sub> produced during the combustion and also to achieve a purity of 95% mol of CO<sub>2</sub>. Where  $y_{i,f}$  represents the CO<sub>2</sub> recovered in the desorber column and  $x_{i,f}$  represents the 95% of the CO<sub>2</sub> produced during the combustion. As well,  $w_{i,f}$  represents the purity achieved at the desorber column and  $u_{i,f}$  represents the purity expected of at least 95% mol of CO<sub>2</sub>.

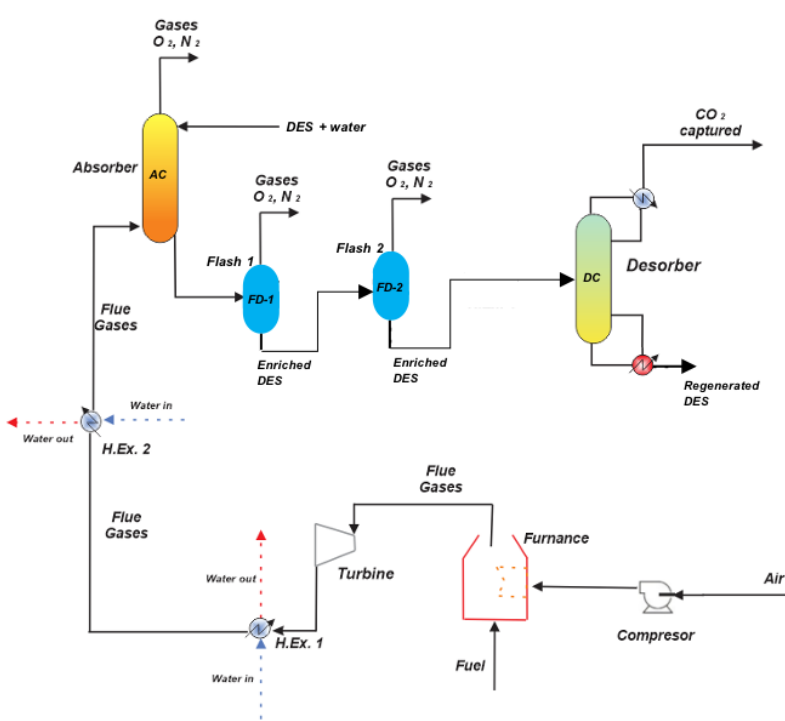


Figure 1. Process diagram of the combustion and PCC proposed scheme.



The decision variables and operation range considered for the optimization are described in Table 1. The optimization method was implemented using a hybrid platform that interconnects the Aspen Plus simulation with Excel through Visual Basic. The variable design vectors generated by the algorithm are fed to the process model, where the rigorous simulation is implemented. The outputs of the process model (flows, thermal loads, etc.) are fed back to Excel, where MODE-TL evaluates the objective functions and proposes new values of input vectors, according to the evolutionary nature of the algorithm, the process begins to iterate. The parameters used for the MODE-TL method were: 120 individuals, 800 maximum number of generations, taboo list of 50% of individuals, taboo radius of  $1 \times 10^{-4}$ , and 0.9 and 0.3 crossover probability and mutation factor, based on previously reported works in literature for optimization of multicomponent separation schemes using the RADFRAC module.

**Table 1.** Decision variables and search intervals

Variable Name	Variable Type	Symbol	Range or Value
Feed air molar flow rate, kmol h <sup>-1</sup>	Continuous	$F_{air}$	10025 – 35000
H <sub>2</sub> O:ChCl/Urea molar proportion	Continuous	$x_{H_2O}$	1.6 – 5.5
DES molar flow rate, kmol h <sup>-1</sup>	Continuous	$F_{DES}$	5000 – 120000
Absorber stages	Discrete	$N_{abs}$	20 – 100
1 <sup>st</sup> packing section ending stage	Discrete	$N_{PS1}$	2 - ( $N_{abs} - 1$ )
2 <sup>nd</sup> packing section starting stage	Discrete	$N_{PS2}$	( $N_{abs} + 1$ ) - $N_{abs}$
Temperature of solvent feed, °C	Continuous	$T_{DES}$	40 - 60
Absorber top pressure, bar	Continuous	$P_{abs}$	1.1 – 14.0
Flash 1 pressure, bar	Continuous	$P_{f1}$	1.1 – 6.0
Flash 1 vapor fraction	Continuous	$\Psi_{F1}$	0 – 1
Flash 2 pressure, bar	Continuous	$P_{f2}$	1.01 – 6.0
Flash 2 vapor fraction	Continuous	$\Psi_{F2}$	0 – 1
Desorber reflux ratio	Continuous	$RR$	0.1 - 2.5
Distillate rate, kmol h <sup>-1</sup>	Continuous	$D$	1378 – 1500
Desorber stages	Discrete	$N_{des}$	4 – 30
Desorber feed stage	Discrete	$N_{feed}$	2 - $N_{des}$
Desorber top pressure, bar	Continuous	$P_{des}$	1.1 - 14.0

### 3. Results

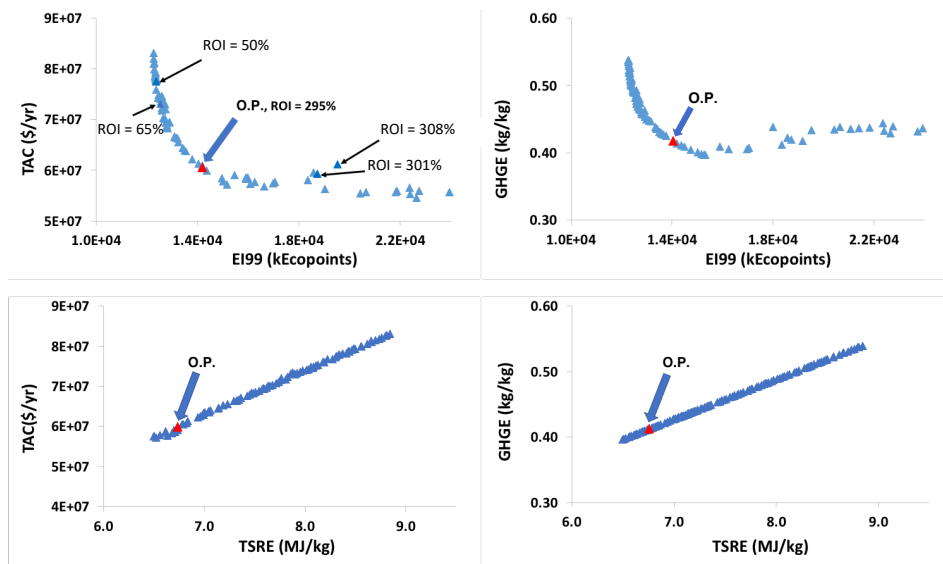
This section provides the multi-objective optimization results for the PCC process. For all designs, CO<sub>2</sub> recovery constraint is set at 95%, and 95% mol purity for the CO<sub>2</sub> at the distillate of the DC.

#### 3.1. Pareto fronts

Two dimensional pareto diagrams obtained for the coal case study are presented in Fig. 2. The form of the pareto fronts for EI99 and TAC exhibits a trend of competing objective

functions. This behavior indicates that the selection of a design with the lowest EI99 causes the TAC to increase, hence the solutions that offer the best trade-offs between the two objectives are those located in the curve zone of the Pareto chart. The pareto front for TSRE and TAC indicates that a design selection with the lowest TSRE causes TAC to diminish, therefore, designs that minimize both objectives are found in the lower left corner. The solvent recovery energy has a direct impact in TAC, given that the high use of vapor in the PCC process represents 70-80% of all annual costs. The pareto front for EI99 and GHGE exhibits a competing objective function trend, similar to EI99 vs TAC, where the best trade-off between objectives can be found in the curve area of the chart.

In terms of the impact of some design variables on the indices considered, a trend between Water:ChCl/Urea proportion and TAC was found. A higher water to ChCl/Urea proportion diminishes CO<sub>2</sub> absorption capacity, increasing the need for solvent flow. This causes process equipment topology to increase, as well as increasing TSRE requirements. As a consequence, GHGE also increases, and from the pareto chart trends in Fig. 2 EI99 decreases. Higher ChCl/Urea concentration will increase environmental impact and pumping costs considerably due to a higher viscosity of a poorly diluted ChCl/Urea. As well, the reflux ratio is directly related to the energy requirements. A higher reflux ratio causes energy requirements for the capture to increase.



**Figure 2.** Pareto fronts for EI99 vs TAC, EI99 vs GHG, TSRE vs TAC and TSRE vs GHGE.

### 3.2. Optimal design performance indices

The optimal point (O.P.) was selected from the pareto fronts, considering a compromise between minimizing the objective functions. In the pareto front for EI99 and TAC, the highest and lowest ROI values found for the optimal designs were highlighted, along with the ROI value for the O.P. To showcase that a good performance in terms of this objective was also taken into consideration when searching for the optimal design.

The optimization results reveal that the Coal process provides the lowest energy usage per unit of CO<sub>2</sub> captured, with a TSRE and GHGE which is 12.3%, 27.7% and

31.7% lower than AG, NG and BG respectively. The energy savings are also reflected in the value of TAC for Coal (8.2%, 15.5% and 16.2% lower per tonne of CO<sub>2</sub> captured, compared with AG, NG and BG respectively). In terms of environmental impact, the Coal process presents a value that is 20.1%, 15.7% and 30.2% lower per tonne of CO<sub>2</sub> captured, as compared with AG, NG and BG respectively. The coal PCC process exhibits the best performance considering all environmental and economic indices simultaneously, although the highest ROI is obtained when AG is used as fuel.

The proposed scheme (Fig. 1) for solvent regeneration using two flash drums and a desorber column, allows the separation of CO<sub>2</sub> from the enriched solvent to occur at pressures between 1 and 5 bar, avoiding a costly vacuum pressure operation.

#### 4. Conclusions

The simulation of the PCC process shows that the CO<sub>2</sub> capture of different flue gases is possible by use of green solvent ChCl/Urea (1:2).

Multi-objective optimization considered simultaneously economic and environmental objectives and resulted in optimal solutions according to sustainable design for the PCC process. Through the MODETL optimization method, design and operating conditions were found as to avoid vacuum pressure operation at the solvent regeneration steps while using a considerably lower absorption pressure (13.9 bar) compared to the PCC process proposed by Luo et al (45 bar).

The optimization results reveal that the Coal process provides the lowest energy use compared, a lower TAC and diminished environmental impact per tonne of CO<sub>2</sub> captured, compared with other PCC processes. The highest ROI was obtained when using AG as fuel.

#### References

- UN General Assembly, Transforming our world: the 2030 Agenda for Sustainable Development, 2015, A/RES/70/1, available at: <https://www.refworld.org/docid/57b6e3e44.html> [accessed 22 November 2022].
- H. Eldardiry, E. Habib, Carbon capture and sequestration in power generation: review of impacts and opportunities for water sustainability. *Energ Sustain Soc* 8, 6 (2018).
- P. Luis, Use of monoethanolamine (MEA) for CO<sub>2</sub> capture in a global scenario: Consequences and alternatives, *Desalination*, Volume 380, 2016, Pages 93-99.
- G. García, S. Aparicio, R. Ullah, and M. Atilhan, Deep Eutectic Solvents: Physicochemical Properties and Gas Separation Applications, *Energy Fuels* 2015, 29, 4, 2616–2644
- Ma, C.; Xie, Y.; Ji, X.; Liu, C.; Lu, X. Modeling, simulation and evaluation of biogas upgrading using aqueous choline chloride/urea. *Appl. Energy* 2018, 229, 1269–1283.
- F. Luo, X. Liu, S. Chen, Y. Song, X. Yi, C. Xue, L. Sun, and J. Li., Comprehensive Evaluation of a Deep Eutectic Solvent Based CO<sub>2</sub> Capture Process through Experiment and Simulation, *ACS Sustainable Chemistry & Engineering* 2021 9 (30), 10250-10265.
- R. Turton, R. C. Bailie, W. B. Whiting, and J. A. Shaeiwitz, *Analysis, Synthesis, and Design of Chemical Processes*, Second Edition, PrenticeHall, 2003.
- M. Goedkoop, R. Spriensma, 2001, *The Eco-Indicator 99: A Damage Oriented Method for Life Cycle Impact Assessment*.
- C. Jiménez-Gonzalez, D. J. C. Constable & C. S. Constable. *Chem. Soc. Rev.*, 2012,41, 1485-1498
- A.J. Gutiérrez, *Diseño de procesos en ingeniería química*, Reverte, 2021
- W. L. Luyben, Compressor heuristics for conceptual process design. *Ind. Eng. Chem. Res.* 2011, 50, 13984–13989.
- J. R. Couper, W. R. Penney, J. R. Fair, S. M. Walas, 2005. *Chemical Process Equipment – Selection and Design*. Gulf Professional Publishing.

# Optimal Retrofitting of Conventional Oil Refinery into Sustainable Bio-refinery under Uncertainty

Lifeng Zhang,<sup>a</sup> Ana Inés Torres,<sup>b</sup> Bingzhen Chen,<sup>a</sup> Zhihong Yuan,<sup>a\*</sup> Ignacio E. Grossmann<sup>b\*</sup>

<sup>a</sup>*Department of Chemical Engineering, Tsinghua University, Beijing, 100084, China*

<sup>b</sup>*Chemical Engineering Department, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA*

Corresponding authors: [zhongyuan@mail.tsinghua.edu.cn](mailto:zhongyuan@mail.tsinghua.edu.cn) (Z. Yuan),  
[grossmann@cmu.edu](mailto:grossmann@cmu.edu) (I.E. Grossmann)

## Abstract

This paper focuses on a novel multistage retrofitting problem from conventional fossil-based refinery into biomass-based refinery with the aim of sustainable development of fuels. Given a typical refinery and potential biomass-based technologies, the problem has an objective to integrate the latter into the former by making use of existing units in the new process(es). A mixed-integer linear programming model that considers a ten-year long retrofit planning is formulated. Furthermore, the deterministic problem is extended to a multi-stage stochastic programming problem under both endogenous yield uncertainty and exogenous demand uncertainty and solved via a Lagrangean decomposition algorithm. The results provide flexible design alternatives by determining the units that should be added or modified for the selected biomass-based technologies. Regarding uncertainty, different schemes with strategic and operations decisions are determined. The results show advantages in considering retrofitting problem with detailed operation constraints for each year.

**Keywords:** Refinery; Biomass; Retrofit; Uncertainty; Lagrangean decomposition.

## 1. Introduction

Significant growth in renewable energy, such as solar and biomass, has taken place over the last 30 years. As the core part of energy supply, the petroleum refinery industry is increasingly gaining importance for producing more sustainable energy instead of fossil fuels. Among these novel energy forms, biomass-based energy is a promising alternative for the refinery industry since the biomass can be used to produce hydrocarbons that are drop-in replacements for the ones obtained from crude oils. This provides opportunities to repurpose existing crude oil refineries into biomass refineries. Such a plan has been under consideration in both industry and academia for a long time and has recently seen renewed interest (see for example Exxon Mobil (Sanicola, 2021), and in academia (Ericson et al., 2019)). The evaluation of such design is often executed with Life Cycle Assessment (LCA) (Garcia-Nunez et al., 2016) under predefined flowsheet and preselected biomass-based technologies. However, such framework usually ignores the flexibility to adopt flowsheets according to expected development in market and industry. This paper proposes a mixed linear programming (MILP) modeling framework to solve the long-term retrofit problem, while considering detailed production constraints in each year. Uncertainty is also addressed by extending the proposed model with stochastic programming.

## 2. Problem Statement

### 2.1. Crude-based Refinery Demonstration

The flowsheet of a prototypical crude-based refinery is shown in Fig. 1. The ranges of hydrocarbons are labelled to represent the corresponding components, typically considered in each stream/processed by each equipment. For example, the crude oil, is first distilled in crude distillation unit (CDU) into liquified petroleum gas (LPG, composed of hydrocarbons  $C_3$  to  $C_5$ ), naphtha (NAP composed of hydrocarbons  $C_5$  to  $C_{10}$ ), etc. These streams are then further hydrotreated, cracked and reformed in secondary processing units, such as continuous reformer unit (CRU), fluid catalytic cracking (FCC) and others. The final products, LPG, gasoline (GASO), jet fuel (JET), diesel (DIESEL) and fuel oil (OIL) are obtained by blending intermediates to satisfy quality specifications.

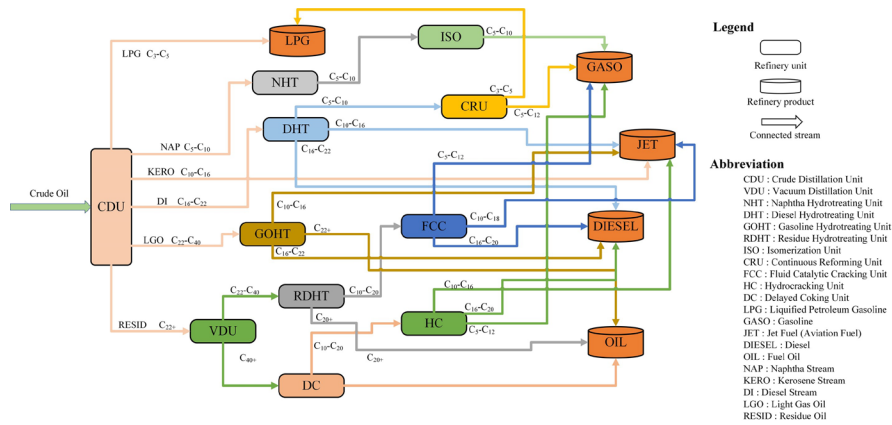


Fig. 1 Flowsheet of a given crude-based refinery. (Presented in Appendix)

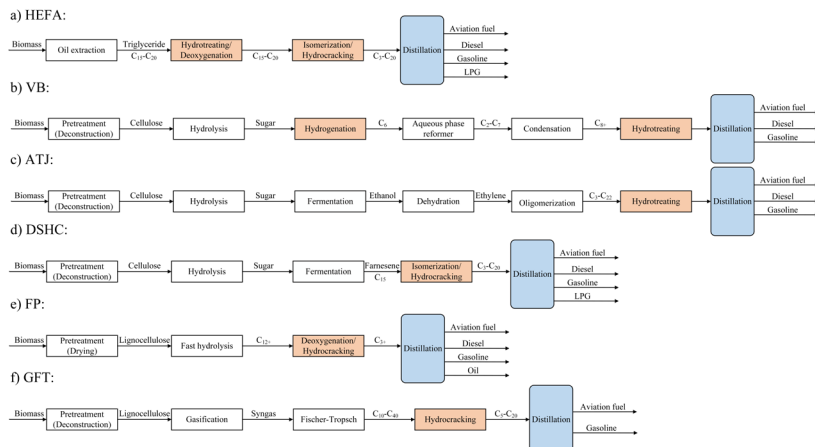


Fig. 2 Detailed process steps for potential biotechnologies. (Presented in Appendix)

### 2.2. Biomass-based Refinery Demonstration

To produce sustainable energy based on biomass, six biomass-based technologies are considered, namely, Hydroprocessed esters and fatty acids (HEFA), Virent's BioForming (VB), alcohol to jet (ATJ), direct sugar to hydrocarbon (DSHC), fast pyrolysis (FP), and gasification and Fischer-Tropsch (GFT) (Tanzil et al., 2021a, 2021b). Details of each technology path can be seen in Fig. 2 with illustration for every step and stream.

For example, the VB technology, which addresses the pine or stover as the feedstocks, first transforms the materials into cellulose. The cellulose is converted into sugar via hydrolysis. Then the sugar is hydrotreated into hydrocarbons and followed by aqueous phase reforming, condensation and hydrotreating. Finally, the biomass-based hydrocarbons are distilled to produce aviation fuel, gasoline and diesel. In the VB technology, the hydrotreating steps can be executed in a newly built unit or in an already existing refinery hydrotreating unit. We compare components of the streams entering each unit to find possible matches. In addition, the distillation step can also be done in the CDU instead of in new-installed units.

### 3. Mathematical Formulation

#### 3.1. Deterministic Model

The goal is to minimize the total cost over ten years by subtracting sales income from the summation of material operation reconnection and new unit installation costs. For each year, the operation constraints on selection of materials and production are also considered. The problem is formulated as a MILP as follows and details for the mathematical formulation can be seen in Appendix.

$$\begin{aligned} \min z &= a^T x + b^T y \\ Ax + By &\leq d \\ x \geq 0, y &\in \{0,1\}^m \end{aligned} \quad \text{MILP}$$

#### 3.2. Multistage Stochastic Programming Model

The problem is extended to a multistage stochastic programming (MSSP) under endogenous yield uncertainty and exogenous demand uncertainty (Apap and Grossmann, 2017). The scenario tree is formulated as Fig. 3, where each stage is formulated as an inner two-stage stochastic programming problem with endogenous  $\xi$  uncertainty and exogenous  $\theta$  uncertainty (Lara et al., 2019). The decisions are divided into strategic decisions and operational decisions. The strategic decisions, including the selection of biomass-based technology, capacity expansion and unit substitution, are assumed only related to endogenous uncertainty. The operational decisions, such as type of crude oil, amount of feedstock and sale strategy for products in each year, are assumed related to demand uncertainty and are independent in each year.

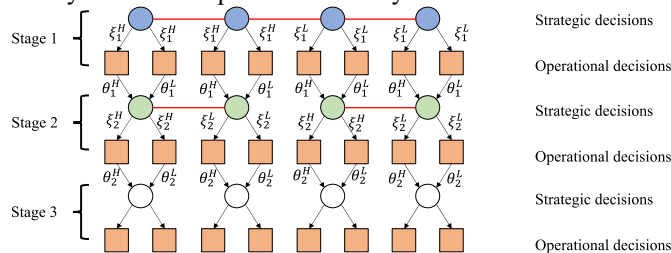


Fig. 3 Scenario tree illustration. (Presented in Appendix)

The mathematical formulation of the MSSP is given as follows. The set  $i \in I$  denotes the outer scenarios corresponding to endogenous yield uncertainty and set  $j \in J$  denotes the inner scenarios related to exogenous demand uncertainty appearing in each year. The problem is solved by reducing scenario pairs first, and then using a Lagrangean decomposition algorithm by dividing it into subproblems based on outer scenario pairs. Details on the modeling and algorithm can be found in Gupta and Grossmann (2014, 2011).

$$\begin{aligned}
 \min z = & \sum_{i \in I} p_i \left( \sum_{t \in T} c_{t,i} y_{t,i} + \sum_{j \in J} p_j d_{t,i,j} x_{t,i,j} \right) \\
 \text{s. t. } & \sum_{\tau \leq t} a_{\tau,i} y_{\tau,i} \leq b_{t,i} \quad \forall t \in T, i \in I \\
 & e_{t,i} y_{t,i} + f_{t,i,j} x_{t,i,j} \leq g_{t,i,j} \\
 & y_{t,i} = y_{t,i'} \quad \forall t \in \{T_1\}, i, i' \in I \\
 & Z_{t,i,i'} \Leftrightarrow F(y_{t1,i}, y_{t2,i} \dots y_{t-1,i}) \quad \forall t \in T \setminus \{T_1\}, i, i' \in I \\
 & \begin{bmatrix} Z_{t,i,i'} \\ y_{t,i} = y_{t,i'} \end{bmatrix} \vee [\neg Z_{t,i,i'}] \quad \forall t \in T \setminus \{T_1\}, i, i' \in I \\
 & y_{t,i} \in \{0,1\} \quad \forall t \in T, i \in I \\
 & x_{t,i,j} \in R \quad \forall t \in T, i \in I, j \in J
 \end{aligned}$$

MSSP

### 4. Results and Discussion

#### 4.1. Computational Performance

Both the deterministic problem and the MSSP problem are solved in Gurobi 9.5.2 via GAMS 41.1.0 on Windows 11 with Intel(R) Core(TM) i7-9700 CPU @ 3.00GHz and 16GB memory. The model statistics and computational results are shown in Tables. 1 - 2 and Fig. 4.

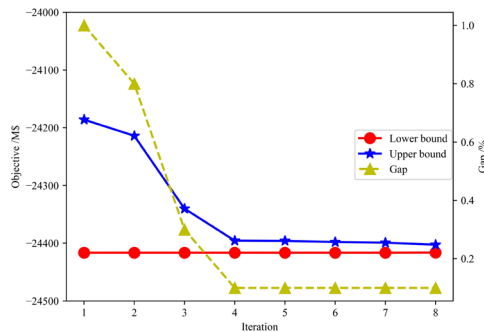


Fig. 4 Computational performance of Lagrangean decomposition algorithm. (Presented in Appendix)

Table.1 Model statistics.

Problem	# of continuous variable	# of binary variable	# of constraints
Deterministic model	14,197	7,146	13,954
MSSP	11,576,705	459,264	15,808,450

Table.2 Optimization results.

Problem	Solver	CPU time /s	Gap /%	Objective /M\$
Deterministic model	CPLEX	138.5	0.01	-1020.85
	Gurobi	16.6	0.01	-1020.85
MSSP	CPLEX	7200.0	/	/
	Gurobi	7200.0	/	/
	Lagrangean decomposition	4511.3	0.1	-24402.6

The deterministic model is solved to optimality quite fast. For the MSSP model, 64 outer scenarios are included related to the combinations of endogenous yield uncertainty and in each year, 27 scenarios related to exogenous demand uncertainty are considered, leading to 270 inner scenarios in total. The scale of the MSSP increases exponentially as shown in Table. 1 and commercial solvers fail to give a feasible solution within 2 h of

CPU time; the Lagrangean decomposition algorithm, solves the problem to an optimality gap of 0.1% in ~ 75 min.

4.2. Results of Deterministic Model

From the deterministic model, HEFA, VB, FP, GFT are selected to produce the desired bioproducts. The retrofitted flowsheets in year 1, year 2, year 5 and year 9 are displayed in Fig. 5. For a better illustration, high qualified figures are presented in Appendix.

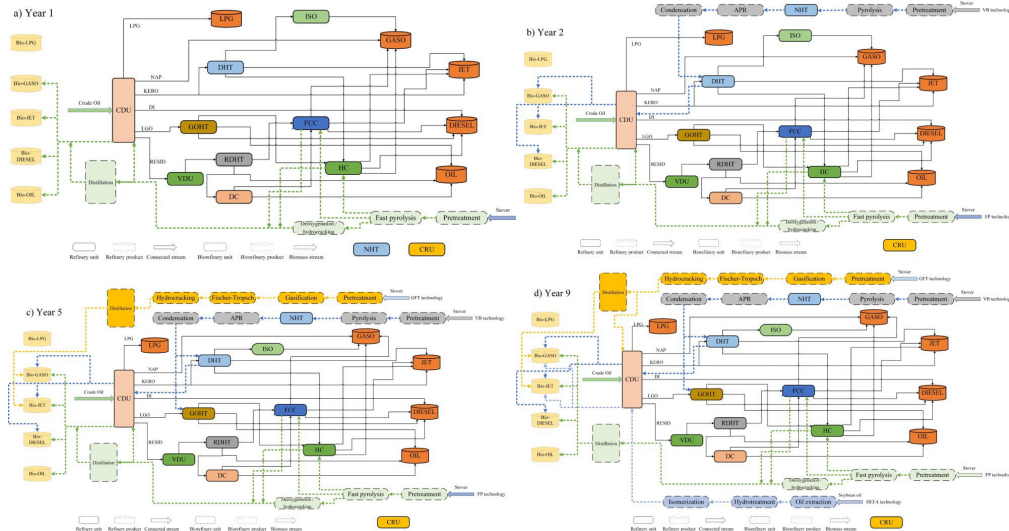


Fig. 5 Retrofitted flowsheet of the biorefinery over several years. (Presented in Appendix)

As can be seen VB is built in the first year to process corn stover. For each step in VB, a new operation unit is built. While in the hydrocracking step, the biomass-based hydrocarbon stream can be directly fed into the FCC and HC units for deep cracking. Although a distillation unit is built for VB, the CDU is still used to obtain bioproducts with the same time until year 7. The NHT and CRU units are removed from the refinery in the first year, and NHT is used in VB technology as the hydrotreating step as well as DHT unit in year 2. Later, the GOHT is also used to process the biomass-based hydrocarbons from VB as the final hydrotreating step in addition to oil-based stream. GFT is implemented in year 5 with all the units installed. In year 9, the HEFA technology is also installed and the products are obtained via the CDU. Note that the continuous reforming unit (CRU) is removed and never used during the time horizon.

Table.3 Statistics of result from MSSP

Bio-technology	Scenarios selecting bio-technology	Bio-technology	Scenarios selecting bio-technology
HEFA	28	DSHC	0
VB	64	FP	64
ATJ	0	GFT	36

4.3. Results of MSSP

With the MSSP model, the biomass-based technology selection statistics are presented in Table. 3. For a total of 64 scenarios, VB and FP are all selected as the main process to produce biomass-based hydrocarbons, while ATJ and DSHC are not. Comparing to VB and FP, the higher expense and lower yield of ATJ and DSHC prevent their applications. To compensate for the production demands, GFT and HEFA are the backup biomass-



based technologies to install in later years. The GFT is preferred since it is selected in 36 scenarios while HEFA is selected in 28 scenarios.

To evaluate the effect of MSSP, the value of stochastic solution ( $VSS$ ) is also calculated by solving the problem stage by stage iteratively (Zhang et al., 2018). The definition of  $VSS$  and  $\overline{VSS}$  are presented as follows.

$$VSS = z^{DP} - z^{MSSP} \quad (1)$$

$$\overline{VSS} = (z^{DP} - z^{MSSP}) / z^{DP} \quad (2)$$

$z^{DP}$  is the optimal solution from solving problem for each stage iteratively.  $z^{MSSP}$  is the solution from MSSP model. The  $VSS$  equals to 522 M\$ and the  $\overline{VSS}$  is 2.15% which reveal the benefit of implementing stochastic programming.

## 5. Conclusion

This paper addressed the retrofit problem from a conventional refinery to a biorefinery by incorporating new biomass-based technologies into an existing oil-based refinery. The problem is extended into a MSSP model under both endogenous and exogenous uncertainties. Each stage is divided into strategic and operational steps. The Lagrangean decomposition algorithm is used to solve such problem. The results show that the FP and VB technologies are preferred, and a cost saving of 522 M\$ can be obtained with stochastic programming with the evaluation of  $VSS$ .

## Appendix

The Appendix can be seen at <https://cloud.tsinghua.edu.cn/f/9d6288101611467d9601/>

## References

- Apap, R.M., Grossmann, I.E., 2017. Models and computational strategies for multistage stochastic programming under endogenous and exogenous uncertainties. *Comput. Chem. Eng.* 103, 233–274.
- Ericson, S.J., Engel-Cox, J., Arent, D.J., 2019. Approaches for Integrating Renewable Energy Technologies in Oil and Gas Operations (No. NREL/TP-6A50-72842). National Renewable Energy Lab. (NREL), Golden, CO (United States).
- Garcia-Nunez, J.A., Rodriguez, D.T., Fontanilla, C.A., Ramirez, N.E., Silva Lora, E.E., Frear, C.S., Stockle, C., Amonette, J., Garcia-Perez, M., 2016. Evaluation of alternatives for the evolution of palm oil mills into biorefineries. *Biomass Bioenergy* 95, 310–329.
- Gupta, V., Grossmann, I.E., 2014. A new decomposition algorithm for multistage stochastic programs with endogenous uncertainties. *Comput. Chem. Eng.* 62, 62–79.
- Gupta, V., Grossmann, I.E., 2011. Solution strategies for multistage stochastic programming with endogenous uncertainties. *Comput. Chem. Eng.* 35, 2235–2247.
- Lara, C.L., Sirola, J.D., Grossmann, I.E., 2019. Electric power infrastructure planning under uncertainty: stochastic dual dynamic integer programming (SDDiP) and parallelization scheme. *Optim. Eng.* 21, 1243–1281.
- Sanicola, L., 2021. EXCLUSIVE Exxon, Chevron look to make renewable fuels without costly refinery upgrades -sources. Available at <https://www.reuters.com/world/middle-east/exclusive-exxon-chevron-look-make-renewable-fuels-without-costly-refinery-2021-08-12/>

- Tanzil, A.H., Brandt, K., Wolcott, M., Zhang, X., Garcia-Perez, M., 2021a. Strategic assessment of sustainable aviation fuel production technologies: Yield improvement and cost reduction opportunities. *Biomass Bioenergy* 145.
- Tanzil, A.H., Brandt, K., Zhang, X., Wolcott, M., Stockle, C., Garcia-Perez, M., 2021b. Production of Sustainable Aviation Fuels in Petroleum Refineries: Evaluation of New Bio-Refinery Concepts. *Front. Energy Res.* 9.
- Zhang, Q., Bremen, A.M., Grossmann, I.E., Pinto, J.M., 2018. Long-Term Electricity Procurement for Large Industrial Consumers under Uncertainty. *Ind. Eng. Chem. Res.* 57, 3333–3347.



# A Combined D-optimal and Estimability Model-Based Design of Experiments of a Batch Cooling Crystallization Process

Xuming Yuan<sup>a</sup>, Brahim Benyahia<sup>a,\*</sup>

<sup>a</sup> *Department of Chemical Engineering, Loughborough University, Epinal Way, Loughborough, LE11 3TU, United Kingdom*  
*B.Benyahia@lboro.ac.uk*

## Abstract

In this work, a systematic methodology is proposed to help develop model-based design of experiments to build robust and reliable mathematical model of a batch crystallization process. The cooling crystallization of paracetamol in water and propanol is used as the case study. The mathematical model consists in the mass balance and a set of population balance equations, involving primary and secondary nucleation, growth, agglomeration, breakage and dissolution kinetics. Firstly, a structural identifiability approach is used to investigate whether the model parameters can be determined uniquely with an idealized input-output behavior of the process. The approach is also critical to determine the minimum set of required observable outputs and help discriminate model candidates. A novel Model-Based Design of Experiments (MBDoE) is then proposed based on the combination of the D-optimality criterion and the estimability analysis. The objective is to reduce the uncertainties in the model parameters by enhancing the data information content and help maximize the estimability potential of all model parameters while reducing correlation amongst them. Moreover, a new operating strategy based on temperature cycling is used in a sequential design of experiment to maximize data information content from one single experiment while reducing the experimental burden and inherent wastes.

**Keywords:** Structural Identifiability, Estimability, Sensitivity Analysis, Model-Based Design of Experiment (MBDoE), Temperature Cycling, Batch Crystallization.

## 1. Introduction

Crystallization is an important separation and purification technique that is widely applied in the pharmaceutical industry. The successful development of crystallization technologies is a painstaking and costly process. The availability of a reliable mathematical models is critical at all development stages to help explore more effectively the design space and deliver robust and cost-effective designs, operation, and control procedures (Benyahia et al., 2021; Liu and Benyahia, 2022). Most importantly, mathematical models can help achieve built-in quality insurance (e.g., Quality-by-Design and Quality-by-Control). However, the prediction capabilities and inherent uncertainties of the models strongly depend on the experimental data obtained at early development stages, their information content, and the parameter estimation procedure.

To obtain the parameter values, parameter estimation is performed, where the difference between the experimental data and the prediction is minimized. However, an excellent matching between the collected data and the prediction does not necessarily indicate that all the model parameter estimates are accurate and precise. For example, it can be found

that wide ranges of parameter values may still result in similar model predictions (Balsa-Canto et al., 2010). Another common issue is the high uncertainty in the model parameter estimates. Such issues originate from the lack of structural identifiability or estimability (practical identifiability) of the model, and they may lead to the high uncertainty in model predictions which consequently affect the final designs and product quality.

To address these issues, extensive studies focusing on estimability have been carried out. In 2003, Yao et al. proposed the use of a sequential orthogonalization algorithm to rank the parameters according to their estimability potential and least correlation. The method became very popular over the last two decades and found many successful applications in various research areas (Benyahia et al. 2011 and 2013). Different estimability methods were also used to capture more effectively the minimum subset of the most estimable parameters using more effective cut-off procedures (Fysikopoulos et al., 2019). Another method based on the mean-squared-error (MSE) and transformed parameters was used by Kim and Lee (2019) to help identify the cut-off value more effectively. Nonetheless, structural identifiability, as a necessary condition of estimability, was usually neglected in previous studies. Hence, Balsa-Canto et al., 2010 performed structural identifiability analysis prior to the estimability analysis to confirm that all model parameters have unique estimates theoretically. Recently, model-based design of experiment (MBDoe) is gaining more attention and it has been implemented in optimal experimental design to obtain high-quality parameter estimates. For instance, Maheshwari et al., 2013 applied to a double feedback gene switching model a multi-objective optimal experimental design that maximized the D-optimality criterion and minimized the correlation between the parameters simultaneously.

In this work, a novel approach that combines the estimability and MBDoe is proposed as a systematic approach to build more predictive models and design more effective and information rich experiments. The aim of this study is to demonstrate the benefits of combining D-optimal design with the estimability to reduce the uncertainty in crystallization models. This work systematically performs structural identifiability analysis, estimability analysis and D-optimal design in a seeded batch cooling crystallization process of paracetamol (PCM), which includes primary and secondary nucleation, growth, agglomeration, breakage and dissolution. Several new experimental strategies are proposed in the MBDoe which includes temperature cycling with holds.

## 2. Methodology

### 2.1 Process Model

A seeded batch cooling crystallization of paracetamol (PCM) is conducted in a 4:1 water/propanol solvent system, and the model is based on the quadrature method of moments (QMOM). Due to the temperature cycling, two mathematical models are used (hybrid model) as described below to capture the dissolution of crystals. When the concentration of PCM is higher than the concentration at saturation, primary nucleation, secondary nucleation, growth, agglomeration and breakage are considered, while dissolution is suppressed. The population balance of the  $k$ th moment is given by

$$\frac{d\mu_k}{dt} = 0^k(k_{b1}S^{b1} + k_{b2}S^{b2}\mu_2^{j2}) + kk_gS^g \sum_{i=1}^{N_q} W_i L_i^{k-1} + \frac{1}{2}K_a \sum_{i=1}^{N_q} \sum_{j=1}^{N_q} W_i W_j (L_i^3 + L_j^3)^{1+\frac{k}{3}} - \frac{1}{2}K_a \sum_{i=1}^{N_q} \sum_{j=1}^{N_q} W_i W_j L_i^k (L_i^3 + L_j^3) + 2^{1-\frac{k}{3}}K_b \sum_{i=1}^{N_q} W_i L_i^{\gamma+k} - K_b \sum_{i=1}^{N_q} W_i L_i^{\gamma+k} \quad (1)$$

where  $N_q = 2$  is the order of quadrature approximation,  $W$  and  $L$  are the weights and abscissa obtained from quadrature approximation, and  $S$  is the absolute supersaturation:

$$S = C - C^* \quad (2)$$

$C^*$  is the concentration at saturation captured by a polynomial function of temperature  $T$ :

$$C^* = p_0 + p_1T + p_2T^2 \quad (3)$$

where  $p_0$ ,  $p_1$  and  $p_2$  are the known polynomial coefficients.

Let  $\bar{d}$ ,  $\rho$  and  $k_v$  be the mean crystal size, the density and the shape factor of PCM crystals. The evolution of the PCM concentration and mean crystal size along with time are mathematically expressed as

$$\frac{dC}{dt} = -3\rho k_v k_g S^g \sum_{i=1}^{N_q} W_i L_i^2 \quad (4)$$

$$\frac{d\bar{d}}{dt} = \frac{1}{\mu_0} \frac{d\mu_0}{dt} - \frac{\mu_1}{\mu_0^2} \frac{d\mu_0}{dt} \quad (5)$$

The cooling/heating rate  $\frac{dT}{dt} = R$  is the decision variables in the optimal experimental design problem. The evolution of concentration at saturation is given by

$$\frac{dC^*}{dt} = (p_1 + 2p_2T) \frac{dT}{dt} \quad (6)$$

In summary, the vectors of state  $\mathbf{x}(t)$ , observables  $\mathbf{y}(t)$ , control variables  $\mathbf{u}(t)$  and unknown model parameters  $\boldsymbol{\theta}$  are given by

$$\mathbf{x}(t) = [\mu_0(t), \mu_1(t), \mu_2(t), \mu_3(t), C(t), \bar{d}(t), C^*(t), T(t)]^T \quad (7)$$

$$\mathbf{y}(t) = [C(t), \bar{d}(t), \mu_0(t)]^T \quad (8)$$

$$\mathbf{u}(t) = \frac{dT(t)}{dt} = R(t) \quad (9)$$

$$\boldsymbol{\theta} = [k_{b1}, b_1, k_{b2}, b_2, j_2, k_g, g, K_a, K_b, \gamma, k_{ds}, ds]^T \quad (10)$$

When the PCM concentration is lower than its concentration at saturation, dissolution is activated; primary/secondary nucleation and growth are suppressed, while agglomeration and breakage still occur. The population balance of the  $k$ th moment becomes:

$$\begin{aligned} \frac{d\mu_k}{dt} = & k k_{ds} (-S)^{ds} \sum_{i=1}^{N_q} W_i L_i^{k-1} + \frac{1}{2} K_a \sum_{i=1}^{N_q} \sum_{j=1}^{N_q} W_i W_j (L_i^3 + L_j^3)^{1+\frac{k}{3}} - \\ & \frac{1}{2} K_a \sum_{i=1}^{N_q} \sum_{j=1}^{N_q} W_i W_j L_i^k (L_i^3 + L_j^3) + 2^{1-\frac{k}{3}} K_b \sum_{i=1}^{N_q} W_i L_i^{\gamma+k} - K_b \sum_{i=1}^{N_q} W_i L_i^{\gamma+k} \end{aligned} \quad (11)$$

In this scenario, the evolution of the PCM concentration becomes

$$\frac{dC}{dt} = -3\rho k_v k_{ds} (-S)^{ds} \sum_{i=1}^{N_q} W_i L_i^2 \quad (12)$$

while the evolution of mean crystal size and concentration at saturation can still be expressed using equations (5) and (6).

## 2.2 Formulation of the MBDoE Optimization Problem

Before performing MBDoE, a structural identifiability analysis was conducted, and it showed that the model is at least structurally locally identifiable. Afterwards, four operating strategies were proposed to achieve the intended systematic MBDoE namely: linear cooling, piecewise linear cooling with continuity, temperature cycling, and temperature cycling with holds (Figure 1 (A)). For the sake of brevity, only the formulation of the optimization problem using temperature cycling with holds is presented here.

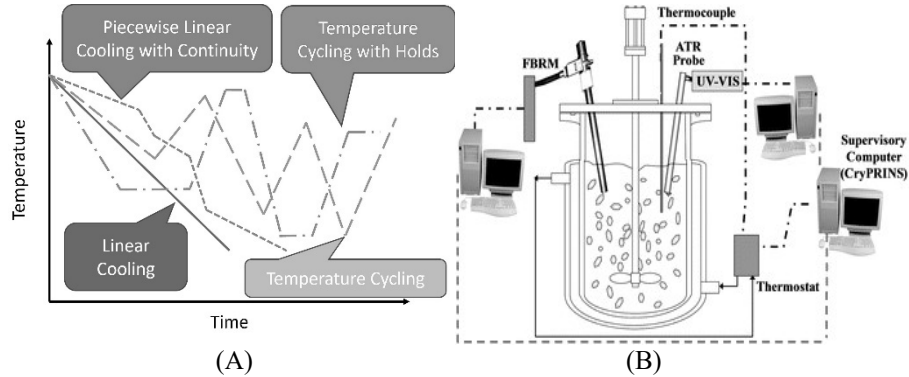


Figure 1. (A) Proposed operating strategies using different temperature trajectories. (B) Experimental set up.

The mathematical formulation of the multiobjective MBDOE is given by

$$\begin{aligned}
 & Obj1: \min_{R, \Delta t_R, t_s} \det(FIM^{-1}); \quad Obj2: \max_{R, \Delta t_R, t_s} \min ||Z_j|| \\
 & \text{s. t. } \dot{\mathbf{x}}(t) = f(t, \mathbf{x}(t; \boldsymbol{\theta}); \boldsymbol{\theta}), \quad \mathbf{x}(t_0) = \mathbf{x}_0(\boldsymbol{\theta}) \\
 & C1: -0.5 \leq R_i \leq 0 \quad (i = 1, 3, 5) \\
 & C2: 0 \leq R_i \leq 1 \quad (i = 2, 4, 6) \\
 & C3: 0 \leq \Delta t_{R,j} \quad (j = 1, \dots, 12) \\
 & C4: \sum_{j=1}^{12} \Delta t_{R,j} = 300 \\
 & C5: 21 \leq 42 + \sum_{i=1}^6 R_i \Delta t_{R,2i-1} \leq 42 \\
 & C6: 0 \leq t_{s,1} \\
 & C7: t_{s,10} \leq 300 \\
 & C8: t_{s,k} - t_{s,k+1} \leq 0 \quad (1 \leq k \leq 9)
 \end{aligned} \tag{11}$$

where  $R$  denotes the cooling/heating rate,  $\Delta t_R$  represents the durations of different cooling/heating stages, and  $t_s$  refers to the sampling times.

The first criterion is the minimization of the D-optimality criterion, which aims at the minimization of the joint confidence regions of the model parameters; the second criterion maximizes the estimability score of the least estimable parameter (Max Min), so that its effect on the observables is enhanced, which may ensure higher quality parameter estimates.

### 3. Results & Discussion

Figure 2 shows a sample of results obtained with the multi-objective optimal experimental design. The Pareto Front in Figure 2 (A) shows the best MBDoE compromises. The star represents one selected Pareto solution that can be experimentally implemented. Figure 2 (B) shows the temperature trajectory of the selected Pareto optimal solution. The estimability analysis revealed that  $k_{b2}$  was the least estimable parameter, while  $\gamma$  was found to be the most estimable.

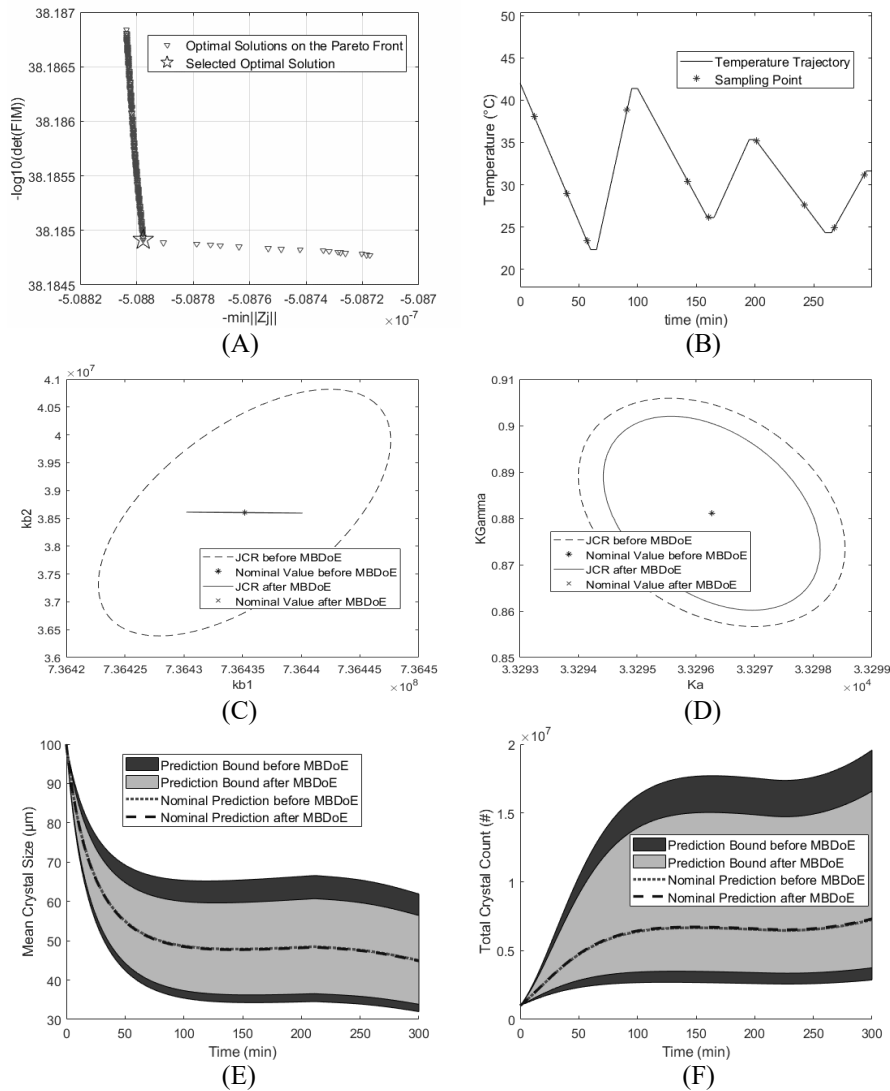


Figure 2. (A) Pareto Front of the optimal experimental design using temperature cycling and holds. (B) Optimal temperature operating trajectory. (C) 95% Joint confidence regions (JCR) of  $k_{b1}$  and  $k_{b2}$  before & after optimal experimental design. (D) 95% Joint confidence regions (JCR) OF  $K_a$  and  $\gamma$  before & after optimal experimental design. (E) Prediction bounds of mean crystal size before & after optimal experimental design. (F) Prediction bounds of total crystal count before & after optimal experimental design.

A basic single temperature cycle experiment was used as a reference case to show the benefits of the proposed MBDoe. Figure 2 (C) clearly shows that the joint confidence interval of  $k_{b2}$  and  $k_{b1}$  are significantly decreased and so is their correlation. The results shown in Figure 2 (D) demonstrate a decrease of the uncertainty in the most estimable model parameter  $\gamma$ . The decrease in parameter uncertainties resulted in more reliable predictions, as shown in Figure 2 (E) and (F), where the nominal predictions refer to the predictions using the nominal parameter estimates before and after applying the optimal



experimental design. Several additional operating strategies were also investigated in this study as outlined in Figure 1 (A) (the results are not shown here for the sake of brevity). However, temperature cycling with holds resulted in the best parameter quality and the lowest uncertainties in model prediction. Thus, this operating strategy has the potential to be implemented in the development of MBDoE for similar crystallization systems.

#### 4. Conclusions

A series of novel MBDoE strategies based on different operating trajectories were proposed and implemented to a batch crystallization process. Both D-optimal design of experiments and the estimability approaches were implemented to generate information-rich experiments and help reduce parameter uncertainties while increasing the estimability potential of the parameters. The combination of both strategies (multi-objective MBDoE) offered more optimal operating options and demonstrated the parameter estimates with reduced prediction uncertainties. The introduction of temperature cycling with holds was shown to be the best operating strategy in this study, which indicates that it has the potential to be a reliable alternative for the development and implementation of systematic MBDoE in crystallization.

**Acknowledgements:** This work was funded by the EPSRC (EP/V062077/1) Digital Medicines Manufacturing Research Centre.

#### References

- Balsa-Canto, E., Alonso, A.A. and Banga, J.R. (2010). An iterative identification procedure for dynamic modeling of biochemical networks. *BMC Systems Biology*, 4(1).
- Benyahia, B., Latifi, M.A., Fonteix, C. and Pla, F. (2011). Modeling of a batch emulsion copolymerization reactor in the presence of a chain transfer agent: Estimability analysis, parameters identification and experimental validation. *Computer Aided Chemical Engineering*, 29, pp.126–130.
- Benyahia, B., Latifi, M.A., Fonteix, C. and Pla, F. (2013). Emulsion copolymerization of styrene and butyl acrylate in the presence of a chain transfer agent. Part 2: Parameters estimability and confidence regions. *Chemical Engineering Science*, 90, pp.110–118.
- Benyahia, B., Anandan, P.D. and Rielly, C. (2021). Robust Model-Based Reinforcement Learning Control of a Batch Crystallization Process. In: *2021 9th International Conference on Systems and Control (ICSC)*, pp.89–94.
- Fysikopoulos, D., Benyahia, B., Borsos, A., Nagy, Z.K. and Rielly, C.D. (2019). A framework for model reliability and estimability analysis of crystallization processes with multi-impurity multi-dimensional population balance models. *Computers & Chemical Engineering*, 122, pp.275–292.
- Liu, J. and Benyahia, B. (2022). Optimal start-up strategies of a combined cooling and antisolvent multistage continuous crystallization process. *Computers & Chemical Engineering*, 159, p.107671.
- Kim, B. and Lee, J.H. (2019). Parameter subset selection and biased estimation for a class of ill-conditioned estimation problems. *Journal of Process Control*, 81, pp.65–75.
- Maheshwari, V., Kandpal, M. and Samavedham, L. (2013). Multi-Criteria Optimization Based Experimental Design for Parameter Estimation of a Double Feedback Gene Switching Model. *IFAC Proceedings Volumes*, 46(31), pp.333–337.
- Yao, K.Z., Shaw, B.M., Kou, B., McAuley, K.B. and Bacon, D.W. (2003). Modelling Ethylene/Butene Copolymerization with Multi-site Catalysts: Parameter Estimability and Experimental Design. *Polymer Reaction Engineering*, 11(3), pp.563–588.

# A novel mathematical formulation for short-term scheduling of multipurpose batch plants in chemical manufacturing

Dan Li,<sup>a</sup> Taicheng Zheng,<sup>a</sup> Jie Li,<sup>a,\*</sup>

<sup>a</sup>*Centre for process Integration, Department of Chemical Engineering, The University of Manchester, Manchester, M13 9PL, United Kingdom*

## Abstract

In this work, a novel mixed-integer linear programming formulation is proposed based on the unit-specific event-based representation to address the short-term scheduling of multipurpose batch plants in network environment. Concepts of material transfers between processing units or storage tankers are introduced to conditionally sequence and align the related production and consumption tasks. In this case, material flows between the processing units can be explicitly tracked, resulting in decreases on appropriate number of event points and improvements on computational efficiency. Additionally, the processing units are allowed to temporarily store a batch after production over multiple event points. Computational studies shows that the developed continuous-time based formulation leads us to fewer numbers of event points in many cases and better objective results with a maximum improvement of 67%. More interestingly, all considered examples reach the optimality without any task spanning over multiple event points.

**Keywords:** Unit-specific event-based formulation, multipurpose batch process, material transfer, short-term scheduling

## 1. Introduction

Process scheduling is a key managerial tool for manufacturing, which seeks to allocate limited resources, determine task assignments and sequences to achieve production targets over time [Rakovitis et al. (2019)]. Multipurpose batch plants [Li et al. (2022)] widely exist in chemical industry for production of low-volume and high-value products. Unit-specific event-based formulation is an efficient continuous-time based approach to address the short-term scheduling of multipurpose batch plants. Numbers of event points have significant effects on the model size and computational performance. Attempts have been made to eliminate some unnecessary event points by relaxing some unconditional sequencing constraints enforced for the consumption and its related production tasks with the same state. Specifically, Seid and Majozi (2012) as well as Vooradi and Shaik (2013) relaxed the unconditional sequencing and alignments between the related production and consumption tasks if there is sufficient state inventory for consumption and storage space for production, respectively. However, the model of Seid and Majozi (2012) leads to real-time storage violation. Vooradi and Shaik (2013) introduced enormous variables, resulting in heavy computational burdens. More importantly, almost all existing models did not allow a batch of materials to be temporarily stored in processing units over multiple event points, leading to inefficient utilization of processing units.

In this work, a novel mixed-integer linear programming (MILP) model is proposed to address above mentioned limitations of the existing unit-specific event-based MILP formulations. Specifically, concepts on material transfer are introduced for the related production and consumption tasks at two adjacent event points. Material flows of states are monitored among units rather than specific tasks to decrease numbers of variables. And temporary holding of batches on units after production are modelled to yield better objective results for problems subject to finite intermediate storage (FIS).

## 2. Problem statement and material transfer

Multipurpose batch plants are formulated based on the state-task network [Kondili et al. (1993)]. Tasks  $i \in \mathbf{I}$  are processed on units  $j \in \mathbf{J}$ . Units  $j$  performing task  $i$  are included in the set  $\mathbf{J}_i$ . States  $s \in \mathbf{S}$  are consumed ( $\rho_{si} < 0$ ) or produced ( $\rho_{si} > 0$ ) by the consumption ( $i \in \mathbf{I}_s^C$ ) or production ( $i \in \mathbf{I}_s^P$ ) tasks with the proportion  $\rho_{si}$ . The formulation is proposed to minimize makespan to fulfill demands ( $D_s$ ) of products ( $s \in \mathbf{S}^P$ ) or maximize profit ( $z$ ) over a given horizon ( $H$ ). Material transfer between facilities is explicitly tracked to conditionally sequence the related production and consumption tasks of the same state. Four scenarios of material transfer (MT) are illustrated in Figure 1. In the first scenario (MT1), materials after production are transferred into its dedicated storage. In MT2, the batch of materials after production are first temporarily held on its processing unit and then flow into the storage. In MT3 and MT4, material after production at  $n$  are consumed by tasks at  $(n + 1)$ . MT3 illustrates the indirect material transfer in which materials available at  $n$  would be first transferred to the storage and then to the downstream unit at  $(n + 1)$  as there are sufficient storage space. In MT4, materials are directly transferred to the downstream units due to the storage limitation, which is termed as the direct material transfer. Two sets of binary variables are introduced to formulate the indirect ( $zI_{jj'n}$ ) and direct ( $zD_{jj'n}$ ) material transfer between units  $j \in \mathbf{J}_s^P$  and  $j' \in \mathbf{J}_s^C$  performing the production and consumption tasks, respectively, of a state  $s$ . Specifically,  $zI_{jj'n} = 1$ , if a state produced on  $j$  at  $(n - 1)$  is indirectly transferred to  $j'$  at  $n$ .  $zD_{jj'n} = 1$ , if a state produced on  $j$  at  $(n - 1)$  is directly transferred to  $j'$  at  $n$ .

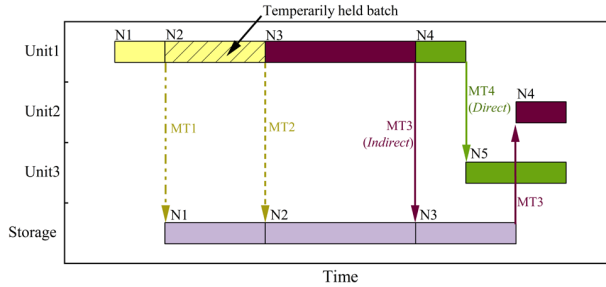


Figure 1. Four scenarios of material transfer between units and storage tanks

## 3. Mathematical model

The binary variable  $w_{ijn'n'}$  is one if task  $i$  is processed on  $j$  from event point  $n$  to  $n'$  ( $n \leq n' \leq n + \Delta n$ ).  $\Delta n$  is the number of event points over which a task can span. Binary variables  $ys_{ijn}$  equal to one if one or partial batch of a task  $i$  is temporarily held on  $j$  at  $n$ . Allocation constraints demand at most one task to be processed or held on  $j$  at a time. Sets  $\mathbf{I}_j$  and  $\mathbf{I}^P$  include tasks that can be processed on  $j$  and held on units, respectively.

$$\sum_{i \in \mathbf{I}_j} \sum_{n - \Delta n \leq n' \leq n} \sum_{n \leq n'' \leq n' + \Delta n} w_{ijn'n''} + \sum_{i \in (\mathbf{I}_j \cap \mathbf{I}^P)} ys_{ijn} \leq 1 \quad \forall j, n \quad (1)$$

Batch sizes  $b_{ijn'n'}$  are bounded by maximum ( $B_{ij}^{\max}$ ) and minimum ( $B_{ij}^{\min}$ ) unit capacity.

$$B_{ij}^{\min} \cdot w_{ijn'n'} \leq b_{ijn'n'} \leq B_{ij}^{\max} \cdot w_{ijn'n'} \quad \forall j, i \in \mathbf{I}_j, n \leq n' \leq n + \Delta n \quad (2)$$

Amounts of the batch of task  $i$  held on unit  $j$  at  $n$  are denoted by  $bs_{ijn}$ . Variables  $ST_{sn}$  express amounts of  $s$  stored at  $n$ , including units and its dedicated storage. Material balance of one state at  $n$  are formulated as Eqs. 3-4, where Eq. 3 is valid for the first event point. Variables  $ST0_s$  and  $bs0_{ij}$  express the initial inventory level of state  $s$  and initial amount for a batch of task  $i$  held in  $j$ , respectively. Amounts of states  $s$  stored at  $n$  should be always no less than its amounts temporarily held on the processing units (see Eq.5).

$$ST_{sn} = ST0_s + \sum_j \sum_{i \in (\mathbf{I}_j \cap \mathbf{I}_s^p)} \rho_{si} \cdot bs0_{ij} + \sum_j \sum_{i \in (\mathbf{I}_j \cap \mathbf{I}_s^c)} \sum_{n \leq n' \leq n + \Delta n} \rho_{si} \cdot b_{ijn'} \quad \forall s, n = 1 \quad (3)$$

$$ST_{sn} = ST_{s(n-1)} + \sum_j \sum_{i \in (\mathbf{I}_j \cap \mathbf{I}_s^p)} \sum_{n-1 - \Delta n \leq n' \leq n-1} \rho_{si} \cdot b_{ijn'(n-1)} + \sum_j \sum_{i \in (\mathbf{I}_j \cap \mathbf{I}_s^c)} \sum_{n \leq n' \leq n + \Delta n} \rho_{si} \cdot b_{ijn'} \quad \forall s, n > 1 \quad (4)$$

$$\sum_j \sum_{i \in (\mathbf{I}_j \cap \mathbf{I}_s^p)} \rho_{si} \cdot bs_{ijn} \leq ST_{sn} \quad \forall s \in \mathbf{S}^{FIS}, n \quad (5)$$

Start and finish time of a unit  $j$  at event point  $n$  are denoted by  $T_{jn}^s$  and  $T_{jn}^f$ , respectively. The unit  $j$  finishes the event point  $n$  after processing task  $i$  that is performed on  $j$  at  $n$ .

$$T_{jn}^f \geq T_{jn}^s + \sum_{i \in \mathbf{I}_j} \sum_{n \leq n' \leq n + \Delta n} (\alpha_{ij} \cdot w_{ijn'} + \beta_{ij} \cdot b_{ijn'}) \quad \forall j, n \quad (6)$$

Sequencing constraints require that on a unit  $j$ , an event point  $(n + 1)$  must start after the finish of its previous event point  $n$  (see Eq.7). Continuous variables  $T_{sjn}$  are the time for state  $s$  produced on unit  $j$  at  $n$  being available. Available time of the same state at the next event point  $(n + 1)$  should be larger than the available time at  $n$ , as Eq.8.

$$T_{j(n+1)}^s \geq T_{jn}^f \quad \forall j, n < N \quad (7)$$

$$T_{sjn} \leq T_{sj(n+1)} \quad \forall s \in \mathbf{S}^{in}, j \in \mathbf{J}_s^p, n < N - 1 \quad (8)$$

As formulated by Eqs. 9-10, the time when a state  $s$  is available at  $n$  should be after the finish of its production task on unit  $j$  and before the start of the next event point  $(n + 1)$ .

$$T_{sjn} \geq T_{jn}^f - M \cdot \left[ 1 - \sum_{i \in (\mathbf{I}_j \cap \mathbf{I}_s^p)} \sum_{n - \Delta n \leq n' \leq n} w_{ijn'n} \right] \quad \forall s \in \mathbf{S}^{in}, j \in \mathbf{J}_s^p, n < N \quad (9)$$

$$T_{sjn} \leq T_{j(n+1)}^s + M \cdot \left[ 1 - \sum_{i \in (\mathbf{I}_j \cap \mathbf{I}_s^p)} \left( \sum_{n - \Delta n \leq n' \leq n} w_{ijn'n} + y_{sijn} \right) \right] \quad \forall s \in \mathbf{S}^{in}, j \in \mathbf{J}_s^p, n < N \quad (10)$$

Indirect material transfer of a state takes place only with sufficient storage space for  $s$ . If it occurs from a unit  $j$  at  $n$  to  $j'$  at  $(n + 1)$ , the unit  $j'$  at  $(n + 1)$  must start not earlier than the finish of  $j$  at  $n$ . Set  $\mathbf{CJ}_j$  includes units  $j'$  consuming a state produced in unit  $j$ .

$$T_{jn}^f \leq T_{j'(n+1)}^s + M \cdot [1 - z_{I_j j'(n+1)}] \quad \forall j, j' \in \mathbf{CJ}_j, j \neq j', n < N \quad (11)$$

Continuous variables  $bT_{ijj'j'n}$  express amounts of a state through indirect material transfer from  $j$  to  $j'$ . For any consumption task, its consumed amounts of  $s$  cannot exceed the total amounts stored in the storage and transferred from the production tasks.

$$\sum_{j'} \sum_{i' \in (\mathbf{I}_{j'}^c \cap \mathbf{I}_j^p)} \left( -\rho_{si'} \cdot \sum_{n \leq n' \leq n + \Delta n} b_{i'j'nn'} \right) \leq ST_{s(n-1)} + \sum_j \sum_{j'} \sum_{i \in (\mathbf{I}_j^p \cap \mathbf{I}_{j'}^p)} \sum_{i' \in (\mathbf{I}_{j'}^c \cap \mathbf{I}_j^p)} bT_{ijj'j'n} \quad \forall s \in \mathbf{S}^{in}, n > 1 \quad (12)$$

$$\rho_{si} \cdot \sum_{n-1 - \Delta n \leq n' \leq n-1} b_{ijn'(n-1)} \geq \sum_{j'} \sum_{i' \in (\mathbf{I}_{j'} \cap \mathbf{I}_s^c)} bT_{ijj'j'n} \quad \forall s \in \mathbf{S}^{in}, j, i \in (\mathbf{I}_j \cap \mathbf{I}_s^p), n > 1 \quad (13)$$

$$-\rho_{si'} \cdot \sum_{n \leq n' \leq n + \Delta n} b_{i'j'nn'} \geq \sum_j \sum_{i \in (\mathbf{I}_j \cap \mathbf{I}_s^c)} bT_{ijj'j'n} \quad \forall s \in \mathbf{S}^{in}, j', i' \in (\mathbf{I}_{j'} \cap \mathbf{I}_s^c), n > 1 \quad (14)$$

Amounts of  $s$  through indirect material transfer at  $n$  should not exceed the production amount finishing at  $(n - 1)$  and also the consumption amounts starting at  $n$  (see Eqs. 13-14). Without indirect material transfer, the corresponding variables on transfer amounts are zero, where  $B_j^{max} = \max_{s, i \in (\mathbf{I}_j \cap \mathbf{I}_s^p)} [\rho_{si} \cdot B_{ij}^{max}]$  and  $B_{j'}^{max} = \max_{s, i' \in (\mathbf{I}_{j'} \cap \mathbf{I}_s^c)} [-\rho_{si'} \cdot B_{i'j'}^{max}]$ .

$$\sum_{i \in \mathbf{I}_j} \sum_{i' \in \mathbf{I}_{j'}} bT_{ij'i'n} \leq \min[B_j^{\max}, B_{j'}^{\max}] \cdot zI_{jj'n} \quad \forall j, j' \in \mathbf{CJ}, j \neq j', n > 1 \quad (15)$$

Constraint (16) enforces that a unit  $j'$  consuming a state at  $(n + 2)$  or higher event points should always start after the state is available on unit  $j$  at  $n$  to avoid storage violation.

$$T_{sjn} \leq T_{j'(n+2)}^s + M \cdot \left[ 1 - \sum_{i' \in (\mathbf{I}_{j'} \cap \mathbf{I}_s^c)} \sum_{n+2 \leq n' \leq n+2+\Delta n} w_{i'j'(n+2)n'} \right] \quad \forall s \in \mathbf{S}^m, j \neq j', j \in \mathbf{J}_s^p, j' \in \mathbf{J}_s^c, n < N - 1 \quad (16)$$

When there is no sufficient storage space to store the production amount of a state  $s$ , this state must be transferred to the downstream unit directly or temporarily held on its processing unit, as indicated by Eq.17. Consequently, the unit  $j'$  should finish the processing at  $n$  before the end of unit  $j$  at  $n$  to avoid overlapping on  $j'$ , if direct material transfer of one state takes place from  $j$  at  $n$  to  $j'$  at  $(n + 1)$ , as constrained by Eq.18.

$$\begin{aligned} \sum_j \sum_{i \in (\mathbf{I}_j \cap \mathbf{I}_s^p)} \left[ \rho_{si} \cdot \sum_{n-1-\Delta n \leq n' \leq n-1} b_{ijn'(n-1)} \right] + ST_{s(n-1)} \\ \leq ST_s^{\max} + \sum_j \sum_{i \in (\mathbf{I}_j \cap \mathbf{I}_s^p)} \rho_{si} \cdot bs_{ijn} \\ + \sum_j \sum_{j'} \sum_{i \in (\mathbf{I}_j \cap \mathbf{I}_s^p)} \sum_{i' \in (\mathbf{I}_{j'} \cap \mathbf{I}_s^c)} bTd_{ij'i'n} \quad \forall s \in \mathbf{S}^{FIS}, n > 1 \end{aligned} \quad (17)$$

$$T_{j'n}^f \leq T_{jn}^f + M \cdot [1 - zD_{jj'(n+1)}] \quad \forall j, j' \in \mathbf{CJ}_j^f, j \neq j', n < N \quad (18)$$

Similar to the indirect material transfer, amounts of a state  $s$  transferred at  $(n + 1)$  through direct material transfer should be bounded by the amount available (i.e. produced or temporarily held) at  $n$  and the consumption amount by tasks at  $(n + 1)$ . And the amounts of state transferred should be zero without direct material transfer.

$$\rho_{si} \cdot \left[ \sum_{n-1-\Delta n \leq n' \leq n-1} b_{ijn'(n-1)} + bs_{ij(n-1)} \right] \geq \sum_{j'} \sum_{i' \in (\mathbf{I}_{j'} \cap \mathbf{I}_s^c)} bTd_{ij'i'n} + \rho_{si} \cdot bs_{ijn} \quad \forall s \in \mathbf{S}^{FIS}, j, i \in (\mathbf{I}_j \cap \mathbf{I}_s^p), n > 1 \quad (19)$$

$$-\rho_{si'} \cdot \sum_{n \leq n' \leq n+\Delta n} b_{i'j'nn'} \geq \sum_j \sum_{i \in (\mathbf{I}_j^p \cap \mathbf{I}_j)} bTd_{ij'i'n} \quad \forall s \in \mathbf{S}^{FIS}, j', i' \in (\mathbf{I}_{j'} \cap \mathbf{I}_s^c), n > 1 \quad (20)$$

$$\sum_{i \in \mathbf{I}_j} \sum_{i' \in \mathbf{I}_{j'}} bTd_{ij'i'n} \leq \min[B_j^{\max}, B_{j'}^{\max}] \cdot zD_{jj'n} \quad \forall j, j' \in \mathbf{CJ}_j^f, j \neq j', n > 1 \quad (21)$$

Equation 22 ensures sufficient storage space for state  $s$  produced on  $j$  at  $n$  when  $s \in \mathbf{S}^{FIS}$  is subject to FIS, by enforcing  $j'$  to finish at  $(n - 1)$  before the available time of  $s$  at  $n$ .

$$T_{sjn} \geq T_{j'(n-1)}^f - M \cdot \left[ 1 - \sum_{i' \in (\mathbf{I}_{j'} \cap \mathbf{I}_s^c)} \sum_{n \leq n' \leq n+\Delta n} w_{i'j'nn'} \right] \quad \forall s \in \mathbf{S}^{FIS}, j \neq j', j \in \mathbf{J}_s^p, j' \in \mathbf{J}_s^c, 1 < n < N \quad (22)$$

$$ST_{sn} \leq ST_s^{\max} + \sum_{j \in \mathbf{J}_s} \left[ \max_{i \in (\mathbf{I}_j \cap \mathbf{I}_s^p)} (\rho_{si} \cdot B_{ij}^{\max}) \right] \quad \forall s \in \mathbf{S}^{FIS}, n \quad (23)$$

$$bs_{ijn} \leq B_{ij}^{\max} \cdot ys_{ijn} \quad \forall j, i \in (\mathbf{I}_j \cap \mathbf{I}^p), n \quad (24)$$

Amount of state  $s \in \mathbf{S}^{FIS}$  stored on units and storage at  $n$  should satisfy Eq. 23. Amount of the batch temporarily held on  $j$  should be no larger than the unit capacity. Tightening constraints (refer Eqs. 34, 35 and 37 of Li et al. 2022) are formulated to handle binary variables on allocation and material transfer. Indirect material transfer tasks place if there

is the direct material transfer between two units (Eq. 25). Bounds are added on variables as Eqs. 26-30. Both  $y_{sijn}$  and  $bs_{ijn}$  are zero if states produced by  $i$  are subject to UIS.

$$zI_{jj'n} \geq zD_{jj'n} \quad \forall j, j' \in \mathbf{CJ}_j^F, j \neq j', n > 1 \quad (25)$$

$$T_{jn}^s \leq H, T_{jn}^f \leq H \quad \forall j, n \quad (26,27)$$

$$w_{ijnn'} = 0, b_{ijnn'} = 0 \quad \forall i, j, n, (n' < n) \cup (i \notin \mathbf{I}_j) \quad (28,29)$$

$$bT_{ijj'j'n} = bT_{ijj'j'n} = zI_{jj'n} = zD_{jj'n} = y_{sijn} = bs_{ijn} = 0 \quad \forall i, i', j, j', n = 1, bs_{0ij} = 0 \quad (30)$$

$$y_{sijn} = 0, bs_{ijn} = 0 \quad \forall i \notin \mathbf{I}^P, j, n \quad (31,32)$$

The objective is to maximize profit (see Eq. 33) or minimize makespan (see Eq. 34-35). Equation (36) is the tightening constraint for minimization of makespan.

$$z = \sum_{s \in \mathbf{S}^P} \left[ p_s \cdot \sum_j \sum_{i \in (\mathbf{I}_j \cap \mathbf{I}_s^F)} \sum_n \sum_{n - \Delta n \leq n' \leq n} (\rho_{si} \cdot b_{ijn'n}) \right] \quad (33)$$

$$T_{jn}^f \leq MS \quad \forall j, n = N \quad (34)$$

$$STO_s + \sum_j \sum_{i \in (\mathbf{I}_j \cap \mathbf{I}_s^F)} \sum_n \sum_{n \leq n' \leq n + \Delta n} (\rho_{si} \cdot b_{ijnn'}) \geq D_s \quad \forall s \in \mathbf{S}^P \quad (35)$$

$$\sum_{i \in \mathbf{I}_j} \sum_n \sum_{n \leq n' \leq n + \Delta n} (\alpha_{ij} \cdot w_{ijnn'} + \beta_{ij} \cdot b_{ijnn'}) \leq MS \quad \forall j \quad (36)$$

Development of the mathematical formulation is completed, which contains of equations (1-33) to maximize profit and (1-32) and (34-36) to minimize makespan.

Table 1. Computational results for examples 1-6 with FIS. (Exs1, 3 and 4 are addressed to minimize makespan. Exs2, 5 and 6 are addressed to maximize profit)

Ex	Model	Events	RMILP	MILP	CPU (s)	Binary	Continuous	Constraints
1	V&S	3	5.00	11.5	0.02	12	32	68
	L&F	3	5.00	11.5	0.02	6	29	43
	SLK2	4	5.00	11.5	0.02	12	83	99
	M	3	8.00	<b>8</b>	0.02	12	37	76
2	V&S	10	1795.48	989.03	2096	569	805	2670
	M	11	2208.35	<b>1201.39<sup>d</sup></b>	3600	511	1801	3091
3	V&S	14	24.24	27.88 <sup>e</sup>	3600	226	371	1031
	M	14	24.24	27.88	<b>1969</b>	213	450	1243
4	V&S	21 <sup>a</sup>	47.38	47.68 <sup>f</sup>	3600	1,080	1567	4955
	M	<b>21</b>	47.38	47.68	<b>82.0</b>	<b>648</b>	2449	3791
5	V&S	6 <sup>b</sup>	400.00	400.00	0.16	132	197	596
	M	<b>6</b>	400.00	400.00	0.05	<b>79</b>	185	536
6	V&S	10 <sup>c</sup>	400.00	400.00	1.77	316	421	1300
	M	<b>10</b>	400.00	400.00	0.16	<b>139</b>	317	956

a:  $\Delta n = 2$ , b:  $\Delta n = 3$ , c:  $\Delta n = 7$ . Relative gap, d: 5.74%, e: 0.08%, f: 0.64%.

#### 4. Computational studies

Comparisons between the proposed model (M) and three existing continuous-time based models [V&S: Vooradi and Shaik (2013), L&F: Li and Floudas (2010), SLK2: Susarla et al. (2010)] are conducted. Considered examples 1 and 2 with FIS are the motivating example 2 and example 12 from Li et al. (2022). Examples 3-6 are examples 1a, 2b, 8 and 9 in their work. Models are solved using CPLEX 12.6.3/GAMS 24.6.1 on a desktop computer with AMD Ryzen 9 3900X 3.8 GHz, 48 GB RAM running Windows 10. Computational results of the examples 1-6 are listed in Table 1. Better objective results are obtained by M in examples 1 and 2 as the processing units can hold materials to share the storage burden under FIS. Computational time is reduced over 45% (1969 s vs. 3600 s) and 98% (82 s vs. 3600 s) by using M relative to model V&S for examples 3 and 4. That is attributed to the decreased number of binary variables. Concerning appropriate

numbers of event points to reach optimality for examples 4-6, it proves that model M does not require any task to span over multiple event points (i.e.  $\Delta n > 0$ ). Four instances (instances 9, 11, 14, and 15 in Vooradi and Shaik 2012) of the Kallrath example, are addressed using the developed model, as presented in Table 2. Results prove that the proposed model always generate lower makespan using less computational time. For instance, the model M obtains a 2.5% (39 h vs. 40 h) lower makespan using 45% (21890 s vs. 40000s) less computational time, compared with V&S.

Table 2. Computational results of the Kallrath examples with FIS to minimize makespan

In	Model	Events	RMILP	MILP	CPU (s)	Binary	Continuous	Constraints
1	V&S	10	16	32	10761	888	1386	5113
	M	10	16	32	<b>5377</b>	591	1267	4229
2	V&S	11	28	40 <sup>a</sup>	40000	984	1530	5664
	M	11	28	<b>39</b>	21890	654	1401	4690
3	V&S	9	21.6	36	447	792	1242	4562
	M	9	21.6	36	348	528	1133	3768
4	V&S	15	40	56 <sup>b</sup>	40000	1368	2106	7868
	M	15	40	<b>54<sup>c</sup></b>	40000	906	1937	6534

Relative gap, <sup>a</sup>: 7.5%, <sup>b</sup>: 14.29%, <sup>c</sup>: 3.70%.

## 5. Conclusion

In this work, a novel unit-specific event-based MILP formulation is proposed to address the short-term scheduling of the multipurpose batch plants. Material transfer of states are explicated to conditionally sequence the related tasks on units. And batches of tasks after production are permitted to be held in the processing units. Computational results demonstrate that the proposed model leads us to better objective results and decreases the computational time up to 98%. More interestingly, the proposed model does not require any task to span over multiple event points to yield the global optima.

## Acknowledgement

Dan Li and Taicheng Zheng appreciate financial support from China Scholarship Council - the University of Manchester Joint Scholarship (201908130170, 202106440020). Jie Li appreciates financial support from Engineering and Physical Sciences Research Council (EP/T03145X/1).

## References

- D. Li, N. Rakovitis, T. Zheng, Y. Pan, J. Li, G. Kopanos, 2022, Novel Multiple Time-grid Continuous-time Mathematical Formulation for Short-term Scheduling of Multipurpose Batch Plants. *Ind. Eng. Chem. Res.*, 61, 16093-16111.
- J. Li, C.A. Floudas, 2010, Optimal event point determination for short-term scheduling of multipurpose batch plants via unit-specific event-based continuous-time approaches, *Ind. Eng. Chem. Res.*, 49, 7446-7469.
- E. Kondili, C.C. Pantelides, R.W.H. Sargent, 1993, A general algorithm for short-term scheduling of batch operations-I MILP formulation. *Comput. Chem. Eng.*, 17, 211-227.
- N. Rakovitis, N. Zhang, J. Li, L.P. Zhang, 2019, A new approach for scheduling of multipurpose batch processes with unlimited intermediate storage policy, *Front. Chem. Sci. Eng.*, 13, 784-802.
- R. Seid, T. Majozi, 2012, A robust mathematical formulation for multipurpose batch plants. *Chem. Eng. Sci.*, 68, 36-53.
- N. Susarla, J. Li, I.A. Karimi, 2010. Novel approach to scheduling multipurpose batch plants using unit-slots. *AIChE J.*, 56, 1859-1879.
- R. Vooradi, M.A. Shaik, 2012, Improved three-index unit-specific event-based model for short-term scheduling of batch plants. *Comput. Chem. Eng.*, 43, 148-172.
- R. Vooradi, M.A. Shaik, 2013, Rigorous unit-specific event-based model for short term scheduling of batch plants using conditional sequencing and unit-wait times, *Ind. Eng. Chem. Res.*, 52, 12950-12792.

# Heat integration and heat exchanger network design for oxyfuel cement plants

Leif E. Andersson,<sup>a</sup> Avinash Subramanian,<sup>a</sup> Mari Voldsund,<sup>a</sup> Rahul Anantharaman,<sup>a</sup> Kristina Fleiger,<sup>b</sup> Francisco Carrasco,<sup>c</sup> Mirko Weber<sup>d</sup>

<sup>a</sup>*SINTEF Energy Research, Sem Sælands vei 11, 7043 Trondheim, Norway*

<sup>b</sup>*VDZ Technology gGmbH, Toulouser Allee 71, 41476 Düsseldorf, Germany*

<sup>c</sup>*HeidelbergCement AG, Berliner Straße 6, 69120 Heidelberg, Germany*

<sup>d</sup>*Holcim Technology Ltd, Grafenauweg 10, 6300 Zug, Switzerland*

*leif.andersson@sintef.no*

## Abstract

The cement sector needs to reduce its CO<sub>2</sub> emissions. An oxyfuel CO<sub>2</sub> capture technology allows to considerably reduce the emission. However, heat recovery and energy efficiency measures are essential to make the technology economically feasible. An approach to design heat exchanger networks applied to a 1<sup>st</sup> generation oxyfuel cement plant is described in this article. The approach consists of two steps: preliminary targeting and heat exchanger network design. For the studied cement plant, the steam Rankine cycle was identified to be superior to organic Rankine cycles. In the ideal case about 10.5 MW of power can be recovered. However, in a cost-efficient simple heat exchanger network recovery of only about 8.7 MW is economically reasonable.

**Keywords:** Heat integration, Heat exchanger network design, Oxyfuel cement plants.

## 1. Introduction

The cement sector is responsible for about 7% global anthropogenic CO<sub>2</sub> emissions (IEA, 2018). Two-third of the CO<sub>2</sub> emission originate from calcination of limestone while one-third come from the combustion of fuel. It is, therefore, impossible to reach CO<sub>2</sub> emission targets with fuel switching alone. CO<sub>2</sub> capture and storage is essential to become carbon neutral. An oxyfuel-based capture process is a promising candidate for capturing CO<sub>2</sub> from a cement plant (Voldsund et al., 2019). In the oxyfuel process the combustion is performed with oxygen mixed with recycled CO<sub>2</sub>. The CO<sub>2</sub> enriched flue gas allows a relatively cost-efficient purification and separation. Nevertheless, CO<sub>2</sub> purification consumes additional power and process heat must be transferred to air streams to dry the raw material energy-efficiently. The temperature levels in an oxyfuel plant are higher than in conventional plants because of the increased oxygen concentration in the combustion, and flue gas has to be cooled before it is recirculated. Waste heat recovery and heat integration are important for an economic implementation of oxyfuel technology in cement plants.

An existing cement plants was investigated for retrofit of the 1<sup>st</sup> generation oxyfuel process. The process simulations are performed with VDZs in-house cement process model. Afterwards, the heat integration is performed using a two-step methodology. This article focusses on a systematic approach to heat integration and applies it to an oxyfuel cement plant. In the following, a brief description of the model setup is given



followed by the introduction of the two-step methodology for heat integration and its application the oxyfuel cement plant is described.

## 2. Methodology

### 2.1. Modelling approach

The retrofit of the 1<sup>st</sup> generation oxyfuel process to cement plant was performed with several models. The clinker burning process was accessed by a kiln process model. The model described the process from the kiln meal feed to the outlet of the clinker from the cooler. It is made up by individual linked models of preheater, calciner, bypass, kiln and cooler, where material and energy balances are calculated (Koring, 2013). The process model outputs are performance data, thermal energy demand, clinker quality and available excess heat. Input data about the process design and some plant specific data to make the process model representative was provided by the cement producer and equipment supplier.

A second model, the heat integration model, access the waste heat recovery and the heat integration of the CO<sub>2</sub> Processing Unit (CPU). Data was iteratively exchanged between the two models.

### 2.2. Heat integration

Stream data from the process engineering model by VDZ was used to create the heat integration model in Aspen HYSYS. In addition, a CPU model was created in Aspen HYSYS. These models were used to access the energy streams of the 1<sup>st</sup> generation oxyfuel cement plant. In the following the two-step methodology consisting of a preliminary targeting step and a Heat Exchanger Network (HEN) design step, is presented.

#### 2.2.1. Preliminary targeting of the heat exchanger network design

The goal of the preliminary targeting is to identify the most promising heat to power cycles for the HEN design. A pinch analysis is performed which identifies the bottle neck of the plant regarding heat integration and allows to estimate an upper bound on the power production of the heat to power cycles. This step is also used to dimension the heat to power cycles. The stream data from this analysis is used in the HEN design phase.

#### 2.2.2. Heat Exchanger Network design

In the second step after the preliminary targeting the HEN is designed which allows cost-efficient design of the heat recovery in the cement plant. The best HEN involves optimizing the trade-off between capital costs determined by the number of heat exchanger units and their areas and the operating costs determined by the amount of hot and cold utilities required by the process. The software tool termed SeqHENS (Sequential Framework for HEN Synthesis) developed at the Norwegian University of Science and Technology and SINTEF Energy was used to design the optimal HEN (Anantharaman, 2011). An overview of the four-step methodology in the toolbox is given next:

- 1) Given stream data on the relevant hot and cold streams given by the preliminary targeting the minimum amount of hot and cold utility required is determined. The Linear Programming (LP) transshipment model is used to solve the optimization (Papoulias et al., 1983).
- 2) The absolute minimum number of heat exchanger units is determined where the heuristic is used that the optimal number is close to the minimum number of

units. This problem is solved with the Mixed-Integer Linear programming formulation (Papoulias et al., 1983).

- 3) Given the numbers of units and using engineering judgement the correct matches between hot and cold streams is determined. This “Steam Match Generator” problem is formulated using the vertical MILP transportation model (Anantharaman, 2011). The result is a Heat Load Distribution which gives the amount of heat exchanged between hot and cold stream.
- 4) Finally, the optimal topology of the HEN is determined. This problem is solved with a nonconvex Nonlinear Program (Floudas et al., 1986). The objective is to minimize the total cost of the HEN, where engineering judgment is used to get a simple HEN.

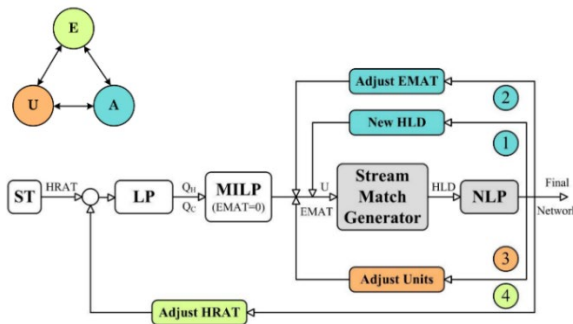


Figure 1. Overview of SeqHENS framework.

### 3. Case Study

The 1<sup>st</sup> generation oxyfuel cement plant consists of a preheater, pre-calciner, rotary kiln and cooler (Figure 2). The preheater exhaust is recirculated to the cooler, where a part of the recirculated stream is extracted and sent to the CPU. The Air Separation Unit (ASU) provides purified oxygen. An air stream is heated and used to dry the raw material.

The studied cement plant has an average annual clinker production of about 5600 t/d. The kiln line is a common dry process, and the raw material moisture is about 2-3%. The estimated drying demand for the raw material is about 300 kJ/kg<sub>clinker</sub>. This heat must be supplied by the excess heat of the process before additional excess heat can be used in a heat to power cycle.

The hot streams in the oxyfuel cement plant are the preheater exhaust stream at a temperature of about 450°C, cooler exhaust stream at a temperature of about 180°C, and the bypass stream, which is pinched with parts of the recycling stream, after which the combined stream has a temperature of 400°C. Additional heat streams are available from the CPU at a temperature of about 120°C. The cold streams are the heat required for drying the raw material, heat required in the CPU and heat for the heat to power cycle.

It is necessary that the recirculated stream is cooled down to about 50°C to remove chlorine, sulfur and water. Moreover, bag filters are used to remove dust. In addition, the air stream existing the raw mill should have required heat left for evaporate water. Therefore, a minimum temperature of 120°C was chosen for this stream.

#### 3.1. Preliminary targeting of 1<sup>st</sup> generation oxyfuel cement plant case study

The preliminary targeting phase was used to 1) find an optimal flue gas recirculation rate, 2) evaluate if heat from the Bypass should be used and 3) compare different heat to power cycles. The first and second step required iterations between the heat integration and

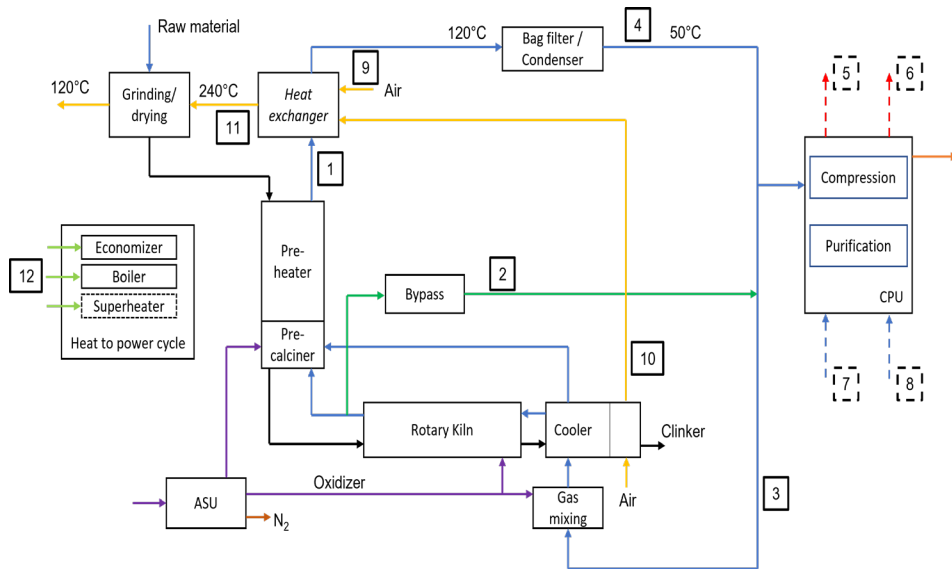


Figure 2. 1st generation oxyfuel cement plant. The numbers indicate hot and cold streams. Temperature constraints are displayed.

updating the process engineering model. It was concluded that no heat should be extracted from the Bypass since it increased the overall fuel use.

Organic Rankine cycles using Benzene or Butane and a steam Rankine cycle with different pressure levels were compared. The best performing Rankine cycle was the steam Rankine cycle at a pressure level of 15 bar, which was used for the HEN design (Table 1).

Table 1. Performance of different Rankine cycles.

Rankine cycle	Power production	Recirculation rate
Steam Rankine cycle (15 bar)	10.5 MW	0.49
ORC – Benzene	10.0 MW	0.49
ORC – Butane	8.4 MW	0.49

### 3.2. Heat exchanger network design of 1<sup>st</sup> generation oxyfuel cement plant case study

The preliminary targeting phase concluded that in the ideal case 10.4 MW power can be produced. However, this performance can only be reached with a complicated and expensive HEN. In the HEN design phase using SeqHENS stream splits and small heat exchangers which are not cost efficient were avoided. This resulted in a HEN which allows the required heat recovery and a power production in the steam Rankine cycle of about 8.7 MW (Figure 3). This is a 15% reduction compared to the ideal case.

In this study a retrofit is investigated, and the CPU will be located several hundred meters away from the kiln line. In the HEN obtained long pipelines are required between the CPU and the kiln line for only a small heat recovery. It was therefore decided to also investigate a case with the CPU excluded from the HEN design to create a simpler

network. The design of the simpler network (Figure 4) increases the overall heat exchanger area but reduces complexity and piping, which is not included in the objective function of SeqHENS. The same amount of power as before (8.7 MW) can be produced.

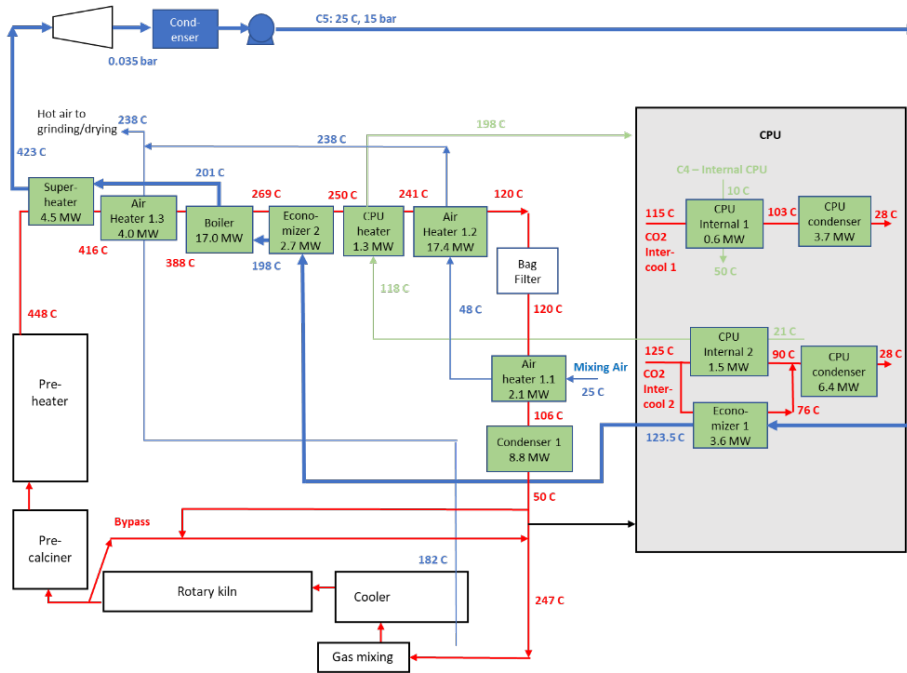


Figure 3. Heat exchanger network including heat integration of CPU.

In both HENs the Air Heater 1.2 is responsible for the majority of costs since it has a large duty, and it is a gas-gas heat exchanger with low heat transfer and large heat exchanger area.

#### 4. Conclusion

This article shows how to apply heat integration and heat exchanger network design to a 1<sup>st</sup> generation oxyfuel cement plant. Heat integration is essential for economically implementation of the oxyfuel technologies to existing cement plants. required heat exchanger is a gas-gas heat exchanger which also is the largest cost driver. This heat exchanger is essential since air must be heated for drying the raw material. The direct use of the preheater exhaust as in conventional cement plants is impossible since it would emit CO<sub>2</sub>. It is possible for this cement plant to recover about 8.7 MW of power which decreases the energy consumption of the plant. In future, a more detailed economic

assessment of the steam Rankine cycle and required heat exchangers must be performed to evaluate the profitability.

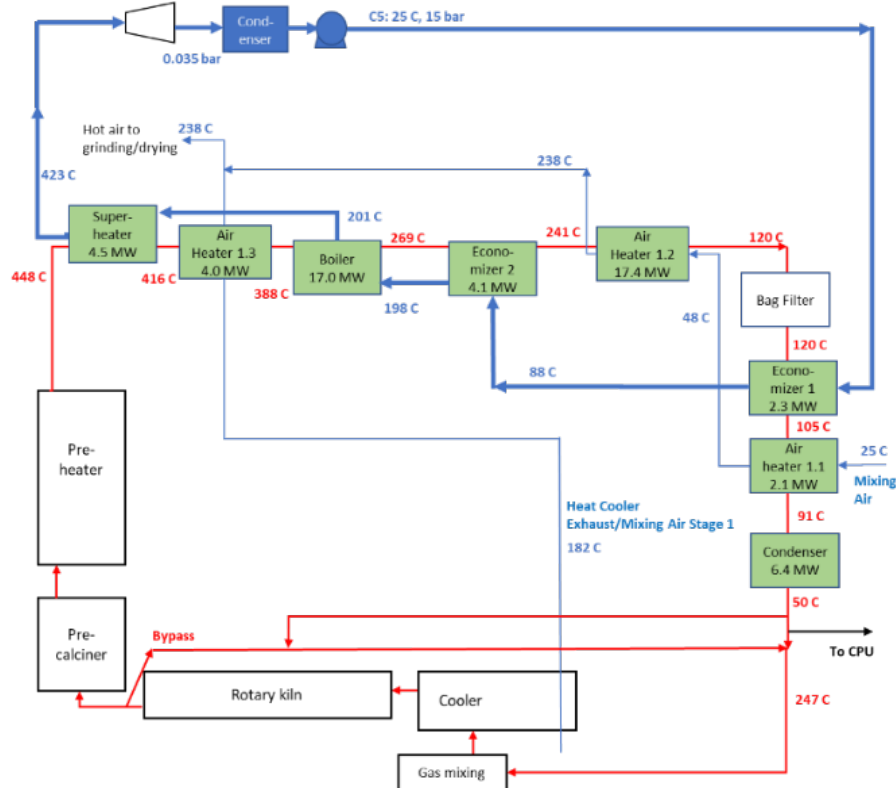


Figure 4. Heat exchanger network without heat integration of CPU.

## Acknowledgements

This research is executed in scope of the AC<sup>2</sup>OCem project which is funded through the ACT program (Accelerating CCS Technologies, Horizon2020 Project No 299663). Financial contributions are made from; The Research Council of Norway, (RCN), Norway, Federal Ministry for Economic Affairs and Energy (BMWi), Germany, Swiss Federal Office of Energy (SFOE), Switzerland, General Secretariat for Research and Development (GSRT), Greece, French Environment & Energy Management Agency (ADEME), France, and TotalEnergies SE.

## References

- International Energy Agency (IEA), 2018, Technology Roadmap: Low-Carbon Transition in the Cement Industry, IEA: Paris, France.
- M. Voldsund et al., 2019, Comparison of Technologies for CO<sub>2</sub> Capture Cement Production – Part 1: Technical Evaluation, *Energies*, 12, 559.
- K. Koring, 2013, Emissionsminderungspotential und technologische Auswirkungen der Oxyfuel- Technologie im Zementklinkerbrennprozess, PhD-Thesis, Verlag Bau + Technik GmbH, Düsseldorf, Germany.
- S. A. Papoulias et al., 1983, A structural optimization approach in process synthesis – II: Heat recovery networks, *Computer & Chemical Engineering*, 7(6), 707-721.
- R. Anantharaman, 2011, Energy Efficiency in Process Plants with emphasis on Heat Exchanger Networks, PhD-Thesis, NTNU, Department of Energy and Process Engineering, Trondheim, Norway.
- C. A. Floudas et al., 1986, Automatic synthesis of optimum heat exchanger network configurations, *AIChE Journal*, 32(2), 276-290.

# Technological Portfolio Optimisation for de-risking within the energy, water and food nexus

Maryam Haji, Mohammad Alherbawi, Sarah Namany, Tareq Al-Ansari

*College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar*

[\\*talansari@hbku.edu.qa](mailto:*talansari@hbku.edu.qa)

## Abstract

Food insecurity is one of the major issues that have loomed as a result of rapid population growth. In recent studies addressing modern-day resource issues, the energy, water and food (EWF) nexus approach was introduced as a holistic method to aid decision-makers in pre-empting unintended, and possibly damaging consequences resulting from an imbalance made in any part of the nexus system. Existing studies have focused on centralised approaches to study the interlinkages amongst EWF systems by focusing mostly on two elements of the EWF nexus. However, the current and future transitions should address the entire EWF system, holistically and sustainably. Besides, the existing indicator frameworks do not expressly capture the key interactions between the EWF resources to address the main challenges and risk factors associated with the EWF systems. Consequently, the aim of this study is to tackle this knowledge gap by integrating the analytical hierarchy process (AHP) tool along with EWF risk-associated optimisation as means to assess the changes in EWF resource availability, and therefore identify the tradeoffs of the large-scale implementation of decentralised energy and water technologies. The proposed framework utilises a nexus approach to optimise the energy and water technological portfolios in an efficient way; thus, supporting the decision-making for resource management in high-risk environments, such as the State of Qatar. The methodology consists of designing a composite risk indicator using the AHP method to determine the weight of different risk criteria for energy and water technologies. The obtained values are then used as an input for multi-objective optimisation that aims at minimising the tradeoffs between two objectives (i.e. risk and GWP). Results of the AHP indicate that wind energy is 11.8% riskier than solar energy, where the key issue is that the availability of these energy sources is far from constant but seasonal. In addition, treated wastewater was the most environmentally friendly, having the lowest impact on the ecosystem and resources, however, treated wastewater must be properly treated as it has the highest risk level (12.5%) and can lead to health issues. Finally, the multi-objective optimization model generates relatively 70 optimal solutions due to the mutual conflict between the two objectives.

**Keywords:** EWF nexus; food security; multi-objective optimisation; AHP.

## 1. Introduction

Energy, water and food systems, which are inherently interdependent, are essential for human health and survival, economic progress and resilience. These interrelated relationships were initially introduced at the Rio+20 Summit in 2012 as the “EWF” nexus. Developing a comprehensive decision-making framework necessitates the implementation of an effective EWF Nexus system that is adequate in tackling the multifaced issues within EWF systems to achieve resilient resource management. A

common approach to deal with risks and uncertainties is building a framework based on composite indicators (Becker, 2017). In the EWF nexus, the establishment of composite indicators is a common method for a quantitative assessment. Despite the availability of numerous indicators within the EWF nexus, developing a resilience indicator using the AHP technique is extremely beneficial; specifically, since they are a valuable tool that deals with interdisciplinary boundaries (Albrecht, 2018). In this context, Haji et al., (2020) used the AHP to design a risk assessment indicator for greenhouse planning in arid regions. While Zhang et al., (2021) developed an integrated approach to assess the composite risk of WEF nexus systems using a generalized copula-based chance-constrained programming model and generate risk-based plans. The proposed approach has been applied to an agricultural WEF nexus system in northern China, where the shortage of water, energy and land affected agricultural outputs. A recent study by Singh et al., (2022) established an assessment framework to rank individual countries whilst still allowing decision-makers with the possibility to allocate resources sufficiently and incorporate other appropriate indicators in countries that struggle in regard to food security, wherein the set of food security indicators was assessed based on the future climate change scenario. In this study, an assessment tool based on the EWF Nexus approach will be developed to support addressing dynamic decision-making, especially in hazardous non-resilience environments, in order to reduce the effect of overall risk, thus improving the national adaptability of EWF systems. This will be achieved through implementing an integrated approach based on the AHP method coupled with a multi-objective optimization model. A composite risk indicator was established using the AHP method to determine the weight of different risk criteria for energy and water technologies. The considered technologies consist of solar, wind and bioenergy for energy generation, while groundwater, treated wastewater and desalinated water as water supplies. The obtained values are then used as an input for multi-objective optimisation that aims at quantifying each EWF nexus system in terms of risk and environmental impact. The multi-objective linear program is formulated to minimise the tradeoffs between two objectives (i.e. risk and GWP).

## **2. Dynamic decision-making framework for EWF Nexus**

### *2.1. Development of Composite Risk Indicators using Analytic Hierarchy Process (AHP)*

The AHP method was used to design a composite risk indicator that will determine the weight of four risk criteria (cost, availability, readiness and applicability) for each energy and water technology that affect the overall goal, as illustrated in Figure 1. The technologies considered consist of solar, wind and bioenergy for energy generation, while groundwater, desalinated water and treated wastewater as water supplies. The assessment process begins with two-level pairwise comparisons between criteria and alternatives by identifying which of the two is riskier and by how much. In which the preference is expressed on a semantic scale of 1 to 9. The information which is related to the importance of each factor in comparison to one and another is obtained from previous literature. For the 2<sup>nd</sup> level, there are 21 pairwise comparisons between seven alternatives considering each risk factor. However, for the 1<sup>st</sup> level, there are 6 pairwise comparisons between four risk factors. Then, the normalized relative weight of each factor and alternative is obtained. Finally, the normalized principal eigenvector of risk factor and alternative was determined by averaging across each row, as demonstrated in Table 1.

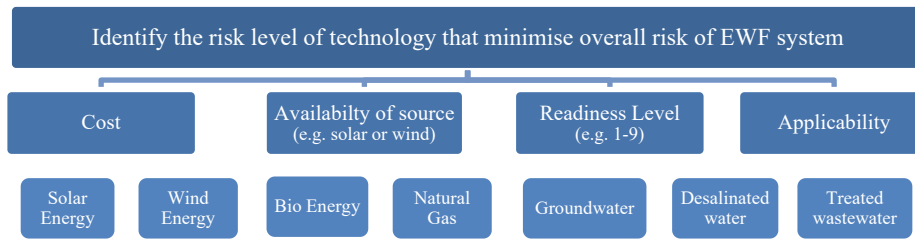


Figure 1: Hierarchy structure for determining the risk level of the energy and water technologies.

Table 1: Eigenvalues of four criteria and seven alternatives generated from the AHP method.

Risk Criteria	Eigen Vector of Criteria	Solar Energy (i=1)	Wind Energy (i=2)	Biomass Energy (i=3)	Natural Gas (i=4)	Groundwater (i=5)	Desalinated water (i=6)	Treated wastewater (i=7)
Cost	0.1067	0.1059	0.1237	0.0567	0.1269	0.1272	0.2133	0.2463
Availability	0.3576	0.1111	0.2390	0.1531	0.1261	0.1485	0.1111	0.1111
Readiness	0.3038	0.1250	0.2500	0.1250	0.1250	0.1250	0.1250	0.1250
Applicability	0.2320	0.1884	0.3259	0.1802	0.0540	0.0798	0.0798	0.0919

### 2.2. Multi-Objective Optimisation Model Development

The multi-objective optimisation model aims at minimising the risk and global warming potential (GWP) of different energy and water technologies required for food production. A summary of the mathematical model formulation is presented in Table 2. Two constraints were introduced to ensure that the sum of all contribution percentages of energy or water technologies must be 100%. The third constraint is to restrict groundwater to participate in the water mix when exceeding the annual renewable rate since groundwater consider as the scarcest water resource in the country studied. Moreover, the optimisation model assumed that the EWF nexus adopted does not involve the food sector explicitly in the technology selection, however, its contribution is translated through the water and energy required to produce food. The risk level of each technology generated from the AHP method will be used to run the minimisation of the risk objective function. However, as of GWP, the required data have been adapted from a study by Namany et al., (2019) which will be used to run the minimisation of the second objective function. A Solver was used to run and solve the optimization problem.

Table 2: Mathematical formulation for the optimisation model.

Objective Function:	
$\text{Min Risk} = \sum_{i=1}^4 R_i^e x_i^e + \sum_{i=5}^7 R_i^w x_i^w$	Identifies the optimal energy and water technologies technology mix that minimises the total risk and global warming potential of the EWF system delivering 40% of Qatar's demand for food.
$\text{Min Global Warming Potential} = \sum_{i=1}^4 GWP_i^e x_i^e$	Identifies the optimal energy technologies mix that minimises GWP
Subject to:	
$\sum_{i=1}^4 x_i^e \leq 100\%$	The sum of all contribution percentages of energy technologies must be 100%.



$\sum_{i=5}^7 x_i^w \leq 100\%$	The sum of all contribution percentages of water technologies must be 100%.
$g x_6^w \leq A_{max}$	Restricts the participation of groundwater in the water mix, being the scarcest water resource in the country studied.
$R_i^e, R_i^w, GWP_i^e, GWP_i^w, x_i^e, x_i^w \geq 0$	It implies that all decision variables must be strictly positive.

*Decision variables:*

$x_i^w$ : the percentage of contribution of each water source.

$x_i^e$ : the percentage of contribution of each energy source

*Parameters:*

$R_i^e$ : the risk level of using energy technology (i) where i=1,2,3 and 4

$R_i^w$ : the risk level of using water technology (i) where i=5,6 and 7

$GWP_i^e$ : the global warming potential of energy source. where i=1,2,3 and 4

$GWP_i^w$ : the global warming potential of the water source. where i=5,6 and 7

$g$ : the amount of groundwater utilised for food production with a 40% self-sufficiency level.

$A_{max}$ : the annual renewable rate of groundwater in Qatar is assumed to be 58 million m<sup>3</sup> per year.

### 3. Result and Discussion

The result from the AHP method (shown in Table 3) indicates that wind energy (25.0%) is the most riskier technology to be used for energy supply, where the key challenge is that the availability of this energy source is far from constant but seasonal. On the other hand, treated wastewater (12.5%) is slightly riskier than groundwater (12.3%) mainly due to the fact that it should be properly treated to avoid any health issues. The risk values obtained from the AHP method and Namany et al., (2019) study which represent the GWP of each technology are summarised in Table 3 and were used as an input to run the multi-objective optimisation model that aims at minimising the tradeoffs between the two objectives (i.e., risk and GWP). The optimisation model has been conducted to determine the optimal energy and water supply share using different technologies within the existing EWF system. The result from the solver was practicing a tradeoff between the two mutually conflicting objectives and generated relatively 70 optimal solutions, as demonstrated in Figures 2 and 3. Thus, the decision maker can pick any of the optimal solutions based on additional criteria, such as which solution achieved the lowest risk/GWP. For instance, Table 4 illustrates that out of 70 solutions, Solution 1 achieved the lowest risk and solution 69 achieved the lowest GWP. However, by averaging all 70 solutions, then Solution 34 can achieve the balanced minimisation between risk and GWP.

Table 3: Risk and environmental data used for the optimisation model.

Technologies	Solar	Wind	Biomass	Natural Gas	Groundwater	desalinated water	Treated wastewater
Risk %	13.3%	25.0%	14.1%	10.9%	12.3%	11.9%	12.5%
GWP (kg of CO <sub>2</sub> eq)	98,328	1,588	-	994,305,184	43,575,206	390,796,923	114,201,322

Table 4: The risk and GWP value of the three optimal solutions.

Solution #	Solar Energy	Wind Energy	Biomass Energy	Natural Gas	Groundwater	Desalinated water	Treated wastewater	Risk	GWP
1	33.3%	33.3%	33.3%	0.0%	4.74%	5.32%	89.84%	0.299	125,500,268
69	0.0%	0.00%	0.00%	99.99%	25.90%	26.08%	47.92%	0.232	1,162,161,486

34	22.9%	1.14%	26.65%	49.29%	8.87%	9.43%	81.68%	0.249	624,080,242
----	-------	-------	--------	--------	-------	-------	--------	-------	-------------

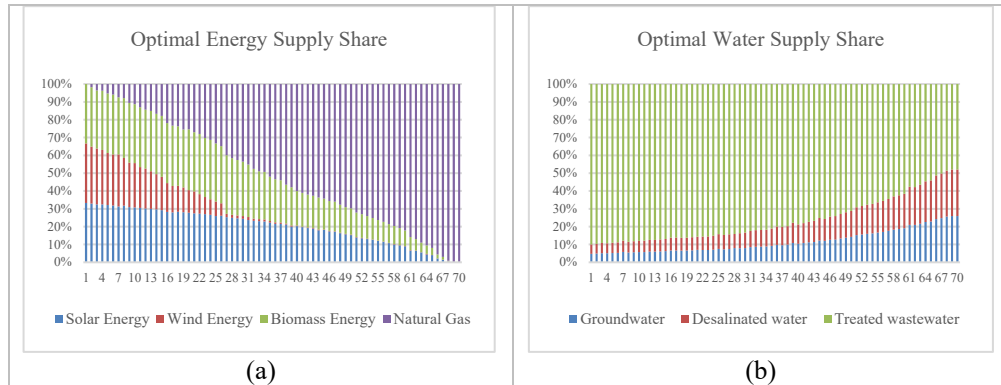


Figure 2: The 70 optimal solutions for energy and water supply share representing (a) energy technologies and (b) water technologies.

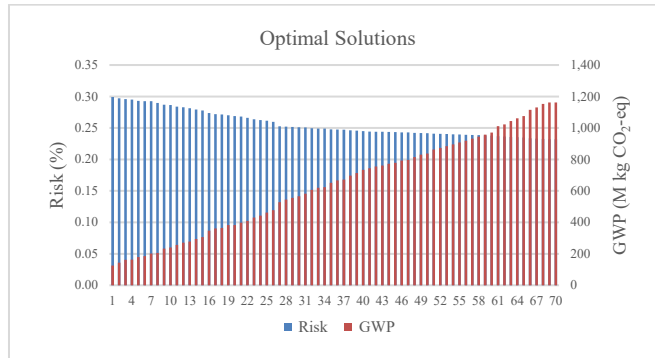


Figure 3: The tradeoff between risk and GWP for the 70 optimal solutions.

Moreover, due to the existence of mutually conflicting objectives, the problem solutions will not be a unique optimal but a set of non-dominated solutions referred to as the Pareto front due to the Pareto dominance concept (Pareto et al., 1896). Moreover, the solution will be Pareto-optimal if there is no other feasible solution that enhances one objective without weakening one another. In this study, the two objectives (risk % and GWP) of the non-dominated solutions set are demonstrated as Pareto Front in Figure 4 using excel.

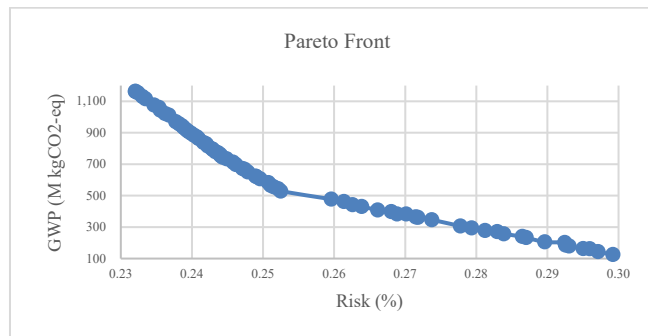


Figure 4: Schema of the Pareto front, for the minimization of two contradictory objectives risk and GWP.

#### 4. Conclusion

This study introduced an EWF Nexus assessment framework to define the technological share needed in optimally supplying energy and water that will minimise the risk and GWP of the overall EWF system using the AHP method and solver. The AHP method has considered financial and operational aspects (e.g., cost, availability, readiness and applicability) of energy and water technologies. The model aimed at providing optimal energy and water supply share from various technologies that minimised overall EWF system risk and GWP. The multi-objective optimization model generates relatively 70 optimal solutions due to the mutual confliction between the two objectives. However, the results from the AHP indicate that wind energy is 11.8% more riskier than solar energy, as the availability of these energy sources is seasonal. Furthermore, the treated wastewater was the most environmentally friendly, having the lowest impact on the ecosystem and resources, however, treated wastewater must be properly treated as it has the highest risk level (12.5%) and can lead to health issues. The assessment tool can support decision makers in understandings possible optimal ways of deploying various technologies to supply energy and water resources while minimizing overall risk and GWP of the EWF system.

#### Acknowledgments

This research was made possible by an Award (GSRA7-1-0407-20014) and supported by proposal number NPRP11S-0107-180216 from Qatar National Research Fund (a member of Qatar Foundation). The contents herein are solely the responsibility of the authors[s].

#### References

- Albrecht, T. C. (2018). The Water-Energy-Food Nexus: a systematic review of methods for nexus assessment. *Environ. Res. Lett.*, 4(13).
- Becker, W. S. (2017). Weights and importance in composite indicators: closing the gap. *Ecol. Indicate.*, (80), 12–22.
- Haji, M., Govindan, R., & Al-Ansari, T. (2020). Novel approaches for geospatial risk analytics in the energy–water–food nexus using an EWF nexus node. *Computers and Chemical Engineering*, 140, 106936. <https://doi.org/10.1016/j.compchemeng.2020.106936>
- Namany, S., Al-Ansari, T., & Govindan, R. (2019). Optimisation of the energy, water, and food nexus for food security scenarios. *Computers and Chemical Engineering*, 129. <https://doi.org/10.1016/j.compchemeng.2019.106513>
- Pareto, V., Politique, C. D., & Lausanne., F. R. *Volume I and II.*, (1896).
- Singh, R. K., Joshi, P. K., Sinha, V. S. P., & Kumar, M. (2022). Indicator based assessment of food security in SAARC nations under the influence of climate change scenarios. *Future Foods*, 5(January). <https://doi.org/10.1016/j.fufo.2022.100122>
- Zhang, T., Tan, Q., Wang, S., Zhang, T., Hu, K., & Zhang, S. (2021). Assessment and management of composite risk in irrigated agriculture under water-food-energy nexus and uncertainty. *Agricultural Water Management*, 262(August 2021), 107322. <https://doi.org/10.1016/j.agwat.2021.107322>

# Prediction of Melt Flow Rate (MFR) in polymer production, with auto-regressive Machine Learning algorithms, for multivariate time-series measured in irregular timestamps

Symon Doe<sup>a</sup>, Christoforos Kassianides<sup>a</sup>, Symeon Kassianides<sup>a</sup>, Christos Christodoulou<sup>b</sup>, Nikos Bakas<sup>b</sup>

<sup>a</sup>*Hyperion Systems Engineering Ltd., 36 Athalassis Street, Geri-Nicosia, 2201, Cyprus*

<sup>b</sup>*Computation-based Science and Technology Research Center, The Cyprus Institute, 20 Konstantinou Kavafi Street, Aglantzia-Nicosia, 2121, Cyprus*

## Abstract

The purpose of this work is to use Machine Learning algorithms to implement soft sensors to predict the material quality (MFR – Melt Flow Rate) at polymer production plants. The goal is to identify off-spec material in real-time without needing specialist online analyzer equipment. The Machine Learning algorithms use historical time series data of production plant measurements. These historical data comprise a wide variety of associated features, potentially driving the evolution of the material over time. The features are the time series values of specific physical parameters, such as temperature, pressure, etc., measured on or before the polymerization reactors. We investigated a variety of mathematical formulations for predicting MFR and Machine Learning Algorithms accordingly. We present the dataset, as well as machine learning models results along with performance metrics and future steps.

**Keywords:** machine learning, multivariate time series, autoregressive model, feature selection, polymer production.

## 1. Introduction

The aim of this work is to predict the material quality (ASTM-MFR) at polymer production plants, based on historical time series data. These historical data comprise a variety of associated features, potentially driving the evolution of the material over time. Accordingly, the given dataset is comprised of approximately three thousand signals measured in different timestamps among them, as well as with the target variable (MFR). The research started with an investigation of the statistical properties of the studied variables, followed by feature selection, and predictive modelling with Machine Learning Algorithms. In any predictive modelling computation, the predicted values always differ from the given ones in the raw dataset (Bakas, 2019). Hence, and despite the extended effort made to create a robust model, the predictions deviate from the actual values of MFR, especially in certain timestamps, due to the high amount of statistical noise existing within the variables. The input variables, exhibit distributions with no known patterns or even chaotic ones, and we present their computational handling, and numerical results.

We had to deal with specific challenges, such as the sparseness of MFR results (low frequency, measured every 4 hours on average), while the corresponding features utilized as an input in the Machine Learning Models are measured more frequently, e.g. per minute or even per second. Furthermore, we had to smooth the signals' outliers and deal with their dynamic nature, as the correlations of the features with the target variable, though strong, vary significantly with time and disorientate the predictive modelling. Furthermore, we present how we approached the cleaning of the data along with heuristic features' selection algorithms (Bakas, et al., 2022), so as to obtain a structured dataset to feed the Machine Learning Models. The amount of noise in the dataset was significant; hence we utilized a variety of autoregressive models and time lags (Bakas, 2019), in order to attain the best possible prediction accuracy. We found a combination of XGBoost (Chen & Guestrin, 2016; Xu & Chen, 2014), with the gradients of the features  $\nabla F_k$  at a time stamp  $k$ , to exhibit adequate accuracy. Furthermore, we investigate various ML models as new data are being obtained constantly. Particularly, we use Polynomial Regression (PR) (AlHamaydeh, et al., 2022), Random Forests (Breiman, 2001; Sadeghi, 2013), and Artificial Neural Networks (Bakas, et al., 2019).

To address the issue of noise and time-shift observations in the raw dataset, a hybrid approach was developed. This will use a simple model of the process equipment, which will precondition the data values by applying correct time offsets and filtering before sending it to the Machine Learning algorithms. This procedure removes the burden of processing noise and applying time offsets from the Machine Learning algorithms. As further development the model could identify different operating conditions, e.g., which catalyst is being used in the reactors, and/or change of raw material feeds to produce High/Low MFR product. This allows Machine Learning models to be trained for each condition with the expectation that each model will be more accurate within its defined range than a single combined model.

## 2. Specific Challenges

### 2.1. Sparseness of MFR results

We aim to predict MFR values which are measured every 4 hours on average (low frequency), while the corresponding features utilized as an input in the Machine Learning Models, are measured more frequently, e.g. per minute or even per second. We investigated a variety of models, initially we keep constant the previously known MFR value for the unknown regions, and ultimately, using only  $t$  timestamps with known values for all predictors and target.

### 2.2. Signal Outliers

The given dataset comprises a rich variety of potential features for the prediction of MFR. Twenty-one features exhibit high Pearson's Correlation with MFR ( $>0.9$ ). However, these correlations change with time strongly, along with the corresponding outliers in the features. Signal outliers disorientate the predictions made by any of the utilized models. To confront this, we used robust-to-noise machine learning models and also tailor-made algorithms for feature selection, to obtain a final set of features which predicts MFR with the best possible accuracy.

## 3. Computational Approach

Exploratory Data Analysis was performed for the fundamental statistical properties, as well as inter-item correlation and clustering among the time-series. Predictive modelling has started using various time-lags and linear models to identify prediction accuracy with the given data. The predictive modelling of the database was implemented by utilizing

four ML methods, and in particular: a. linear, b. higher-order regression, c. random forests, and d. gradient boosting. Machine learning methods exhibit diverse performance on the studied dataset, with respect to the error metrics each time utilized. The utilized error metrics were the Root Mean Squared Error, the Mean Absolute, the Mean Absolute Percentage Error, the Maximum Absolute Percentage Error, the slope of the Predicted vs Actual values, as well as the Pearson Correlation Coefficient. The studied data set contains a rich variety of input variables. However, many of them comprise extreme outliers. In order to identify an optimal subset of features, we use optimization algorithms (Bakas, et al., 2022; Plevris, et al., 2021).

### 3.1. Mathematical Formulation

Let  $Q$  be the sought Quality of MFR, which is intended to be predicted. We use symbol  $k$  as the index for the timestamps of  $Q$ .  $k$ , may iterate  $\in \{1, 2, \dots, m\}$  values of all time stamps where the features are being measured, or iterate  $\in \{1, 2, \dots, n\}$  known values of  $Q$ , with  $n \ll m$ . Hence at each time stamp  $k$  we have a particular value  $Q_k$ , while the entire vector of the target variable is

$$Q := Q_k \quad (1)$$

Accordingly, we define  $F_{k,j}$ , with  $j \in \{1, 2, \dots, p\}$ , for  $p$  features in total, the value of an input variable (feature) at time  $k$ , such that the vector of features at a time stamp  $k$  is

$$F_k := \{F_{k,1}, F_{k,2}, \dots, F_{k,p}\}, \quad (2)$$

and hence the entire matrix of features is

$$F := F_{k,j} \quad (3)$$

Let  $\mathcal{M}: R^p \rightarrow R$  be an continuous function of the features  $F_k$  returning the Quality values  $Q_k$ . We approximate  $\mathcal{M}$  with Machine Learning Algorithms, such that

$$Q_k = \mathcal{M}(F_k) + \epsilon_k, \quad (4)$$

and

$$\epsilon_k \sim \mathcal{U}(0, \sigma), \quad (5)$$

the regression errors, which should ideally follow a distribution  $\mathcal{U}$  exhibiting a Gaussian shape with zero mean.

This is the baseline approach. We may enhance this model, by utilizing the moving average of each feature  $j$ , select a subset  $o$  of the  $p$  features, add the time-lagged features as new features, and also add the differences of  $F_{k,j}$  along with the time passed until a difference occurs, as per the following formulation.

We define the differences of  $Q$ , at time  $k$  as

$$\Delta Q_k := Q_k - Q_{k-1}, \quad (6)$$

and the corresponding differences of  $F_{k,j}$ ,

$$\nabla F_k := \{\Delta F_{k,1}, \Delta F_{k,2}, \dots, \Delta F_{k,p}\}. \quad (7)$$

Let  $\mathcal{L}: R^{2p+1} \rightarrow R$  be an continuous function of the features  $F_k$ , such that

$$\Delta Q_k = \mathcal{L}(F_k, \nabla F_k, \Delta t_k) + \epsilon_k. \quad (8)$$

We now identify  $\mathcal{L}$  with Machine Learning, taking into consideration the distribution of the errors  $\epsilon_k$  as above-mentioned.

Afterwards, we may predict  $Q_k$ , by using

$$Q_k = Q_{k-1} + \Delta Q_k. \quad (9)$$

For the real-time, prediction stage, we iterate for each minute (or appropriate real-time interval), and sequentially add to the previous  $Q_{k-1}$ , the computed  $\Delta Q_k$ , as predicted, with  $\Delta t_k = 1$ , and  $k \in \{1, 2, \dots, m\}$ .

### 3.2. Latest data and Model

We retrain the models, after the data engineering as was automated by the HYPPOS program, enriched by new data, by using all raw features at MFR measurements timestamps as an input for the ML models, and target variable MFR. Particularly, this version is for ALL MFR values and both Catalysts. For Nonlinear Regression (NLR), XGBoost (XGB), and Random Forests (RF), we use exhaustive grid search with cross validation. The search parameters for tuning are presented in Table 1. For example, for XGBoost model, we run 36 combinations of Hyper-Parameters, times 100 intermediate rounds = 3600 different XGBoost models in total, and we select the best one. Furthermore, for each model, we run it 5 times in order to do cross-validation with a random permutation of the Train set. Hence the actual models are 18000.

*Table 1. Hyperparameter Tuning Search Space.*

<b>Nonlinear Regression</b>	
number of NLR folds	100
Percentage of NLR Cross Validation Samples per fold	0.8
Maximum Number of NLR Rounds	10
Maximum Number of NLR Features	$\infty$
Polynomial Degree	2
<b>XGBoost</b>	
number of XGB folds	5
Percentage of XGB Cross Validation Samples per fold	0.8
Minimum Number of XGB Rounds	10
Maximum Number of XGB Rounds	1000
Eta	0.05, 0.2, 0.5
Percentage of Combinations to Check	1
Number of Best Models to keep	10
Depth	1, 7, 15
Number of Intermediate Rounds	100
colsample bytree	0.5, 1
Subsample	0.5, 1
<b>Random Forests</b>	
number of RF folds	5
Percentage of RF Cross Validation Samples per fold	0.8
Percentage of RF Combinations to Check	1
Number of Best Models to keep	10
n trees	10, 100, 1000
n subfeatures	0.25, 0.5, 0.75
partial sampling	0.25, 0.5, 0.75
max depth	1, 10, 100
min samples leaf	1, 10, 100

## 4. Results

The best performing model for the new data after data engineering, was a combination of parameters of the Random Forests (Table 2). In Figure 1, we present the error analysis in the Test Set for Random Forests model. We may see that the errors are low on average, with some specific regions of high outliers (Figure 1).

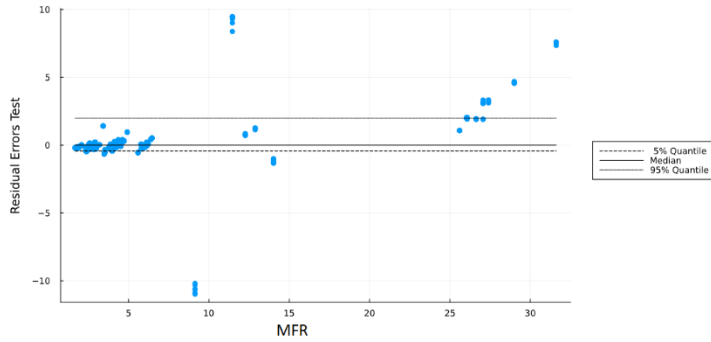


Figure 1: Prediction Errors in the Test Set, using the Random Forests model.

In order to explain statistically the drives of outliers in Figure 1, we performed an error analysis by following the next steps: (a) we split the Train Set into 1000 bins, equally distributed from the minimum value of MFR, up to its maximum value, (b) we counted the number of MFR values found per each particular bin, (c) for each bin, we computed the Mean Absolute Error (MAE), among the given MFR values, and the predicted ones, in the Test Set, (d) we plotted the results in a scatter plot (Figure 2), along with the 95% & 99% Quantiles of the MAE.

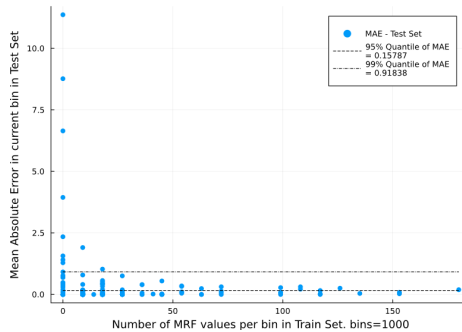


Figure 2: Number of MFR values per bin in the Train Set vs MAE in the Test Set.

Interestingly, we observe an inverse association among the number of MFR values per bin in the Train Set, and the corresponding MAE in the Test Set. Hence, we empirically deduce that the density of MFR values per MFR range corresponds to low errors and vice versa. Hence, e.g. for high values of MFR (>25) the high errors occur as a result of low number of MFR values in the training dataset. We foresee in the future steps of this work to gather more data and train new models, to resolve the data adequacy issue.

Table 2: Performance Metrics for all Models Used

Model	Dataset	Pearson	M.A.P.E.	M.A.E.	R.M.S.E
Linear	Train	0.988	0.165	0.791	1.176
Linear	Test	0.946	0.310	1.246	2.120
Polynomial	Train	0.992	0.127	0.650	0.998
Polynomial	Test	0.967	0.234	0.907	1.598
XGBoost	Train	1.000	0.000	0.001	0.002
XGBoost	Test	0.983	0.075	0.585	1.462
ANNBN	Train	0.999	0.054	0.260	0.350
ANNBN	Test	0.418	0.277	1.838	12.939
Random Forests	Train	1.000	0.005	0.032	0.080
Random Forests	Test	0.970	0.071	0.548	1.555



## 5. Conclusions

In this work, we developed machine learning models for the MFR time series prediction on real industrial conditions. The input for the models is a wide variety of industrial measurements, as obtained in the timestamps MFR was measured. We investigated various mathematical formulations, along with machine learning models. We found that hyperparameter tuning with exhaustive search, and tree-base models (XGBoost and Random Forests) yielded the best possible performance. The average errors are low for entirely out-of-sample data points. However, as outliers also exist in specific MFR timestamps, we foresee retraining the models with new data that will be obtained gradually in real-time industrial conditions. Despite the vast noise in data, irregular timestamps of MFR measurements and limited dataset, we developed a model exhibiting adequate accuracy on average.

## Acknowledgements

This work received financial support by the EuroHPC-JU project EuroCC (G.A. 951732) of the European Commission, and the Research and Innovation Foundation of Cyprus programs for research, technological development and innovation “RESTART 2016–2020” (RIF PROPOSAL NUMBER: ENTERPRISES/0521/0175). Parts of the runs were performed on the Cyclone machine hosted at the HPC National Competence Center of Cyprus at the Cyprus Institute (project pro21b103).

## References

- M. AlHamaydeh, G. Markou, N. Bakas, M. Papadrakakis, 2022, AI-based shear capacity of FRP-reinforced concrete deep beams without stirrups, *Engineering Structures*, Volume 264.
- N. Bakas, A. Langousis, M. Nicolaou, S. Chatzichristofis, 2019, A Gradient Free Neural Network Framework Based on Universal Approximation Theorem, s.l.:<https://arxiv.org/abs/1909.13563>.
- N. Bakas, 2019, Numerical Solution for the Extrapolation Problem of Analytic Functions, *Research - Science Partner Journal*, Volume 2019.
- N. Bakas, V. Plevris, A. Langousis, S. Chatzichristofis, 2022, ItsO: A novel inverse transform sampling-based optimization algorithm for stochastic search, *Stochastic Environmental Research and Risk Assessment*, Volume 36, p. 67–76.
- L. Breiman, 2001, Random forests, *Machine Learning*, Volume 45, p. 5–32.
- T. Chen, C. Guestrin, 2016, XGBoost: A Scalable Tree Boosting System. s.l.:*Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.
- V. Plevris, N. Bakas, S. German, 2021, Pure Random Orthogonal Search (PROS): A Plain and Elegant Parameterless Algorithm for Global Optimization, *Applied Sciences*, Volume 11, p. 5053-5081.
- B. Sadeghi, "DecisionTree.jl," 2013. [Online]. Available: <https://github.com/JuliaAI/DecisionTree.jl>.
- B. Xu and T. Chen, "XGBoost.jl," 2014. [Online]. Available: <https://github.com/dmlc/XGBoost.jl>.

# Achieving Operational Excellence by Combining Material Tracking and On-line Polymer Analysis and Classification data in an all-in-one Integrated Decision Support System

Symeon Kassianides<sup>a</sup>, Symon Doe<sup>a</sup>, Sanja Micunovica<sup>a</sup>, Artemis Theodoropoulou<sup>a</sup>, Charis Komodromos<sup>a</sup>, Sofia Louloudi<sup>a</sup>

<sup>a</sup>*Hyperion Systems Engineering Ltd., 36 Athalassis Street, Geri-Nicosia, 2201, Cyprus*

## Abstract

The HYPPOS software has been developed to encapsulate experience gained over 15 years of building bespoke solutions for the polymer industry. HYPPOS not only is configurable and easily maintainable, but it also uniquely provides the ability to combine data analytics and Artificial Intelligence to support decision making.

We developed a mathematical material tracking system that discretizes continuous manufacturing processes into identifiable slices of material (“Quanta”) and track them as they move through the manufacturing stages. To achieve this, HYPPOS is integrated with production and business management ICT systems. HYPPOS transforms data collected into information and applies different inferencing and Machine Learning algorithms, which calculate critical quality parameters, then tracking them in real-time. As the Quanta move past sensors, they collect KPI’s and quality parameters making the history of the material visible.

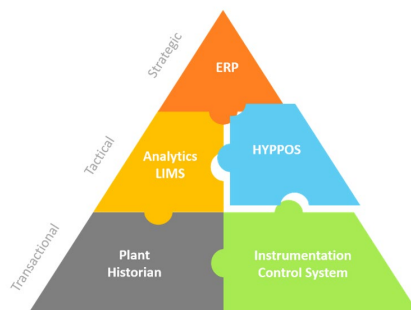
**Keywords:** petrochemicals, sustainability, real-time monitoring, waste reduction, decision support

## 1. Introduction

Today polymer production and forming processes account for more than 80% of the carbon footprint of plastic manufacturing (Hydrocarbon Processing, 2018). That presents detrimental environmental challenges that need immediate attention and intervention. The International Energy Agency ‘s (IEA) “Future of Petrochemicals” report (IEA, 2018) recommends building a more sustainable and efficient petrochemicals industry. Early identification of potential process issues and their correct handling is critical to reducing material waste and energy consumption. This can be achieved through specialized software technology that enables automatic and real-time analysis. However, current software platforms do not offer the required specialization, flexibility, and information without extensive customization.

Identifying the gap in the currently available solutions, we developed HYPPOS - Hyperion Predictive Production Online Software. HYPPOS as presented at the 11<sup>th</sup> International Conference on Chemistry and Chemical Engineering (ICCCE, 2020) and at the 9<sup>th</sup> International Conference on Chemical and Process Engineering Materials (ICCPE, 2020) is an all in-one Artificial Intelligence driven decision support tool that

addresses specific challenges of the polymers industry including quality traceability, real-time analysis, rapid and accurate reporting, process visualization and quality consistency. HYPPOS is positioned in Industry 4.0 Tactical Management Level, it collects data from systems at the Transactional Management Level, feeds and receives data from systems at the Strategical Management Level (e.g. Enterprise Resource Planning Systems such as SAP) and co-exists and interacts with other systems at the Tactical Management Level (e.g., Laboratory Information Management Systems).



*Figure 1: HYPPOS Positioning as per ISA 95 Industry Standard.*

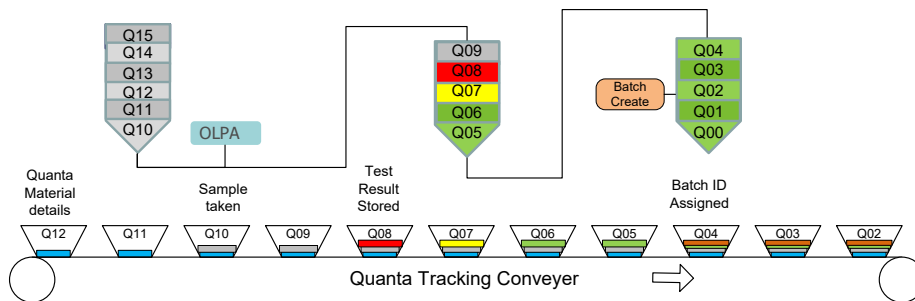
Polymer manufacturing requires quality monitoring at the different production stages from the polymerization reactors through to extrusion and batch assembly in blender silos. For this purpose, specialized On-Line Polymer Analysis equipment (OLPA) is available. Nevertheless, the majority of polymer plants are not equipped with such on-line analyzers primarily due to their high cost. Instead, process engineers and plant operators usually rely on off-line laboratory measurements of quality critical parameters. This does not allow real-time monitoring and early corrective actions, and as a result it leads to off-spec polymer product. Therefore, the aim of this work is to embed Machine Learning (ML) algorithms into HYPPOS software to create Soft Sensors to predict quality in real-time at early production stages eliminating the need for on-line physical sensors.

The Melt Flow Rate (MFR) is one of the most frequently used parameters in polymer manufacturing to control the material quality, and it is a measure of the viscosity of the polymer. Therefore, the MFR value range constitutes the most fundamental quality specification of the final polymer product since it determines the product grade.

## 2. Method

HYPPOS combines material flow models with AI/ML to create a digital twin of the production equipment and process. Its monitoring and inferential capabilities allow the online tracking of product quality as it moves through the production process and can forecast the final quality of polymer being produced enabling early corrective real-time actions. It uses a proprietary new method that allows the discretization of continuous processes and formulations of material into small ‘Quanta’ (depicted as  $Q_i$  in Figure 2), which act as a discreet container of trackable information enabling the traceability of material and its properties throughout the entire production process. The integrated tracking system is based on a unique, in-house developed process flow model which moves the Quanta through the equipment on a FIFO basis. Using an analogy, the ‘Quanta’

are located on a conveyor belt that helps them move through the process equipment advancing at the same speed as the production stream.



*Figure 2: Analogy for Quanta Tracking System.*

The model covers the main production stream from process feeds through to final product. The model is block based, with Equipment Blocks, Movements and Modifier blocks which merge/separate streams. The Movements transfer are driven by plant measurements, and transfer Quanta between Equipment using this information. The exact transfer rate calculation depends on the availability of plant measurements. Where instrumentation is missing, special proprietary algorithms calculate the movement rates based on measurements in surrounding equipment. The Equipment Block maintains details of the material holdup within the process equipment. Additionally analytical processes have been developed to process the information collected by the model and produce detailed reports for the final product.

When real time OLPA instrumentation is available, the Quanta can collect and carry this information and can be used to display the product quality properties as it advances through the production process.

Efforts to predict the polymer quality through Soft Sensors that use ML models based on plant measurements have been reported before (Shi, et al., 2006; Jumari & Mohd-Yusof, 2016; Liu, et al., 2018; Abeykoon, 2018; Yan, et al., 2023;). The innovation of this work lies in embedding the ML models in the HYPPOS Quanta algorithm that is implemented at industrial scale and it allows the real-time tracking of material together with its quality characteristics as it moves through the production equipment.

We have identified the following challenges when implementing ML algorithms to a continuous process such as polymer production. The process measurements used as input in the ML model are recorded continuously in real-time, whereas the target variable (MFR) is measured on average every 4 hours. Since MFR measurements come off-line from the lab, they become available with a delay of around 2 hours. Therefore, they need to be time shifted to various points in the process.

Furthermore, the plants are using relatively infrequent Lab test results for the quality parameter measurements, and this makes it difficult to obtain useful quantities of data to train the ML system to predict these parameters.

While initial tests using information from the plant were promising, they were not working well in all situations, especially during process transients. To simplify the ML algorithms task, we developed a pre-conditioning algorithm which is based on a simple model description of the process stages. The model defines the measurement signals, and the time duration of each stage. This allows the pre-conditioner to correctly time shift the measurements and apply noise reduction techniques to the values.

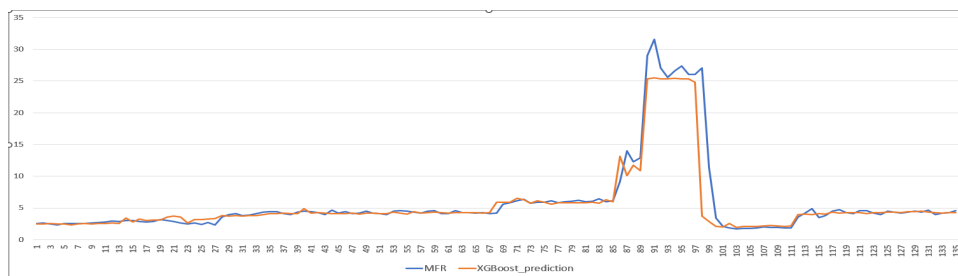
### 3. Case Study Results and Discussion

The objective was to predict the MFR on the polymerization reactor of an industrial scale polypropylene plant. To train the ML system, 5 months of raw process data from the plant including Lab measurements were used. Features selection algorithm was implemented, and 47 process variables were found to correlate the best with MFR measurements including among others the hydrogen feed rate to the polymerization reactors, doner concentration, teal concentration, the catalyst type, the reactor pressure and reactor temperature, and input material flow rates.

The raw data was run through a specially developed tool that converted the raw data into a training data set. The training data set contained the target variable values (MFR lab measurements) and all the pre-conditioned values that affected the production at that instant.

Using the trained ML algorithms, predictions were made using pre-conditioned process test data. This resulted in a marked improvement in prediction accuracy and better handling of process transients.

Figure 3 below shows XGBoost prediction of MFR parameter without using pre-conditioned data. It clearly shows that the predictions anticipate changes in MFR, therefore the Orange prediction line generally changes before the Blue measured line.



*Figure 3: MFR Predictions without Preconditioning.*

Figure 4 shows XGBoost prediction of MFR parameter with pre-conditioning. This shows great improvement in time shift and overall accuracy.

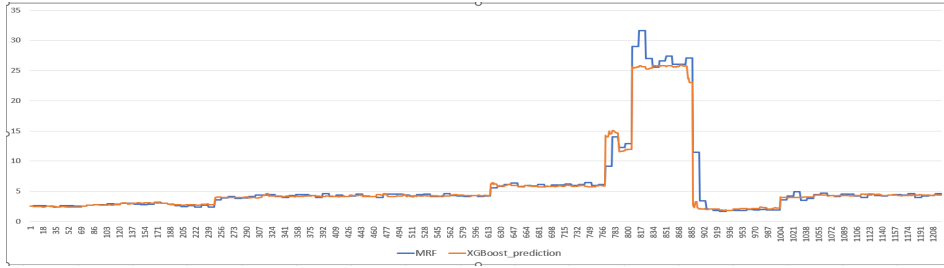


Figure 4: MFR Predictions with Preconditioning.

It is evident from the above graphs that there is a significant improvement in terms of tracking changes in operating conditions as now the Orange line is not time displaced from the Blue as in previous graph. There is also an improvement in prediction accuracy which is shown by the performance metrics calculated for each case as shown in the Table below.

Table 1. Performance Metrics for Methodologies Used.

	Pearson Correlation	Mean Absolute Percentage Error
No Preconditioning	0.92846	0.09523
With Preconditioning	0.98381	0.06697
Improvement	5.96%	-29.67%

The pre-conditioning technology has now been developed into a module within the HYPPOS software. This uses the process data values collected by HYPPOS in real-time, and at specified intervals it produces a table of pre-conditioned values. This table is passed to the ML prediction algorithm, which predicts the MFR, and the result is returned and passed to HYPPOS for use as part of the product quality management functionality.

It should be noted that these results were achieved with a relatively small quantity of training data, and that the installed system is expected to improve further on accuracy as it retrains itself based on continuous data collection.

#### 4. Conclusions

Based on this, HYPPOS is expected to have a significant positive impact on production efficiency, profitability, and competitiveness while also reducing waste, with a consequential contribution in lowering the carbon footprint.

The implementation of ML pre-conditioning has been a significant step forward in improving accuracy of the quality parameters. And it demonstrates that it can produce useful results with relatively little training data, allowing rapid deployment on new plants. This technology gives HYPPOS users a real-time view of the quality of the product being produced, enabling rapid intervention and reduction in waste and off-spec production.

## Acknowledgement

This work received financial support by the Horizon 2020 – 2nd Opportunity program of the European Commission, and the Research and Innovation Foundation of Cyprus programs for research, technological development, and innovation “RESTART 2016–2020” (RIF PROPOSAL NUMBER: OPPORTUNITY/0916/SME-II/0005).

## References

- <https://www.hydrocarbonprocessing.com/news/2018/10/petrochemicals-set-to-be-the-largest-driver-of-world-oil-demand>
- <https://www.iea.org/news/petrochemicals-set-to-be-the-largest-driver-of-world-oil-demand-latest-iea-analysis-finds>
- S. Louloudi, 2020, Software for Real-time Tracking, Analysis & Visualization of Polymer Data in Production Plants, Presentation at the 9th International Conference on Chemical and Process Engineering Materials (ICCPE 2020), 20-22 May, <http://www.iccpe.org/ICIEM&ICCPE%20Conference%20Program-2020.pdf>
- S. Louloudi, 2020, Software for Real-time Tracking, Analysis & Visualization of Polymer Data in Production Plants, Presentation at the 11th International Conference on Chemistry and Chemical Engineering (ICCCE 2020), 08-10 July 2020.
- Y. M. C. Abeykoon, 2018, Design and Applications of Soft Sensors in Polymer Processing: A Review, *Sensors Journal, IEEE*, [10.1109/JSEN.2018.2885609], <https://doi.org/10.1109/jsen.2018.2885609>
- W. Yan, T. Dong, Y. Zhou, Z. Luo, 2023, Computational modeling toward full chain of polypropylene production: From molecular to industrial scale, *Chemical Engineering Science*, Volume 269.
- N. Jumari, K. Mohd-Yusof, 2016, Comparison of Product Quality Estimation of Propylene Polymerization in Loop Reactors Using Artificial Neural Network Models, *Jurnal Teknologi*.
- J. Shi, X. Liu, Y. Sun, 2006, Melt index prediction by neural networks based on independent component analysis and multi-scale analysis, *ScienceDirect*, Volume 70, pp. 280-287.
- S. Liu, X. Gao, W. Qi, S. Zhang, 2018, Soft sensor modelling of propylene conversion based on a Takagi-Sugeno fuzzy neural network optimized with independent component analysis and mutual information, *Transactions of the Institute of Measurement and Control*.

# Markov Chain Monte Carlo simulation-based optimization for a production scale milk drying process

Adrián Ferrari<sup>a</sup>, Soledad Gutiérrez<sup>a</sup>, Gürkan Sin<sup>b</sup>

<sup>a</sup>*Chemical & Process Systems Engineering Group – Chemical Engineering Institute - Engineering School – Universidad de la República, Julio Herrera y Reissig 565, PC 11300, Montevideo, Uruguay*

<sup>b</sup>*Process and Systems Engineering Centre - Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søtofts Plads, Building 229, DK-2800 Kgs. Lyngby, Denmark*  
[aferrari@fing.edu.uy](mailto:aferrari@fing.edu.uy)

## Abstract

Spray drying is widely used for dehydration of dairy products and among, the most energy-intensive unit operation in this field. An optimization problem for a production scale milk drying system was implemented and solved, considering plant capacity (maximization) and energy consumption (minimization) as the objective. Decision variables were inputs to the spray unit namely the inlet dry bulb air temperature, concentrate moisture, and dry solids flow rate. Product stickiness conditions and moisture content were the main constraints, which were modeled using mass and energy balances. A non-deterministic derivative free based optimization technique, namely Markov Chain Monte Carlo (MCMC) algorithm was chosen to solve the problem. The results showed that throughput maximization is achieved at the expense of a relative energy consumption penalization in the spray and revealed that the structure of the problem seems to be convex. This study shows a promising non-conventional use of MCMC algorithms in optimization studies.

**Keywords:** Milk drying process, process optimization, derivative free optimization, MCMC.

## 1. Introduction

The most widely used technique for dehydration of dairy products is spray drying after evaporation. These processes preserve food properties and allow storage of powders at an ambient temperature (Schuck et al., 2008). A drying plant normally presents three stages: 1) spray chamber; 2) internal fluid bed at the conical base of the spray chamber; and 3) external fluid bed to fine tune the outfeed product stream, as it is shown in Figure 1 (Schuck et al., 2005). As spray drying is the most energy-intensive unit operation in the dairy industry, process optimization becomes of critical importance to increase efficiency and reduce production costs (Petersen et al., 2015). As a continuation of previous works in a production scale system about modelling, uncertainty, identifiability, and sensitivity analysis from the point of view of identifying influential inputs/parameters and quantifying confidence intervals in model simulations (Ferrari et



al., 2015; Ferrari et al., 2016; Ferrari et al., 2017), an optimization problem for the entire milk drying process is here implemented and solved.

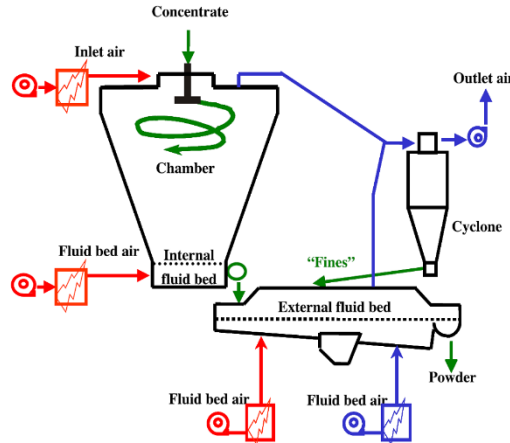


Figure 1. Physical system of three stage dryer scheme.

Maximization of plant capacity and minimization of energy consumption are considered in the objective function. Decision variables are all inputs at chamber inlet streams:

- $T_{A,1}$ : Chamber inlet dry bulb air temperature [K], which is feasible to be manipulated at the feed to the spray unit.
- $M_{P,1}$ : Concentrate moisture [ $\text{kg}_{\text{water}} \cdot \text{kg}_{\text{total}}^{-1}$ ], which is feasible to be manipulated by the output from the evaporation plant upstream of the drying processes.
- $S$ : Dry solids flow rate [ $\text{kg}_{\text{dry solids}} \cdot \text{h}^{-1}$ ], which is feasible to manipulate via the feed to the evaporation plant.

Product stickiness conditions at internal fluid bed and moisture content in final powder are assumed as the main constraints in the problem. A global sensitivity analysis of the optimization results is also performed respect to normal and not controlled variations in the chamber inlet air humidity ( $H_{A,1}$ ; range:  $0.003 - 0.020 \text{ kg}_{\text{water}} \cdot \text{kg}_{\text{dry air}}^{-1}$ ), finding the corresponding result for a different  $H_{A,1}$  value at a time. A non-deterministic/derivative free based technique (Markov Chain Monte Carlo algorithm) is chosen to solve the problem to overcome issues when using a deterministic gradient-based method. The MCMC use is also interesting for trying to visualize non-convexity patterns in the model finding the posterior density functions of decision variables, representing a non-conventional use of MCMC. Details and results about the deterministic approach are not covered in this paper because of length restrictions. The work is structured as follows: first the conceptual optimization problem is described, then the optimization technique is introduced, afterwards the obtained results are analyzed and discussed.

## 2. Materials and methods

### 2.1. Optimization problem

To consider plant capacity maximization and energy consumption minimization, the objective function (OF) is structured as follows:

$$OF = \alpha \cdot T_{A,1} - \beta \cdot P_{out} - \gamma \cdot W_{dried} \quad (1)$$

Being:

$$P_{out} = S \cdot \left[ 1 + \frac{M_{P,4}}{1 - M_{P,4}} \right] \quad (2)$$

$$W_{dried} = S \cdot \left[ \frac{M_{P,1}}{1 - M_{P,1}} - \frac{M_{P,4}}{1 - M_{P,4}} \right] \quad (3)$$

Where  $P_{out}$  is the total flow rate at system output [ $\text{kg}\cdot\text{h}^{-1}$ ];  $W_{dried}$ , the total water dried in the system [ $\text{kg}_{\text{water}}\cdot\text{h}^{-1}$ ];  $M_{P,4}$ , the final powder moisture [ $\text{kg}_{\text{water}}\cdot\text{kg}_{\text{total}}^{-1}$ ]; and  $\alpha/\beta/\gamma$ , dimensionless coefficients to have balanced terms in the objective function. It will be assumed  $\alpha = 20$ ,  $\beta = \gamma = 1$ , all constant during the complete study. The main inequality constraints to be considered are related to the product stickiness [using glass transition temperature as indicator (Silalai and Roos, 2010)] at the internal fluid bed, and also with the residual moisture content in final powder, in accordance to Eq. (4) and (5).

$$T_{P,3} \leq T_{g,3} \quad (4)$$

$$M_{P,4} \leq M_{P,max} \quad (5)$$

Where  $T_{P,3}/T_{g,3}$ , are the product one and its glass transition temperature at the internal fluid bed [ $^{\circ}\text{C}$ ]; and  $M_{P,max}$ , is the maximum allowed powder moisture at the end of the system [ $\text{kg}_{\text{water}}\cdot\text{kg}_{\text{total}}^{-1}$ ]. It is assumed in this work a value of  $0.030 \text{ kg}_{\text{water}}\cdot\text{kg}_{\text{total}}^{-1}$  for this parameter (commercial constraint for whole milk powder). The decision variables for the optimization are  $T_{A,1}$ ,  $M_{P,1}$ , and  $S$ , with the following bounds to screen an important range of operating conditions in the system.

$$453 \text{ K} \leq T_{A,1} \leq 503 \text{ K} \quad (6)$$

$$0.47 \text{ kg}_{\text{water}}\cdot\text{kg}_{\text{total}}^{-1} \leq M_{P,1} \leq 0.53 \text{ kg}_{\text{water}}\cdot\text{kg}_{\text{total}}^{-1} \quad (7)$$

$$5,000 \text{ kg}_{\text{dry solids}}\cdot\text{h}^{-1} \leq S \leq 6,600 \text{ kg}_{\text{dry solids}}\cdot\text{h}^{-1} \quad (8)$$

The optimization problem will minimize energy consumption and maximize throughput in the system, changing  $T_{A,1}$ ,  $M_{P,1}$ , and  $S$  as decision variables, subject to mass and energy balances with also powder stickiness and residual moisture constraints. The detailed process and product stickiness models are described in Ferrari et al. (2015), Ferrari et al. (2016), and Ferrari et al. (2017). The complete structure of the optimization problem is summarized as follows:

$$\begin{aligned} & \min_{T_{A,1}; M_{P,1}; S} OF \\ & \text{s. t:} \\ & T_{P,3} \leq T_{g,3} \\ & M_{P,4} \leq M_{P,max} \\ & T_{A,1}; M_{P,1}; S \in R \end{aligned}$$

## 2.2. Numerical methods

As a non-deterministic/derivative free technique, a Markov Chain based algorithm was used for solving the optimization problem and for detecting non-convex behaviours. This represents a non-conventional use of Markov Chain based algorithms, but it is interesting to analyze its capabilities and obtain first insights for finding posterior density functions of decision variables and trying to visualize non-convexity patterns. The Metropolis random walk algorithm (Gelman et al., 2004; Weinzierl, 2000) was also

used as the method to accept/reject new samples from transition distributions. The algorithm proceeds as follows:

Step 1: randomly select a feasible initial vector of decision variables,  $\theta^0$ . A uniform distribution is assumed for all of them.

Step 2: create a new trial vector,  $\theta^* = \theta^{k-1} + \Delta\theta$ , where  $\Delta\theta$  is randomly sampled from a jumping distribution  $q(\Delta\theta)$ . This is assumed as multivariate normal centered at current iteration and calculated as follows:

$$\Delta\theta = u \cdot \frac{2,4}{\sqrt{l(\theta)}} \quad (9)$$

Where  $u$  is a unit normal random number drawn for each decision variable; and  $l(\theta)$ , the length of  $\theta$  vector.

Step 3: calculate the adapted Metropolis ratio ( $r$ ):

$$r = \frac{OF_{Augmented}(\theta^k)}{OF_{Augmented}(\theta^{k-1})} \quad (10)$$

$$OF_{Augmented} = \left\{ \begin{array}{ll} OF & \text{if } \theta \in \text{Feasible Region} \\ OF_{Max} & \text{Otherwise} \end{array} \right\} \quad (11)$$

$$OF_{Max} = \alpha \cdot T_{A,1} \quad (12)$$

To compute the constraints of the optimization problem, an augmented objective function with a hard penalization term was defined in accordance to Eq. (11).

Step 4: accept the new vector by the following rule.

$$\theta^k = \left\{ \begin{array}{ll} \theta^* & \text{if a random probability} < r \\ \theta^{k-1} & \text{Otherwise} \end{array} \right\} \quad (13)$$

To report the optimal values for decision variables, the following criterion is applied: if chains show that a bound constraint is active, such bound value is assumed as the optimal; otherwise, the mean for all chains data is reported as the optimal value. All mathematical models and numerical methods were implemented in Matlab R2015a using an Intel(R) Core(TM) i7-3770 CPU@3.4 GHz with 8 GB RAM memory.

### 3. Results and discussion

The results are obtained using between 10 and 20 Markov Chains with 5,000 samples per chain. Optimal  $T_{A,1}$  and  $S$  values reach their upper bounds for all  $H_{A,1}$  range (default values in the plant are 468 K and 5,760  $\text{kg}_{\text{dry solids}} \cdot \text{h}^{-1}$  respectively), so throughput maximization is achieved at the expense of a relative energy consumption penalization in the spray. Respect to  $M_{P,1}$  results (between 0.525 – 0.529  $\text{kg}_{\text{water}} \cdot \text{kg}_{\text{total}}^{-1}$  covering the  $H_{A,1}$  range), the values are slightly higher than the default one (0.52  $\text{kg}_{\text{water}} \cdot \text{kg}_{\text{total}}^{-1}$ ) so this is aligned with water dried maximization in the system. Numerical issues were found during the deterministic approach, but the results here obtained confirmed main tendencies for decision variables. Figure 2 shows an example of the evolution for 20 chains, in which the convergence is achieved and requiring in most of the cases less than 5,000 samples. At the end of the chains upper bound is reached for  $T_{A,1}$  and  $S$ . The

computational burden was around 9 hours per optimization run and more representative results obtained from this example are presented in Table 1. Despite the acceptance ratio being low (0.021), most likely because of the constrained nature of the optimization problem, convergence was achieved for all decision variables (R-scale values close to 1), and the Monte Carlo estimation error presented low values for all cases (Gelman et al., 2004) as shown in Table 1.

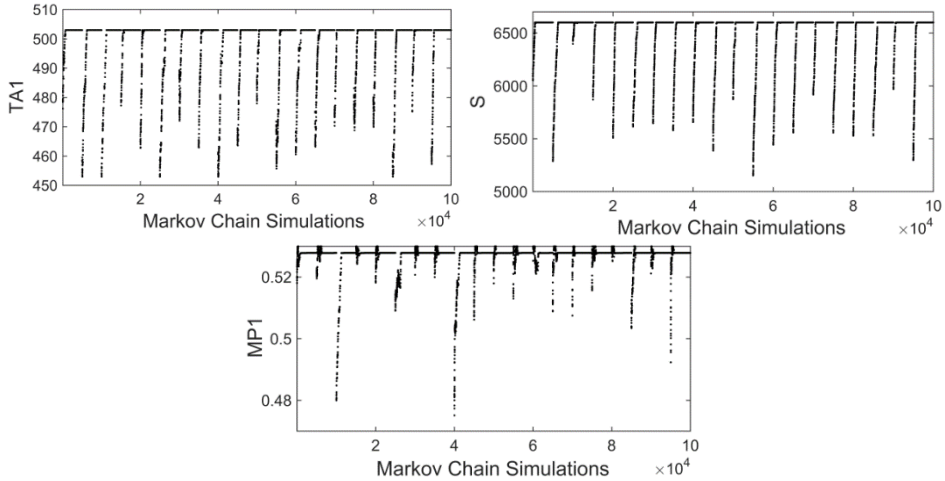


Figure 2. Simulation results for an example of 20 chains.  $[T_{A,1}] = K$ ;  $[M_{P,1}] = \text{kg}_{\text{water}} \cdot \text{kg}_{\text{total}}^{-1}$ ;  $[S] = \text{kg}_{\text{dry solids}} \cdot \text{h}^{-1}$

Table 1. Summary of performance for the example shown in Figure 2

	$T_{A,1}$	$M_{P,1}$	$S$
<b>R-scale</b>	1.0331	1.0348	1.0240
<b>Monte Carlo error</b>	0.0007 % of mean	0.0006 % of mean	0.0004 % of mean

Figure 3 shows the densities obtained for decision variables in accordance with the example shown in Figure 2.

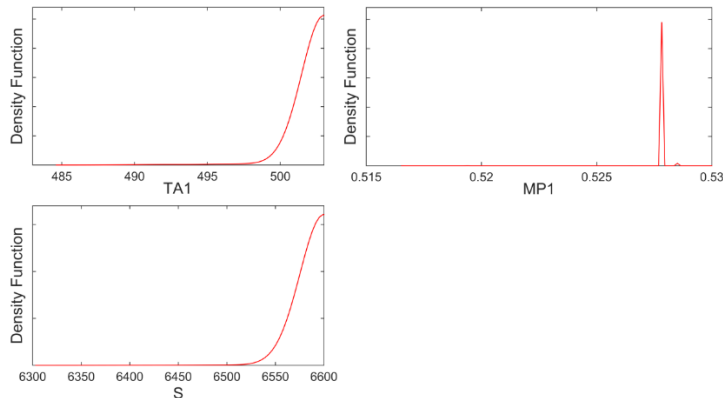


Figure 3. Density functions for decision variables.  $[T_{A,1}] = K$ ;  $[M_{P,1}] = \text{kg}_{\text{water}} \cdot \text{kg}_{\text{total}}^{-1}$ ;  $[S] = \text{kg}_{\text{dry solids}} \cdot \text{h}^{-1}$

As can be interpreted, the nature of the problem seems to be convex since only one peak appeared in density functions. Hence, local optimum would also be global. In general MCMC explore the decision space using a random walk algorithm, which makes it robust against the initial conditions of decision variables. On the other hand, a prior

distribution is required for them and the slow convergence can be improved using different methods such as differential evolution, parallel computing, etc. (Vrugt, 2016).

#### 4. Conclusions

An optimization problem for a production scale milk spray and fluid bed drying was implemented and solved in Matlab. Results from a non-deterministic based technique (Markov Chain based random walk algorithm) revealed that dry bulb air temperature and dry solids flow rate could be controlled close to their upper bounds. These insights are aligned with the current situation in the production scale plant which this work is based on. It can also be observed that optimal concentrate moisture values are higher than the default one for the base problem, being this aligned with the maximization of total water dried in the system. The results involve an improvement in plant throughput but with a penalization in relative energy consumption in the spray dryer compared with the base operational point. The non-deterministic based results revealed that the structure of the problem seems to be convex. Despite it needs further theoretical analysis, the application case study shows promising results as a novel technique for non-deterministic/derivative free global optimization. This is especially suited for industrial applications, where the process models are not amenable for calculation of derivatives (Magnússon et al., 2020). Trying other non-deterministic tools such as Genetic Algorithms, Particle Swarm, etc., and doing a more comprehensive study on the multi-objective optimization problem, can be explored as future works. The developed model is expected to be a promising tool for decision-making at industrial scale.

#### References

- P. Schuck; S. Mejean; A. Dolivet; and R. Jeantet, 2005, Thermohygrometric sensor: A tool for optimizing the spray drying process, *Innovative Food Science & Emerging Technologies*, 6(1), 45-50.
- P. Schuck; A. Dolivet; S. Méjean; and R. Jeantet, 2008, Relative humidity of outlet air: the key parameter to optimize moisture content and water activity of dairy powders, *Dairy Science & Technology*, 88, 45-52.
- L. N. Petersen; J. B. Jorgensen; and J. B. Rawlings, 2015, Economic optimization of spray dryer operation using nonlinear model predictive control with state estimation, *IFAC-PapersOnLine*, 48(8), 507-513.
- A. Ferrari; S. Gutiérrez; and G. Sin, 2015, A comprehensive sensitivity and uncertainty analysis of a milk drying process, *Computer Aided Chemical Engineering*, 37, 2225-2230.
- A. Ferrari; S. Gutiérrez; and G. Sin, 2016, Modeling a production scale milk drying process: parameter estimation, uncertainty and sensitivity analysis, *Chemical Engineering Science*, 152, 301-310.
- A. Ferrari; S. Gutiérrez; and G. Sin, 2017, Powder stickiness in milk drying: uncertainty and sensitivity analysis for process understanding, *Computer Aided Chemical Engineering*, 40, 2743-2748.
- A. F. Magnússon, R. Al, G. Sin, 2020, Development and application of simulation-based methods for engineering optimization under uncertainty, *Computer Aided Chemical Engineering*, 48, 451-456.
- A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, 2004. *Bayesian data analysis. Texts in statistical science series.*
- S. Weinzierl, 2000. Introduction to Monte Carlo methods. arXiv preprint hep-ph/0006269.
- J. A. Vrugt, 2016. Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation, *Environmental Modelling & Software*, 75, 273-316.
- N. Silalai, and Y. H. Roos, 2010. Roles of water and solids composition in the control of glass transition and stickiness of milk powders, *Journal of Food Science*, 75(5), 285-296.

# Polymer Grade Transition Control via Reinforcement Learning Trained with a Physically Consistent Memory Sequence-to-Sequence Digital Twin

Zhen-Feng Jiang,<sup>a</sup> David Shan-Hill Wong,<sup>a\*</sup> Jia-Lin Kang,<sup>b\*</sup> Yuan Yao,<sup>a\*</sup> Yao-Chen Chuang,<sup>c</sup>

<sup>a</sup>*Department of Chemical Engineering, National Tsing Hua University, Hsinchu 30043, Taiwan*

<sup>b</sup>*Department of Chemical and Materials Engineering, National Yunlin University of Science and Technology, Yunlin 64002, Taiwan*

<sup>c</sup>*Center for Energy and Environmental Research, National Tsing Hua University, Hsinchu 30043, Taiwan*

\*[dshwang@che.nthu.edu.tw](mailto:dshwang@che.nthu.edu.tw), [jlakang@yuntech.edu.tw](mailto:jlakang@yuntech.edu.tw), [yyao@mx.nthu.edu.tw](mailto:yyao@mx.nthu.edu.tw)

## Abstract

In this work, a memory layer sequence-to-sequence digital twin (ML-StSDT) of a high-density polyethylene (HDPE) reactor simulated by ASPEN Dynamics<sup>TM</sup> was constructed using simulated grade transition and steady-state operating data. A reinforcement learning control (RLC) algorithm was developed by training with the ML-StSDT. The RLC was able to control both grade transition and steady-state operation of the simulated plant. The RLC performs better or equally well when compared with the direct application of ML-StSDT in nonlinear model predictive control (NLMPC) but substantially reduces the computation load. Our results demonstrate the feasibility of deep learning models serving as a digital twin for RLC training in nonlinear process control applications.

**Keywords:** Reinforcement learning; Sequence-to-Sequence with Memory Layer; Model Predictive Control, Grade transition.

## 1. Introduction

Model predictive control (MPC) has been used in the process industry to maximize the productivity of the process and reduce energy consumption. As the scope of the process being controlled increases, the following issues become more challenging. First, the number of state variables to be identified and decision variables to be optimized increases. Dynamic prediction horizons also become more and more complex and nonlinear. Hence a nonlinear model is required. Computation load becomes a challenging issue and control policies within the sampling period were increasingly difficult and suboptimal decisions may be made (Mayne et al. 2014) and The computation difficulty still limits control policy optimization over a less satisfactory horizon. The similarity between reinforcement learning (RL) and MPC has long been recognized, studied, and

reviewed by G6rges (2017). RL can drastically reduce online computing time. According to Nian et al. (2020), while RL methods such as deep deterministic policy gradient (DDPG) are suitable for online control, application of RL, application of RL to industrial control is mostly limited to a simulation environment. The Monte Carlo method is suitable for generating an infinite set of operations with a clear system terminal state, and the average return can be obtained after each trajectory calculation and does not require estimating values through previously estimated values to avoid the value function falling into local optima (Yoo et al., 2021).

## 2. Memory Layer Sequence-to-Sequence Model

In our previous work, Jiang et al. (2022), we developed a sequence-to-sequence with memory layer (ML-StS) model (Figure 1) for a high-density polyethylene slurry reactor simulator (Figure 2) using ASPEN Dynamics™, based on the kinetic model by Khare et al. (2002). The sensors, manipulation and controlled variables (SV, MV, and CV) are shown in Table 1. The operator employed catalyst flow and hydrogen flow as MV to control the hydrogen-to-ethylene ratio (HER) and pressure. The ML-StS model was able to predict dynamic behavior with high accuracy as well as consistent directionality of MV/CV pairs u shown in Table 2.

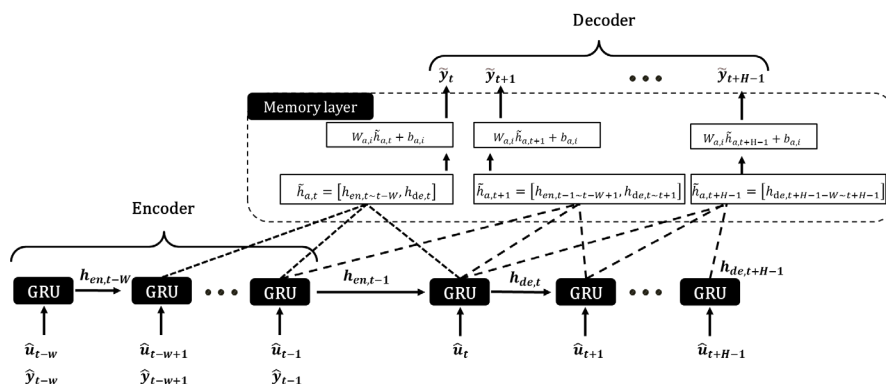


Figure 1. Structure of an StS model with the memory layer

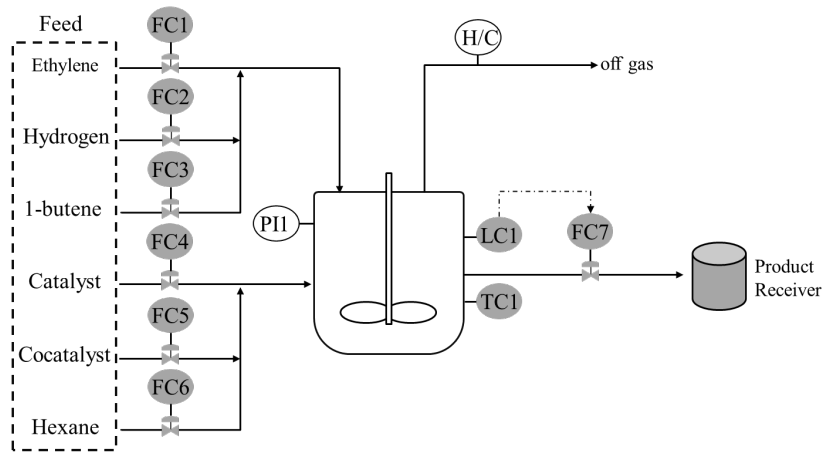


Figure 2. The schematic of the HDPE process.

Table 1: List of variables

Type	Tag
MV	The flow rate of ethylene
MV	The flow rate of hydrogen
MV	The flow rate of 1-butene
MV	The flow rate of catalyst
MV	The flow rate of cocatalyst
MV	The flow rate of hexane
MV	Temperature of reactor
CV	Pressure of reactor
CV	HER of purge gas

Table 2: Directionality of gain

	HER	Pressure
The flow rate of c	+	-
The flow rate of hydrogen	+	+

### 3. Nonlinear Model Predictive Control

The ML-StS model was then used in navigation as the predictive model used by nonlinear model predictive control (NLMPC) as shown in Figure 3. The differential evolution algorithm (DEA) was used to plan control actions in the future horizon and according to the future trajectory objective and operating constraints. While the NLMPC scheme is successful for both grade transition and steady-state operation, the computation load is too heavy to prevent frequent updating of manipulative variables.



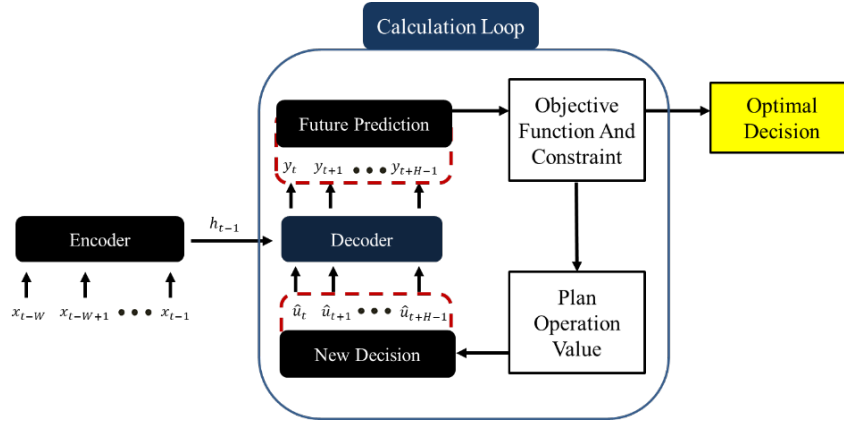


Figure 3. The schematic of the MPC calculation process.

#### 4. Reinforcement Learning Control

Instead of direct application in NLMPC, the ML-StS model can also serve as a digital twin (DT) environment for training a reinforcement learning control (RLC). In this study, the “state” in RL was defined as the input to ML-StS encoder and future CV target setpoints trajectory. The “actor” was defined as the control action of the immediate future. The Monte Carlo Deep Deterministic Policy Gradient (MC-DDPG) was used to train the state actor relation using the ML-StSDT as the “environment” (Figure 4). Eq. 1~8 are utilized to calculate the reward of action. Eq. 1 is the main target of the hydrogen-ethylene ratio. Eq. 2,3 make sure the current pressure (Pre.) is between the desired upper and lower limits (UL, LL). Eq. 4~7 limits the MV, catalyst flow (Cat.), and hydrogen flow (hyd.), within the allowable operating range. Eq. 8 allows RL models to achieve setpoint with minimal MV changes. After training using the ML-StSDT, RLC is then tested by the physical simulator.

$$R_1 = -|(\mathbf{HER}_{setpoint,state} - \tilde{\mathbf{y}}_{HER,t})| * 10^5 \quad (1)$$

$$R_2 = -|(\mathbf{Pre}_{UL} - \tilde{\mathbf{y}}_{Pre,t})| * 10^7, \text{ if } \tilde{\mathbf{y}}_t > \mathbf{Pre}_{UL} \quad (2)$$

$$R_3 = -|(\mathbf{Pre}_{LL} - \tilde{\mathbf{y}}_{Pre,t})| * 10^7, \text{ if } \tilde{\mathbf{y}}_{Pre,t} < \mathbf{Pre}_{LL} \quad (3)$$

$$R_4 = -|(\mathbf{Hyd}_{UL} - \mathbf{u}_{hyd,t})| * 10^6, \text{ if } \mathbf{u}_{hyd,t} > \mathbf{Hyd}_{UL} \quad (4)$$

$$R_5 = -|(\mathbf{Hyd}_{LL} - \mathbf{u}_{hyd,t})| * 10^6, \text{ if } \mathbf{u}_{hyd,t} < \mathbf{Hyd}_{UL} \quad (5)$$

$$R_6 = -|(\mathbf{Cat}_{UL} - \mathbf{u}_{cat,t})| * 10^6, \text{ if } \mathbf{u}_{cat,t} > \mathbf{Cat}_{UL} \quad (6)$$

$$R_7 = -|(\mathbf{Cat}_{LL} - \mathbf{u}_{cat,t})| * 10^6, \text{ if } \mathbf{u}_{cat,t} < \mathbf{Cat}_{UL} \quad (7)$$

$$R_8 = -|\Delta \mathbf{u}_t| * 10^1 \quad (8)$$

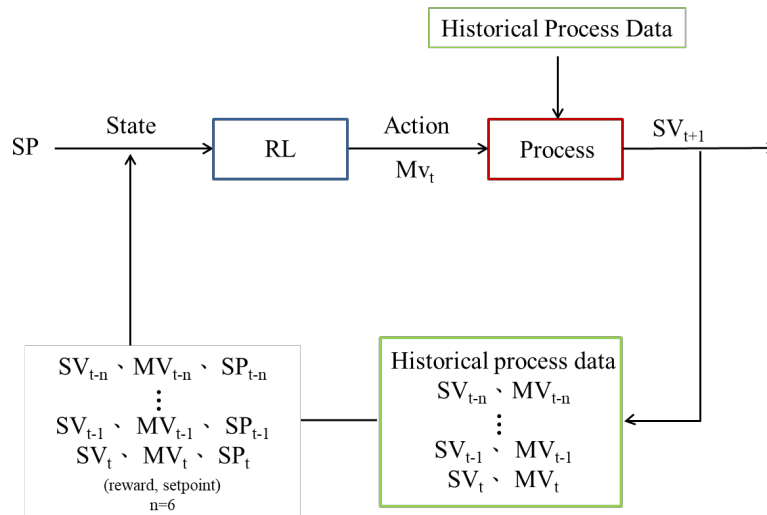


Figure 4. The schematic of the RL interacting with the HDPE process.

## 5. Results

Figure 5 and Figure 6 compared the results of two cases of grade transition directed by controlled LC and NLMPC. As shown in Figure 5 (a) the HER IAE of NLMPC is 34% lower in NLMPC but RLC produces a much smaller overshoot. In Figure 6 (a) the HER IAE of RLC and NLMPC are similar. However, NLNMC results in the oscillation of NLMPC being much more severe while RLC provides a smoother transition. Figure 5(b, d) and Figure 6 (b, d) showed that the action of NLMPC is much more aggressive while RLC provides a much smoother action due to higher frequency updating of control action. The changes in cumulative total and individual rewards of the RLC and NLMPC as shown in Figure 7. RLC also shows a much smoother trajectory. RLC computation time for each decision-making (18 milliseconds) is only about 1/1000 compared with the cost time of calculation of MPC (18 seconds). The 99.9% time reduction of RLC demonstrates the superior potential for the real complex process.

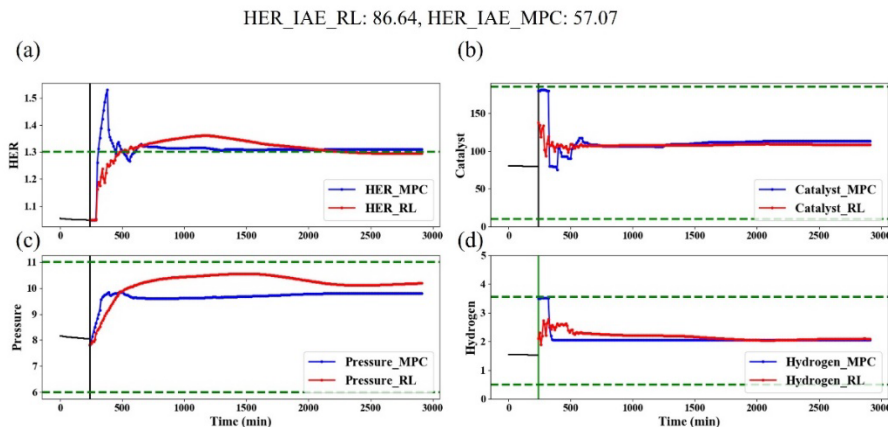


Figure 5. The result of MPC and RL of (a) the H<sub>2</sub>/C<sub>2</sub>H<sub>4</sub> ratio (b) the flow rate of the catalyst (c) the pressure of the reactor (d) the flow rate of hydrogen.

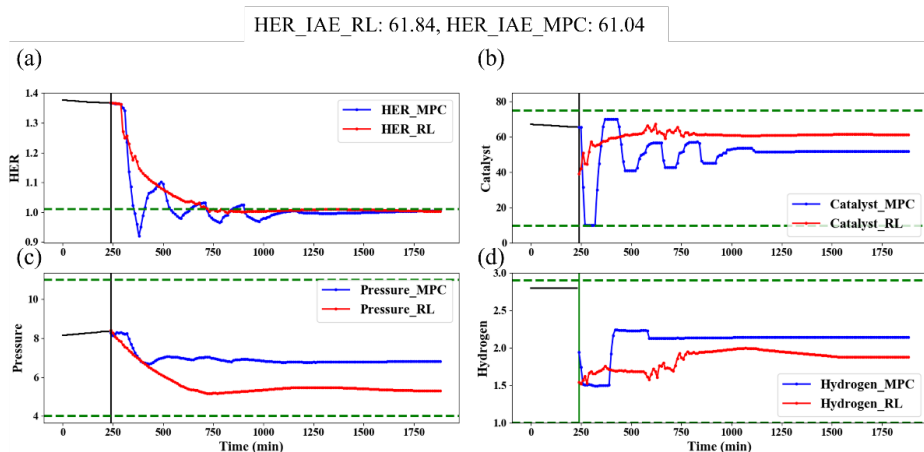


Figure 6. The result of MPC and RL of (a) the H<sub>2</sub>/C<sub>2</sub>H<sub>4</sub> ratio (b) the flow rate of the catalyst (c) the pressure of the reactor (d) the flow rate of hydrogen.

## 6. Conclusion

In this work, we demonstrated that a memory layer sequence-to-sequence digital twin (ML-StSDT) of a high-density polyethylene (HDPE) reactor can serve as a model for NLMPC as well as a surrogate environment for training an RLC. The RLC performs better or equally well when compared with direct application in NLMPC but the online computation load is much reduced.

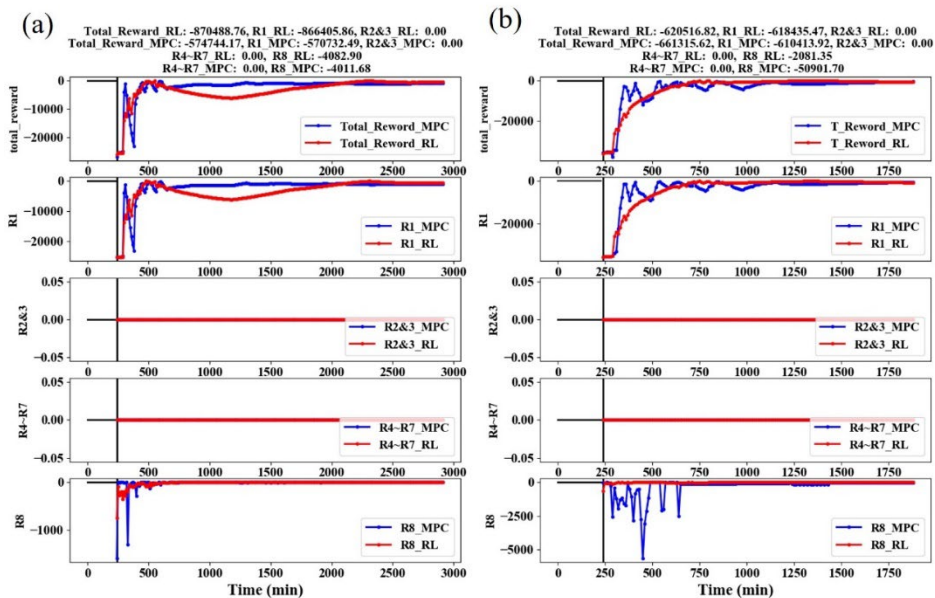


Figure 7. The result of total reward and R1~R8 reward of two different directions of grade transitions of MPC and RL.

## **Acknowledgment**

This work was supported by the National Science and Technology Council of Taiwan (contract nos. NSTC 111-2221-E-007-005).

## **References**

- Görges, D. (2017). Relations between model predictive control and reinforcement learning. *IFAC-PapersOnLine*, 50(1), 4920-4928.
- Jiang, Z. F., Wei, X. Z., Wong, D. S. H., Yao, Y., Kang, J. L., Chuang, Y. C., ... & Ou, J. D. Y. (2022). Model Predictive Control of Grade Transition with Attention Base Sequence-to-Sequence Model. In *Computer Aided Chemical Engineering* (Vol. 49, pp. 367-372). Elsevier.
- Khare, N. P., Seavey, K. C., Liu, Y. A., Ramanathan, S., Lingard, S., & Chen, C. C. (2002). Steady-state and dynamic modeling of commercial slurry high-density polyethylene (HDPE) processes. *Industrial & engineering chemistry research*, 41(23), 5601-5618.
- Mayne, D. Q. (2014). Model predictive control: Recent developments and future promise. *Automatica*, 50(12), 2967-2986.
- Nian, R., Liu, J., & Huang, B. (2020). A review on reinforcement learning: Introduction and applications in industrial process control. *Computers & Chemical Engineering*, 139, 106886.
- Yoo, H., Kim, B., Kim, J. W., & Lee, J. H. (2021). Reinforcement learning based optimal control of batch processes using Monte-Carlo deep deterministic policy gradient with phase segmentation. *Computers & Chemical Engineering*, 144, 107133.



# A Rolling Horizon Approach for Ethylene Production Scheduling with Daily Inventory Constraints

Haoran Li,<sup>a,b</sup> Tong Qiu,<sup>a,b</sup>

<sup>a</sup> *Department of Chemical Engineering, Tsinghua University, Beijing 100084, China*

<sup>b</sup> *Beijing Key Laboratory of Industrial Big Data System and Application, Tsinghua University, Beijing 100084, China*  
*qiutong@tsinghua.edu.cn*

## Abstract

The supply chain for ethylene production is becoming unstable, especially after Covid-19. This situation poses new challenges for ethylene plants to handle frequently changing feed supply conditions with more responsive scheduling strategies. This paper proposes a rolling horizon approach combined with moving average supply forecasting to assist in dynamic decision-making for ethylene production scheduling. When applied to an industrial case study, the proposed model shows superior calculation speed as well as the capability to find fairly good scheduling solutions under uncertain supply conditions.

**Keywords:** rolling horizon, scheduling, ethylene, inventory management

## 1. Introduction

Ethylene is one of the most important chemicals in the world and ethylene plants form one of the most profitable industrial chains (Li et al., 2022b). Owing to Covid-19, the global supply chain is experiencing ongoing turmoil, which increases the feedstock supply uncertainty in the ethylene industry. This brings pressing needs for ethylene plants to adopt a dynamic scheduling strategy to respond to unstable upstream supply conditions to guarantee optimal economic profits over time.

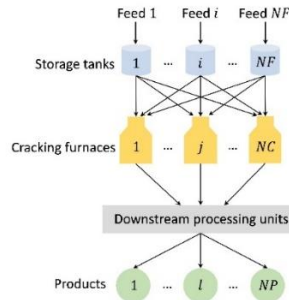
In recent years, ethylene production scheduling has been studied intensively. Ethylene cracking furnace systems process multiple feeds in multiple furnaces to produce ethylene and other olefin products. As the product yields vary with the furnace and the cracking duration as a result of coking inside the coil, the scheduling strategy, which determines the feedstock allocation between furnaces and the batch lengths in each furnace, has vital influences on the economic profits of the ethylene plants. And and Grossmann (And and Grossmann, 1998) first proposed a mixed-integer nonlinear programming (MINLP) model for the ethylene scheduling problem. After that, many modifications, for example, the non-simultaneous constraints (Liu et al., 2010) and integration with operation optimization (Li et al., 2022a) have been made to improve the model.

However, the dynamic response to the changing feedstock supply conditions has been overlooked in these previous scheduling studies. Facing this challenge, this paper developed a novel dynamic scheduling framework including (1) a Rolling horizon approach for the self-update of decision variables (2) Feed supply forecasting by moving average estimation (3) Mathematical formulation of the scheduling model with daily inventory constraints. This framework enables the model to receive new supply information from the environment and make quick responses by updating the scheduling

variables. To demonstrate the efficacy of the developed new dynamic scheduling framework, a reference industrial case study taken from the literature is presented.

## 2. Problem statement

Fig. 1 shows the mass flow in a typical ethylene cracking furnace system (Li et al., 2022a). Feed  $i$  ( $1, 2, \dots, NF$ ) is cracked in furnace  $j$  ( $1, 2, \dots, NF$ ) and finally produces product  $l$  ( $1, 2, \dots, NP$ ) after quenching and separating.



**Figure 1.** Mass flow of the ethylene cracking furnace system

However, because the feedstock supply conditions change frequently and decision-makers can't foresee the perfect information about the future feedstock supply conditions, updates of the following scheduling decision variables: (1) the inlet feed type in each furnace; (2) whether to perform decoking in a furnace given the visible future supply conditions, the inventory conditions, and the furnace conditions on the current day  $n$  are thus needed.

## 3. Methods

### 3.1. Rolling horizon approach

The rolling horizon approach is demonstrated in Fig. 2. The core idea is to optimize over a fixed horizon based on visible parameters and re-optimize the model repeatedly once the time step moves forward and new parameters are received. On the current day, the model will receive visible feed supply parameters from the market and optimize the schedule for the next  $H_1$  days (lookahead horizon, LH). Before the optimization, the variables for the first  $H_2$  days (fixed horizon, FH) will be fixed at variable solution levels from the previous schedule. In real cases, the time horizon of visible feed supply will be shorter than LH (referred to as imperfect information), which requires the model to estimate the future supply parameters based on the current information and generate the schedule. After the optimization, the schedule on the current day will be executed and the rest schedules in FH will be passed to the next optimization.



**Figure 2.** Illustration of rolling horizon approach

### 3.2. Supply forecasting by moving average estimation

Due to the existence of randomness, it's impossible to accurately foresee the long-term feed supply conditions needed for the scheduling optimization beforehand. We assume the feed supply conditions for the next  $H_2$  days will be known (perfect information). One possible solution to this is to estimate the future supply based on the current visible information. Here, we proposed the moving average estimation to achieve this. As shown in Eq. (1), on the current day, the supply parameters will be observed as the true value for the first  $H_2$  days. While for days between  $H_2$  and  $H_1$ , the supply amount will be estimated as the average supply amount of the first  $H_2$  days, as shown in Eq. (2).

$$supply_{i,k} = supply_{i,k}^{obs}, \quad \forall i, 1 \leq k \leq H_2 \quad (1)$$

$$supply_{i,k} = \frac{1}{H_2} \sum_{k=1}^{H_2} supply_{i,k}^{obs}, \quad \forall i, H_2 < k \leq H_1 \quad (2)$$

### 3.3. Rolling horizon scheduling optimization model

#### 3.3.1. Objective function

The objective of the rolling horizon scheduling model is to maximize the net profits over the next  $H_1$  days. As shown in Eq. (3), the net profits are the product revenue minus the raw material costs, operation costs, and the decoking costs. Binary variable  $y_{i,j,k}$  indicates whether feed  $i$  is processed in furnace  $j$  on day  $k$ . Binary variable  $x_{j,k}$  denotes whether to perform decoking for furnace  $j$  on day  $k$ .  $D_{i,j}$  is the flow rate of feed  $i$  in furnace  $j$ .  $P_l$  (the price of product  $l$ ),  $Cr_i$  (raw material cost of feed  $i$ ),  $Cv_{i,j}$  (operation cost for processing feed  $i$  in furnace  $j$ ),  $Cs_j$  (decoking costs for furnace  $j$ ) are price-related parameters.

$$Max \quad J = \sum_i \sum_j \sum_k \sum_l P_l y_{i,j,k,l} y_{i,j,k} D_{i,j} - (Cr_i + Cv_{i,j}) y_{i,j,k} D_{i,j} - Cs_j x_{j,k} \quad (3)$$

#### 3.3.2. Fixed horizon constraints

$invent_{i,k}$  denotes the inventory level of feed  $i$  on day  $k$ . The integer variable  $day_{i,j,k}$  is an indicator for the current batch lengths so far on day  $k$ . For each of the rolling horizon scheduling models, the decision variables in the fixed horizon will be fixed at the variable levels from the last scheduling model, as shown in Eq. (4). This guarantees decision consistency between adjacent days.

$$\begin{aligned} & invent_{i,k}, day_{j,k}, y_{i,j,k}, x_{j,k} \\ & = invent_{i,k}^{Fixed}, day_{j,k}^{Fixed}, y_{i,j,k}^{Fixed}, x_{j,k}^{Fixed}, \forall i, j, 1 \leq k \leq H_2 \end{aligned} \quad (4)$$

#### 3.3.3. Feedstock allocation and decoking constraints

When allocating the feeds in different furnaces, only one type of feed is allowed to be cracked in a furnace at one time. If the furnace undergoes decoking, no feed will be allowed, as shown in Eq. (5). Besides, simultaneous decoking is restricted (Eq. (6)). Additionally, furnaces are not supposed to execute decoking on two consecutive days (Eq. (7)). Eq. (8) assures that if a furnace will not change the inlet feed type without performing a decoking operation.

$$\sum_i y_{i,j,k} \leq 1 - x_{j,k}, \quad \forall j, k \quad (5)$$



$$\sum_j x_{j,k} \leq 1, \forall k \quad (6)$$

$$x_{j,k-1} + x_{j,k} \leq 1, \forall j, k \quad (7)$$

$$(1 - x_{j,k-1} - x_{j,k})(y_{i,j,k} - y_{i,j,k-1}) = 0, \forall i, j, k \quad (8)$$

### 3.3.4. Batch length and yields constraints

If decoking happens on day  $k$ , the indicator  $day_{i,j,k}$  will be 0. Otherwise, it will be  $day_{i,j,k-1}$  plus 1 day (Eq. (9)). Owing to the limits of coking thickness, the batch lengths are not allowed to exceed the upper limit  $tup_{i,j}$ , as shown in Eq. (10). Besides, to avoid overly short batches, decoking is not allowed before the current batch length reaches the lower limit  $tlo_{i,j}$  (Eq. (11)). The product yields  $yd_{i,j,k,l}$  are determined by the exponential yielding model in Eq. (12).  $a_{i,j,l}$ ,  $b_{i,j,l}$ ,  $c_{i,j,l}$  are yield-related parameters.

$$day_{i,j,k} = y_{i,j,k}(day_{i,j,k-1} + 1), \forall i, j, k \quad (9)$$

$$day_{i,j,k} \leq tup_{i,j}, \forall i, j, k \quad (10)$$

$$day_{i,j,k} x_{j,k} (day_{i,j,k-1} - tlo_{i,j}) \geq 0, \forall i, j, k \quad (11)$$

$$yd_{i,j,k,l} = c_{i,j,l} + a_{i,j,l} e^{b_{i,j,l} day_{i,j,k}}, \forall i, j, k, l \quad (12)$$

### 3.3.5. Daily inventory constraints

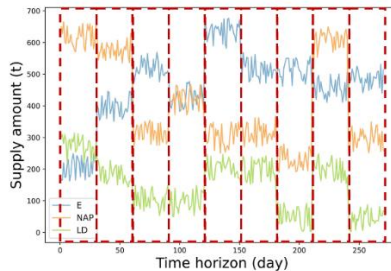
The inventory level for each type of feed is determined by the inventory level of the previous day  $invent_{i,k-1}$ , the arrival supply on the current day  $supply_{i,k}$  and the processed amount on the current day  $\sum_j y_{i,j,k} D_{i,j}$ , as shown in Eq. (13). Note that  $invent_{i,k-1}$  should be non-negative.

$$invent_{i,k} = invent_{i,k-1} + supply_{i,k} - \sum_j y_{i,j,k} D_{i,j}, \forall i, k \quad (13)$$

## 4. Case study

### 4.1. Feedstock supply conditions Parameters

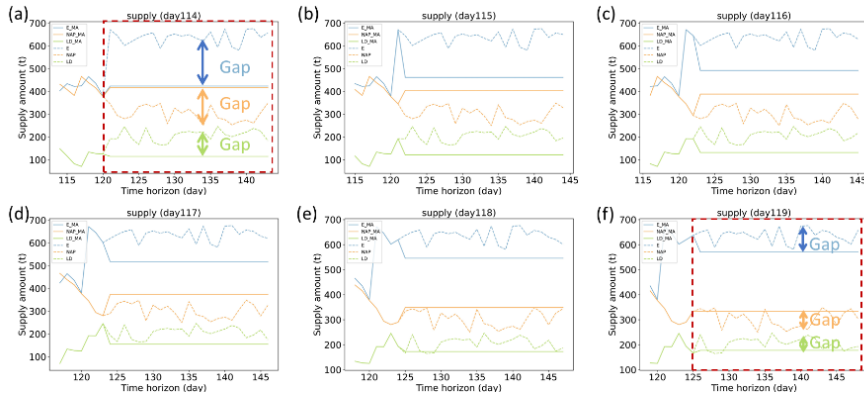
To evaluate the efficacy of the proposed model, a case study taken from the literature (Liu et al., 2010) with modification of feed supply conditions is used in this work. As shown in Fig. 3, the supply mode varies every 30 days. The other parameters remain the same as in the original literature.



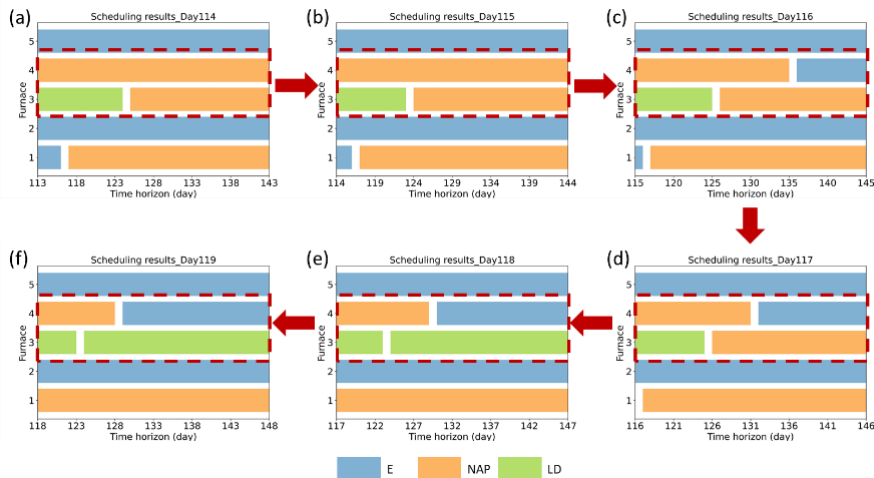
**Figure 3.** Feed supply conditions in 270 days.

### 4.2. Scheduling results

To demonstrate how our proposed rolling horizon model can respond to feed supply changes, Fig. 4 shows the supply parameters from day 114 to day 119. The solid line in each subplot represents the estimated parameters using moving average estimation, while the dotted line represents the future true supply parameters which are invisible on the current day.

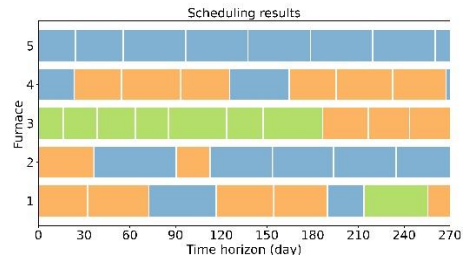


**Figure 4.** Feed supply parameters from day 114 to day 119.



**Figure 5.** Scheduling results from day 114 to day 119.

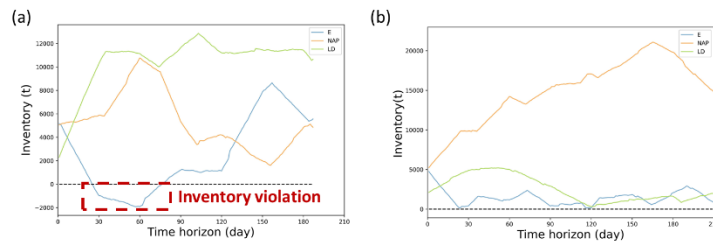
As can be observed in Fig. 4, the feed supply mode experiences a shift around day 120 with the ethane and light diesel supply increasing and the naphtha supply decreasing. On day 114, the estimated supply isn't aware of the change. As time step moving forward, the supply estimation for ethane and light diesel gradually upshifts because more observations are available now. Accordingly, the optimal schedule adjusts itself. As shown in Fig. 5(a) and (c), Furnace 4 planned to process naphtha from day 113 to 143 originally, but the schedule changed on day 116 by executing decoking on day 136 and processing ethane from day 137. Similarly, Furnace 3 planned to process light diesel followed by naphtha on day 114 but changes to process light diesel followed by light diesel on day 118 (Fig. 5(e)). The schedule changes in the two furnaces show how our model can track the supply conditions change and update the schedule accordingly. The executed optimal schedule for a long horizon of 270 days is shown in Fig. 6.



**Figure 6.** The executed scheduling results over 270 days.

#### 4.3. Comparison of model size and solution with the original literature

The original model from the literature and our proposed model are developed in GAMS v38.2 on an Intel Core i7-9700CPU @ 3.00 GHz CPU with 16.0 GB RAM. Without daily inventory constraints, there's no guarantee that inventory constraints are satisfied at all times, as shown in Figure 7. Instead, the rolling model ensures feasible inventory levels.



**Figure 7.** Inventory comparison of the original solution and rolling-horizon solution. The model comparison results are listed in Table 1. The number of variables increases in the rolling model but the solution time is shortened. The objective value is quite close to the original solution, meaning good solution quality with supply uncertainties.

**Table 1.** Model size and solution comparison.

Variable	Original model	Rolling horizon model
Equations	8560	4442
Variables	1855	3392
Solution time	~ 32 min	~ 40 s
Objective value	151204 \$/day	148495 \$/day

## 5. Conclusion

This paper proposed a rolling horizon framework to deal with the ethylene production scheduling problem under feed supply uncertainty. When applied to a case study from the literature, our model shows superiority in reducing the model size as well as guaranteeing a good scheduling solution with consideration of daily inventory management.

## References

- And, V.J., Grossmann, I.E., 1998. Cyclic scheduling of continuous parallel-process units with decaying performance. *AIChE J.* 44, 1623–1636.

- Li, H., Zhang, S., Qiu, T., 2022a. Two-Level Decoupled Ethylene Cracking Optimization of Batch Operation and Cyclic Scheduling. *Ind. Eng. Chem. Res.* acs.iecr.2c01682.
- Li, H., Zhang, S., Zhou, J., Wang, Z., Zhao, X., Qiu, T., 2022b. Toward carbon-neutral ethylene production: assessment of the application potential of bio-ETHYLENE production pathways in China. *Biofuels Bioprod Bioref* bbb.2406.
- Liu, C., Zhang, J., Xu, Q., Li, K., 2010. Cyclic scheduling for best profitability of industrial cracking furnace system. *Computers & Chemical Engineering* 34, 544–554.



# Simulation and optimization of fixed bed and moving bed TSA processes for post-combustion CO<sub>2</sub> capture.

Shreenath Krishnamurthy<sup>a\*</sup>, Richard Blom<sup>a</sup>, Giorgia Mondino<sup>a</sup>, Schalk Cloete<sup>b</sup>, John Morud<sup>b</sup>, Abdelghafour Zaabout<sup>b</sup>, Philip Llewelyn<sup>c</sup>, Cecile Pereira<sup>d</sup>, and Veronique Pugnet<sup>c</sup>.

<sup>a</sup>*Process chemistry and functional materials, SINTEF Industry, Oslo, 0373, Norway*

<sup>b</sup>*Flow technology, SINTEF Industry, Trondheim, 7031, Norway*

<sup>c</sup>*TotalEnergies, Pau, 64000, France*

<sup>d</sup>*TotalEnergies, Palaiseau, 91120, France*

\**Shreenath.Krishnamurthy@sintef.no*

## Abstract

In this work three different TSA process configurations were chosen, and these processes were optimized using genetic algorithm. One of the processes was a moving bed TSA process while the other two process were fixed bed processes using a packed bed containing pellets and hollow fiber modules. With Lewatit as the sorbent chosen, the cycle configurations were optimized to identify minimum energy and maximum productivity values for CO<sub>2</sub> capture from a natural gas combustion cycle (NGCC) flue gas. The hollow fiber TSA process had the highest productivity of 4.25 mol CO<sub>2</sub>/m<sup>3</sup> ads s, followed by the fixed bed and the moving bed TSA processes with values of 2.03 and 1.68 4.25 mol CO<sub>2</sub>/m<sup>3</sup> ads s, respectively

**Keywords:** TSA process, optimization, CO<sub>2</sub> capture

## 1. Introduction

Adsorption processes using solid sorbents rely on the differences in the affinities of the different gases towards the solid and has been implemented on a large scale for several gas separation applications. Adsorption processes are broadly classified into two types based on the mode of regeneration of the sorbent: Pressure/vacuum swing adsorption (PSA/VSA) where the sorbent is regenerated by lowering the pressure of the column and temperature swing adsorption (TSA) process where the sorbent is regenerated by the supply of heat directly or indirectly. TSA processes are better suited to low concentrations of CO<sub>2</sub> than PSA/VSA processes such as a natural gas combustion (NGCC) cycle, where it is possible to achieve purity values >95% from a feed containing 3-5% CO<sub>2</sub>. However, TSA processes applied to conventional fixed bed columns are usually slow due to the long heating and cooling times, and this is especially a drawback as it lowers the productivity of the process. Such a drawback could be mitigated through a combination of direct heating and the use of structured adsorbents like hollow fibres with better mass transfer and lower pressure drop, resulting in improved performance of the process in comparison with traditional packed bed systems [1,2]. At the core of an adsorption process is the adsorbent and the performance of an adsorbent depends on the specifics of the process. Therefore, each adsorbent must be mapped towards the best process configuration [3]. This information can only be obtained from detailed simulations and optimization of different adsorbent-process combination in a standardized manner to

accurately reflect competitiveness. The present work employs a homogenized model to study the following processes: 1. A fixed TSA process with pellets 2. A fixed bed TSA process with hollow fibers 3. A moving bed TSA process with pellets. The adsorbent chosen for this study is a commercially sold benzyl amine polymer, Lewatit for a representative NGCC cycle with 5% CO<sub>2</sub> and 85% N<sub>2</sub>. The three cycle configurations were optimized to identify minimum energy and maximum productivity values for purity-recovery targets of 95% and 90%, respectively.

## 2. Cycle design

For the fixed bed a 4-step TSA process was chosen. The four steps are as follows and the schematic of the process is shown in Figure 1.

1. Adsorption with feed: Feed enters the column at 313 K. Adsorption of the CO<sub>2</sub> takes place while nitrogen is collected at the product end. The outlet of the column is maintained at 1 atm. A part of the nitrogen is used to purge the column in the light reflux step.
2. Heavy reflux step with light reflux product (HRLRP): This step uses the product of the light reflux step. The duration of this step is the same as the light reflux (LR) step since the entire product of the latter is recycled back to the column. The CO<sub>2</sub> concentration in the column is enhanced during this step. The column is also heated up slightly with the gases coming from the LR step.
3. Desorption with steam purge: The desorption of the CO<sub>2</sub> is aided by counter-currently purging the column with steam/water at a high temperature.
4. Light reflux step (LR): The remaining CO<sub>2</sub> in the column is then purged with the nitrogen product of the adsorption step from the counter-current direction. The product is fully recycled back in to the HRLRP step.

A similar concept was also considered for the moving bed temperature swing adsorption (MBTSA) process shown in Figure 1. The moving bed process is a simplified process with the adsorption and desorption sections. In the adsorption step, the solid flows counter-current to the gas. Once the adsorption step is complete, the HRLRP step is carried out with the solid staying constant. Then the solid is circulated to the desorption section. Here steam is fed to the column in a counter current direction and CO<sub>2</sub> is collected as a product. The solid also flows in the same direction as the steam. Once this step is completed the solid flow is stopped and the LR step is carried out with adsorption product. The product of the LR step is sent to the HRLRP step and the solid is now moved back into the adsorption step.

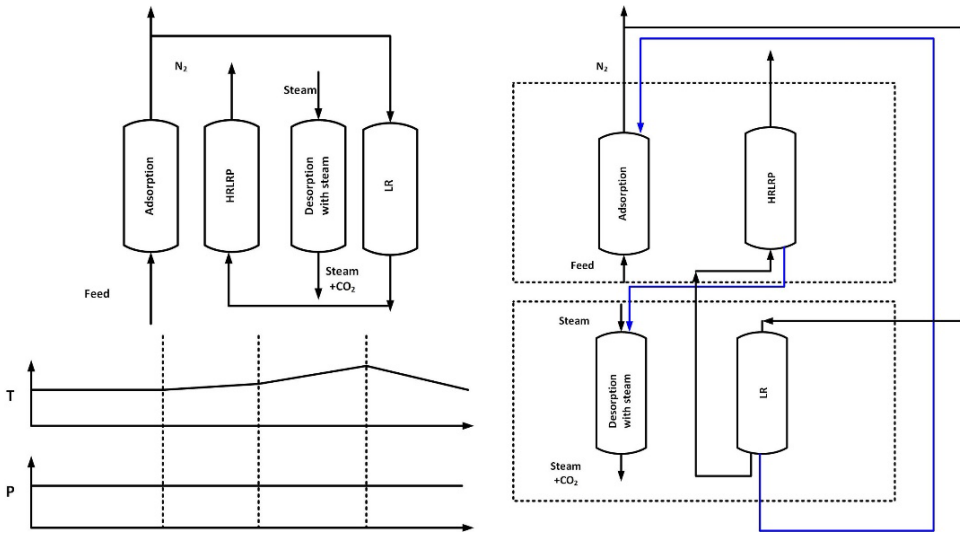


Figure 1: Schematic of the 4-step TSA process (left) and the 4-step MBTSA process (right). Solid blue lines represent the movement of solid, while black solid lines represent gas flow. The dotted lines in the figure on the right is for visual aid to separate the adsorption and desorption sections.

### 3. Model equations

The model equations used in this study are similar to the authors' earlier studies [4] on packed beds with the inclusion of additional terms to account for the solid movement. The equations are described below

$$\frac{\partial y_i}{\partial t} + \frac{y_i}{P} \frac{\partial P}{\partial t} - \frac{y_i}{T} \frac{\partial T}{\partial t} = \frac{T}{P} D_L \frac{\partial \left( \frac{P \partial y_i}{T \partial z} \right)}{\partial z} - \frac{T}{P} \frac{\partial \left( \frac{y_i P}{T} v \right)}{\partial z} - \frac{RT(1-\epsilon)}{P \epsilon} \frac{\partial \bar{q}_i}{\partial t} - \frac{T}{P} V_S \frac{\partial \left( \frac{y_i P}{T} v \right)}{\partial z} \quad (1)$$

$$\frac{1}{P} \frac{\partial P}{\partial t} - \frac{1}{T} \frac{\partial T}{\partial t} = -\frac{T}{P} \frac{\partial \left( \frac{P}{T} v \right)}{\partial z} - \frac{(1-\epsilon) RT}{\epsilon P} \sum_{i=1}^n \frac{\partial \bar{q}_i}{\partial t} - \frac{T}{P} V_S \frac{\partial \left( \frac{P}{T} v \right)}{\partial z} \quad (2)$$

$$\frac{\partial \bar{q}_i}{\partial t} = k_{LDF_i} (q_i^* - \bar{q}_i) - \frac{T}{P} V_S \frac{\partial \bar{q}_i}{\partial z} \quad (3)$$

$$\left[ \frac{(1-\epsilon)}{\epsilon} (\rho_s C_{ps} + C_{pa} \sum_{i=1}^n \bar{q}_i) \right] \left( \frac{\partial T}{\partial t} + V_S \frac{\partial T}{\partial z} \right) = \frac{K_z}{\epsilon} \frac{\partial^2 T}{\partial z^2} - \frac{c_{pg} v}{R} \frac{\partial T}{\partial z} - \frac{c_{pg}}{R} \frac{\partial P}{\partial t} + \frac{(1-\epsilon)}{\epsilon} \sum_{i=1}^n (-\Delta H_i - C_{pa} T) \frac{\partial \bar{q}_i}{\partial t} - \frac{2h_i}{\epsilon r_i} (T - T_w) \quad (4)$$

$$\rho_w C_{pw} \frac{\partial T_w}{\partial t} = k_w \frac{\partial^2 T_w}{\partial z^2} + \frac{2r_i h_i}{r_o^2 - r_i^2} (T - T_w) - \frac{2r_o h_o}{r_o^2 - r_i^2} (T_w - T_a) \quad (5)$$

$$\rho_f C_{p,f} \frac{\partial T_f}{\partial t} = -u_f \rho_f c_{p,eff} \frac{L_z}{L_x} \frac{\partial T_f}{\partial z} - \alpha_{f-HXW} h_f (T_f - T_s) \quad (6)$$

Equations 1-6 represent the component mass balance, total mass balance, adsorption rate, the energy balance, wall temperature and fluid temperature balances. The same set of equations were used for the fixed bed and the moving bed processes, with the solid velocity ( $V_S$ ) being zero in the fixed bed case. Between the hollow fiber and the packed bed, the differences are in the equations used for the pressure drop and axial dispersion terms as shown in Table 1. For the MBTSA process the velocity of the solid was obtained from the desorption/adsorption sections and the duration of the adsorption and steam purge steps. For simplicity in comparison, the length of the column was kept constant in all the simulations. This is also true for the MBTSA process where the length of the adsorption and desorption sections were fixed to be 1m. Appropriate boundary conditions



were used for the different steps along with the initial conditions for each step to simulate a cyclic process. The TSA and MBTSA cycles were simulated until cyclic steady state (CSS) condition was achieved.

The model equations were converted to a dimensionless form and discretized using the finite volume method. The resultant system of equations was solved using the ode 15s solver in MATLAB. The performance of the process was defined by 4 performance indicators, namely CO<sub>2</sub> purity, CO<sub>2</sub> recovery, energy consumption and productivity. Detailed optimization of the different cycle configurations was carried out using genetic algorithm. In total 3600 simulations were carried out for the different cases

Table 1: Correlations for axial dispersion and pressure drop in the simulations

Property	Hollow fiber	Packed bed/moving bed
Pressure drop	$\frac{-dP}{dZ} = 32 \frac{\varepsilon v \mu}{d_c^2}$	$\frac{-dP}{dZ} = 180 \left( \frac{(1-\varepsilon)^2}{\varepsilon} \right) \frac{v \mu}{d_p^2}$
Axial dispersion	$D_L = \frac{20D_{mol}}{\varepsilon} + 0.5v d_p$	$D_L = D_{mol} + \frac{(d_{channel} v_0)^2}{192D_{mol}}$

#### 4. Results and discussion

Figure 2 shows the pareto fronts for the fixed bed and the moving bed temperature swing adsorption processes. All the points on the pareto fronts satisfy 95% purity and 90% recovery constraints. The specific energy consumption is obtained from the energy consumed by the blower and compressor for the adsorption and HRLRP step and the sensible heat required to heat up the solid to the required temperature. It should be highlighted that the energy required to produce the steam is not accounted in the calculations. Based on the pareto plots, the hollow fiber TSA process was the best performing process in terms of productivity. Comparing the cycle times from Figure 3a reveal that the hollow fiber TSA process was able to operate at faster cycle times in comparison with the packed bed TSA process. This is due to the lower pressure drop of the hollow fiber process which helped to operate the cycle at higher flowrates. To minimize the CO<sub>2</sub> losses, the adsorption time is shorter and thereby the cycle time was reduced which improved the productivity. The energy consumption was higher for the hollow fibers due to the values of the regeneration temperatures chosen by the optimizer as seen from Figure 3b.

For the MBTSA process, the high energy and productivity values are a consequence of the high cycle time and higher steam temperature. In the MBTSA process, the feed velocity values were around 2.16-2.25 m/s. whereas for the packed bed it was between 1.8-2.3 m/s. The solid velocities for the MBTSA process were between 4-4.2 mm/s. The operating conditions corresponding to the minimum energy and maximum productivity values are shown in Table 2.

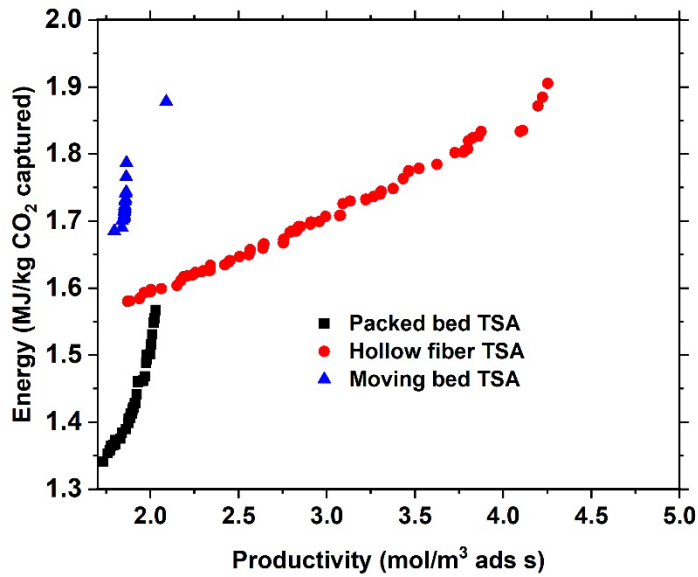


Figure 2: Specific energy vs Productivity plots for the different TSA processes

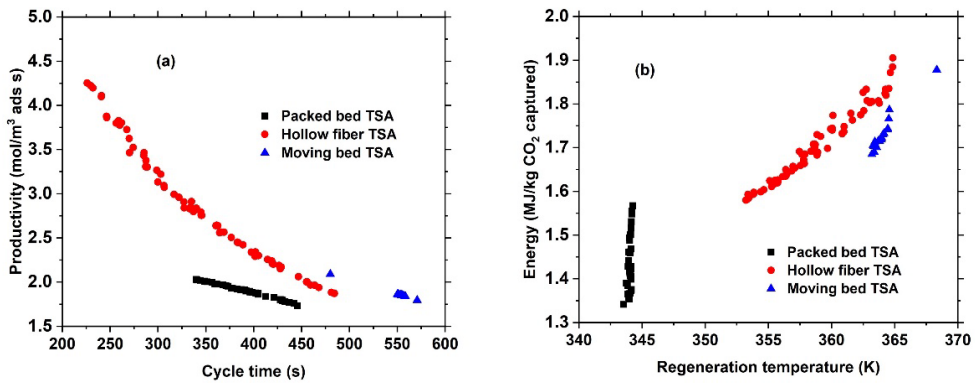


Figure 3: (a) cycle time vs productivity and (b) steam temperature vs Energy consumption

Process	$t_{ads}$ (s)	$t_{LR}$ (s)	$t_{steam-purge}$ (s)	$v_{feed}$ (m/s)	$v_{steam}$ (m/s)	$T_{steam}$ (K)
Minimum energy conditions						
Packed bed	242.10	4.93	193.55	1.79	0.99	343.53
Hollow Fiber	116.58	3.37	360.96	4.23	0.78	353.21
MBTSA	253.01	35.65	246.38	2.16	1.65	363.19
Maximum productivity conditions						
Packed bed	142.64	5.85	185.47	2.70	1.56	344.27
Hollow Fiber	111.60	1.51	111.47	4.66	2.13	364.85
MBTSA	240	20	200	2.25	1.56	368.32

## 5. Conclusions

Here we have used a homogenized model to simulate fixed bed and moving bed TSA processes. Detailed optimization of these processes revealed that the Lewatit sorbent was able to achieve 95% purity and 90% recovery values in all the three processes. The energy consumption was between 1.3-1.9 MJ/kg CO<sub>2</sub> captured for the different processes. It should be highlighted here that the MBTSA process is not the same cycle as that of Mondino et al., [5] and Morales Ospino et al., [6] where the solid moves between the adsorption, desorption and cooling sections. Further work is ongoing to improve the MBTSA process in order to simulate the existing cycles in literature. Additionally other optimization techniques like Bayesian optimization methods are also being explored by the authors

**Acknowledgements:** We acknowledge Gassnova SF and TotalEnergies E&P NORGE AS for financial support through the "Disruptive CO<sub>2</sub> capture" project (Grant nos. 620036 and 4100P20013). Discussions with Carlos Adolfo Grande are acknowledged.

## References

1. Rezaei et al., 2014, Modeling of rapid temperature swing adsorption using hollow fiber sorbents, *Chem Eng Sci*, 113, 62-76
2. Keller et al., 2019, Electrical swing adsorption on functionalized hollow fibers, *Chem Eng J*, 371, 107-117
3. Rajagopalan et al., 2016, Do adsorbent screening metrics predict process performance? A process optimisation based study for post-combustion capture of CO<sub>2</sub>, *Int J Greenh gas con*, 46, 76-85
4. Krishnamurthy, S. 2022 Vacuum swing adsorption process for post-combustion carbon capture with 3D printed sorbents: Quantifying the improvement in productivity and specific energy over a packed bed system through process simulation and optimization, *Chem Eng Sci* 253 117585
5. Mondino et al., 2020, Production of MOF Adsorbent Spheres and Comparison of Their Performance with Zeolite 13X in a Moving-Bed TSA Process for Postcombustion CO<sub>2</sub> Capture, *Ind Eng Chem Res*, 59 7198-7211
6. Morales-Ospino et al., 2021, Parametric Analysis of a Moving Bed Temperature Swing Adsorption (MBTSA) Process for Postcombustion CO<sub>2</sub> Capture *Ind Eng Chem Res*, 60, 10736-10752

# AI-enhanced BoM Optimization with MILP for the EMS Industry

Hossein Mostafaei,<sup>a,\*</sup> Eero Hiltunen,<sup>a</sup> Antti Liski<sup>a</sup>

<sup>a</sup>*AI and Machine Learning, Elisa IndustriIQ, Helsinki, Finland*

\**hossein.mostafaei@elisa.com*

## Abstract

This work explores how artificial intelligence and machine learning (ML) together with rigorous optimization models can help in optimizing electronic bill of materials. We propose a mixed-integer linear programming (MILP) model that considers different real-life restrictions required to complete a customer BoM. Furthermore, we implement a data-driven approach to find a set of alternates for a given component based on data mining and machine learning techniques. With these techniques, the BoM is completed with the most similar alternates that have the same form, fit, and functionality as the given component. The results from a real-life use case illustrate that the proposed AI-enhanced optimization model leads to superior results in terms of solution quality (lower price and shorter lead times) in comparison to the standalone MILP model, with the highlight being a 4.4% improvement in the total cost and a 24% reduction in the lead time.

**Keywords:** EMS industry, Electronic BoM, Optimization, Machine learning, MILP.

## 1. Introduction

The number of electronics used in day-to-day devices has increased in recent years which in conjunction with the global pandemic has caused disruptions in the supply chain of electronic components. According to industry experts, only one-third of components are available for order and the lead times have increased from days to several months. Obviously, this is problematic for companies whose business is dependent on sales of goods that include any electrical components. Therefore, the electronics manufacturing services (EMS) industry is looking for new ways to ensure order completion while maintaining the low costs of electronics manufacturing.

A common method for lowering costs in manufacturing industries is to use mathematical optimization. The quality, quantity, lead time, and price of goods from suppliers provide a profound impact on the overall performance of the supply chain (Pan, 2015) and can all be optimized using MILP models. The information is usually listed in the Bill of Materials (BoM) which is considered one of the critical documents for production planning (Tao and Chang, 2006). An optimized BoM can help companies to plan the purchases of raw materials at minimal prices, avoid excess amounts, manage inventory, stay on schedule, and control the budget.

Machine learning (ML) and data mining techniques have become increasingly useful in finding insights and value from a massive amount of data. ML models can learn patterns and rules automatically from the data that would be impossible for a human. For this reason, ML models are utilized in such applications where rules are not known or are too complex to understand and formulate. Data mining methods refer to techniques that

utilize data to solve problems with algorithms (Leskovec et al., 2010). This definition covers methods ranging from clustering to different types of distance metrics.

Product matching is a well-known problem in e-commerce where vendors try to find duplicate items from their site and conduct price comparisons to other sites with different descriptions and part numbers of items. Proposed solutions incorporate both text-based (Shah et al., 2018) and image-based (Kiapour et al., 2015) models to find matching products. However, none of the proposed solutions address the problem of matching products in a domain where the same form, fit, and functionality of the matches are as crucial as in the electronics supply chain.

This paper proposes a MILP enhanced with an ML-based solution to optimize an electronic BoM. The optimization model strives to minimize the cost while fulfilling all restrictions, whereas the ML solution enables sufficient search space for components that are in stock or available with a short lead time. To experiment with the effect of this, we conduct a case study with real data from the EMS industry and compare our proposed AI-enhanced optimization to a standalone optimization model.

## 2. Problem Statement

A bill of materials (BoM) is a list of the raw materials, assemblies, components, parts, and quantities that are needed to manufacture a product. BoMs can also be enriched by some other information e.g., prices, lead times, minimum quantity that can be ordered, etc. In this work, we consider an enriched electronic BoM, as shown in Table 1. The BoM features 20 line items, with some having a few alternates. For example, line item I1 indicated in the light grey is an alternate (already determined by the customer) for  $I_{EMK107BJ105KA-T}$  (with the subscript indicating the manufacturing part number (MPN)), whereas the one in the dark grey is the one we are trying to find from a massive amount of data using ML techniques, which we will discuss later in Section 3.

Table 1: Example electronic BoM considered in this work.

LI <sup>a</sup>	QTY <sup>b</sup>	Supplier	MPN	MFr <sup>c</sup>	In-stock	Min	Mult	Price (\$)	M-LT <sup>d</sup>
I1	9000	Digikey	EMK107BJ105KA-T	Taiyo Yuden	2400	500	1	0.018	250
I1	9000	Digikey	EMK107BJ105KA-T	Taiyo Yuden	2400	2000	200	0.015	250
I1	9000	Digikey	EMK107BJ105KA-T	Taiyo Yuden	2400	4000	4000	0.011	250
I1	9000	Digikey	GRM188R61C105KA12D	Murata	20000	4000	4000	0.008	300
I1	9000	Myarrow	GRT188R61C105KE13D	Murata	3000	4000	100	0.010	100
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
I20	5200	Avent	0603YC104KAT2A	Kyocera AVX	12000	2000	500	0.320	120

<sup>a</sup>Line item, <sup>b</sup>quantity required, <sup>c</sup>manufacturer, <sup>d</sup>manufacturing lead time. Note: Numbers are dummy and just for indications.

For the line item  $I_{EMK107BJ105KA-T}$ , we have  $Min = (500, 2000, 4000)$ ,  $Mult = (1, 100, 4000)$ , and  $price = (0.018, 0.015, 0.011)$ , meaning that any order for  $I_{EMK107BJ105KA-T}$  from Digikey should be a function of the Min and Mult of form  $Min + n \times Mult$  where  $n$  is a positive integer number ( $n \geq 0$ ). For example, if we want to order 2050 from  $I_{EMK107BJ105KA-T}$ , we need to order 2000 with a unit price of \$0.015 and the remaining 50 with a unit price of \$0.018. Note that all 2050 cannot be ordered with the price of \$0.015, because there is no integer number  $n$  for which  $2050 = 2000 + 100n$ . The requested quantity for I1 is 9000 and if we order all of them: i) from  $I_{GRT188R61C105KE13D}$  supplied by Digikey, then we need to order 12000 ( $4000+4000 \times 2$ ), i.e., 3000 units more than the requested quantity, which is called *excess quantity*, ii) from  $I_{GRT188R61C105KE13D}$  supplied by Myarrow, we need to wait at least 100 days (manufacturing lead time = 100) as the current in-stock (3000) is not sufficient. In this work, we will develop a MILP to consider all combinations for order quantities while minimizing BoM cost.

### 3. Models and Architecture

The following section describes the models and architecture of our proposed system. First, we define the MILP model that is used to minimize the BoM while considering the restrictions. After that, we describe our alternative search solution that can find alternative components with the same form, fit, and functionality for the reference component.

#### 3.1. MILP model

Here we develop a mixed integer linear programming (MILP) model to optimize a given BoM in terms of costs subject to some restrictions.

The binary variable  $X_{i,m,s}$  is one if component  $i$  with MPN  $m$  is ordered from supplier  $s$  (hereafter as line item  $(i, m, s)$ ). If line item  $(i, m, s)$  is not ordered, then its order quantity ( $Q_{i,m,s}$ ) should be equal to zero, as imposed by Eq. (1).

$$Q_{i,m,s} \leq M \cdot X_{i,m,s} \quad \forall (i, m, s) \in IMS \quad (1)$$

Note that the tuple  $(i, m, s)$  is based on BoM's columns (LI, MPN\_i+MFr, Supplier), in which  $m$  is built by first indexing MPN column and then combining it with MFr column. For example, for the BoM in Table 1, we have  $IMS = \{(I1, EMK107BJ105KA-T_1+Taiyo Yudan, Digikey), (I1, EMK107BJ105KA-T_2+Taiyo Yudan, Digikey), (I1, EMK107BJ105KAT_3+Taiyo Yudan, Digikey), (I1, GRM188R61C105KA12D_4+Murata, Digikey), \dots\}$ .

If line-item  $(i, m, s)$  is ordered, the order quantity should be a function of minimum ( $q_{i,m,s}^{\min}$ ) and mult ( $q_{i,m,s}^{\text{mult}}$ ) quantities. In Eq. (2)  $Z_{i,m,s}$  is a continuous integer variable.

$$Q_{i,m,s} = q_{i,m,s}^{\min} \cdot X_{i,m,s} + q_{i,m,s}^{\text{mult}} \cdot Z_{i,m,s} \quad \forall (i, m, s) \in IMS \quad (2)$$

As expressed in Eq. (3), the amount of line-item  $i$  ordered from all suppliers should be at least as large as the demanded quantity ( $dem_i$ ) by the customer, otherwise, the backordered demand (or shortage) could take a positive value and result in delays in production. Note that Eq. (3) also considers the excess amount.

$$\sum_{(m,s) \in MS | (i,m,s) \in IMS} Q_{i,m,s} + ST_i \geq dem_i \quad \forall i \quad (3)$$

Eq. (4) controls the inventory, meaning that the total amount of component  $i_m$  that can be ordered must not exceed the in-stock at supplier  $s$ .

$$\sum_{i \in I} Q_{i,m,s} \leq in_{m,s} \quad \forall (m, s) \in MS \quad (4)$$

The objective function will minimize the component cost and backorder cost. Notice the shortages coefficient  $bc = N \times \max_{m,s} \{cc_{m,s}\}$ , where  $cc_{m,s}$  is the unit cost of a component, and  $N$  is an integer number ( $N > \max_{i,m,s} |dem_i - q_{i,m,s}^{\min}|$ ).

$$\min z = \sum_{(i,m,s) \in IMS} cc_{m,s} \times Q_{i,m,s} + \sum_{i \in I} bc \times ST_i \quad (5)$$

#### 3.2. Alternatives search

Components that can be considered alternatives have the same form, fit, and functionality as the reference component. Such components are often called *Exact* or *Direct crosses*. The exact crosses are components that differ only in the manufacturer or packaging of the component, whereas direct crosses can contain alternate components that have higher specifications or compliances. Alternative components may have same or different MPN, description, and manufacturer compared to the known component.

Our alternatives search solution (Figure 1) utilizes rule-based search logic, machine learning and fundamental data mining techniques to find alternative components for reference component. Our system contains several machine learning models (e.g., random forest) and data processing steps which aim to improve the quality of the results. First, the incoming MPN is cleaned from characters that are known to be irrelevant for MPNs (Kolhase 2019) and the MPN is normalized to same form as rest of the data. This means that only lower-case characters are allowed, and extra spaces are removed. Next, the system infers component *Description* from the database of known descriptions for the MPN. If the MPN is not found in the database, it matched is approximately to the closest MPN based on character level string distance metric (Gonzalo 2001). In our experiments, Levenshtein distance worked the best.

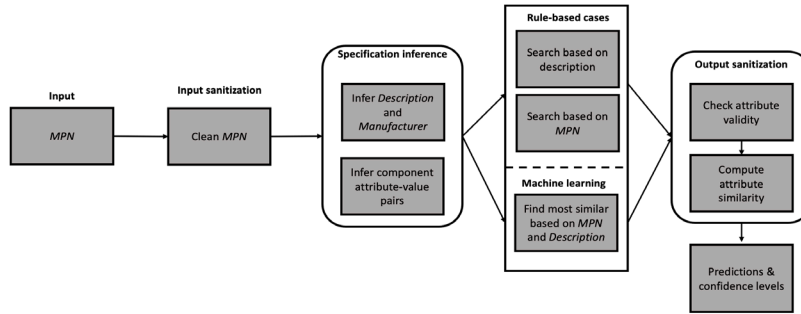


Fig. 1: Flowchart of our proposed alternatives search solution

After the component is inferred and both description and manufacturer for the MPN are found, the system extracts attribute-value pairs from the database, such as {resistance: 100-ohm, tolerance: 5%}. The inferred information is then utilized to find matching components based on *Rule-based cases* and two *Machine learning models*. The rule-based cases include: i) components that have same description but different MPN or manufacturer than the original component, and ii) components that have same MPN but different manufacturer.

Two machine learning models, one trained for description similarity and the other for MPN similarity, are utilized to find most similar components from the database of known MPNs and descriptions. The confidence of both models is averaged to obtain overall ML confidence for a candidate component.

$$Conf_{ML} = 0.5 \times (Conf_{Description} + Conf_{MPN}) \quad (5)$$

Rule-based and machine learning based matching components are then combined and output sanitization is carried out. First, components with conflicting attribute-value pairs (e.g., 1-kohm vs. 10-kohm) are removed from the set of matching components and attribute similarity is computed for each matching component. The attribute similarity is utilized in the computation of the final confidence for a matching component:

$$Conf_{Final} = 0.5 \times (Conf_{ML} + Similarity_{Attributes}) \quad (6)$$

The output is then ordered according to the  $Conf_{Final}$  which is used in the subsequent processing steps.

### 3.3. Solution procedure

Figure 2 depicts the solution procedure that consists of three steps: *preprocessing*, *alternatives addition*, and *optimization*. In the first step, the BoM is cleaned from items that are not viable in terms of client-specific restrictions, such as components from

blacklisted suppliers or with incorrect packaging type. In addition, preprocessing step removes any items that have higher lead time than the client specified maximum lead time. This decreases the size of optimization space while maintaining the quality of the results.

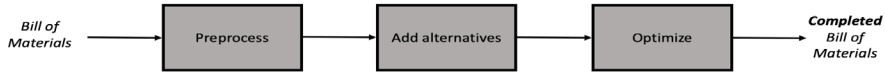


Fig 2: Illustration of three steps of our proposed solution procedure

Next, the BoM is complemented with new alternatives based on our alternatives search solution. Alternatives are queried for any component that have lower total inventory than total demand or no alternatives at all. Highest confidence alternative is picked from alternative search results that satisfies the requirements specified in preprocessing step. Finally, the MILP solver is used to complete the BoM while minimizing the cost and considering specified restrictions, such as inventory and lead time. The output of the solution is a completed BoM that may contain backorders.

#### 4. Case Study

The optimization model was implemented with a Python-based software package *Pyomo* 6.4.2 and solved using *Cbc* 2.10.8 (Coin-or branch and cut) and *GLPK* 4.32 (GNU Linear Programming Kit) open-source MILP solvers. We consider a real customer BoM with 181 components, some having a few alternates already determined by the customer. In total, the BoM consists of 5752 line items. We assume the client-specified maximum lead time is 1 year. Being solved to optimality using *Cbc* in 10s, the standalone optimization (with 0 lead time) involves 1702 constraints, and 10877 (7668) total variables (integers) and corresponds to an optimal cost of \$183695.50 and a total backorder of 16671. Note that *GLPK* could not provide any solutions in 1 hour. After adding the alternate components using the techniques discussed in section 3, the BoM ended up with 7940 line items and the optimal solution (with 0 lead time) decreased by 44.4%. However, the solution CPU time was increased from 10s to 189s, due to enlarging the solution space of the MILP model. Figure 3 and Table 2 show the computational results for the case study (with different lead times) for the standalone MILP and its AI-enhanced counterpart. Summarizing the results, one can observe from Table 2 that the completed BoM with the standalone MILP would feature a lead time of at least 300 days, whereas the AI-enhanced MILP completes the BoM with a lead time of 240 days while still providing slight improvement in the BoM cost.

Table 2: Results obtained using standalone MILP optimization and its AI-enhanced version

Lead time	Optimization		AI + Optimization		The relative improvement in cost (%)
	Backorder	Cost (\$)	Backorder	Cost (\$)	
0	16671	183695	8384	102080	44.4
30	16671	183695	8384	102080	44.4
60	16671	183695	8384	102080	44.4
90	16671	183695	8384	101817	44.5
120	13082	148201	4932	67078	54.7
150	13082	148201	4932	67078	54.7
180	11082	147142	2932	65897	54.8



210	1250	50375	100	38272	24.0
240	250	40771	0	37256	8.6
270	250	40770	0	37256	8.6
300	0	38825	0	37151	4.3
330	0	38771	0	37096	4.3
360	0	38493	0	36818	4.4

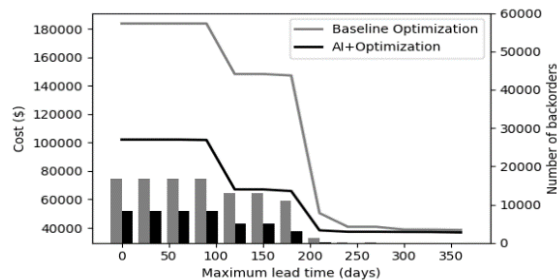


Fig 3: Comparison between standalone MILP and its AI-enhanced version

## 5. Conclusions

In this paper, we developed an AI-enhanced optimization model to complete an electronic bill of material with minimum cost and shorter lead time. The mathematical optimization is based on a mixed-integer linear programming (MILP) model that could consider different real-life restrictions required to complete a customer BoM. Furthermore, we implemented a data-driven approach to find a set of alternates for a given component with ML-powered alternative search solution. With this solution, the BoM is completed with the most similar alternates that have the same form, fit, and functionality as the given component. We considered a customer BoM and implemented the resulting optimization model in Pyomo. Compared to the stand-alone MILP model, the AI-enhanced MILP model provides better solutions in terms of solution quality and lead times but might deteriorate the computational time due to enlarging the solution space.

## References

- F. Pan, (2015), The optimization model of the vendor selection for the joint procurement from a total cost of ownership perspective, *Journal of Industrial Engineering and Management*, 8, 4, 1251-1269
- H. Tao, C. Chen, (2006), Production Planning Based on Bom Optimization, *Journal of Electronic Science and Technology of China*, 4, 1, 89-92
- J. Leskovec, A. Rajaraman, J.D. Ullman, (2010), Mining of Massive Datasets, <http://mmds.org/>
- K. Kolhase, (2019), Part Number Building Tips, <https://forum.digikey.com/t/part-number-building-tips/3460>
- K. Shah, S. Koprur, J-D. Ruvini, (2018), Neural Network based Extreme Classification and Similarity Models for Product Matching, *In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Volume 3 (Industry Papers), 8-15
- M. H. Kiapour, X. Han, S. Lazebnik, A. C. Berg and T. L. Berg, (2015), Where to Buy It: Matching Street Clothing Photos in Online Shops, *IEEE International Conference on Computer Vision (ICCV)*, 3343-3351
- N. Gonzalo, (2001), A guided tour to approximate string matching, *ACM Computing Surveys*. 33 (1): 31-88

# Reduced cleaning model for highly viscous non-Newtonian fluids in pipelines

Vera Liebmann<sup>a</sup>, Matti Heide<sup>b</sup>, Hannes Köhler<sup>b</sup>, Christian Golla<sup>a</sup>, Jochen Fröhlich<sup>a</sup>, Frank Rüdiger<sup>a</sup>

<sup>a</sup>*TU Dresden, Chair of Fluid Mechanics, George-Bähr-Str. 3c, 01069 Dresden, Germany*

<sup>b</sup>*TU Dresden, Chair of Processing Machines and Processing Technology, Bergstraße 120, 01069 Dresden, Germany*  
*vera.liebmann@tu-dresden.de*

## Abstract

To clean highly viscous fluids from pipework, flushing processes are used. The cleaning time for a pair of fluids, one being flushed and one used for flushing the other out, is typically estimated by using expensive experiments or costly fully resolved numerical simulations. In this work the authors propose a new approach to determine an accurate approximation of the cleaning time for a flushing process at substantially lower cost. This is done through a simplification of the cleaning process. The process is composed of two phases: core removal and layer removal. Sub-models for each phase are developed and then combined yielding the cleaning model. Subsequently, this model is validated using highly resolved CFD data of a cleaning process. The cleaning model is able to predict the cleaned radius over time with an error of less than 8% during core removal and less than 2% during layer removal compared to the CFD data. The computing time is negligible which makes the model highly suitable for practical application.

**Keywords:** Simulation, Cleaning, Highly Viscous Fluid, Flushing, Modelica.

## 1. Introduction

Cleaning is a vital process step during production to ensure that manufactured goods are safe for consumers and that they contain no contaminations from any other goods produced on the same line. During manufacturing of highly viscous products, such as oil, food and cosmetics, flushing processes are often employed to clean pipes of the product and at the same time replace it with a different product, thus ensuring quick product changeover, resulting in lower downtimes. The majority of such products are non-Newtonian fluids, i.e., they exhibit a viscosity dependent on the shear rate. A possible approach to classifying non-Newtonian fluids is to divide them into fluids with and without yield stress  $\tau_0$  and by shear-thinning and shear-thickening behavior. They can be modelled accordingly by using appropriate rheological models, e.g. Bingham, Herschel-Bulkley (Macosko, 1994) and Windhab model (Eischen and Windhab, 2002). For the majority of these models, analytical solutions of laminar single-phase flow exist in the literature, e.g. Bird et al. (1983) for Bingham, Kim and Lapushina (1968) and Grinchik and Kim (1972) for Herschel-Bulkley and Pitsillou et al. (2020) for Windhab.

Each flushing process is characterized by two fluids. The first is the flushed fluid, that is supposed to be displaced and the second the flushing fluid, which is used to displace the first. In the literature, flushing processes have been studied for various combinations of fluids. Hasnain et al. (2017) studied buoyant displacement of two Newtonian fluids with different densities and identical viscosity, Oladosu et al. (2020) and Swain et al. (2015)

studied the displacement of a non-Newtonian fluid by Newtonian fluid with different densities. Typically, highly complex experiments or simulations of flushing processes and their results are only transferable to systems, where the process parameters and fluids involved are similar. Since both experiments and simulations are time consuming, it is not practical to conduct and develop them for each individual flushing process. To obtain fast and yet accurate results, the current work proposes a new method, using only the fluids parameters, the geometry of the pipe and a mass flow rate as input to determine the flushing time necessary to obtain a cleaned radius  $r_c$  or a remaining mass of previous product.

The flushing process of two chocolates is selected as an example to showcase the applicability of the method. Before developing the method in more detail in section 3, the problem and its limitations are detailed in section 2. The obtained results are discussed in section 4 and finally, a conclusion is provided in section 5.

## 2. Problem description

In the present work flushed and flushing fluid are assumed to be incompressible, non-Newtonian and of similar density. The algorithm is developed and applied for the fully developed axisymmetric flow of two chocolates at  $\vartheta = 40$  °C. They exhibit non-Newtonian behavior as described by the Windhab model (Eischen and Windhab, 2002), which relates shear stress  $\tau$  and shear rate  $\dot{\gamma}$  using a specific function  $\tau(\dot{\gamma})$ .

The pipe to be cleaned has a radius of  $R = 13$  mm and a length of  $L = 0.5$  m. White chocolate initially fills the entire pipe and is then flushed out by dark chocolate. Based on experience with detailed simulations, the cleaning process is divided into two phases. First, the core material is removed (Figure 1 a). This phase is characterized by a strong dependency of the cleaned radius  $r_c$  on the position  $x$  inside the pipe. The second phase occurs, when the cleaned radius  $r_c$  is nearly independent of  $x$  and the remaining layer of the previous fluid is slowly thinned out.

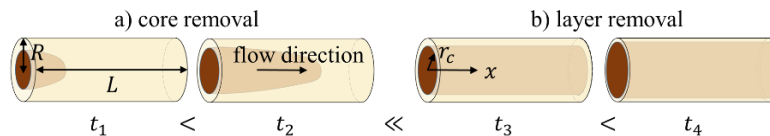


Figure 1 Schematic of the considered flushing process. a) first phase of cleaning – core removal; b) second phase of cleaning – layer removal

Following the distinction of the two phases, the authors first develop a model for core removal and then for layer removal, which is done in sections 3.1 and 3.2, respectively. A condition for the transition between the two phases is described in section 3.2.3.

## 3. Reduced Model

### 3.1. Core removal

The velocity profile  $u(r)$  for an axisymmetric flow of a single fluid with Windhab viscosity can be derived analytically from the Navier-Stokes equations using the approach by Pitsillou et al. (2020). The Windhab model contains a parameter  $\tau_0$  to describe the yield stress of the fluid. For  $\tau_0 > 0$  a plug flow region forms in the center of the pipe, i.e., for  $r \leq r_0$ , where  $r_0 = 2\tau_0/f$  is called the plug flow radius, with  $f = \Delta p/L$  being the pressure loss over pipe length. Starting from the wall, where a no-slip boundary condition is applied,  $u(r = R) = 0$ , the velocity  $u(r)$  increases reaching a maximum velocity  $u_{max}$  at  $r = r_0$ , which remains constant until the center of the pipe at  $r = 0$ . Starting from  $r =$

$R$ , the velocity gradient  $u'(r)$  decreases, from the maximum at the wall to zero within the plug flow radius  $r \leq r_0$ .

For  $f \leq 2\tau_0/R$  no flow develops in the pipe, because the yield stress of the fluid is not overcome. With  $f$  increasing above that threshold the mass flow rate  $\dot{m}$  and the maximum velocity  $u_{\max}$  continuously increase, while the plug flow radius  $r_0$  decreases.

The convective transport of a scalar  $\alpha$  in one dimension can be described by

$$\partial_t \alpha + u_{\text{tr}} \partial_x \alpha = 0, \quad (1)$$

with the transport velocity  $u_{\text{tr}}$ . The method of characteristics may be used to solve this partial differential equation. For this paper only a characteristic of eq. (1), along which  $\alpha = \text{const.}$  is used and given by

$$x = u_{\text{tr}} t. \quad (2)$$

Eq. (1) can be used to describe the core removal in a flushing process, by using  $\alpha$  as a phase indicator. Then  $\alpha(x, t) = 1$  indicates the presence of the first fluid at time  $t$  and location  $x$ , while  $\alpha = 0$  indicates the absence of the first fluid and presence of the second fluid. Assuming both fluids have identical physical properties, the pipe flow is axisymmetric, laminar and no diffusion between the two fluids occurs, the transport velocity  $u_{\text{tr}}$  in eq. (2) can be replaced by the analytical solution  $u(r)$  of the flushing fluid described above. The interface between the two fluids is identified here with  $\alpha = 0.5$ . The cleaned radius  $r_c$  for a given location  $x_0$  along the flow axis  $x$ , can then be determined using with  $x_0 = u(r_c) t$  (see eq. (2)). Defining a non-dimensional time using  $\hat{t} = u_{\max} t / x_0$  with the maximum velocity  $u_{\max}$  and the spatial coordinate  $x_0$  yields a universal curve for  $r_c(\hat{t})/R$  as shown in Figure 2.

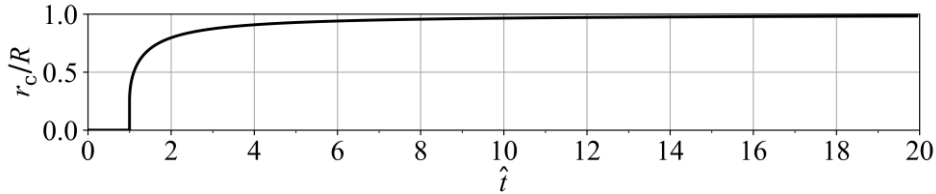


Figure 2 Cleaned radius  $r_c/R$  over dimensionless time  $\hat{t} = u_{\max} t / x_0$ .

Cleaning starts after an offset time  $\hat{t} = 1$ . This is the time needed for the plug flow region of the flushing fluid traveling at  $u_{\max}$  to reach the cross section at  $x_0$ . When the plug flow region passes, the cleaned radius instantly increases from  $r_c = 0$  to  $r_c = r_0$ . From then on, the cleaned radius increases further. The rate of change, however, decreases since  $u(r)$  decreases towards the wall.

### 3.2. Layer removal

#### 3.2.1. Basic algorithm

Once the core has been removed, the remaining layer is transported away. A linear velocity profile is assumed in the remaining layer, i. e. for  $r_c < r < R$

$$u_L(r) = u_c \left( 1 - \frac{r_c - r}{r_c - R} \right), \quad (3)$$

where  $u_c$  is the velocity on the interface located at  $r_c$ . Integrating eq. (3) over the cross section yields the mass flow rate in the layer  $\dot{m}_L$  as shown in eq. (4)

$$\dot{m}_L = 2\pi\rho \int_{r_c}^R u_L(r) r dr = \frac{\pi}{3} \rho u_c (R - r_c) (R + 2r_c) \quad (4)$$

The mass of the fluid in the remaining layer  $m_L$  is calculated using eq. (5).

$$m_L = \pi\rho(R^2 - r_c^2)L \quad (5)$$

Given these three equations, a simple algorithm to determine the cleaned radius  $r_c$  over time is derived. The problem is initialized with radius  $r_c^0 = r_c(t = t_0)$ . The determination of the starting value is discussed in section 3.2.3. Then, four steps are performed for each time step  $t^n$ . First, the velocity  $u_c^n$  is determined from the given cleaned radius at that time  $r_c^n$ , as detailed in section 3.2.2. Second, the mass flow rate  $\dot{m}_L^n$  is determined using eq. (4). Third, the mass of the remaining layer at the next time step is calculated from  $m_L^{n+1} = m_L^n - \dot{m}_L^n(t^{n+1} - t^n)$ . Finally, the height of the remaining layer at the next time step,  $r_c^{n+1}$ , is determined by rewriting eq. (5). These steps are repeated until a desired radius is cleaned or a maximum time is reached.

### 3.2.2. Determining $u_c$

Figure 1 shows that, during layer removal the radius  $r_c$  is nearly independent of  $x$ , so that the flushing fluid is assumed to create a cylinder with radius  $r_c$  concentric to the pipe with radius  $R$ . Assuming a smooth interface between the two fluids, the flow of the flushing fluid can be treated like the flow inside a second pipe with a pipe radius equal to  $r_c$  and a velocity at the wall  $u(r_c) = u_c$ . The analytical solution by Pitsillou et al. (2020) is modified accordingly to yield the velocity profile  $u_{fl}(r)$  inside the flushing fluid. Finally, the value of  $u_c$  is determined from the condition that the shear stress at the interface is the same for the remaining layer of flushed and flushing fluid, i.e.,  $\tau_L(r_c) = \tau_{fl}(r_c)$ . This implies a smooth interface between the phases. The shear stress at the interface is determined via the Windhab law  $\tau(\dot{\gamma})$ . For the flushing fluid the shear rate  $\dot{\gamma}_{fl}$  obtained by differentiating the velocity  $u_{fl}$  in the core region is used. A second expression for the shear stress at the interface is obtained from using the shear rate obtained from the linear velocity profile in the remaining layer (eq. (3)) to  $\dot{\gamma}_L$ . Combining both expressions yields the unknown  $u_c$ .

### 3.2.3. Defining a starting value for $r_c$

The initial value of  $r_c$  to start the algorithm in section 3.2.1 must be sufficiently close to  $R$  and sufficiently far away from the plug flow radius  $r_0$  to justify the approximation of a linear velocity profile inside the remaining layer. To find the starting value  $r_c^0$  the analytical solution for the mass flow  $\dot{m}_{L,an}$  and the solution found by the linear approximation  $\dot{m}_L$  from (4) are compared. The former is found by integrating the analytical solution for the velocity profile  $u(r)$  from Pitsillou et al. (2020) to

$$\dot{m}_{L,an} = 2\pi\rho \int_{r_c}^R u(r) r dr. \quad (6)$$

Figure 3 shows the relative error  $\varepsilon = \dot{m}_L/\dot{m}_{L,an}$  for different total mass flow rates. Each mass flow rate corresponds to a different plug flow radius  $r_0$ , which is related to the cleaned radius  $r_c$ . The latter is varied as the choice affects the error  $\varepsilon$ . If  $r_c/r_0 \approx 1$ , the

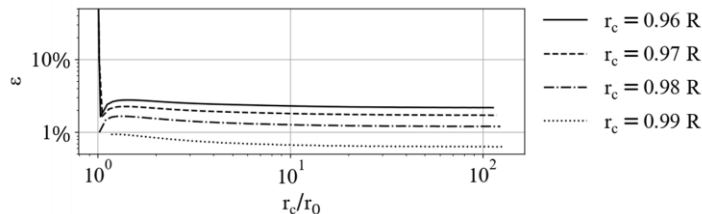


Figure 3 Influence of the choice of  $r_c$  on the error  $\varepsilon$  of the mass flow rate in the remaining layer that arises due the linear approximation of the velocity there. Parameters from Liebmann et al. (2022).

plug flow radius and the cleaned radius are approximately equal. In that case the error  $\varepsilon$  is high for all values of  $r_c$ . This is due to the fact, that at the plug flow radius  $r_0$  the velocity  $u(r)$  transitions from the high gradient at the wall to zero gradient within the plug flow region, which is not well approximated by a linear function. For values of  $r_c/r_0 \gg 1$ , the error is smaller, since the cleaned radius is further away from the plug flow radius, where the velocity gradient transitions to zero. The higher the ratio  $r_c/r_0$  and the lower the value for  $r_c$  the more pronounced this effect becomes. For the second phase, the authors propose an initial cleaned radius of  $r_c^0 = 0.97 R$  which provides errors  $\varepsilon < 2\%$  in this configuration. For more accurate results a higher value of  $r_c$  may be chosen.

### 3.3. Implementation in Modelica

The two-stage model described above was implemented in OpenModelica v1.19.2. The input parameters are the radius  $R$ , the length  $L$  of the pipe and the mass flow rate of the flushing fluid. For the first phase, the physical properties of the flushing fluid are used. For the second phase the flushed and flushing fluids properties are used. The values for the two chocolates are taken from Liebmann et al. (2022). The mass flow rate is used to determine the analytical solution of the Windhab velocity profile  $u(r)$ . The cleaned radius  $r_c(t)$  during the core removal phase is determined according to section 3.1. Once the cleaned radius  $r_c^0 = 0.97 R$  described in section 3.2.3 is obtained during core removal, the algorithm in section 3.2.1 is used for layer removal.

## 4. Results

Figure 4 depicts the results obtained from a fully resolved OpenFOAM simulation, which took 20 hours to complete on 16 CPUs (see Liebmann et al., 2022) and the results obtained from the reduced model in OpenModelica. The latter required only 0.05 seconds on a single CPU to complete the calculation. Core removal starts when the plug flow region reaches the cross section at the considered position  $x$ . The time needed for that is shorter for smaller distances  $x$ . After that,  $r_c$  increases instantly by the plug flow radius and then continually grows. The cleaned radius in the OpenFOAM simulation is up to 8% higher during core removal. This may be due to additional effects at the interface not considered within the reduced model or because the physical properties of the two fluids are similar but not equal. At  $r_c = 0.97 R$  layer removal begins. During this phase, the reduced model shows very good agreement with the OpenFOAM simulations both in quality and quantity. The reduced model predicts  $r_c$  up to 2% smaller than the results from OpenFOAM. The model in OpenModelica runs about seven orders of magnitudes faster, offering a considerable speedup compared to the fully resolved simulation, with little error arising from the use of the reduced model, especially during layer removal.

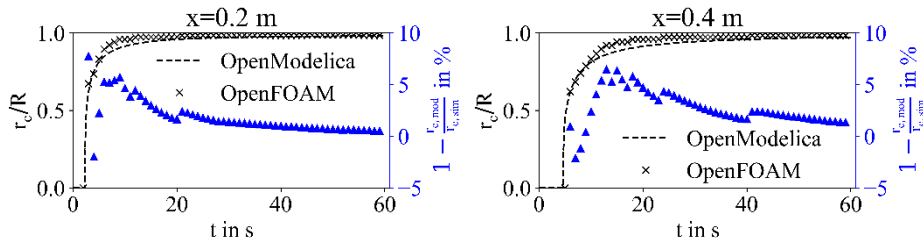


Figure 4 Comparison of results for the cleaned radius  $r_c(t)$  obtained from a fully resolved simulation using OpenFOAM and the reduced model in OpenModelica introduced in this paper. The blue triangles with the right axis indicate the relative error between both results.

## 5. Conclusion and Outlook

The paper proposes an approach for a reduced model of a flushing process of two non-Newtonian fluids with yield stress in a straight pipe. The flushing process is divided into two phases. During the first phase (core removal) the presented results for the cleaned radius  $r_c(t)$  are within 8% of the results obtained using a fully resolved numerical simulation. During the second phase (layer removal) this error is further reduced. The reduced model implemented in OpenModelica offers accurate results after less than a second. Compared to the current state of the art, i.e. conducting fully resolved simulations or experiments, which are time consuming, expensive and complex, the model can be applied in the rapid design of industrial flushing processes. It offers a method of designing new or adapting old processes to new process parameters while reducing the time and cost compared to current methods. Processes not adapted so far due to excessive cost of classical methods can now be optimized, thus saving resources, costs and time.

The approach presented for layer removal is only applicable if the velocity profile in the remaining layer is approximately linear. For cases where the layer is thicker, the authors were able to successfully employ an alternative approach to determine  $u_c$  from the analytical solution of the velocity within the remaining layer. However, this approach is specific to a particular combination of rheological models for the flushed and flushing fluid and, hence, not generally applicable. For these cases and for  $r_c < r_0$  further development of the presented method is necessary. The approach can be extended for other combinations of fluids using the literature provided in section 1. This requires two adaptations of the present model: one for the velocity profile of the flushing fluid and one for the rheological behavior  $\tau(\dot{\gamma})$  of the two fluids. The method has already been employed for a set of different fluid combinations. Further tests will be conducted to determine the limitations in the ratio of the two fluids physical properties.

## References

- R.B. Bird, G.C. Dai, B.J. Yarusso, 1983, The rheology and flow of viscoplastic materials. *Reviews in Chemical Engineering*, 1, 1, 1-70
- J.-C. Eischen, E. J. Windhab, 2002, Viscosity of cocoa and chocolate products, *Applied Rheology*, 12, 1, 32-34
- I.P. Grinchik, A.K. Kim, 1972, Axial flow of a nonlinear viscoplastic fluid through cylindrical pipes. *Journal of Engineering Physics*, 23, 2, 1039–1041
- A. Hasnain, E. Segura, K. Alba, 2017, Buoyant displacement flow of immiscible fluids in inclined pipes, *Journal of Fluid Mechanics*, 824, 661-687
- A.K. Kim, B.I. Lapushina, 1968, The motion of a liquid with variable rheological characteristics in a circular cylindrical tube, *Journal of Engineering Physics*, 14, 2, 159–161
- V. Liebmann, M. Heide, K. Schoppmann, H. Köhler, J. Fröhlich, F. Rüdiger, 2022, Aspects of modelling the cleaning of a chocolate with yield stress in a pipe using CFD, *Fouling and Cleaning in Food Processing 2022 Proceedings*
- C. W. Macosko, 1994, *Rheology – Principles, Measurements and Applications*, Wiley-VCH, 1
- O. Oladosu, J. Bhakta, K. Alba, 2020, Density stable yield-stress displacement flow of immiscible fluids in inclined pipes, *Journal of Non-Newtonian Fluid Mechanics*, 275, 4, 104203
- R. Pitsillou, G.C. Georgiou, R.R. Huilgol, 2020, On the use of the Lambert function in solving non-Newtonian flow problems, *Phys. Fluids*, 32, 9, 093101
- P.A.P. Swain, G. Karapetsas, O.K. Matar, K.C. Sahu, 2015, Numerical simulation of pressure-driven displacement of a viscoplastic material by a Newtonian fluid using the lattice Boltzmann method, *European Journal of Mechanics B/Fluids*, 49, A, 197-207

# Optimal Retrofit of Simple Distillation Sequences to Thermally Coupled Side-Stream Configurations

Anna S. Horsch,<sup>a</sup> Andres R. Acevedo,<sup>a</sup> Mirko Skiborowski<sup>a</sup>

*<sup>a</sup>Hamburg University of Technology - Institute of Process Systems Engineering, Am Schwarzenberg-Campus 4, 21073 Hamburg, Germany  
Mirko.Skiborowski@tuhh.de*

## Abstract

Retrofitting existing plant equipment to a more energy efficient setup is a promising and sustainable way to cut down on the energy demand and consequently the operating costs of a chemical processing plant. While this opportunity is attractive in terms of low capital expenditure and construction time, it is also highly challenging for the design engineers. Degrees of freedom in grassroot design transform into constraints in the retrofit project and finding an optimal design within the constraints of the existing equipment, requires a sophisticated model-based approach. In this work, such an approach is presented for the retrofit of distillation column sequences to thermally coupled side-stream configurations. Based on this approach, an easy evaluation of retrofit opportunities in terms of investment and operating costs is facilitated. The potential of this retrofit is illustrated in the current contribution by means of a case study, which compares the retrofit option to a fully thermally coupled dividing wall column grassroot design. While the dividing wall column generally allows for higher energy savings, the investment in the retrofit is significantly smaller, which yields a considerably quicker amortization.

**Keywords:** Distillation, Thermal Coupling, Retrofit, Optimization

## 1. Introduction

In light of rapidly increasing energy and resource prices as well as sociopolitical pressure due to the quickly progressing climate crisis, the arguably biggest goal of the chemical industry is to become more energy-efficient and sustainable. A focus in this effort lies on distillation, the most frequently applied fluid separation technique in the chemical industry, which is accountable for a major share of the energy consumption in chemical plants (Sholl and Lively 2016). A well-established approach to reduce the energy demand of distillation sequences is the thermal coupling of distillation columns in the form of side rectifiers (SR), side strippers (SS) and fully thermally coupled dividing wall columns (DWC). Out of these three configurations, the DWC is unmatched in terms of operating costs and investment, which sets the DWC as one of the frontrunners of distillation innovation with more than 300 industrial applications (Lukač et al. 2019). However, constructing an energy and cost-efficient grassroot design is not the only promising option in the effort to re-engineer the chemical industry. Retrofitting existing plant equipment can facilitate a considerable decrease in energy consumption and related operating costs, combined with the opportunity to increase process capacity, with comparably little capital expenditure, effort and construction time. This becomes possible due to commonly applied design margins on plant equipment (Rangaiah 2016). Retrofitting and revamping projects account for the major share of design projects



conducted in the chemical industry, offering potential for innovative retrofit options and optimal design methods (Rangaiah 2016).

One particularly interesting concept in that context is the so-called Liquid-Only-Transfer (LOT) configuration proposed by Agrawal (2000), which transforms the bidirectional vapor and liquid transfer in thermal couplings to an unidirectional liquid transfer. This can be achieved for SR and SS configurations by extending the side-columns with an additional column section and heat exchanger, which facilitates the elimination of the intercolumn vapor transfer. The resulting LOT configuration provides an excellent retrofitting option of an existing column sequence through integration of an additional side stream.

Although we are not aware of a dedicated industrial implementation of the LOT concept, it offers several advantages over the classical thermal coupling. Here, it is important to note that the LOT-configurations are thermodynamically equivalent to their thermally coupled counterparts (Madenoor Ramapriya et al. 2014) and as such hold the same potential energy savings, which can amount up to 30 % (Amminudin et al. 2001) for a ternary separation. In addition to these energy savings, the elimination of the vapor connection between the column allows for the hydrodynamic decoupling of both columns, enabling the application of thermal coupling to columns operated at different pressures. This presents the option for further heat integration in the thermally coupled columns through targeted pressure variation, providing potentially significant energy savings (Skiborowski 2020). Furthermore, replacing the generally uncontrolled vapor stream or vapor split with a controllable manipulated variable like the heat duty or reflux ratio, allows for a more precise control of the internal vapor streams and as such supports the control at the energy optimal operating point (Madenoor Ramapriya et al. 2014). Lastly, the structural similarity of LOT-sequences to uncoupled (in)direct sequences consisting of two regular distillation columns connected via a liquid stream, makes them perfect candidates for retrofitting simple distillation sequences with minimal effort.

In this contribution, a framework for the optimal retrofit of an uncoupled distillation sequence to thermally coupled LOT-side stripper and side rectifier configurations is presented. While these do generally not allow for the same energy savings as a fully thermally coupled DWC, they are oftentimes thermodynamically more efficient than the DWC (Agrawal and Fidkowski 1998). However, exploiting these benefits in an optimal retrofit design is a complex task, since structural constraints for the given equipment, including column height, vapor capacity, pressure drop and available heat exchanger area, have to be adhered in the design process. The current contribution proposes a superstructure based design approach, which builds on a modified mixed integer nonlinear programming (MINLP) problem originating from the grassroot design of column sequences as presented by Waltermann and Skiborowski (2019). In an exemplary case study, the performance of this developed methodology is showcased, addressing the question if a retrofit design can compete with an optimally designed fully thermally coupled grassroot DWC.

## 2. Methods

In this chapter, the concept of thermally coupled Liquid-Only-Transfer distillation sequences is explained based on the example of the side rectifier configuration. Subsequently, the optimization procedure for the grassroot design and the retrofit is described.

### 2.1. Liquid-Only-Transfer (LOT) Sequences

As introduced by Agrawal (2000) and theoretically proven by Madenoor Ramapriya et al. (2014), a thermodynamically equivalent LOT configuration can be developed through elimination of the vapor transfer in thermal couplings in three steps: First, the stripping section of the second column ( $S2$ ) including the reboiler is duplicated ( $S2^*$ ) and placed at the bottom of the first column, where the recycle vapor stream ( $V$ ) enters the first column. As the second step, the intercolumn vapor stream  $V$  is redirected to originate from the duplicated stripping section ( $V^*$ ). Lastly, the duplicated section  $S2^*$  is connected to the stripping section of the first column ( $S1$ ). This procedure is visualized in Figure 1. The development of the LOT-side stripper and the LOT-DWC is done in a similar manner.

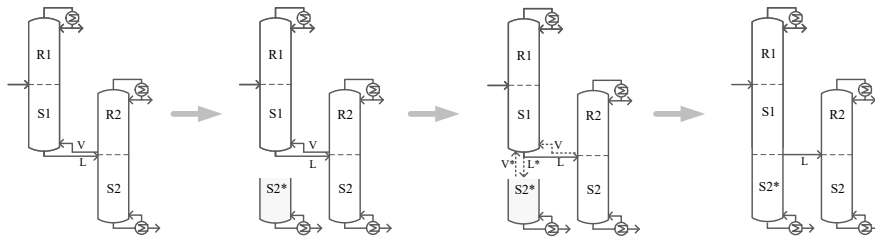


Figure 1: Development of thermodynamically equivalent LOT-side rectifier

### 2.2. Optimization-based design

Both the optimal grassroot design as well as the optimal retrofit design are obtained based on an equation-oriented superstructure-based optimization approach implemented in GAMS 31.2.0. The superstructure builds on an equilibrium-stage model of the distillation columns on the basis of the MESH-equations (**M**ass balance, **E**quilibrium conditions, **S**ummation constraint, **E**nthalpy balance). The thermodynamic properties are calculated by means of an external function programmed in C which communicates with GAMS via a dynamic link library, while the optimization problems are solved as a series of successively relaxed MINLP problems (Skiborowski et al. 2015). The grassroot design procedure is described in detail by Waltermann and Skiborowski (2019). For retrofit design this approach is extended to account for pressure drop and separation efficiency and is subsequently adapted to calculate optimal retrofit designs for a predefined column sequence with given height and diameter.

#### 2.2.1. Modelling of pressure drop and separation efficiency

In the original MESH-based grassroot design model described by Waltermann and Skiborowski (2019), the pressure drop is assumed to be negligible and the height equivalent to one theoretical plate (HETP) is considered to be constant. The column height and diameter are determined in accordance with the required number of equilibrium stages and the vapor flowrates. While this is a reasonable simplification for conceptual grassroot designs, the pressure and separation efficiency need to be modelled more precisely for retrofit calculations, considering the utilization of existing equipment and the resulting changes in gas and liquid loads.

The pressure drop per height ( $\Delta p/h$ ) as well as the separation efficiency, which is represented by the HETP, are calculated based on the vapor load. The vapor load is described by the F-factor  $FF = \dot{V}/A_{column} \cdot \sqrt{\rho_V}$  which is calculated as the product of the superficial gas velocity, given as the ratio of vapor  $\dot{V}$  per column section  $A_{column}$  and the square root of the vapor density  $\rho_V$ .

In this work, the properties of the Montz B1-500 packing are utilized. The correlation for the pressure drop is approximated with a third-degree polynomial  $\Delta p/h = 0.83F^3 - 1.3F^2 + 1.6F - 0.3$  ( $R=0.999$ ) and the *HETP* is approximated with a second-degree polynomial  $HETP = 0.01F^2 + 0.01F + 0.24$  ( $R=0.999$ ). Both correlations are fitted to the vendor given and packing specific pressure drop and separation efficiency diagrams (Julius Montz GmbH 2022).

### 2.2.2. Optimal Retrofit

Feed properties, purity constraints and a maximum stage number of the superstructure are input information for both the grassroot design procedure as well as the retrofit design. Furthermore for the retrofit, structural constraints for the available column heights, column diameters and heat exchanger areas are considered. The model initialization and heat duty minimization is performed based on initial flash calculations and successive model refinement as described by Waltermann and Skiborowski (2019).

The grassroot design and the retrofit design optimization differ as described in the following. For the grassroot design, the total annualized costs (*TAC*)

$$TAC = C_{op} \cdot t_{op} + C_{inv} \cdot \alpha \quad (1)$$

are minimized, considering an optimization of the equipment sizing and operating conditions, based on variable column heights, feed and side stream locations and heat duties. The *TAC* include the operating costs ( $C_{op}$ ) considering heat duties and utility prices with an annual operation time ( $t_{op}$ ) of 8000 h, and the investment costs ( $C_{inv}$ ) which are based on a correlation from Biegler et al. (1997) with a depreciation factor  $\alpha = f(z, t_a) = 0.37$ , based on an interest rate ( $z$ ) of 6 % and depreciation time ( $t_a$ ) of 3 years. In the retrofit procedure only the operating costs  $C_{op}$  are minimized, if the existing equipment suffices. Therefore, the column structure is optimized on the basis of the superstructure model such that structural constraints on the available column height and diameter are obeyed and variable side stream and feed location result in maximum energy savings. In order to allow for more flexibility, a rearrangement of the columns for the LOT sequence is considered in the design process. If based on the resulting design the heat exchanger area of the available heat exchangers is insufficient, an additional, parallel heat exchanger is considered and depreciation costs are added to the objective. The heat exchanger costs are calculated based on the same correlation as in the grassroot case. The main steps of the grassroot design and the retrofit procedure are summarized in Figure 2.

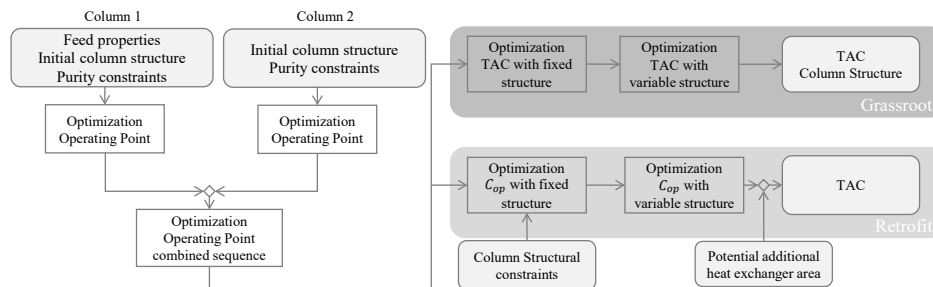


Figure 2: Modeling and solution approach for grassroot design and retrofit; “Optimization Operating Point” for individual columns and combined sequence according to Waltermann and Skiborowski (2019)

## 3. Case Study

As previous shortcut screening calculations highlighted the importance of feed compositions for the potential energy savings (Skiborowski, 2020), two exemplary feed compositions are evaluated for the separation of benzene, toluene and xylene, an

equimolar feed  $F_1$  and a second feed with low amount of mid boiling component  $F_2=(0.45/0.10/0.45)$  mol/mol. For both cases, a saturated liquid feed of 1 kmol/s at 1 atm shall be separated with product purities of at least 0.99 mol/mol. The liquid phase activities are modelled with the NRTL (Non-Random-Two-Liquid) model, while the vapor phase is assumed to behave ideal.

As a basis for the retrofit, an optimal grassroot design of a column sequence for the separation task is calculated. Since the direct sequence results in lower TAC than the indirect sequence, it is used as the reference case for both case studies. Design safety factors are applied to the height (+10 % stages), the cross-sectional area (80% of flooding velocity) and the heat exchanger area (+10 %) (Rangaiah 2016). The operating costs of this uncoupled sequence are the benchmark that will be compared to the grassroot DWC design and the retrofit design of the LOT-side rectifier. The costs for the direct sequence, as well as the retrofit and DWC for both study cases are summarized in Table 1.

Table 1: Summary of operating costs, investment, total annualized costs and possible amortization period for the initial direct sequence, the grassroot DWC and the retrofit LOT-SR.

		Operating Cost $C_{op} \cdot t_{op}$ [k€]	Investment $C_{inv}$ [k€]	TAC [k€]	Amortization [a]
Case 1 $F_1$	DS	133	466	307	-
	Retrofit	124	18	131	2.2
	DWC	93	394	240	9.9
Case 2 $F_2$	DS	112	420	269	-
	Retrofit	90	30	101	1.3
	DWC	74	361	209	9.5

For the grassroot design DWC, both operating and investment cost and consequently the TAC are below that of the direct sequence for both cases. To guarantee comparability, the DWC design also contains safety factors of the same magnitude. The grassroot DWC is excellently suited for both cases with 30 % ( $F_1$ ) and 34 % ( $F_2$ ) of operating cost savings, which equals a saving potential  $S$  of 40 k€/year ( $F_1$ ) and 38 k€/year ( $F_2$ ). Based on the necessary investment, this yields an amortization period

$$A = C_{inv}/S \tag{2}$$

of 9.9 years and 9.5 years, respectively. Thus, it takes almost 10 years until the 30 %-34 % savings in operating costs amount to the necessary investment for the DWC.

For the retrofit LOT-side rectifier designs, investments of 18 k€ and 30 k€ are determined based on additional heat exchanger area that needs to be added for the retrofit. The retrofit reduces the operating costs by 6 % ( $F_1$ ) and 20 % ( $F_2$ ).

As to be expected, the energy saving potential of the retrofit remains lower than that of the fully thermally coupled DWC. However, since the necessary investment is much lower, the amortization period is noticeably shorter for the retrofit in both cases with only 2.2 years and 1.3 years, respectively. As a consequence, the TAC of the retrofit are considerably below the TAC of the DWC for both cases. The setup of the direct sequence, the retrofitted LOT-sequence and the DWC for the equimolar case  $F_1$  is visualized in Figure 3.

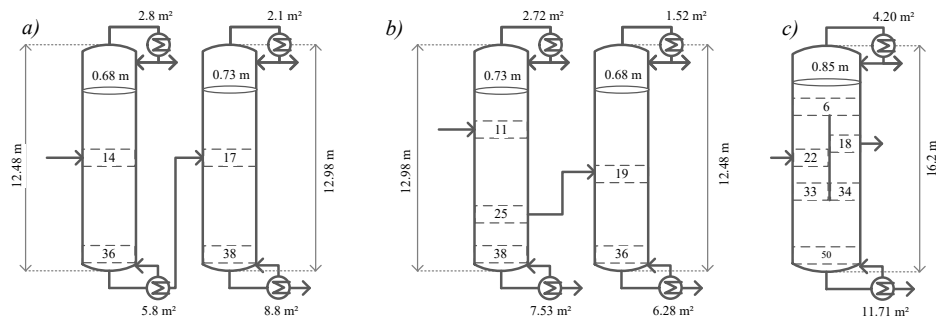


Figure 3: Results of (a) the grassroot design of the direct sequence, (b) the resulting retrofitted LOT-sequence and (c) the grassroot design of the DWC.

#### 4. Conclusion

The current work presents an optimization-based methodology for the retrofit of distillation sequences to thermally coupled side stream configurations and illustrates the benefits of these configurations for the separation of two exemplary feed compositions of a benzene-, toluene- and xylene mixture. While the implementation of a new dividing wall column offers larger energy savings than the simple retrofit, the drastically lower investment of the side stream retrofit results in much lower pay out times. Considering the frontrunner status of the dividing wall column, this showcases the potential of such Liquid-Only-Transfer configurations when used for retrofit. For a final implementation of such a retrofit option further cost factors for reassembly and lost earnings for plant shutdown, which depend on the given plant, should be considered. However, the resulting retrofit configuration also enables further potential for process intensification by means of heat integration, which can even exceed the energy savings that are enabled by dividing wall columns, as has been illustrated by Skiborowski (2020).

#### 5. References

- Agrawal (2000): *AIChE J.* 46 (11), pp. 2198–2210.  
 Agrawal; Fidkowski (1998): *Ind. Eng. Chem. Res.* 37 (8), pp. 3444–3454.  
 Amminudin; Smith; Thong; Towler (2001): *Chem. Eng. Res. Des.* 79 (7), pp. 701–715.  
 Biegler; Grossmann; Westerberg, (1997): Upper Saddle River, NJ: Prentice-Hall (Prentice-Hall international series in the physical and chemical engineering sciences).  
 Julius Montz GmbH (20.10.2022): Separation Efficiency and Pressure Drop - Montz B1-500. Available online at <https://www.montz.de/standard-blechpackung-typ-b1#&gid=1033229313&pid=2>.  
 Lukač; Halvorsen; Olujić; Dejanović (2019): *Chem. Eng. Res. Des.* 79 147, pp. 367–377.  
 Madenoor Ramapriya; Tawarmalani; Agrawal, (2014): *AIChE J.* 60 (8), pp. 2949–2961.  
 Rangaiah (Ed.) (2016): Chichester, UK: John Wiley & Sons, Ltd.  
 Sholl; Lively (2016): *Nature* 532 (7600), pp. 435–437.  
 Skiborowski (2020): 30th European Symposium on Computer Aided Process Engineering, vol. 48: Elsevier (Computer Aided Chemical Engineering), pp. 991–996.  
 Skiborowski; Harwardt; Marquardt (2015): *Comput. Chem. Eng.* 72, pp. 34–51.  
 Waltermann; Skiborowski (2019): *Comput. Chem. Eng.* 129, p. 106520.

# Computer aided-design of energy efficiency tools on a microalgae biorefinery scheme

Araceli Guadalupe Romero-Izquierdo <sup>a</sup>, Claudia Gutiérrez-Antonio <sup>a</sup>, Fernando Israel Gómez-Castro <sup>b</sup>, Salvador Hernández <sup>b\*</sup>, Juan Fernando García-Trejo <sup>a</sup>

<sup>a</sup> *Facultad de Ingeniería, Universidad Autónoma de Querétaro, Cerro de las Campanas S/N, Querétaro, Querétaro. 76010, México.*

<sup>b</sup> *Departamento de Ingeniería Química, Universidad de Guanajuato, Noria Alta S/N, Guanajuato, Guanajuato. 36050, México. [hernasa@ugto.mx](mailto:hernasa@ugto.mx)*

## Abstract

The search of new alternatives of processing for the production of sustainable products, designed under energy efficiency tools, has gained the interest of scientific community. In this work is proposed the conversion of a microalgae (*Chlorella vulgaris*) through a biorefinery scheme to produce biojet fuel, other biofuels and value-added products. Process intensification and energy integration tools were applied as energy efficiency strategies to improve the system. The biorefinery is modelled in Aspen Plus. The application of both strategies increases the equipment cost up 7 %, but reduces in 70 % the thermal energy consumption and 45% the CO<sub>2</sub> emissions released.

**Keywords:** microalgae biorefinery, biojet fuel, process intensification, energy integration.

## 1. Introduction

Microalgae are micro-organisms able that covert sunlight and CO<sub>2</sub> to biomass, from which biofuels and value-added products can be produced; they do not require fertile soil, and can grow in wastewater (Romero-Izquierdo et al., 2022). In particular, the use of microalgae for the sustainable aviation fuel production has gained attention; this biofuel is the most resilient strategy to achieve the sustainable recovery of the sector (IEA, 2021). Various reports present the use of microalgae oil to produce biojet fuel (Gutiérrez-Antonio et al., 2017; O'Neil et al., 2019; Martínez-Villarreal et al., 2022), while a few ones present the conversion of whole-microalgae strains into a biorefinery scheme (de Carvalho et al., 2022; Siddiki et al., 2022; Elkelawy et al., 2022); however, the biojet fuel generated in these processes is not competitive with its fossil counterpart (NESTE, 2022). The microalgae strain cost and the high energy requirements for the processing are the main issues to reach its profitability (Romero-Izquierdo, 2020). Thus, the efforts must be oriented to the application of strategies to make more efficient the use of energy, such as process intensification and energy integration. Therefore, in this work the computer-aided design of intensified and integrated routes for the conversion of whole-microalgae (*Chlorella vulgaris*) to biojet fuel, other biofuels and high-value products in a biorefinery scheme is presented; the implementation of such strategies over the energetic, economic and environmental indicators of the production route is analyzed.

## 2. Modelling and simulation of microalgae biorefinery

*Chlorella vulgaris* is cultivated by 14 days in wastewater, estimating a microalgal biomass after the harvesting of 442,795.64 kg by production lot (Romero-Izquierdo, 2020). This biomass is composed by proteins (46 wt %), carbohydrates (18 wt %) and lipids (27 wt %) (Becker, 2007). The biorefinery (Figure 1) is divided in three zones: oil extraction (Z1), press cake treatment (Z2) and microalgal oil treatment (Z3); modelling each one using NRTL for Z1 and Z2, and Peng-Robinson and BK10 for Z3. Z1 considers the modelling of oil extraction by ultrasonication; Natarajan et al. (2014) reported a maximum cell disruption of 75 % is reached with 0.8 kWh/L. This process is modelled with *CFuge* module, from which oil and press cake streams are obtained. The press cake is processed in Z2: 10 wt % enters to pyrolysis process, 10 wt % to anaerobic digestion and 80 wt % to ethanol production. The pyrolysis is realized at 500 °C and air, obtaining bio-oil (mainly composed by levoglucosan, toluene and hexadecenoic acid) and biochar modelled by the empirical formula  $C_{1.7}H_{8.2}N$  (Brown et al., 2012). This process was modelled by *RStoic* module. The anaerobic digestion is carried out at 55 °C with a water-biomass ratio of 1:1 (Li et al., 2014), and represented by the Buswell equation (Symons and Buswell, 1933). The biogas obtained is fed to isentropic turbine (module *Compr*) to produce power. In the case of ethanol production, the available biomass enters to a separate saccharification and fermentation process (SHF). The first step is the acid hydrolysis at 121 °C with 1 % of sulfuric acid; its outlet stream enters to the fermentation reactor (*RStoic* module) at 30 °C adding *Z. mobilis*. The non-converted biomass is removed (by *Sep* module) of the outlet stream from fermentation reactor to be fed to a distillation train with three columns. From the first column, 35 stages without condenser, the CO<sub>2</sub> is obtained by the top, extracting side stream rich in ethanol, which enters to second column, 13 stages and partial-vapor condenser, concentrating the ethanol at the top. The ethanol-rich stream enters to the third column, 15 stages and total condenser, breaking the ethanol-water azeotrope with glycerol. All columns are modelled with *RadFrac* module, reaching 99 % recovery for key components. The produced ethanol is inserted to the ATJ process on Z3, where are three consecutive reactors (modelled by *RStoic* modules): ethanol dehydration, oligomerization, and hydrogenation. The first one at 450 °C operates with a superheated steam stream with a 1:1 ethanol-steam ratio, reaching 99.4 % yield of ethylene; the second one at 120 °C and 35 bar reached 99 % yield to alkenes, and the last one operates at 110 °C with 99 % conversion to paraffins. The ethylene purification before the second reactive stage is realized using a heat exchanger (Heat-X module) and two flash tanks (Flash2 modules), reaching 99 % of ethylene purity. The outlet stream is conditioned until 30 °C and 1 bar before the separation, along with the hydrocarbons from hydroprocessing. The hydrotreating process uses 20 wt % of microalgae oil from Z1, and operates at 410 °C and 50 bar adding 1500 mL of hydrogen per oil mL (Verma et al., 2011), using a *RStoic* module. The outlet stream is conditioned until 27 °C and 1 bar using an isentropic turbine (*Compr* module); then, it is mixed with the ATJ hydrocarbons stream, to separate the hydrocarbons in a unique distillation train with two columns (conventional scheme). The first column, 52 stages, separates the naphtha fuel (C4-C7) by the top, whilst the bottom enters to second column, 106 stages, separating the biojet fuel (C8-C16) by the top and the green diesel (C17-C20) at the bottom. Both columns were designed with *DSTW* module, and later rigorously simulated with *RadFrac* module. The thermodynamic model used in both cases was BK10. The 80 wt % of the oil from Z1 is transesterified with methanol, at 62 °C, 1 atm and sodium methoxide as catalyst (Farooq et al., 2016), using *REquil* module and UNIFAC model. The biodiesel refining takes place on three stages; in the first one a

Decanter module is used to separate the FAME from glycerol, water and methanol. In the second stage, the FAME stream enters to a washing column (*RadFrac* module and UNIF-DMD model) with 2 stages, to remove the alcohol and glycerol traces. The washed stream enters to an evaporator (*Flash* module and UNIFAC model) to remove the water traces. Two energy efficiency tools are implemented on the biorefinery scheme: process intensification (PI) on separation zone, and energy integration (EI) for the whole biorefinery. In this work, the PI through the Rong and Errico (2012) methodology is applied on the separation zone of hydrocarbons from ATJ and hydrotreatment processes. On the other hand, EI application is realized by the pinch point methodology (Kemp, 2006) on whole biorefinery, considering the scheme with the intensified distillation train.

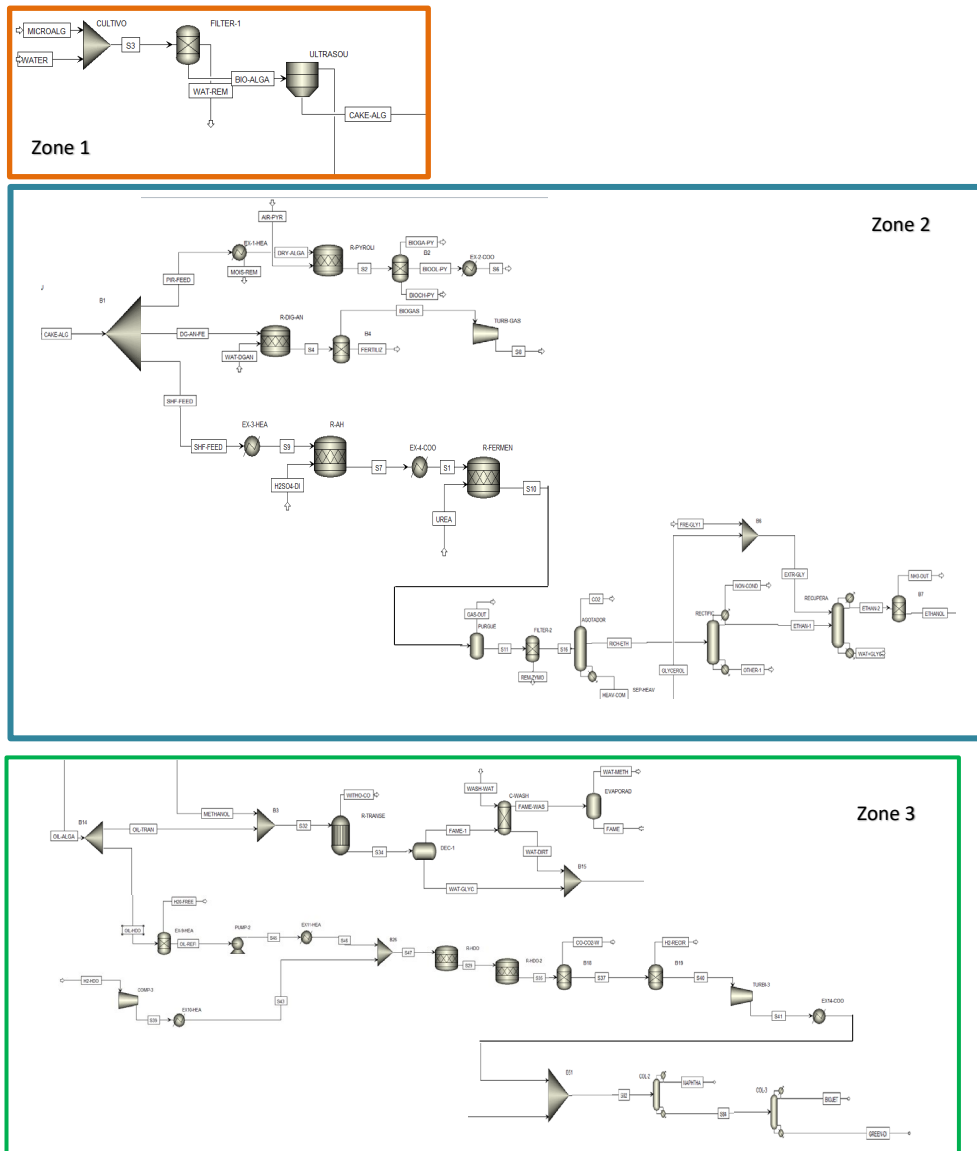


Figure 1. Biorefinery of microalgae by processing zones, non-intensified scheme.



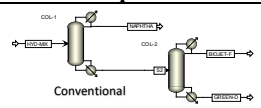
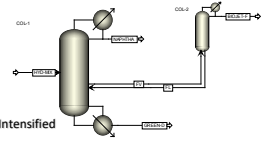
### 3. Assessment of the biorefinery scheme

The assessment of the biorefinery is carried out with the total annual cost (TAC) and CO<sub>2</sub> emissions by processing. The TAC involves capital cost, calculated by Aspen Economics, adding 18% (A1) and 61% (A2) due to contingencies, installation fees and equipment maintenance (Turton et al., 2012). The operating cost involving utilities cost (heating and cooling), raw material cost (microalgae, hydrogen, methanol) and additional reagents cost (glycerol, urea, H<sub>2</sub>SO<sub>4</sub>, etc), excluding catalyst and filters. The CO<sub>2</sub> emissions considers the steam and electricity requirements, according to Gutiérrez-Antonio et al. (2016).

### 4. Analysis of results

In this section the main results obtained for the microalgae biorefinery are described. This biorefinery allows to obtain biojet fuel (16,538.18 kg/h), naphtha (10,055.994 kg/h), green diesel (277.025 kg/h) and biodiesel as fuels; also, glycerol (5,130.888 kg/h), acetic (113,329.49 kg/h), succinic (734.855 kg/h) and lactic (910.893 kg/h) acids, along with bio-oil (8,130.41 kg/h), bio-char (8,463.90 kg/h) of pyrolysis, power and water with nutrients. As can be seen the total biojet fuel produced is 22,844.07 kg/h (58 % from hydrotreatment and 42 % from ATJ); also, acetic acid, succinic and lactic acids, respectively, are produced as value-added products. Regarding to PI applied on the separation zone from ATJ and hydrotreating, the results of the conventional and intensified schemes are presented in Table 1.

Table 1. Results of conventional and intensified sequence.

Sequence	Q tot con (kW)	Q tot reb (kW)	No. Stages
 <p>Conventional</p>	5,052.76	5,797.92	158
 <p>Intensified</p>	5,007.29	5,722.55	157

The energy savings for the condenser and reboiler are 1.3 % and 0.9 %, regarding the conventional scheme. Also, the intensified scheme has two less stages than the conventional, thus the equipment cost could be reduced. With respect to energy integration, the biorefinery scheme was analyzed through two routes: the first one (R1) including the Z1 and Z2 processes, and the route 2 (R2) for Z3 processes. R1 has 5 and 4 cold and hot streams, respectively; whilst R2 6 and 5 cold and hot streams respectively. According to the pinch point methodology, and considering  $\Delta T_{min}$  of 10°C, it is possible to reach 23 % and 8.89 % of savings for heating and cooling, respectively, for R1. In the case of R2, the savings for heating and cooling are 75.43 % and 74.76 %, respectively. The exchanges between the process streams for R1 and R2 are presented in Figure 2. Both designs were constructed taking into account the minimum number of heat exchangers. The TAC estimation of the scheme with energy efficiency tools is 61,487.335 million USD/y. In the case of equipment cost, it is 7 % higher (30.479 million USD/y), regarding to original scheme due to the additional heat exchangers and the size increase of intensified column. Nevertheless, economic savings of 72.31 % (0.375 million USD/y) and 55 % (0.036 million USD/y) are obtained for heating and cooling costs. Additional

factors involved in the estimation of TAC are the same for both schemes. Moreover, the original scheme releases 1,284.54 MTon CO<sub>2</sub>/y, due to its steam requirements; whilst the scheme with energy efficiency tools releases 706.497 MTon CO<sub>2</sub>/y, representing savings of 45 %.

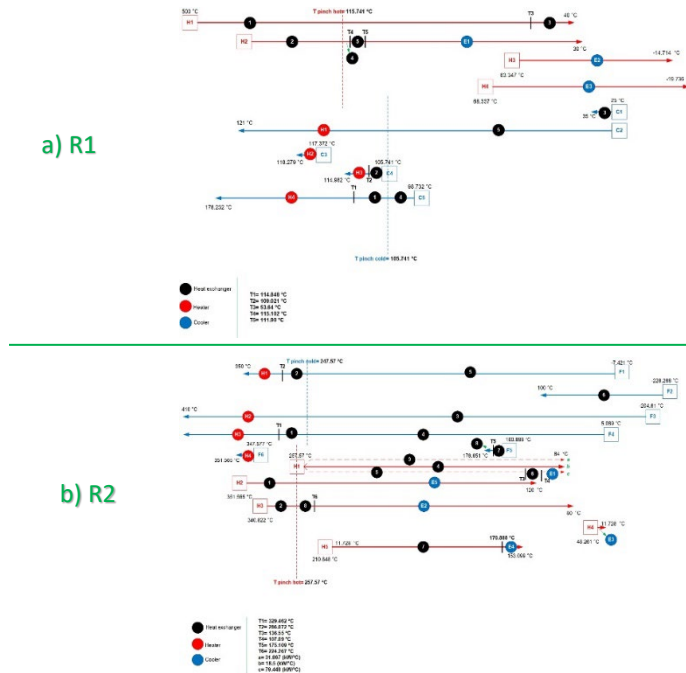


Figure 2. Heat exchanger networks: a) R1, b) R2.

## 5. Conclusions

The modeling and simulation of a microalgae biorefinery proposal with energy efficiency tools has been presented. The biorefinery produces biojet fuel, other biofuels and high value products. Based on results, the application of PI on the separation zone reduced 1.3 % of energy requirements at the reboilers. Likewise, the EI allowing reaching savings until 23 % and 75.43 % of heating utilities for R1 and R2, respectively. These energy savings enhanced the operational cost until 72.31 %, regarding to original scheme. However, microalgae cost has the higher contribution, which is the main opportunity area.

## Acknowledgments

Financial support provided by CONACYT for the doctoral and post-doctoral studies of Araceli Guadalupe Romero-Izquierdo is acknowledged.

## References

- E. Becker, 2007. Micro-algae as a source of protein. *Biotechnol. Adv.* 25, 207–210.
- R. Brown, R.C., Sidhu, S.S., Martinez, L., Wang, K., Homsy, S., 2012. Fast pyrolysis of microalgae remnants in a fluidized bed reactor for bio-oil and biochar production. *Bioresour. Technol.* 127, 494–9.
- J.de Carvalho, D.T.M. Aulestia, M.A. de Carvalho, E.B. Sydney, A.I. Magalhães, C.R. Soccol, A. Ravishankar, R. Ambati, , 2022. 15 - Biorefinery approaches for integral use of microalgal

- biomass, in: Jacob-Lopes, E., Zepka, L.Q., Severo, I.A., Maroneze, M.M. (Eds.), 3rd Generation Biofuels, Woodhead Publishing Series in Energy. Woodhead Publishing, 321–344.
- M. Elkelawy, H. Bastawissi, A. Radwan, M. Ismail, M. El-Sheekh, 2022. Chapter 15 - Biojet fuels production from algae: conversion technologies, characteristics, performance, and process simulation, in: El-Sheekh, M., Abomohra, A.E.-F. (Eds.), Handbook of Algal Biofuels. Elsevier, 331–361.
- A. Farooq, U. Amin, Y. Abdullah, 2016. Transesterification of oil extracted from different species of algae for biodiesel production. *African J. Environ. Sci. Technol.* 7, 358–364.
- C. Gutiérrez-Antonio, F. Gómez-Castro, J. de Lira-Flores, S. Hernández, 2017. A review on the production processes of renewable jet fuel. *Renew. Sustain. Energy Rev.* 79, 709–729.
- C. Gutiérrez-Antonio, A. Romero-Izquierdo, F. Gómez-Castro, S. Hernández, 2016. Energy Integration of a Hydrotreatment Process for Sustainable Biojet Fuel Production. *Ind. Eng. Chem. Res.* 55, 8165–8175.
- IEA.2021. IEA – COVID-19 Topics [WWW Document]. Int. Energy Agency. URL <https://www.iea.org/topics/covid-19> (accessed 9.1.22).
- I. Kemp, 2006. Pinch Analysis and Process Integration, Pinch Analysis and Process Integration. Elsevier Ltd.
- H. Li, Z. Liu, Y. Zhang, B. Li, H. Lu, N. Duan, M. Liu, Z. Zhu, B. Si, 2014. Conversion efficiency and oil quality of low-lipid high-protein and high-lipid low-protein microalgae via hydrothermal liquefaction. *Bioresour.*
- S. Martínez-Villarreal, A. Breitenstein, P. Nimmegeers, P. Perez Saura, B. Hai, J. Asomaning, A. Eslami, P. Billen, S. Van Passel, D. Bressler, D. Debecker, C. Remacle, A. Richel, 2022. Drop-in biofuels production from microalgae to hydrocarbons: Microalgal cultivation and harvesting, conversion pathways, economics and prospects for aviation. *Biomass and Bioenergy* 165, 106555.
- R. Natarajan, W. Ang, X. Chen, M. Voigtmann, R. Lau, 2014. Lipid releasing characteristics of microalgae species through continuous ultrasonication. *Bioresour. Technol.* 158, 7–11.
- NESTE, 2022. SAF for airlines and airports, [WWW Document]. Int. Energy Agency. URL <https://www.neste.com/products/all-products/saf/key-benefits> (accessed 9.1.22)
- G. O'Neil, G. Knothe, C. Reddy, 2019. Chapter 15 - Jet biofuels from algae, in: Pandey, A., Chang, J.-S., Soccol, C.R., Lee, D.-J., Chisti, Y. (Eds.), *Biofuels from Algae* (Second Edition), Biomass, Biofuels, Biochemicals. Elsevier, pp. 359–395.
- A. Romero-Izquierdo, 2020. Diseño, modelado y simulación de un esquema de biorefinería para el aprovechamiento integral de mezclas de materias primas renovables. Univ. Guanajuato PhD Thesis.
- A. Romero-Izquierdo, C. Gutiérrez-Antonio, F. Gómez-Castro, S. Hernández, J. García-Trejo, 2022. 27 - Production of renewable aviation fuel from microalgae, in: Jacob-Lopes, E., Zepka, L.Q., Severo, I.A., Maroneze, M.M. (Eds.), 3rd Generation Biofuels, Woodhead Publishing Series in Energy. Woodhead Publishing, 639–664.
- B. Rong, M. Errico, 2012. Synthesis of intensified simple column configurations for multicomponent distillations. *Chem. Eng. Process. Process Intensif.* 62, 1–17.
- S. Siddiki, M. Mofijur, P. Kumar, S. Ahmed, A. Inayat, F. Kusumo, I. Badruddin, T. Khan, L. Nghiem, H. Ong, T. Mahlia, 2022. Microalgae biomass as a sustainable source for biofuel, biochemical and biobased value-added products: An integrated biorefinery concept. *Fuel.* 307, 121782.
- G. Symons, A. Buswell, 1933. The Methane Fermentation of Carbohydrates 1,2. *J. Am. Chem. Soc.* 55, 2028–2036.
- R. Turton, R. Bailie, W. Whiting, J. Shaeiwitz, D. Bhattacharyya, 2012. Analysis, Synthesis, and Design of Chemical Processes, Fourth Ed. ed.
- D. Verma, R. Kumar, B. Rana, A. Sinha, 2011. Aviation fuel production from lipids by a single-step route using hierarchical mesoporous zeolites. *Energy Environ. Sci.* 4, 1667–1671.

# Optimization of Simulated Moving Bed Chromatographic Processes using Surrogate Models

Rojiar Pishkari<sup>a</sup>, Marcus Fechtner<sup>a</sup>, Tobias Keßler<sup>a</sup>, Achim Kienle<sup>a,b</sup>

<sup>a</sup>*Otto von Guericke University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany*

<sup>b</sup>*Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany*

*Rojiar.pishkari@ovgu.de*

## Abstract

In this paper, we investigate a surrogate-based optimization of Simulated Moving Bed (SMB) Chromatography with Langmuir adsorption isotherm using an iterative approach. Artificial neural networks are fitted in each iteration based on randomly distributed sampling points around the optimal solution of the previous iteration. Crucial (Hyper)parameters of this surrogate-based optimization are related to the sampling region, e.g. the size, the position, and the number of samples within. It is shown that for highly efficient chromatographic columns with a large number of theoretical stages, the surrogate-based optimization is much faster than the numerical optimization of the full-blown model.

**Keywords:** SMB, Surrogate model, Optimization, Neural network, Chromatography

## 1. Introduction

SMB is an advanced technology that realizes continuous chromatographic separations (Schmidt-Traub et al., 2020). In the last decade, biotechnology, fine chemistry, and pharmaceutical industries have benefited greatly from SMB applications. The main advantages of this process over conventional batch processes are increased productivity and reduced solvent consumption. Generally, chromatographic processes contribute significantly to the overall cost of the corresponding process chains. It is therefore essential to optimize their design and control through efficient process optimization.

Figure 1 illustrates a standard binary separation configuration. The process consists of four zones, each of which consist of at least one chromatographic column. They are interconnected and form a ring shape. The component with the higher affinity to the solid phase is obtained in the extract, whereas the component with the lower affinity to the solid phase is found in the raffinate. In order to achieve continuous separation, cyclic switching of the inlet- and outlet-ports in the direction of the fluid flow is used to simulate counter-current flow of the solid phase. SMBs are complex processes that involve periodic nonlinear dynamics governed by partial differential equations (PDEs). It is crucial to design and operate SMBs optimally, as this allows to maximize the economic potential of the process and helps with the successful implementation on an industrial scale.

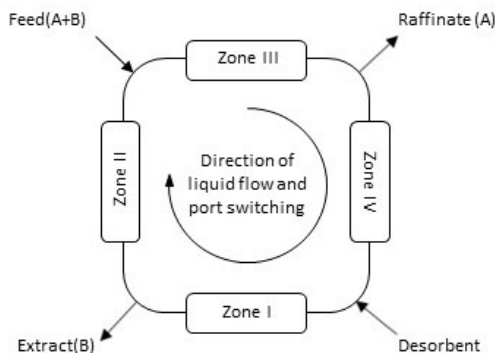


Figure 1. Schematic illustration of 4-zone SMB

The economic potential is maximized by finding optimal operating conditions regarding productivity and solvent consumption. Additionally, the process needs to fulfill certain purity requirements. For total separation, the well-known triangle theory developed by Storti et al. (1993) can be used in a first approximation to identify the optimal operating conditions. It is based on the assumptions of a true moving bed process, neglects axial dispersion and assumes thermodynamic equilibrium between the solid and the liquid phase. Productivity can be increased substantially if purity requirements can be relaxed, which depends on the specific type of application. For reduced purity requirements, Kaspereit et al. (2007) presented an extension of the triangle theory. However, the extension is involved, has limitations and is based on the same simplifying assumptions as mentioned above. By using genetic algorithms, Zhang et al. (2003) have optimized the productivity and purity of the SMB process based on a dynamic model, and compared it to the Varicol process, however computational time is rather high. In Li et al. (2014), a surrogate model was used to optimize SMB processes. They used two different types of surrogate models. The proper orthogonal decomposition (POD) method is employed to derive cost-efficient reduced-order models (ROMs) for the SMB process and the other one is a coarse model. To create a low-fidelity DAE model, they applied the Finite Element Method to the SMB model with coarser spatial discretization. In the present paper, a numerical optimization of a more detailed model is proposed instead. Due to the complexity of the used SMB model, rigorous numerical optimization is highly time-consuming. This paper adapts the methodology from Kessler et al. (2019) to develop a surrogate-based iterative approach by using a simple feed forward artificial neural network, requiring only minimal data, thereby increasing its numerical efficiency.

## 2. Description of the Optimization Problem

There are linear and nonlinear solvers in the optimizer toolbox in MATLAB (MathWorks, 2023) of which the 2021 version was used here. In this work the following types of optimization problems are considered:

$$\min J(\mathbf{x}), \quad (1)$$

$$\begin{aligned} \mathbf{h}(\mathbf{x}) &= 0, \\ \mathbf{g}(\mathbf{x}) &\leq 0, \quad x_i \in R, \\ lb_i &\leq x_i \leq ub_i, \quad \text{for } i=1,2, \end{aligned}$$

Where  $J(\mathbf{x})$ , is the objective function,  $\mathbf{h}(\mathbf{x})$  are the equality constraints, e.g. material balances,  $\mathbf{g}(\mathbf{x})$  are the inequality constraints, such as purity requirements, and  $\mathbf{x}$  are the decision variables, which are the ratios between the liquid and solid phase flow rate. The well-known equilibrium dispersion model can be used to describe chromatographic columns.

$$\left\{ \begin{aligned} \varepsilon \frac{\partial c_i}{\partial t} + (1 - \varepsilon) \frac{\partial q_i}{\partial t} + \varepsilon v \frac{\partial c_i}{\partial z} &= D_{ax} \varepsilon \frac{\partial^2 c_i}{\partial z^2} \\ q_i &= \frac{H_i c_i}{1 + \sum_{i=1}^{n_c} b_i c_i} \end{aligned} \right. \quad (2)$$

In the first equation, which is the equilibrium model,  $v$  represents the interstitial velocity, while  $D_{ax}$  represents the apparent dispersion coefficient, which lumps together all effects that contribute to band broadening. Furthermore,  $c_i$  is the fluid phase concentration of component  $i$ , while  $q_i$  is the solid phase concentration and  $\varepsilon$  is the volume fraction of the fluid phase. In the second equation, which is the Langmuir isotherm,  $H$  and  $b$  are the adsorption constants of each component and  $n_c$  is the number of components. According to the adsorption isotherm,  $q_i$  depends on the fluid phase composition since the model assumes thermodynamic equilibrium between the two phases. There are linear and nonlinear solvers in the optimizer toolbox in MATLAB. In this case, we are using `fmincon`. The objective function to be maximized is the productivity for a given minimum product purity. Our decision variables,  $\mathbf{x}$ , are the  $m$ -values, which are the dimensionless ratios between the liquid and the solid phase flowrates in each column. Maximum productivity is obtained by maximizing the difference between  $m_3$  and  $m_2$ , which is equivalent to find a feasible operating point farthest from the diagonal  $m_2 = m_3$ . At the beginning, the Henry constants are used as the starting point, i.e. the first  $m_3$  and  $m_2$ . Since they are the optimal point of the linear adsorption isotherm for total regeneration, they can be used as initial guess for Langmuir adsorption isotherm. The domain for the optimization area is given by  $m_2 \in [3, 7]$  and  $m_3 \in [5.5, 9.5]$ . In order to maintain total regeneration,  $m_1$  and  $m_4$  have a safety margin of 10% from Henry constants. A nonlinear inequality constraint of a minimum purity of 0.85 for both products is used for this maximization. In a cyclic steady state system, the purity of the raffinate and extract streams is defined as follows:

$$\left\{ \begin{aligned} pu_{raf} &= \frac{\int_t^{t+tsw} c_2^{raf} dt}{\int_t^{t+tsw} c_1^{raf} dt + \int_t^{t+tsw} c_2^{raf} dt} \\ pu_{ex} &= \frac{\int_t^{t+tsw} c_1^{ex} dt}{\int_t^{t+tsw} c_1^{ex} dt + \int_t^{t+tsw} c_2^{ex} dt} \end{aligned} \right. \quad (3)$$

Therein,  $tsw$  is the switching time. Variables  $pu_{raf}$  and  $pu_{ex}$  indicate the purity of raffinate and purity of extract, respectively, while  $c$  refers to the concentration of the components. The numerator indicates the quantity of target component collected during the switching period, and the denominator indicates the sum of two components collected during the switching period. Table 1 shows the parameters and operating conditions for SMB simulation.  $n_{col}$  is the number of columns,  $\epsilon$  is the porosity and c-feed is the concentration of the feed. Optimizing is performed using Matlabs multi-start algorithm with 100 starting points and the sqp algorithm. The step tolerance option is set to  $1e^{-8}$ . An extension of the optimization is possible for all m-values.

Table 1. Model Parameters of the reference model

SMB Parameters					
$n_c$	2	c-feed <sub>1</sub>	0.5[gr. L <sup>-1</sup> ]	$b_2$	0.3
$n_{col}$	4	c-feed <sub>2</sub>	0.4[gr. L <sup>-1</sup> ]	$m_1$	8.25
$n_z$	1000	$H_1$	5	$m_2$	$x_1$
$tsw$	99.5 s	$H_2$	7.5	$m_3$	$x_2$
$\epsilon$	0.74	$b_1$	0.15	$m_4$	4.5

### 3. Surrogate Model

A surrogate model is developed to replace the detailed but expensive full-order SMB model for optimization purposes. In this work an artificial neural network is used as a surrogate model. Using the methodology introduced in Kessler et al. (2019), this paper proposes a surrogate-based iterative approach, requiring minimal data, which increases the numerical efficiency. The initial guess defined by the Henry constants and an area around it is defined. Some sample points are generated inside the area based on the Halton set algorithm from Kocis et al. (1997). Halton set creates points based on the Halton sequence in a quasi-random manner. To fill the space uniformly, the Halton sequence uses different prime bases in each dimension. The location of these points inside the sampling area is used as m-values ( $m_2, m_3$ ). A simulation is performed for each point inside the area and the output data, such as purity and productivity, are saved for every point (pair of m-values). Here 150 points are generated. The reference model is used to calculate the purity of the products, as well as the productivity for each point. The purity must be greater than 0.85 to be considered for the further calculation. Only the points that meet the purity criteria are used for the surrogate model optimization. During each iteration, the sample points are saved and reused if they are in the sampling area of the next iteration as well. Each iteration may have more or fewer samples than 150. The data obtained from this calculation are used to train the neural net. The neural net is feed forward and has one layer with 15 neurons. The surrogate model uses the same optimization algorithm as the full model, multi-start with 100 random starting points and step tolerance option set to  $1e^{-8}$  to find an optimal point in each iteration. If the deviation of an optimal point at each iteration to the previously calculated optimum is bigger than the termination condition which is 0.019, the next calculation starts, otherwise the final optimal point is found. Using the current optimal point as an initial guess, the following formula defines the new sampling region (Kessler et al. (2019)):

$$lb/ub = x_{opt,k-1} \pm \frac{1}{2k} \quad (4)$$

The integer variable  $k$  corresponds to the iteration counter. The optimal point found by the previous iteration is the center of the current calculation area. The area is shrinking according to equation 4 and the density of samples inside it is increasing. In the present calculation procedure, the algorithm terminated after 4 iterations. The progress is shown in Figure 2. The first figure shows the optimal point as red star for the first iteration, while the red circle is illustrating the first initial guess. The black dots are the whole sample data generated by the Halton set, and the blue ones are the samples that are matching the purity requirement. A convex hull is created that includes all blue dots, and the optimizer searches for the solution inside this new area. It is clear that the shrinking effect is significant only in the first iteration indicating the importance of the first sampling region. The big difference from first to second iteration is due to the usage of equation 4. It can be seen from the second figure that the area decreases at a constant rate. Due to the reduced purity requirement that defines the convex hull, the optimal point lies on its boundary with maximal distance to the diagonal  $m_2 = m_3$ . In the last figure, the final optimal point found by surrogate-model (red star) almost coincides with the optimal point found by full-model optimization (yellow star). Further, the deviation from the previous optimum is less than the termination condition. The same initial guess point is used as the full-model optimization in the surrogate model.

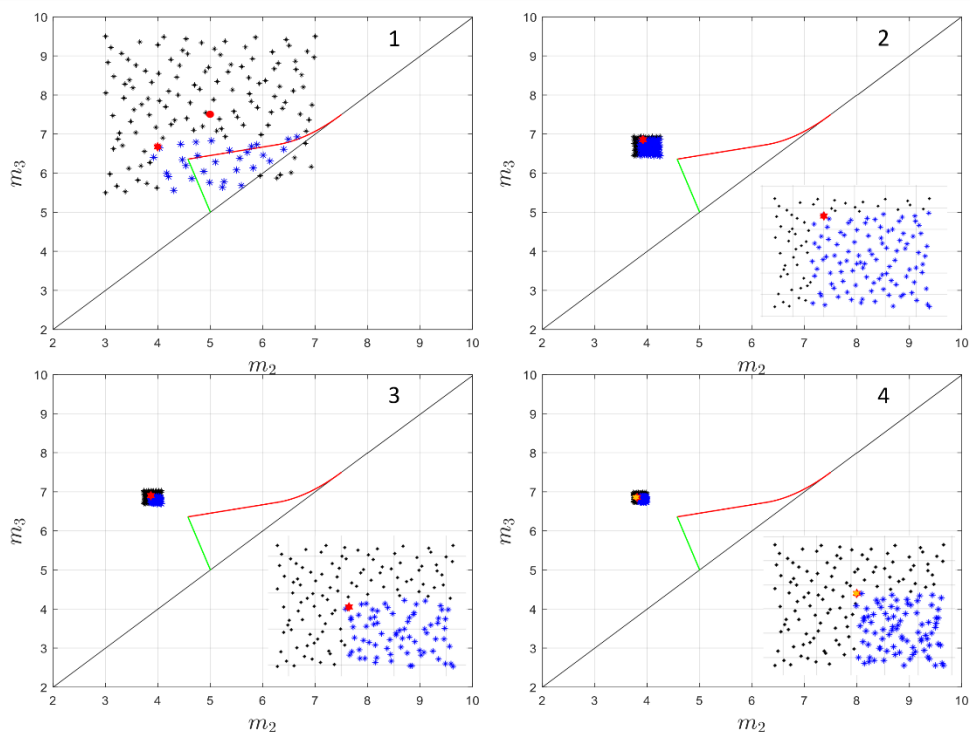


Figure 2. Optimal point of every iteration and how it converges to the final optimum.



#### 4. Results and Conclusion

A surrogate based optimization approach for the SMB process is proposed in this work. In place of the time-consuming rigorous model, a surrogate model is used for optimization. Table 2 compares the final optimal points obtained by both models to show the effectiveness of this approach. Using the surrogate model, SMB productivity was optimized resulting in a reasonable computational effort. It was expected that the optimal point would be outside the triangle because of reduced purity. The concentration of the extract as well as the raffinate and the productivity are calculated by the reference model for the optimal points found by both models. It is shown that the purity requirements are met. Though the full-model's optimal point produces higher productivity than the surrogate-model, the difference is marginal. Despite almost identical results, the surrogate model's computational time is much faster than the full model for the presented optimization set-up, making it a better choice.

Table 2. CPU time comparison

	$x_{opt}$	cpu-time	$pu_{ex}$	$pu_{raf}$	productivity
Sur-model	3.8582, 6.8806	5423.75 s	0.8501	0.8504	3.0224
Full-model	3.8578, 6.8826	152453 s	0.8500	0.8500	3.0249

The proposed methodology could be particularly useful if deterministic global optimization is applied as in Kessler et al. (2019), which relies on a very large number of model evaluations but was beyond the scope of this paper. Future work will be concerned with minimizing the solvent consumption, which is achievable by minimizing  $m_1 - m_4$ . Furthermore, optimization of the SMB with Bi-Langmuir isotherm is also of interest.

#### References

- M. Kasperit, A. Seidel-Morgenstern, A. Kienle, (2007). Design of simulated moving bed process under reduced purity requirements. *Journal of Chromatography A*, 1162, 2-13.
- T. Kessler, C. Kunde, K. McBride, N. Mertens, D. Michaels, K. Sundmacher, A. Kienle, (2019). Global Optimization of Distillation Columns using Explicit and Implicit Surrogate Models. *Chemical Engineering Science*, 197, 235-245.
- L. Kocis, W. J. Whiten, (1997). Computational investigations of low-discrepancy sequences. *ACM Transactions on Mathematical Software* 23 (2), 266–294.
- S. Li, L. Feng, P. Benner, & A. Seidel-Morgenstern, (2014). Using surrogate models for efficient optimization of simulated moving bed chromatography. *Computers & Chemical Engineering*, 67, 121-132.
- MathWorks, (2023) MATLAB, <https://uk.mathworks.com/products/matlab.html> (Last accessed: 29.01.2023).
- H. Schmidt-Traub, M. Schulte, A. Seidel-Morgenstern, (2020) *Preparative Chromatography*, 3rd edition.
- G. Storti, M. Mazzotti, M. Morbidelli, S. Carrà, (1993). Robust design of binary countercurrent adsorption separation processes. *AIChE J.* 39(3), 471–492.
- Z., Zhang, M., Mazzotti & M. Morbidelli, (2003). Multiobjective optimization of simulated moving bed and Varicol processes using a genetic algorithm. *Journal of chromatography A*, 989(1), 95-108.

# Multi-Objective Model-based Design of Experiments of Pharmaceutical Tableting Process

Ilias Bouchkira<sup>a</sup>, Brahim Benyahia,<sup>a,\*</sup>

<sup>a</sup> *Department of Chemical Engineering, Loughborough University, Epinal Way, Loughborough, LE11 3TU, United Kingdom*  
[B.Benyahia@lboro.ac.uk](mailto:B.Benyahia@lboro.ac.uk)

## Abstract

The development of high-fidelity and predictive models is one of the cornerstones of process engineering. Often mathematical models involve several unknown parameters to be identified from experiments. Designing the minimum set of information rich experiments for precise estimate is a critical to reduce the examination costs and help improve model prediction capabilities. In this work, we present a novel Multi-Objective Model-Based Design of Experiments (MOMBDoE). The proposed approach is based on the simultaneous maximizing the D-optimal design of experiments criterion and the estimability potential of the model parameters (Estimability). Global sensitivity analysis is used to build the Fisher Information Matrix (FIM), which allows the maximization or minimization of MOMBDoE criteria. The Pareto optimal solutions which represent the best experimental compromises were ranked using a multicriteria decision aiding method to help identify the best alternatives for experimental validation. To validate the proposed MOMBDoE framework, a tablet lubrication process is used as a case study. Kushner and Moore's model is used to predict the tensile strength and hardness of the tablets as one of the main critical quality attributes in tablet manufacturing.

**Keywords:** Multi-objective Optimization, Estimability Analysis, Model-Based Design of Experiments, Tablet Lubrication Process.

## 1. Introduction

Pharmaceutical tablets are the most important dosage forms which commonly contain one or more active pharmaceutical ingredients and several excipients. The optimization of the process performance and product quality is painstaking as the aim is to produce tablets that meet a set of optimal critical quality attributes such as tensile strength, hardness, friability, dissolution, etc. The development of robust digital technologies in pharma requires systematic methodologies to build reliable and predictable mathematical models, soft sensors, and digital twins to help optimize the process/plant design and operation while delivering quality assurance. One of the most effective approaches to address these challenges is to implement a Model-Based Design of Experiments (MBDDoE) to generate the minimum set of information-rich experiments.

Model-based design of experiments have been widely investigated for such purposes. Most of these methods are based on the Fisher Information Matrix which can be obtained from the sensitivity analysis. While the current model-based design of experiments aims at designing experiments with rich information, the estimability analysis is mainly used to assess whether the available experimental measurements contain sufficient information

to reliably estimate all model parameters or only a subset of them. Combining these two approaches (i.e., MBDoe and estimability analysis), can consequently consolidate the advantages of each method and allow the design of a minimum set of experiments for precise estimates.

This work presents a novel multi-objective model-based design of experiments which combines a D-optimal design of experiments and global estimability analysis. The approach is formulated as a constrained multi-objective optimization problem that accounts for two objectives, the first is based on the D-optimal MBDoe criterion and the second is based on the estimability. The optimization problem is also subject to a set of operating constraints and bounds. The set of Pareto optimal solutions obtained from multi-objective optimization provide the best experimental compromises or tradeoffs. Here, the solutions are ranked using a multicriteria decision aiding method to help identify the best candidates for experimental implementation. To validate the proposed approach, a pharmaceutical tablet lubrication model is considered.

## 2. Model and methods

### 2.1. Lubrication model

In tablet press, tensile strength and hardness are very important Critical Quality Attributes (CQAs) which need to be tightly monitored and controlled. These CQA can be related the solid fraction of tablets to the powder lubrication extent through several expressions among which, are the ones developed by Kushner and Moore (2010) and Pitt et al. (1988):

$$\frac{ts}{ts_{sf=0.85,0}} = (1 - \beta) + \beta \exp(-\gamma k) \quad (1)$$

$$H = \frac{t\pi D^2}{10} \left( 2.84 \frac{\sigma}{D} - 0.126 \frac{\sigma}{\sigma - 2\sigma_c} + 3.15 \frac{\sigma - 2\sigma_c}{D} + 0.01 \right) \quad (2)$$

Where  $ts_{sf=0.85,0}$  is the initial tensile strength at 0.85 solid fraction,  $\gamma$  is the lubrication rate constant of the blend, and  $\beta$  is the total fraction of tensile strength that can be lost due to lubrication.  $D$  is the diameter of the tablet,  $\sigma$  is the total thickness of the tablet, and  $\sigma_c$  is the thickness of the convex cups.

To avoid dependence on the initial solid fraction, the following empirical equations were introduced (Cenci et al., 2022):

$$ts_{sf=0.85,0} = a_1 \exp(b_1(1 - sf)) \quad (3)$$

$$\beta = a_2(1 - sf) + b_2 \quad (4)$$

The resulting model captures the impact of two main factors or inputs namely the solid fraction  $sf$  and lubricant extent  $k$ , and involves a vector of five unknown parameters to be identified from experiments ( $\theta = [a_1(MPa); b_1(-); a_2(-); b_2(-), \gamma(dm^{-1})]$ ).

### 2.2. Multi-objective MBDoe

The proposed multi-objective model-based design of experiments relies on the Fisher Information Matrix (FIM), which is computed from the global sensitivity indices of the model unknown parameters for a given set of experimental recipes or operating conditions:

$$Z = \begin{pmatrix} S_{1,1}|_{m_1} & \cdots & S_{1,n_\theta}|_{m_1} \\ \vdots & \ddots & \vdots \\ S_{n_y,1}|_{m_x} & \cdots & S_{n_y,n_\theta}|_{m_x} \end{pmatrix}, \quad (5)$$

where  $m_i$  is the  $i^{th}$  sampling point or experiment.  $n_\theta$  is the number of model parameters and  $n_y$  is the number of observed/measured outputs.  $S_{i,j}$  is the individual sensitivity coefficient of the  $i^{th}$  output with respect to the  $j^{th}$  parameter.

The FIM matrix is given by:

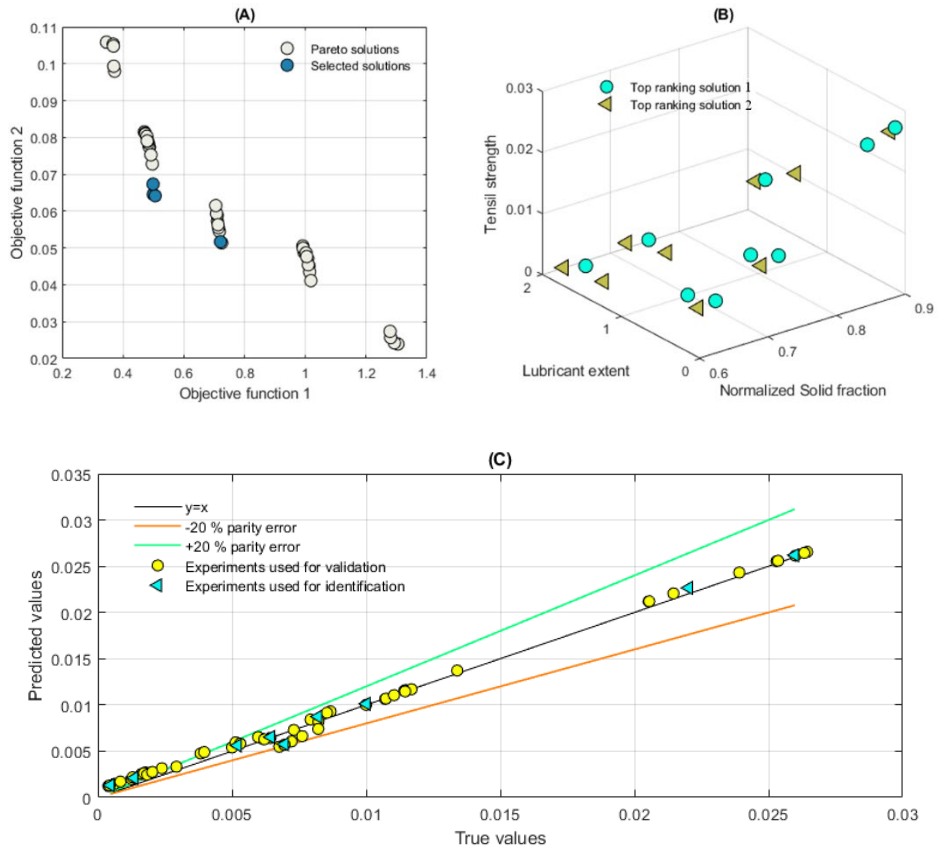
$$\mathbf{FIM} = Z \cdot Z^T \quad (6)$$

$S_{i,j}$  is the individual sensitivity coefficient of the  $i^{th}$  output with respect to the  $j^{th}$  parameter which can be calculated using the algorithm developed by Saltelli (2012). The sensitivity matrix is then used to generate the FIM to optimize the D-optimal MBDoe criterion and to calculate the estimability magnitude of the least estimable parameter (Benyahia et al. 2011a; Fysikopoulos et al., 2019; Bouchkira et al., 2023). A regularization test is firstly performed to determine whether the FIM is invertible. If this condition is satisfied a multi-objective optimization is performed to determine the Pareto optimal solutions which can be ranked using a multicriteria decision-aiding method.

The mathematical formulation of the proposed MOMBDoe problem is given by:

$$\begin{aligned} \max_{\mathbf{u}} \mathbf{F} &= [f_1, f_2]^T \\ f_1 &= \det(\mathbf{FIM}) \\ f_2 &= \min \sum_i \|r_{i,j}\| \\ \text{St. } 0 &< Sf_{j-1} \leq Sf_j, \quad j \in [1; n] \\ 0 &< k_{r-1} \leq k_r, \quad r \in [1; m] \\ Sf_{j-1} - Sf_j &\leq \omega_{sf} \\ k_{r-1} - k_r &\leq \omega_k \\ \text{Cond}(\mathbf{FIM}) &\leq \delta \\ \mathbf{u} &= [Sf_1, \dots, Sf_n, k_1, \dots, k_m] \end{aligned} \quad (\text{S.1})$$

where  $\det$  and  $\text{Cond}$  refer respectively to square diagonal of FIM and conditioning number of FIM.  $r_{i,j}$  are orthogonalization residuals calculated as in Benyahia et al. 2013,  $\omega_{sf}$  and  $\omega_k$  represent the difference between two solid fractions and lubricant extents, used to avoid duplicated experiments, and  $\delta$  is a cut-off criterion to guarantee that FIM is invertible, in this work  $\delta = 10$  is considered as in Cenci et al., 2022.



**Figure 1:** MOMBDoE results. (A) Pareto front and top-ranking solutions. (B) Experimental strategies / recipes for the 2 top ranking solutions from Pareto front. (C) Model validation showing normalized tensile strength and hardness data.

**Table 1:** Sample of four top ranking Pareto optimal solutions (ranked using MAUT).

$k_1$	$k_2$	$k_3$	$Sf_1$	$Sf_2$	$Sf_3$	Objective 1	Objective 2
0.132	0.368	1.863	0.615	0.704	0.848	0.499	0.064
0.179	1.376	1.883	0.618	0.710	0.868	0.724	0.051
0.287	1.058	1.753	0.726	0.807	0.886	1.304	0.023
0.249	1.373	1.818	0.640	0.737	0.897	1.011	0.045

### 3. Results and discussions

The global sensitivity analysis based on Sobol approach was first performed. A Quasi-Monte Carlo sampling of size 8000 is used to achieve convergence of the global sensitivity indices. Sensitivity indices are then used to build the Z sensitivity matrix and FIM. The latter is used in the proposed MOMBDoE. The resulting multi-objective optimization problem was solved using a Genetic Algorithms solver under MATLAB to generate a large set of Pareto optimal solutions which represent the best experimental compromises.

The optimal profiles i.e.,  $X^{opt}$  from the multi-objective optimization approach are used to generate information rich data, to be used for the identification of the unknown model parameters and model validation. The approach used in this work consists of identifying model unknown parameters from two sets of experiments: (i) from true measurements generated using  $X^{opt}$ . (ii) and from noisy measurements, then to compare model performances in both scenarios to see if it keeps very good prediction capabilities, even from noisy measurements. It is worth noting that that noisy data used in this work is created by adding gaussian noise to true data from Cenci et al., 2022.

The optimal solutions (i.e.,  $X^{opt}$ ) are given in the Pareto front as shown in Figure 1A. Indeed, in the Pareto front, each point corresponds to an optimal profile. For demonstration purposes, four points labeled “1-4” in the Pareto front are given in Table 1. The Pareto solutions represent the best tradeoffs which are commonly considered as equally good. However, the experimental validation requires a small set of solutions to be implemented. Consequently, it is important to implement a decision-aiding method to select the best profile.

The Multi-Attribute-Utility-Theory (MAUT), which has been successfully used as a robust decision aiding method in many engineering problems (Benyahia et al. 2011b; Bouchkira et al., 2022) was used to rank the Pareto optimal solutions based on the utility functions that capture the distance from the best single objective and its criticality. The Pareto solutions were ranked accordingly from the best to the worst which allows the identification of the best experimental strategies /recipes to demonstrate the benefits of the proposed MOMBDoE approach.

The Pareto solutions are shown in Figure 1A with a sample of the top-ranking solutions highlighted in a different color. The experimental conditions/recipes associated with 2 selected top-ranking solutions (Table 1) are shown in Figure 1B. Each of the 2 identified optimal experimental campaigns require 9 experiments at various lubricant extents and solid fractions. The top-ranking experimental campaign, identified as the best experimental candidate using the proposed approach above, was implemented to generate true and noisy data, based on which model performances are compared in Figures 1C. It is shown that the predicted values from the novel MOMBDoE are in good agreement with the true values from the literature (Cenci et al., 2022). The data in blue triangles are used in the parameter identification, while the data in yellow circles represents measurements that were not used in the identification process. The parity errors between the two compared sets of outputs lie between -20% and 20% which reflects good model performances and highlight the reliability of the proposed approach.

#### 4. Conclusion

In this work, a novel multi-objective model-based design of experiments is presented and implemented to a case study relevant to a pharmaceutical tableting process. The proposed MOMBDoE relies on global sensitivity analysis to combine traditional MBDoE with global estimability analysis. This new approach consequently allows the design of experiments which guarantee that all model unknown parameters are estimable and exhibit higher estimate precision (lower uncertainties). As the approach proposes several optimal solutions, decision-aiding is used to select a set of experimental candidates to be implemented. The results from the case study showed that the model developed based on the proposed MOMBDoE approach shows enhanced prediction capabilities and accuracy.

**Acknowledgements:** This work was funded by the EPSRC (EP/V062077/1) Digital Medicines Manufacturing Research Centre.

#### References

- A. Saltelli, (2002), Making best use of model evaluations to compute sensitivity indices, *Computer physics communications*, 145(2), 280-297.
- B. Benyahia, M. A. Latifi, C. Fonteix, F. Pla, (2011a). Modeling of a batch emulsion copolymerization reactor in the presence of a chain transfer agent: Estimability analysis, parameters identification and experimental validation. *Computer Aided Chemical Engineering*, 29, 126-130.
- B. Benyahia, M. A. Latifi, C. Fonteix, F. Pla, (2013). Emulsion copolymerization of styrene and butyl acrylate in the presence of a chain transfer agent. Part 2: Parameters estimability and confidence regions. *Chemical Engineering Science*, 90, 110-118.
- B. Benyahia, M. A. Latifi, C. Fonteix, F. Pla. (2011b). Multicriteria dynamic optimization of an emulsion copolymerization reactor. *Computers & Chemical Engineering*, 35(12), 2886-2895.
- D. Fysikopoulos, B. Benyahia, A. Borsos, Z. K. Nagy, C. D. Rielly, (2019), A framework for model reliability and estimability analysis of crystallization processes with multi-impurity multi-dimensional population balance models. *Computers & Chemical Engineering*, 122, 275-292.
- F. Cenci, G. Bano. C. Christodoulou, Y. Vueva, S. Zomer, M. Barolo, F. Bezzo, P. Facco, (2022). Streamlining tablet lubrication design via model-based design of experiments, *International Journal of Pharmaceutics*, 614, 121435.
- I. Bouchkira, A. M. Latifi, L. Khamar, S. Benjelloun, (2021), Global sensitivity based estimability analysis for the parameter identification of Pitzer's thermodynamic model, *Reliability Engineering & System Safety*, 207, 107263.
- I. Bouchkira, A. Latifi, L. Khamar, S. Benjelloun, 2022. Modeling and multi-objective optimization of the digestion tank of an industrial process for manufacturing phosphoric acid by wet process. *Computers & Chemical Engineering*, 156, 107536.
- I. Bouchkira, S. Benjelloun, L. Khamar, A. Latifi, 2023. Thermodynamic modeling and parameter estimability analysis of a wet phosphoric acid process with impurities. *Fluid Phase Equilibria*, 564, 113594.
- IV, J. Kushner, F. Moore, (2010), Scale-up model describing the impact of lubrication on tablet tensile strength, *International journal of pharmaceutics*, 399(1-2), 19-30.
- K. G. Pitt, J. M. Newton, P. Stanley. Tensile fracture of doubly-convex cylindrical discs under diametral loading, *Journal of materials science* 23 (1988), 2723-2728.

# Water bodies restoration as an optimal control problem with integrated internal and external strategies. The inclusion of artificial floating islands

Amira Siniscalchi<sup>a,b</sup>, Laura J. Fritz<sup>a</sup>, Vanina Estrada<sup>a,b</sup>, Maria Soledad Diaz<sup>a,b</sup>

<sup>a</sup> *Planta Piloto de Ingeniería Química (PLAPIQUI CONICET-UNS), Camino La Carrindanga km. 7, Bahía Blanca, Argentina*

<sup>b</sup> *Departamento de Ingeniería Química, Universidad Nacional del Sur (UNS), Bahía Blanca, Argentina*

## Abstract

In this work, we propose an integrated water quality model of an artificial reservoir that is the drinking water source for two cities and undergoes recurrent algal blooms. The integrated model is formulated within a dynamic optimization framework as an optimal control problem to address its restoration through the combination of different restoration techniques. The model includes a global mass balance that takes into account inflows from tributaries and rain, and outflows for drinking water and evaporation. Composition gradients are considered along the water column, rendering a Partial Differential Algebraic Equations (PDAE) system. Dynamic mass balances are formulated for the main biogeochemical variables in both reservoir and wetland, as well as artificial floating islands (AFIs) platform. The PDAE is then transformed into a DAE. Numerical results allow analyzing the performance of the different restoration strategies and their joint application on a eutrophic reservoir, showing the efficiency of AFIs as complement of artificial wetlands.

**Keywords:** eutrophication control, optimal control problem, AFI, dynamic optimization

## 1. Introduction

Eutrophication is a severe environmental problem caused by anthropogenic nutrient enrichment (mainly nitrogen and phosphorus) that causes Cyanobacteria Harmful Algal Blooms (CyanoHABs), and biodiversity and economic losses (WHO, 2020). In 2007 a severe algal bloom in Taihu Lake (China) caused economic losses of US\$ 6.5 billions (Tao and Xin, 2014). CyanoHABs affect public health, commercial and sports fishing, tourism and recreational activities around the world. The potential production of toxins by cyanobacteria (cyanotoxins) is nowadays one of the main environmental and public health concerns (Figueiredo *et al.* 2004). Many water quality restoration strategies have been applied to attempt to mitigate eutrophication and control cyanobacteria growth (Hamilton *et al.* 2016). Since the 90s, there has been a lot of debate about the relative importance of control phytoplankton blooms by nutrients decrease (bottom-up control) (Siniscalchi *et al.* 2020) or by zooplankton grazing (top-down control) (Beklioglu, 1999; Estrada *et al.* 2011). Constructed Wetlands (CW) and Artificial Floating Islands (AFIs) has been both proposed as bottom-up strategies. CW are portions of land covered with



macrophytes (plants that can live in flooded lands), built next to the banks of the tributaries to reduce nitrogen and phosphorus loading by runoff. On the other hand, in AFIs, the macrophytes are sown in buoyancy structures on the surface of the water body and the N and P (and micronutrients) are up-taken by roots directly from the water column.

In this work, we propose an integrated water quality mechanistic model of an artificial reservoir (Siniscalchi *et al.*, 2020) that is the drinking water source for more than 450,000 people and undergoes recurrent algal blooms. The integrated model includes the water quality model for Paso de las Piedras Reservoir (PPR), CWs model for the two reservoir's tributaries, and the AFIs model. It is formulated within a dynamic optimization framework as an optimal control problem to address its restoration through the combination of CW and AFIs strategies. The optimization integrated model includes CWs and AFIs areas as design variables and the fractions of tributaries that are diverted through the wetlands as control variables.

## 2. Case Study

Paso de las Piedras Reservoir, is an eutrophic freshwater reservoir, that supplies drinking water to more than 450,000 inhabitants of two cities and for industrial purposes at a petrochemical complex nearby. It has two tributaries that run through an important agricultural area in the country, with a drainage basin area of 1620 km<sup>2</sup>. The lake has a surface area of 36 km<sup>2</sup>, while its mean depth is 8.2 m and its retention time is 4 years. The eutrophic characteristic of the reservoir is mainly caused by the external discharge of nutrients (mainly P and N). This fact causes recurrent algal blooms, which produce several problems in the potabilization process.

## 3. Mathematical Model description

A first principle-based ecological water quality model for Paso de las Piedras Reservoir (PPR model) was previously formulated, calibrated and validated (Estrada *et al.*, 2009 a,b, 2011; Di Maggio *et al.*, 2016). This model includes dynamic mass balances for main biogeochemical components in the water body: three taxonomical groups of phytoplankton (cyanobacteria, chlorophytes and diatoms), two groups of zooplankton (copepoda and cladocera), three nitrogen species (organic nitrogen, nitrate and ammonia), two phosphorus species (organic phosphorus and orthophosphate), dissolved and particulate organic carbon and dissolved oxygen. and calibrate the model with experimental data (Estrada *et al.*, 2009a), respectively. The model was formulated within an optimization framework to evaluate the application of top-down restoration strategies (Estrada *et al.*, 2011). Also, the model includes a global mass balance that takes into account inflows from tributaries and rain, and outflows for drinking water, and evaporation. Composition gradients are considered along the water column, rendering Partial Differential Algebraic Equations (PDAE) system. The PDAE is then transformed into an ordinary Differential Algebraic Equation system (DAE) by spatial discretization in horizontal layers. Siniscalchi *et al.* (2020) integrated the PPR model with a constructed wetland model (CWs model) which includes the possibility of installing CW in both tributaries of the lake. The model includes mass balances for: macrophyte biomass, three nitrogen species (organic nitrogen, nitrate and ammonia), two phosphorus species (organic phosphorus and orthophosphate), dissolved and particulate organic carbon and dissolved oxygen. For a detailed description of this model see Siniscalchi *et al.* (2020). In this work we formulated the AFIs model integrated with PPR model and CWs model in order to evaluate different restoration strategies. AFIs model includes a mass balance

for macrophytes (Eq. 1). The experimental data to set the parameters used for plant growth and harvest both in AFIs and in CWs correspond to the species *Senecio bonariensis*, a native plant with known remedial potential (Sinscalchi, 2013). The *S. bonariensis* biomass balance ( $C_{Biomass}$  in  $\text{mgC L}^{-1}$ ) includes a growth term ( $R_{Biomass, growth}$ ) ( $\text{d}^{-1}$ ) (Eq. 2) that considers the maximum growth of the planted species multiplied by the main factors that affect photoautotrophic organisms: temperature ( $f(T)$ ) (Eq. 3), solar radiation ( $f(I)$ ) (Eq. 4), P and N concentrations ( $f(N)$ ) (Eq. 5) from the upper layer of the reservoir; a mortality term ( $R_{Biomass, death}$ ) ( $\text{d}^{-1}$ ) (Eq. 6) dependent on temperature, and a harvest term ( $R_{Biomass, harvest}$ ) ( $\text{d}^{-1}$ ) (Eq. 7) that is performed only when the plant reaches its maximum growth.

$$\frac{dC_{Biomass}}{dt} = r_{Biomass} = R_{Biomass, growth} - R_{Biomass, death} - R_{Biomass, harvest} \quad \text{Eq. 1}$$

$$R_{Biomass, growth} = k_{Biomass, growth} f(T)f(I)f(N)C_{Biomass} \quad \text{Eq. 2}$$

$$f(T) = \left(\frac{T_{emp}}{T_{opt}}\right) \exp\left(1 - \left(\frac{T_{emp}}{T_{opt}}\right)\right) \quad \text{Eq. 3}$$

$$f(I) = \left(\frac{I_o}{I_{opt}}\right) \exp\left(1 - \left(\frac{I_o}{I_{opt}}\right)\right) \quad \text{Eq. 4}$$

$$f(N) = \left(\frac{C_{upper layer, PO_4}}{K_P + C_{i, PO_4}}\right) \left(\frac{C_{upper layer, NO_3}}{K_N + C_{i, NO_3}}\right) \quad \text{Eq. 5}$$

$$R_{Biomass, death} = k_{Biomass, death} \theta_{Biomass, death}^{(T-20)} C_{Biomass} \quad \text{Eq. 6}$$

$$R_{Biomass, harvest} = \begin{cases} 0, & \text{if } R_{Biomass, growth} < k_{Biomass, growth max} \\ \beta * C_{Biomass}, & \text{if } R_{Biomass, growth} = k_{Biomass, growth max} \end{cases} \quad \text{Eq. 7}$$

Were  $Temp$  and  $T_{opt}$  are the water temperature and the optimal growth temperature ( $^{\circ}\text{C}$ ), respectively;  $I_o$  and  $I_{opt}$  are irradiance and optimal irradiance for growth ( $\text{wm}^{-2}$ ), respectively;  $K_P$  and  $K_N$  are the half saturation constants ( $\text{mgL}^{-1}$ ) for P and N uptake;  $k_{Biomass, growth}$  and  $k_{Biomass, death}$  are the maximum growth rate and the mortality rate ( $\text{d}^{-1}$ );  $\theta_{Biomass, death}$  is the temperature for mortality rate; and  $\beta$  is the proportion of harvested biomass. As the Paso de las Piedras Reservoir model is spatially discretized in two horizontal layers, and the AFIs implementation leads to a vertical difference in components concentrations (i.e., Areas covered and non-covered by AFIs), the model was again discretized including two vertical compartments in the upper layer.

#### 4. Optimization Problem

The objective function (Eq 8, Table 1) follows the dual criteria (P and N) recommended by Environmental Protection Agency of U.S. (EPA, 2015) and is the minimization of the sum of integrals of the square differences between inorganic phosphorus (P- $\text{PO}_4$ ) and inorganic nitrogen (N- $\text{NO}_3$ ) concentrations and a desired value below eutrophication (0.02 for P and 0.3  $\text{mg L}^{-1}$  for N) (Wetzel, 1983). As degrees of freedom of the control problems are the area of the constructed wetlands for El Divisorio Stream ( $CW_{ED}$ ) and Sauce Grande River ( $CW_{SG}$ ) (design variables) (Eqs. 11 and 12, Table 1), the area covered by AFIs ( $A_{AFI}$ ) (design variable) (Eq. 18, Table 1) (Saviolo Osti *et al.*, 2020), and the fraction of the tributaries diverted to the wetlands ( $F_{CW_{SG}}$  and  $F_{CW_{ED}}$ ) (control variables) (Eqs. 14 and 15 Table 1). Also, a path constraint is included (Eq- 13, Table 1) to satisfy that the sum of the areas ( $A_{CW, Total}$ ) of both wetlands ( $A_{CW_{SG}}$  and  $A_{CW_{ED}}$ ) does not exceed a quarter of the total area of the reservoir (9  $\text{km}^2$ ). To solve the formulated optimal control

and design problems, a control vector parameterization approach is applied within an equation-oriented framework (gPROMS, PSEnterprise Ltd., 2019).

**Table 1.** Design and optimal control problems

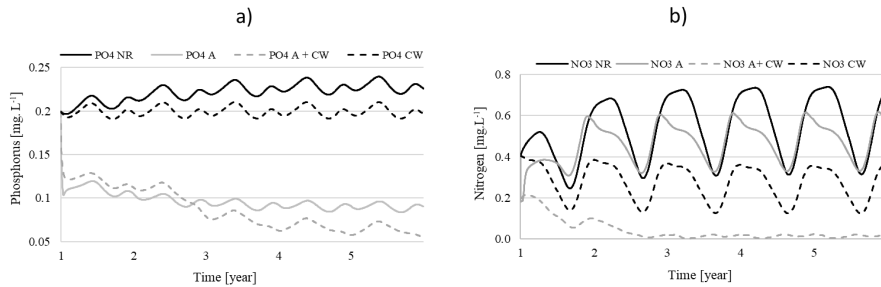
$\min z \int_0^{t_f} (C_P(t) - 0.02)^2 dt + \int_0^{t_f} (C_N(t) - 0.3)^2 dt$		Eq. 8	
Subject to			
DAE Water Quality Model (PPR model) + CWs model + AFIs model			
$C(0) = C^0$		Eq. 9	
$C^L \leq C \leq C^U$		Eq. 10	
CWs constraints		AFIs constraints	
$V_{CW_{SG}} = A_{CW_{SG}} h_{CW_{SG}}$	Eq. 11	$A_{AFI\_Total} = A_{AFI} - A_{no\_AFI}$	Eq. 16
$V_{CW_{ED}} = A_{CW_{ED}} h_{CW_{ED}}$	Eq. 12	$V_{AFI} = A_{AFI} * h_{upper\_layer}$	Eq. 17
$0 \leq A_{CW_{SG}} + A_{CW_{ED}} \leq 0.25 A_{CW\_Total}$	Eq. 13	$0 \leq A_{AFI} \leq 0.5 A_{Total}$	Eq. 18
$0 \leq F_{CW_{SG}} \leq 0.1$	Eq. 14		
$0 \leq F_{CW_{SG}} \leq 0.5$	Eq. 15		

## 5. Numerical Results

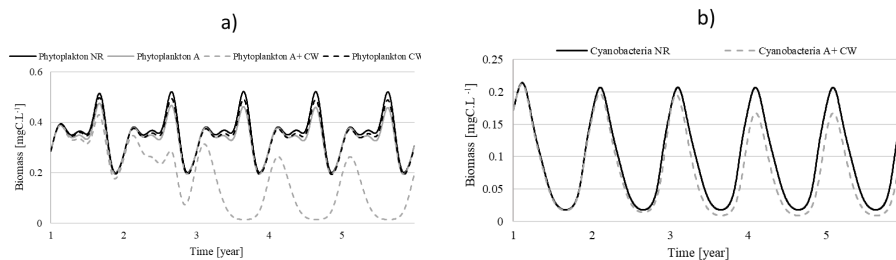
The resulting integrated (PPR+AFIs+CWs) model has 64 differential equations and more than 200 algebraic ones. The model is implemented in a personal computer with an Intel® Core™ i7-4700 K processor, operating at a CPU frequency of 3.6 GHz, and with 8 Gb RAM. The total CPU time is 4.469s. Numerical results provide optimal areas of AFIs and CWs, as well as nutrients, phytoplankton, zooplankton, fish, dissolved oxygen, *S. bonariensis* biomass in both CWs and AFIs profiles along a five-year time horizon. Numerical results (Figures 1 and 2, Table 2) allow assesses the performance of the different bottom-up strategies and their combination to restore the water quality in a eutrophic reservoir, showing the synergy between both combined strategies (CWs and AFIs). With the application of this approach the theoretical reduction of inorganic P and N is 61 % and 92 %, respectively. It is important to note that when only CWs strategy is applied, the reduction of 10 % in P (Table 2) is achieved with 3 and 6 km<sup>2</sup> of CW<sub>SG</sub> and CW<sub>ED</sub>, respectively, which seems inapplicable in practice. However, with an area covered by AFIs of 1.8 km<sup>2</sup> (5% of the total area of the reservoir, Saviolo Osti *et al.*, 2020) the necessity for CWs area is halved and it is only necessary to construct a wetland on El Divisorio Stream. With this combination P reduction is 56 % (Table 2). Figures 1 and 2 shows that the best strategy is the combination of restoration techniques (AFIs + CWs) approach for the substantial reduction in nutrients and, also, for phytoplankton and cyanobacteria concentration.

**Table 2.** Nutrient removal and decrease in phytoplankton and cyanobacteria biomass concentrations for each restoration strategy in percentage, as compared to no restoration

	Phosphorous	Nitrogen	Phytoplankton	Cyanobacteria
AFIs (%)	56	14	5	1
CW <sub>SG</sub> and CW <sub>ED</sub> (%)	10	48	3	1
AFIs + CW <sub>ED</sub> (%)	61	92	50	20



**Figure 1:** Concentration time profiles for **a)** inorganic phosphorus and **b)** inorganic nitrogen concentrations. NR: without restoration; A: AFIs; CW: Constructed wetlands.



**Figure 2:** Concentration time profiles for **a)** total phytoplankton biomass and **b)** cyanobacteria biomass concentrations. NR: without restoration; A: AFIs; CW: Constructed wetlands.

## 6. Conclusions

In this work we study the water quality improvement after combined restoration strategies application formulated as optimal control problems. The integrated optimization model includes a water quality model of Paso de las Piedras Reservoir, and models for two bottom-up restoration strategies, constructed wetlands and floating artificial islands. Numerical results indicate that implementation of AFIs would be an effective green technology to restore contaminated ecosystems by reducing excess of nutrients and controlling cyanobacteria and total phytoplankton growth in combination with constructed wetlands. Finally, integrated models used with dynamic optimization approaches are powerful management tools for rational evaluation and planning of bottom-up strategies for water quality restoration.

## References

- M. Beklioğlu, 1999, A review on the control of eutrophication in deep and shallow lakes. Turkish Journal of Zoology, 23, 327-336.
- J. Di Maggio, C. Fernández, E.R. Parodi, M.S. Díaz, V. Estrada, 2016, Modeling phytoplankton community in reservoirs. A comparison between taxonomic and functional groups-based models, Journal of Environmental Management, 165, 31-52.

- V. Estrada, E.R. Parodi, M.S. Diaz, 2009a, Determination of biogeochemical parameters in eutrophication models with simultaneous dynamic optimization approaches, *Computers and Chemical Engineering*, 33, 1760-1769.
- V. Estrada, E.R., Parodi, M.S., Diaz, 2009b, Addressing the control problem of algae growth in water reservoirs with advanced dynamic optimization approaches, *Computers and Chemical Engineering*, 33, 1598-1613.
- V. Estrada, J. Di Maggio, M.S. Díaz, 2011, Water sustainability: a systems engineering approach to restoration of eutrophic lakes, *Computers and Chemical Engineering*, 35, 1598-1613.
- EPA, 2015, Preventing eutrophication: scientific support for dual nutrient criteria. Environmental Protection Agency EPA- 820-S-15-001.
- D.P. Hamilton, N. Salmaso, H.W. Paerl, 2016, Mitigating harmful cyanobacterial blooms: strategies for control of nitrogen and phosphorus loads, *Aquatic Ecology*, 50, 351-366
- D.R. de Figueiredo, U.M. Azeiteiro, S.M. Esteves, F. Goncalves, M.J. Pereira, 2004, Microcystin-producing blooms—a serious global public health issue, *Ecotoxicology and Environmental Safety*, 59, 151–163.
- PSEnterprise Ltd., 2019. gPROMS Advanced User Guide-release 2.3. Process Systems Enterprise Ltd., London.
- J.A. Saviolo Osti, C. Ferreira do Carmo, M. A. Silva Cerqueira, M.T. Duarte Giamas, A.C. Peixoto, A. Vaz-dos-Santos, C.T. Mercante, 2020, Nitrogen and phosphorus removal from fish farming effluents using artificial floating islands colonized by *Eichhornia crassipes*, *Aquaculture Reports*, 17, 100324.
- A. Siniscalchi, 2013, Fitorremediación de un arroyo eutrofizado mediante el cultivo de dos especies autóctonas: *Senecio bonariensis* (Compositae) y *Cladophora surera* (Chlorophyta). <http://repositoriodigital.uns.edu.ar/handle/123456789/507>.
- A. Siniscalchi, J.A. Di Maggio, V. Estrada, M.S. Díaz, 2020, Integrated mathematical models for drinking water reservoirs and constructed wetlands as a tool for restoration planning. *Journal of Hydrology*, 586, 124867.
- T. Tao, K.L. Xin, 2014, A sustainable plan China's drinking water, *Nature*, 511, 527-528.
- R.G. Wetzel, 1983, *Limnology*, second ed. Saunders College Publishing, pp. 860.
- WHO, 2020, Cyanobacterial toxins: microcystins. Background document for development of WHO Guidelines for drinking-water quality and Guidelines for safe recreational water environments. Geneva: World Health Organization; 2020. Licence: CC BY-NC-SA 3.0 IGO. Switzerland.

## Developing robust hybrid-models

Peter Jul-Rasmussen<sup>a,b</sup>, Xiaodong Liang<sup>a</sup>, Xiangping Zhang<sup>b</sup>, Jakob Kjøbsted Huusom<sup>a</sup>

<sup>a</sup>*Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søtofts Plads 228A, 2800 Kgs. Lyngby, Denmark.*

<sup>b</sup>*Institute of Process Engineering, Chinese Academy of Sciences, 1 North 2nd street, Zhongguancun, Haidian District, Beijing, PR China*

### Abstract

With the increased availability of process data, data-driven methods are becoming more appealing to use in process modelling. Data-driven methods are especially interesting in systems where the existing first-principles models are not adequate in describing the full system dynamics. This work compares a serial semi-parametric hybrid-modelling approach, integrating Machine Learning and first-principles modelling, to a stochastic greybox modelling approach proposed by (Kristensen et al. 2004). Through a CSTR case study both modelling approaches shows accurate predicting performance and robustness towards noisy data. The stochastic greybox modelling is found to be more robust towards measurement frequency as the true model structure is provided when fitting the parameters. Through this work, it is found that good modelling performance can be achieved without providing an explicit model structure by utilizing process data.

**Keywords:** Process Simulation, Machine Learning, Hybrid-Modelling, Kalman Filter

### 1. Introduction

With Industry 4.0 the accessibility and quality of data from the production lines in the chemical and biochemical industries is improving due to digital advances such as better data infrastructure and more advanced sensors for on-line/at-line measurements. The increased availability of process data along with improved computational power and algorithmic developments is improving the possibility for introducing data-driven modelling methods in chemical engineering applications (Udugama et al. 2020). The use of data-driven methods enables the modelling of systems where the existing physical and chemical knowledge is limited and, as a result, the existing first-principles models are inaccurate. A problem arising when using data-driven methods is the process data not being truly “big-data” (Venkatasubramanian 2019), a usual requirement for the development of accurate Machine Learning (ML) models. Problems often also arise in predicting the states, interpreting the models, and extrapolating the models when using data-driven modelling. The introduction of semi-parametric hybrid-models, in which process knowledge is combined with data-driven methods, has shown promising performance in mitigating the issues in “purely” data-driven modelling. Semi-parametric hybrid-models shows improvements in state prediction, interpretability, and extrapolability compared to “purely” data-driven modelling when using the same data (von Stosch et al. 2014). Semi-parametric hybrid-models have previously been used in chemical reactor modelling (Bellos et al. 2005) and for modelling flocculation processes (Nazemzadeh et al. 2021). Experimental knowledge can also be included in modelling through stochastic greybox modelling in which maximum likelihood or maximum a

posteriori parameter estimates are used for modelling using a system of stochastic differential equations (Kristensen et al. 2004). Stochastic greybox modelling has proven accurate in modelling several different processes including fed-batch cultivation (Rasmussen et al. 2006) and alternating activated sludge process (Halvgaard et al. 2015).

## 2. Serial semi-parametric hybrid-modelling

Lack of process knowledge is mitigated in semi-parametric hybrid-models by modelling unknown or uncertain phenomena in the system using ML. In this work, a serial semi-parametric hybrid-model is introduced using the workflow given in figure 1.

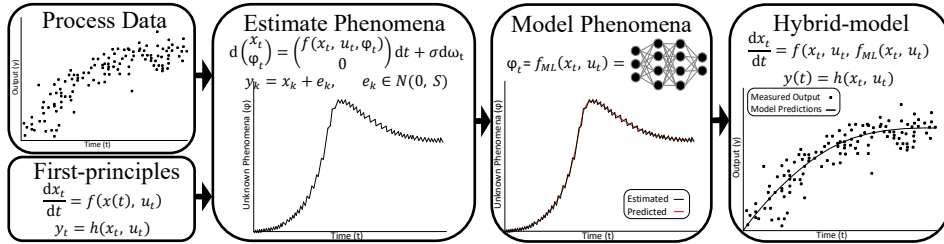


Figure 1 Workflow for serial semi-parametric hybrid-modelling using ML model trained on unknown phenomena,  $\varphi_t$ , estimated for training and validation data using a Continuous-Discrete Extended Kalman Filter.

Considering a process, which is described by a set of ODEs as given in equation (1).

$$dx_t = f(x_t, u_t, \varphi_t)dt \quad (1)$$

In which  $t \in \mathbb{R}$  is the time,  $x_t \in \mathbb{R}^n$  is the state variables,  $u_t \in \mathbb{R}^m$  is the input variables,  $f(\cdot) \in \mathbb{R}^n$  is a non-linear function based on first-principles, and  $\varphi_t \in \mathbb{R}^p$  is some unknown phenomena in the system. The unknown phenomena,  $\varphi_t$ , is modelled by introducing a ML model,  $f_{ML}(\cdot) \in \mathbb{R}^p$ , such that  $f_{ML}(x_t, u_t) = \varphi_t$  resulting in the system equations (2).

$$dx_t = f(x_t, u_t, f_{ML}(x_t, u_t))dt \quad (2)$$

A problem arise in training the ML model using supervised learning as the unknown phenomena,  $\varphi_t$ , often cannot be measured. This problem is mitigated by estimating  $\varphi_t$  for the training and validation data before training the ML model.  $\varphi_t$  can be estimated using a Continuous-Discrete Extended Kalman Filter (CDEKF) in which  $\varphi_t$  is included as an additional state (Cox 1964) with system equation (3) and measurement equation (4).

$$d \begin{pmatrix} x_t \\ \varphi_t \end{pmatrix} = \begin{pmatrix} f(x_t, u_t, \varphi_t) \\ 0 \end{pmatrix} dt + \sigma(u_t, \varphi_t) d\omega_t \quad (3)$$

$$y_k = h(x_k, u_k, \varphi_k) + e_k \quad (4)$$

In which  $k = 0, \dots, N$  is the discrete time index,  $y_k \in \mathbb{R}^l$  is the measured output variables,  $\sigma(\cdot) \in \mathbb{R}^{n \times n}$  and  $h(\cdot) \in \mathbb{R}^l$  are functions,  $\omega_t$  is a  $n$ -dimensional standard Wiener process, and  $e_k \in N(0, S)$  is an  $l$ -dimensional white noise process. A ML model can now be trained using the CDEKF estimates of  $\varphi_t$  before introducing the ML model to the system equations (2).

### 3. Case Study: Synthetic CSTR

To investigate the performance of different data-driven modelling methods, a CSTR case study (Figure 2) is considered using simulated data. A first-order exothermic conversion of  $A$  to  $B$  is taking place in the CSTR system with the material- and energy balance (5) being the governing equations.

$$d \begin{pmatrix} C_A \\ C_B \\ T \end{pmatrix} = \begin{pmatrix} \frac{F}{V} (C_{A,f}(t) - C_A(t)) - r(t) \\ -\frac{F}{V} C_B(t) + r(t) \\ \frac{F}{V} (T_f(t) - T(t)) - \frac{\Delta H}{\rho c_p} r(t) \end{pmatrix} dt \quad (5)$$

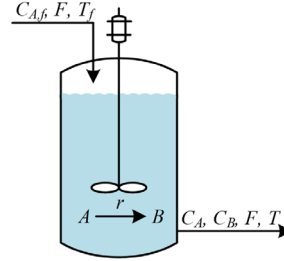


Figure 2 Systematic diagram of CSTR

In which  $C_A$  and  $C_B$  are the concentrations of  $A$  and  $B$  respectively,  $T$  is the temperature in the CSTR,  $F$  is the feed flowrate,  $V$  is the volume of the CSTR.  $C_{A,f}$  is the concentration of  $A$  in the feed,  $T_f$  is the temperature in the feed.  $\Delta H$  is the heat of reaction,  $\rho$  is the density of the mixture, and  $c_p$  is the heat capacity coefficient of the mixture. For the CSTR case, both stochastic greybox modelling and serial semi-parametric hybrid-modelling is applied, introducing data-driven methods for modelling  $r(t)$ .

#### 3.1 Stochastic Greybox Modelling

Three different model structures are introduced for the reaction rate in the stochastic greybox modelling: 1) assuming 0<sup>th</sup> order reaction rate,  $r(t) = r$ , 2) assuming 1<sup>st</sup> order reaction,  $r(t) = k \cdot C_A(t)$ , and 3) using an Arrhenius expression,  $r(t) = k_0 \exp\left(-\frac{E_a}{RT(t)}\right) C_A(t)$ .

Table 1 Maximum likelihood estimates of parameters for stochastic greybox models.

Model	Parameter
0 <sup>th</sup> Order	$r = 1.78 \cdot 10^{-1}$
1 <sup>st</sup> Order	$k = 1.81 \cdot 10^{-1}$
Arrhenius	$k_0 = 3.50 \cdot 10^7$
	$E_a = 1.18 \cdot 10^4$

Plotting the model predictions along with the measurements (Figure 3) it is evident that the models using 0<sup>th</sup>- and 1<sup>st</sup> order reaction rates are performing significantly worse than the model using an Arrhenius expression.

#### 3.2 Serial semi-parametric Hybrid-modelling

A semi-parametric hybrid-modelling structure is constructed by introducing a feedforward artificial neural network ( $f_{NN}$ ) for modelling the reaction rate resulting in the system equations (6).

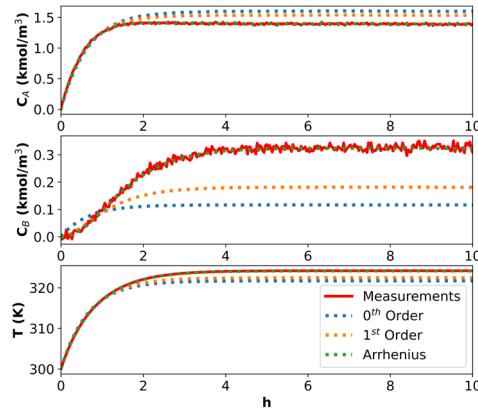


Figure 3 Stochastic greybox model predictions of output variables.



$$d \begin{pmatrix} C_A \\ C_B \\ T \end{pmatrix} = \begin{pmatrix} \frac{F}{V} (C_{A,f}(t) - C_A(t)) - f_{NN}(C_A(t), T(t)) \\ -\frac{F}{V} C_B(t) + f_{NN}(C_A(t), T(t)) \\ \frac{F}{V} (T_f(t) - T(t)) - \frac{\Delta H}{\rho c_p} f_{NN}(C_A(t), T(t)) \end{pmatrix} dt \quad (6)$$

The reaction rate is first estimated using a CDEKF in which the reaction rate is included as an additional state resulting in the system- and measurement equations (7).

$$d \begin{pmatrix} C_A \\ C_B \\ T \\ r \end{pmatrix} = \begin{pmatrix} \frac{F}{V} (C_{A,f}(t) - C_A(t)) - r(t) \\ -\frac{F}{V} C_B(t) + r(t) \\ \frac{F}{V} (T_f(t) - T(t)) - \frac{\Delta H}{\rho c_p} r(t) \\ 0 \end{pmatrix} dt, \quad \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}_k = \begin{pmatrix} C_A \\ C_B \\ T \end{pmatrix}_k + e_k \quad (7)$$

The CDEKF is applied to the training data, for training a neural network (NN) with one hidden layer, and to the validation data, for selecting hyper-parameters using a grid search. The possible number of neurons,  $n$ , are specified in the grid search to be  $n \in [2, 4, 8, 16, 32, 64, 128]$ . The best prediction performance on the validation data set was achieved in a NN with 2 neurons in the hidden layer (Figure 4).

Predictions of the three output variables are obtained after introducing the trained NN model to the system equations (6). Plots of the output predictions (Figure 5) shows that the serial semi-parametric hybrid-model provides accurate predictions for all three output variables.

#### 4. Robustness

When modelling real production processes the available data may be noisy and the measurement frequency may be low, especially when using off-line measurements. In the following an assessment of the robustness of the two different modelling approaches towards noise and measurement is provided.

##### 4.1. Noise

White noise is added in the generated data sets as both output measurement noise and input noise on the three input parameters,  $C_{A,f}$ ,  $T_f$ , and  $F$ , with the same standard deviation

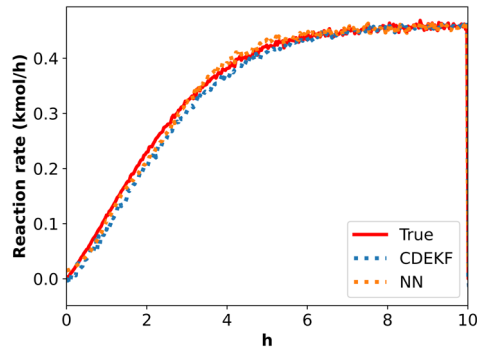


Figure 4 Predictions, estimates, and true reaction rate.

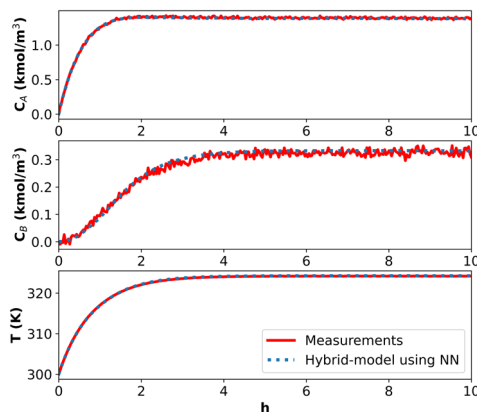


Figure 5 Serial semi-parametric hybrid-model predictions of output variables.

for all noise terms in the individual datasets. Three different levels of noise is investigated using standard deviation of 0.01, 0.02, and 0.05. Both stochastic greybox models using an Arrhenius expression as the reaction rate, and serial semi-parametric hybrid-models using a NN as reaction rate are trained using each of the noisy data sets.

Table 2 fitness of prediction with different noise levels for the three output variables using both stochastic greybox model with Arrhenius expression as reaction rate and serial semi-parametric hybrid-model.

Noise std.	Arrhenius			Serial semi-parametric hybrid-model		
	$C_A$ fitness	$C_B$ fitness	$T$ fitness	$C_A$ fitness	$C_B$ fitness	$T$ fitness
	MAE ( $R^2$ )	MAE ( $R^2$ )	MAE ( $R^2$ )	MAE ( $R^2$ )	MAE ( $R^2$ )	MAE ( $R^2$ )
0.01	0.008 (0.99)	0.008 (0.87)	0.014 (0.99)	0.009 (0.99)	0.009 (0.85)	0.053 (0.99)
0.02	0.017 (0.97)	0.016 (0.69)	0.016 (0.99)	0.018 (0.97)	0.018 (0.66)	0.095 (0.99)
0.05	0.041 (0.88)	0.039 (0.38)	0.041 (0.99)	0.041 (0.88)	0.040 (0.36)	0.093 (0.99)

The MAE is approximately the same as the noise standard deviation for all output predictions at each of the noise levels with the exception of the  $T$  predictions when using a serial semi-parametric hybrid-model (Table 2). This indicates that the two different approaches are able to predict the trend in the data, while the increase in error is a result of the increase in the noise.

#### 4.2. Measurement Frequency

The robustness towards measurement frequency of the two modelling approaches is investigated using data sets generated using different sampling time. The NN in the serial semi-parametric hybrid-model is sensitive to sparse data, posing a problem when training the model. This problem is mitigated in the data sets with low measurement frequency by interpolating between the measurements using cubic smoothing splines when preprocessing the training data, generating more data for training.

Table 3 fitness of prediction with different measurement frequency for the three output variables using both stochastic greybox model with Arrhenius expression as reaction rate and serial semi-parametric hybrid-model.

Freq. ( $hr^{-1}$ )	Arrhenius			Serial semi-parametric hybrid-model		
	$C_A$ fitness	$C_B$ fitness	$T$ fitness	$C_A$ fitness	$C_B$ fitness	$T$ fitness
	MAE ( $R^2$ )	MAE ( $R^2$ )	MAE ( $R^2$ )	MAE ( $R^2$ )	MAE ( $R^2$ )	MAE ( $R^2$ )
30	0.008 (0.99)	0.008 (0.87)	0.014 (0.99)	0.009 (0.99)	0.009 (0.85)	0.053 (0.99)
15	0.008 (0.99)	0.008 (0.87)	0.008 (0.99)	0.010 (0.99)	0.009 (0.84)	0.066 (0.99)
3	0.007 (0.99)	0.008 (0.89)	0.008 (0.99)	0.008 (0.99)	0.009 (0.86)	0.041 (0.99)
1	0.006 (0.99)	0.007 (0.92)	0.012 (0.99)	0.018 (0.99)	0.016 (0.86)	0.190 (0.99)

The stochastic greybox model is found to be robust towards measurement frequency, as the fitness does not decrease when lowering the measurement frequency (Table 3). The serial semi-parametric hybrid-model is found to be less robust towards measurement frequency, as the MAE is increasing when the measurement frequency is low. The higher robustness of the stochastic greybox model towards measurement frequency may be a result of the higher utilization of process knowledge through the introduction of the “true” structure for the reaction rate in the model. Effectively data is only used to fit two parameters in the stochastic greybox model, whereas the full structure of the reaction rate has to be learned by the NN in the serial semi-parametric hybrid-model.

## 5. Conclusion

Two different approaches for applying hybrid-models, stochastic greybox modelling and serial semi-parametric hybrid-modelling, were used in this paper. The two approaches were applied to a CSTR case study using simulated data. Both approaches were successful in developing models with accurate output predictions and robustness towards noisy data, with the stochastic greybox model being slightly more accurate. The stochastic greybox model was found to be more robust towards measurement frequency due to the “true” structure of the reaction rate being provided for the model. While the stochastic greybox modelling provides a better model, it has higher requirements for process knowledge, as an appropriate model structure has to be provided for good performance to be achieved. It is found that good prediction performance can be achieved without providing an explicit model structure by introducing a ML part for modelling unknown or uncertain phenomena to the first-principles model.

## Acknowledgements

This study was financially supported by the Sino-Danish Center for Education and Research (SDC).

## References

- G. D. Bellos, L. E. Kallinikos, C. E. Gounaris, N. G. Papayannakos, 2005, Modelling of the performance of industrial HDS reactors using a hybrid neural network approach, *Chemical Engineering and Processing: Process Intensification*, 44 (5), 505-515
- H. Cox, 1964, On the estimation of state variables and parameters for noisy dynamic systems, *Ieee Transactions on Automatic Control*, Volume AC-9 (1), 5-12
- R. F. Halvgaard, L. Vezzaro, M. Grum; T. Munk-Nielsen, P. Tychsen, H. Madsen, 2015, Stochastic Greybox Modeling for Control of an Alternating Activated Sludge Process, 9th IWA Symposium on Systems Analysis and Integrated Assessment (Watermatex 2015)
- N. R. Kristensen, H. Madsen, S. B. Jørgensen, 2004, A Method for Systematic Improvement of Stochastic Grey-Box Models, *Computers and Chemical Engineering*, 28 (8), 1431-1449
- N. Nazemzadeh, A. A. Malanca, R. F. Nielsen, K. V. Gernaey, M. P. Andersson, S. S. Mansouri, 2021, Integration of first-principle models and machine learning in a modeling framework: An application to flocculation, *Chemical Engineering Science*, 245, 116864
- J. K. Rasmussen, H. Madsen, S. B. Jørgensen, 2006, Grey-box stochastic modelling of industrial fed-batch cultivation, *Computer Aided Chemical Engineering*, 21 (C), 421-426
- I. A. Udogama, C. L. Gargalo, Y. Yamashita, M. A. Taube, A. Palazoglu, B. R. Young, K. V. Gernaey, M. Kulaçci, C. Bayer, 2020, The Role of Big Data in Industrial (Bio)chemical Process Operations, *Industrial and Engineering Chemistry Research*, 59 (34), 15283-15297
- V. Venkatasubramanian, 2019, The promise of artificial intelligence in chemical engineering: Is it here, finally?, *Aiche Journal*, 65 (2), 466-478
- M. von Stosch, et. al., 2014, Hybrid semi-parametric modeling in process systems engineering: Past, present and future, *Computers and Chemical Engineering*, 60, 86-101

# A gPC-based Global Sensitivity Analysis for Phosphate Slurry Flow in Pipelines

Marwane Elkarii,<sup>a,b</sup> Radouan Boukharfane,<sup>b</sup> Saad Benjelloun,<sup>b</sup> *Chakib Bouallou,<sup>a</sup> Nabil Elmocayd,<sup>c,d</sup>*

<sup>a</sup> *Mohammed VI Polytechnic University (UM6P), MSDA Group, Benguerir, Morocco*

<sup>b</sup> *MINES ParisTech, PSL Research University, Center of Energy Efficiency of Systems (CES), Paris, France*

<sup>c</sup> *Mohammed VI Polytechnic University (UM6P), IWRI & IAP, Benguerir, Morocco*  
*marwane.elkarii@mines-paristech.fr*

## Abstract

The present work focuses on a new application of the Generalized Polynomial Chaos (gPC) approach: the stochastic response of a three-dimensional simulation of slurry pipe flows in pipelines subject to parametric uncertainties. Initial and boundary conditions (e.g., the slurry flowrate used at the pipe entrance, initial solid concentration), material properties (e.g., particles size), model parameters (e.g., specular coefficient when the kinetic theory of granular flow (KTGF) is coupled with the Eulerian-Eulerian model), and geometry-related factor (e.g., pipe inclination) are considered as random parameters. gPC surrogate model is built through a least angle regression (LAR) methodology in order to perform uncertainty quantification and global sensitivity analysis following a variance-based approach. The use of gPC is motivated based on its ability to estimate Sobol' indices efficiently. These variance-based sensitivity indices are effective to perform sensitivity analysis without any assumptions about the model's linearity or monotony. Retaining the gPC technique has the advantage of giving the global sensitivity Sobol' indices in a straightforward manner at a lower computing cost than the usual Monte Carlo (MC) method. The first order and total Sobol' indices of the pressure drop along the pipe are calculated and their inspection show that the variability of the pressure gradient is mainly due to the principal effects of the inlet velocity, followed by the inclination of the pipe and then the size of particles. Within the framework of uncertainty quantification, the gPC expansions will also be applied as a surrogate model, as its objective is to recreate the global behavior of the CFD model in a manner that is consistent with a polynomial decomposition.

**Keywords:** Phosphate slurry flow, Generalized polynomial chaos, CFD, Global sensitivity analysis

## 1. Introduction

With the advancement of computational methods, numerical simulation has become a popular method for predicting the behavior of physical systems. Upon its simulation with a numerical model, uncertainties are ubiquitous and arise from a variety of sources that may contain substantial randomness in their description. Hence, the results in model outcomes are necessarily subject to uncertainty. To improve prediction reliability, it is consequently necessary to quantify the related uncertainties. This is generally achieved through the non-linear propagation of the initial parameter's uncertainties in CFD simulations (Walters and Huyse, 2002). Global Sensitivity Analysis (GSA) can be carried out to attribute to each input its contribution to the resulting output uncertainty. In the

previous two decades, several methodologies have been developed to assess the uncertainty propagation in CFD simulation. The reader is referred to Saltelli et al. (2006) who provides a thorough state-of-the-art of the methods. As sensitivity metrics, the Sobol' indices are extensively used in this context. Several approaches can be employed for calculating the Sobol indices (Sobol, 2001; Kucherenko et al., 2005). Nevertheless, the cornerstone of this method is the highly expensive cost of computations needed to estimate these indices (Sudret, 2008). To overcome the computational requirement associated with the stochastic simulations, new methodologies based on the use of surrogate models have been introduced. gPC expansions have been demonstrated to be effective for Uncertainty Quantification (UQ) and GSA (Blatman and Sudret, 2011), among other options.

To this purpose, the issue of slurry flows in pipes is tackled using GSA together with the gPC expansion approach, to rank the input parameters based on their influence on quantity of interest. As model inputs, five variables are investigated and sampled within a physical range. Initially, we address the physical parameters defining the transport of the intended slurry, including the solids particle size distribution, solid concentration in the flow, and the entrance velocity of the flow mixture. In slurry transport, the inclination of the pipe, a geometry-related factor, is also taken into consideration as a crucial component. Furthermore, the specular coefficient, which is a wall boundary coefficient related to the kinetic theory for granular flows (Zhong et al., 2015), is considered as a modeling parameter. After identifying the model's uncertain inputs and modeling them within a probabilistic framework, the second stage involves propagating the uncertainty across the model. Finally, a sensitivity analysis is performed, in which the input parameters are ranked according to their effect on the prediction uncertainty.

## 2. Polynomial Chaos Expansion for Sensitivity Analysis

The gPC expansion aims at recreating the global behavior of a simulation in a manner that is consistent with a polynomial decomposition. The latter are multivariate orthogonal polynomials and serve as the basis functions that are selected in accordance with the joint probability distributions of the stochastic input variables following the so-called Askey scheme of polynomial (Xiu and Karniadakis, 2002). In the present study, if our variables (shown in Tab. 1) are stored in the vector  $\mathbf{X} \in R^d$ , and  $\mathcal{Y}$  represents the model responses discussed earlier, then one may write, according to the gPC expansion:

$$\mathcal{Y} = \mathcal{M}(\mathbf{X}) = \sum_{\alpha \in \mathbb{N}^d} \pi_{\alpha} \Psi_{\alpha}(\mathbf{X}), \quad (1)$$

Where  $\mathbb{N}^d$  denotes a multi-index  $\alpha = \{\alpha_1, \dots, \alpha_d\}$ ,  $\{\pi_{\alpha}, \alpha \in \mathbb{N}^d\}$  are the expansion coefficients that has to be computed.  $\{\Psi_{\alpha}(\mathbf{X}), \alpha \in \mathbb{N}^d\}$  are multivariate polynomials that are orthonormal with regards to the joint (pdf)  $f_{\mathbf{X}}$  of  $\mathbf{X}$ , i.e.,  $E[\Psi_{\alpha}(\mathbf{X})\Psi_{\beta}(\mathbf{X})] = 1$  if  $(\alpha = \beta)$  and 0 otherwise. The multivariate polynomials  $\Psi_{\alpha}$  are constructed by taking the tensor product of their corresponding univariate polynomials, i.e.,

$$\Psi_{\alpha}(\mathbf{X}) = \prod_{i=1}^d \phi_{\alpha_i}^{(i)}(X_i), \quad (2)$$

where  $\phi_{\alpha_i}^{(i)}$  is a polynomial in the  $i$ -th variable of degree  $\alpha_i$ . gPC expansion in Eq.1 must be truncated after  $p$  terms for computational purposes. Typically, the polynomials  $\Psi_{\alpha}$  with

total degree up to  $p$  are maintained as  $\Psi_\alpha(\mathbf{X}) = \sum_{\alpha \in \mathcal{A}^{d,p} \subset \mathbb{N}^d} \pi_\alpha \Psi_\alpha(\mathbf{X})$ . The set  $\mathcal{A}^{d,p}$  is defined here based on a total polynomial degree.

The next step is the computation of the polynomial chaos coefficients. The regression technique can be used to estimate the spectral coefficients  $\pi_\alpha$ . This could be achieved by solving a least-square (LS) minimization problem in  $\ell_2$ -norm (Berveiller et al., 2006). We first define the metric error  $\epsilon_{PC}$  as the difference between the (exact) model's evaluation ( $\mathcal{Y}$ ) and the gPC surrogate estimates for a finite training set of randomly sampled input variables  $\mathbf{X}$ . A set of  $N_{ts}$  realization of the input vector,  $\mathbf{X} = X(1), \dots, X(N_{ts})$ , is then needed, called experimental design (ED). The input ED is achieved in the present work using the well-known method Latin Hypercube sampling (LHS) approach. The regression problem is solved using a Least-angle-regression (LAR) approach, which offers a sparse gPC representation. This allows us to overcome the curse of dimensionality. Here, we follow the procedure as suggested in Blatman and Sudret (2011). Evaluating its quality is of crucial interest, therefore, Leave-One-Out (LOO) error is employed as it gives an excellent measure of accuracy by enabling error estimates at a reasonable processing. The definition of the relative LOO error is as follows:

$$\epsilon_{LOO} = \frac{\sum_{i=1}^N \left( \frac{\mathcal{M}(x^{(i)}) - \mathcal{M}^{PC}(x^{(i)})}{1 - h_i} \right)^2}{\sum_{i=1}^N (\mathcal{M}(x^{(i)}) - \widehat{\mu_Y})^2} \quad (3)$$

where  $h_i$  is the  $i^{th}$  diagonal term of matrix  $\Psi(\Psi^T\Psi)^{-1}\Psi^T$ , where  $\Psi = \{\Psi_{ij} = \Psi_j(X^i)\}$  and  $(\widehat{\mu_Y}) = \frac{1}{N} \sum_{i=1}^N \mathcal{M}(x^{(i)})$ . Once the PCE is built, the mean  $\mu$  and the total variance  $D$  can be obtained using properties of the orthogonal polynomials, such that  $\mu = y_0$  and  $D = \sum_{\alpha \in \mathcal{A} \setminus 0} y_\alpha^2$ . As mentioned above, the Sobol' indices of any order can be computed in a straightforward manner. The first order and total Sobol' indices are then given by  $S_i = \sum_{\alpha \in \mathcal{A}_i} y_\alpha^2 / D$  and  $\mathcal{A}_i = \{\alpha \in \mathcal{A}: \alpha_i > 0, \alpha_{j \neq i} = 0\}$ , and  $S_i^T = \sum_{\alpha \in \mathcal{A}_i^T} y_\alpha^2 / D$  with  $\mathcal{A}_i^T = \{\alpha \in \mathcal{A}: \alpha_i > 0\}$ , respectively.

### 3. CFD modeling procedure

The system under consideration is a three-dimensional pipeline of diameter  $D = 0.103m$  and of length  $L = 140D$ , thus ensuring a fully developed flow. The transported slurry consists of a mixture of phosphate and water with densities of  $\rho_s = 2600$  and  $\rho_w = 1000$   $kg/m^3$ , respectively. The current CFD modeling approach is based on the Euler-Euler two-fluid formulation, which treats both solid and fluid phases as interpenetrating continua (Gidaspow, 1994). We explore this two-phase slurry flow in pipes numerically using the open-source continuum mechanics suite OpenVFOAM®, which is a C++ toolkit for developing numerical solutions. The CFD simulations are run using the solver TwoPhaseEulerFoam, which has been adapted to accommodate mixtures of solids and liquids (Elkarii et al., 2020). The reader is directed to (Elkarii et al., 2022) for further information on the current CFD modeling technique as well as to review the numerical validation. In the present work, the head loss parameter will be estimated using the CFD simulations. This will serve as the quantity of interest for the GSA, since it constitutes a good proxy for the energy requirements of the pipe flow.

#### 4. UQ methodology

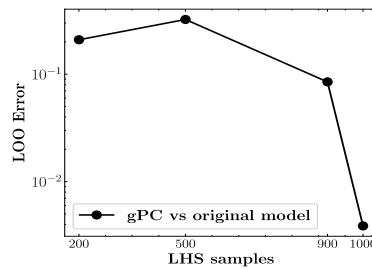
To perform UQ and GSA for the problem of slurry flows in pipelines, one should identify the various sources of uncertainty, which will be considered as model inputs. As stated in the introduction, the slurry pipe flows behavior is often defined through the properties of the materials (El Moçayd and Seaid (2021)). This motivates the choice of  $\mathbf{X} = \{\phi_s, u_m, d_p, \theta, SC\}$  as random parameters. These random variables are presented according to their respective pdf's (cf. table1). In order to assess the quality of the gPC expansion, four design of experiments of sizes  $N = \{200, 500, 900, 1000\}$  are considered. They are generated using the LHS method. Once the design of experiments are defined it remains to run the computational model for each sampling node in those sets to obtain the vector  $\mathbf{Y}$  of observations of the quantity of interest (QOI).

Table 1: Phosphate slurry flow: parameters of the model

Parameters	Type of PDF	Orthogonal polynomials	Range of variability
Inlet velocity $u_m$ [m/s]	Uniform	Legendre	[1.75, 5]
Solid concentration $\phi_s$ [-]	Uniform	Legendre	[20%, 40%]
Particles size $d_p$ [ $\mu\text{m}$ ]	Uniform	Legendre	[44, 250]
Pipe inclination $\theta^\circ$	Gaussian	Hermite	[0, 4]
Specularity coefficient SC[-]	Uniform	Legendre	[0, 1]

#### 5. Results and discussion

To ensure results convergence, the LOO error (Eq. 3) is calculated for each design of experiment and portrayed in figure 1. The LOO error convergence rate indicates that 1000 samples are sufficient to build a high quality gPC expansion.



A further investigation of gPC accuracy is performed. First, the sensitivity of the gPC with respect to the polynomial degree is evaluated as depicted in Fig. 2a. A polynomial degree  $p = 6$  is sufficient to carry on UQ and GSA using the surrogate model. Then, the ability of the surrogate model to mimic the physical behavior of the CFD model is presented in Fig. 2b. This constitutes a good validation for the gPC ability to serve as a low-cost meta-model.

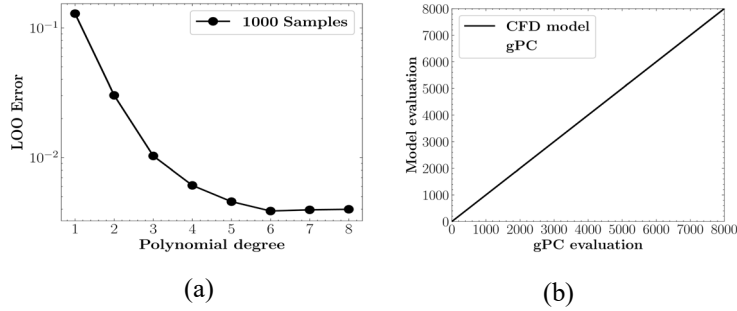
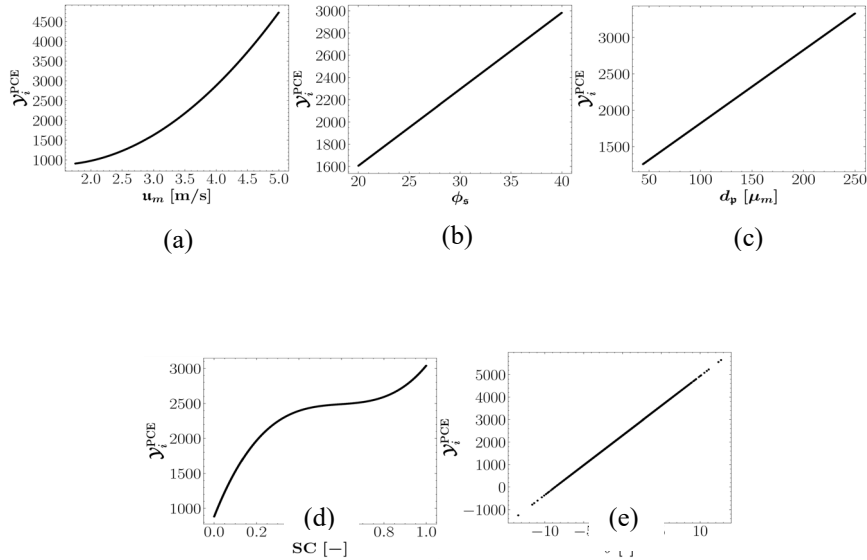


Figure 2: (a) Optimal gPC degree with respect to gPC degree and (b) Comparison between gPC and the exact model using 1000 samples.

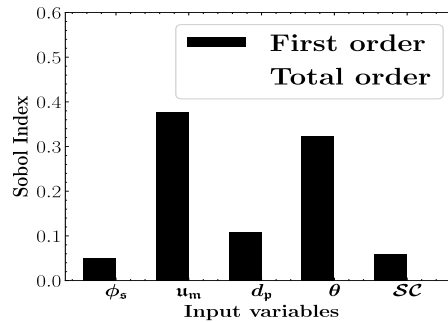
Next, we examine the marginal effect of the uncertain parameters on the model response. It is observed that there is a good fit, which is highlighted by a minor LOO error (less than 0.001) (Fajraoui et al., 2017), and which corresponds to an optimal PCE of degree  $P = 6$ . We carry out this analysis on the total head loss, in order to better understand how energy loss in pipelines is impacted, since the total head loss is a measure of the energy reduction produced by the sum of elevation head, velocity head and frictions head of the fluid as it passes through a fluid system. The sensitivity of the head loss to the variability of the random parameters  $\{\phi_s, u_m, d_p, \theta, \mathcal{SC}\}$ , can be assessed by means of the Sobol’ indices. Figure 3 shows bar plots of the computed first and total Sobol’ indices.



around  $u_m \leq 1.75$ . This occurs as a result of approaching the limit deposition velocity, at which point a stationary regime emerges, culminating in total particle sedimentation. As for  $\theta$  (cf Fig. 3e),  $\Delta H_{Total}$  increases with positive slopes and decreases on the other hand with the negative ones. At positive slope, the longitudinal resultant force acts as an extra



resistivity force for the wall contact layer and tends to reduce its velocity compared to horizontal flow. However, as a suspension goes down a pipe, flow is increased due to elevation effects, and pressure levels rise downstream. This extra kinetic energy is utilized to compensate for frictional consequences. With respect to the solids concentration and the particle sizes,  $\Delta H_{Total}$  grows as both of these variables rise, as noticed in Figs. 3b and 3c. Finally, when it comes to the specular coefficient  $SC$ , it is remarked that the pressure drop increases with the latter (cf Fig. 3d). We recall here that  $SC$  quantifies the momentum and energy transfer caused by colliding particles with the pipe wall. As a result, a larger specular coefficient implies that it is fair to expect a considerable loss of kinetic energy once particles collide with the pipe wall, resulting in an increase in pressure drop. Finally, we evaluate the sensitivity of the head loss uncertainty to the variability of the random parameters  $\{\phi_s, u_m, d_p, \theta, SC\}$ . This will be assessed using the Sobol' indices resulting from the gPC expansion. Figure 4 shows bar plots of the computed first and total Sobol' indices. Inspection of the sensitivity indices showed that the variability of the total head loss  $\Delta H_{Total}$  is mainly due to the principal effects of  $u_m$  followed by  $\theta$  and then  $d_p$ . A small but not negligible influence of  $\phi_s$  and  $SC$ , is also observed. There is a notable difference between first and total order sobol' indices, indicating interactions among the random parameters.



## 6. Conclusion

In this study, a sensitivity analysis of turbulent phosphate slurry flow in the pipeline has been performed. The effects of our input parameter on the pressure drop along the pipe have been investigated. The choice of the parameters is performed with the aim to improve the physical understanding of the considered flow configuration. Results showed that the variability of the pressure drop is largely influenced by the effect of  $u_m$ ,  $\theta$ , and  $d_p$ . It can be also deduced, that a good calibration of  $SC$  in the numerical validation step is crucial, to ensure reliable predictions of the pressure drop along the pipe.

## References

- M. Berveiller, B. Sudret, M. Lemaire, 2006. Stochastic finite element: a non intrusive approach by regression. European Journal of Computational Mechanics/Revue Européenne de Mécanique Numérique 15 (1-3), 81–92.
- G. Blatman, B. Sudret, 2011. Adaptive sparse polynomial chaos expansion based on least angle regression. Journal of computational Physics 230 (6), 2345–2367.
- N. El Moc ayd, M. Seaid, 2021. Data-driven polynomial chaos expansions for characterization of complex fluid rheology: Case study of phosphate slurry. Reliability Engineering & System Safety 216, 107923.

- M. Elkarii, C. Bouallou, A. Ratnani, 2020. Towards modelling a diphasic flow using the CFD technique to achieve a digital twin of a phosphate slurry piping process. *Chemical Engineering Transactions*.
- M. Elkarii, R. Boukharfane, S. Benjelloun, C. Bouallou, 2022. A cfd-based surrogate model for predicting slurry pipe flow pressure drops. *Particulate Science and Technology*, 1–11.
- D. Gidaspow, 1994. *Multiphase flow and fluidization: continuum and kinetic theory descriptions*. Academic press.
- S. Kucherenko, et al., 2005. Global sensitivity indices for nonlinear mathematical models, review. *Wilmott Mag* 1, 56–61.
- A. Saltelli, M. Ratto, S. Tarantola, F. Campolongo, et al., 2006. Sensitivity analysis practices: Strategies for model-based. *Reliability Engineering & System Safety* 91 (10-11), 1109–1125.
- M. Sobol, 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation* 55 (1-3), 271–280.
- B. Sudret, 2008. Global sensitivity analysis using polynomial chaos expansions. *Reliability engineering & system safety* 93 (7), 964–979.
- R. W. Walters, L. Huyse, 2002. Uncertainty analysis for fluid mechanics with applications.
- D. Xiu, G. E. Karniadakis, 2002. The wiener–askey polynomial chaos for stochastic differential equations. *SIAM journal on scientific computing* 24 (2), 619–644.
- H. Zhong, X. Lan, J. Gao, Y. Zheng, Z. Zhang, 2015. The difference between specularly coefficient of 1 and no-slip solid phase wall boundary conditions in cfd simulation of gas–solid fluidized beds. *Powder Technology* 286, 740–743.



# Hybrid dynamic model of monoclonal antibody production using CHO cells

Mariana Monteiro,<sup>a</sup> Cleo Kontoravdi<sup>a</sup>

*<sup>a</sup>Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, Exhibition Road, London SW7 2AZ, United Kingdom*

## Abstract

Mammalian cells are used to produce up to 80% of the commercially available therapeutic proteins, with Chinese Hamster Ovary (CHO) cells being the main production host. Antibody production occurs in a train of bioreactors typically operated in fed-batch, i.e., semi-continuous, mode. Real-time monitoring of antibody concentration is costly and involves sampling and offline analytics. Current reactor models lack extrapolation power while being highly reliant on extensive amounts of data for parameterization. This work aims to leverage knowledge of the cell's metabolism and incorporate it in the production reactor model, thus creating a hybrid formulation with increased predictive capability and generality. This novel formulation can make predictions without requiring kinetic parameter estimations. The applicability of the proposed formulation is demonstrated with a soft sensor for antibody production.

**Keywords:** Biomanufacturing, Hybrid Modeling, Fed-Batch, CHO cells, monoclonal antibodies

## 1. Introduction

The global market of biologics reached an annual value of US\$265 billion in 2020, with an expected increase to US\$856 billion by 2030 (Precendence Research, 2021). Mammalian cells are used to produce up to 80% of the commercially available therapeutic proteins, with Chinese Hamster Ovary (CHO) cells representing the major fraction of that (Fan, *et al.*, 2015), because they produce safe and efficacious protein products. As a result of high demand, there is pressure to optimize the manufacturing process even further. Current upstream optimization techniques focus on medium composition, feeding strategies and process conditions (e.g., temperature and pH) (Bassem, *et al.*, 2021). Given the number of possible operating parameters to manipulate, it is both time-consuming and costly to perform experiments covering the entire design space. However, this burden can be lessened using a combination of predictive process models and targeted experimentation.

Current mathematical models describing the production reactor assume saturation kinetics and typically treat the cell as a black box. Although this formulation tends to fit cell culture data well, it offers little extrapolating power given that most parameters heavily rely on the process conditions for which the model was parameterized, such as temperature and pH. In this work, we propose a hybrid modeling approach comprising a stoichiometric model of cell metabolism and a dynamic reactor model, in which the rate

of change of each metabolite concentration is calculated using the metabolic model instead of treating the cell as a black box.

Metabolic models depict the network of chemical reactions occurring inside the cell and describe the consumption and production of metabolites. CHO cell metabolic network models can range from a few tens of reactions (Nolan and Lee, 2011) to over 6000 reactions and 4000 metabolites (Hefzi, *et al.*, 2016). Even when using a relatively small-scale network, it is not possible to experimentally obtain reaction rates of all reactions. As such, given that the system is underdetermined, it is necessary to solve the system of equations as an optimization problem and constrain the solutions to feasible and biological relevant spaces. The typical way to solve these problems is to define a feasible objective function that describes the cell's objective; examples include the maximization of growth rate or ATP production, minimization of ATP consumption, or minimization of NADH production (Carvalho *et al.*, 2019; Höffner *et al.*, 2019).

In this work we take advantage of our knowledge of cell metabolism to define a metabolic network model and we incorporate it in a dynamic model of the fed-batch production reactor. Since the core reactions in a cell's metabolic network always hold true, our novel formulation offers increased generality. We apply this approach to describe an antibody-producing CHO cell culture and demonstrate its applicability as a soft-sensor of antibody concentration (one of the costliest sensors).

## 2. Methodology

### 2.1. Hybrid Model

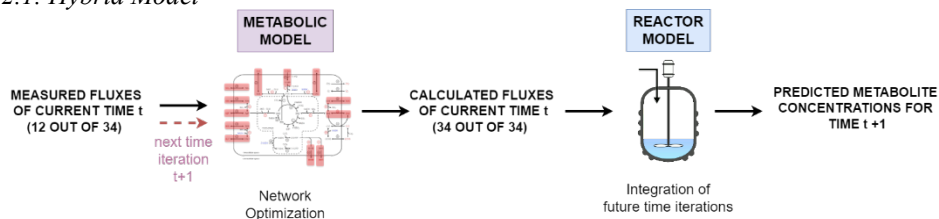


Figure 1 Proposed hybrid model

Figure 1 depicts the proposed hybrid modeling algorithm. The dynamic simulation loops over the following steps:

*Step 1.* We calculate the measured fluxes by subtracting the current (time  $t$ ) and past (time  $t-1$ ) concentrations of each metabolite. As shown in figure 1, we have available a subset of the total number of fluxes required (12 out of the 34 fluxes), none of which are intracellular.

*Step 2.* We perform the metabolic network optimization using the measured fluxes (graphically represented as the pink boxes in the metabolic model). As a result, we calculate the fluxes of the entire network for the current time  $t$ .

*Step 3.* We perform an integration of future time steps, in the reactor model, to calculate the predicted metabolite concentrations for time  $t+1$ , using the 34 calculated fluxes.

*Step 4.* We repeat each step iteratively, for each daily measurement available. Python library Scipy (Virtanen, *et al.*, 2020) was used for both optimization and simulation of the proposed hybrid model.

### 2.2. Metabolic model

The implemented optimization routine is shown below (equation 1), and it is formulated as a Metabolic Flux Analysis (MFA) problem, i.e., a minimization of the squared difference between measured and calculated fluxes:

$$\min_{v_k^{calc}} \sum \left( \frac{v_k^{calc} - v_k^{meas}}{v_k^{meas}} \right)^2 \quad \forall k \in \text{measured fluxes} \quad (1)$$

$$S \cdot v = 0$$

$$0 < v_{irrev}$$

where  $v_k^{calc}$  are the calculated fluxes,  $v_k^{meas}$  are the measured fluxes,  $S$  is the stoichiometric matrix,  $v$  are the fluxes of the entire networks,  $v_{irrev}$  are the irreversible fluxes and  $k$  represents the number of the flux. The measured fluxes,  $v_k^{meas}$  are the inputs for the optimization and the calculated fluxes,  $v_k^{calc}$ , are the optimization decision variables. The stoichiometric matrix is a  $m \times n$  matrix, where  $m$  is the total number of metabolites and  $n$  is the total number of reactions of the network. As such, there is a row per metabolite and a column per reaction. The stoichiometric matrix used contains 24 metabolites and 34 reactions. The network and reactions were taken from previous published work (Nolan and Lee, 2011).

This optimization routine is subject to two constraints. The first one enforces a *quasi-steady* state for each metabolite, meaning that it is assumed that the total rate of production equals the total rate of consumption for each metabolite, during each short time iteration. The second constraint provides information of which fluxes are irreversible and bounds them to biologically meaningful values.

### 2.3. Fed-batch model

The fed-batch model is implemented as a system of differential and algebraic equations, describing the mass balances for each metabolite, of the following form:

$$\frac{d\text{Metabolite}}{dt} = ([\text{Metabolite}]_{t=0} - [\text{Metabolite}]_{t=t}) * \frac{F(t)}{V} - \text{flux}_{\text{Metabolite}}(t) \quad (2)$$

$$* X_v(t)$$

where  $X_v$  is the cell volumetric density,  $F(t)$  is the inlet flow to the fed-batch and  $V$  is the working volume. Given that there is no outlet flow of the reactor, the working volume is solely a function of the inlet:

$$\frac{dV}{dt} = F_{in}(t) \quad (3)$$

### 2.4. Soft Sensor Validation

The applicability of the hybrid model formulation to soft sensing is demonstrated using published data from a fed-batch experiment (Kyriakopoulos and Kontoravdi, 2014). The cell line used is CHO GS46 and the expressed protein is a chimeric IgG4 monoclonal antibody. The cells were grown in a working volume of 250 mL for 14 days and fed on alternate days. Available daily measurements include viable cell density, viability and

extracellular glucose, lactate, ammonia, and amino acid concentrations. The soft sensor capability is demonstrated by rerunning the implemented algorithm with no measurement data of antibody concentration except for the initial time point.

### 3. Results

In this paper, we describe the formulation of a hybrid model comprising a metabolic network and a reactor model, which can use early measurement data to predict key metabolite concentrations throughout the culture. The model is then used as a soft sensor for antibody concentration. Figure 2 compares the hybrid model's predictions for key extracellular metabolite concentrations to experimental data. In addition, the graphs show the soft sensor results, for which no antibody measurements were used to solve the metabolic model except for the initial time point.

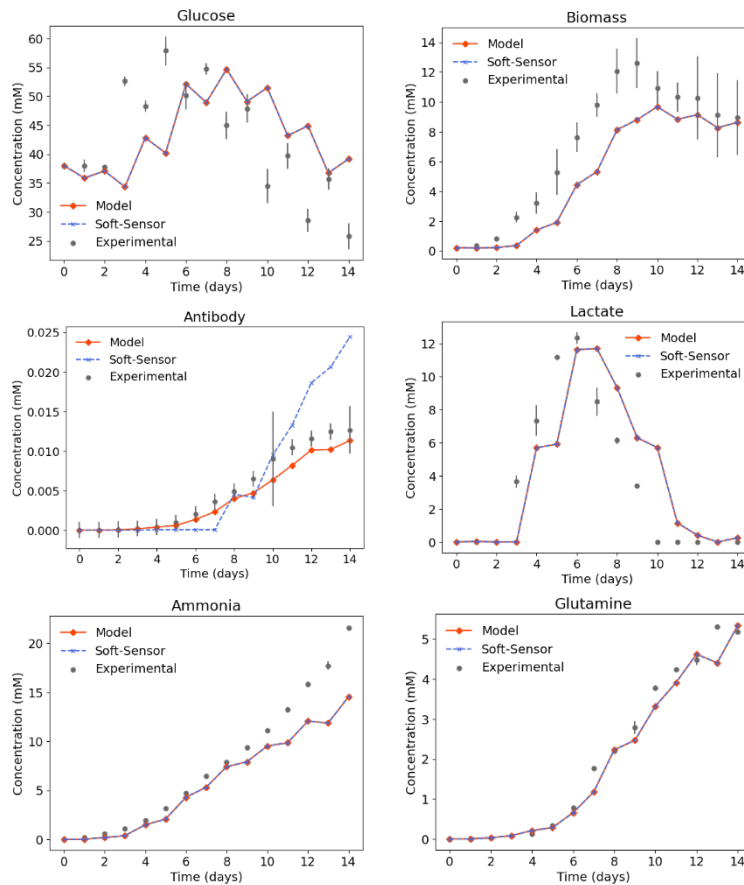


Figure 2 Model Validation Results: Measured vs Predicted Concentrations (in mM) of glucose biomass, antibody, lactate, ammonia, and glutamine, with (Hybrid) and without (Soft-Sensor) antibody, as a function of fed-batch time (in days)

Overall, the proposed hybrid model shows good agreement for the exponential growth phase (initial stages of the batch). Furthermore, the model can capture the lactate shift from consumption to production. The model shows deviations in glucose consumption in the later stages of the culture, which have a cascading effect on the other metabolites. Given that the model is predicting a lower glucose consumption, the cells do not produce as much energy, so they cannot grow as much or produce as much antibody product. The main reason for this underprediction of glucose lies in the *quasi*-steady state assumption, which is one of the constraints of metabolic network optimization. This assumption is valid for the early stages of the culture, when there is a high-nutrient uptake, resulting in fast cell growth. However, it proves insufficient to explain cellular behavior in the later stages, when cells stop growing at a high rate and significant metabolic changes occur, which are dependent on process conditions. As such, the assumption of *quasi*-steady state for the duration of one day (in the metabolic model) leads to higher deviations in the model predictions (Coulet, *et al.*, 2022). Additionally, metabolic network models are known to underpredict glucose consumption as there are energy requirements not captured in the stoichiometry of the reactions of the model, resulting in an underprediction of energy demands. As an improvement, we hope to remove the glucose steady-state constraint in the later stages of the batch and replace it with a minimum glucose uptake rate, backed by experimental data of multiple batches.

The soft sensor is able to estimate monoclonal antibody concentration in the first half of the batch, overpredicting it in the later stages of the batch. The overprediction is linked to an underestimation of the energy requirements of the cell. Given the simplicity of the metabolic model, the entirety of glucose consumed is directed mostly to cell growth and antibody production. There are, however, other metabolic functions that also take up energy, that are not accounted for in the model used herein. This shortcoming could be overcome by using a more comprehensive metabolic network.

The soft-sensor formulation shows the highest deviation from experimental values. When we compare the deviation of the soft sensor from the data and that from the hybrid model, the difference is not as pronounced, reflecting how much closer the hybrid model predictions are to the data. The soft sensor does not show any significant changes when compared to the original hybrid model formulation, for the remaining amino acids. This seems to indicate that removing the measured flux of antibody does not increase the deviation error of the predictions for the remaining amino acids. A possible reason for the low impact of antibody is its low flux value when compared to the remaining amino acids in the network optimization.

The novelty of the proposed model is the estimation of antibody from metabolite concentrations readily available online without the requirement of kinetic parameters estimation. We envisage this formulation as a technological tool towards Industry 4.0, as it paves the way for advanced process control without the need for online monitoring of hard-to-measure data. Our formulation can be integrated into a real-time optimization framework and act as feedback for a control system. Defining control objectives such as maximizing yield or minimizing the batch duration can be possible with the proposed framework, as the soft-sensor would provide the current state of the product.



#### 4. Concluding Remarks

In this paper, we propose a hybrid formulation which integrates a metabolic network in the reactor model that avoids parametrization by using knowledge of cell metabolism to predict key metabolite concentrations. Initial results show that our hybrid formulation offers good predictions for the early stages of the batch, where cells are showing a high nutrient consumption and growing at a fast pace. We demonstrate the applicability of the proposed formulation as a soft sensor for antibody concentration.

#### Acknowledgements

The authors gratefully acknowledge funding from United Kingdom's Biotechnology and Biological Sciences Research Council (BBSRC) and GlaxoSmithKline (GSK).

#### References

- M. Carvalho, A. Nikdel, J. Riesberg, D. Lyons, H. Budman, 2019, Identification of a dynamic metabolic flux model for a mammalian cell culture, 12th IFAC Symposium on Dynamics and Control of Process Systems
- M. Coulet, O. Kepp, G. Kroemer, S. Basmaciogullari, 2022, Metabolic Profiling of CHO Cells during the Production of Biotherapeutics, *Cells*, 11. 10.3390/cells11121929.
- Y. Fan, I. Jimenez Del Val, C. Müller, J. Wagtberg Sen, S. K. Rasmussen, C. Kontoravdi, D. Weilguny, M.R. Andersen, 2015, Amino acid and glucose metabolism in fed-batch CHO cell culture affects antibody production and glycosylation. *Biotechnol. Bioeng.*, 112: 521-535.
- H. Hefzi, K.S. Ang, M. Hanscho, *et al.*, 2016, A consensus genome-scale reconstruction of chinese hamster ovary cell metabolism, *Cell Systems*, 3(5):434–443
- K. Höffner, S. M. Harwood, P. I Barton, 2019, A reliable simulator for dynamic flux balance analysis, *Biotechnology and bioengineering*, 110(3):792–802, 2013.
- Y. M. Huang, W. Hu, E. Rustandi, K. Chang, H. Yusuf-Makagiansar, T. Ryll, 2010, Maximizing productivity of CHO cell-based fed-batch culture using chemically defined media conditions and typical manufacturing equipment, *Biotechnology progress*, 26(5), 1400–1410
- S. Kyriakopoulos, C. Kontoravdi, 2014, A framework for the systematic design of fed-batch strategies in mammalian cell culture, *Biotechnology and bioengineering*, 111(12), 2466–2476
- R. P. Nolan, K. Lee, 2011, Dynamic model of CHO cell metabolism, *Metabolic engineering*, 13(1), 108–124
- Precedence Research. Biopharmaceuticals Market - Global Industry Analysis, Size, Share, Growth, Trends, Regional Outlook, and Forecast 2021 – 2030, 2021, <https://www.precedenceresearch.com/biopharmaceutical-market> accessed on 30/11/2022
- P. Virtanen, R. Gommers, T. E. Oliphant, *et al.*, 2020, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17(3), 261-272.
- B. B. Yahia, L. Malphettes, E. Heinzle, 2021, Predictive macroscopic modeling of cell growth, metabolism and monoclonal antibody production: Case study of a CHO fed-batch production, *Metabolic Engineering*, 66, 204-216

# Combined Bayesian optimization and global sensitivity analysis for the optimization of simulation-based pharmaceutical processes

Niki Triantafyllou<sup>a</sup>, Nilay Shah<sup>a</sup>, Maria M. Papathanasiou<sup>a</sup>, Cleo Kontoravdi<sup>a,\*</sup>

<sup>a</sup> *The Sargent Centre for Process Systems Engineering, Imperial College London, London, United Kingdom, SW72AZ*

\**cleo.kontoravdi98@imperial.ac.uk*

## Abstract

We propose an efficient framework that employs Bayesian optimization and global sensitivity analysis for the optimization of detailed pharmaceutical flowsheets. Global sensitivity analysis based on quasi-random sampling is utilized to reduce the dimensionality of the problem by identifying critical process and economic parameters that contribute significantly to the variability of Key Performance Indicators (KPIs) such as batch size and OpEx. Then, Bayesian optimization is performed in the previously identified critical input space based on gaussian process surrogate models and a number of different acquisition functions to find the optimal critical operating conditions that minimize the aforementioned KPIs. We apply this framework to the manufacture of plasmid DNA (pDNA), which is a critical raw material for advanced therapeutics, leading to a surge in demand for pDNA for clinical or commercial use. Optimized manufacturing recipes identified with the proposed framework are projected to achieve an up to 170% increase in the batch size and a 34.7% decrease in the OpEx per batch.

**Keywords:** flowsheet optimization, Bayesian optimization, global sensitivity analysis

## 1. Introduction

Plasmid DNA (pDNA) is a carrier of genetic information, which makes it a critical raw material in the manufacturing and delivery of Advanced Therapy Medicinal Products (ATMPs). Some of the main pharmaceuticals that use pDNA include DNA and mRNA vaccines and cell and gene therapies (Gosse *et al.*, 2022). The fast-growing interest in ATMPs has led to a growing demand for the production of pDNA for clinical and commercial use. Scaling up the pDNA manufacturing process to meet current and future demands is one of the major challenges in the cell and gene therapy market (Ohlson, 2020). Simultaneously, pDNA production costs and manufacturing timeframes need to be minimized in an attempt to decrease the cost of goods (COGs) and maximize the production of these innovative products. The pDNA manufacturing process is described by long batch times, high capital, operating and labour expenditures. Nevertheless, pDNA manufacturing is considered a well-known process and limited process development and optimization are observed compared to other similar products such as viral vectors and monoclonal antibodies.

Manufacturers need to develop strategies for the establishment of scalable and cost-effective manufacturing processes that will accelerate the wider adoption of ATMPs. Process systems engineering tools can be employed for the development of (bio)pharmaceutical processes, which mostly rely on batch processes that may not be

operated at optimal conditions (Boukouvala & Ierapetritou, 2013). Techno-economic analysis using flowsheet modelling is an efficient way to verify the feasibility of a process and compare different technologies. In addition, flowsheet optimization can significantly improve process efficiency by identifying the optimal operation conditions that minimise/maximise key performance indicators (Wang *et al.*, 2017), while data-driven techniques can be employed for simulation-based flowsheet optimization (van de Berg *et al.*, 2022).

In this study, we present a framework for (bio)pharmaceutical process flowsheet optimization that utilizes Bayesian Optimization (BO) and variance-based Global Sensitivity Analysis (GSA). GSA is used as means for uncertainty quantification in the flowsheet. This information is utilized to rank variables and decrease the dimensionality of the black-box model that is then optimized with BO. BO is a sample-efficient optimization method that utilizes a probabilistic model to represent the uncertainty in the objective function. The proposed framework is used for the identification of optimal operating conditions in the production of pDNA.

## 2. Materials & Methods

The proposed framework combines techno-economic assessment (TEA), global sensitivity analysis and Bayesian optimization. The underlying algorithm is displayed below:

---

### Algorithm 1 Combined TEA, GSA, and BO.

---

**Start:** Design process and gather process and economic data.

**Input:** Converged flowsheet, an initial set of  $N$  uncertain parameters  $\mathbf{P} = \{p_1, p_2, \dots, p_N\}$  and their acceptable ranges  $[\mathbf{P}_{min}, \mathbf{P}_{max}]$ .

**while**  $threshold \leq S_{T,min}$  **do**

1. *Screening:* Select subset  $n \subseteq N$  of uncertain parameters  $\mathbf{P} = \{p_1, p_2, \dots, p_n\}$ .
2. *Sampling:* Draw 4,096 input samples  $P$  based on Sobol sequence and run the converged flowsheet to obtain the target outputs  $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$ .
3. *GSA:* Compute the total sensitivity indices  $S_T$  based on the RS-HDMR metamodel.

**end while**

**BO initialization:** train  $\mathcal{GP}^0$  with an initial data set  $D_0 = (\mathbf{P}_0, y_0)$

**for**  $i = \text{NumberOfIterations}$  **do**

1. Find  $\mathbf{P}_i$  by optimizing the acquisition function over the  $\mathcal{GP}^i$ , where  $\mathbf{P}_i = \text{argmin}_p u(\mathbf{P}|D_{1:i-1})$ .
2. Run the flowsheet to sample the objective function  $y_i = f(\mathbf{P}_i) + \varepsilon_i$ .
3. Augment the data set  $D_{1:i} = \{D_{1:i-1}, (\mathbf{P}_i, y_i)\}$  and update the  $\mathcal{GP}^i$ .

**end for**

---

### 2.1. Techno-economic Modeling & Global Sensitivity Analysis

A fed-batch production process for pDNA is modelled using SuperPro Designer (Intelligen, Inc) (Ferreira & Petrides, 2021). SuperPro Designer uses built-in process and economic models based on sets of algebraic and differential equations to compute equipment sizes, operation scheduling, CapEx and OpEx. The pharmaceutical pDNA production flowsheet includes cell cultivation, pDNA recovery, purification, formulation and fill & finish as shown in Fig. 1. Information on process and economic parameters and

their proposed accepted ranges are based on literature, cGMP manufacturers, suppliers and expert feedback for the design of the specific process at hand.

Global sensitivity analysis is then performed to quantify manufacturing uncertainty based on the contribution of uncertain input parameters to the output variance of specific KPIs, namely batch size, formulation concentration, throughput, batch time, cycle time, CapEx, OpEx, and production cost. GSA was performed using a SobolGSA (Kucherenko, 2013)-Matlab-Component Object Model (COM)-SuperPro interface with 4,096 quasi-randomly generated samples of the uncertain input variables based on Sobol quasi-random sequence within the pre-defined ranges shown in Table 1. Critical parameters have, for at least one of the KPIs, total sensitivity indices exceeding a user-specified threshold ( $S_{T,min} = 1\%$ ). A higher  $S_{T,min}$  would lead to a lower-dimensional design space but for the purpose of this work, the capabilities of the algorithm are tested for a high-dimensional space. Non-influential parameters are fixed after each global sensitivity run until the algorithm converges and all the critical parameters are identified. The use of GSA enables the ranking of uncertain parameters and therefore dimensionality reduction of the design space by only considering a subset of the most influential inputs.

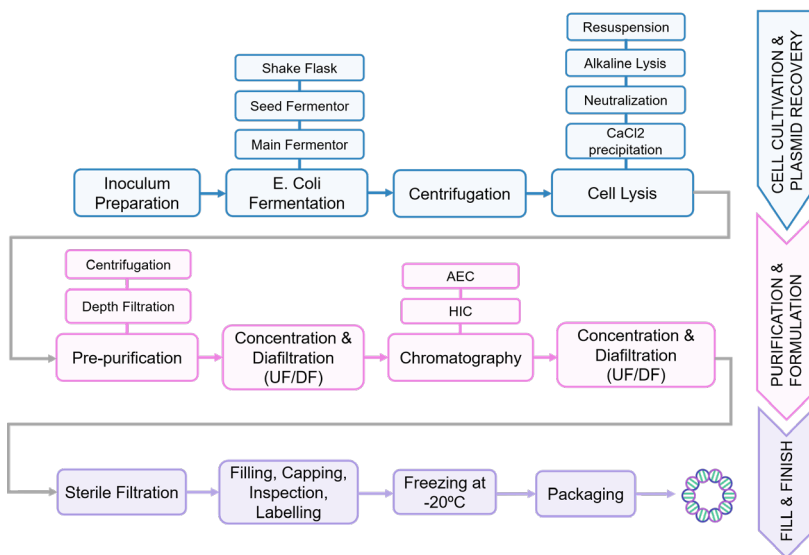


Figure 1. Process flowsheet for the production of plasmid DNA.

## 2.2. Bayesian Optimization (BO)

Once the critical design space has been identified, BO is used to derive the optimal manufacturing recipes. The pDNA flowsheet is treated as a black-box function ( $f(\mathbf{P})$ ) that can be sampled at the critical inputs ( $\mathbf{P} \in \mathbb{R}^n$ ) to obtain the optimal solution ( $\mathbf{P}^*$ ). Gaussian processes (GPs) with zero mean and the Matérn kernel function are used as the surrogate model. GPs are chosen because they can reliably quantify uncertainty and point to unexplored regions of the search space and due to their ability to incorporate prior knowledge into the model (Rasmussen & Williams, 2006). Initially, the GP surrogate is trained with 5 samples based on Sobol sequence and then BO iteratively updates the GP based on adaptive sampling. Adaptive sampling is achieved through the optimization of an acquisition function that can balance exploration versus exploitation. In this work, we test four of the most well-known acquisition functions—probability of improvement (PI),

expected improvement (EI), upper (or lower) confidence bound (UCB or LCB), and max-value entropy search (MES). UCB and LCB are tested for two different confidence values that control the level of exploration. Before the GP is trained, data features are normalized, and data outputs are standardized. The BO part of the algorithm is built using the Python packages PyTorch and BoTorch (Balandat *et al.*, 2020).

### 3. Results & Discussion

The screening for critical parameters considers an initial set of more than 25 parameters including task lengths, conversions, buffer volumes, working volumes, rejection and concentration coefficients, failure rates, etc. The parameters that had an insignificant effect on the output metrics were fixed to their nominal values without adding extra uncertainty to the model. The GSA part of the algorithm was repeated until all remaining parameters were critical. The final subset of critical parameters that are responsible for at least 1% of the output variability for at least one output and their corresponding acceptable ranges are displayed in Table 1.

Table 1. Critical parameters in pharmaceutical pDNA production, their acceptable ranges, nominal values, and optimal values for the maximization of the batch size and the minimization of OpEx.

Input	Process step	Range	Nominal	Optimal <sub>batch-size</sub>	Optimal <sub>OpEx</sub>	Unit
Fermentation duration (seed)	Seed Fermentor	12–30	30	23.25	12	<i>h</i>
Fermentation duration (main)	Main Fermentor	18–40	40	20.75	18	<i>h</i>
Fed-batch medium volume		1.5–4.5	3	2.63	1.5	<i>L</i>
Starting working volume		8.1–24.2	16.1	14.1	8.1	<i>L</i>
Resuspension buffer volume	Resuspension	22.78–68.36	45.57	28.48	22.79	<i>L</i>
Lysis buffer volume	Lysis	22.78–68.36	45.57	28.48	22.79	<i>L</i>
Neutralization buffer volume	Neutralization	22.78–68.36	45.57	28.48	22.79	<i>L</i>
Diafiltration volumes 1	1 <sup>st</sup> UF/DF	10–50	50	42	10	–
Flush volume per filter area 1		1.5–4.5	3	4.12	1.5	<i>L/m<sup>2</sup></i>
Diafiltration volumes 2	2 <sup>nd</sup> UF/DF	10–50	40	43	10	–
Flush volume per filter area 2		0.67–2	1.33	1.17	0.67	<i>L/m<sup>2</sup></i>

The BO algorithm is performed for all critical parameters in a 9-dimensional space. However, the critical parameters are ranked and contribute differently to each output metric. Fig. 2 illustrates the variance as defined by the GSA algorithm by summarizing the first ( $S_i$ ) and second ( $S_{ij}$ ) order sensitivity indices for two of the output metrics. Fig. 3 illustrates the total variance of the two output metrics when all critical inputs are simultaneously varied within their acceptable ranges.

The BO algorithm leads to (a) an increase of 170% in the batch size and (b) a decrease of 34.7% in the OpEx per batch from the nominal points while operating within the acceptable ranges for all critical inputs (Fig. 3). Optimal manufacturing recipes are shown

in Table 1. The nominal point of the OpEx lies far up from the median because the base case flowsheet operates on the upper bounds of the seed and main fermentor task durations.

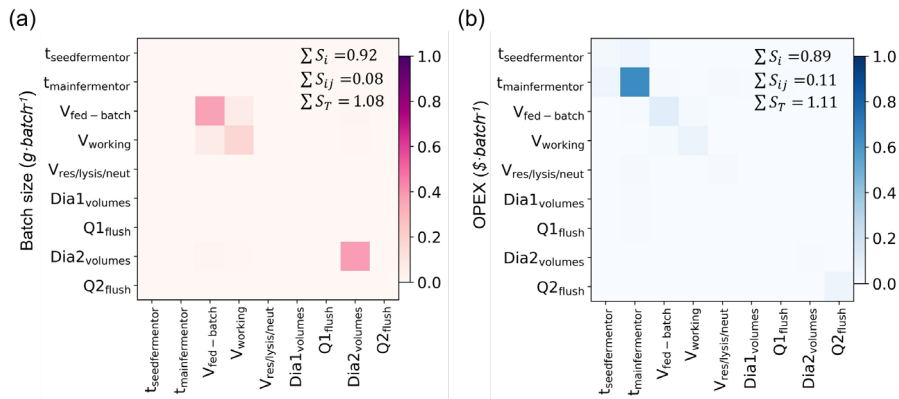


Figure 2. GSA of the nine critical input parameters on the batch size and OpEx for the production of pharmaceutical pDNA. First-order effects ( $S_i$ ) are displayed in the diagonal and second-order effects ( $S_{ij}$ ) are presented in the upper and lower triangular.

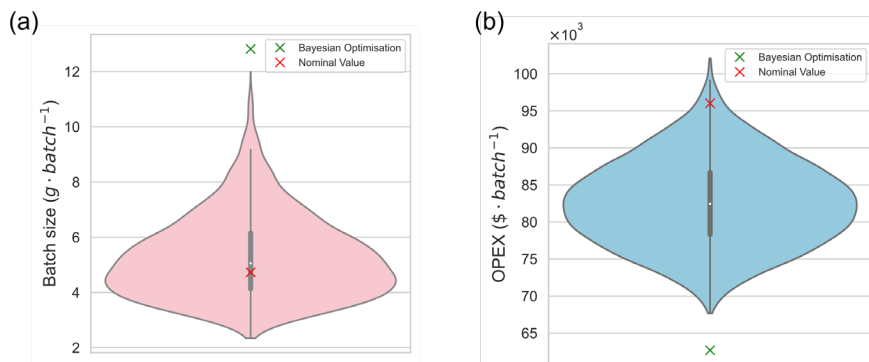


Figure 3. Violin plots for two of the key output metrics. The black bar in the center is the interquartile range, the white dot inside the black bar is the median value, and the black lines stretched from the bar are the lower/upper adjacent values. Bayesian optimization is used to (a) maximize the batch size and (b) minimize the OpEx.

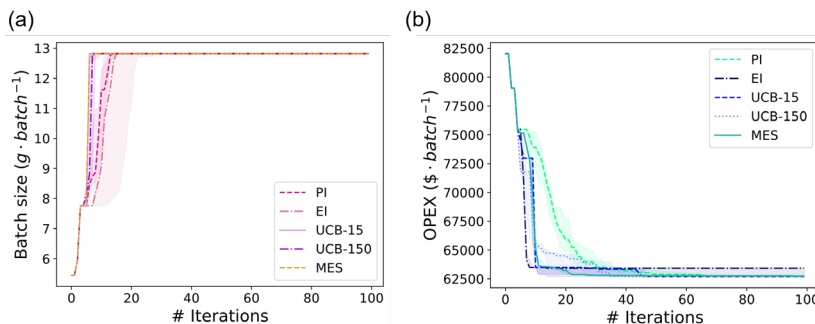


Figure 4. Bayesian optimization progress for different acquisition functions for (a) batch size maximization and (b) OpEx minimization. Plot presents the median and confidence intervals ( $1^{st}$  and  $3^{rd}$  quartiles) from 10 random seeds.

Fig. 4 shows the progress of the BO algorithm for the different acquisition functions for 100 iterations. Most acquisition functions converge to the optimum after a few iterations. In the batch size maximization problem, the best-performing acquisition function is UCB, for both confidence values, resulting in convergence to the optimum in less than 10 iterations. On the contrary, EI performs worse compared to the others, converging after 17 iterations. In the OpEx minimization, MES performs the best, reaching the optimum after 25 iterations, with UCB-150 following at 20 iterations. It can be observed that the algorithm requires more iterations to find the optimum for OpEx because this metric has a wider range of critical variables contributing to the overall variability (Fig. 2).

#### 4. Conclusions

In this work, we present a framework for black-box optimization of simulation-based pharmaceutical processes that relies on the principles of global sensitivity analysis and Bayesian optimization. The case study shows that the proposed algorithm performs reliably well and leads to optimum solutions after a few iterations. We use four different acquisition functions that use different strategies for exploration and exploitation and the algorithm converges in all cases. The right acquisition function for each output metric is key for optimum algorithm performance, with cases of acquisition functions performing differently for different objective functions being observed.

#### Acknowledgements

This research is funded by the Department of Health and Social Care using UK Aid funding and is managed by the Engineering and Physical Sciences Research Council (EPSRC; grant EP/R013764/1). The views expressed in this publication are those of the author(s) and not necessarily those of the Department of Health and Social Care.

#### References

- M. Gosse, C. Jones, D. Jesus, S. D'Costa, 2022, Regulatory & supply chain implications for plasmids as critical starting materials in the manufacture of viral vector gene therapy products. *Cell and Gene Therapy Insights*, 8(2), 279–286.
- J. Ohlson, 2020, Plasmid manufacture is the bottleneck of the genetic medicine revolution, *Drug Discovery Today*, 25(11), 1891–1893.
- F. Boukouvala, M. Ierapetritou, 2013, Surrogate-Based Optimization of Expensive Flowsheet Modeling for Continuous Pharmaceutical Manufacturing, *Journal of Pharmaceutical Innovation*, 8, 131–145.
- Z. Wang, M. Sebastian Escotet-Espinoza, M. Ierapetritou, 2017, Process analysis and optimization of continuous pharmaceutical manufacturing using flowsheet models, *Computers and Chemical Engineering*, 107, 77–91.
- D. van de Berg, T. Savage, P. Petsagkourakis, D. Zhang, N. Shah, E. A. del Rio-Chanona, 2022, Data-driven optimization for process systems engineering applications, *Chemical Engineering Science*, 248, 117135.
- R. Ferreira, D. Petrides, 2021, Plasmid DNA (pDNA) Large Scale Manufacturing – Process Modeling and Techno-Economic Assessment (TEA) using SuperPro Designer, doi: 10.13140/RG.2.2.12780.28800.
- S. Kucherenko, 2013, SobolHDMR: A general-purpose modeling software. *Methods in Molecular Biology*, 1073, 191–224.
- C. E. Rasmussen, C. K. I. Williams, 2006, Gaussian Processes for Machine Learning, *The MIT Press*.
- M. Balandat, B. Karrer, D. R. Jiang, S. Daulton, B. Letham, A. Gordon Wilson, E. Bakshy, 2020, BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization, *Advances in Neural Information Processing Systems*, 33.

# Competitive adsorption of copper, nickel, and chromium ions onto EDTA modified SBA-15

Bawornpong Pornchuti,<sup>a</sup> Yuttana Phoochahan,<sup>a</sup> Prarana Padma,<sup>a</sup> Suchada Ruengrit,<sup>a</sup> Pravit Singtothong,<sup>b</sup>

<sup>a</sup>*Department of Process and Industrial Engineering, School of Engineering and Industrial Technology, Mahanakorn University of Technology, Bangkok 10530, Thailand*

<sup>b</sup>*Department of Chemistry, Faculty of Science, Mahanakorn University of Technology, Bangkok 10530, Thailand*

## Abstract

Mathematical modeling is an important tool in chemical engineering. Not only to understand processes, it is also used in many fields such as design and scale-up, process control, optimization, experimental design, trouble shooting, etc. One application in adsorption process is to represent adsorption isotherm. In contrast to the real processes, most of researches studied single-component isotherm. Wastewater from electroplating industry always comprises of several heavy metal ions which are harmful to human and ecosystem. In this work, EDTA modified SBA-15 was used to treat copper, nickel, and chromium ions in multicomponent system. It was found that adsorption of heavy metals preferred higher pH. Moreover, adsorption of nickel ions was very low. The multicomponent isotherm models were applied. The RMSEs (Residual root mean square errors) of Freundlich, Langmuir, extended Langmuir, and modified competitive Langmuir models were 2.02, 3.37, 9.12, and 10.72, respectively.

**Keywords:** adsorption, heavy metals, multicomponent isotherm, EDTA, SBA-15.

## 1. Introduction

The COVID-19 pandemic extremely affected this world. Million people died. Many cities were locked down. Human activities dropped dramatically. This pandemic still had positive impacts such as bluer skies, fewer crimes, lower accidents, better hygiene, stronger relationships, etc. Considering environmental remediation, it was evident that pollution was mainly caused by human. Many pollutants were contaminated in air, water, and soil.

Heavy metals are hazardous pollutants. They are always contaminated in wastewater from electroplating processes. One effective way treating heavy metal ions is adsorption. Adsorbents are the key of adsorption processes. Novel adsorbents are developed and investigated.

Surface modification is used to improve adsorption performance. Silica is favored in adsorbent production because it is easy to adjust its surface and has large surface area. Silica materials can have different pore structures. However, our previous study found that adsorption capacities of modified silica do not depend on pore structure (Pornchuti et al., 2021). As a consequence, SBA-15 was selected in this work since it is convenient



to synthesize. In addition, EDTA was grafted onto SBA-15 surface to catch copper, nickel, and chromium ions in multicomponent system.

Generally, real wastewater contains various heavy metal ions. There are interactions among these ions (Girish, 2017). Consequently, the study of multi-ion adsorption is essential. In case of single ion adsorption, Freundlich and Langmuir models are always chosen for the representation of adsorption isotherm. These two models are simple but in case of multi-ions system they do not include the effect of other ions. For multicomponent system, extended Langmuir and modified competitive Langmuir models were developed to cover their interaction. As a consequence, Freundlich, Langmuir, extended Langmuir, and modified competitive Langmuir models were selected in this study. In summary, this work was set up to study the modeling of competitive adsorption of copper, nickel, and chromium ions onto EDTA modified SBA-15.

## 2. Materials and Methods

### 2.1. Materials

TEOS (Tetraethyl orthosilicate), Pluronic P123, APTES (Aminopropyltriethoxysilane), EDTA disodium salt, ethanol, toluene, HCl, acetic acid, and ammonium hydroxide were purchased from Sigma Aldrich or Merck.

### 2.2. Synthesis of SBA-15

SBA-15 was synthesized by following the method of Naik et al. (2011). In brief, 6.64 g of Pluronic P123 and 13.5 mL of HCl solution were agitated with 202 mL of deionized water. Then 13.86 g of TEOS was dropped into the solution. The mixture was kept for 24 h at 90°C. After vacuum filtration, the solid was washed and dried. SBA-15 was obtained by calcination at 550°C for 3 h.

### 2.3. Integration of amino functional groups

Amino functional group was grafted onto SBA-15 surface by applying the method of Parida and Rath (2009). At first, 2 g of SBA-15 was suspended in 60 mL of toluene. Next, 1.2 mL of APTES was added dropwise. The mixture was refluxed at 110°C for 8 h. The solid was dried after washing with ethanol and deionized water.

### 2.4. Incorporation of EDTA

The method of Repo et al. (2009) was adapted to the incorporation of EDTA. Briefly, 2 g of amine functionalized SBA-15, 5.4 g of EDTA disodium salt, and 42.0 mL of ethanol-acetic acid solution were mixed together. The solid was filtered and washed with ammonium hydroxide solution, ethanol-acetic acid solution, deionized water, and ethanol, respectively. The EDTA modified SBA-15 was received after dried at 105°C.

### 2.5. Multicomponent Adsorption

Normally, multi-ion adsorption experiments were conducted as follows. 100 mg of EDTA modified SBA-15 was shaken in 50 mL of multi-ion solution at 105 rpm for 48 h. The solid was separated by vacuum filtration. The filtrate was analyzed by atomic absorption spectroscopy (AAS). Eq. (1) was applied to calculate adsorption capacity.

$$q_e = \frac{(C_i - C_e)V}{m} \quad (1)$$

$q_e$  is equilibrium adsorption capacity;  $C_i$  is metal concentration at start;  $C_e$  is equilibrium metal concentration;  $V$  is the volume of metal solution;  $m$  is the adsorbent mass. To compare the suitability between models, the RMSE was determined from Eq. (2).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (q_{i,exp} - q_{i,model})^2} \quad (2)$$

$q_{i,exp}$  is the experimental adsorption capacity of component  $i$  and  $q_{i,model}$  is the adsorption capacity of component  $i$  calculated from the model.

### 3. Results and Discussion

#### 3.1. Effect of pH

The amount of metal ions adsorbed on EDTA modified SBA-15 at various pH was shown in Figure 1. Effect of pH on competitive adsorption was similar to those of single metal adsorption. Metal adsorption was hindered by hydrogen ion at low pH. As a result, the amount of adsorbed metals increased with an increase of pH.

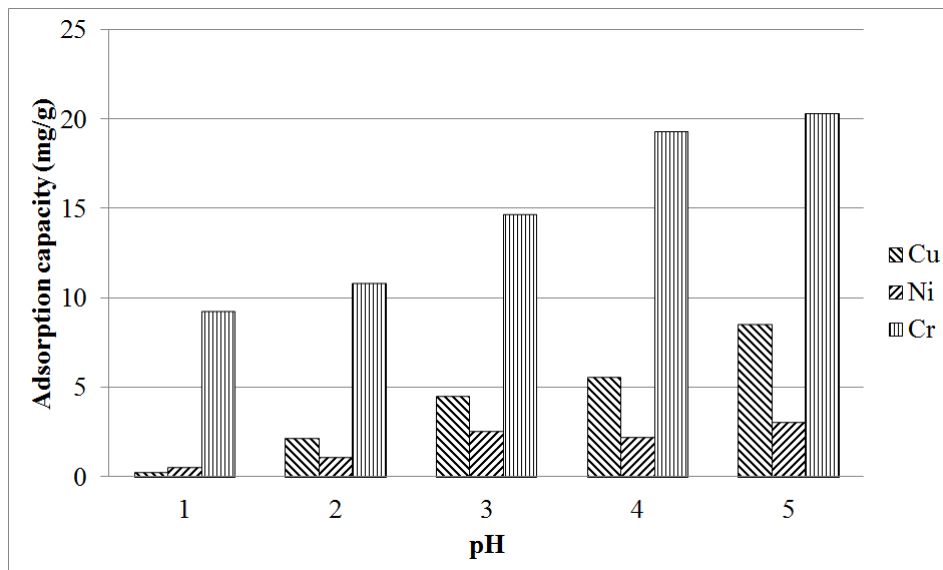


Figure 1. pH Effect.

#### 3.2. Mathematical Modeling

Adsorption isotherm was illustrated as Figure 2. It can be seen that adsorption of nickel ion was very low comparing to copper and chromium ions. It was suppressed by other ions. Other adsorbents were also found this antagonistic effect such as adsorption by amine functionalized SBA-15 (Pornchuti et al., 2022), olive stones, date stones, peat, sphagnum moss, laterite, and bentonite (Girish 2018). Since the amount of nickel adsorbed was low, the mathematical models were applied only with copper and chromium ions.

Although Freundlich and Langmuir models are designed for single component adsorption, their utilization in multicomponent adsorption can be found such as the investigation of Qi and Pichler (2017). Freundlich model can be represented as Eq. (3).

$$q_{e,i} = K_{F,i} C_{e,i}^{1/n_i} \quad (3)$$

$K_F$  and  $n$  are the Freundlich constants, while  $i$  indicates the component. Langmuir model is expressed as Eq. (4).

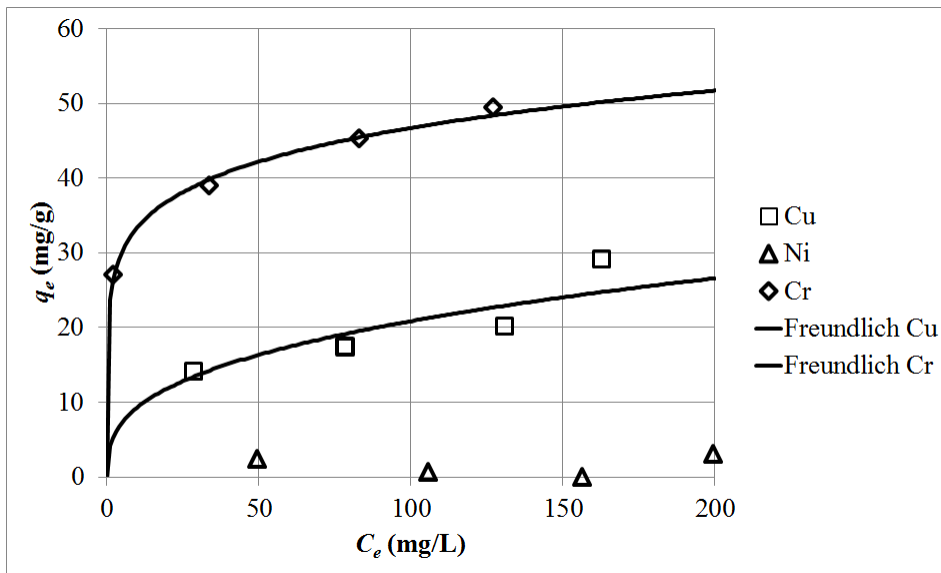


Figure 2. Adsorption isotherm.

$$q_{e,i} = \frac{q_{m,i} b_{L,i} C_{e,i}}{1 + b_{L,i} C_{e,i}} \quad (4)$$

$q_m$  is the maximum loading and  $b_L$  is the Langmuir constant. The linearized form of both models were applied to compute their parameters and listed in Table 1.

Table 1. Parameters of isotherm models.

Parameters of isotherm models	Cu	Cr
Freundlich isotherm		
$K_{F,i}$	4.1505	23.7739
$n_i$	2.8531	6.8213
Langmuir isotherm		
$q_{m,i}$	26.5252	45.2489
$b_{L,i}$	0.0382	0.6500
Extended Langmuir isotherm		
$q_{m,i}$	23.6173	206.3216
$b_{L,i}$	13.1572	5.9948
Modified competitive Langmuir isotherm		
$q_{m,i}$	97.0874	163.9344
$b_{L,i}$	0.0258	0.0667
$\eta_{L,i}$	5.8333	8.9594

Extended Langmuir model is shown as Eq. (5).

$$q_{e,i} = \frac{q_{m,i} b_{L,i} C_{e,i}}{1 + \sum_{j=1}^N (b_{L,j} C_{e,j})} \quad (5)$$

Total number of ions is represented as  $N$ , while  $j$  indicates the component. The parameters of this model were computed by minimization of the error in non-linear regression analysis and shown in Table 1. Modified competitive Langmuir model can be written as Eq. (6).

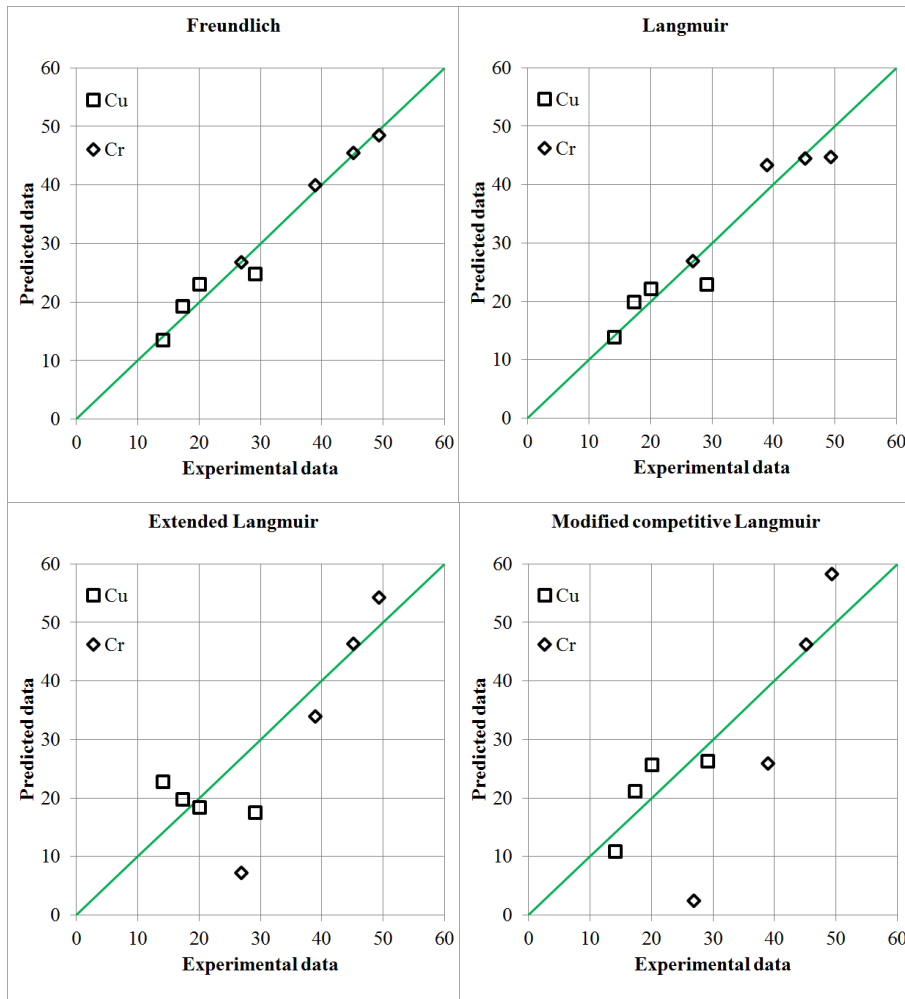


Figure 3. Comparison between predicted adsorption capacities and experimental data.

$$q_{e,i} = \frac{q_{m,i} b_{L,i} (C_{e,i} / \eta_{L,i})}{1 + \sum_{j=1}^N (b_{L,j} (C_{e,j} / \eta_{L,j}))} \quad (6)$$

$q_m$  and  $b_L$  in this model is Langmuir parameters of single-ion adsorption. They were obtained from our previous study (Pornchuti et al., 2021).  $\eta_L$  is interaction factor. They were computed from experimental data of multi-ion adsorption by minimization of the

error in non-linear regression analysis. All parameters of this model were listed in Table 1.

Each isotherm models were utilized to predict adsorption capacities and compare with the experimental data as demonstrated in Figure 3. The RMSEs of Freundlich, Langmuir, extended Langmuir, modified competitive Langmuir models were 2.02, 3.37, 9.12, and 10.72, respectively. It obviously indicated that Freundlich and Langmuir were more appropriate in the representation of experimental data in this case. Many researches illustrated isotherm in 2D like Figure 2 which lead to erroneous interpretation. The concentration of each ion was related to the others. It should be realized that both models cannot explain the effect of other ions. For extended Langmuir and modified competitive Langmuir models, they took the concentration of other ion in the model but gave high RMSEs. This might be caused by the complexity of the models and more observational data were required to improve their RMSEs.

#### 4. Conclusions

Simultaneous adsorption of copper, nickel, and chromium ions onto EDTA modified SBA-15 was conducted. It was figured out that adsorption of these ions favored high pH. Moreover, nickel ion was hindered by other ions. Finally, fitting of Freundlich model was better than those of Langmuir, extended Langmuir, and modified competitive Langmuir models.

#### References

- C. R. Girish, 2017, Various Isotherm Models for Multicomponent Adsorption: A Review, *Journal of Civil Engineering and Technology*, 8, 80-86
- C. R. Girish, 2018, Multicomponent Adsorption and the Interaction between the Adsorbent and the Adsorbate: A Review, *International Journal of Mechanical Engineering and Technology*, 9, 177-188
- B. Naik, V. Desai, M. Kowshik, V. S. Prasad, G. F. Fernando, N. N. Ghosh, 2011, Synthesis of Ag/AgCl-Mesoporous Silica Nanocomposites using a Simple Aqueous Solution-Based Chemical Method and a Study of Their Antibacterial Activity on *E. coli*, *Particuology*, 9, 243-247
- K. M. Parida, D. Rath, 2009, Amine Functionalized MCM-41: An Active and Reusable Catalyst for Knoevenagel Condensation Reaction, *Journal of Molecular Catalysis A: Chemical*, 310, 93-100
- B. Pornchuti, B. Pongpattananurak, D. Sutthiard, P. Singtothong, 2021, Adsorption of Copper, Nickel, Chromium Ions Using EDTA Modified Silica Adsorbents, The 30<sup>th</sup> TIChE Conference (TIChE2021), Nakhon Ratchasima, Thailand, 6-7 May 2021, 744-749
- B. Pornchuti, Y. Phoochahan, P. Padma, S. Ruengrit, P. Singtothong, 2022, Competitive Adsorption of Copper, Nickel, and Chromium Ions onto Amine Functionalized SBA-15, *Computer Aided Chemical Engineering*, 49, 2065-2070
- P. Qi, T. Pichler, 2017, Competitive Adsorption of As(III), As(V), Sb(III), and Sb(V) onto Ferrihydrite in Multi-Component Systems: Implications for Mobility and Distribution, *Journal of Hazardous Materials*, 330, 142-148
- E. Repo, T. Agustiono Kurniawan, J. K. Warchol, M. E. T. Sillanpää, 2009, Removal of Co(II) and Ni(II) Ions from Contaminated Water using Silica Gel Functionalized with EDTA and/or DTPA as Chelating Agents, *Journal of Hazardous Materials*, 171, 1071-1080

# Auto-extractive reactive distillation for the production of fatty esters

Hernán D. Muñoz,<sup>a,b</sup> Christian Hoffmann,<sup>b</sup> Markus Illner,<sup>b</sup> Álvaro Orjuela,<sup>a</sup>  
Jens-Uwe Repke<sup>b</sup>

<sup>a</sup>*Department of chemical and Environmental Engineering, Universidad Nacional de Colombia, Bogotá D.C. 111321, Colombia*

<sup>b</sup>*Technische Universität Berlin, Sekr. KWT 9, Process Dynamics and Operations Group, Str. des 17. Juni 135, D-10623 Berlin, Germany*

[hdmunozg@unal.edu.co](mailto:hdmunozg@unal.edu.co)

## Abstract

This work proposes a systematic methodology based on graphic methods for the conceptual design of simultaneous reactive distillation-extractive distillation processes in one single apparatus (named extractive-reactive distillation). The configuration is useful to overcome azeotropes that reactive distillation alone cannot. A key assumption of the methodology is the use of an entrainer that is part of the components consumed in the reaction (i.e. an auto-entrainer). In the methodology, the analysis of the statics is used to generate a reactive distillation configuration, while an iso-volatility and bifurcation analysis is used to generate an extractive distillation configuration. For the combined system, we propose to extend the analysis of the statics to include the auto-entrainer effect in reactive distillation. The methodology is applied to the production of isopropyl myristate from acid oleochemical streams. A new column configuration that has the potential to reduce additional azeotropic separation units is proposed.

**Keywords:** Reactive distillation; extractive distillation; isopropyl myristate

## 1. Introduction

Reactive distillation (RD) is a well-known method for process intensification. In RD, reactants are fed to a reactive section and the products are separated via distillation in the same apparatus at the same time. In some special cases, however, an RD process may also show an auto-extractive behavior (Coker, 2010), i.e., one of the reactants behaves like an internal mass separation agent (entrainer), so that at least one of the azeotropes in the system is overcome. In such cases, the entrainer is the component with the highest boiling point in the system and does not form azeotropes with the other components. The double functionality reactant-entrainer generates a synergetic effect, which is not always obvious. This effect can be exploited in the synthesis of intensified process configurations that have so far not been reported in the literature. In spite of this potential benefit, available methodologies only consider the reactive phenomena (Almeida-Rivera et al., 2004). This contribution presents a methodology that explicitly takes the auto-extractive behavior into account. The methodology was applied exemplarily on the esterification of fatty acids, e.g., myristic acid (MA), with isopropanol (IPA) to produce isopropyl myristate (I-MA). Besides its auto-extractive behavior, the system is of interest because it can be used for the valorization of acid oleochemical streams, which are an industrial

waste rich in free fatty acids. In these systems, the alcohol (i.e., IPA) forms a minimum-boiling azeotrope with water (W) produced by the reaction (Bock et al, 1997). In addition, the system contains a liquid phase split that is considered during the analysis.

**Table 1.** Methodology outline.

Step	Purpose	Input	Output
<b>Topology</b> (quaternary system)	Determination of the thermodynamic characteristics of the system.	Component boiling points. Thermodynamic model.	Graphical representation of phase equilibria. Residue curves. Characteristic distillation maps.
<b>Analysis of the statics</b> (quaternary system)	Determination of the initial single-feed structure of the RD..	Stoichiometric coefficients. Chemical equilibrium. Residue curves.	Limiting steady states. Trial trajectory. Localization of the reactive zone.
<b>Isovolatility</b> (ternary subsystem)	Determination of the extractive distillation configuration (direct or indirect).	Thermodynamic model.	Locus of the constant relative volatility in the composition space.
<b>Bifurcation analysis</b> (ternary subsystem)	Determination of feasible entrainer flows in extractive distillation	Pinch equations. Initial guess entrainer/feed ratio.	Pinch branches. Pitchfork bifurcation.
<b>Coupled synthesis</b> (quaternary-ternary system)	Determination of the single-feed reactive distillation- extractive distillation configuration	Feasible reactive distillation trial trajectories. Feasible extractive distillation mass balance lines.	Reactive distillation feed composition that matches bottoms composition of extractive distillation in the same composition space. Feasible entrainer to feed ratio

## 2. Methodology

An outline of the proposed methodology is presented in **Table 1**, in which the motivation, the input, and the output of each step is given. The individual steps are described in more detail in the following sections.

### 2.1. Thermodynamic Topological analysis

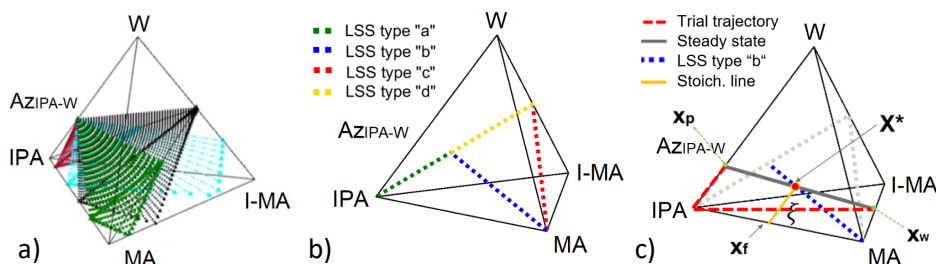
Most of the graphic methods in distillation begin with a thermodynamic characterization of the systems based on its fixed points and boundaries. The case study in this contribution is a reactive quaternary system: IPA (21,15°C), W (100,00°C), I-MA (314,85°C) and MA (326,20°C). There are four ternary subsystems, from which the subsystem IPA-W-MA presents a residue curve map type 1-0-0 according to the nomenclature reported by Matsuyama et al. (1977). This is the characteristic map of extractive distillation with heavy entrainer and without additional azeotrope formation (Foucher et al., 1991). There are six binary subsystems, one of them with a homogeneous binary azeotrope (IPA-W). The map does not include first-order separatrices (distillation boundaries), but it has a second-order separatrix (geometric) that subdivides the composition space in two subregions with a different saddle node each one. There is a liquid-liquid equilibrium in three of the four ternary subsystems and in the quaternary region near the water node.

### 2.2. Analysis of the statics (AS)

The method is based on the geometric theory of distillation and was recently reviewed by Cardona et al (2020). The AS determines the limiting steady states (LSS) of a reactive distillation system, which are defined as the steady states that maximize the yield of a target product. The AS uses a reactive distillation process model that assumes a) kinetically reversible chemical reaction in the liquid phase, b) an infinitely high separation efficiency ( $\infty/\infty$ ), c) a sufficiently high volume of the reaction zones in the column, and d) catalyst placed anywhere in the column. The AS requires little initial information, such

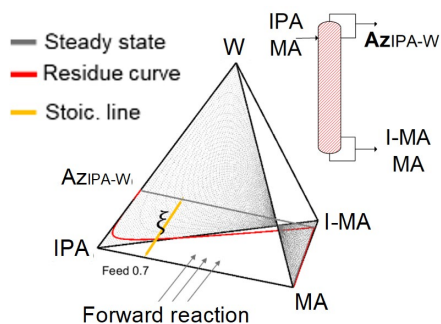
as the reaction stoichiometry and parameters for the physical and chemical equilibria. The procedure that is followed in the analysis of the statics consists of the following steps. For more details, the reader is please referred to (Cardona et al, 2020) or the original paper by Pisarenko et al. (2001) in which the terms are explained more extensively.

1. *Complete set of steady states* These steady states are lines representing mass balances of distillation columns where for a defined top or bottoms, the conjugated product falls in the boundary of the composition space. Each steady state is determined for a given feed composition and conversion. For a direct arrangement in the production of isopropyl myristate, a representative sample of the steady states is shown in **Figure 1a**.
2. *Limiting steady states based on complete set of steady states.* From the lever rule it is known that the steady states have a corresponding relation of product flows P/W (top flow/bottom flow). This relation depends on conversion and feed composition. By computing P/W as function of the conversion, it is possible to find the LSS. In the case study, the LSS were grouped in the a, b, c, d types shown in **Figure 1b**.
3. *Segregation of LLS based on feasibility criterion of distillation to ensure the occurrence of all chemical species for the separation.* For the case study, only the LSS type “b” were found to be feasible because only this line contains all four components. The other three lie on the ternary sub-planes.
4. *Determination of the reactive distillation process design able to develop the feasible LSS.* For the case study, this is exemplified in **Figure 1c** for the feasible LSS corresponding to a stoichiometric feed composition  $X_F$ . This feed follows a reaction along a stoichiometric line until the feasible steady state is achieved, which is characterized by the conversion,  $\xi$ . For this steady state,  $X^*$  represents the overall composition of the column products in mixture that are joined by a mass balance between the top product  $X_p$  and the bottom product  $X_w$  of the column. A “trial distillation trajectory” (Pisarenko et al., 2001) joins the top and bottoms product of the mass balance. This trial trajectory is a simplified representation of a residue curve starting at the top composition, and therefore it is similar to the trajectory followed by a column with infinite reflux and number of separation stages  $\infty/\infty$ .
5. *Flowsheet synthesis.* The synthesis of the RD flowsheet begins with the definition of the reactive and non-reactive sections of the column. The location of the trial trajectory with respect to the equilibrium defines those profile sections, in which the reaction should be promoted with a catalyst addition. In the case study, the entire trial trajectory lays in a region of the composition space where the presence of catalyst promotes the forward reaction. Consequently, the whole column is a catalytic section, see **Figure 2**.

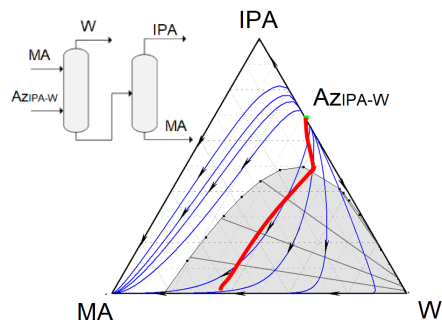


**Figure 1:** Analysis of the statics for the esterification of myristic acid with isopropanol, a) set of steady states, b) set of limiting steady states, c) trial trajectory for the limiting steady state type b that corresponds to an equimolar feed.





**Figure 2:** Smoothing of the trial trajectory to a residue curve, configuration.



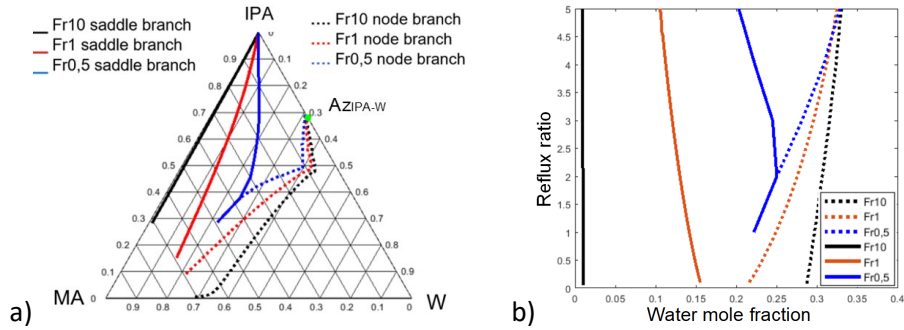
**Figure 3:** Residue curve map (blue lines) for the ternary system IPA-W-MA. Univolatility line (red).

### 2.3. Iso-volatility curve

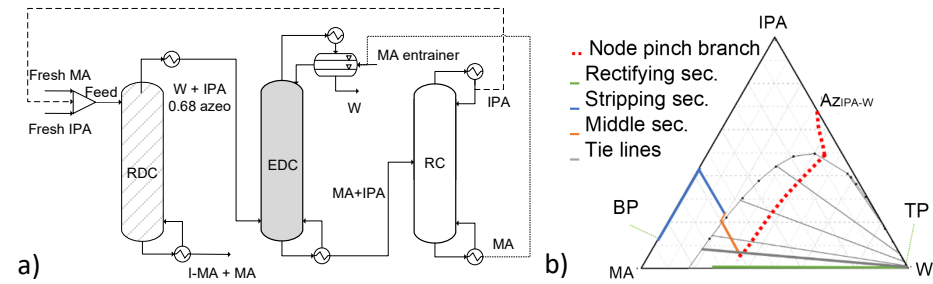
An iso-volatility curve is the locus of the ternary composition space in which the relative volatility of two defined components is constant. The iso-volatility curve of constant value one is a uni-volatility line that starts in the azeotrope composition and ends in one of the opposite edges of the ternary composition space. In the ternary system IPA-W-MA, the univolatility line falls into the edge W-MA, see **Figure 3**. This indicates that the distillate of the extractive distillation column is almost pure water (indirect configuration). Consequently, the bottoms product is a mixture IPA-MA that requires an additional fractionation column for separation. The two-column arrangement is shown in the upper left side of the **Figure 3**. For a given MA/azeotrope feed ratio, it is possible to define a global mass balance and specify all the streams. However, the MA/azeotrope feed ratio is sometimes difficult to obtain and requires the more detailed analysis introduced in the next section.

### 2.4. Bifurcation analysis

A method for the definition of the minimum entrainer flow (minimum MA/azeotrope feed ratio) that is required for separation via extractive distillation was previously proposed by Knapp et al. (1994). The method is based on the bifurcation analysis of the pinch branches of the middle section in an extractive distillation column. The bifurcation analysis observes the behavior of the middle section pinch branches with respect to small additions of entrainer. Thus, the minimum amount of entrainer is determined with the help of the phenomenon called “pitchfork”, which precedes a strong change in the topology of the pinch branches (please refer to Figure 14 presented by Knapp et al. (1994) for a visualization of this behavior). For the ternary system IPA-W-MA, different additions of entrainer generated the pinch branches changes shown in **Figure 4**. For an entrainer-to-feed ratio  $Fr=0.5$ , **Figure 4** shows that the end points of the pinch branches have achieved a reflux ratio of zero, and even with more additions of entrainer it is not possible to find the pitchfork phenomena. This indicates that for the case study, extractive distillation alone is not feasible because it is not possible to find a continuous profile joining the stripping-middle section to the rectifying section of the column in absence of the pitchfork phenomena (see **Figure 5,b**). However, if the rectifying section is replaced by a top decanter as is shown in **Figure 5a**, the LLE helps to join the top product water in a heterogeneous extractive distillation column profile (**Figure 5,b**). As the bottom product is mainly a mixture of IPA-MA, it is possible to separate this mixture and send the IPA back to the RDC and the MA to EDC (**Figure 5a**).



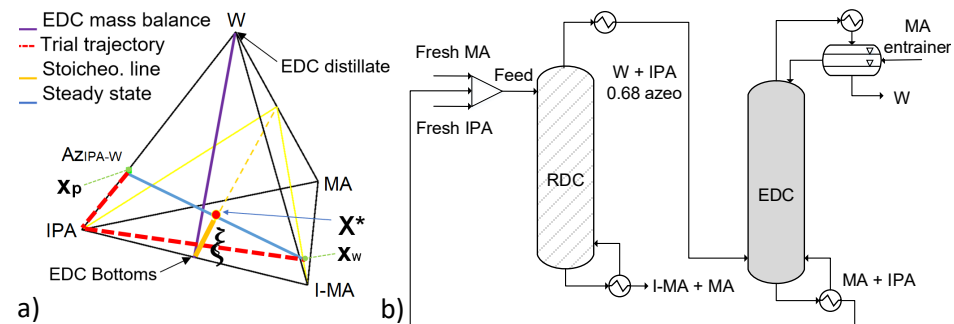
**Figure 4:** a) Pinch branches at entrainer to feed ratios of 0.5, 1 and 10; b) reflux ratio along the pinch branches



**Figure 5:** a) Heterogeneous extractive distillation column (EDC) with recovery column (RC) to overcome the azeotrope formed in the reactive distillation column (RDC) in **Figure 2**. b) Composition profiles of the EDC (BP=Bottom product; TP=top product).

2.5. Coupled system

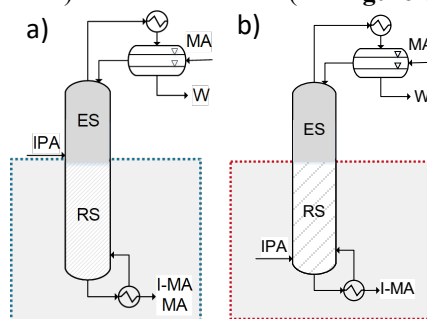
Both the reactive distillation and the extractive distillation can be represented in the same quaternary composition space. Assuming the top product of the extractive column is almost pure water, it is possible to plot a mass balance line that starts at the water vertex and ends in the binary edge IPA-MA, see **Figure 6a**. As the bottom composition of the EDC is similar to the feed composition of the RDC, they can be recirculated, and with this, the recovery column is not necessary anymore, see **Figure 6b**. The bottoms binary composition depends on the amount of entrainer that has been added in the extractive distillation column, but this amount can be varied to achieve different stoichiometric lines in the reactive column.



**Figure 6:** a) Representation of the overall extractive distillation mass balance and a feasible reactive distillation steady state, b) corresponding flowsheet.

### 3. Results and discussion

Based on the proposed methodology, the coupled arrangement showed in **Figure 6** reduces one of the columns in the **Figure 5** (the RC). On the other hand (see **Figure 6**), as the bottoms product of the EDC is similar to the RDC feed, and the top product of the RDC is similar to the EDC feed, the two totally integrated configurations shown in **Figure 7a** and **b**, are proposed. The main difference between both configurations is the IPA feed location in the reactive section (RS). When the feed stage is located above the reactive section (**Figure 7a**), the section is a single-feed reactive section and follows the behavior of a reactive residue curve. Such a behavior limits the quality of the products because the I-MA falls in a saddle node, and therefore it cannot be obtained as pure component. This drawback is avoided by feeding the alcohol at the bottom of the reactive section (**Figure 7b**). In this case a double-feed reactive section is generated (the section is defined by the location of the catalyst) and, in analogy to extractive distillation, the area of possible products is expanded.



**Figure 7:** a) Extractive distillation with reactive stripping, b) reactive distillation with extractive rectification.

### 4. Conclusions

A method for designing an auto-extractive reactive distillation column has been presented. This method contributes to synthesis of configurations able to overcome azeotropes that reactive distillation alone cannot. The method was applied for the esterification of myristic acid with isopropanol. It was demonstrated that this system presents an auto-extractive phenomenon. Given the heterogeneous nature of such phenomenon, a decanter was considered by the methodology. An extension of the analysis of the statics was presented to include the auto-extractive phenomenon. The analysis was used to generate a process flowsheet that avoids the use of a secondary column in the recovery of the auto-entrainer. The extended analysis of the statics was also used to propose a configuration with a two-feed, heterogeneous, auto-extractive reactive distillation that ensures pure product streams. To our understanding, this configuration has not been proposed in previous studies for the production of IMA. A reduction in capital and energy cost is expected with the new configuration.

### References

- Almeida-Rivera, C.P., Swinkels, P.L.J., Grievink, J., 2004. Designing reactive distillation processes: Present and future. *Comput Chem Eng* 28, 1997–2020.
- Bock, H., Wozny, G., Gutsche, B., 1997. Design and control of a reaction distillation column including the recovery system. *Chemical Engineering and Processing* 36, 101–109.
- Cardona Alzate, C., Ortiz Sanchez, M., Pisarenko, Y.A., 2020. *Reactive Separation for Process Intensification and Sustainability*. CRC Press, New York.
- Foucher, E.R., Doherty, M.F., Malone, M.F., 1991. Automatic Screening of Entrainers for Homogeneous Azeotropic Distillation. *Ind Eng Chem Res* 30, 760–772.
- Knapp, J.P., Doherty, M.F., 1994. Minimum entrainer flows for extractive distillation: a bifurcation theoretic approach. *AIChE Journal* 40, 243–268.
- Matsuyama, H., Nishimura, H., 1977. Topological and thermodynamic classification of ternary vapor-liquid equilibria. *J. Chem. Eng. Japan* 10, 181–187.
- Pisarenko, Y.A., Serafimov, L.A., Cardona, C.A., Efremov, D.L., Shuwalov, A.S., 2001. Reactive distillation design: analysis of the process statics. *Reviews in Chemical Engineering* 17, 253.

# Long-term Capacity Management of Pharmaceutical Manufacturing Networks

Simon B. Lindahl,<sup>a,b</sup> Deenesh K. Babi,<sup>b</sup> Gürkan Sin<sup>a</sup>

<sup>a</sup>*Process and Systems Engineering Center (PROSYS), Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 228 A, 2800 Kgs. Lyngby, Denmark*

<sup>b</sup>*PS API Manufacturing Development BIO, Novo Nordisk A/S, Novo Allé 1, 2880 Bagsværd, Denmark*

## Abstract

This work presents a stochastic capacity and production planning model for network environments with multiple stages and shared manufacturing lines. Network-wide capacity decisions are modelled under uncertainties in demand and objective function coefficients with uncorrelated capacity expansions. A framework is presented for evaluating the model under uncertainty and it is applied to a case study from pharmaceutical manufacturing. First, the relevant uncertainties are defined through a sensitivity analysis and their effects on the optimal capacity decisions are explored. A sensitivity analysis of the input distributions is used to systematically study the required locations and total expansion amounts subject to different means, standard deviations, trends, and distribution types for the relevant uncertain demands. It is shown that capacity requirements are underestimated when the uncertainties are not considered, and that uncertainty sets with larger demand volatilities lead to larger capacity expansions.

**Keywords:** Capacity expansion, production planning, stochastic programming, sensitivity analysis

## 1. Introduction

Capacity planning (CP) is essential to ensure the long-term ability to deliver medicines and increase revenues from exclusive sales. Most pharmaceuticals are made in multi-stage processes where intermediates are produced for storage and each manufacturing line can produce multiple products. As such, production planning (PP) details should be considered in the CP problem to describe the allocation of time on shared manufacturing lines and the inventories of all materials (Marques et al. (2020)). Capacity decisions are resource intensive and irreversible, so the decisions should be made while considering relevant parameter uncertainties. An early work used 2-stage stochastic programming to define which new facilities to construct in a single-stage environment under clinical trial uncertainty (Levis & Papageorgiou (2004)). C. M. Marques et al. (2018) used Monte Carlo Simulations to find candidate process designs for a simplified flowsheet which were passed on to a second model that decided the optimal design per period for each product. Uncertainties in clinical trials and product demands were considered. Vieira et al (2020) applied a 2-stage stochastic programming model for design and scheduling of multiproduct plants under demand uncertainty to achieve a trade-off between profits and financial risk. Chou et al. (2014) used a dynamic programming model to determine capacity requirements subject to uncertain demands with constant trends and different

volatilities. They conclude that higher demand volatilities increase the capacity requirements for problems with demand satisfaction. Sensitivity and uncertainty analysis can be used to identify how individual parameter uncertainties affect the model output, which is the sensitivity analysis problem, or what uncertainty in the model outputs can be expected subject to an uncertain input space, which is the uncertainty analysis problem (Saltelli et al. (2019)). Different methods have been developed to consider uncertainties directly in optimization models (Gabrel et al. (2014)) but even if a representation of uncertainty is considered that representation might itself be uncertain. This work therefore studies the effects of different uncertainty sets on the here-and-now capacity decisions to provide a recommendation on where and how much to expand capacity.

### 1.1. Problem Statement

The engineering context for the problem is defined as follows: a company wants to determine the required capacity expansions in its manufacturing network to provide a tradeoff between inventory risk and expansion cost. The studied system is a general network with multiple stages and multiple lines on each stage. Each line can produce a set of materials with given initial capacities. The goal is to determine the location and amount of capacity expansion as well as material inventory profiles subject to uncertain demands and capacity expansion costs. The company policy requires that the inventories be maintained above the safety level and as close to the strategic level as possible. Capacity expansions for different materials are assumed independent with costs defined through economies of scale.

## 2. Methodology

### 2.1. Model

The developed model follows the state-task-network (STN) structure (Kondili et. al (1993)) that has been expanded to include variable capacities and uncertainty through its formulation as a stochastic program. The main variables are inventory for material  $m$  at time  $t$  in scenario  $s$ ,  $I_{mts} \in \mathbb{R}^{\geq 0}$ , safety inventory violation,  $B_{mts} \in \mathbb{R}^{\geq 0}$ , inventory distance above/below the strategic inventory,  $I_{mts}^+/I_{mts}^- \in \mathbb{R}^{\geq 0}$ , manufacturing amount of task  $i$  on line  $j$ ,  $P_{ijts} \in \mathbb{R}^{\geq 0}$ , fractional time allocation  $W_{ijts} \in \mathbb{R}^{\geq 0}$  and the choice of capacity level  $n$ ,  $Y_{ijn} \in \{0,1\}$ . The allocation variable from the original STN has been relaxed and a variable capacity has been added to choose the required capacity level  $\kappa_{ijn}$  through the binary variable  $Y_{ijn}$  for all task-line combinations under demand and cost coefficient uncertainties. Capacity decisions are modelled as here-and-now variables that are chosen once for each task-line combination across all scenarios and time periods. This approach assumes independent capacity upgrades or in other words no synergies between capacity expansions. The PP details are wait-and-see variables under the assumption that the production scheduling on each line can be modified based on realized demand.

$$\text{minimize } C^I + C^r \quad (1)$$

$$I_{mts} = I_{m,t-1,s} + \sum_{i \in I_m^+} \sum_{j \in J_i} \rho_{im}^+ P_{i,j,t-\tau_i,s} - \sum_{i \in I_m^-} \sum_{j \in J_i} \rho_{is}^- P_{i,j,t-\tau_i,s} + B_{mts} - B_{m,t-1,s} - \delta_{mts}, \forall mts \quad (2)$$

$$\sigma_{mt}^{sfty} \leq I_{mts}, \quad \forall mts \quad (3)$$

$$1 - \frac{I_{mts}}{\sigma_{mt}^{strat}} \leq I_{mts}^-, \quad \forall m \in \mathbf{M}_p, ts \quad (4)$$

$$\frac{I_{mts}}{\sigma_{mt}^{strat}} - 1 \leq I_{mts}^+, \quad \forall m \in \mathbf{M}_p, ts \tag{5}$$

$$\sum_{i \in \mathbf{I}_j} \sum_{n \in \mathbf{P}_{ij}} \sum_{t'=t}^{t-\tau_i+1} W_{ijt'ns} \leq 1, \quad \forall jts \tag{6}$$

$$P_{ijts} \leq \sum_{n \in \mathbf{P}_{ij}} \kappa_{ijn} W_{ijtns}, \quad \forall j, i \in \mathbf{I}_j, ts \tag{7}$$

$$\sum_{n \in \mathbf{P}_{ij}} Y_{ijn} \leq 1, \quad \forall j, i \in \mathbf{I}_j \tag{8}$$

$$\sum_{t \in \mathbf{T}} W_{ijtns} \leq Y_{ijn} |\mathbf{T}|, \quad \forall j, i \in \mathbf{I}_j, n \in \mathbf{P}_{ij}, s \tag{9}$$

$$\sum_{m \in \mathbf{M}_p} \sum_{t \in \mathbf{T}} B_{mts} = 0, \forall s \tag{10}$$

$$C^I = \sum_{s \in \mathbf{S}} \alpha_s \sum_{m \in \mathbf{M}_p} \sum_{t \in \mathbf{T}} I_{sts}^+ v_m^{I+} + I_{sts}^- v_m^{I-} + B_{mts} v_m^{I,bl} \tag{11}$$

$$C^R = \sum_{j \in \mathbf{J}} \sum_{i \in \mathbf{I}_j} \sum_{n \in \mathbf{P}_{ij}} Y_{ijn} v_{ij}^{cb} \kappa_{ijn} v_{ij}^{ce} \tag{12}$$

The objective function is defined in Eq. (1) as the minimization of costs for inventory and capacity expansion. Eq. (2) sets the inventory balance with slack variables to allow inventories to go below the safety level in Eq. (3). Eqs. (4-5) define the distance below/above the strategic safety level. Eq. (6) limits the maximum allocation to each line, Eq. (7) constraints the maximum production based on the chosen capacity level, Eq. (8) ensures that only one capacity level is chosen, and Eq. (9) forces all capacity allocations for capacity levels not chosen to 0. Eq. (10) ensures that safety violations are only possible for products. Eq. (11) defines the inventory cost for not being at setpoint with three terms for being either above/below the strategic inventory or below the safety inventory. Individual parameters are used to define their relative importance where it generally applies that  $v_m^{bl} > v_m^{I-} > v_m^{I+}$ . Finally, Eq. (12) defines the capacity expansion cost through a sum of individual capacity expansions in which an economy of scale exponent  $v_{ij}^{ce}$  is used to ensure that few larger expansions are preferred over many small expansions.

### 2.2. Uncertainty Treatment

The model uses an uncertainty set composed of discrete scenarios for demands, base capacity costs and economy of scale exponents as input. Its solution provides here-and-now CP decisions which are shared across all scenarios and wait-and-see PP decisions for each individual scenario. The demands are time series, and the cost factors depend on the required changes to achieve a given capacity expansion. The available information on the cost factors depends on the amount of work that has been carried out to investigate expansion possibilities while for demands, both the levels, trends and volatilities are uncertain. In this work a 4-step solution framework is used: (1) baseline deterministic model solution with an uncertainty set of one scenario with expected values for all parameters, (2) sensitivity analysis of the deterministic model with different single-scenario uncertainty sets to determine parameters for further analysis, (3) baseline stochastic programming solution for an uncertainty set with scenarios sampled from expected parameter distributions, (4) sensitivity analysis of the stochastic program with different uncertainty sets sampled from different parameter distributions to provide recommendations in the presence of incomplete distribution information. Scenarios are created by combining samples from each parameter distribution. For demands, a sample is drawn for each time point and Figure 1 shows 100 example paths created by

independent and identically distributed samples from a normal distribution with mean 1 and standard deviation 0.2. The paths with highest, lowest, and median mean demands are highlighted and the right plot shows the distribution of mean demands, which approaches normal distribution as sampling number increases.

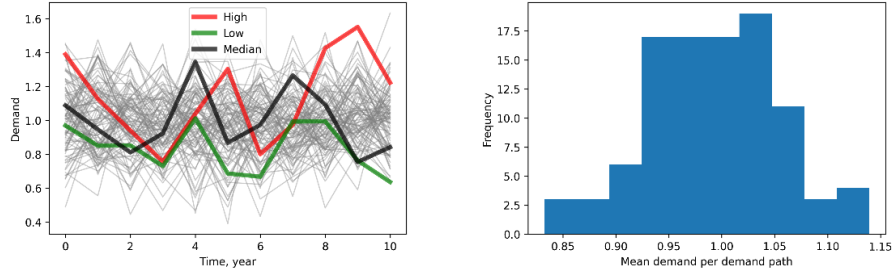


Figure 1: Demand paths (left) and mean demand distribution (right)

### 3. Case Study

The proposed model is applied to a case study abstracted from a real pharmaceutical manufacturing network consisting of three stages with three lines on the first stage, five lines on the second stage and two lines on the third stage. Six products are produced, and the total number of materials is 14. The time horizon is 10 years divided in a discrete yearly time grid. Capacity expansions of 0%, 10% or 20% were allowed.

#### 3.1. Baseline Deterministic Solution

The baseline deterministic model solution was created by solving the stochastic program with the nominal scenario. In this case, no capacity expansions are required.

#### 3.2. Sensitivity Analysis of Deterministic Model

The sensitivities to product demands were tested by sequentially solving the deterministic model with demands increased by a factor 10% or 20% at all time points for one product at a time. No demand decreases were tested as no capacity expansions were seen in the baseline solution and lower demands require less capacity. The sensitivities to the capacity base cost  $v_{ij}^{cb}$  and the capacity exponent  $v_{ij}^{ce}$  were tested by varying them around their nominal point for each demand increase of each product. Table 1 shows the total capacity expansion for each combination of increased demand and cost parameters. The results show relevant demand uncertainties for products B, C and F. Both capacity cost uncertainties are relevant for demand increases of 20%.

Table 1: Total capacity expansion requirement for increased demand of each product and different cost parameters (change from nominal value given in parenthesis)

Demand Factor	Capacity Base cost	Capacity Exponent	Product demand varied					
			A	B	C	D	F	G
	$v_{ij}^{cb}$	$v_{ij}^{ce}$	<b>Total capacity expansion, %</b>					
+10%	150	0.8	0	10	10	0	0	0
+10%	125 (-25)	0.8	0	10	10	0	0	0
+10%	175 (+25)	0.8	0	10	10	0	0	0
+10%	150	0.7 (-0.1)	0	10	10	0	0	0
+10%	150	0.9 (+0.1)	0	10	10	0	0	0
+20%	150	0.8	0	10	20	0	0	0
+20%	125 (-25)	0.8	0	10	20	0	10	0
+20%	175 (+25)	0.8	0	10	20	0	0	0

+20%	150	0.7 (-0.1)	0	40	20	0	10	0
+20%	175 (+25)	0.9 (+0.1)	0	10	20	0	0	0

3.3. Baseline Stochastic Programming Solution

Without specific distribution information, the costs were assumed uniformly distributed with the bounds from Table 1 and the demands for products B, C and F at each time point were created by multiplying the nominal demand  $\delta_{mt}^{nom}$  with a sample from a uniform distribution with a lower/upper bound of 0.8/1.2. An uncertainty set with 100 equiprobable scenarios was created by random sampling, and the stochastic program was solved which resulted in no capacity expansions. Also, the deterministic model was solved for each scenario, giving in an average total expansion of only 2.4% distributed over two different product-line combinations which indicates that capacity expansions might be achieved through optimizations of the existing process and equipment.

3.4. Sensitivity Analysis of Uncertain Parameter Distributions

A sensitivity analysis was carried out on the demand distributions to identify locations in scope for capacity expansions and to test the robustness of the recommendations. A computational study with 54 uncertainty sets was defined with variations along the demand dimensions of distribution type  $D$  (uniform, normal), base level  $\mu$  (1.0, 1.1, 1.2), standard deviation  $\sigma$  (baseline, +25%, +50%) and trend  $\psi$  (1.0, 1.01, 1.02). The demand in each period was created as  $\delta_{mts} = \delta_{mt}^{nom} \cdot \psi^t \cdot D(\mu, \sigma)$  and the stochastic program was solved for each uncertainty set while the deterministic model was solved for each scenario in each set. Figure 2 shows the total capacity expansion for each uncertainty set based on the overall mean and standard deviation of all the uncertain demands combined. Higher mean demands require larger expansions and the middle region for mean demands is affected by larger demand volatilities. The uniform and normal distributions produce similar results although larger expansions are needed in three cases for uniformly distributed demands. The deterministic model results are comparable to those of the stochastic model, but for some of the uncertainty sets it underestimates the capacity expansion due to its larger flexibility which contrasts with its higher capacity increase in section 3.3. Now, larger capacity expansions are preferred by the stochastic program to benefit the high demand scenarios while lower demand scenarios will experience overcapacity. The effect of demand trends at a given base demand can be seen by comparing the three points on the upward diagonal starting from any point with mean demand in the set (1.0, 1.1, 1.2). As such, demand trends become important if the base demand level is also increased.

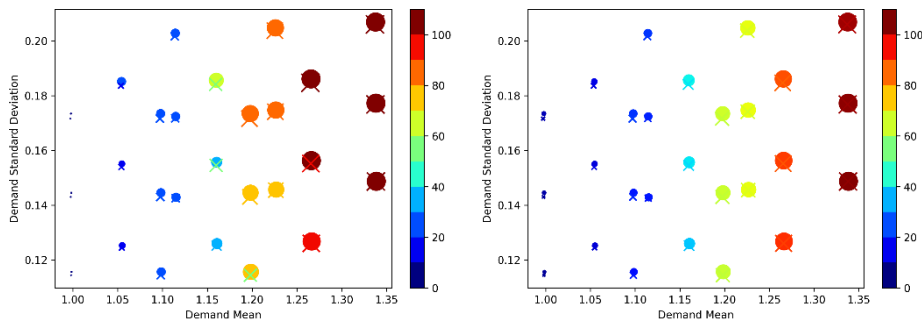


Figure 2: Total capacity expansion for uniform (o) and normal (x) demands at different means and standard deviations. Stochastic model (left) and deterministic model (right).

The model determines the locations that should be expanded, and we suggest to use this information to guide investigations into optimization or retrofit projects. Table 2 shows



the frequency of expansion choices based on the 54 uncertainty set solutions of the stochastic program. Table 2 shows that five lines may require capacity expansions if demand increases are realized, and it provides the product on that line that is relevant to expand. Naturally, an expansion for one product on a line might affect the other products as well but this interaction can be included in subsequent solutions of the provided model.

Table 2: Frequency of expansion size choice across parameter combinations

Line	A1	A2	B4	C1	C2
Product	B	C	F	B	C
0%	0.57	0.44	0.67	0.19	0.13
10%	0.0	0.26	0.17	0.37	0.37
20%	0.43	0.3	0.17	0.44	0.50

#### 4. Conclusion and Perspectives

A stochastic programming model was presented to determine the location and size of capacity expansions under uncertainty. A framework was used to find and explore the input uncertainties that influence the capacity decisions. The relevant parameters were determined through sensitivity analysis of a deterministic model. Then, a sensitivity analysis of the stochastic program with uncertainty sets from different distributions was used to study the decisions subject to incomplete distribution information. The framework was applied to a case study which shows that the capacity expansion is underestimated by the deterministic model and that higher demand volatilities increase the expansion need at moderate demand increases. Generally, the same expansion was needed regardless of the distribution type but in some cases, the requirements were larger for uniformly versus normally distributed demands. The proposed framework can be used to identify where to expand capacity to meet long-term demand given limited information on the input uncertainties. As such, the results can be used to identify the need for retrofit and optimization projects as well as determine which product demands that should be monitored more closely since their increase will require capacity expansions.

#### References

- Y. C. Chou, W. C. Sung, G. Lin & J. Jahn, (2014). "A comparative study on the performance of timing and sizing models of capacity expansion under volatile demand growth and finite equipment lifetime", *Computers & Industrial Engineering*, 76, 98-108.
- V. Gabrel, C. Murat, A. Thiele, (2014), "Recent advances in robust optimization: An overview", *European Journal of Operational Research*, 235, 471-483.
- A. A. Levis & L. G. Papageorgiou, (2004), "A hierarchical solution approach for multi-site capacity planning under uncertainty in the pharmaceutical industry", *Computers & Chemical Engineering*, 28(5), 707-725
- C. M. Marques, S. Moniz and J. P. de Sousa, (2018). "Strategic decision-making in the pharmaceutical industry: A unified decision-making framework." *Computers & Chemical Engineering*, **119**, 171-189
- C. M. Marques, S. Moniz, J. P. de Sousa, A. P. Barbosa-Póvoa & G. Reklaitis, (2020), "Decision-support challenges in the chemical-pharmaceutical industry: Findings and future research directions", *Computers & Chemical Engineering*, 134, 106672.
- A. Saltelli, K. Aleksankina, W. Becker, P. Fennell, F. Ferretti, N. Holst, S. Li & Q. Wu, (2019), "Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices", *Environmental Modelling & Software*, 114, 29-39.
- M. Vieira, H. Paulo, T. Pinto-Varela and A. P. Barbosa-Póvoa, (2020), "Assessment of financial risk in the design and scheduling of multipurpose plants under demand uncertainty", *International Journal of Production Research*, 59 (20), 6125-6145

# Modeling of a trickle bed reactor: Study of arabinose oxidation on gold catalyst extrudates

Mouad Hachhach<sup>a,\*</sup>, Vincenzo Russo<sup>a,b</sup>, Dmitry Yu. Murzin<sup>a</sup>, Tapio Salmi<sup>a</sup>

<sup>a</sup> *Laboratory of Industrial Chemistry and Reaction Engineering (TKR), Johan Gadolin Process Chemistry Centre (PCC), Åbo Akademi University, Åbo-Turku FI-20500, Finland*

<sup>b</sup> *Università di Napoli Federico II, Chemical Sciences, Complesso Universitario Monte S. Angelo, Napoli IT-80126, Italy*

\*Corresponding Author: [mouad.hachhach@abo.fi](mailto:mouad.hachhach@abo.fi)

## Abstract

A comprehensive trickle bed reactor model for simulation purposes was developed in this work. As an industrially relevant case study, the model was applied on selective sugar oxidation to sugar acid on gold catalyst extrudates. The obtained results revealed that the maximum conversion for extrudates is only 25% for a reactor length of 0.5 m and diameter of 1 cm, which is due to slow mass transfer inside the catalyst layer and external mass transfer limitations. Variation of the residence time and the reactor length revealed that a reactor length of 2m minimum is required to achieve complete conversion of the sugar (arabinose) to the corresponding sugar acid (arabinonic acid).

**Keywords:** Trickle Bed Reactor, Oxidation, Extrudates, gProms, Dynamic model.

## 1. Introduction

Trickle bed reactors (TBR) are continuous packed beds where liquid moves in downwards through a catalyst packing whereas gas can move either in a downward or upward direction (Lange et al., 2004). These reactors are solid-gas-liquid contacting devices in which the gas-liquid stream is in continuous contact with the catalyst particles (Hamdan and Abdullah, 2022). Trickle beds are the classical workhorses of chemical industry: in fact these reactors have been used extensively in a wide range of chemical industries from petro-chemical and refinery processes such as hydrodesulfurization and hydrogenation, to the oxidation of biomass and organics in wastewaters, volatile organic compound abatement in air pollution control, enzymatic reactions and bio-methanation (Aleksandrov et al., 2022; Dudukovic et al., 1999; Jønson et al., 2022), in addition to the oxidation of sugar (glucose) (Diallo et al., 2017; Tsukamoto et al., 1982). Trickle beds offer several advantages: simplicity in operation, flexibility for a variety of materials, high catalyst loading per unit volume, and low capital and operating costs as well as lower energy consumption and higher gas conversion rate and higher productivities compared to tank and bubble column reactors (Klasson et al., 1992; Li et al., 2013; Ranade et al., 2011). In contrast to slurry reactors, trickle beds enable safe continuous operation, whereas slurry reactors are in practise limited to batch and semibatch operation.

In this work, the goal was to develop a general and comprehensive trickle bed reactor model, and use it to study the catalytic oxidation of arabinose on gold based catalyst extrudates. The model can be used as framework for the simulation of various reaction

systems as well as a tool to assist process design, scale-up and life cycle assessment for arbitrary reaction systems carried out in trickle beds.

## 2. Materials and methods

The trickle bed reactor model was developed to be as general as possible. The concentration profiles inside the reactor tube and within the catalyst particles were computed by solving simultaneously the mass and energy balances of the components including rate equations (kinetic laws) as well as semi-empirical correlations for pressure drop and fluid dynamics. It was assumed that all the reactions take place inside the catalyst particles only, therefore the rate term is present only in the mass balance of the solid phase, and the active metal of the catalyst is assumed to be distributed evenly inside the catalyst particles. The kinetics of sugar oxidation from batch experiments (Herrero Manzano et al., 2021) was used as an input for the model and realistic values of physical properties were selected. The most essential parameters are summarized in Table 1.

Table 1. List of parameters that characterize the simulation

Parameters	Value	Unit	Parameters	Value	Unit	Parameters	Value	Unit
$\rho_{cat}$ : Catalyst density	400	Kg/m <sup>3</sup>	Reactor radius	5	cm	Reactor length	0.5	m
$\rho_L$ : Density of liquid phase	950	Kg/m <sup>3</sup>	CO <sub>2</sub> S :saturation concentration of O <sub>2</sub>	0.103	mol/m <sup>3</sup>	Effective diffusivity ( $D_{eff}$ ) of O <sub>2</sub>	2.403 10 <sup>-9</sup>	m <sup>2</sup> /s
m(O <sub>2</sub> ): Molar mass	32	g/mol	$Pe_G$ : Péclet number of gas phase	100		$k$ : rate constant	53.99	m <sup>6</sup> /(s mol.kgcat)
$\sigma_{H_2O}$ : Water surface tension	72.8 10 <sup>-3</sup>	N/m	$Pe_L$ : Péclet number of liquid phase	10		Effective diffusivity ( $D_{eff}$ ) of Arabinose	2.474 10 <sup>-9</sup>	m <sup>2</sup> /s
$\eta_{H_2O}$ : Water viscosity	0.2 10 <sup>-3</sup>	Pa s	Effective diffusivity ( $D_{eff}$ ) of arabinic acid	9.851 10 <sup>-9</sup>	m <sup>2</sup> /s	Arabinose diffusivity in water ( $D_L$ )	4 10 <sup>-8</sup>	m <sup>2</sup> /s
$a^*$ :limiting activity factor	0.47282							

## 2.1. Mass balances

The mass balance of the gas-phase components displayed in equation (1) affirm that the accumulation term is the sum of gas-liquid mass transfer with a negative sign, axial and radial dispersion alongside the convection terms (Russo et al., 2015).

$$\frac{\partial C_{i,g}(t,z,r)}{\partial t} = \frac{-k_{i,L}a}{\varepsilon_G} (C_{i,L}^* - C_{i,L}) - \frac{\partial(u_G(z,r)C_{i,g}(t,z,r))}{\varepsilon_G \partial z} + D_{z,g}(z,r) \frac{\partial^2 C_{i,g}(t,z,r)}{\partial z^2} + D_{r,g}(z,r) \left( \frac{\partial^2 C_{i,g}(t,z,r)}{\partial r^2} + \frac{1}{r} \frac{\partial C_{i,g}(t,z,r)}{\partial r} \right) \quad (1)$$

The liquid-phase mass balance equation is shown in equation (2): the accumulation is the sum of gas-liquid mass transfer term with a positive sign in opposition to the gas phase, in addition to the convection and axial and radial dispersion terms, and finally the mass transfer flux from/to the solid catalyst phase as represented in the last term.

$$\frac{\partial C_{i,L}(t,z,r)}{\partial t} = \frac{k_{i,L}a}{\varepsilon_L} (C_{i,L}^* - C_{i,L}) - \frac{uL}{\varepsilon_L} \frac{\partial(C_{i,L}(t,z,r))}{\partial z} + D_{z,L} \frac{\partial^2 C_{i,L}(t,z,r)}{\partial z^2} + D_{r,L} \left( \frac{\partial^2 C_{i,L}(t,z,r)}{\partial r^2} + \frac{1}{r} \frac{\partial C_{i,L}(t,z,r)}{\partial r} \right) - \frac{s D_{eff,i}}{\varepsilon_L R_p} \frac{\partial C_{i,s}(t,z,r,r_p)}{\partial r_p} \Big|_{r_p} = R_p \quad (2)$$

For the solid catalyst phase, the concentration field depends on the effective diffusivity, catalyst shape factor and the holdups of the liquid and solid phases. Moreover, the reaction term is included comprising the catalyst density, ratio of holdups, stoichiometric coefficients and reaction rate, as shown in equation (3) below.

$$\frac{\partial C_{i,s}(t,z,r,r_p)}{\partial t} = \frac{D_{eff,i}}{\varepsilon_p} \left( \frac{\partial^2 C_{i,s}(t,z,r,r_p)}{\partial r_p^2} + \frac{s}{r_p} \frac{\partial C_{i,s}(t,z,r,r_p)}{\partial r_p} \right) + \rho_{cat} \sum (\vartheta_{i,j} r_j(t,z,r,r_p) \frac{\varepsilon_L}{\varepsilon_p}) \quad (3)$$

## 2.2. Physical properties and reactor parameters

The liquid viscosity, which is required to calculate the Reynolds number, was included in the model, the correlation adopted in equation (4) was based on arabinose hydrogenation experiments (Sifontes Herrera et al., 2016) and taken into account the dependence of the viscosity on both temperature and arabinose concentration ( $x_a$ ).

$$\eta_{fluid} = 10^{-3} \exp \left( \frac{1.54x_a}{T_{in,L}-273} + 3.81 \cdot 10^{-4} x_a^2 - 1.10 \cdot 10^{-2} x_a - 2.85 + \frac{194}{T_{in,L}-273} - \frac{3890}{(T_{in,L}-273)^2} \right) \quad (4)$$

The gas-liquid mass transfer coefficient is taken into account in the mass balance as shown previously in equations (1) and (2). The equation 6 adopted from Ranade et al. (2011) was used in this study.

$$k_L a(z,r) = \gamma 0.45 \left( \frac{DL}{dK^2} \right) (X_G(z,r))^{0.65} ReL^{1.04} WeL^{0.26} ScL^{0.65} \left( \frac{\alpha sp dK}{1-\varepsilon_p} \right)^{0.325} \quad (5)$$

The axial dispersion coefficient was calculated by using the Péclet number as shown in equations (7) and (8), while the radial dispersion coefficient was taken equal to one third of the axial one as shown in the equations (9) and (10) as has been reported previously

(Russo et al., 2015). According to our previous work by (Durante et al., 2014; Kilpiö, 2013) the order of magnitude of the liquid phase Péclet number was 2-10, while for the gas phase the Péclet number can go up to 100. For numerical simulations, the Péclet number for liquid-phase was taken 10 and for the gas phase 100.

$$Dz = uL \cdot \frac{L}{Pe} \quad (6) \qquad DzG(z, r) = uG(z, r) \cdot \frac{L}{PeG} \quad (7)$$

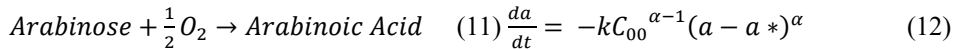
$$Dr = \frac{Dz}{3} \quad (8) \qquad DrG = \frac{DzG(z, r)}{3} \quad (9)$$

The shape factor is evaluated by equation (10) proposed by Salmi et al., (2019), where  $V_p$  and  $A_p$  are respectively the volume and surface area of the catalyst particle.

$$s = 1 + R_p \frac{A_p}{V_p} \quad (10)$$

### 2.3. Kinetics of the case study

The case study is the selective oxidation of arabinose to arabinic acid on gold catalyst (Herrero Manzano et al., 2021; Kusema et al., 2012, 2010). The kinetics has been described by classical Langmuir-Hinshelwood rate expression as shown in the equations (12) and (13). The kinetic expression and parameters values were adapted from the work of Herrero Manzano et al. (2021).



$$r(t, z, r, r_p) = a^2 \frac{K(t, z, r, r_p)C_{O_2, s}(t, z, r, r_p)C_{Ar, s}(t, z, r, r_p)}{1 + K_{Ar, s}C_{Ar, s}(t, z, r, r_p) + K_{O_2, s}C_{O_2, s}(t, z, r, r_p)} \quad (13)$$

The model equations were implemented in the software gPROMS, which was chosen due to its high convergence speed and its user-friendliness: In fact, in this case study the convergence time was in order of 5-10 minutes. The partial differential equations (PDEs) describing the mass balances (1)-(3) were converted to ordinary differential equations (ODEs) by discretization of the radial and axial coordinates with finite differences. A stiff differential equation algorithm was used for solving the ODEs, the initial value problem.

## 3. Results and discussion

The obtained results are summarized in Figure 1. The simulations reveal that the maximum conversion for the catalyst extrudates at the reactor output is only 25% for a reactor length of 0.5 m and diameter of 1 cm, which can be due to slow mass transfer inside the catalyst particles but also the existence of external mass transfer limitations, which can be surmounted by using smaller particle sizes, slurry reactor technology or structured catalysts, such as monoliths or solid foams. The results show that the radial position is not affecting the sugar conversion as illustrated by Figure 1, where the conversion is practically the same for various radial positions across the reactor tube, in contrast to the axial position, where the sugar conversion increases as the length of the reactor increases.

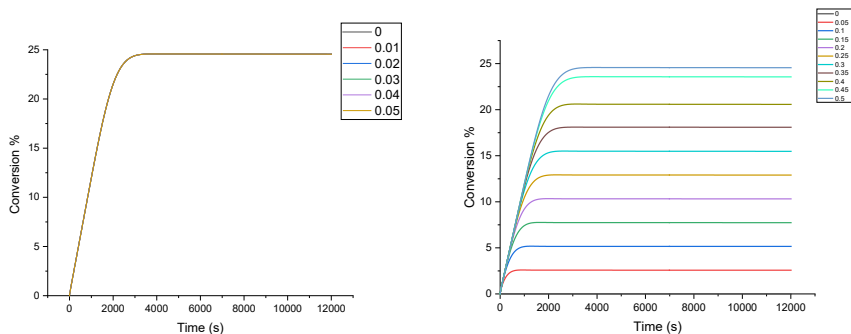


Figure 1: Conversion rate at various axial (left) and radial (right) positions

In order to increase the conversion, the residence time and the reactor length were varied and the conversion at the reactor outlet was computed as displayed in Figure 2, which confirms that a reactor length of 2m minimum is required to achieve complete conversion of the sugar (arabinose) to the corresponding sugar acid (arabinonic acid).

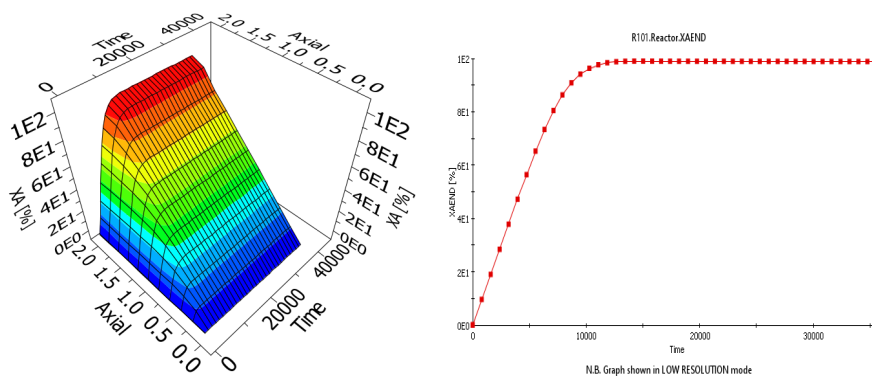


Figure 2: Effect of Simultaneous increase of reaction time and length of the reactor on the Conversion Rate (left). Conversion rate at the outlet of the scaled-up reactor (2m) (right)

The obtained results demonstrate quite interesting modelling stability that might pave the way to the investigation of the scale-up of various reaction systems, although more accurate simulations would require further kinetic and flow experiments using trickle bed reactors.

## Acknowledgment

This work is partially financed by the Academy of Finland, the Academy Professor grant 319002 (T. Salmi). The financial support obtained from PCC, Finnish Cultural Foundation (SKR), Walter Ahlstrom Foundation, and from The Society of Swedish Literature in Finland (SLS) is gratefully acknowledged (M. Hachhach).

## References

- Aleksandrov, P. V., Reshetnikov, S.I., Bukhtiyarova, G.A., Noskov, A.S., 2022. Deep hydrodesulfurization of gas oils with high sulfur content: Experiment and kinetic modeling. Chem. Eng. J. 446, 137059. <https://doi.org/https://doi.org/10.1016/j.cej.2022.137059>
- Diallo, M.M., Derrien, E., Marion, P., Perret, N., Pinel, C., Besson, M., 2017. Base-free aerobic oxidation of

- glucose to glucaric acid over Pt-Au/ZrO<sub>2</sub> catalysts in batch and trickle-bed reactors.
- Dudukovic, M.P., Larachi, F., Mills, P.L., 1999. Multiphase reactors - revisited. *Chem. Eng. Sci.* 54, 1975–1995. [https://doi.org/10.1016/S0009-2509\(98\)00367-4](https://doi.org/10.1016/S0009-2509(98)00367-4)
- Durante, D., Kilpiö, T., Suominen, P., Herrera, V.S., Wärnå, J., Canu, P., Salmi, T., 2014. Modeling and simulation of a small-scale trickle bed reactor for sugar hydrogenation. *Comput. Chem. Eng.* 66, 22–35. <https://doi.org/10.1016/j.compchemeng.2014.02.025>
- Hamdan, H.Y., Abdullah, G.H., 2022. A novel trickle bed electrochemical reactor design for efficient hydrogen peroxide production. *Chem. Eng. Process. - Process Intensif.* 181, 109123. <https://doi.org/https://doi.org/10.1016/j.cep.2022.109123>
- Herrero Manzano, M., Eranen, K., Freitas Aguilera, A., Wärnå, J., Franz, S., Peurla, M., Serna, J.G., Murzin, D., Salmi, T., 2021. Interaction of intrinsic kinetics, catalyst durability and internal mass transfer in the oxidation of sugar mixtures on gold nanoparticle extrudates. *Ind. Eng. Chem. Res.* 60, 6483–6500. <https://doi.org/10.1021/acs.iecr.0c05305>
- Jønson, B.D., Tsaekos, P., Tahir Ashraf, M., Jeppesen, M., Ejbye Schmidt, J., Bastidas-Oyanedel, J.-R., 2022. Pilot-scale study of biomethanation in biological trickle bed reactors converting impure CO<sub>2</sub> from a Full-scale biogas plant. *Bioresour. Technol.* 365, 128160. <https://doi.org/https://doi.org/10.1016/j.biortech.2022.128160>
- Kilpiö, T., 2013. Mathematical modeling of laboratory scale three-phase fixed bed reactors. Doctoral thesis, Åbo Akademi University.
- Klasson, K.T., Ackerson, M.D., Clausen, E.C., Gaddy, J.L., 1992. Bioconversion of synthesis gas into liquid or gaseous fuels. *Enzyme Microb. Technol.* 14, 602–608. [https://doi.org/10.1016/0141-0229\(92\)90033-K](https://doi.org/10.1016/0141-0229(92)90033-K)
- Kusema, B.T., Campo, B.C., Mäki-Arvela, P., Salmi, T., Murzin, D.Y., 2010. Selective catalytic oxidation of arabinose - A comparison of gold and palladium catalysts. *Appl. Catal. A Gen.* 386, 101–108. <https://doi.org/10.1016/j.apcata.2010.07.037>
- Kusema, B.T., Mikkola, J.-P.P., Murzin, D.Y., 2012. Kinetics of L-arabinose oxidation over supported gold catalysts with in situ catalyst electrical potential measurements. *Catal. Sci. Technol.* 2, 423–431. <https://doi.org/10.1039/c1cy00365h>
- Lange, R., Schubert, M., Dietrich, W., Grünewald, M., 2004. Unsteady-state operation of trickle-bed reactors. *Chem. Eng. Sci.* 59, 5355–5361. <https://doi.org/10.1016/J.CES.2004.09.007>
- Li, D., Li, Z., Li, W., Liu, Q., Feng, Z., Fan, Z., 2013. Hydrotreating of low temperature coal tar to produce clean liquid fuels. *J. Anal. Appl. Pyrolysis* 100, 245–252. <https://doi.org/10.1016/J.JAAP.2013.01.007>
- Ranade, V., Chaudhari, R., Gunjal, P.R., 2011. Trickle Bed Reactors, Trickle Bed Reactors. Elsevier. <https://doi.org/10.1016/C2009-0-16553-4>
- Russo, V., Kilpiö, T., Di Serio, M., Tesser, R., Santacesaria, E., Murzin, D.Y., Salmi, T., 2015. Dynamic non-isothermal trickle bed reactor with both internal diffusion and heat conduction: Sugar hydrogenation as a case study. *Chem. Eng. Res. Des.* 102, 171–185. <https://doi.org/10.1016/j.chemd.2015.06.011>
- Salmi, T., Mikkola, J.-P., Wärnå, J., 2019. Chemical Reaction Engineering and Reactor Technology, Chemical Reaction Engineering and Reactor Technology. Chapman and Hall/CRC. <https://doi.org/10.1201/9781315200118>
- Sifontes Herrera, V.A., Rivero Mendoza, D.E., Leino, A.R., Mikkola, J.P., Zolotukhin, A., Eränen, K., Salmi, T., 2016. Sugar hydrogenation in continuous reactors: From catalyst particles towards structured catalysts. *Chem. Eng. Process. - Process Intensif.* 109, 1–10. <https://doi.org/10.1016/J.CEP.2016.07.007>
- Tsakamoto, T., Morita, S., Okada, J., 1982. Oxidation of Glucose on Immobilized Glucose Oxidase in a Trickle-Bed Reactor : Effect of Liquid-Solid Contacting Efficiency on the Global Rate of Reaction. *Chem. Pharm. Bull. (Tokyo)*. 30, 1539–1549. <https://doi.org/10.1248/cpb.30.1539>

# Numerical simulation of flow in porous media using the SPH method and Ergun force

Carlos E. Alvarado-Rodriguez,<sup>a,b</sup> Lamberto Diaz-Damacillo,<sup>c</sup> Eric Plaza,<sup>d,e</sup>

<sup>a</sup>*Departamento de Ingeniería Química, DCNE, Universidad de Guanajuato, Noria Alta S/N, Guanajuato, 3605, México.*

<sup>b</sup>*Consejo Nacional de Ciencia y Tecnología, Avenida Insurgentes Sur 1582, Crédito Constructor, Ciudad de México, 03940, México.*

<sup>c</sup>*Departamento de Ciencias Básicas, Universidad Autónoma Metropolitana Azcapotzalco (UAM-A), Av. San Pablo 180, Ciudad de México 02200, Mexico.*

<sup>d</sup>*ESPOL Polytechnic University, Escuela Superior Politécnica del Litoral, ESPOL, Departamento de Física, Laboratorio de Fluidos, km. 30.5, Vía Perimetral Guayaquil, 09-01-5863, Ecuador.*

<sup>e</sup>*Laboratorio de Física de Fluidos y Plasmas, Instituto Venezolano de Investigaciones Científicas (IVIC), Altos de Pipe, 21827, Venezuela.*

## Abstract

This work presents the implementation of the Ergun force term on the DualSPHysics code, to simulate flow in porous media using a representative element volume (REV). In the study of microscopic seepage and fluid-solid interactions, a typical approach is a pore-scale simulation due to its precision; however, it involves a high computation cost, making it more suitable to adopt the REV scale. Here are presented nine simulations comparing the velocity, at the outlet of the system, obtained by a pore-scale geometry and the Ergun force term and REV simulations using SPH. The results have a good fit between the pore-scale simulations and the Ergun force term simulations, where the lowest and highest errors are  $2.92 \times 10^{-3}$  and  $1.43 \times 10^{-2}$ , respectively. These results present novel and alternative simulations in the study of the flow in porous media using the SPH method where the pore-scale geometry is usually used.

**Keywords:** Porous media, SPH, Ergun force, CFD.

## 1. Introduction

Smoothed particle hydrodynamics (SPH) is a fully Lagrangian particle method introduced by Lucy (1977). Since then, due to its ease of implementation and flexibility, the method has spread into numerous branches of computational physics and engineering. SPH has a Lagrangian nature making enabling simulations of complex systems that would be very difficult to carry out with the use of more conventional grid-based schemes. In SPH, the discretization is carried out by computing an interpolating kernel function that uses the properties of the surrounding particles over the studied one to compute its physical properties. Due to the characteristics of discretization in SPH, the dynamic of the fluid volume element is controlled by the local variations of pressure and density, making it possible to simulate the direct inclusion of boundaries, treatment of the free surface, and recently the treatment of multiphase flows. SPH can be used in several branches of science, including the flow in porous media (Tartakovsky et al., 2007; Kunz et al., 2015; Klapp et al., 2022), which presents a great interest to the industry due to its connection with crude oil extraction, the dynamics of flow in aquifers, agricultural irrigation



processes, and into materials that can be used as filters. In general, the dynamic of fluid is controlled by the principles of conservation of momentum, energy, and mass. These differential equations do not allow analytical solutions for this kind of problem, which opens the way for the introduction of numerical simulations as a tool for prediction. In the case of flow in porous media, Darcy's Law has been widely used, when the Reynolds number is small. This equation, obtained through experimental measurements, approximates the behavior of the fluid in the porous media, using empirical relationships that associate the permeability of the medium and geometric variables. The purpose of this work is to extend the empirical relation of Darcy and provide a more accurate representation of flow in porous media. To do this, the Ergun term was added to the movement equation, which includes a friction expression and allows for the quantification of the pressure drop and energy required for pumping the flow, (Ergun, 1952). A novel SPH model was developed to simulate flow in porous media, making use of the Ergun force term in the momentum equation. The model avoids the need to depict the internal shapes of the porous media in the simulation and instead uses parameters that describe the media's specifications. This enables the simulation of large length scales and provides a better understanding of the internal flow in porous media. The work is structured as follows: Section 2 presents the complete SPH numerical model. Section 3 validates the velocity profile in a Poiseuille flow. Section 4 compares numerical cases to demonstrate the effectiveness of using the Ergun force term in the momentum equation. Finally, the conclusions of the work are presented in Section 5.

## 2. The SPH Numerical Model

Smoothed Particle Hydrodynamics (SPH) is a meshless Lagrangian method that recently has been applied to simulate a wide range of applications within the field of Computational Fluid Dynamics.

For viscous incompressible flows, the governing equation is the Navier–Stokes expressed in SPH form as

$$\frac{d\mathbf{v}_a}{dt} = - \sum_b m_b \left( \frac{p_a + p_b}{\rho_a \rho_b} + \Gamma \right) \nabla_a W_{ab} + \mathbf{g} \quad (1)$$

where  $\rho$  is the density,  $p$  pressure,  $\mathbf{v}$  is the velocity vector,  $W$  is the kernel function,  $\mathbf{g}$  is the gravity acceleration and  $\Gamma$  is the viscosity dissipative term. In a standard SPH formulation, where the pressure is given as a function of the density, local variations of the pressure gradient may induce local density fluctuations in the flow. Therefore, flow is modelled by an artificial fluid that is slightly compressible. The particle mass remains constant and only is associated with density fluctuations. Such density fluctuations are calculated by solving the continuity equation expressed in Eq. (2). The dynamical pressure  $p$ , is calculated using the relation expressed in Eq. (3).

$$\frac{d\rho_a}{dt} = \sum_{b=1}^N m_b (\mathbf{v}_a - \mathbf{v}_b) \cdot \nabla_a W_{ab}, \quad (2)$$

$$p = p_0 \left[ \left( \frac{\rho}{\rho_0} \right)^\gamma - 1 \right] \quad (3)$$

where  $\gamma=7$ ,  $p_0 = c_0^2 \rho_0 / \gamma$ ,  $\rho_0$  is a reference density, and  $c_0$  is the sound speed at the reference density.

The viscous stress tensor  $\Gamma$  is described using the laminar viscosity model and the Sub Particle Scale (SPS) model, proposed by Darlympe & Rogers, (2006) and reported in the DualSPHysics code as

$$\Gamma_{ab} = \sum_b m_b \left( \frac{4v_0 \mathbf{r}_{ab} \cdot \nabla_a W_{ab}}{(\rho_a + \rho_b)(r_{ab}^2 + \eta^2)} \right) \mathbf{v}_{ab} + \sum_b m_b \left( \frac{\tilde{\tau}_{ab}^j}{\rho_b^2} + \frac{\tilde{\tau}_{ab}^i}{\rho_a^2} \right) \nabla_a W_{ab}, \quad (4)$$

where  $v_0$  is la kinematic viscosity and  $\tilde{\tau}_{ab}^j$  is the shear stress.

In the field of fluid dynamics in porous media, numerical simulations are a crucial tool for predicting fluid flow behavior. When conducting these simulations, two scales are typically considered: the pore scale and the Representative Elementary Volume (REV) scale. Pore-scale simulations provide more accurate results and have an advantage in the study of microscopic seepage and fluid-solid interactions. However, they are computationally expensive and not feasible for large systems, such as those of an oil reservoir. In these cases, REV-scale simulations are the more practical option, despite their limitations in terms of accuracy.

Since the geometry of the porous media is not used in the REV-scale simulations, the porous media are modeled by adding the resistance term in the momentum equation (Eq. 5), and the interaction of the solid matrix with the fluid is out of consideration, which reduces the computation time.

$$\frac{d\mathbf{v}_a}{dt} = -\sum_b m_b \left( \frac{p_a + p_b}{\rho_a \rho_b} + \Gamma \right) \nabla_a W_{ab} + \mathbf{g} + \frac{\mathbf{F}_a}{m_a}. \quad (5)$$

The Ergun formula is widely used to calculate the additional resistance term due to its simple form, which is

$$\mathbf{F} = -\frac{\epsilon v}{K} \mathbf{v} - \frac{\epsilon F_e}{\sqrt{K}} |\mathbf{v}| \mathbf{v}, \quad (6)$$

where  $\mathbf{v}$  is the velocity vector,  $v$  is the viscosity of the fluid;  $K$  is the permeability of the porous medium,  $\epsilon$  is the porosity and  $F_e$  is a structure-function, the parameters shown in Eq.(6) are reported in detail by Lai et al., (2020).

In this work, Eq. (6) is applied in the last term of Eq. (5) by particle, for the cases of a REV scale. The implementation was performed using version 5 of the DualSPHysics code (Dominguez et al., 2022).

### 3. Validation test.

The SPH model presented in section 2 is validated with the simulation of the Poiseuille flow which has an analytical solution. The SPH numerical results are compared with the analytical solution of the system. For this case was considered the geometry shown in Figure 1, where the distance between plates is 1 mm and the length is 5 mm. The properties of the fluid are according to the liquid water,  $\rho = 100\text{kg/m}^3$ ,  $\nu = 1 \times 10^{-6} \text{ Pa*s}$ , and the particles have an acceleration of  $1 \times 10^{-4} \text{ m/s}^2$  in the right direction. This case was simulated with different resolutions changing the initial distance between particles  $dp = 0.1, 0.05, 0.025, \text{ and } 0.01 \text{ mm}$ . The velocity profiles obtained from the numerical results using the model reported in section 2 and their comparison with the analytical solution are shown in Figure 1. The errors (RMSE) for each simulation are reported in Table 1 where is clear that the error decrease when the resolution is increased.

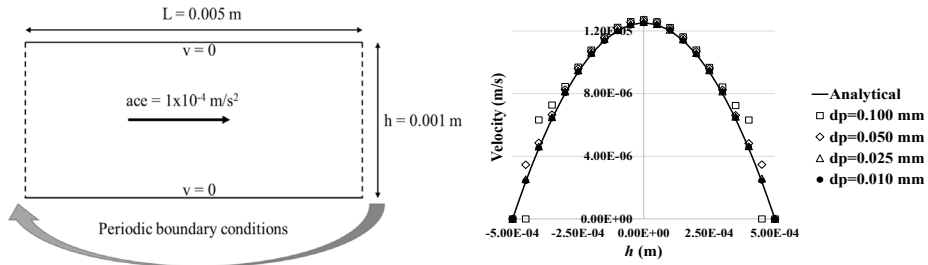


Figure 1. The left figure represents the dimensions and boundary conditions for the Poiseuille flow simulation, and the right one the velocity profiles for the different initial distances between particles ( $dp$ ).

Table 1. Root mean square error between the numerical results and the analytical solution in the validation test cases.

$Dp$ (mm)	0.1	0.05	0.025	0.01
RMSE	$1.01403 \times 10^{-6}$	$4.04818 \times 10^{-7}$	$1.10592 \times 10^{-7}$	$5.28235 \times 10^{-8}$

### 4. Comparative cases.

The pore-scale simulations were conducted in 2D with porosities of 0.55, 0.44, and 0.3, represented by three different arrays of circles named square, diagonal, and random (as shown in Figure 2). A total of 9 study cases were carried out. The porous media were inserted in the center of a rectangular channel with dimensions of 10 cm x 5 cm. The inlet boundary was located at the left side of the channel and was set to a constant velocity profile (as shown in Figure 2). The outlet boundary was located at the right side of the channel, where the velocity value of the particles was computed (as shown in Figure 2). The top and bottom boundaries were no-slip boundary conditions, using the dynamics particle method described in (Crespo et al., 2007). The fluid was characterized with the same parameters used in the validation case, with an initial particle distance of  $dp = 0.001$  mm for all cases reported in this section. In the simulations using the Ergun Force, the geometry of the porous media was changed to a rectangular section (as shown in Figure 2) and the Ergun force (Eq. 6) was imposed in the momentum equation (Eq. 4), while maintaining the same parameters and channel dimensions as the pore-scale simulations.

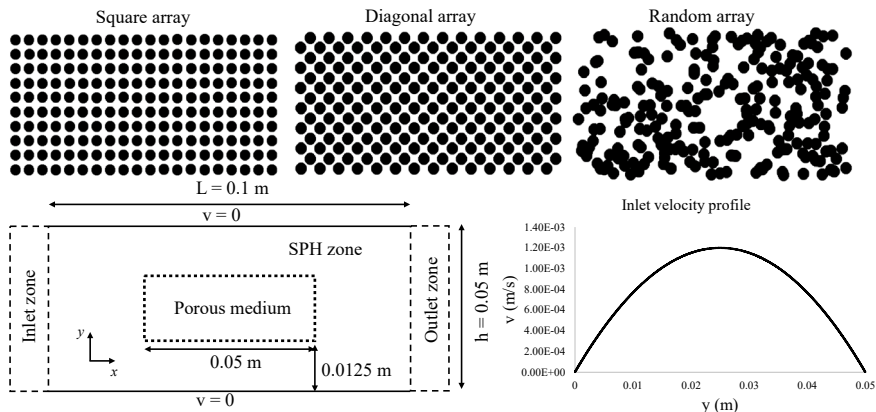


Figure 2. The three different left arrays, shown an example of the PM used for each porosity value. Right frame: at top is shown the geometry used for all cases, and the bottom is the inlet velocity profile.

The velocity profiles at the outlet obtained in the pore-scale and REV scale were compared to prove the accuracy of the Ergun force to represent the flow in different porous media. The velocity fields obtained from both scale simulations are shown in Figure 3, only for the case with  $\epsilon = 0.55$  and the random array, both results are very similar in qualitative analysis. The quantitative fit was obtained using the velocity profiles at the outlet shown in Figure 4 for all cases. Here, the root-mean-square error (RMSE) is used as the error of the models. RMSE is a measure of accuracy, typically used to compare forecasting errors of different models for a particular dataset and not between datasets, as it is scale dependent. The results presented in Figure 4 have very similar profiles, between the simulations carried using the Ergun force term and those using the porous geometry for different porosities and geometries. The error between the profiles shown in Figure 4 are reported in Table 3. The errors reported in Table 3 are two order of magnitude less than the maximum velocity shown in Figure 4 for all cases.

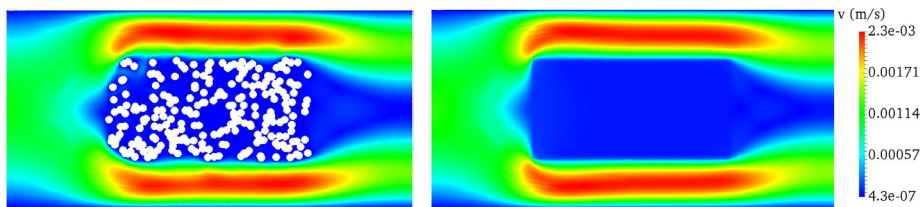


Figure 3. Comparison of the velocity field between the pore-scale and REV simulations: left frame pore-scale, right frame Ergun force.

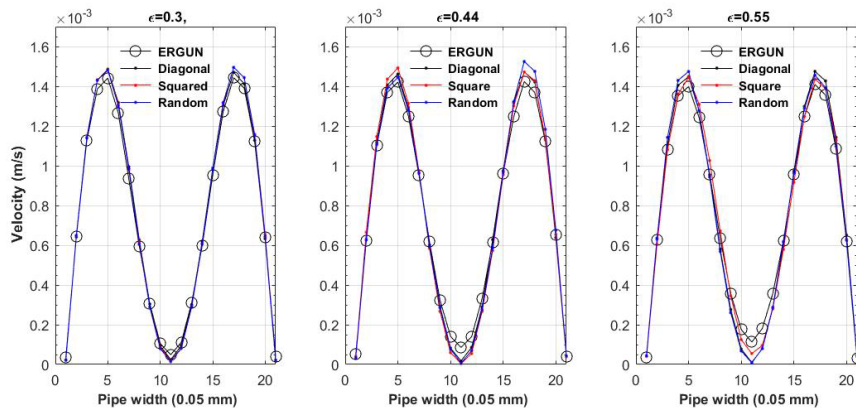


Figure 4. Velocity profiles comparison for all cases reported in the section 4.

Table 2. Root mean square error between the numerical results and the Ergun force term simulations.

	$\epsilon = 0.33$	$\epsilon = 0.44$	$\epsilon = 0.55$
RMSE (Squared)	$2.84 \times 10^{-5}$	$3.82 \times 10^{-5}$	$4.36 \times 10^{-5}$
RMSE (Diagonal)	$3.20 \times 10^{-5}$	$5.23 \times 10^{-5}$	$5.77 \times 10^{-5}$
RMSE(Random)	$3.19 \times 10^{-5}$	$5.12 \times 10^{-5}$	$5.91 \times 10^{-5}$

## 5. Conclusions.

In this work was implemented the Ergun force term in the momentum equation using the Smoothed Particle Hydrodynamics (SPH) method. The implementation was validated through a Poiseuille flow simulation and a comparison of the numerical results with the

analytical solution. Additionally, a convergence test was conducted and the error between the numerical and analytical data decreased as the resolution increased, reaching a minimum error of  $5.28235 \times 10^{-8}$  in the simulation with a particle distance of 0.01 mm and 501501 total particles. The results demonstrate the successful implementation of the Ergun force using the DualSPHysics code. Comparing the pore-scale and REV simulations, the velocity profiles at the outlet were similar, with a maximum error of  $5.91 \times 10^{-5}$  in the case of a porosity of 0.55 and the random array and a minimum error of  $2.84 \times 10^{-5}$  in the case of a porosity of 0.3 and the square array. The square geometry resulted in better emulation of the Representative Elementary Volume (REV) area, with a lower error in all cases, due to the similarity of the squared area to the area where the Ergun force is applied. The error increased with the porosity, but further simulations with a wider range of porosities are necessary to confirm the proper application of the Ergun force in porous media simulations.

## References

- L. Lucy, 1977, A numerical approach to testing the fission hypothesis, *Journal Astronomical*, 82:1013–1024.
- A. M. Tartakovsky, P. Meakin, T. D. Scheibe, B. D. Wood, 2007, A smoothed particle hydrodynamics model for reactive transport and mineral precipitation in porous and fractured porous media. *Water Resour. Res.* 43, W05437 (18 pages).
- J. Klapp, L. D. G. Sigalotti, C. E. Alvarado-Rodríguez, 2022, Approximately consistent SPH simulations of the anisotropic dispersion of a contaminant plume. *Comp. Part. Mech.* 9, 987–1002.
- P. Kunz, I. M. Zarkos, N. K. Karadimitriou, M. Huber, U. Nieken, S. M. Hassannizadeh, 2015, Study of multiphase flow in porous media: comparison of SPH simulations with Micro-model experiments, *Transport in Porous media*, Springer.
- S. Ergun, 1952, Fluid flow through packed columns. *Chemical Engineering Progress*, 48, 89–92.
- A. M. Tartakovsky, A. Panchenko, 2016, Pairwise Force Smoothed Particle Hydrodynamics model for multiphase flow: Surface tension and contact line dynamics, *Journal of Computational Physics*, 305, 1119–1146.
- R. A. Dalrymple, B. Rogers, 2006, Numerical modeling of water waves with the SPH method. *Coastal Engineering*, 53:141–147.
- A. J. C. Crespo, M. Gómez-Gesteira, R. A. Dalrymple, 2007a, Boundary conditions generated by dynamic particles in SPH methods, *Computers, Materials & Continua*, 5, 173–184.
- J. M. Domínguez, G. Fourtakas, C. Altomare, R. B. Canelas, A. Tafuni, O. García-Feal, I. Martínez-Estévez, A. Mokos, R. Vacondio, A. J. C. Crespo, B. D. Rogers, P. K. Stansby, M. Gómez-Gesteira, 2022, DualSPHysics: from fluid dynamics to multiphysics problems. *Computational Particle Mechanics*, 9(5), 867–895.
- T. Lai, X. Liu, S. Xue, J. Xu, M. He, Y. Zhang, 2020, Extension of Ergun equation for the calculation of the flow resistance in porous media with higher porosity and open-celled structure, *Applied Thermal Engineering*, 173, 115262.

# Impact of structural change in electricity and hydrogen end use on low-carbon transition of an energy system: A case study of China

Nuobei Zhang,<sup>a</sup> Pei Liu,<sup>\*</sup> Zheng Li

<sup>a</sup>*State Key Lab of Power Systems, Department of Energy and Power Engineering, Tsinghua-BP Clean Energy Center, Tsinghua University, Beijing 100084, China*  
*liu\_pei@tsinghua.edu.cn*

## Abstract

Driven by global targets to reduce greenhouse gas emissions, energy systems are expected to undergo fundamental changes, and the proportion of electricity and hydrogen may greatly increase. Scenarios analysis based on optimal planning is beneficial to exploring how this growth affects the structure of an energy system, whilst at the same time reducing costs, better fulfilling the carbon emission target and ensuring energy security. This study sets two groups of different scenarios based on share of electricity and hydrogen in end use respectively, and uses China Regional Energy Supply System Optimization Model (CRESOM) as an analysis tool. China is chosen as a case study due to the large scale, complex structure, fast growth rate, strict yet changing emissions constraints of its energy system. Results show that with an increase in electrification rate, cumulative carbon dioxide emissions over a certain planning horizon could reduce, at a priced of increased unit reduction cost. By increasing the share of electricity and hydrogen in the end use from 71% to 80% in 2060, the unit carbon reduction cost rises by 17%. Increasing share of hydrogen could move the carbon emissions peak time by approximately five years.

**Keywords:** Scenarios analysis; Low-carbon transition; China; Electricity; Hydrogen

## 1. Introduction

Low carbon development has become a worldwide consensus, but the task is arduous. Fossil energy is still the main part of the global energy consumption. According to BP's scenario forecast (BP, 2020), the proportion of fossil energy will drop from 84% (2018) to 21.7% (2050) to achieve near zero emissions. The global energy system is expected to undergo a profound low-carbon transition inevitably. In the next 30 years, with the promotion of decarbonization of the energy system, the end use of electricity and hydrogen could grow significantly (BP, 2020). Therefore, it is necessary to conduct research to get guidance such as determining the transition objectives and reducing the total cost in the transition process.

At present, there are many studies on transition of the energy system. Dai et al. (2016) built two scenarios from the perspective of renewable development, and assessed the impact of large-scale renewable development on economy and environment by 2050. Guo et al. (2017) established the LoMLoG model, gave the most likely scenario in the future, proved the importance of the goal of clean energy for reducing carbon emissions in the power sector and promoting the use of renewable energy. Guo et al. (2021) conducted scenario analysis on building energy consumption, taking China as a case, and pointed out that the carbon peak time is expected to be 2020-2035. Duan et al. (2021) compared

the results of various models, pointed out that China will reduce carbon emissions by 90% with the target of 1.5 °C under the scenario of policies, explained the importance of negative emission technology in the future. Zhang et al. (2022) set different scenarios, analyzed the contribution of emission reduction measures in different periods, and analyzed the uncertainty of key parameters.

However, previous studies paid little attention to the impact of the increase in the end-using electricity proportion and end-using hydrogen proportion in the energy system. Ignoring these two factors may well leads to unclear description of energy substitution process, so it is necessary to consider in the setting of scenarios.

Based on the existing optimization model, this study analyzes the impact of the increase in the end use of electricity and hydrogen in the transition of energy structure by setting different scenarios.

The structure of this paper is as follows. The methodology will be introduced in Section 2. The case study and the results will be introduced in Section 3, and Section 4 is the conclusions.

## 2. Methodology

Scenario analysis based on optimization model is chosen as a method. Through scenario analysis, we can get some valuable transition guidance for countries with fossil energy as the main energy.

### 2.1. The introduction of CRESOM

The model used in this study is China Regional Energy Supply System Optimization Model (CRESOM) developed by Tsinghua-BP Clean Energy Research and Education Center, which is mainly used for the optimal planning of energy supply system under the given low-carbon transformation strategy.

The main inputs of the model include the historical data used to describe the current energy supply and demand and infrastructure construction, prospective data used to describe future economic growth, energy intensity and energy technology cost, scenario data used to describe the emission reduction policies and carbon policies.

The model first calculates the energy demand by category of each terminal energy department in each region from 2016 to 2060 through the terminal energy demand forecasting model, then the final planning scheme is obtained by minimizing the total cost through the power-thermal system optimization model.

The output data includes the forecast of the future and the planning scheme at different stages. The structure of the model is shown in Figure 1.

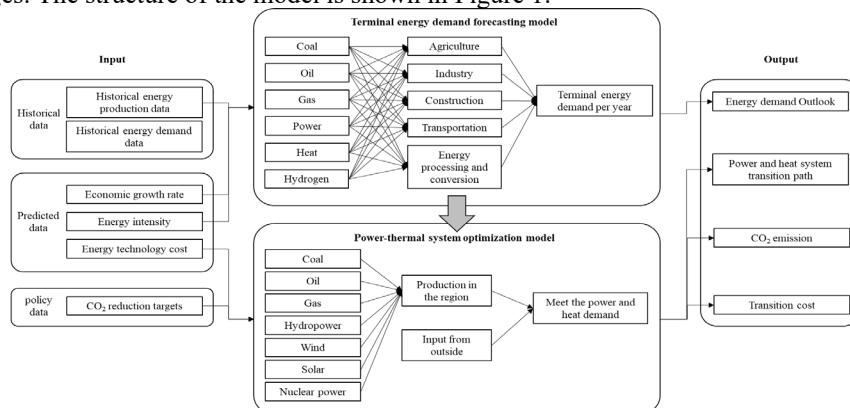


Figure 1. The structure of CRESOM

2.2. Scenario setting

This study focuses on the impact of electricity and hydrogen in the process of energy transition, and constructs two groups of scenarios. One is the high proportion electricity scenario, the other one is high proportion hydrogen scenario. The relevant settings of the scenarios are shown in the tables below.

Table 1. The setting of high proportion electricity scenario

Scenario	EH65	EH68	EH71	EH74	EH77	EH80
End-using electricity proportion in 2060 (%)	55	57.5	60.1	62.6	65.2	67.7
End-using hydrogen proportion in 2060 (%)	10	10.5	10.9	11.4	11.8	12.3
Total proportion (%)	65	68	71	74	77	80

Table 2. The setting of high proportion hydrogen scenario

Scenario	H12	H15	H18	H21	H24	H27	H30
End-using electricity proportion in 2060 (%)	65	62	59	56	53	50	47
End-using hydrogen proportion in 2060 (%)	12	15	18	21	24	27	30
Total proportion (%)							77

3. Case Study

As the largest developing country in the world, China’s energy structure needs to be transformed pressingly, thus its energy transition research is essential. This study takes China as a case to study the impact of the increase of end-using electricity and hydrogen proportion on the energy system through multi scenario calculations, aiming to provide guidance for its transition and provide experience for other countries with fossil energy as the main energy source.

3.1. High proportion electricity scenario results

3.1.1. Total emissions and system costs

As shown in Figure 2A, the carbon emissions show a decline after reaching the peak in 2025, and the promotion of electrification makes the carbon emissions decline faster.

The *costyear* is the sum of the costs of all subsystems each year, as shown in Eq. (1). The *costyear* increases by 26.4% in 2060 (EH65 to EH80), mainly due to the large-scale energy substitution in 2040-2060 (Figure 2B).

$$costyear = costyear_c + costyear_o + costyear_g + costyear_h + costyear_e \tag{1}$$

The *cost<sub>CUTCRC</sub>* (Cumulative Unit Transition Carbon Reduction Cost) is shown in Eq. (2). The molecular part is the total cost of all scenarios compared with the scenario without energy substitution policy. The denominator part is the carbon dioxide reduction of all scenarios compared to the scenario without energy substitution policy. It reflects the cost performance of the system's carbon reduction.

$$cost_{CUTCRC} = \frac{\sum_{2020}^{2060} costyear_{EH} - \sum_{2020}^{2060} costyear_{BAU}}{\sum_{2020}^{2060} CO_{2BAU} - \sum_{2020}^{2060} CO_{2EH}} \tag{2}$$



As the end-using electricity proportion increases, the  $cost_{CUTCRC}$  continues to rise, especially when the total proportion is greater than 74%, where the proportion of end-using electricity in EH80 is 6% higher than that in EH74, but the  $cost_{CUTCRC}$  is 17% higher. Therefore, excessive promotion of electrification leads to a decline in the emission reduction cost performance.

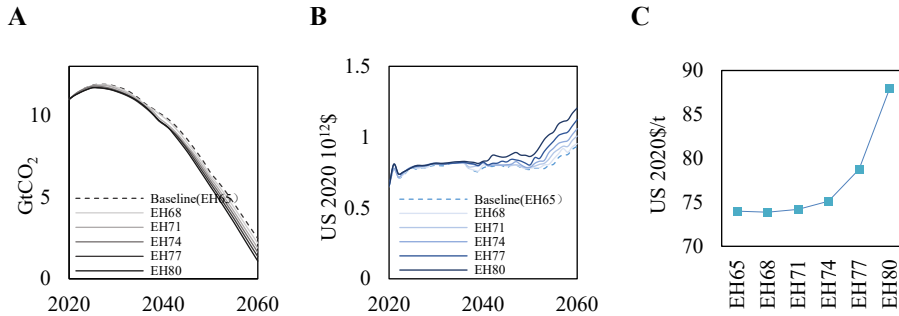


Figure 2 Changes in carbon emissions and system costs under the high proportion electricity scenario. (A) Total carbon emission. The darker the color, the higher the end-using electricity proportion. The following figures are the same. (B)  $cost_{year}$ . (C)  $cost_{CUTCRC}$ .

### 3.1.2. Power supply structure

Non fossil power shows a significant increase, as shown in Figure 3A. The EH80 (2060) and the EH65 (2060) have increased by 9.3 times and 8.5 times respectively compared with 2020. For fossil power, the capacity of fossil power remains stable from 2020 to 2040, and declines after 2040 (Figure 3B). However, with the increase of end-using electricity proportion, the installed capacity of fossil power declines slowly. In 2060, fossil power installations are dominated by coal power with CCS and gas power with CCS, which mainly play a peak shaving role. This phenomenon indicates that the rising non fossil power installations lead to the growth of power fluctuations. From the perspective of cost, the fossil power installation with CCS brings more construction costs and operating costs than the non fossil power installation, thus it brings an increase in the system cost of unit power generation (Figure 3C), and this cost of EH80 is 20.3% higher than that of EH65.

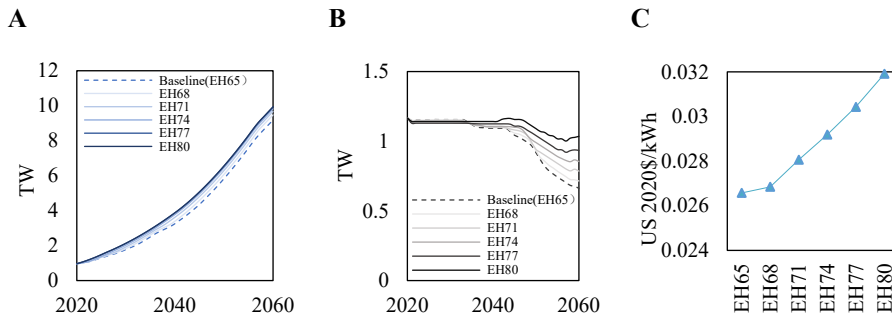


Figure 3: Changes in power supply structure under the high proportion electricity scenario. (A) Non fossil power installations capacity (B) Fossil power installations capacity (C) System cost of unit power generation

3.1.3. High proportion hydrogen scenario results

3.1.4. Total emissions and system costs

Under the setting of high proportion hydrogen scenario, the sum of end-using electricity and hydrogen proportion among different scenarios is the same, so there is little difference in carbon emission trajectory and annual system cost close to 2060. The difference in results is mainly reflected in the advance of peak year due to the increase of hydrogen (Figure 4A) and the change of total accumulated cost ( $cost_{TAC}$ ). The expression of  $cost_{TAC}$  is as follows:

$$cost_{TAC} = \sum_{2020}^{2060} (cost_{year_c} + cost_{year_o} + cost_{year_g} + cost_{year_h} + cost_{year_e}) \quad (3)$$

The  $cost_{TAC}$  of H15 decreased by 0.5% compared with H12, while the  $cost_{TAC}$  of H30 increased by 2.7% compared with H12 (Figure 4B). From the perspective of cost performance of emission reduction. The  $cost_{CUTCRC}$  increases first and then decreases with the increase of hydrogen, it shows that both lower and higher end-using hydrogen proportions have good emission reduction cost performance.

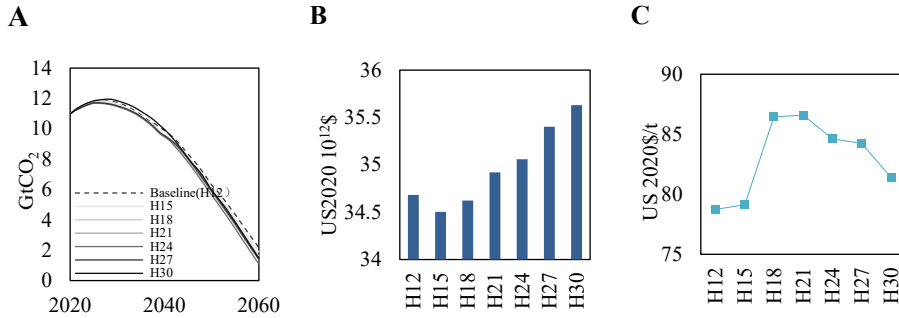
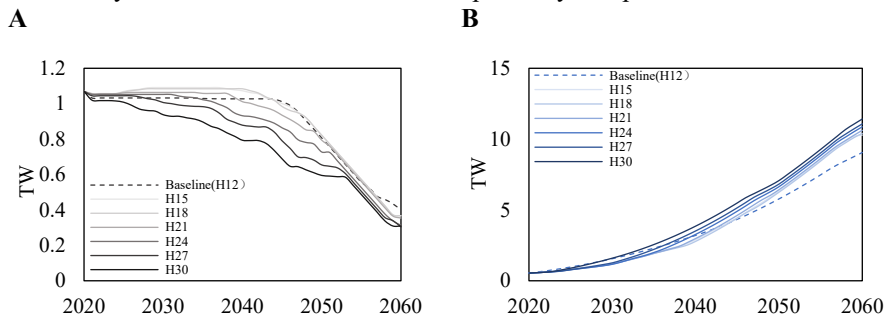


Figure 4: Changes in carbon emissions and system costs under the high proportion hydrogen scenario. (A) Carbon emissions. (B)  $cost_{TAC}$  (C)  $cost_{CUTCRC}$

3.1.5. Power supply structure

The installed fossil capacity shows a decline after a steady change in time (Figure 5A). With the increase of hydrogen, the decline time of the fossil capacity is advanced. The decline time of H12 is 2046, while the decline time of H30 is 2025. At the same time, the decline speed of the scenario with a higher hydrogen proportion is relatively slow. A higher hydrogen proportion means a higher demand for wind power and photovoltaic installation (Figure 5B). The installed capacity in H12 (2060) and H30 (2060) has increased by 15.76 times and 20.15 times respectively compared with that in 2020.



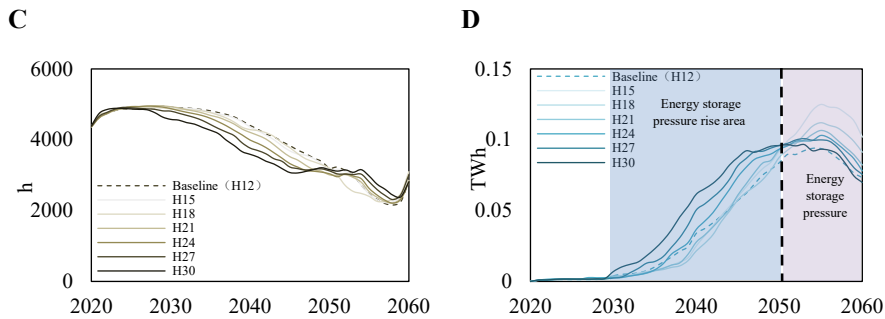


Figure 5: Changes in power supply structure under the high proportion hydrogen scenario. (A) Fossil capacity. (B) Non fossil capacity. (C) Operating hours of coal power. (D) Energy storage capacity

The operating hours of coal power decrease in advance with the increase of hydrogen (Figure 5C), which means that hydrogen promotes the transition of coal power from the main power generation energy to the energy with peak shaving function. This advance also shows that the development of hydrogen in 2030-2050 brings a large number of wind power and photovoltaic installation that need a large amount of energy storage for peak shaving (Figure 5D). From 2050 to 2060, it can be seen that electric energy storage decreases with the increase of hydrogen in the last ten years. Therefore, the increase of hydrogen causes the increase of energy storage pressure from 2030 to 2050, but the decrease of electricity reduces the energy storage pressure from 2050 to 2060.

#### 4. Conclusions

Based on CRESOM, this study reveals the impact of electricity and hydrogen on the low-carbon transition of energy supply system through multi scenario calculations.

Under the high proportion electricity scenario, the promotion of electrification can promote the reduction of carbon emissions, whilst excessive promotion of the end-using electricity proportion brings about an increase in the  $cost_{CUTCRC}$ , this also means the reduction of the cost performance of emission reduction the system. The higher non fossil power installations capacity, the higher the demand for peak shaving. Therefore, more fossil power with CCS is retained, but system cost of unit power generation rises.

Under the high proportion hydrogen scenario, the increase of hydrogen brings forward the peak year of carbon. For  $cost_{CUTCRC}$ , both lower and higher end-using hydrogen proportions have good emission reduction performance. The increase of wind power and photovoltaic installation brought about by the increase of hydrogen demand has led to more peak shaving demand. Fossil energy has changed to peak shaving energy in advance, and the demand for energy storage has increased in advance.

#### References

- BP, 2020, BP Energy Outlook 2020 edition.
- H. Dai, X. Xie, Y. Xie, et al., 2016, Green growth: The economic impacts of large-scale renewable energy development in China[J]. Applied Energy, 162: 435-449.
- Z. Guo, L. Ma, P. Liu, et al., 2017, A long-term multi-region load-dispatch model based on grid structures for the optimal planning of China's power sector. Computers & Chemical Engineering, 102:52-63.
- S. Guo, D. Yan, S. Hu, et al., 2021, Modeling building energy consumption in China under different future scenarios[J]. Energy, 214: 119063.

H. Duan, S. Zhou, K. Jiang, et al., 2021, Assessing China's efforts to pursue the 1.5°C warming limit. *Science*, 372: 378–385.

X. Zhang, X. Huang, D. Zhang, et al., 2022, Research on the Pathway and Policies for China's Energy and Economy Transformation toward Carbon Neutrality. *Management World*, 1:35-51.



# Optimal planning for thermal power plants with CCS under source–sink matching constraints: A case study of China

Binglin Du,<sup>a</sup> Pei Liu,\* Zheng Li

*<sup>a</sup>State Key Lab of Power Systems, Department of Energy and Power Engineering, Tsinghua-BP Clean Energy Center, Tsinghua University, Beijing 100084, China  
[liu\\_pei@tsinghua.edu.cn](mailto:liu_pei@tsinghua.edu.cn)*

## Abstract

The low-carbon transition pathway of the power sector, especially thermal power plants is significant for reducing greenhouse gas emissions worldwide. Carbon capture and storage (CCS) technology is identified as a critical decarbonization technology, especially in a thermal power-dominated power system. Previous studies have focused on the potential of CCS in low-carbon transformation of large-scale coal-fired power plants, but rarely considered the fact that many thermal power plants locate far from suitable CO<sub>2</sub> storage places, which may lead to deviations in the CCS potential. To quantify the application potential of CCS technology in power sector at a regional level, a long-term power generation expansion planning model is proposed in this paper, featuring detailed technical and economic characteristics of various source–sink matching options, retrofitting options and carbon emission constraints, which can get the optimal planning for thermal power plants with CCS. The power supply system of China is selected as a case study, due to its large capacity and relatively young average life of thermal power plants. Results show that with consideration of source–sink matching, technically and economically viable capacity of power plants that can be retrofitted with CCS can be reduced by 50 percent with a more concentrated distribution in North China.

**Keywords:** Carbon capture and storage; Coal-fired power plants; Source–sink matching; China

## 1. Introduction

The decarbonization of power sector is decisive to tackle climate change challenge. Carbon emissions from electricity and heat generation sector account for 40 % of global energy-related carbon emissions, and the share of electricity in energy consumption will keep rising in the future (IEA, 2019). Besides, the power sector is projected to attain carbon neutral or even negative carbon emissions earliest among all sectors, due to its higher concentration of carbon emissions and more advanced carbon emission reduction technologies (Zhang et al.,2013).

Carbon capture and storage (CCS) technology is identified as a critical decarbonization technology, especially in a thermal power-dominated power system, for the reason that CCS is the only technology choice to achieve net-zero emission of fossil energy. If the large fleet of existing thermal power plants, especially coal-fired power plants are not retrofitted with carbon capture and storage (CCS) technologies, they will

retire earlier than designed lifetime, which will bring about high stranding cost (Zhang et al., 2021). Besides, when penetration of renewable energy becomes high, more flexibility is needed to ensure the reliability of the power supply (Chen et al., 2021). CCS technology is an important technical choice for coal-fired and gas power stations to participate in balancing the fluctuation of renewable energy in the carbon-neutral power system (Xiao et al., 2022).

Many researchers have studied the potential of CCS from the perspective of power sector development and carbon emission constraints. In the Intergovernmental Panel on Climate Change (IPCC) Special Report on Global Warming of 1.5°C, it is proposed that CCS technology plays an important role in the three scenarios with a 1.5°C target (IPCC, 2018). For the power industry, CCS can contribute 14% of the emission reduction by 2050 (IEA, 2017). However, these studies pay little attention to the fact that many thermal power plants locate far from suitable CO<sub>2</sub> storage places, which may lead to rather high costs for transportation and storage of CO<sub>2</sub> (Wei et al. 2021). The lack of CCUS source-sink matching constraints will lead to deviations in the CCS potential.

Relevant studies rarely focus on CCUS source-sink matching compared with other topics, such as cost and investment evaluation (Fan et al., 2020). Wang et al. (2020) proposed a source-sink matching model for optimal CCUS deployment in China's existing coal power plants to achieve the 2 °C target, and found 175 GW coal-fired power plants need to be retrofitted with CCS under the 2 °C constraint. Fan et al. (2020) evaluated the deployment potential of CCUS in existing coal-fired power plants by a high precision CCUS source-sink matching model. However, these studies on CCUS source-sink matching lack consideration of the whole power system development, such as the power supply mix and the scale of power transmission.

To address the research gap, a long-term power generation expansion planning model with source-sink matching options is proposed, which can give the optimal planning of thermal power plants with CCS with the overall consideration of power sector development and source-sink matching constraints. Seven scenarios are set to quantitatively analyze the impact of source-sink matching constraints on coal power development.

The structure of this paper is as follows. Section 2 introduces the methodology, section 3 introduces the case study and its results, and section 4 is the conclusion.

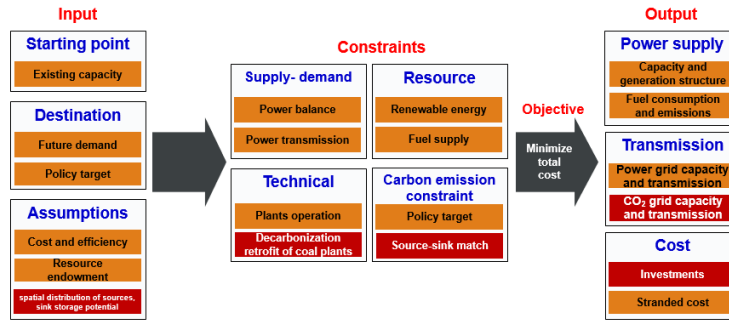
## 2. Methodology

### 2.1. Model structure and assumptions

The long-term power generation expansion planning model with source-sink matching options proposed here aims to optimize long-term planning both for power systems and CCS. The model structure is shown in Fig.1. The input part includes the existing installed capacity, costs, and efficiency of different power generation technologies, predicted future power demand, policy target, and resource endowment of renewable energy and fossil fuels. The constraints can be divided into four parts: (1)the balance of power supply and demand; (2)the potential of each energy source; (3)the technical factors of power plants; (4)the carbon target. The model objective is to minimize the total system cost in the planning horizon. The output can provide the power supply mix, power, and CO<sub>2</sub> transmission, the annual investment. Among them, the spatial distribution of sources, sink storage potential and the cost of CO<sub>2</sub> transport cost constitute the regional source-sink matching options, which will get the distribution of CCS technology and the transmission of CO<sub>2</sub>.

The model includes sixteen different power generation technologies: Sub and super-

critical coal power plants (SPC), Ultra-supercritical coal power plants (UPC), Sub and super-critical coal power plants with CCS(SPCC), Ultra-supercritical coal power plants with CCS(UPCC), Sub and super-critical coal-biomass co-firing power plants with CCS(SPCCOCCS), Ultra-supercritical coal-biomass co-firing power plants with CCS(UPCCOCCS), Natural gas combined cycle(NGCC), Nuclear(NU), Hydropower(HD), Wind onshore(WD), Wind offshore(WDOFF), Centralized solar photovoltaic(PV), Distributed solar photovoltaic(PVDIS), Biomass(BE), Biomass with CCS(BECCS). Among them, coal-fired power plants and natural gas combined cycle power plants can be retrofitted with CCS or decommissioned earlier than the anticipated lifetime, and all power plants with CCS will be constrained by source-sink matching. The model is a multi-regional model that includes 17 regions based on resources and grid structure. It also takes into account short-term planned cross-region transmission lines. Each year in the planning horizon is split into 24 time slices in the temporal module, with



four typical days for each season.

Fig.1 Model structure

## 2.2. Mathematical formulation

The six subscripts (t, r, b, g, f, s) are set respectively to cover different dimensions (year, region, basin, technology, fuel type, and time slice) for the parameters and variables.

### 2.2.1. Objective function

The model objective is minimizing the accumulated total system cost (*atoc*) in the planning horizon (from 2020 to 2060), including capital cost  $tocac_{r,t}$ , capital cost for retrofit  $torfac_{r,t}$ , operation, and maintenance cost  $toomc_{r,t}$ , fuel cost  $tofc_{r,t}$ , power transmission cost  $totrc_{r,t}$ , and CO<sub>2</sub> transmission cost  $toctrc_{r,t}$ , as shown in Eq. (1) – (7).

$$atoc = \sum_r \sum_{t=2020}^{2060} \left( \frac{tocac_{r,t} + torfac_{r,t} + toomc_{r,t} + tofc_{r,t} + totrc_{r,t} + toctrc_{r,t}}{(1+I)^{t-2020}} \right) \quad (1)$$

$$tocac_{r,t} = \sum_g \sum_{t'=t-ELT_g+1}^t \left( CAC_{r,t',g} \cdot nbca_{r,t',g} \cdot \frac{I \cdot (1+I)^{-1}}{1 - (1+I)^{-ELT_g}} \right) \quad (2)$$

$$torfac_{r,t} = \sum_g \sum_{t'=t-ELT_g+1}^t \sum_{t''=t'+1}^t \left( RFCAC_{r,t',g} \cdot rf_{r,t',g} \cdot \frac{I \cdot (1+I)^{-1}}{1 - (1+I)^{-(t'+ELT_g-t'')}} \right) \quad (3)$$

$$toomc_{r,t} = \sum_g OMC_{t,g} \cdot inca_{r,t,g} \quad (4)$$

$$tofc_{r,t} = \sum_f FEULP_{f,r,t} \cdot \sum_{g,s} feuld_{f,r,t,g,s} \quad (5)$$

$$totrc_{r,t} = \sum_s \sum_{r' \neq r} idpt_{r,r',t,s} \cdot PTRC_{r,r'} \quad (6)$$

$$toctrc_{r,t} = \sum_s \sum_b cot_{r,b,t,s} \cdot CTRDIS_{r,b} \cdot CTRC_{r,b} \quad (7)$$



### 2.2.2. Operational constraints

Regional power demand is needed to satisfy by regional power supply, which equal regional power generation  $pge_{r,t,g,s}$  (installed capacity  $inca_{r,t,g}$  multiply hours of operation  $H_{r,t,s,g}$ ) coupled with net power import  $npi_{r,t,s}$  (power import  $pit_{t,r',r}$  minus power export  $pit_{t,r'',r}$ ) and the net electricity provided by energy storage (electricity discharged  $dchar_{r,t,s}$  minus electricity charged  $char_{r,t,s}$ ), as expressed in Eq. (8)–(10). Ideal power transmission from  $r$  to  $r'$  equals the negative value of that from  $r'$  to  $r$  and is limited by transmission capacity.

$$PD_{r,t,s} \leq \sum_g egs_{r,t,g,s} + npi_{r,t,s} + dchar_{r,t,s} - char_{r,t,s} \quad (8)$$

$$egs_{r,t,g,s} \leq inca_{r,t,g} * H_{r,t,s,g} \quad (9)$$

$$npi_{r,t,s} = \sum_{r'',r' \neq r} [pit_{t,r',r} - (1 - TRLOSS_{r,r'}) * pit_{t,r'',r}] \quad (10)$$

### 2.2.3. Construction constraints

The installed capacity of technologies equals the sum of newly built capacity ( $nbca_{r,t,g}$ ) each year during the past years (i.e. lifetime of each power plants) minus the early retired capacity ( $rfca_{r,t,g}^{t'}$ ) and retrofitted capacity ( $erca_{r,t,g}^{t'}$ ) plus the capacity retrofitted by other power plants ( $erfca_{r,t,g}^{t''}$ ), as shown in Eq. (13).

$$inca_{r,t,g} = \sum_{t'=t-ELT_g+1}^t nbca_{r,t',g} - \sum_{t'=t-ELT_g+1}^t \sum_{t''=t'+1}^t (rfca_{r,t',g}^{t''} + erca_{r,t',g}^{t''}) + \sum_{t'=t-ELT_g+1}^t \sum_{t''=t'+1}^t (erfca_{r,t',g}^{t''}) \quad (11)$$

Other inequality constraints such as the upper bound of regional total installed capacity of renewable energy, annual newly-built capacity are also considered.

### 2.2.4. Carbon emission and source-sink matching constraints

The annual net CO<sub>2</sub> emission ( $tocoe_t$ ) equals CO<sub>2</sub> emission from fuel energy combustion minus the amount of CO<sub>2</sub> capture, which equals the sum of CO<sub>2</sub> stored in each basin. The total emission in the planning horizon cannot exceed the 1.5°C carbon target ( $COET$ ). Each basin's injection capacity is limited, including storage potential ( $CSP_b$ ) and annual injection rate ( $ACSP_b$ ).

$$tocoe_t = \sum_f tofueled_{f,t} \cdot COEF_f - \sum_r \sum_s \sum_b cot_{r,b,t,s} \quad (12)$$

$$\sum_{t=2020}^{2060} tocoe_t \leq COET \quad (13)$$

$$\sum_t \sum_r \sum_s cot_{r,b,t,s} \leq CSP_b \quad (14)$$

$$\sum_r \sum_s cot_{r,b,t,s} \leq ACSP_b \quad (15)$$

## 3. Case Study

The power supply system of China is selected as a case study, due to its large capacity and relatively young average life of thermal power plants. At the same time, the source-sink matching of China's CCS technology is not ideal, which makes it difficult to reduce fossil energy emissions and achieve carbon neutrality target. This study will apply the above method to explore the impact of consideration and non-consideration of CCS source-sink matching and the number of storage sites and transportation costs on the trajectory of thermal power unit development during the planning period of 2020–2060.

The economic and technical parameters of each power plant are incorporated from previous work (Chen et al., 2021). The transmission cost of carbon dioxide is set based on the report about China's CCUS development (Zhang et al., 2021). And the injection

capacity of each basin and the distance from emission source to sink refer to the research of current China CCUS source–sink matching situation (Chen et al.,2021).

7 scenarios (S1–S7) are set in this study. S1 don't take the CCS source-sink matching restrictions and relative cost into account. S2 consider the CCS source-sink matching constraints with only 9 basins (Tarim, Junggar, Ordos, Songliao, Erlian, Tuopan-Hami, Sanjiang, Qaidam, Hailaer), which are far from populated areas for security reasons. S3-S7 consider the CCS source-sink matching with 16 basins, which also includes 7 basins in Central and South China (Sichuan, Qinshui, Bohai Bay, Suibei, Hehuai, Jiangnan, Nanxiang ). In S2-S7, the cost of transporting CO2 will decrease by 50%, 90%, 70%, 50%, 30%, and 10% in 2060 compared to 2020. For modelling and optimization, the General Algebraic Modelling System (GAMS, GAMS Development Corporation, Washington, DC, USA) Linear Programming Solver was employed.

3.1. Installed capacity of thermal power plants with CCS in 2060

As shown in Figure 1 the constraint of source-sink matching will have a great impact on the total installed capacity of thermal power plants with CCS. When there are only 9 optional basins and the cost of CO2 transportation drops by 50 %, the total installed capacity falls from 592 GW to 145 GW, with a decrease of 76.1 %, and the overall installed capacity was 236GW when there were 16 chosen basins, with a decrease of 60.1%. Among these thermal power plants, the most affected is the UPCC, with a decreased by 86.3 % and 77.8 % in S2, S5 relative to S1. Analysis of S3-S7 shows that for every 20 % reduction in carbon dioxide reduction rate, the installed capacity of thermal power CCS units also decreases by 29-36 GW.

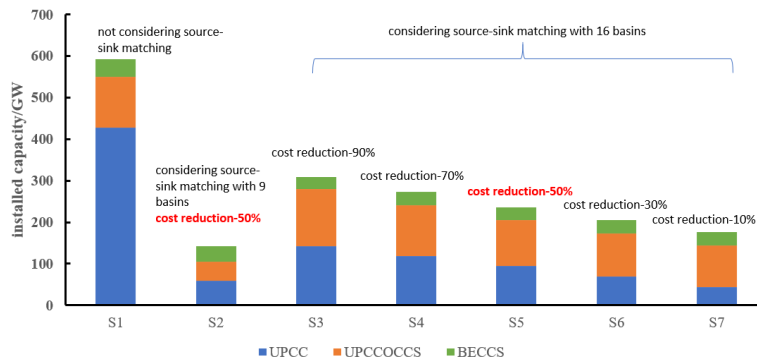


Figure 1 Installed capacity of thermal power plants with CCS in 2060

3.2. Spatial distribution of thermal power plants with CCS in 2060

With the consideration of CO2 transmission cost, the installed capacity of coal plants with CCS in Inner Mongolia and East and BECCS in Chuan-Yu, Central and Henan will decline significantly. There will be more thermal power plants with CCS in places near to storage basins, such as Xinjiang, Northeast and Shandong.

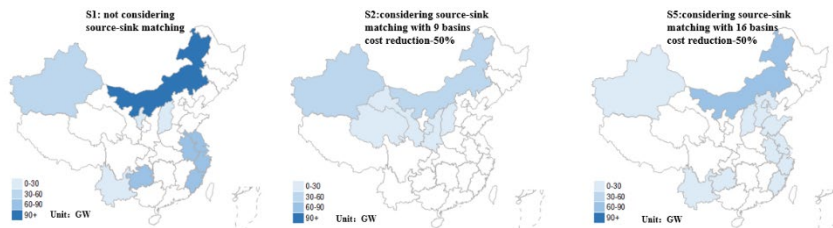


Figure 2 Spatial distribution of coal plants with CCS (UPCC and UPCCOCCS) in 2060  
 Figure 3 Spatial distribution of BECCS in 2060

### 3.3. CO<sub>2</sub> storage amount in each basin

As shown in Figure 4, CO<sub>2</sub> will be buried in Ordos, Junggar, Songliao, Erlian, and Turpan basins in the S2 scenario. The two basins with the most CO<sub>2</sub> storage are Ordos Basin and Junggar Basin, with 4.3 Gt and 2.9 Gt, respectively. CO<sub>2</sub> will also be stored in



Sichuan, Qinshui, Bohai Bay, Subei, and Nanxiang basins in the S5 scenario. Among them, the Ordos Basin, Sichuan Basin, and Bohai Bay Basin have the largest sequestration amount, with 4.3Gt, 3.7Gt, and 3.5Gt respectively.

Figure 4 CO<sub>2</sub> storage amount in each basin

## 4. Conclusions

This study proposes a low-carbon transition model of power system with source-sink matching module, which can be used to quantitatively analyze the development path of thermal power plants with CCS with minimum transition costs based on both the power system development and CCS source-sink matching constraints. The results of scenario analysis show that the total installed capacity of thermal power plants with CCS will drop significantly when source-sink matching is taken into account. The spatial distribution of CCS will change greatly, and the distance to the storage sinks will also become a key factor in the layout of CCS. The Ordos, Junggar and Bohai Bay Basin will become important CO<sub>2</sub> storage locations in the future.

## References

- IEA, 2019, World Energy Outlook.
- D. Zhang, P. Liu, L. Ma, Z. Li, 2013, A multi-period optimization model for planning of China's power sector with consideration of carbon dioxide mitigation-The importance of continuous and stable carbon mitigation policy, *Energy Policy*, 58, 319-328.
- X. Zhang, K. Li, Q. Ma, J. Fan, 2021, Orientation and prospect of CCUS development under carbon neutrality target, *China Population, Resources and Environment*, 31, 29-33.
- S. Chen, Z. Li, W. Li, 2021, Integrating high share of renewable energy into power system using customer-sited energy storage, *Renewable & Sustainable Energy Reviews*, 143.
- K. Xiao, B. Yu, L. Cheng, F. Li, D. Fang, 2022, The effects of CCUS combined with renewable energy penetration under the carbon peak by an SD-CGE model: Evidence from China, *Applied Energy*, 321.
- IPCC, 2018, Summary for Policymakers in: *Global Warming of 1.5°C*.
- IEA, 2017, *Energy Technology Perspectives*.
- Y.-M. Wei, J.-N. Kang, L.-C. Liu, Q. Li, P.-T. Wang, J.-J. Hou, Q.-M. Liang, H. Liao, S.-F. Huang, B. Yu, 2021, A proposed global layout of carbon capture and storage in line with a 2 degrees C climate target, *Nature Climate Change*, 11.

- J.-L. Fan, M. Xu, S. Wei, S. Shen, Y. Diao, X. Zhang, 2021, Carbon reduction potential of China's coal-fired power plants based on a CCUS source-sink matching model, *Resources Conservation and Recycling*, 168.
- P.-T. Wang, Y.-M. Wei, B. Yang, J.-Q. Li, J.-N. Kang, L.-C. Liu, B.-Y. Yu, Y.-B. Hou, X. Zhang, 2020, Carbon capture and storage in China's power sector: Optimal planning under the 2 degrees C constraint, *Applied Energy*, 263.
- C. Wen-Hui, L. Xi, 2022, The optimal layout of CCUS clusters in China's coal-fired power plants towards carbon neutrality, *Climate Change Research*, 18, 261-271.



# On discrete time chemical production scheduling MILP models containing record keeping variables

Amin Samadi,<sup>a</sup> Nathan Adelgren,<sup>b</sup> Christos T. Maravelias,<sup>a,b</sup>

<sup>a</sup>*Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08540, USA*

<sup>b</sup>*Andlinger Center for Energy and the Environment, Princeton University, Princeton, NJ 08540, USA*

## Abstract

In this paper, we present strategies for reformulating discrete time-based mixed-integer programming models for chemical production scheduling. We introduce several new integer variables, which we refer to as record keeping variables, that mixed-integer linear programming solvers are able to exploit in order to reduce total solution time. We consider these record keeping variables in the context of both batch process networks and continuous process networks. Results of several computational tests are provided that demonstrate the utility of incorporating record keeping variables in chemical production scheduling models.

**Keywords:** batch process, continuous process, scheduling, discrete time, reformulation.

## 1. Introduction

We consider a chemical processing facility that converts raw materials into higher value products. In this context, production scheduling involves assigning tasks to processing units and determining the times at which the processing of these tasks should begin, possibly subject to a set of restrictions. As these task assignments and timing decisions ultimately govern a significant portion of a facility's operations, scheduling is a crucial decision-making step with applications in a broad range of systems, from batch production of low-volume products such as pharmaceuticals (Papavasileiou et al., 2007) to crude oil blending (Castro and Grossmann, 2014). Determining an appropriate schedule requires knowledge of the processing duration of each task, any limitations on the units to which a given task can be assigned, material conversion coefficients, and many other complex, application-specific details. As such, much research in the field of chemical production scheduling has been devoted to developing strategies for modeling various process characteristics such as limits on utility and/or resource availability (Méndez et al., 2001), changeovers (Wolsey, 1997), and material transfer restrictions (Giménez et al., 2009). In considering this wide array of process characteristics, researchers aim to develop highly general models that can be readily employed in a wide range of industrial applications. However, discrete-time models grow at least linearly with the number of time periods (Velez et al., 2017). As such, a major research direction in production scheduling has been the development of computationally efficient procedures for solving existing scheduling models. Research efforts to reduce the computational costs of solving chemical production scheduling models have included the development of reformulations (Velez and Maravelias, 2013a), decomposition-based algorithms (Harjunkoski and Grossmann, 2002), parallel computing tools

(Subrahmanyam et al., 1996), and tightening methods using valid inequalities (Velez, 2013b). In this work we present a reformulation strategy for discrete time-based mixed-integer linear programming (MILP) models for production scheduling. We divide our discussion into two categories: (i) reformulations for MILP models related to batch processes, and (ii) reformulations for MILP models related to developments for continuous processes.

## 2. Preliminaries

### 2.1. Batch and Continuous Process Modeling Differences

There are significant modeling differences between batch and continuous processes, primarily due to the fact that continuous processes produce (output) and consume (input) materials continuously and simultaneously, whereas batch processes are assumed to consume all required materials at the start of a task and produce all outputs at the end. The amount of material produced/consumed in continuous production depends on both the rate and duration of task execution, where the duration is not strictly fixed. This implies a different number of degrees of freedom (DOF). Batch processes have one DOF: the batch size. Continuous processes, however, have two DOF: the processing rate and duration of a task. Because of this, MILP models for continuous processes are often carefully constructed extensions of MILP models for batch processes. Hence, we first discuss model reformulations for batch process and later discuss them for continuous processes.

### 2.2. Notation

We employ the following convention for notation: (i) sets are indicated using bold, upper-case, Roman letters, (ii) indices are indicated using lower-case Roman letters, (iii) parameters are indicated using lower-case Greek letters, and (iv) variables are indicated using non-bold, upper-case, Roman letters. Throughout the work we utilize  $i \in \mathbf{I}$  to represent tasks,  $j \in \mathbf{J}$  to represent processing units,  $k \in \mathbf{K}$  to represent materials, and  $n \in \mathbf{N}$  to represent discrete time points. We note that, while the typical use of  $n \in \mathbf{N}$  is to represent a discrete time *point* that is  $n\delta$  time units beyond the start of the scheduling horizon, it can also be used to represent the time *period*  $[(n-1)\delta, n\delta)$ . For a given  $i \in \mathbf{I}$ , we use  $\mathbf{J}_i \subseteq \mathbf{J}$  to represent the set of units capable of processing task  $j$ . Additionally, for a given  $k \in \mathbf{K}$ , we use  $\mathbf{I}_k^+$ ,  $\mathbf{I}_k^- \subseteq \mathbf{I}$  to represent the sets of tasks producing and consuming material  $k$ , respectively. The MILP models we employ also rely on the parameters:  $\eta$  - the scheduling horizon;  $\delta = \frac{\eta}{|\mathbf{N}|}$  - the discretization time step;  $\tau_{i,j}$  - the time required to process task  $i \in \mathbf{I}$  in unit  $j \in \mathbf{J}_i$ ;  $\beta_j^{\text{MAX}}/\beta_j^{\text{MIN}}$  - the maximum/minimum capacity of unit  $j \in \mathbf{J}$ ;  $\chi_k^{\text{MAX}}$  - the maximum amount of material  $k \in \mathbf{K}$  that can be stored;  $\rho_{i,k}$  - the conversion coefficient of material  $k \in \mathbf{K}$  produced or consumed by task  $i \in \mathbf{I}$ ; and  $\xi_{k,n}$  - the net shipment of material  $k \in \mathbf{K}$  at time  $n \in \mathbf{N}$ . We also use the variables:  $X_{i,j,n} \in \{0,1\}$ , which equals 1 if task  $i \in \mathbf{I}$  begins in unit  $j \in \mathbf{J}_i$  at time point  $n \in \mathbf{N}$ ;  $B_{i,j,n} \in \mathbb{R}_+$ , the batch size of task  $i \in \mathbf{I}$  processed in unit  $j \in \mathbf{J}_i$  at time point  $n \in \mathbf{N}$ ; and  $S_{k,n} \in \mathbb{R}_+$ , the inventory level of material  $k$  during time period  $t$ . The MILP models we employ can utilize various objectives, but all include at least the following sets of constraints:

$$\sum_{i \in \mathbf{I}_j} \sum_{t=0}^{\lceil \tau_{i,j}/\delta \rceil - 1} X_{i,j,n-t} \leq 1 \quad \forall j \in \mathbf{J}, n \in \mathbf{N}$$

$$\beta_{i,j}^{\text{MIN}} X_{i,j,n} \leq B_{i,j,n} \leq \beta_{i,j}^{\text{MAX}} X_{i,j,n} \quad \forall i \in \mathbf{I}, j \in \mathbf{J}_i, n \in \mathbf{N}$$

$$S_{k,n+1} = S_{k,n} + \sum_{i \in \mathbf{I}_k^+} \sum_{j \in \mathbf{J}_i} \rho_{i,k} B_{i,j,n} + \sum_{i \in \mathbf{I}_k^-} \sum_{j \in \mathbf{J}_i} \rho_{i,k} B_{i,j,n} + \xi_{k,n} \leq \chi_k \quad \forall k \in \mathbf{K}, n \in \mathbf{N}$$

### 3. Batch Processes

Beginning with the MILP model employing the three sets of constraints given above, Velez and Maravelias (2013a) propose a reformulation that involves adding an integer variable  $N_{i,j}$  that represents the number of times task  $i$  is carried out in unit  $j$ . Specifically, for each  $i \in \mathbf{I}$  and  $j \in \mathbf{J}_i$  the authors add the constraint  $\sum_{n \in \mathbf{N}} X_{i,j,n} = N_{i,j}$  and bound  $N_{i,j}$  as  $0 \leq N_{i,j} \leq \left\lfloor \frac{\eta}{\tau_{i,j}} \right\rfloor$ .

We propose the use of several additional integer variables and associated constraints. Consider the variables  $N_i \in \mathbb{Z}$  - the number of times task  $i$  is performed;  $N_j \in \mathbb{Z}$  - the number of times unit  $j$  performs a task;  $N_n \in \mathbb{Z}$  - the number of tasks performed at time  $n$ ; and  $N \in \mathbb{Z}$  - the total number of tasks performed. These can be incorporated into the MILP model using the following constraints and bounds:

$$N_i = \sum_{j \in \mathbf{J}_i} \sum_{n \in \mathbf{N}} X_{i,j,n}; \quad 0 \leq N_i \leq \sum_{j \in \mathbf{J}_i} \left\lfloor \frac{\eta}{\tau_{i,j}} \right\rfloor \quad \forall i \in \mathbf{I},$$

$$N_j = \sum_{i: j \in \mathbf{J}_i} \sum_{n \in \mathbf{N}} X_{i,j,n}; \quad 0 \leq N_j \leq \left\lfloor \frac{\eta}{\min_{i: j \in \mathbf{J}_i} \{\tau_{i,j}\}} \right\rfloor \quad \forall j \in \mathbf{J},$$

$$N_n = \sum_{i \in \mathbf{I}} \sum_{j \in \mathbf{J}_i} X_{i,j,n}; \quad 0 \leq N_n \leq |\mathbf{J}| \quad \forall n \in \mathbf{N},$$

$$N = \sum_{i \in \mathbf{I}} \sum_{j \in \mathbf{J}_i} \sum_{n \in \mathbf{N}} X_{i,j,n}; \quad 0 \leq N \leq \min \left\{ \sum_{i \in \mathbf{I}} \sum_{j \in \mathbf{J}_i} \left\lfloor \frac{\eta}{\tau_{i,j}} \right\rfloor, \sum_{j \in \mathbf{J}} \left\lfloor \frac{\eta}{\min_{i: j \in \mathbf{J}_i} \{\tau_{i,j}\}} \right\rfloor \right\}.$$

As each of the variables  $N_{i,j}$ ,  $N_i$ ,  $N_j$ ,  $N_n$ , and  $N$  serves to *keep record* of a quantity of interest, we refer to each as a *record keeping variable* (RKV). We now summarize the results of a study in which we compare the performance of branch-and-bound (BB) when a carefully chosen subset of these RKVs are included within our model. We note that this, as well as the remainder of the tests described in this work, were conducted using GAMS v36.1 for modeling and CPLEX 20.1 as the MILP solver. The tests described here are performed using a combination of randomly generated instances and instances obtained from minlp.org and from Velez and Maravelias (2013a). In total, we utilize 115 instances and consider cost minimization and profit maximization objectives. For brevity, and because the results for both objective types displayed relatively similar patterns, we only present results for cost minimization here. We report results for all instances that at least one of the considered formulations was able to solve in under 5 hours. The results are summarized in the performance profile displayed in Figure 1.

In Figure 1, we use the following notation: O - no RKVs are added; B -  $N_{i,j}$  is added; I -  $N_i$  is added; BIA -  $N_{i,j}$ ,  $N_i$ , and  $N$  are added; BIJA -  $N_{i,j}$ ,  $N_i$ ,  $N_j$ , and  $N$  are added; and BIJNA -  $N_{i,j}$ ,  $N_i$ ,  $N_j$ ,  $N_n$ , and  $N$  are added. Here O is selected as it represents the original formulation with no RKVs, B is selected as it represents the most promising reformulation considered in (Velez and Maravelias, 2013a), and the rest of the considered reformulations are selected for their relatively strong performance. Of the 115 instances



considered, there were 11 that no considered formulation was able to solve in 5 hours, 8 that were solved in under 5 hours by B and BIJA but not by O (B - 4667.8s avg; BIJA - 170.4s avg), and 22 that were solved in under 5 hours by BIJA but not be either O or B (BIJA - 399.4s avg). Moreover, of the remaining 74 instances, there were 44 that at least one of the formulations O, B, or BIJA took more than 5 seconds to solve, and on average BIJA solved these in 1.16% of the time used by O and 2.63% of the time used by B.

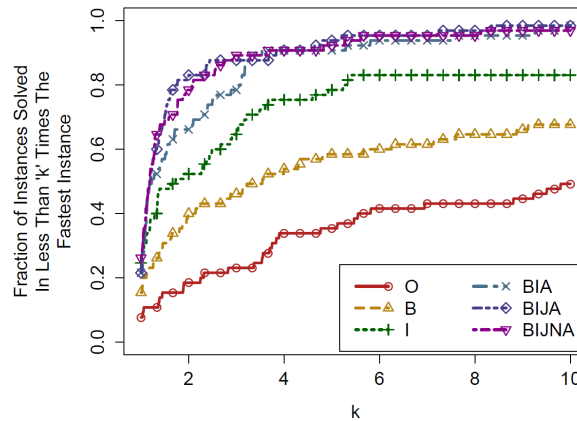


Figure 1: Performance Profile for Batch Process Formulations

#### 4. Continuous Processes

In the field of chemical production scheduling, less focus has been given to continuous processes relative to batch processes. The simultaneous production and consumption of materials in conjunction with the flexibility to choose the duration of task execution adds a layer of processing complexity to continuous process models. Strategies to more efficiently model continuous processes, such as introducing RKVs, would prove to be valuable, especially when considering transient operations such as startups, shutdowns, and direct transition tasks that notably escalate model complexity; transient operations can result in the addition of up to four more binary variables per task as well as the ancillary constraints needed for their modeling. A general optimization framework that accurately represents system dynamics, specifically transient operations, is utilized; however, the reader is directed to the literature for the extensive list of modeling constraints (Wu and Maravelias, 2021).

From a modeling perspective, a continuous task is represented as a series of identical subtasks, each having a processing time of one discrete time step. Any set of consecutive executions of a single subtask are then referred to as a *run* of the associated continuous task. The minimum/maximum processing rates of a task  $i \in \mathbf{I}$  executed in a unit  $j \in \mathbf{J}_i$  during each time period are comparable to batch sizes in the batch process model, so they are represented similarly ( $\beta_{i,j}^{\text{MIN}}/\beta_{i,j}^{\text{MAX}}$ ). In terms of run duration, parameters to enforce run length restrictions ( $\tau_{i,j}^{\text{MIN}}/\tau_{i,j}^{\text{MAX}}$ ) are also applied. In addition to  $X_{i,j,n}$ , two new binary variables containing information about the start and end times of a run are employed;  $Y_{i,j,n}^{\text{S}}$  and  $Y_{i,j,n}^{\text{E}}$  indicate that a run of task  $i$  in unit  $j$  starts, or ends, at time  $n$ , respectively. The values of these variables are enforced as:

$$Y_{i,j,n}^{\text{S}} = X_{i,j,n} - X_{i,j,n-1} + Y_{i,j,n}^{\text{E}} \quad \forall i \in \mathbf{I}^{\text{P}}, j \in \mathbf{J}_i, n \in \mathbf{N}$$

The above equation ensures that  $Y_{i,j,n}^S = 1$  whenever a run of a production task  $i \in \mathbf{I}^P \subseteq \mathbf{I}$  begins. We note that production tasks are simply tasks that are not transition tasks, i.e., their role is to produce material, as opposed to transitioning to/from other tasks.

The same RKVs introduced in Section 3, i.e., those defined in terms of the  $X_{i,j,n}$  binary variables, can be written for continuous process models. However, with the availability of the  $Y_{i,j,n}^S$  binary variable, an additional class of RKVs can be as follows:

$$\begin{aligned}
 N_{i,j}^Y &= \sum_{n \in \mathbf{N}} Y_{i,j,n}^S; & 0 \leq N_{i,j}^Y &\leq \left\lfloor \frac{\eta}{\tau_{i,j}^{\text{MIN}} + \alpha_{i,j}^{\text{MIN}}} \right\rfloor & \forall i \in \mathbf{I}^P, j \in \mathbf{J}_i, \\
 N_i^Y &= \sum_{j \in \mathbf{J}_i} \sum_{n \in \mathbf{N}} Y_{i,j,n}^S; & 0 \leq N_i^Y &\leq \sum_{j \in \mathbf{J}_i} \left\lfloor \frac{\eta}{\tau_{i,j}^{\text{MIN}} + \alpha_{i,j}^{\text{MIN}}} \right\rfloor & \forall i \in \mathbf{I}^P, \\
 N_j^Y &= \sum_{i: j \in \mathbf{J}_i} \sum_{n \in \mathbf{N}} Y_{i,j,n}^S; & 0 \leq N_j^Y &\leq \left\lfloor \frac{\eta}{\min_{i: j \in \mathbf{J}_i} \{\tau_{i,j}^{\text{MIN}}\}} \right\rfloor & \forall j \in \mathbf{J}, \\
 N_n^Y &= \sum_{i \in \mathbf{I}} \sum_{j \in \mathbf{J}_i} Y_{i,j,n}^S; & 0 \leq N_n^Y &\leq |\mathbf{J}| & \forall n \in \mathbf{N}, \\
 N^Y &= \sum_{i \in \mathbf{I}} \sum_{j \in \mathbf{J}_i} \sum_{n \in \mathbf{N}} Y_{i,j,n}^S; & 0 \leq N^Y &\leq \min \left\{ \sum_{i \in \mathbf{I}} \sum_{j \in \mathbf{J}_i} \left\lfloor \frac{\eta}{\tau_{i,j}^{\text{MIN}} + \alpha_{i,j}^{\text{MIN}}} \right\rfloor, \sum_{j \in \mathbf{J}} \left\lfloor \frac{\eta}{\min_{i: j \in \mathbf{J}_i} \{\tau_{i,j}^{\text{MIN}}\}} \right\rfloor \right\}.
 \end{aligned}$$

In the above,  $\alpha_{i,j}^{\text{MIN}}$  represents the minimum transition time needed to transition to and from a production task  $i$ . If a production task has no transitions associated with it,  $\alpha_{i,j}^{\text{MIN}} = 0$ . To test the utility of continuous models that incorporate RKVs, we include various combinations of the above RKVs that track the number of runs of a task ( $Y_{i,j,n}^S$ ) and the RKVs that track the task execution ( $X_{i,j,n}$ ) within our cost-minimization MILP model. A total of 96 instances were solved to optimality by all formulations within a 24-hour time limit, and a performance chart of the results is given in Figure 2.

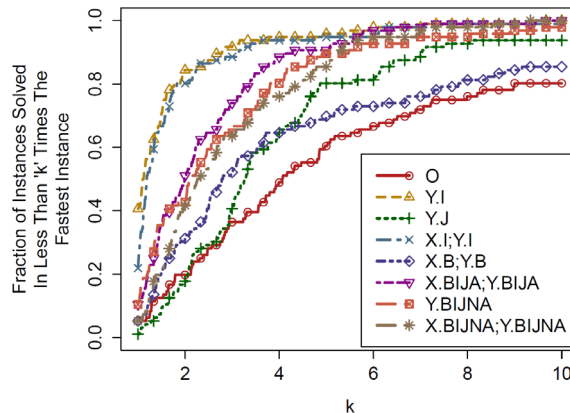


Figure 2: Performance Profile for Continuous Process Formulations

In Figure 2 we use similar notation to that of Figure 1, but we prefix the run-related variables with “Y.” and the task execution-related variables with “X.” We note that we considered several combinations of RKVs beyond those for which results are

displayed. However, for brevity, only formulations that showed strong performance relative to the original formulation with no RKVs are illustrated in the figure. From our results, it appears that the run-related RKVs perform better than the task execution-related RKVs. Additionally, we observe from the curves labeled “X.I;Y.I” and “Y.I” in Figure 2 that  $N_i^Y$  was the most impactful of our considered RKVs. We also point out that formulation Y.I solved about 40% of the instances the fastest. One possible explanation for these results is that by branching on  $N_i^Y$ , the solver is able to bound the number of runs of each task  $i$ , thus efficiently eliminating suboptimal or infeasible schedules and closing the optimality gap more quickly.

## 5. Conclusion

We have presented results that demonstrate the large computational improvements that can be achieved by simply incorporating RKVs into chemical production scheduling MILP models. Moreover, we have provided evidence that such improvements can be obtained in the context of both batch processes and continuous processes. However, this reformulation technique is not limited to only chemical production scheduling models but can also be extended to various other types of MILP models that utilize a large quantity of integer variables.

## References

- P. M. Castro and I. E. Grossmann, 2014, Global optimal scheduling of crude oil blending operations with RTN continuous-time and multiparametric disaggregation, *Industrial and Engineering Chemistry Research*, 53, 39, 15127–15145.
- D. M. Giménez, G. P. Henning, and C. T. Maravelias, 2009, A novel network-based continuous-time representation for process scheduling: Part I. Main concepts and mathematical formulation. *Computers and Chemical Engineering*, 33, 9, 1511–1528.
- I. Harjunkoski and I. E. Grossmann, 2002, Decomposition techniques for multistage scheduling problems using mixed-integer and constraint programming methods. *Computers and Chemical Engineering*, 26, 1533–1552.
- C. A. Méndez, G. P. Henning, and J. Cerdá, 2001, An MILP continuous-time approach to short-term scheduling of resource-constrained multistage flowshop batch facilities, *Computers and Chemical Engineering*, 25, 701–711.
- V. Papavasileiou, A. Koulouris, C. Siletti, and D. Petrides, 2007, Optimize manufacturing of pharmaceutical products with process simulation and production scheduling tools, *Chemical Engineering Research and Design*, 85, 7A, 1086–1097.
- S. Subrahmanyam, G. K. Kudva, M. H. Bassett, and J. F. Pekny, 1996, Application of Plant Distributed Design and Computing to Batch Scheduling, *AIChE Journal*, 42, 6, 1648–1661.
- S. Velez and C. T. Maravelias, 2013, Reformulations and branching methods for mixed-integer programming chemical production scheduling models, *Industrial and Engineering Chemistry Research* 52, 10, 3832–3841.
- S. Velez, A. Sundaramoorthy, and C. T. Maravelias, 2013, Valid Inequalities Based on Demand Propagation for Chemical Production Scheduling MIP Models, *AIChE Journal*, 59, 3, 872–887.
- S. Velez, Y. Dong, and C.T. Maravelias, 2017, Changeover formulations for discrete-time mixed-integer programming scheduling models, *Eur J Oper Res*, 260, 3, 949–963.
- L. A. Wolsey, 1997, MIP modelling of changeovers in production planning and scheduling problems, *European Journal of Operational Research* 99, 154–165.
- Y. Wu and C. T. Maravelias, 2021, A general framework and optimization models for the scheduling of continuous chemical processes, *AIChE Journal*, 67, 10, 1–15.

## ESTAN – A toolbox for global sensitivity based estimability analysis

Ilias Bouchkira,<sup>a,b</sup> Abderrazak M. Latifi,<sup>b,\*</sup> Brahim Benyahia,<sup>a,\*</sup>

<sup>a</sup> *Department of Chemical Engineering, Loughborough University, Epinal Way, Loughborough, LE11 3TU, United Kingdom.*

<sup>b</sup> *Laboratoire Réactions et Génie des Procédés, CNRS-ENSIC, Université de Lorraine, 1 rue Grandville, BP20451, Nancy Cedex, 54001, France.*

[B.Benyahia@lboro.ac.uk](mailto:B.Benyahia@lboro.ac.uk) [abderrazak.latifi@univ-lorraine.fr](mailto:abderrazak.latifi@univ-lorraine.fr)

### Abstract

Recently in process engineering field, there is an increasing demand for high fidelity, large and multi-scale mathematical models. In most cases, these models involve several unknown parameters whose identifiability from experimental measurements is often not guaranteed. It is therefore necessary to carry out an estimability analysis to determine which parameters can be reliably estimated. This task is however laborious and is still neglected in most studies. Most importantly, its wide adoption is hampered by the lack of standardized tools or methodologies. To address these issues, a new estimability toolbox, ESTAN, was developed to make the estimability analysis accessible to a broader community of specialist and non-specialist users. ESTAN can handle different types of mathematical models including dynamic and non-dynamic models. It uses a Quasi-Monte Carlo method to sample the unknown model parameters within their range of variation. Then, depending on whether the studied model is computationally cheap or expensive, global sensitivity indices are calculated using either the Sobol method or the Fourier Amplitude Sensitivity Test. The sensitivities are exploited within an orthogonalization algorithm to rank the parameters from the most to the least estimable followed by the identification of the subset of the most estimable parameters based on a preset estimability threshold. Finally, more reliable parameter estimates are obtained for the subset of the most estimable parameters. To validate the toolbox and demonstrate its capabilities, ability analysis of three models is performed using the developed toolbox. They are given by a non-dynamic, a dynamic, and a computationally expensive model. The results for the case studies are found to be very promising, showing how the presented toolbox simplifies the investigation of the estimability analysis, and significantly improves the model's precision.

**Keywords:** Sensitivity analysis, Estimability analysis, Orthogonalization, Toolbox.

### 1. Introduction

For many years, mathematical models have been employed extensively in various branches of science and engineering. They are frequently utilized in the design of processes and plants, risk analysis, process optimization, and control. These models typically consist of thermodynamic, kinetic, and hydrodynamic equations as well as momentum, mass, and energy balances. Recently, with the advent of industry 4.0 and inherent digital transformation, there is an increasing need for more precise, high fidelity, large, and multi-scale models. However, these models always involve several parameters

to be identified from available experimental measurements. The presence of a large number of model parameters to be identified combined with poor data information content and data redundancy, model structure, correlation between parameters, non-linearity with respect to model states and parameters etc., result in poor prediction capabilities, and poor model interpolation and extrapolation. To address some of these critical issues, it is necessary to implement an estimability analysis approach to assess whether the model parameters can be identified from the available or planned experimental data. The current local/global estimability approaches involve complex and laborious calculations, among which, Monte/Quasi-Monte Carlo sampling, estimation of local and/or global sensitivity indices, assessing sensitivity indices convergence, orthogonalization etc. These techniques are difficult to develop, implement, or verify, and quite often, do not necessarily belong to the modeler's skills. Consequently, the estimability part is commonly neglected by simply assuming that the available data contain the necessary information to identify all the unknown model parameters, which is very often not true.

Several estimability approaches have been investigated in the literature (Bouchkira et al., 2021, Fysikopoulos et al., 2017). These techniques are not unified, and currently, no single toolbox is available. ESTAN toolbox addresses these gaps and suggests a user-friendly interface to help address estimability in the cases of dynamic and non-dynamic models (i.e., ODEs, DAEs, Algebraic expressions). The objective is also to propose a versatile and standard tool for all engineering and scientific subjects and make the estimability analysis accessible to non-specialist users (i.e., modelers with only basic background in sensitivity and estimability analysis and/or MATLAB). The toolbox is proposed as an open-source package to allow experienced modelers to easily use it, customize and even integrate it to other digital tools.

## 2. ESTAN structure

Figure 1 shows the framework of the parameters estimability analysis toolbox ESTAN. It is based on computing the global sensitivity indices of the model inputs (with respect to the outputs) and then process the sensitivity matrix  $Z$  using an orthogonalization algorithm to determine which parameters can be reliably estimated, and in which order. The estimability matrix  $Z$  below represents the global sensitivity indices of the model inputs with respect to the outputs at a given set of operating conditions, for steady models, and sensitivity indices at given sampling points for the dynamic models.

$$Z = \begin{pmatrix} S_{1,1}|_{m_{1,k}} & \cdots & S_{1,d}|_{m_{1,k}} \\ \vdots & \ddots & \vdots \\ S_{n,1}|_{m_{n,k}} & \cdots & S_{n,d}|_{m_{n,k}} \end{pmatrix}, k \in [1; M] \quad (1)$$

where  $m_{i,k}$  is the  $i^{th}$  sampling point of the  $k^{th}$  experiment (dynamic models) or the set of operating conditions (steady state models) and  $M$  is number of observed/measured outputs,  $n$  is the number of model parameters and  $d$  is either the number of sampling points (dynamic) or number of repeated experiments (steady state or algebraic models).  $S_{i,j}$  is the individual sensitivity coefficient of the  $i^{th}$  output with respect to the  $j^{th}$  parameter.

Among several sensitivity analysis approaches available in the literature, Saltelli's approach (Saltelli, 2008) and Fast approach (Cukier et al., 1973) have been adopted in the proposed toolbox, as they are very reliable, precise, and widely accepted. In both

approaches, the variance-based sensitivity indices are given by partial and total output variances:

$$S_{i,k} = \frac{V_i}{Var(Y)} |_{k}, \quad k \in [1; M] \tag{2}$$

ESTAN toolbox requires several user inputs for sensitivity analysis performance, among which, the number of parameters and measured outputs, the variability bounds of the unknown parameters, the initial conditions etc. Before performing sensitivity analysis, well informed variability bounds are required which can be obtained from the literature for existing models. Moreover, depending on whether the model to be evaluated is computationally cheap or expensive, the toolbox computes the  $S_{i,k}$  either based on the approach developed by Saltelli (2008), using Quasi-Monte Carlo sampling and statistical estimates to approximate the average and the variance of the model outputs, or the Fast method developed by (Cukier et al., 1973), which is based on Fourier series to evaluate the partial and total variances in  $S_{i,k}$  (Test 1). Once the sensitivity indices  $S_{i,k}$  are computed, they are used to build the Z estimability matrix, to be processed by the orthogonalization algorithm as in (Fysikopoulos et al., 2019). The later allows to determine the most estimable parameters as well as their ranking according to their estimability from the available experimental measurements (Test 2).

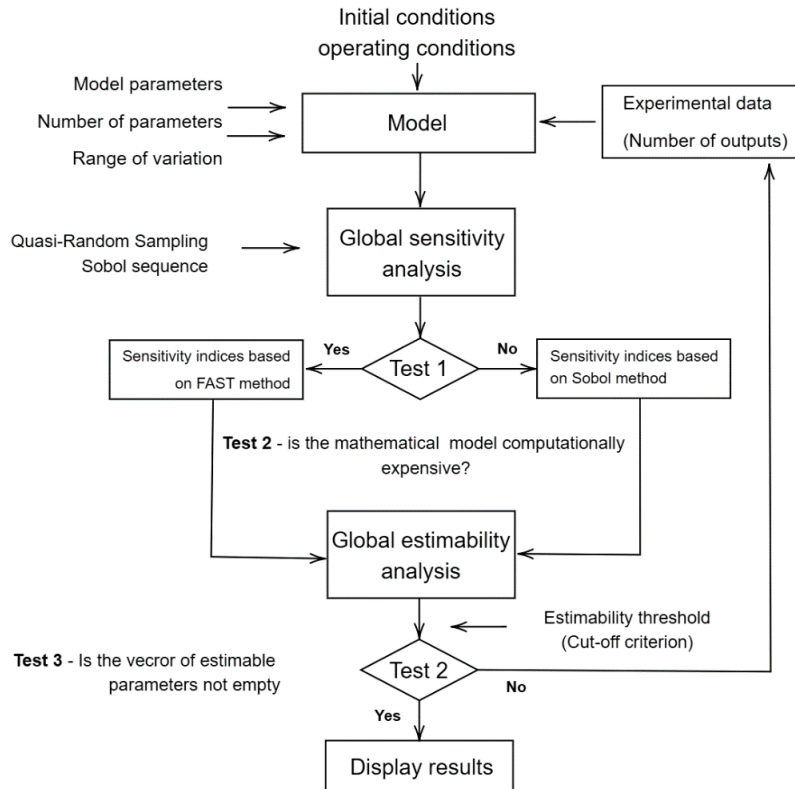
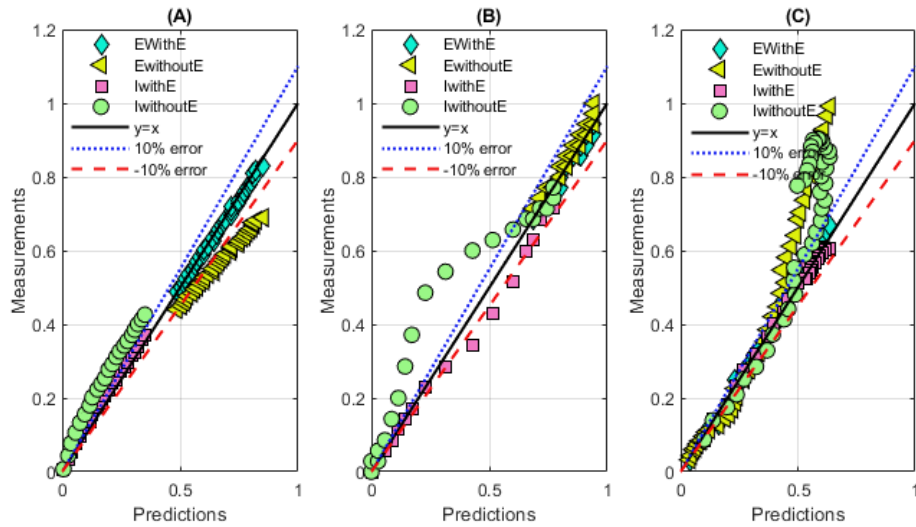


Figure 1: Simplified Framework of the parameters estimability analysis using the toolbox.



**Figure 2:** Results for model extrapolation and Interpolation. (A): Case of the non-dynamic model. (B): Case of the dynamic model. (C): Case of an expensive model. E/Iwith/withoutE: Extrapolation/Interpolation with/without estimability

**Table 1:** Summary of the outcomes of the GSA and estimability analysis.

	Interpolation	Extrapolation
Non-dynamic model		
Estimable parameters	3 out of 6	
Fixed values (3)	(Bedel et al., 2017)	
CPU time (Sobol)	173.02 seconds	
MPE with estimability (%)	4.114	4.541
MPE without estimability (%)	33.10	13.45
Dynamic model		
Estimable parameters	3 out of 4	
Fixed values (1)	(Benyahia et al., 2017)	
CPU time (Sobol)	187.21 seconds	
MPE with estimability	0.226	0.441
MPE without estimability	0.238	39.51
Expensive model		
Estimable parameters	4 out of 10	
Fixed values (6)	(Bouchkira et al., 2022;2023)	
CPU time (FAST)	359 seconds	
MPE with estimability	2.331	2.121
MPE without estimability	17.91	31.74

### 3. Case studies

#### 3.1. Models

To validate the developed toolbox, three case studies are considered. They include (1) an adsorption isothermal model from (Bedel et al., 2017) involving 6 unknown parameters. (2) A dynamic cooling batch crystallization model from (Benyahia et al., 2021) involving 4 unknown parameters. (3) A thermodynamic model from (Bouchkira et al., 2023) involving 10 unknown parameters.

#### 3.2. Experimental measurements

A database of experimental measurements was developed mainly from our previous works. For the estimability of the parameters of the adsorption isothermal model, measurements of carbon dioxide adsorption on a commercial adsorbent are taken from (Bedel et al., 2017). For the estimability of the batch cooling crystallization model, experimental measurements concerning particle size distribution of crystallized paracetamol are taken from (Anandan et al., 2021). Finally, for the thermodynamic model, equilibrium measurements are taken from (Bouchkira et al., 2022;2023).

#### 3.3. Computational method

The present toolbox was developed based on App Designer MathWorks and involves several MATLAB built-in functions which are available only in MATLAB v.2020 onwards. A reasonable computational power is required to cope with GSA and inherent simulation (e.g., Intel or AMD x86-64 processor and > 4 GB RAM). All calculations in this work were performed using an ASUS workstation with 12th Gen Intel(R) Core (TM) i9-12900K 3.20 GHz 128GB Ram processor.

### 4. Results and discussion

The estimability analysis results for the three case studies based on ESTAN toolbox are summarized in Table 1 and Figure 2. The sensitivity analysis for the first two cases was performed based on Saltelli (2008) in view of their non-expensive computing nature. To this end, a Quasi-Monte Carlo sampling of size 5000 was used. Whilst, for the last case, where the model is moderately expensive, FAST method based on Cukier et al. (1973) was investigated. The CPU times to achieve convergence of global sensitivity indices were estimated to 173 seconds for the non-dynamic model, 187 seconds for the dynamic model, and 359 for the algebraic expression-based model. The computed global sensitivity indices were then used to fill the Z estimability matrices for all tree cases, and these matrices were used to determine the number of estimable parameters in each case study, as well as their estimability-based ranking. Based on the collected data, it has been found that 3 parameters out of 6 are estimable in the first case, 3 parameters out of 4 are estimable in the second case and 4 parameters out of 10 are estimable in the last case. For each case, two scenarios are carried out for comparison purposes. In the scenarios “with estimability”, only estimable parameters are identified from the available measurements and non-estimable parameters are fixed from previous works. While in the scenario “without estimability”, all model parameters are identified neglecting the estimability analysis results. Moreover, for both scenarios, in each case study, the resulting models are used for model interpolation and extrapolation.

As expected, the results show how carrying out estimability analysis can significantly improve the model precision. This can be shown quantitatively from the computed Mean Percentage Errors (MPE) between models’ predictions and experiments. It has been found



that taking into account estimability analysis results reduced the MPE from 33.10% to 4.11% and from 13.45% to 4.54% respectively for interpolation and extrapolation in the case of the non-dynamic model, from 0.24% to 0.22% and from 39.51% to 0.44% respectively for interpolation and extrapolation in the case of the dynamic model, from 17.91 to 2.33% and from 31.74% to 2.12% respectively for interpolation and extrapolation in the case of the algebraic expression based model. These results can finally be presented in a normalized parity diagrams as in figure 2, where it can be seen that the models keep good prediction capabilities when taking estimability analysis results into account, while they fail to keep good interpolation and extrapolation performances when neglecting estimability.

## 5. Conclusion

ESTAN toolbox was developed to standardize the estimability analysis, and to help improve model prediction capabilities and precision. As demonstrated in the three case studies, ESTAN can be used for both cheap and expensive computationally models, through its ability to switch between Sobol and FAST methods for the calculation of sensitivity indices, to help enhance model prediction capabilities. Moreover, it is developed to avoid all laborious tasks involved in estimability analysis. It is worth mentioning that the current ESTAN version uses an arbitrary estimability criterion to determine the set of estimable parameters, and it is possible to implement a systematic approach to determine the optimal value of this cutoff estimability criterion as in (Bouchkira et al. 2021). This is a very interesting perspective that we will address in our future works. The estimability open-source toolbox will be freely available from the authors for non-commercial research and educational purposes.

## References

- A. Saltelli, 2002, Making best use of model evaluations to compute sensitivity indices. *Computer physics communications*, 145(2), 280–297.
- B. Benyahia, M. A. Latifi, C. Fonteix, F. Pla, (2011). Modeling of a batch emulsion copolymerization reactor in the presence of a chain transfer agent: Estimability analysis, parameters identification and experimental validation. In *Computer Aided Chemical Engineering* (Vol. 29, pp. 126-130). Elsevier.
- B. Benyahia, P. D. Anandan, & C. Rielly, (2021). Robust Model-Based Reinforcement Learning Control of a Batch Crystallization Process. In *2021 9th International Conference on Systems and Control (ICSC)* (pp. 89-94). IEEE.
- D. Fysikopoulos, B. Benyahia, A. Borsos, Z. K. Nagy, C. D. Rielly, (2019). A framework for model reliability and estimability analysis of crystallization processes with multi-impurity multi-dimensional population balance models, *Computers & Chemical Engineering*, 122, 275-292.
- I. Bouchkira, A. Latifi, L. Khamar, S. Benjelloun, 2021, Global sensitivity based estimability analysis for the parameter identification of Pitzer's thermodynamic model. *Reliability Engineering & System Safety*, 207, 107263.
- I. Bouchkira, A. Latifi, L. Khamar, S. Benjelloun, 2022. Modeling and multi-objective optimization of the digestion tank of an industrial process for manufacturing phosphoric acid by wet process. *Computers & Chemical Engineering*, 156, 107536.
- I. Bouchkira, S. Benjelloun, L. Khamar, A. Latifi, 2023. Thermodynamic modeling and parameter estimability analysis of a wet phosphoric acid process with impurities. *Fluid Phase Equilibria*, 564, 113594.
- R. Cukier, C. Fortuin, K. Shuler., A. Petschek, J. Schaibly, 1973. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. *The Journal of chemical physics*, 59(8), 3873–3878.
- S. Bedel, C. Vallieres, A. Latifi, 2017, Parameters estimability analysis and identification for adsorption equilibrium models of carbon dioxide. *Adsorption*, 23(2), 373–380.

# Multi-technology separation network synthesis

Garry S.P. Taifan,<sup>a</sup> Christos T. Maravelias,<sup>a,b</sup>

<sup>a</sup>*Department of Chemical and Biological Engineering, Princeton University, 50-70 Olden St, Princeton and 08540, USA*

<sup>b</sup>*Andlinger Center for Energy and Environment, Princeton University, 86 Olden St, Princeton and 08540, USA*  
*maravelias@princeton.edu*

## Abstract

Considering multiple technologies when synthesizing separation networks can lead to a more efficient separation. We present a superstructure-based approach to synthesize separation networks while considering multiple technologies. The approach uses key separation properties to generate potential separation sequences for each technology and employs richly connected superstructure to represent numerous potential configurations. Using a generalized problem statement, the proposed model can handle variable inlet and outlet streams thereby allowing the coupling with reactor network synthesis models.

**Keywords:** process synthesis, separation, global optimization, superstructure.

## 1. Introduction

In many chemical processes, a sequence of separation tasks is typically performed to recover unconverted raw materials and to purify final products. Synthesizing an appropriate separation network can result in significant savings for the overall process. Most of the available studies on separation network synthesis revolve around distillation with other technologies being an afterthought. Consequently, the benefits of having multiple technologies in the separation network are not fully exploited. Nevertheless, attention toward non-thermal separations has recently gained traction due to their potential for energy-efficient processes and positive environmental impact (Sholl and Lively, 2016). The incorporation of different technologies in separation network synthesis expands the search space, and thus allows potentially better networks to be found.

The separation network synthesis problem has been addressed using various approaches. Heuristic approaches rely on rules based on engineering knowledge and judgment, whereas evolutionary approaches generate a network by sequentially improving several initial flowsheets based on heuristics. Alternatively, most potentially useful units and their relevant interconnections can be embedded in a superstructure (Sargent and Gaminibandara, 1976). Superstructure-based approaches consider multitudes of promising network configurations in an optimization problem. Agrawal (2003) introduced a systematic method to generate all possible basic distillation configurations, and Heckl et al. (2006) considered different technologies explicitly. However, the former approach used a problem statement where complete information on the feed stream is required. Ryu et al. (2020) discussed a generalized problem statement which allows the seamless integration of the separation network and reactor network synthesis.

Accordingly, we propose a superstructure-based approach for separation network synthesis with multiple technologies where the feed streams may be variable, enabling coupling with reactor network synthesis. The approach employs a superstructure that

encompasses numerous promising configurations comprising multiple separation technologies. For each technology, we construct a separation matrix (Shah and Agrawal, 2010) based on the key properties that enable separation. We identify all possible separations and connections using these matrices and generate the corresponding superstructure. The synthesis problem is formulated as a mixed-integer nonlinear programming (MINLP) model. We demonstrate the applicability of the proposed approach through a multicomponent mixture separation case study.

## 2. Framework

### 2.1. Problem statement

For all technologies, we consider simple sharp split separators with no distributed components. This separation is performed by exploiting the difference in the magnitude of some key properties (e.g., volatility for distillation). Consequently, a component ranking system based on a specific key property is used to generate possible separation splits. In zeotropic distillation network synthesis, for example, all possible sequences of the separations can be enumerated if the ranking of components' relative volatilities is given. Analogously, we can generate sequences of separations for other technologies. The difference in the rankings between technologies provides an opportunity for more efficient separation. If two components are adjacent in technology A's ranking but not in technology B's ranking, separating them using technology B may be easier. It is worth noting that the actual separation may favor the former technology depending on the magnitudes of the key properties. If components have the same ranking, then it would be impractical to separate these components using that technology.

Traditionally, reactor and separation systems are synthesized sequentially; therefore, the conventional separation network synthesis problem requires full information on the inlet stream. When these two systems are considered simultaneously, the flowrates of the separation network's inlet stream vary depending on the decisions made in the reactor system. In this case, the inlet stream is termed variable inlet stream. The generalized separation network synthesis problem can be stated as follows. We are given variable inlet and outlet streams (with some specifications), the ranking of postulated components for all technologies, and cost functions for the separation tasks. The objective is to find the least cost separation network.

### 2.2. Superstructure generation

We use the concepts of nodes and arcs to represent the superstructure (Kong and Maravelias, 2020). We introduce set  $\mathbf{N}$  to denote all nodes and set  $\mathbf{I} = \{A, B, C, \dots\}$  to denote the components in the superstructure. Set  $\mathbf{N}$  can be split into three subsets: (1) source nodes  $\mathbf{N}^{SO}$  (inlets to the network); (2) sink nodes  $\mathbf{N}^{SI}$  (outlets of the network); and mixture nodes  $\mathbf{N}^M$  (nodes where separation or direct transfer to the sink nodes can take place). Each mixture node  $n$  is uniquely associated with a set of present components  $i \in \mathbf{I}_n^M$ . Throughout the manuscript, mixture nodes are labeled by the present components in alphabetical order (e.g.,  $\mathbf{I}_{ABC}^M = \{A, B, C\}$ ).

To enumerate potential sequences from an initial mixture, we utilize the matrix method of Shah and Agrawal (2010), originally developed to find a distillation sequence to separate a  $|\mathbf{I}|$ -component mixture into  $|\mathbf{I}|$  pure product streams. The construction of the matrix for technology  $t \in \mathbf{T}$  requires the ranking of the components,  $\rho_{t,i}$  and an initial mixture. For an element (or equivalently mixture) located in row  $j$  and column  $k$ , the candidate top products are in the same row (i.e.,  $j' = j, k' > k$ ) while the candidate

bottom products are in the same diagonal (i.e.,  $j' = j + k' - k, k > k$ ). To illustrate, in Fig 1 (left panel), we construct matrices for initial mixture ABCD for three technologies with different rankings. The solid/dotted arrows represent top/bottom arcs connecting a mixture to a candidate top/bottom product, and the dashed arrows connect equivalent mixtures in different matrices. One possible separation sequence for mixture ABCD is split AC/BD in technology T2 followed by split A/C in technology T1 and split B/D in technology T3. Note that since component B (or A) is missing from the top (or bottom) product of split AC/BD, mixture AC (or BD) is equivalent to mixture ABC (or ABD) in the matrix for technology T1 (or T2). The equivalent sequence in the aggregated matrix is shown in Fig 1 (right panel). Here, both mixtures AC and BD are assigned unique labels.

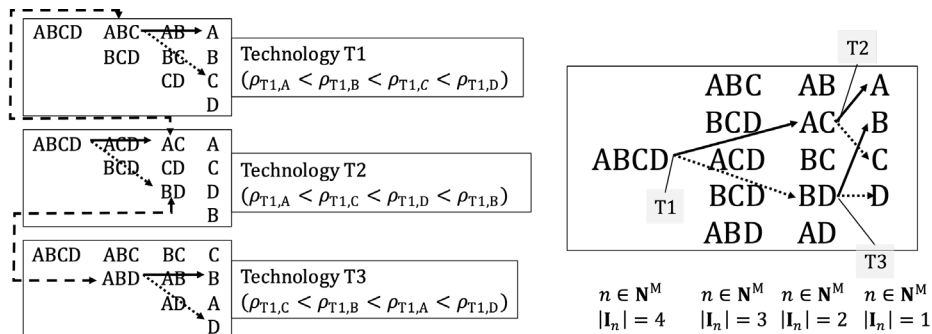


Figure 1: Different technology separation matrices (left) and aggregated matrix (right). Similar to the aggregated matrix, in the final superstructure, mixtures from different technology matrices are combined into unique nodes,  $n \in \mathbf{N}^M$ , (see Fig 2); and the top and bottom arcs are included in set  $\mathbf{A}_t^{\text{TECH},T}/\mathbf{A}_t^{\text{TECH},B}$ . The overall top/bottom arcs are contained in set  $\mathbf{A}^T/\mathbf{A}^B$ , the union of  $\mathbf{A}_t^{\text{TECH},T}/\mathbf{A}_t^{\text{TECH},B}$  over  $t \in \mathbf{T}$ .

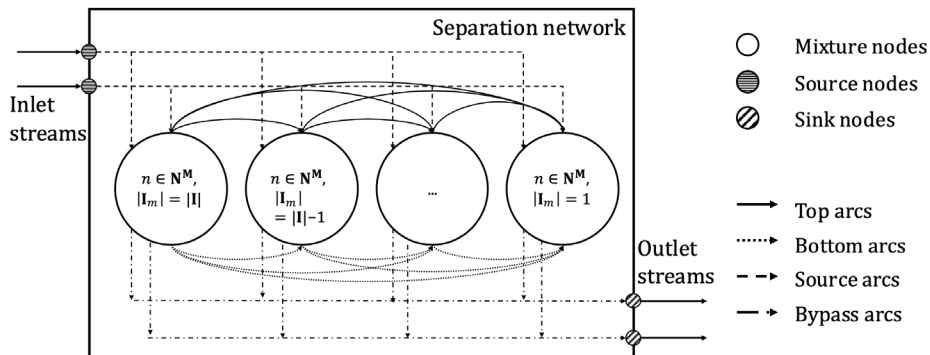


Figure 2: Superstructure representation of the separation network.

### 2.2.1. Mixture nodes

In each mixture node, streams from (1) sources, (2) top products, and (3) bottom products are mixed (see Fig. 3, left panel). A fraction of this mixed stream can be sent to the outlets directly, thereby bypassing unnecessary separations, while the remaining stream is sent to a separation block where technology  $t$  is performed leading to the outgoing top and bottom products. A separation block in this approach represents a set of operations that enables a sharp split separation. In solvent extraction, for instance, a separation block consists of an extraction column and the solvent recovery system (see Fig. 3, right panel). Similarly, in reverse osmosis, we utilize multi-stage membrane units to achieve a sharp

split separation. We also introduce subset  $\mathbf{N}^{M,P} \subset \mathbf{N}^M$  to denote mixture nodes containing only pure component (i.e.,  $|\mathbf{I}_n^M| = 1$ ). In these nodes, no separation will be performed; hence, the mixed stream is directly sent to the outlets.

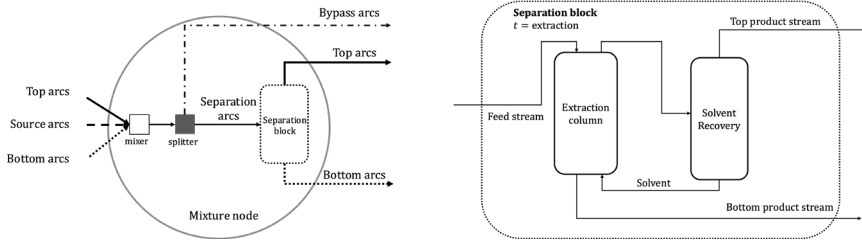


Figure 3: Mixture node (left) and separation block for  $t = \text{extraction}$  (right). Separation block and its streams are relevant only for  $n \in \mathbf{N}^M \setminus \mathbf{N}^{M,P}$ .

### 2.2.2. Connectivity

Using the previously defined sets, we employ  $X_n^M \in \{0,1\}$  to indicate if mixture node  $n$  is active and  $Y_{\hat{n},n}^0 \in \{0,1\}$  to denote the activation of the source arc connecting source node  $\hat{n} \in \mathbf{N}^{SO}$  to mixture node  $n$ . From each source node, there can only be one active arc entering an active mixture node. We employ  $Z_i \in \{0,1\}$  to denote the existence of component  $i$  in the separation network, and using this variable, we deactivate several mixture nodes when a component is not present.

To represent connections between mixture nodes, we employ  $Y_{n,\hat{n}}^{T/B} \in \{0,1\}$  to denote the active top/bottom arc connecting mixture node  $n$  to its top/bottom product mixture node  $\hat{n}$ . The active pairs of top and bottom arcs represent the performed separation splits in the network. We employ  $Y_{n,\tilde{n}}^{BP} \in \{0,1\}$  to denote active arcs from the splitter which bypasses separation and directly goes to sink node  $\tilde{n} \in \mathbf{N}^{SI}$ . For node  $n \in \mathbf{N}^M \setminus \mathbf{N}^{M,P}$ , the remaining mixed stream can undergo separation which produces top and bottom products. The selection of separation technology in this node is done via variable  $Y_{n,t}^C \in \{0,1\}$ . On the other hand, all outgoing arcs from mixture node  $n$  will go to only sink nodes  $\tilde{n} \in \mathbf{N}^{SI}$ .

### 2.2.3. Active streams

Arcs in the superstructure essentially represent streams in the network; hence, we associate each arc (binary variable) with component molar flowrates (continuous variables). The component molar flowrate of the inlet stream from source node  $\hat{n}$ ,  $F_{i,\hat{n}}^0 \in \mathbb{R}^+$ , is disaggregated into  $F_{i,\hat{n},n}^{0,D} \in \mathbb{R}^+$  for each mixture node  $n$ . The value of  $F_{i,\hat{n},n}^{0,D}$  is positive only when its corresponding arc is active (i.e.,  $Y_{\hat{n},n}^0 = 1$ ). We use the inlet streams' flowrates to detect the existence of a component in the network. Moreover, we disaggregate the top and bottom product streams,  $F_{i,n}^T \in \mathbb{R}^+$  and  $F_{i,n}^B \in \mathbb{R}^+$ , into  $F_{i,n,\hat{n}}^{T,D} \in \mathbb{R}^+$  and  $F_{i,n,\hat{n}}^{B,D} \in \mathbb{R}^+$ , respectively, to represent the material transfer from one mixture node to another. Similarly, only flow variables associated with active  $Y_{n,\hat{n}}^{T/B}$  have positive values. Finally, we use split fraction  $K_{n,\tilde{n}}^{BP} \in [0,1]$  to calculate the amount of the mixed stream which bypasses separation and goes directly to the sink node  $\tilde{n}$ ,  $F_{i,n,\tilde{n}}^{BP} \in \mathbb{R}^+$ . Both  $F_{i,n,\tilde{n}}^{BP}$  and  $K_{n,\tilde{n}}^{BP}$  can be deactivated by variable  $Y_{n,\tilde{n}}^{BP}$  via big-M constraints.

### 2.3. Separation block model

Since only sharp separations are considered, we characterize a separation by the selection of a pair of key components in a mixture. The candidate pairs can vary depending on the chosen separation technology. We adopt the terms “light” and “heavy” keys and

generalize them by defining the light/heavy key as the component in the separation block feed stream with the highest/lowest ranking  $\rho_{t,i}$  that is not present in the bottom/top product stream. In each mixture node, we find candidate light and heavy keys according to a specific technology. We introduce  $Y_{i,n}^{\text{LK/HK}} \in \{0,1\}$  to indicate if component  $i$  is the light/heavy key in the separation performed in mixture node  $n$ . Using these binary variables, we write constraints to describe the relationship between the light and heavy keys and the top and bottom arcs. Since we consider only adjacent sharp splits, components with a ranking lower/higher than or equal to the ranking of the light/heavy key will not be present in the bottom/top products.

### 3. Examples

Our approach can be applied to separations where alternative technologies are considered. One such example is xylene isomers separation for which distillation is energy-intensive and thus non-thermal technologies (e.g., reverse osmosis and extraction) are investigated. All MINLP models are solved using BARON 38.3.0 to global optimality (optcr =  $10^{-6}$ ).

#### 3.1. Conventional separation network synthesis

A mixture of 1 kmol/s containing four components (15% A, 30% B, 35% C, 20% D) is to be separated into pure components (Fig 3, left panel). We consider three technologies: distillation, reverse osmosis, and solvent extraction. Components A and C have similar polarities and molecular sizes, which makes separation split between these two components impractical. The objective is to minimize the annualized separation cost:

$$\Phi_1 = \sum_{n,t} \gamma_{n,t}^{\text{FC}} \gamma_{n,t}^{\text{VC}} + \sum_{i,t,n} \gamma_{i,t,n}^{\text{VC}} F_{i,n,t}^{\text{F,LK}} \quad (1)$$

where  $F_{i,n,t}^{\text{F,LK}}$  is the molar flowrate of the stream undergoing separation using technology  $t$  with  $\hat{i}$  as the light key, and  $\gamma_{n,t}^{\text{FC}}/\gamma_{i,t,n}^{\text{VC}}$  is the parameter for fixed/variable cost. The cost parameters can be obtained by doing offline calculations.

The resulting model has 2,567 equations and 1,115 variables (236 binaries) and is solved in 4 s. In the optimal solution (Fig 3, right panel), all three technologies are used. The reverse osmosis block separates B and sends the bottom product (ACD) to the distillation block. The distillation then separates A taking advantage that component B is missing, resulting in an easier separation. Finally, mixture CD is separated using solvent extraction.

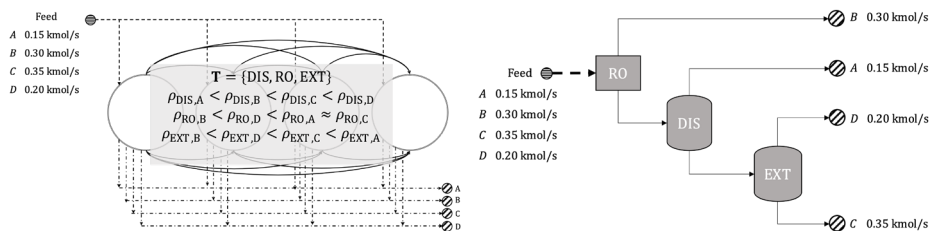


Figure 4: Conventional separation network synthesis problem (left) and its optimal solution (right). DIS: distillation, RO: reverse osmosis, EXT: solvent extraction.

#### 3.2. Integrated reactor-separation network synthesis

Raw materials  $B$  and  $C$  are purchased and fed into the reactor network to undergo one of three reactions:  $2C + B \rightarrow 3A$  (80% conversion);  $C + 2B \rightarrow 3D$  (90% conversion);  $C + B \rightarrow A + D$  (80% conversion) (Fig 4, left panel). The reactor effluent is then fed into the separation network to recover valuable product  $A$  or  $D$ . For each product, the reactor can only produce up to 10 kmol/s. Depending on the chosen reaction, the reactor effluent may contain different combinations of components at different concentrations. Due to the

partial conversions, there will be unconverted reactants that can be recycled. The technologies along with the associated rankings are the same as in the previous example. Our goal is to identify the raw material purchases, reactor selection, and separation network which result in the maximum profit:

$$\Phi_2 = \sum_{i,\tilde{n}} \gamma_{\tilde{n}}^{\text{SELL}} F_{i,\tilde{n}}^{\text{P}} - \left( \sum_i \gamma_i^{\text{RM}} R_i^{\text{RM}} + \sum_{rx} \gamma_{rx}^{\text{RX}} R_{rx}^{\text{FEED}} + \Phi_1 \right) \quad (2)$$

where  $R_i^{\text{RM}}$  denotes the purchased raw material,  $R_{rx}^{\text{FEED}}$  denotes the flowrate of the feed undergoing reaction  $rx$ , and  $\gamma_{\tilde{n}}^{\text{SELL}}/\gamma_i^{\text{RM}}/\gamma_{rx}^{\text{RX}}$  are the product price/raw material price/reaction cost. The influence of feed compositions on the separation cost is considered via the use of cost parameter  $\gamma_{i,t,n,i}^{\text{VC}}$  which depends on the component molar flowrate in the separation block feed stream.

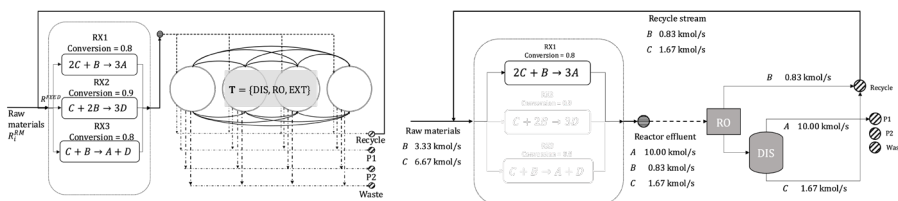


Figure 5: Integrated reactor-separation network synthesis problem (left) and optimal solution (right).

The resulting MINLP model has 2,623 equations and 1,134 variables (239 binaries) and is solved within 6 s. In the solution (Fig 5, right panel), reaction 1 is chosen, and the effluent (mixture ABC) is first separated by reverse osmosis followed by distillation. Unconverted reactants are recovered, mixed, and recycled. Finally, raw materials purchases are such that, given the recycle stream, the reactor feed stream is sufficient to produce 10 kmol/s of A.

#### 4. Conclusion

We proposed a superstructure-based approach to synthesize separation networks with multiple technologies. We utilize rankings based on key properties to generate possible separations using each technology. To account for interactions between different separation technologies, rich interconnections between mixture nodes are considered (e.g., bypass streams). The adopted generalized problem statement allows variable feed to the network, enabling the simultaneous synthesis of separation and reactor networks.

#### References

- R. Agrawal, 2003, Synthesis of multicomponent distillation column configurations, *AIChE journal*, 49, 2, 379-401.
- I. Hechl, Z. Kovács, F. Friedler, L. T. Fan, J. Liu, 2007, Algorithmic synthesis of an optimal separation network comprising separators of different classes, *Chemical Engineering and Processing: Process Intensification*, 46, 7, 656-665.
- L.Kong, C. T. Maravelias, 2020, Expanding the scope of distillation network synthesis using superstructure-based methods, *Computers & Chemical Engineering*, 133, 106650.
- J. Ryu, L. Kong, A. E. P. de Lima, C. T. Maravelias, 2020, A generalized superstructure-based framework for process synthesis, *Computers & Chemical Engineering*, 133, 106653.
- R. W. H. Sargent, K. Gaminibandara, 1976, *Optimum design of plate distillation columns*, in *Optimization in action*, New York: Academic Press, 267-314.
- V. H. Shah, R. Agrawal, 2010, A matrix method for multicomponent distillation sequences, *AIChE journal*, 56, 7, 1759-1775.
- D. S. Sholl, R. P. Lively, 2016, Seven chemical separations to change the world, *Nature*, 532, 7600, 435-437.

# A Reinforcement Learning Development for The Exact Guillotine with Flexibility on Cutting Stock Problem

Jie-Ying Su,<sup>a</sup> Chia-Hsiang Liu,<sup>c</sup> Cian-Shan Syu,<sup>c</sup> Jia-Lin Kang,<sup>b\*</sup> Shi-Shang Jang,<sup>a\*</sup>

<sup>a</sup>*Department of Chemical Engineering, National Tsinghua University, Hsinchu, 300, Taiwan ROC*

<sup>b</sup>*Department of Chemical and Materials Engineering, National Yunlin University of Science and Technology, Yunlin 64002, Taiwan ROC*

<sup>c</sup>*Department of Industrial Engineering and Engineering Management, National Tsinghua University, Hsinchu, 300, Taiwan ROC*

\* [jlakang@yuntech.edu.tw](mailto:jlakang@yuntech.edu.tw); [ssjang@mx.nthu.edu.tw](mailto:ssjang@mx.nthu.edu.tw)

## Abstract

The two-dimensional cutting stock problem (CSP) is critical in several industries. Reinforcement learning (RL) is a novel method to obtain a quality solution of two-dimensional CSP in a short computation time. In this research, we applied a model-free off-policy RL algorithm to an industrial example of exact guillotine two-dimensional CSP, and compared the results with mixed-integer programming (MIP), which is a common traditional mathematical method for optimization. The results showed that RL had a much lower computation time than MIP with a solution closed to optima, and the ability to make a trade-off between waste, inventory level, and back order.

**Keywords:** Reinforcement learning, machine learning, cutting stock problem

## 1. Introduction

The CSP is a problem of minimizing stock material waste while cutting stock material into pieces of required sizes. The two-dimensional CSP has been applied to a variety of industrial applications, such as paper, metal, and glass manufacturing. It is critical for industries to minimize costs, including waste costs, inventory costs, and back-order costs. The two-dimensional CSP is divided into guillotine cutting and non-guillotine cutting. Guillotine cutting can be classified into an exact guillotine and a non-exact guillotine (Andrade et al. (2016)), which increases the complication and difficulty of the two-dimensional CSP.

The computation time of traditional mathematical optimization methods such as MIP grows rapidly as the amount of data increases and the problem becomes more complex. The exact guillotine that caused some items cannot be scheduled, leading to back-order increases when there are no matching items due to the inflexible limitations of traditional MIPs. For addressing the scheduling problem, RL is a promising method that is expected to provide solutions close to optimal within a very short period after training. (Hubbs. et al. (2020)) However, only one previous research is committed to solving CSP by RL, but for one-dimensional. (Pitombeira-Neto et al. (2022))

The purpose of this study is to provide a RL model named Monte Carlo-Deep Deterministic Policy Gradient (MC-DDPG) to solve a two-dimensional exact guillotine CSP in the paper-cutting factory, with the flexibility to switch to the non-exact



guillotine mode for no matching items to reduce back-order and accelerate the scheduling. In the industrial case, there are several bins with different length, width, and weight per area (called base weight) to cut into various items which corresponding to the orders for daily scheduling. Each order may contain one or more items with varying specifications, and total weight of the items finally produced must meet the order's demand. Furthermore, the width requirement of the items is fixed, but the length may be a range, thus the length is also a variable that must be determined. There are several constraints that must be satisfied due to physical and procedural limitations.

This paper is divided into four sections. Section 2 describes the details of our proposed approach. Section 3 and 4 discuss the results and conclusions.

## 2. Method

### 2.1. Reinforcement learning (RL)

RL involves an agent to interact with the environment. The agent determines an *action* according to a given *state*, and the environment returns next state and *reward* (Fig 1). Sets of state, action, reward, and next state at time step  $t$  can be denoted as the following form:  $(s_t, a_t, r_t, s_{t+1})$ . The  $(s_t, a_t, r_t, s_{t+1})$  sequence of an episode with several time steps is called a *trajectory*. The goal of the agent is to maximize the reward, thus RL develops a *policy* for choosing action depending on the state.

RL can be divided into on-policy and off-policy. The main difference between on-policy and off-policy is that whether the policy which is improved (target policy) is same as the policy that is used to select actions (behavior policy). Same for on-policy and not for off-policy. Thus, current data and historical data can both be used for off-policy training.

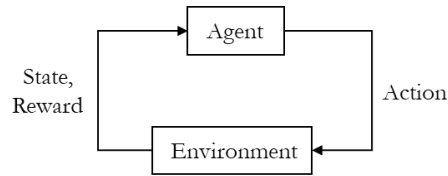


Fig 1. Reinforcement learning system diagram.

### 2.2. Monte Carlo- Deep Deterministic Policy Gradient (MC-DDPG)

DDPG is an off-policy method in RL, composed of two neural networks: actor and critic. Actor determines the actions, and the critic estimates the expected return value which indicates how good are the actions of corresponding states. Actor is trained by maximizing the value output by critic, and the critic is trained by minimizing the difference with the expected return value. (Casas (2017))

Monte Carlo (MC) is a common method for critic updating that estimate the expected return value as:

$$G_t = \sum_{k=t}^{\infty} \gamma^k r_k \quad (1)$$

However, we assume that the current action has a negligible effect on the reward of future actions in our problem, thus we proposed an expected return value function which is explained in the later chapter instead. We store the transition  $(s_t, a_t, r_t, s_{t+1})$  after the end of an episode instead of step by step due to the calculation of  $G_t$ .

### 2.3. Purposed method

The proposed framework of MC-DDPG is composed of two parallel parts: replay buffer management (solid line in Fig 2) and critic and actor training (dashed line in Fig 2). After an episode of interaction between the actor and the environment, the expected return values  $G_t$  were calculated, and the trajectory was stored into the replay buffer. It should be noted that whether an episode ends or not is according to the parameter  $D_{target}$ , which was adjusted when replay buffer stored more than a set amount of data for the training to start, ensuring that there was enough data with action that met the

constraints to improve the policy. Once the training began, a batch size of data was sampled from replay memory at each time step and used to update both actor and critic.

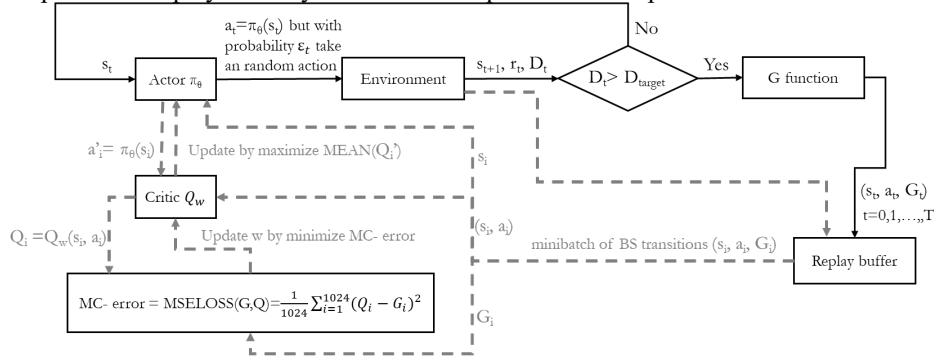


Fig 2. Proposed process diagram.

2.3.1. Environment (Sur et al. (2022))

At the beginning of each episode, we sample M bins and N orders from the database which is provided by a real-world factory, where  $M = \text{DiscUnif}(1,100)$  and  $N = \text{DiscUnif}(1,50)$ .

Fig 3 and Table 2 showed detailed information in state and action provided by actor. Noted that the size of state is fixed to 190 and the size of action is fixed to  $6 \times 18 = 108$ . There have 4 types of action:

I. Slit: Select one of the 15 items to schedule on the current bin in the width direction (called as slit), and also determine the length  $l$  of the item. To determine the length, we define a variable called ratio, then the length of item would be:

$$l = \text{ratio} \times (h_{\max} - h_{\min}) + h_{\min} \tag{2}$$

We may select a ratio of 0, 0.25, 0.5, 0.75, or 1. In addition, we have another option of setting  $l$  equal to the length of the previous slit item without using the ratio.

II. Cut: Finish the scheduling in width direction, and cut the length at the longest length of the slit items after previous cut.

III. Finish: Finish the scheduling of the current bin.

IV. Abandon: Abandon the current bin without scheduling.

The constraints of the system were categorized as physical constraints, process constraints, and environmental constraints. Physical constraints contained the total width of slit items should be no larger than the width of the bin, the total length of cuts should be no larger than the length of the bin, and the item length should be between maximum and minimum length of the item. Process constraints contained the number of slit items should be no larger than 6 and the total produced weight of the order should be no larger than the demand weight. Environmental constraints contained there should be no action if the current bin is to be abandoned, there should be a cut action just before finish action, there should be some slit items before cut action, there should be a

Table 1. Design parameters and variables in process diagram.

parameter/variable	value/statement
$D_{\text{target}}$	$\begin{cases} 2, \text{ while replay buffer size} < 50000 \\ 1, \text{ otherwise} \end{cases}$
$\epsilon_t$	$\begin{cases} 1, \text{ while replay buffer size} < 50000 \\ 0.1, \text{ while } \epsilon_{t-1} \times 0.999 \leq 0.1 \\ \epsilon_{t-1} \times 0.999, \text{ otherwise} \end{cases}$
BS	batch size = 1024
$\pi_\theta$	actor network following parameter matrix $\theta$
$Q_w$	critic network following parameter matrix $w$
T	total time step of the episode

previous slit item if 1 is assigned the length of the previous slit item, the item should exist if it is selected to slit.

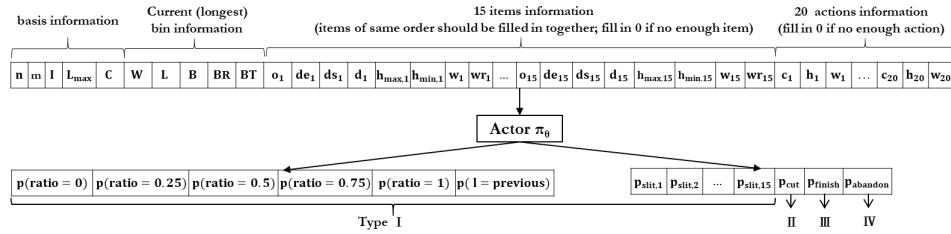


Fig 3. State and action of our proposed method.  
Table 2. Definitions of variables in state.

variable	statement
$n/m$	number of items/bins in state
$I$	number of the slit items after previous cut
$L_{max}$	the maximum length in the slit items after previous cut
$C$	number of cuts in the current bin
$W/L$	residual width/ length of the current bin after $I$ slits/ $C$ cuts
$B/BT$	base/ total weight of the current bin
$BR$	residual weight of the current bin after $C$ cuts
$\sigma_i/de_i$	number/ allowable demand error of corresponding order of $i$ -th item
$ds/d_i$	total produced/ demand weight of corresponding order of $i$ -th item
$h_{max}/h_{min,i}$	maximum/minimum length of $i$ -th item
$w_i$	width of $i$ -th item
$wr_i$	total produced weight of $i$ -th item / $d_i$
$c_k$	number of cuts at $k$ -th action (slit or cut)
$h_k/w_k$	length/width of items chose by $k$ -th action, if $k$ -th action=slit -1, if $k$ -th action=cut

A cut reward would be observed if a cut action is executed, cut reward can be formulated as:

$$r_{cut} = \frac{\text{utilization area of current cut}}{\text{total area of current cut}} \quad (3)$$

A bin reward would be observed when taking a finish action, bin reward can be formulated as:

$$r_{bin} = \frac{\text{utilization area of current bin}}{\text{total area of current bin}} \quad (4)$$

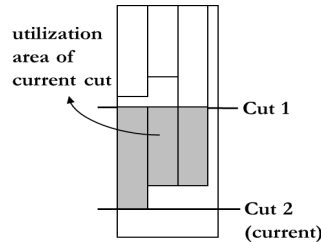


Fig 4. Schematic diagram of the scheduled bin.

A reward of -1 would be observed if the action violates the constraints, and 0 otherwise.

There is a variable  $D_t$  to determine whether the episode ends. The value of  $D_t$  is 2 if all of the bins or items are done, 1 if the action violates the constraints, and 0 otherwise.

The expected return value of the environment is estimated as:

$$G_{j,k} = \bar{r}_{cut,j} \times r_{bin,j}, \bar{r}_{cut,j} = \frac{\sum_z^{Z_{T,j}} L_{z,j} \times r_{cut,z,j}}{\sum_z^{Z_{T,j}} L_{z,j}} \quad (5)$$

where  $G_{j,k}$  is the expected return value for  $k$ -th action in bin  $j$ ;  $\bar{r}_{cut,j}$  is the mean utilization rate of all of the cuts in bin  $j$  based on length;  $Z_{T,j}$  is the total number of cuts in bin  $j$ ;  $L_{z,j}$  is the length of  $z$ -th cut in bin  $j$ ;  $r_{cut,z,j}$  is utilization rate of  $z$ -th cut in bin  $j$ .

### 3. Results and discussion

#### 3.1. Training curve.

The higher the mean bin used ratio the better ability for agent to continuously provide actions that satisfied the constraints. After training, the mean bin used ratio achieved about 0.4 (Fig 5-(a)), and the mean utilization rate achieved about 0.7 which is the mean value in the factory (Fig 5-(b)).

#### 3.2. Testing

4 weeks data was used to evaluate the performance of our proposed RL agent. To efficiently reduce the inventory level and the back-order, we force the agent to select the action with the largest probability and satisfying the constraint, so that the agent would never terminate by taking an action that violates the constraints. We compare the result of inventory level, back-order, and utilization rate with MIP. There is an addition constraint in MIP that the utilization rate of each bin should be larger than 0.7, to ensure the mean utilization rate achieve 0.7. To compare our proposed method with MIP fairly, we also add the constraint to our agent and called as “RL\_0.7”.

Although we forced the agent to take an action, avoiding violate the constraints, there are still some reasons lead to a bin used ratio less than 1, such as all of the orders are satisfied and some bins are abandoned (type IV action). Surprisingly, our RL agent solved most bins with a mean utility rate closed to 0.7 (Fig 6).

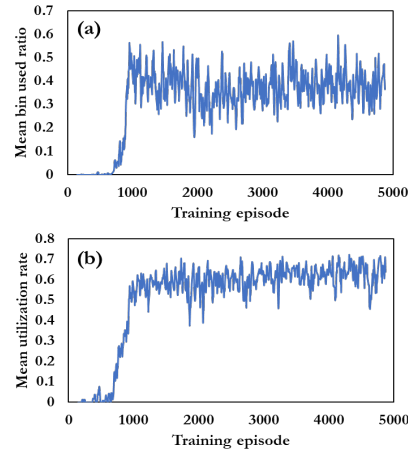


Fig 5. Training curve of (a) mean bin used ratio and (b) mean utilization rate (average of the last 20 episodes) for each training episode.

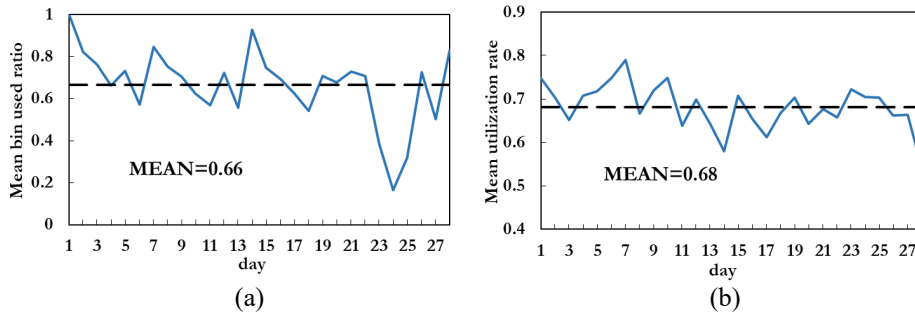


Fig 6. RL testing result of (a) Mean bin used ratio and (b) mean utilization rate for each day.

According to Fig 7-(a) and Table 3, MIP had the highest mean utilization rate, but also the highest computation time. RL\_0.7 showed no utilization rate lower than 0.8 with quite lower computation time than MIP which is important in real-world factory for rescheduling. Although RL had the lowest mean utilization rate, the value was also close to 0.7 which is the mean value in the factory and with the lowest computation time. The inventory level of MIP gradually increased every day. RL and RL\_0.7 both showed a lower inventory level compared to

Table 3. Computation time of testing.

Model	Computation time
MIP	23min 50s
RL	2min 36.4s
RL_0.7	9min 11s

MIP and ability to decrease inventory from a high level. (Fig 7-(b)). Fig 7-(c) showed that RL produced the most orders result in the least back-order. RL\_0.7 had less back-order compared to MIP.

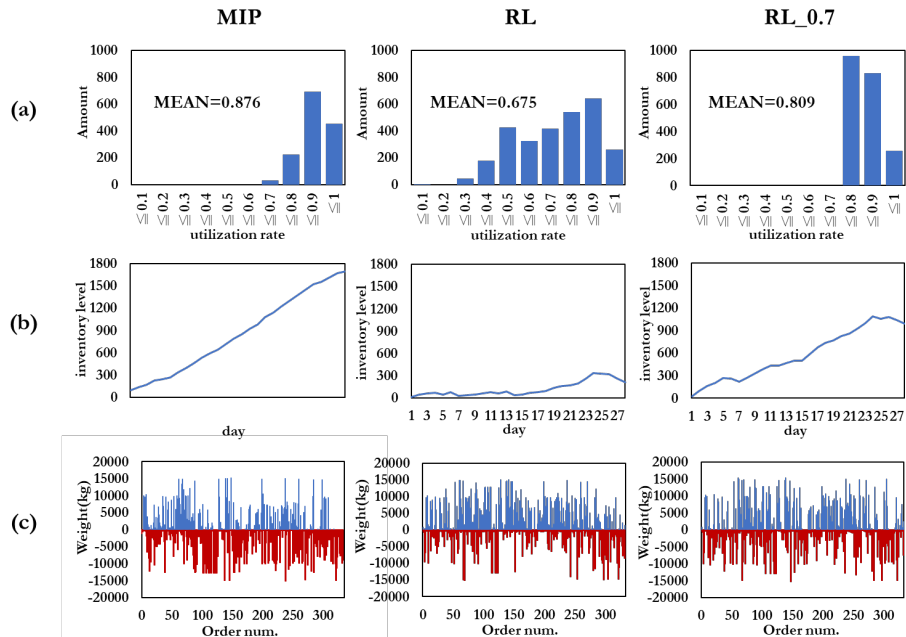


Fig 7. (a) Utilization rate statistics histogram (b) Inventory level for each day (c) Satisfied demand weight (blue) and back-order weight (red) for each order of MIP, RL and RL\_0.7.

#### 4. Conclusion

The proposed RL showed a much better result of inventory level and back-order compared to MIP. Although the mean utilization rate of MIP is the best, the computation time of MIP is also the longest. Furthermore, RL may easily and flexibly make a trade-off between waste, inventory level, and back order, but MIP must add more constraints to take these issues into account, which may make the problem more difficult to solve. The surprising result indicated that RL ability to deal with complex practical industrial problems.

#### References

- Andrade, R., Birgin, E. G., & Morabito, R. (2016). Two-stage two-dimensional guillotine cutting stock problems with usable leftover. *International Transactions in Operational Research*, 23(1-2), 121-145.
- Hubbs, C. D., Li, C., Sahinidis, N. V., Grossmann, I. E., & Wassick, J. M. (2020). A deep reinforcement learning approach for chemical production scheduling. *Computers & Chemical Engineering*, 141, 106982.
- Pitombeira-Neto, A. R., & Murta, A. H. (2022). A reinforcement learning approach to the stochastic cutting stock problem. *EURO Journal on Computational Optimization*, 10, 100027.
- Casas, N. (2017). Deep deterministic policy gradient for urban traffic light control. *arXiv preprint arXiv:1703.09035*.
- Sur, G., Ryu, S. Y., Kim, J., & Lim, H. (2022). A Deep Reinforcement Learning-Based Scheme for Solving Multiple Knapsack Problems. *Applied Sciences*, 12(6), 3068.

# Real-time optimization of a chemical plant with continuous flow reactors via reinforcement learning

Min Wu<sup>a</sup>, Furkan Elmaz<sup>b</sup>, Ulderico Di Caprio<sup>a</sup>, Dries De Clercq<sup>c</sup>, Siegfried Mercelis<sup>b</sup>, Peter Hellinckx<sup>b</sup>, Leen Braeken<sup>a</sup>, Florence Vermeire<sup>d</sup> and M. Enis Leblebici<sup>a,\*</sup>

<sup>a</sup>*Center for Industrial Process Technology, KU Leuven, Agoralaan Building B, 3590 Diepenbeek, Belgium*

<sup>b</sup>*Faculty of Applied Engineering, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium*

<sup>c</sup>*Ajinomoto Bio Pharma Services, Cooppallaan 91, 9230 Wetteren, Belgium*

<sup>d</sup>*Chemical Reactor Engineering and Safety, KU Leuven, Celestijnenlaan 200f, 3001 Leuven, Belgium*

\**muminenis.leblebici@kuleuven.be*

## Abstract

Reinforcement learning (RL) has many new applications in recent years, and its results often exceed human performance, especially in environments where the action space is discrete. However, it is challenging to use RL in the chemical industry, where variables are often continuous and various constraints are complex. This study applies RL with continuous actions to maximize the productivity of a continuous process. The RL agent provides optimal setpoints of flow rates and temperatures while the concentrations of raw materials are changing. Two environments with one and six actions were established after the sensitivity analysis. In the one-action environment, the agents SAC, PPO and A2C showed similar performances, but A2C needed fewer timesteps for training. SAC outperforms PPO and A2C in the environment with six actions. This paper shows the successful RL applications in a continuous process and the high applicability of SAC in both low-dimension and high-dimension environments.

**Keywords:** reinforcement learning, real-time optimization, continuous flow reactors, continuous action space

## 1. Introduction

Reinforcement learning (RL) is a machine learning technique that learns an optimal decision policy based on defined rewards and observations of an environment. Because of its flexibility, it gained popularity in many applications in recent years. The RL results often exceed human performance, especially in the environments (e.g. game) where the action spaces are discrete and the reward is straightforward. However, chemical processes are more difficult to use RL as the common settings should be approached in a continuous way (e.g. flow rates, temperatures). Usually, more trials are done and more time is required compared to discrete spaces at the training stage. Nevertheless, the calculation of the RL agent at the application stage is much faster than classic methods (e.g. linear or nonlinear programming), as no models are used. In the open literature, a RL agent needs 87.7% less calculation time than the nonlinear programming when optimizing a

continuous stirred tank reactor (CSTR) (Powell et al., 2020). Moreover, various constraints in chemical plants are challenging to define in the reward function. Because the RL agent encourages the actions which can lead to higher reward values, the reward function should be designed wisely to meet different constraints. Pèrè et al. (2022) have studied the impact of rewards for the microgrid control. However, the open literature lacks the studies about real-time optimization of continuous flow reactors (CFR) via reinforcement learning and the corresponding designs of reward functions.

This study used RL to optimize the productivity of a chemical plant with six continuous action spaces. The applicability of three RL algorithms, namely Soft actor-critic (SAC) (Haarnoja et al., 2018), proximal policy optimization (PPO) (Schulman et al., 2017), and advantage actor-critic (A2C) (Mnih et al., 2016), in the low- and high-dimension environments was investigated. The following section explains the details of the chemical plant and how RL works.

## 2. Methodology

The plant consists of four identical CFRs in series (see Figure 1). The first three are equipped with jackets, and the last one is air-cooled. The real-time optimization provides six setpoints, which are raw materials A and B flow rates ( $F_{A,SP}$ ,  $F_{B,SP}$ ), feed temperature ( $T_{f,SP}$ ) and three jacket temperatures ( $T_{J1,SP}$ ,  $T_{J2,SP}$ ,  $T_{J3,SP}$ ). The optimization aims to maximize the throughput of product C. A simulation model of this plant for steady states was developed via swarm intelligence (Wu et al., 2022), which was used as the environment in the RL training phase.

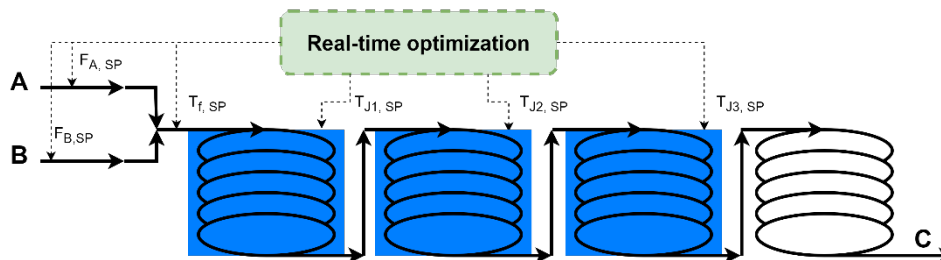


Figure 1: A schematic drawing of the chemical plant. This plant consists of four identical flow reactors. In this system, the feeding raw materials are A and B, and the product is C. The real-time optimization provides six setpoints: A and B flowrates ( $F_{A,SP}$ ,  $F_{B,SP}$ ), feed temperature ( $T_{f,SP}$ ) and three jacket temperatures ( $T_{J1,SP}$ ,  $T_{J2,SP}$ ,  $T_{J3,SP}$ ).

RL consists of two main parts: Agent and Environment. The agent can be structured via different RL algorithms. In this paper, SAC, PPO and A2C are utilized and compared. The simulation environment is only required during the training phase. When the deployment of the trained RL agent is done, the simulation model is not needed any longer (i.e. model-free algorithms). Also, the agent calculation time is usually much less than classic model-based methods because of the high computational complexity of accurate models involved in model-based approaches (Alizadeh et al., 2020). The interaction loop between the agent and the environment in RL is shown in Figure 2.

At each timestep, the agent provides actions after getting observations from the environment. After actions are executed, the environment gives a reward and transits to a new state with new observations. The observations in this paper are set to the changing concentrations of raw materials A and B. The reward value is the normalized outlet

throughput of product C, which is the multiplication of the C concentration and the total flow rate.

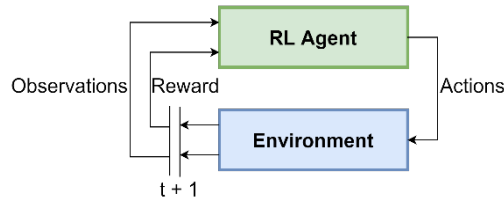


Figure 2: Interactions between the environment and the agent in reinforcement learning. At each time step ( $t$ ), after getting observations from the environment, the agent provides actions. After the actions are executed, the environment gives a reward and transits to a new state.

Because of the environment complexity, the RL training can require more than 100,000 environment evaluation steps. Therefore, to reduce the system complexity, a sensitivity study is conducted to check the importance of six actions to the productivity (Li et al., 2010). The RL agent is first trained for the most important action. Then with similar structures, the agent for six actions is developed. The calculation time and the agent performance are also analyzed to obtain the suitable methodology to develop the RL agent. The environments are encoded in “OpenAI Gym v0.21.0” (Brockman et al., 2016) and the python library “Stable-Baseline3 v1.6.2” is utilized (Raffin et al., 2021).

### 3. Results

A steady-state model for the studied was used to establish the training environment. The agent learning is based on the environment, and the utilized model accuracy is crucial. The model in the paper has high accuracy and fast calculation as demonstrated in the previous work. To reduce the model system complexity, a sensitivity study is conducted to check the importance of six actions.

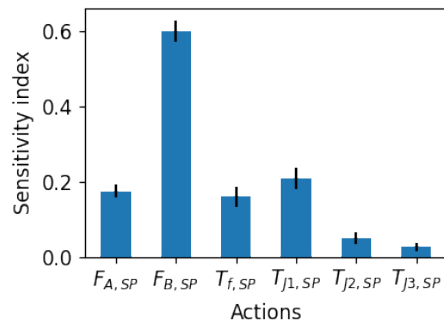


Figure 3: Sensitivity indices of actions to the productivity in the environment. The black vertical lines at the top of the bars show the confidence intervals, and the small ranges of the confidence interval confirm the robustness of the utilized simulation model.

#### 3.1. Sensitivity analysis of available actions

The aim of the global sensitivity analysis is to rank the importance of six available actions with respect to the productivity. The actions with higher sensitivity indices are more important than the others, which means these actions should be considered first if the computation power or the time is limited. Latin Hypercube Sampling generated 1,000 analysis samples within 30% deviations from the baseline setpoints. The baseline setpoints were provided by plant operators and they are 97 kg/h ( $F_{A,SP}$ ), 80 kg/h ( $F_{B,SP}$ ),



20 °C ( $T_{f,SP}$ ), 32.5 °C ( $T_{J1,SP}$ ), 32.5 °C ( $T_{J2,SP}$ ), and 32.5 °C ( $T_{J3,SP}$ ). Moreover, the feeding raw material concentrations ( $C_{f,A}$ ,  $C_{f,B}$ ) are kept constant at 91 wt% and 40 wt%.

The calculated global sensitivity indices of actions are shown in Figure 3. The analysis returned  $F_{B,SP}$  to be the most important action. Physically, this is because B is the raw material in both main and side reactions.  $T_{J1,SP}$  gives a relatively high sensitivity index resulting from that the fact that most reactions are fast and can complete in the first reactor. Therefore,  $F_{B,SP}$  is selected as the most important action in this paper to test three algorithms – SAC, PPO, A2C firstly. After the initial development for the one-action environment, the RL agent for six actions in the high-dimension environment is developed as well.

### 3.2. Performance analysis of agents

The training progresses of SAC, PPO, A2C with one and six actions are illustrated in Figure 4a. The reward value is the normalized productivity, and in all cases the RL training stops when the total timesteps reach 100,000. For the one-action system, all three algorithms reached the maximum reward at the end of the training, and A2C already approached it with less than 20,000 steps. However, A2C got the lowest ending reward in the six-action environment where SAC outperforms the other two algorithms.

Figure 4b gives the required training time to reach the maximum reward value in different approaches. It can be seen that the agents trained in the one-action environment reached the maximum reward faster than the ones trained in the high-dimension environment with six actions, which meets the initial hypotheses. For A2C, the six-action agent requires 25.1% more time than the one-action agent, and that number is 154.5% in SAC. PPO did not even reach the maximum reward value at the end of the training. It shows the high applicability of SAC in both low- and high-dimension environments. A2C is more suitable for the low-dimension environments compared to SAC and PPO, which requires a low amount of training time.

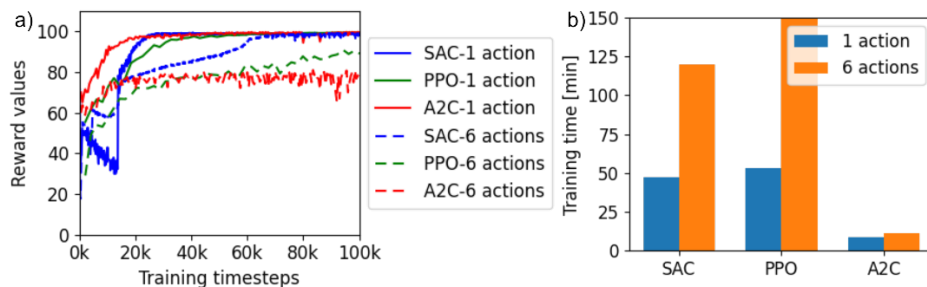


Figure 4: a) The training progress of SAC, PPO, A2C with one and six actions. The reward value is the normalized C productivity. The maximum number of training steps is the same, which is 100,000. b) The required training time of SAC, PPO, A2C with one and six actions to reach the maximum reward values. The used processor is Intel® Xeon® Gold 5220 CPU @2.20GHz.

After completing the training, the developed agent was applied, where the concentrations of the raw materials A and B ( $C_{f,A}$ ,  $C_{f,B}$ ) changed randomly over 50 hours, as shown in Figure 5a. The agents proposed one or six actions to reach a high reward value, and the corresponding reward values are given in Figure 5b and Figure 5c. In the one-action system, SAC, PPO and A2C offered similar performances, which follows the findings in their training progress. Only SAC gave a slightly low reward at the beginning. In the six-action system, SAC outperformed PPO and A2C. In other words, SAC returned a better optimization than PPO and A2C in the high-dimension environment. Figure 6 shows the

behaviors of the six actions from SAC agent. As the concentrations of raw materials changed, the SAC agent gave six setpoints in order to obtain the highest productivity at the new conditions.

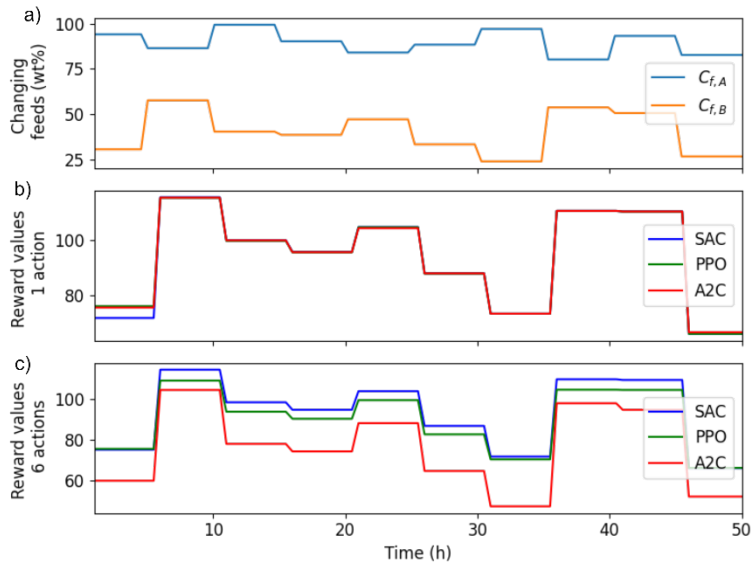


Figure 5: The reward values of RL agents for one and six actions over 50 hours. The upper plot shows how the concentrations (in weight mass percentage wt%) of the raw materials A and B ( $C_{f,A}$ ,  $C_{f,B}$ ) change with time. The RL agents were developed via three algorithms (SAC, PPO, A2C) for one and six actions, respectively.

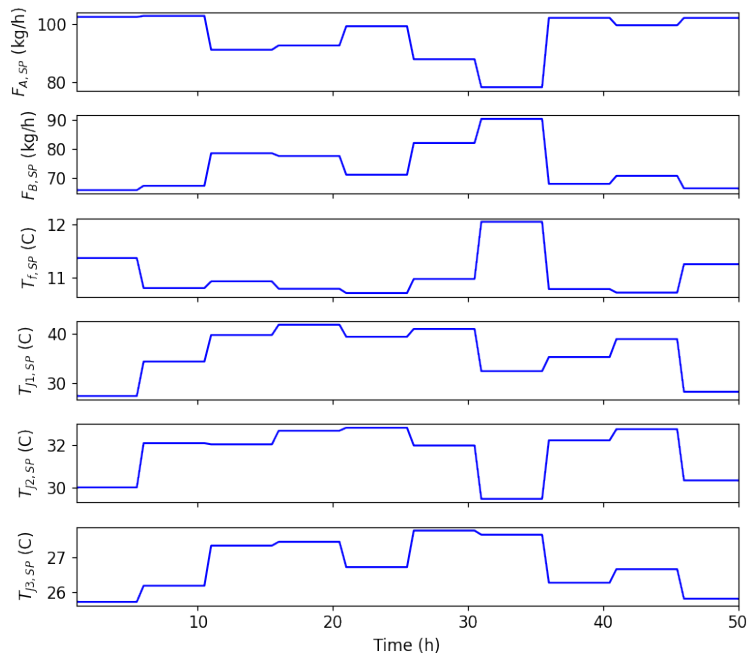


Figure 6: The provided six actions ( $F_{A,SP}$ ,  $F_{B,SP}$ ,  $T_{f,SP}$ ,  $T_{J1,SP}$ ,  $T_{J2,SP}$ ,  $T_{J3,SP}$ ) from SAC agent, which obtains higher reward values than PPO and A2C.

#### 4. Conclusions

This study successfully applied reinforcement learning (RL) in a chemical plant with continuous flow reactors for the real-time optimization. In the plant, the concentrations of the raw materials changed over time. In order to obtain the maximum productivity, six setpoints of flow rates and temperatures were available to be optimized. The developed RL agents provided optimal setpoints based on their learned strategies. The agent learning or training process relies on the interactions between the environment and the agent. Two environments with one and six actions and three RL agent algorithms (soft actor-critic (SAC), proximal policy optimization (PPO), and advantage actor-critic (A2C)) were investigated. The one-action environment was established after the sensitivity analysis. In the one-action environment, SAC, PPO and A2C showed similar performances, but A2C needed less training time. This indicates that A2C is more suitable in the low-dimension environment due to its high efficiency. SAC outperformed PPO and A2C in the environment with six actions, which illustrates the high applicability of SAC in both low-dimension and high-dimension environments. This paper showed the performance of RL agents in the chemical industry. The future work will compare the employed RL algorithms with classic model-based methods. Also, the applied reward functions in both environments for steady states are simply the normalized productivity in this paper. The future work will design new reward functions considering complex constraints, and will study the transit behaviors of the plant from dynamic models.

#### Acknowledgements

The authors acknowledge the funding from VLAIO, DAP<sup>2</sup>CHEM: Real-time data-assisted process development and production in chemical applications (HBC.2020.2455).

#### References

- Alizadeh, R., Allen, J.K., Mistree, F., 2020. Managing computational complexity using surrogate models: a critical review. *Res. Eng. Des.* 31, 275–298.
- Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., Zaremba, W., 2016. OpenAI Gym 1–4.
- Haarnoja, T., Zhou, A., Abbeel, P., Levine, S., 2018. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. *arXiv*.
- Li, G., Rabitz, H., Yelvington, P.E., Oluwole, O.O., Bacon, F., Kolb, C.E., Schoendorf, J., 2010. Global Sensitivity Analysis for Systems with Independent and / or Correlated Inputs. *J. Phys. Chem. A* 6022–6032.
- Mnih, V., Badia, A.P., Mirza, L., Graves, A., Harley, T., Lillicrap, T.P., Silver, D., Kavukcuoglu, K., 2016. Asynchronous methods for deep reinforcement learning. 33rd Int. Conf. Mach. Learn. ICML 2016 4, 2850–2869.
- Père, V., Baillon, F., Milhe, M., Dirion, J.L., 2022. The Impact of Reward Shaping in Reinforcement Learning for Agent-based Microgrid Control. *Comput. Aided Chem. Eng.* 51, 1459–1464.
- Powell, B.K.M., Machalek, D., Quah, T., 2020. Real-time optimization using reinforcement learning. *Comput. Chem. Eng.* 143, 107077.
- Raffin, A., Hill, A., Gleave, A., Kanervisto, A., Ernestus, M., Dormann, N., 2021. Stable-baselines3: Reliable reinforcement learning implementations. *J. Mach. Learn. Res.* 22, 1–8.
- Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O., 2017. Proximal Policy Optimization Algorithms 1–12.
- Wu, M., Di Caprio, U., Elmaz, F., Metten, B., De Clercq, D., Van Der Ha, O., Mercelis, S., Hellinckx, P., Braeken, L., Leblebici, M.E., 2022. A comparative study of swarm intelligence and artificial neural networks applications in modeling complex reaction processes, *Comput. Aided Chem. Eng.* 51, 175-180

# From problem specification to converged flowsheet simulations: Integrated workflow in an industrial context

Mirko Skiborowski<sup>a</sup>, Norbert Asprion<sup>c</sup>, Sergej Blagov<sup>c</sup>, Michael Bortz<sup>b</sup>,

Dennis Manuel Heim<sup>b</sup>, Kai Fabian Kruber<sup>a</sup>, Thulasi Rolland<sup>a</sup>, Tobias Seidel<sup>b</sup>

<sup>a</sup>*Hamburg University of Technology, Institute of Process Systems Engineering, Am Schwarzenberg-Campus 4, 21073 Hamburg, Germany*

<sup>b</sup>*Fraunhofer Institut für Techno- und Wirtschaftsmathematik, Optimierung, Fraunhofer Platz 1, 67663 Kaiserslautern, Germany*

<sup>c</sup>*BASF SE, Process Modeling & Cheminformatics, Carl-Bosch-Str. 38, 67063 Ludwigshafen am Rhein, Germany*

*mirko.skiborowski@tuhh.de*

## Abstract

Although distillation is considered the most mature and computationally easiest to analyze and optimize fluid separation process, support for flowsheet synthesis for the separation of multicomponent azeotropic mixtures is still limited. Commercial simulation software like Aspen Plus mainly offers a graphical analysis of residue curve maps for ternary (sub)systems or an iterative evaluation based on tedious simulation studies. The current work presents an integrative algorithmic approach that pursues an automatic synthesis of distillation processes for azeotropic mixtures, building on a topological analysis of the distillation regions and evaluation of residue curves, pinch lines, and rectification bodies. The derived flowsheet structures and the estimated minimum reflux ratios are the basis for optimization-based initialization of rigorous column models and full-scale simulations of the integrated flowsheets in the industrial in-house simulation tool Chemasim. Thus, the method provides an algorithmic fast-track from a thermodynamic mixture model to a rigorous simulation of automatically derived flowsheets.

**Keywords:** process synthesis, simulation, distillation, conceptual design, azeotropic mixtures.

## 1. Introduction

The conceptual design of separation processes has been subject to continuous research for more than 40 years, with various developments, especially for the design of distillation-based processes. While the separation of zeotropic mixtures is well understood and all basic and complex column sequences can be generated (Madenoor Ramapriya et al., 2018) and evaluated efficiently by systematic and optimization-based methods (Jiang et al., 2019), the synthesis of distillation processes for the separation of azeotropic mixtures appears to present a more challenging problem, which has not been satisfactorily solved yet. The most established tools for the generation of process concepts rely on graphical analysis by means of residue curves or distillation line maps, which represent the vapor-liquid equilibrium characteristics, including the location of azeotropes and the illustration of distillation boundaries. While available in many

flowsheet simulators, such as Aspen Plus, such tools are limited in applicability to ternary and quaternary mixtures. Few algorithmic methods have been proposed to automate this process since the early 2000s. Superstructure optimization methods either consider linear approximations of distillation boundaries and the treatment of azeotropes as pseudo-components (Yang et al., 2011) or, e.g., a limitation to binary azeotropes, which can be separated by extractive distillation (Tsai and Ward, 2020). A more general concept builds on an analysis of split feasibility considering topological distillation regions, as proposed by Rooks et al. (1998), which was the basis for the synthesis methods proposed by Tao et al. (2003) and the work of Thong et al. (2004), who illustrated the potential application to multicomponent mixtures with several azeotropes and distillation regions. Yet, only the method of Waslykiewicz and Castillo (2001), which was developed as part of DISTIL, was shortly available in Aspen Plus after the acquisition of Hyprotech by Aspen Tech. Yet, to our knowledge, there is no documented case of an industrial application of these methods. One potential reason for the lack of adoption is the lack of a consistent connection between the process synthesis methods and a rigorous process simulation, which is the main tool used for process development in industrial practice and which, given sufficiently accurate thermodynamic models, can be seen as a validation of the feasibility of the derived process concept.

In order to bridge this gap, the current study presents an integrated workflow for the automated synthesis of distillation processes for the separation of homogeneous azeotropic mixtures that allows for a direct transfer and validation of the results by means of a converged rigorous steady-state simulation of the derived flowsheets. Flowsheet alternatives are generated using a recently developed automatic process synthesis (APS) tool (Sasi and Skiborowski, 2020), building on a topological analysis, residue curve and pinch line computations, and the rectification body method (Bausa et al., 1998), providing an estimate of the minimum energy demand of the individual separations. The resulting flowsheet configurations, mass balance information, and energy estimates are further processed in a FlowsheetExplorer, which determines initial composition profiles and actual reflux ratios for individual columns with finite stage numbers that are finally transferred to the BASF in-house flowsheet simulator Chemasim, where the full process flowsheet is generated and optimized providing a converged rigorous flowsheet simulation, which satisfies the product specifications. The method is demonstrated for the separation of a well-known quaternary mixture of acetone-chloroform-benzene and toluene, which exhibits a strongly curved distillation boundary. The results illustrate the potential benefits of an algorithmic flowsheet synthesis and the possible speed-up of the conceptual design process.

## 2. Method

The proposed integrated workflow that allows for a direct application in an industrial context is schematically illustrated in Figure 1. It consists of two main steps for an automated process flowsheet synthesis (APS Tool, Fig. 1 left) and an optimization-based column and flowsheet initialization (FlowsheetExplorer, Fig. 1 right), which are pursued in a sequence of steps further outlined in the following subsections. The combination allows for the derivation of converged flowsheet simulations on the basis of rigorous equilibrium tray models of the individual distillation columns, which satisfy the desired product specifications, starting from a definition of the separation problem and the thermodynamic property models.

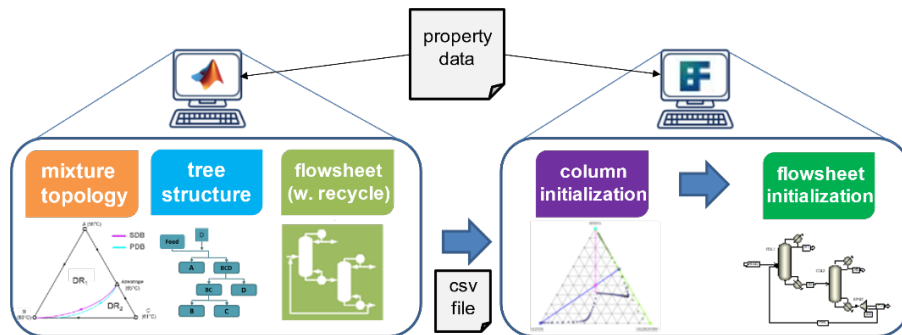


Figure 1: Illustration of the individual steps of the topology-based process synthesis (APS Tool) and optimization-based initialization and flowsheet simulation (FlowsheetExplorer).

### 2.1. Automated process synthesis

The APS tool aims at the generation of feasible process configurations based on the specification of a thermodynamic model for the phase equilibria of the considered mixture and the specification of the feed and desired products. In the first step, the topology of the mixture is analyzed, i.e., the computation of azeotropes and the characterization of all singular points regarding their stability, as well as the identification of the topological distillation regions (DR), based on the concept of the adjacency and reachability matrices (Rooks et al., 1998). The knowledge of the topological distillation regions is the foundation on which process variants are derived in the subsequent step, building on an algorithmic split feasibility check, which combines repeated computations of residue curves (total reflux distillation) and pinch lines (reversible distillation) to check if the products are located in the same DR for at least one mode of limiting operation. Further application of the rectification body method (RBM) (Bausa et al., 1998) evaluates the minimum reflux ratio of individual splits, given an intersection of the linear approximation of the potential manifolds of the individual column section profiles. Based on the derived tree structure, potential recycle streams and locations are identified following the ideas of Tao et al. (2003), closing the respective recycle loops in an iterative approach with additional RBM evaluations. The final flowsheets are thereby characterized by the respective mass balances for the individual columns and the estimates of the required heat duties, which not only allow for an initial ranking of alternative flowsheet configurations, but also provide essential information for the initialization of more rigorous flowsheet simulations based on consistent equilibrium tray models. This information is further transferred to the FlowsheetExplorer for the subsequent initialization and flowsheet optimization in a specific CSV file format. Refer to the article of Sasi and Skiborowski (2020) for further details on the process synthesis method, including potential pressure modifications. The MATLAB tool can be evaluated via the open GitLab project (APS tool), including respective examples.

### 2.2. Optimization-based column and flowsheet initialization

The second part of the integrated workflow aims at the initialization and subsequent optimization of rigorous flowsheet models in BASF's in-house flowsheet simulation software Chemasim. The results of the variant generation provide the necessary connectivity and mass balance information to build the individual flowsheet models for the derived candidates in the FlowsheetExplorer.

Since the minimum heat requirements derived from the RBM build on the computation of pinch points alone, they do not provide continuous composition and temperature profiles for the individual columns. To derive appropriate stage numbers for each column

section and to provide reasonable estimates for the composition and temperature profiles, individual optimization problems are first solved for all columns. These optimizations aim to find a composition profile such that the compositions of the bottom and top streams of the column are close to the compositions found by the APS tool. The optimizations are constructed as described by Seidel et al. (2020). The profiles are partially calculated using tray-to-tray calculations starting from the top and bottom streams and calculated towards the feed stage. The equations of the feed stage are included as constraints in the optimization problem. This means that after the optimization has converged, all equations for the rigorous simulation of the column are satisfied. For the optimization of each column, a constant molar overflow is assumed, and the optimizer can vary the number of stages and the reflux ratio for each individual column.

The goal of the individual column optimizations is to find good initial profiles quickly and robustly. However, in the transfer, the compositions calculated by the APS tool and this initialization procedure will not match perfectly, such that the feed streams of columns in the flowsheet will potentially change, and the profiles calculated by a rigorous flowsheet simulation may deviate due to the connection of the different columns. In order to further improve the flowsheet initialization, a second optimization can be started in the FlowsheetExplorer to find an initial profile that matches the concentrations, but this time with respect to the connected closed-loop flowsheet. Since the first optimization already found good initial profiles, this second optimization can be done within Chemasim. The FlowsheetExplorer automatically sets up and executes this second optimization, which not only verifies the feasibility of the derived process configurations, but also allows for further derivation of more informative cost estimates for the different process variants.

### 3. Case Study

As a representative case study, the separation of the four-component mixture of acetone, chloroform, benzene, and toluene (ACBT) with an equimolar composition and a flow rate of 10 mol/s, as investigated in the article of Liu et al. (2004), is considered. While the algorithmic approach of Liu et al. (2004) was able to derive a variety of flowsheets with at least four distillation columns, Kraemer et al. (2009) proposed three additional process configurations with only three distillation columns by means of graphical analysis of the pitchfork distillation boundary (PDB), which were shown to be favorable in terms of the minimum energy demand estimates of the RBM. The APS tool automatically derives 12 process configurations as feasible flowsheets, including the three configurations with only three distillation columns proposed by Kraemer et al. (2009). This becomes feasible due to the consideration of both limiting operation modes by computation of residue curves and pinch lines. The computations take less than five minutes of computational time on a standard desktop computer. Figure 2 summarizes the respective minimum heat duties estimated by means of the RBM in the APS tool (right) and illustrates the four most promising process variants (left), which include all three of the previously mentioned three-column configurations. All remaining process variants consist of either four or five column flowsheets, with an increase in the accumulated heat requirement of up to 135%.

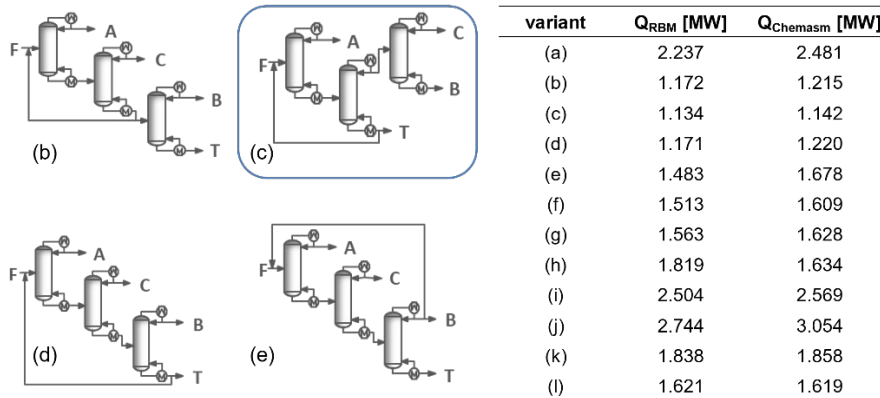


Figure 2: Illustration of the four most promising process flowsheets for the ACBT separation (left) and overview of the cumulative heat requirement of all 12 derived process variants (right).

Based on the results of the APS tool, adequate tray numbers and column profiles are first determined for each individual column and then for the entire flowsheets by means of the FlowsheetExplorer, which takes around 16 minutes of computational time. Converged closed-loop flowsheet simulations were obtained for all 12 process variants. Figure 3 exemplarily illustrates the composition profiles of the individual components in the respective columns 1-3 for process configuration (c), which provides the lowest energy requirements. The resulting flowsheets and the obtained starting profiles were passed over to Chemasim to pursue a heat minimization with rigorous equilibrium tray models and product purities of 99 mol% set as constraints. The optimization of all 12 flowsheets takes around 52 minutes. The excellent agreement of the resulting accumulated heat duties of the rigorous flowsheet optimization and the initial shortcut evaluation are illustrated by the side-by-side comparison in Figure 2 (right).

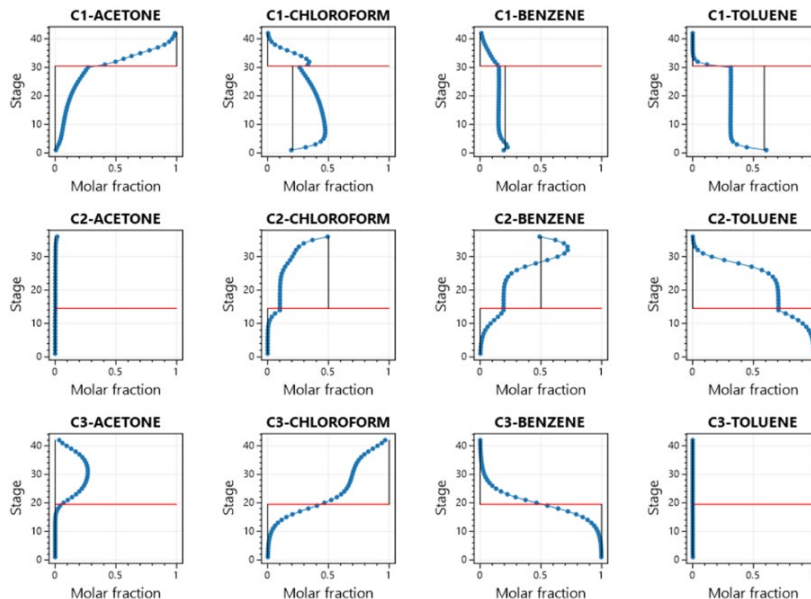


Figure 3: Illustration of the resulting composition profiles of the three columns in the flowsheet variant (c) after initialization and optimization in Chemasim.



Especially for the three best variants, a mean deviation of only 3% is reached. Overall, even larger deviations, which can be explained by the linear approximations of the RBM and variations in the recycle compositions, are limited (< 15%) and do not significantly affect the ranking of the variants. Furthermore, it can be noted that the results of the rigorous flowsheet evaluation of the best flowsheet variant fit very well with the optimization results reported by Kraemer et al. (2009).

#### 4. Conclusion and outlook

The current article illustrates how a set of multi-purpose tools for the conceptual design of distillation-based separation processes can be integrated for a fully automated synthesis and rigorous simulation of distillation processes to separate homogeneous azeotropic mixtures. The basis for the integrative approach is a specification of the separation problem and a consistently applied thermodynamic model, which is the fundamental requirement for any meaningful flowsheet simulation. Since the applied topological synthesis tool of Sasi and Skiborowski (2020), which also allows for the consideration of pressure variations, does not build on constant relative volatilities, a graphical interpretation or a geometric approximation of the distillation boundaries, it is practically not limited with respect to the number of components or complexity of the mixture. The derived flowsheet structures and the estimated minimum reflux ratios build the basis for the subsequent optimization-based initialization of rigorous equilibrium tray models and the final simulation of the integrated closed-loop flowsheet in Chemasim.

Future investigation will not only evaluate the application to much larger multicomponent mixtures, but also extend the design space to consider liquid-liquid separations, as in heteroazeotropic distillation (Sasi and Skiborowski, 2020) and liquid-liquid extraction (Sasi and Skiborowski, 2022), as well as extractive distillation. Furthermore, the possibility for process intensification and energy integration may also be explored in an automated form by further extending the synthesis method with adapted shortcut-screening methods (Skiborowski, 2018). Also, to speed up the closure of multiple recycle streams simultaneously, we are currently exploring the option to adaptively determine points on the distillation boundary and formulate a linear optimization problem.

#### References

- APS tool GitLab project, 2023, <https://collaborating.tuhh.de/v-4/psi-public/aps-tool>.
- J. Bausa, R.v. Watzdorf, W. Marquardt, 1998, *AIChE J.*, 44(10), 2181-2198
- Z. Jiang, T.J. Mathew, H. Zhang, J. Huff, U. Nallasivam, M. Tawarmalani, R. Agrawal, 2019, *Comput. Chem. Eng.*, 126, 249-262.
- K. Kraemer, S. Kossack, W. Marquardt, 2009, *Ind. Eng. Chem. Res.*, 48, 6749-6764.
- G. Liu, M. Jobson, R. Smith, O. Wahnschafft, 2004, *Ind. Eng. Chem. Res.*, 43, 3908-3923.
- G. Madenoor Ramapriya, M. Tawarmalani, R. Agrawal, 2018, *AIChE J.*, 64, 649-659.
- R.E. Rooks, V. Julka, M.F. Doherty, M.F. Malone, 1998, *AIChE J.*, 44, 1382-1391.
- T. Sasi, M. Skiborowski, 2020, *Ind. Eng. Chem. Res.*, 59(47), 20816-20835
- T. Sasi, M. Skiborowski, 2020, *Comp. Aid. Chem. Eng.*, 48, 1009-1014
- T. Sasi, M. Skiborowski, 2022, *Comp. Aid. Chem. Eng.*, 51, 751-756
- M. Skiborowski, 2018, *Chem. Eng. Trans.*, 69, 199-204.
- L. Tao, M.F. Malone, M. F. Doherty, 2003, *Ind. Eng. Chem. Res.*, 42, 1783-1794.
- D. Thong, G. Liu, M. Jobson, R. Smith, 2004, *Chem. Eng. Proc.*, 43, 239-250.
- T. Tsai, J.D Ward, 2020, *Ind. Eng. Chem. Res.*, 59, 17964-17976
- S.K. Wasylkiewicz, F.J.L. Castillo, 2001, *Comput. Aided Chem. Eng.*, 9, 591-596.
- X. Yang, H.G. Dong, I.E. Grossmann, 2011, *AIChE J.*, 58, 1487-1502.
- T. Seidel, A. Hoffmann, M. Bortz, A. Scherrer, J. Burger, N. Asprion, K.-H. Küfer, H. Hasse, 2020, *Chem. Eng. Sci. X*, 7, 100063

# A Hybrid Modelling Framework for Dynamic Modelling of Bioprocesses

Haiting Wang<sup>a</sup>, Cleo Kontoravdi<sup>a</sup>, Ehecatl Antonio del Rio Chanona<sup>a</sup>

<sup>a</sup> *Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, London, United Kingdom SW72AZ*

## Abstract

One of the main hurdles in the computer-aided design and optimization of industrial bioprocesses is the limited capability of models to accurately represent biosystems. On one hand, mechanistic models have drawbacks in terms of expressing all potential biomechanisms. On the other hand, data-driven models have limited extrapolation ability. Thus, hybrid models may represent the best of both worlds. However, it is challenging to build a hybrid model that can accurately balance its data-driven and first principles components, hence allowing for extrapolation ability and accuracy.

In this work, we propose a framework for the development of hybrid models that includes a mechanistic backbone for its construction and judiciously adds the data-driven components. Statistical methods e.g., Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC) and Hannan Quinn Criterion (HQC), are employed to choose the statistically best hybrid model structure. The proposed framework is tested on a simplified microalgae cultivation case study and shows good prediction capabilities under different noise levels, especially by applying BIC for hybrid model selection.

**Keywords:** Model selection, Hybrid modelling, Bioprocess modelling, Information criteria.

## 1. Introduction

Industrial bioprocesses have experienced fast development in many fields. However, most bioprocesses suffer from low production yields, and experimental optimization of these systems is both resource-intensive and time-consuming. Therefore, building mathematical models which can accurately describe the dynamic behavior of biosystems is crucial for process design, scale-up and optimization. Constructing a mechanistic (first principles) model based on physical and biological knowledge has been widely investigated with very good success, but also room for improvement (Tsopanoglou and Jiménez del Val, 2021). Because investigating and quantifying cellular mechanisms is very challenging, especially considering the intracellular metabolic reaction networks of living cells and how they are influenced by variations in operating conditions. To address these challenges, data-driven models have attracted a lot of attention given that they do not require physical knowledge. However, the biggest drawback of data-driven modelling is the requirement for high quantity and quality of measured data for model training, and their inability to extrapolate, which is critical for optimization and control.

To take advantage of both mechanistic and data-driven models while avoiding their drawbacks, hybrid models are considered good potential solutions. They have mechanistic backbones to promote extrapolation performance. At the same time, complicated biomechanisms are modelled with properly selected data-driven components

that avoid heavy mechanistic expressions. Therefore, when training a hybrid model with the same dataset, it has better extrapolation and interpretation ability compared to the data-driven model, and it is easier to construct than a mechanistic model.

However, to build a good performing hybrid model, it is important to balance the weight of the data-driven sub-model and the mechanistic backbone. To achieve this, empirical model selection methods such as the L1 and L2 regularizations have been widely applied (Tibshirani, 1996). However, these methods are non-statistical, and the model selection results are not consistent. Alternatively, in this work, statistically-derived information criteria are applied for model selection, specifically, the Akaike Information Criterion (AIC) (Stoica and Selen, 2004), Bayesian Information Criterion (BIC) (Stoica and Selen, 2004), and Hannan Quinn Criterion (HQC) (Hannan and Quinn, 1979).

AIC is derived from information theory, which defines the best model as the one that can minimize the information divergence between the observed data and simulation results. The equation of AIC can be represented as below:

$$J = 2n_{\theta} + 2NLL(\theta) \quad (1)$$

Where  $NLL(\theta)$  is the negative log-likelihood (NLL) function based on parameter  $\theta \in \mathbb{R}^{n_{\theta}}$ , which represents the goodness of fit between the model prediction and the measured data, and the total number of parameters in the model candidate,  $n_{\theta}$ , is penalized.

On the other hand, BIC is derived from the Bayes' rule, which assumes that a 'true' model exists among all candidate models, and the best model is the one which has the maximum posterior probability value to be the best model (Stoica and Selen, 2004):

$$J = 2n_{\theta} \ln(n_d) + 2NLL(\theta) \quad (2)$$

It can be seen from Eq.2 that BIC penalises model complexity by considering not only  $n_{\theta}$ , but also  $n_d$  which is the number of data points in the training set. Like the BIC, the HQC is derived to select the 'true' model instead of the one that minimizes the information loss from data, and the expression can be shown below:

$$J = 2n_{\theta} \ln(\ln(n_d)) + 2NLL(\theta) \quad (3)$$

In terms of penalty terms, the HQC also considers the impact of the training dataset size, while it penalises the model complexity less strongly than BIC. By using these information criteria for model selection, the best model is chosen by minimizing the value of these information criteria.

Because these information criteria are derived from different statistical perspectives, they show different model selection behaviors. It can be found from the literature that BIC tends to consistently choose the 'true' model and penalizes model candidates with higher model complexity more than the 'true' model, while AIC prefers to select the model with better fitting performance even if it is more complex than the 'true' model (Vrieze, 2012). In this work, we propose a hybrid modelling framework with the application of statistical information criteria to efficiently construct hybrid models with good extrapolation ability for the dynamic modelling of bioprocesses.

## 2. Hybrid modelling framework

The proposed hybrid modelling framework comprises 4 steps: mechanistic backbone construction, data-driven sub-modelling, hybrid model integration and model validation as shown in Figure 1. To find the best hybrid model that balances the fitting and extrapolation performance, AIC, BIC and HQC are used for statistical model selection and embedded into the framework. During this step, model structure selection and parameter estimation are conducted. In the end, the parameter uncertainty analysis and robustness test are carried out on the constructed hybrid model.

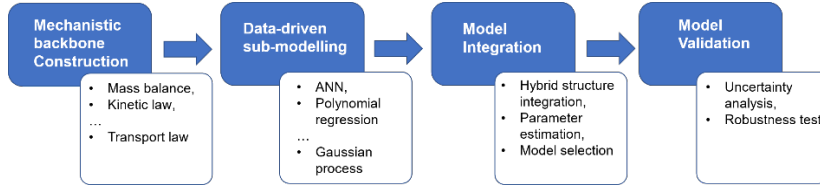


Figure 1 Illustration of the hybrid modelling construction framework.

### 3. Case study –Microalgae Cultivation

#### 3.1. Microalgae cultivation system

A simplified microalgae cultivation process under fed-batch operating conditions was used as a case study. This model is a simplification from the model which describes the biomass growth and nitrate consumption in a bioreactor (del Rio-Chanona et al., 2015). A mechanistic model is introduced to describe the dynamic change of biomass concentration  $C_X$  and nitrate concentration  $C_N$  with a constant feeding rate  $F_{in}N_{feed}$ :

$$\frac{dC_X}{dt} = \mu_M C_X \frac{C_N}{C_N + K_N} - \mu_d C_X^2 \quad (4)$$

$$\frac{dC_N}{dt} = -Y_{N/X} \mu_M C_X \frac{C_N}{C_N + K_N} + F_{in}N_{feed} \quad (5)$$

Where  $\mu_M$  is the specific biomass growth rate,  $\mu_d$  is the biomass death rate,  $Y_{N/X}$  is the specific nitrate uptake ratio, and  $K_N$  is the half-velocity coefficient.

#### 3.2. In-silico data generation

The mechanistic model above is used for in-silico data generation for model training. In this work, three computational experiments are generated with different initial concentrations for  $C_X$  and  $C_N$ . These computational experiments simulate the dynamic change of  $C_X$  and  $C_N$  in 384 hours of operational time, and the generated data for  $C_X$  and  $C_N$  is collected in 10-hour intervals. Different levels of Gaussian noise are applied to the training datasets to mimic real-life measurements. Noise levels of  $C_X$  and  $C_N$  are certain percentages of their mean values respectively. At the same time, three validation datasets are generated with different initial concentrations, but the same data collection intervals compared to training datasets.

#### 3.3. Hybrid model construction

Following the first step in the framework, the mechanistic backbone is designed first. By applying simple mass balance on  $C_X$  and  $C_N$ , the mechanistic backbone is shown below:

$$\frac{dC_X}{dt} = \mu_m C_X - \mu_d C_X^2 \quad (6)$$

$$\frac{dC_N}{dt} = -Y_{N/X} \mu_m C_X + F_{in}N_{feed} \quad (7)$$

Among all variables in the mechanistic backbone, it is challenging to accurately describe the biomass growth rate  $\mu_m$  dependence on metabolite concentrations and bioreactor operating conditions. Thus, to avoid the complex mechanistic expression, we choose a data-driven method to represent it in terms of the variable it depends on.

In this work, the data-driven sub-model aims to determine the relationship between  $\mu_m$ ,  $C_X$  and  $C_N$ . Third-order polynomial regression is chosen for this purpose as shown below:

$$\mu_m = p_0 + y_0 p_1 C_{X\_norm} + y_1 p_2 C_{N\_norm} + y_3 p_4 C_{X\_norm}^2 + \dots + y_8 p_9 C_{N\_norm}^3 \quad (8)$$

$$C_{X\_norm} = \frac{C_X}{C_{Xmax}} \quad (9)$$

$$C_{N_{norm}} = \frac{C_N}{C_{N_{max}}} \quad (10)$$

Where  $p_i$  are coefficients of each polynomial term and  $y_i$  are binary variables. Binary variables are used to describe whether their corresponding polynomial term is selected or not when carrying out statistic model selection. Moreover,  $C_{X_{max}}$  and  $C_{N_{max}}$  are the maximum measured value of  $C_X$  and  $C_N$  respectively and they are introduced to normalize the dependencies. Other data-driven methods can also be applied instead of polynomial regression, especially for highly nonlinear systems, such as Artificial Neural Network (ANN), Partial Least Square regression (PLS) and Gaussian Process (GP).

By integrating the mechanistic backbone with the polynomial expression of  $\mu_m$  we can get the basic structure of the hybrid model. Then, to find the best structure of the hybrid model, statistical model selections are applied, which is the core of the model integration step. During this step, parameter estimation and model selection can be carried out by solving a MINLP problem to obtain the best hybrid model. Finally, the best hybrid model is validated using 3 sets of validation data as described in section 3.2.

#### 4. Results and Discussion

First, a simple task is carried out to demonstrate the concept of statistical selection of the best-performing hybrid model. In this case, we assume the ideal behavior of each state variable is described by the mechanistic model that we used for data generation, thus the best hybrid model is the one that gives the best agreement with the mechanistic model even if it is trained by noisy data. The agreement is judged based on the mean square error from validation datasets. Then, to further test the proposed hybrid model construction framework, we assess its performance when it is trained with datasets with different noise levels, from 5% of the mean of  $C_X$  and  $C_N$  to 20%.

##### 4.1. Proof of concept

To empirically study whether the information criteria can effectively select the best hybrid model structure that can avoid overfitting and achieve the best extrapolation capability, we generate 9 hybrid model structures by adding polynomial terms in the sequence in Eq.8. The existence of polynomial terms is represented by setting their corresponding  $y_i$  to 1. Then, each hybrid model is trained with a dataset with 5% noise. All three information criteria and NLL values are estimated for each structure. Moreover, each model is validated using validation datasets and the results are shown in Figure 2:

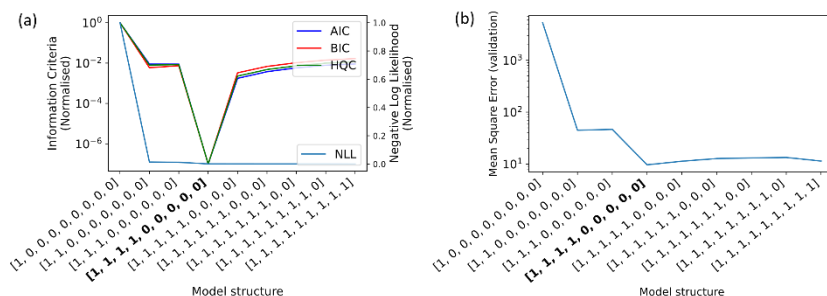


Figure 2. Training and validation results of 9 hybrid model structures. (a) Estimation of AIC, BIC, HQC and NLL for each hybrid model structure. (b) Validation results of each hybrid model structure.

As shown in Figure 2, the increase of polynomial terms in sequence of the hybrid model decreases the corresponding NLL values, which indicates an improvement in the fitting. However, the improvement becomes much less significant after a certain stage, indicating the occurrence of overfitting. By calculating the value for AIC, BIC and HQC, we can see from Figure 2(a) that all three information criteria start to penalize the over-complexed structure when the model has more than the first 4 polynomial terms. The validation results show a strong agreement on the best hybrid model structure selected by information criteria among these 9 candidates, highlighting the effectiveness of statistical model selection methods.

4.2. Robustness test under different noise levels

As mentioned in Section 1, AIC, BIC and HQC have different statistical properties. Thus, we need to evaluate their robustness, especially under different noise levels since low-quality data is one of the most common problems for bioprocess modelling. Therefore, 4 training datasets are generated with the noise level of 5%, 10%, 15% and 20% respectively. Also, under each noise level, 10 datasets are generated and used for hybrid model construction to quantify result uncertainties.

The total number of polynomial terms ( $n_p$ ) in Eq.8 can vary from 1 to 8. Therefore, we can first find 7 potential best hybrid models with certain  $n_p$  value (models with  $n_p = 8$  are significantly overfitted and not presented here). Then the corresponding information criteria's values and the validation results can be estimated for each model with a certain standard deviation as shown in Figure 3.

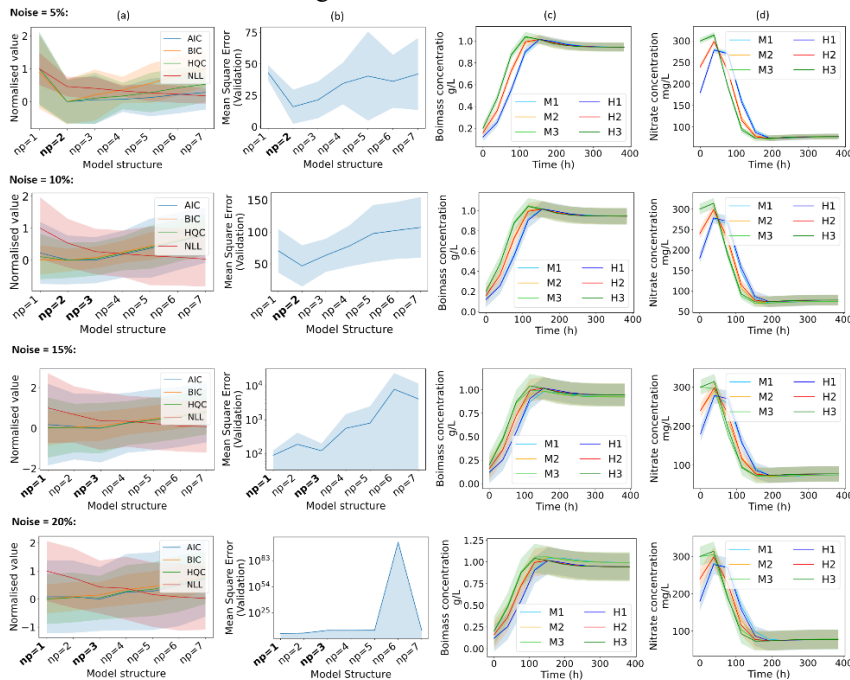


Figure 3. Training and validation results of all potential hybrid model structures under different noise levels. (a) Estimation of AIC, BIC, HQC and NLL for each hybrid model structure. (b) Validation of each hybrid model structure. (c) and (d) Validation results of biomass and nitrate concentration prediction given by the best hybrid model respectively. (M1, 2 & 3 represent validation datasets from the mechanistic model, while H1, 2 & 3 represents validation results given by the hybrid model)

It can be seen from the validation results in Figure 3(b) that the best hybrid model structure varies with the increase of noise level because the model is easier to overfit when data is highly noisy. The best hybrid model is the one with  $n_p = 2$  under 5% and 10% noise while it becomes the one with  $n_p = 1$  when noise increases to 15% and 20%.

Statistical model selection results are shown in Figure 3(a) at each noise level. It can be seen that all information criteria choose the same model structure under 5% noise. However, they show different model selection results under different noise levels. When noise increases to 10%, BIC and HQC still successfully select the best hybrid model structure, while AIC chooses a model structure with worse performance ( $n_p = 3$ ). Under a 15% noise level, both HQC and AIC select the wrong model structure ( $n_p = 3$ ) while BIC still selects the correct one. By increasing the noise level to 20% only BIC and HQC can find the best model structure.

In summary, our results suggest that BIC is the most robust information criterion which shows stable performance on choosing the best model structure, while HQC and AIC are not as reliable since they put more weight on getting the better fitting performance and penalise model structure complexity weaker. Additionally, it can be seen from Figure 3 (c) and (d) that by using information criteria for model selection, even if the hybrid model is trained under strong noise, the model can maintain good extrapolation ability and shows good agreement with the purely mechanistic model.

## 5. Conclusion

In this work, a hybrid model construction framework was proposed and tested on a biomass cultivation case study. The results show that by applying this framework, and especially by using statistical information criteria, the best hybrid model structure can be efficiently selected. Among all three information criteria, BIC shows the most consistent performance against the increase of noise level from the training dataset. In the future, we will apply this framework to more complex case studies such as the photo-production process, and the polynomial regression will be replaced with a more complex data-driven model such as ANN. Finally, an uncertainty analysis of the constructed hybrid model will be carried out.

## References

- DEL RIO-CHANONA, E. A., DECHATIWONGSE, P., ZHANG, D., MAITLAND, G. C., HELLGARDT, K., ARELLANO-GARCIA, H. & VASSILIADIS, V. S. 2015. Optimal Operation Strategy for Biohydrogen Production. *Industrial & Engineering Chemistry Research*, 54, 6334-6343.
- HANNAN, E. J. & QUINN, B. G. 1979. The determination of the order of an autoregression. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41, 190-195.
- STOICA, P. & SELEN, Y. 2004. Model-order selection: a review of information criterion rules. *IEEE Signal Processing Magazine*, 21, 36-47.
- TIBSHIRANI, R. 1996. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58, 267-288.
- TSOPANOGLOU, A. & JIMÉNEZ DEL VAL, I. 2021. Moving towards an era of hybrid modelling: advantages and challenges of coupling mechanistic and data-driven models for upstream pharmaceutical bioprocesses. *Current Opinion in Chemical Engineering*, 32, 100691.
- VRIEZE, S. I. 2012. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological methods*, 17, 228.

# Simultaneous selection and optimization of reaction pathways and downstream separation network

Shuang Xu<sup>a</sup>, Selen Cremaschi<sup>a</sup>, Mario R. Eden<sup>a</sup> and Anjan K. Tula<sup>b\*</sup>

<sup>a</sup>*Department of Chemical Engineering, Auburn University, Auburn, AL 36849, USA*

<sup>b</sup>*College of Control Science & Engineering, Zhejiang University, Hangzhou, China*

\**anjantula@zju.edu.cn*

## Abstract

Synthesis of separation networks has been researched for many years because of their potential in improving profit and reducing carbon emissions. However, most of the presented work concentrated on identifying novel solutions for the downstream section such as improving thermal efficiency by using complex heat exchanger networks, applying advanced separation techniques, and optimizing the design/operating parameters, etc., without considering the effect of upstream reaction pathways. This work aims to develop a mixed integer non-linear programming (MINLP) optimization model, which integrates reaction pathway selection and downstream separation optimization. To ensure both process profit and sustainability, the objective is to maximize the process profit per carbon emission. Decision variables such as binary variables, which represent the selection of reaction pathway and separation task, and continuous variables, which influence the separation mass/energy balance, are considered. By giving parameters like reaction conversion rate, component relative volatility, component enthalpy, raw material/product price, etc., this optimization model can simultaneously identify the optimal reaction pathways and their separation network. The optimization model was applied to isobutylene and methanol utilization case studies. The results indicate that the model could identify the best product and the associated best performing process.

**Keywords:** Reaction pathway identification, separation synthesis, optimization.

## 1. Introduction

The downstream separation network is defined as a process that comprises different unit operations, such as separators, heat exchangers and mixers, etc., to handle multicomponent separation and produce the desired products. In general, several configurations based on different separation techniques and separation sequences can be generated, and the selected options significantly influence the overall process cost and sustainability. Therefore, the optimal configuration has to be identified to ensure overall process profitability and sustainability. One established approach is the superstructure-based model, where all feasible separation configurations are generated, and the optimal choice is identified through mathematical optimization. Despite significant research on this approach, most work has been limited to downstream separation optimization, such as adding intensified equipment into the separation configurations (Caballero and Grossmann 2001, 2013), systematically identifying novel intensified unit operations (Demirel et al., 2017), simultaneous optimization of distillation networks and heat exchanger networks (Yeomans and Grossmann, 2000), but the effect of upstream conditions/decisions was not considered explicitly. Ryu et al. (2020) pointed out that the



separation network synthesis efforts should also consider the upstream reaction because the choice of the reaction influences the decisions on the downstream separation.

Integrating different reaction pathways and separation process configurations to determine the optimal products/feedstocks and process configuration has been widely researched in the biochemistry area. Kong and Shah (2016) proposed a mixed integer linear programming (MILP) model to determine the optimal reaction and separation process for utilizing bio-based materials. Zondervan et al. (2011) reported a superstructure-based optimization model to determine the optimal reaction and separation steps for biorefinery systems under different optimization objectives. Kelloway and Daoutidis (2014) developed an optimization model to identify the optimal feedstocks, products, and process configurations by maximizing the net present value and carbon efficiency for the biorefinery systems. However, with the progress of sustainable design, it is also necessary to construct a generalized optimization model for the conventional process industry to simultaneously select/optimize reaction pathways and downstream separation networks.

## 2. Methodology

The main goal is to construct an optimization model to identify the best process for producing a specific product or utilizing raw materials, which includes two parts: 1) selection of reaction pathways that convert raw materials into products. 2) identification of separation configurations that separate the reaction outlet into the final product. An example superstructure representation of this problem is illustrated in Figure 1. Given a low-value raw material, various reaction pathways can be applied to convert the raw material into different high-value products, and the outlet mixture can be separated by passing through the separation networks. The objective is to identify the most valuable product and the optimal separation configuration that ensure both process profit and sustainability.

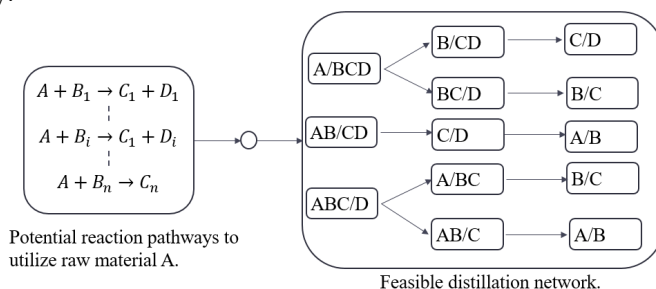


Figure 1. Superstructure representation of the optimization model.

### 2.1. Optimization model formulation

The objective is to identify the optimal reaction pathway and separation network that has the maximum process profit per carbon emission (Eq. 1). The whole process profit involves product sales and process operating costs. Also, to ensure the process sustainability, carbon emissions is considered. Reaction pathways selection and distillation mass balance constraints are given in Eqs. (2) and (3). To calculate the distillation operating cost, a short-cut distillation model (Ryu et al., 2020) that is based on distillation column vapor flowrate and enthalpy balance is given in Eqs. (4) – (7). Note here, because the number of inlet components varies according to the selection of different reaction pathways, a pseudo-component that has zero inlet flowrate is defined

in this model. To find feasible solutions of the short-cut model when the inlet component has zero flowrate, a disjunctive constraint is formulated in Eq. (4). The carbon emission constraint is given in Eq. (8). Here, natural gas is set as the primary fuel for generating the utility. By converting the heat duty into the amount of standard natural gas, the carbon emissions can be calculated based on the standard natural gas conversion factor (50.15 Kg CO<sub>2</sub>/GJ) (Shi et al., 2022). In this model, the purity of the product is specified at 99.5%, which is given in Eq. (9).

$$\max: \frac{\sum_k P_k F_k - \sum_j Q_j}{\sum_j C_j} \quad (1)$$

$$s. t. \quad \sum_i y_i = 1, \quad y_i \in \{0,1\} \quad (2)$$

$$y d_j F_{j,k} = F_{j,k,0} + F_{j,k,1}, \quad y d_j \in \{0,1\} \quad \forall j \in J, k \in K \quad (3)$$

$$\left[ \begin{array}{l} F_{j,k} F_{j,k+1} = 0 \\ \phi_l = 0, l = k \end{array} \right] \vee \left[ \begin{array}{l} F_{j,k} F_{j,k+1} \neq 0 \\ \sum_k \frac{\alpha_k F_{j,k}}{\alpha_k - \phi_l} = 0 \\ \alpha_k < \phi_l < \alpha_{k-1} \end{array} \right], \quad \forall k \in K, j \in J, l \in L \quad (4)$$

$$V_{j,0} \geq \sum_k \frac{\alpha_k F_{j,k,0}}{\alpha_k - \phi_l}, \quad \forall j \in J, l \in L \quad (5)$$

$$V_{j,1} \geq \sum_k \frac{-\alpha_k F_{j,k,1}}{\alpha_k - \phi_l}, \quad \forall j \in J, l \in L \quad (6)$$

$$Q_j = U_{j,0} V_{j,0} H_{j,0} + U_{j,1} V_{j,1} H_{j,1}, \quad \forall j \in J \quad (7)$$

$$C_j = \eta V_{j,1} H_{j,1}, \quad \forall j \in J \quad (8)$$

$$g(F_k) \geq 0.995, \quad \forall k \in K \quad (9)$$

Here, the sets  $k \in K, j \in J, i \in I$ , and  $l \in L$  represent components, separation tasks, reaction pathways, and active roots.  $P$  is the product price.  $F$  is flowrate.  $Q$  is distillation heat duty.  $U$  is the utility price (Kalakul et al., 2014).  $C$  is the amount of carbon emission.  $y$  is a binary decision variable for reaction pathway selection.  $y d$  is a binary decision variable for separation task selection.  $\alpha$  is relative volatility.  $\phi$  is active root.  $V$  is the vapor flowrate inside the stripping (1) and rectifying (0) section.  $H$  is latent heat.  $\eta$  is carbon emission coefficient.

### 3. Case study

The developed optimization model was applied to two case studies, which aim to find the best product and process configurations for utilizing isobutylene and methanol. In these two case studies, two different optimization objectives (process profit and process profit per carbon emission) are considered. All the optimization problems are solved by using BARON. Note here, for all the case studies, the separation pressure is adjusted to guarantee that cooling water is used for condensation. For temperature-sensitive products such as pivalic acid, the separation pressure is close to vacuum.

#### 3.1. Isobutylene utilization

Isobutylene was widely applied for methyl t-butyl ether (MTBE) production, which is an important gasoline additive. However, with the development of the electric car, the market demand for isobutylene may drop in the future, so three alternative isobutylene utilization reaction pathways that produce pivalic acid (PA), prenol, and tert-butylamine,

are considered in this case study. The parameters like reaction conversion rate, component relative volatility, price data ([www.echemi.com](http://www.echemi.com), [www.sunsirs.com/uk](http://www.sunsirs.com/uk)), and latent heat, which are the input for the optimization model, are listed in Table 1. In this case study, the isobutylene inlet flowrate is 150 kmol/h and the co-reactant inlet flowrate is calculated based on reaction stoichiometry.

Table 1. Parameters for the isobutylene utilization case study.

	Components	Conversion	Rel. Vol.	Price, \$/kmol	Latent heat, kJ/mol
R1	A: isobutylene	0.8	5548.36	51.79	25.31
	B: water	-	25.5	0	44.02
	C: PA	-	1	510.65	58.91
	D: <i>pseudo</i>	-	1	0	0
R2	A: formaldehyde	-	16.86	5.76	18.63
	B: isobutylene	0.8	7.82	51.79	16.99
	C: prenol	-	1	301.46	36.11
	D: <i>pseudo</i>	-	1	0	0
R3	A: ammonia	-	5.94	13.42	19.86
	B: isobutylene	0.52	2.61	51.79	16.99
	C: tert-butylamine	-	1	234.048	20.97
	D: <i>pseudo</i>	-	1	0	0

The optimal solution, including the reaction pathway, process configuration, detailed stream flowrate, carbon emissions, etc., are shown in Figure 2 and Table 2. The results indicate that using isobutylene for PA production has both the highest overall process profit and the highest process profit per amount of carbon emission (291.6 \$/(h, kg CO<sub>2</sub>)). The corresponding optimal separation configuration is based on the sequence of component boiling point.

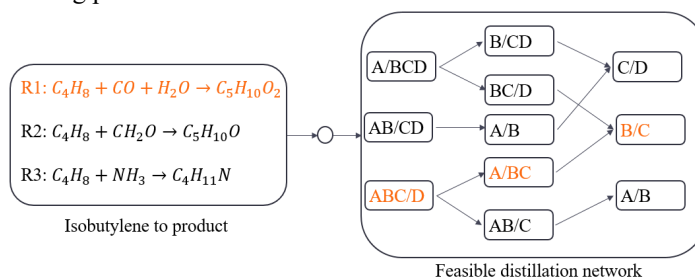


Figure 2. Optimal reaction pathway and configuration for isobutylene utilization.

Table 2. Optimization results of using isobutylene for PA production.

		Separation task					
		ABC/D		A/BC		B/C	
		Distillate	Bottom	Distillate	Bottom	Distillate	Bottom
Component flowrate, kmol/h	A	30	0	29.85	0.15	0.15	0
	B	30	0	0.15	29.85	29.25	0.6
	C	120	0	0	120	0	120
	D	0	0	0	0	0	0
Carbon emissions, kg CO <sub>2</sub> /h		0		84.23		104.35	
Operating cost, \$/h		0		31.44		33.62	

3.2. Methanol utilization

In this case study, four reaction pathways are considered. R1 uses methanol for dimethyl ether (DME) production. R2 converts methanol into gasoline with liquefied petroleum (LPG) gas as a by-product. R3 directly converts methanol into acetic acid (AA). And R4 utilizes methanol and propylene carbonate (PC) to produce dimethyl carbonate (DMC) with by-product propylene glycol (PG). The model input parameters are listed in Table 3. In this case study, the methanol inlet flowrate is 150 kmol/h and the co-reactant inlet flowrate is calculated based on reaction stoichiometry.

Table 3. Parameters for the methanol utilization case study.

	Components	Conversion rate	Relative volatility	Component price, \$/kmol	Latent heat, kJ/mol
R1	A: DME	-	81.28	25.7	17.91
	B: methanol	0.82	3.38	11.7	30.24
	C: water	-	1	0	36.9
	D: <i>pseudo</i>	-	1	0	0
R2	A: LPG	-	10	39.1	16.58
	B: methanol	0.98	3.27	11.7	29.56
	C: gasoline	-	1	139	35.45
	D: <i>pseudo</i>	-	1	0	0
R3	A: methanol	0.95	5.91	11.7	35.27
	B: AA	-	1	26.4	23.38
	C: <i>pseudo</i>	-	1	0	0
	D: <i>pseudo</i>	-	1	0	0
R4	A: methanol	0.78	175.95	11.7	35.27
	B: DMC	-	37.94	81.5	33.67
	C: PG	-	7.1	103.9	52.14
	D: PC	-	1	15.8	49.14

The solution for this case study is shown in Figure 3 and Table 4. The results show that using methanol for gasoline production has the highest process profit per amount of carbon emission (78.2 \$/(h, kg CO<sub>2</sub>)). The process configuration, which separates gasoline first, has the best performance. However, without considering the process carbon emission, using methanol for DMC production has the highest overall process profit.

Table 4. Optimization results of using methanol for gasoline production.

	Component	Separation task					
		AB/CD		A/B		C/D	
		Distillate	Bottom	Distillate	Bottom	Distillate	Bottom
flowrate, kmol/h	A	13.08	0	13.07	0.01	0	0
	B	2.7	0.3	0.07	2.64	0.3	0
	C	0	58.8	0	0	58.8	0
	D	0	0	0	0	0	0
Carbon emissions, kg CO <sub>2</sub> /h		58.83		29.96		0	
Operating cost, \$/h		11.75		5.99		0	

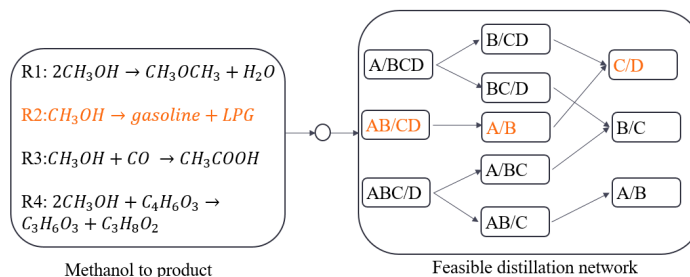


Figure 3. Optimal reaction pathway and configuration for methanol utilization.

## Conclusions

This paper developed a generalized MINLP model for simultaneous reaction pathway selection and separation network optimization. Given a product/raw material and the potential reaction pathways, this model can find the most valuable product or promising raw materials and the optimal separation configuration that ensures both process profit and the sustainability.

The model was applied to isobutylene and methanol utilization case studies. The results show that the best process for isobutylene utilization, which guarantees both process profit and process sustainability, is to convert isobutylene into PA and separate the mixture based on the sequence of boiling point. For the methanol utilization case study, the best product is gasoline, and the configuration, which separates gasoline first, gives us the most promising process performance. Future work will focus on solving multistep reaction pathway selection and separation network optimization problems.

## References

- Caballero, J. A., & Grossmann, I. E. (2001). Generalized disjunctive programming model for the optimal synthesis of thermally linked distillation columns. *Industrial & Engineering Chemistry Research*, 40(10), 2260-2274.
- Caballero, J. A., & Grossmann, I. E. (2013). Synthesis of complex thermally coupled distillation systems including divided wall columns. *AIChE Journal*, 59(4), 1139-1159.
- Demirel, S. E., Li, J., & Hasan, M. F. (2017). Systematic process intensification using building blocks. *Computers & Chemical Engineering*, 105, 2-38.
- Yeomans, H., & Grossmann, I. E. (2000). Disjunctive programming models for the optimal design of distillation columns and separation sequences. *Industrial & Engineering Chemistry Research*, 39(6), 1637-1648.
- Ryu, J., Kong, L., de Lima, A. E. P., & Maravelias, C. T. (2020). A generalized superstructure-based framework for process synthesis. *Computers & Chemical Engineering*, 133, 106653.
- Kong, Q., & Shah, N. (2016). An optimisation-based framework for the conceptual design of reaction-separation processes. *Chemical Engineering Research and Design*, 113, 206-222.
- Shi, T., Liu, Y., Yu, H., Yang, A., Sun, S., Shen, W., & Ren, J. (2022). Improved design of heat-pump extractive distillation based on the process optimization and multi-criteria sustainability analysis. *Computers & Chemical Engineering*, 156, 107552.
- Kalakul, S., Malakul, P., Siemanond, K., & Gani, R. (2014). Integration of life cycle assessment software with tools for economic and sustainability analyses and process simulation for sustainable process design. *Journal of cleaner production*, 71, 98-109.

# Workflow for adaptation, analysis and application of mechanistic models for experimental planning of protein refolding processes

Jan Niklas Pauk,<sup>a,b</sup> Chika Linda Igwe,<sup>a,b</sup> Christoph Herwig,<sup>b</sup>

<sup>a</sup>*Competence Center Chase GmbH, Altenberger Straße 69, 4040 Linz, Austria*

<sup>b</sup>*Institute of Chemical, Environmental and Bioscience Engineering, TU Wien, Gumpendorfer Straße 1a, 1060 Vienna, Austria*

## Abstract

Mechanistic models are an important tool as they allow storage of process knowledge and enable model-based techniques to plan, monitor, and control biological processes. Fundamental to these operations are models that can accurately describe their in vivo counterpart. Therefore, good modelling practices have been described including different analysis steps to ensure and maintain identifiability of model parameters during model calibration and the model life cycle (Daume et al., 2020).

Here, a workflow to analyze, adapt and apply a mechanistic model for experimental planning is suggested and its application demonstrated with a set of fed-batch refolding experiments conducted in the controlled environment of a stirred tank reactor. The model is based on the folding kinetics of a recombinant protein solubilized from aggregated proteins accumulated inside bacterial inclusion bodies. Along the adaptation cycle the model is analyzed for unique identifiability of model parameters both from simulations and noisy measurements.

**Keywords:** Mechanistic Model, Identifiability, Fisher Information, Protein Refolding, MBDOE

## 1. Introduction

In light of the current quality by design (QbD) paradigm and the process analytical technology (PAT) framework, introduced by the FDA, the use of process models for process monitoring and control are drawn into focus (FDA, 2004). Such models can either be black box data-driven models based on large data sets. Alternatively, they can be more deterministic models such as mechanistic models based on first principles or mass-, energy and momentum balances, or hybrid combinations thereof.

The focus of this work are purely mechanistic models, as they offer some important advantages: i) They require significantly smaller data sets compared to data-driven models. This is relevant, as measurements of biological processes are often expensive, involve long evaluation times such as chromatographic methods and sometimes, as is the case with cell specific reaction rates, there are no direct measurements available. ii) They serve as a good storage medium for process knowledge since they contain easily interpretable information due to correlation of data and physical parameters and physical equations. iii) Furthermore, this knowledge can be used for process monitoring (Kager et al., 2018; Sinner et al., 2019), process control (Kager et al., 2020; Ulonska et al., 2018) and optimal experimental design.

Model-based design of experiments (MBDOE) aims to generate optimal and valid models from as few experiments as possible. Thus, techniques exist to either discriminate different model variants statistically against each other to find the model that best suits the process (Brik Ternbach et al., 2005; Franceschini and Macchietto, 2008) or to design an experiment with maximal information content - usually done by maximizing D- or E-optimality of the fisher-information-matrix - to uniquely identify model parameters and achieve a valid model with good fit (Daume et al., 2019; Franceschini and Macchietto, 2008).

MBDOE generally starts with the use of a basic model structure to define different possible model variants, however, despite the advantages mentioned above, mechanistic models often require expert knowledge for the setup of the model. This contribution focuses on the generation of different model or kinetic substructures for mechanistic models. Therefore, a workflow is developed on how to adapt and improve a mechanistic model in accordance with defined good modelling practices including analysis of model sensitivity and model parameter identifiability. This workflow is applicable during the whole life cycle of the model, including the initial process development, adaptation to new products or process modes, integration of new measurement systems leading to improved parametrization potential and reoccurring model validations in the production of biopharmaceuticals. As an example, a fed-batch protein refolding process was used, optimization potential of the model was analyzed, and an experiment was designed to specifically analyze dependencies on the reaction kinetics. Furthermore, the superiority of this optimized process compared to the classical fed-batch process was shown.

This publication is organized as follows: First, the experimental and computational methods are explained. Then the model adaptation workflow is introduced, and its application discussed using the protein refolding example. Finally, some thoughts about the further application potential are given.

## 2. Methods

### 2.1. Protein Refolding Model

The basic non-linear process model of the protein refolding process consists of four protein states: solubilized protein S, folding intermediates I, native protein N and aggregated protein A. It is assumed, that solubilized protein instantly starts the folding process upon entering the refolding buffer and thus state S is removed from the model. Furthermore, extension of the model by the changing denaturant concentration D and reaction volume V due to the feeding during the fed-batch dilution is necessary. The reaction kinetics are described by the following rates depending on the concentration of the denaturant:

$$k_n = a_n \cdot (1 + D)^{b_n} \quad a_n: [0, \infty), \quad b_n: (-\infty, 0] \quad (1)$$

$$k_a = a_a \cdot (1 + D)^{b_a} \quad a_a: [0, \infty), \quad b_a: (-\infty, 0] \quad (2)$$

Here,  $k_n$  denotes the reaction rate of native protein formation and  $k_a$  the reaction rate of aggregate formation. For a more thorough review of the model see (Pauk et al., 2021).

### 2.2. Model Analysis & Optimization

*Model Fitting:* Fitting of the basic model on experimental data is performed using the genetic algorithm from the global optimization toolbox in MATLAB (2020a, MathWorks, Natick, USA) by minimizing the residual sum of squares between simulation and measured data, weighted by the measurement variances:

$$WRSS(\theta) = \sum_{p=1}^m \sum_{k=1}^d \frac{(y_{p,k} - \hat{y}_{p,k})^2}{\sigma_{p,k}^2} \quad (3)$$

Here,  $y_{p,k}$  denotes the measured concentration,  $\hat{y}_{p,k}$  denotes the simulated concentration,  $\sigma_{p,k}^2$  denotes the variance of state  $p$  at time point  $k$ .  $m$  is the number of states and  $d$  is the number of sampling times.

*Structural & practical identifiability:* Structural and practical identifiability are important attributes of a model describing if fitting of model parameters yields a unique solution from the model structure itself or the noisy measurements, respectively. Local structural identifiability analysis is performed by calculating the discrete-time scaled sensitivity matrix  $S$  with entries,

$$s_{ij} = \theta_j \cdot \frac{\partial y_i}{\partial \theta_j} \quad (4)$$

and checking for full matrix rank (Daume et al., 2020). Here,  $s_{ij}$  denotes the local sensitivity of model parameter  $j$  on the model output  $y$  at measurement  $i$ . Practical identifiability focuses on the assessment of measurement noise on the unique identifiability of all model parameters. Furthermore, correlations between model parameters resulting in multiple possible solutions where parameter value changes counteract each other need to be analyzed. One possibility to calculate parameter co-linearity is to calculate the fisher information matrix (FIM) from the sensitivity matrix and measurement variance. From the Cramer-Rao theorem follows, that the minimum parameter covariance matrix is the inverse of the FIM and following normalization yields the parameter correlation matrix (Daume et al., 2020). Here, values close to one or negative one on the off-diagonal entries imply high correlation between these parameters and therefore problems with the practical identifiability. (Brun et al., 2001) suggested two parameters which can help finding suitable parameter subsets to facilitate practical parameter identifiability. The parameter importance index can be used to rank model parameters based on their influence on model outputs due to changes in the parameter. The co-linearity index quantifies the linear dependency of model parameters of a given subset of the model parameters in regard of their practical identifiability.

*Experimental optimization & model fitting:* Optimization of the experiment was done using the genetic algorithm NSGA2. It is embedded in the pyOptSparse framework, can handle non-convex and non-smooth global optimization problems using a tournament selection-based strategy (Wu et al., 2020).

### 2.3. Protein Refolding

Protein refolding was conducted via dilution in fed-batch mode. Solubilized lactate dehydrogenase (7 g L<sup>-1</sup> protein, 4 M guanidine hydrochloride (GuHCl), 150 mM phosphate buffer, pH 6.0) originating from bacterial fermentation with recombinant Escherichia coli BL21 (DE3) was fed with constant feed-rate of 6.48 mL h<sup>-1</sup> for 9 hours into a Labfors 5 stirred tank reactor (Infors AG, Bottmingen, Switzerland) containing 0.8 L refolding buffer (150 mM phosphate buffer, 1 mM EDTA, 20  $\mu$ M NADH, pH 6.0). Refolding was carried out under controlled conditions at a temperature of 10 °C, constant stirring with a Rushton impeller at 200 rpm, headspace aeration with pressurized air at a flow-rate of 1 vvm. The pH was controlled to 6.0 by addition of 0.2 M HCl and 0.5 M NaOH.

Offline samples for the analysis of folding intermediates and native protein in the soluble fraction were measured via SEC-HPLC (Dionex UltiMate 3000, Thermo Fisher, Waltham, USA) with a BEH 200A SEC 1.7  $\mu$ m, 4.6 x 300 mm, 3.5  $\mu$ m column (Waters Corporation, Milford, USA). 2  $\mu$ L of sample were injected and eluted at a flow-rate of



0.5 mL min<sup>-1</sup> via isocratic elution (80 mM phosphate buffer, 250 mM KCl, pH 6.8) at a runtime of 18 minutes. The column oven was controlled at 25 °C and absorbance was monitored at a wavelength of 214 nm and 280 nm. The samples were further analyzed for the concentration of soluble proteins via RP-HPLC on a Polyphenyl BioResolve-RP-mAb 2.7 µm 3.0 x 100 mm column (Waters Corporation, Milford, USA) according to Kopp et al. (2020). The concentration of aggregated proteins was determined by BCA-assay as the difference of unprocessed and centrifugated sample.

### 3. Results & Discussion

#### 3.1. Workflow

The workflow illustrated in Figure 1 shows a cyclic model-based optimization approach to find suitable model structures during the model development process. An initial model either from literature or preliminary experiments (1) is used to analyze for optimization potential of the process model (2). The model-based optimization is performed (3) and the experiment conducted (4). During the evaluation process the model is analyzed regarding sensitivity, identifiability and suitable parameter subsets before the new model structure is fitted on the data set (5). The fit can now be compared to the original model using the WRSS, NRMSE of single states or a specifically targeted behavior (6). Improved fits result in adapted process models, otherwise the process can be analyzed for new optimization potential.

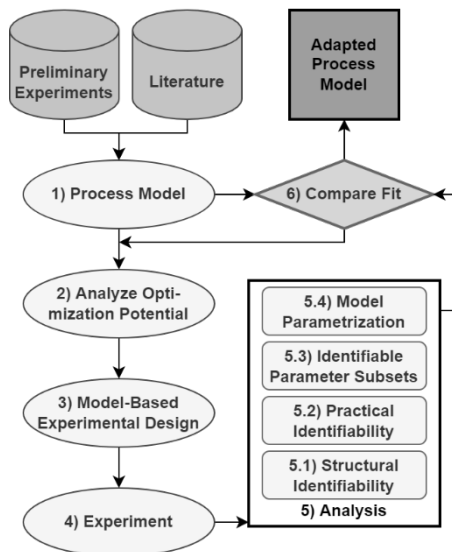


Figure 1: Experimental design cycle.

#### 3.2. Adaptation of a Protein Refolding Model

From a preliminary refolding experiment (data not shown), it was observed that the refolding rate started to stagnate after a total protein concentration of 0.6 g L<sup>-1</sup> and a denaturant concentration of 0.2 mol L<sup>-1</sup> was reached while aggregation started to increase. From equation 1 follows, that indeed the rate decreases with increasing denaturant concentration. However, a clear dependency with the total protein concentration is not described. Therefore, the rates are extended by a term penalizing increasing total protein concentration,

$$k_n = a_n \cdot (1 + D)^{b_n} \cdot (1 + P)^{c_n} \quad c_n: (-\infty, 0] \quad (5)$$

$$k_a = a_a \cdot (1 + D)^{b_a} \cdot (1 + P)^{c_a} \quad c_a: (-\infty, 0] \quad (6)$$

with  $P$  the total protein concentration in the refolding vessel. Furthermore, an experiment is designed to analyze the dependency of rate and protein concentration by changing one of the parameters while keeping the other constant. The process consists of three phases: (I) an initial fed-batch phase, (II) a phase of constant protein and denaturant concentration, and (III) a phase with constant denaturant but decreasing protein concentration (see Figure 2, left). These phases are realized by using three different reservoirs and feed pumps with reservoir one containing the solubilized protein (7 g L<sup>-1</sup>, 4 M GuHCl), reservoir two containing pure refolding buffer, and reservoir three containing refolding

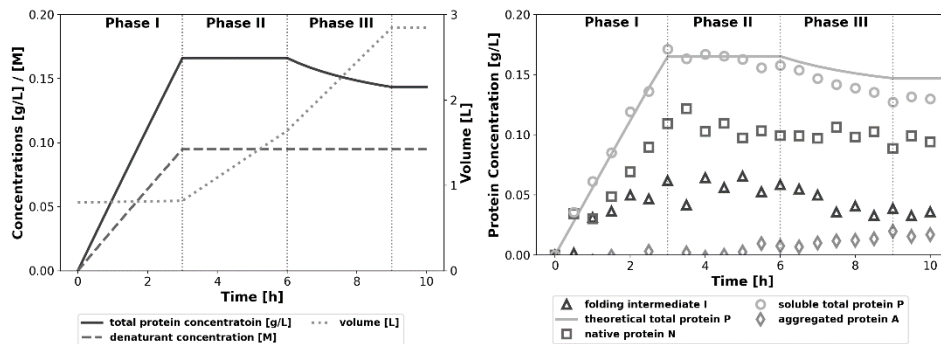


Figure 2: Feeding strategy (left) and data of the experiment optimized through model-based optimization (right). Three phases (vertical dashed lines): initial fed-batch phase, phase with constant protein and denaturant concentrations, phase with constant denaturant and declining protein concentration.

buffer spiked with 0.031 M GuHCl. The first feed rate and the concentrations in reservoir one was set according to section 2.3, the other two feed rates and the concentration of the denaturant (GuHCl) in reservoir three were optimized by the NSGA2 algorithm ( $F_{R2} = 0.2667 \text{ L h}^{-1}$ ;  $F_{R3} = 0.3962 \text{ L h}^{-1}$ ).

Figure 2 (right) shows measurements of the protein states and the total soluble protein concentration as symbols and the theoretical total protein concentration as solid line. Although the total soluble protein concentration decreases during the later stages of the process due to aggregation, the protein concentration follows the optimized trajectory well and the process results in a final yield of 64 % compared to 28 % in the preliminary experiment.

Analysis of the model by calculating the rank of the sensitivity matrix and the FIM suggested that all model parameters are structurally and practically identifiable. However, the correlation matrix and calculation of the co-linearity index showed moderate correlation of the model parameters  $a_n$  to  $b_n$  and  $a_a$  to  $b_a$  in case of the initial rates as well as high correlation of  $b_n$  to  $c_n$  and  $b_a$  to  $c_a$  for the adapted rates. Thus, a subset analysis or an optimization of the information density by optimizing sampling times would be necessary to accurately estimate the model parameters (Daume et al., 2019).

Table 1: NRMSEs comparing the fit of the model with initial (I) and adapted rates (II).

States	I	N	A	D	V
NRMSE (I)	0.1437	0.0878	0.1410	0.0027	0.0167
NRMSE (II)	0.1414	0.0775	0.1443	0.0027	0.0165
Improvement	1.61 %	11.73 %	-2.33 %	0.73 %	1.09 %

Nonetheless, the fit of the model with old and adapted rates is compared using the NRMSE of each state (see Table 1). A slight overall improvement is visible and a large improvement specifically in state N, the native protein. N is the most important state regarding a protein refolding process, therefore, the conducted optimized experiment results in a sufficiently different adapted process model structure for further analysis in the model discrimination process. The designed experiment was able to identify the potential of the adapted rates by decoupling protein and denaturant concentrations during the different phases of the experiment.

#### 4. Conclusion

This paper suggested a workflow to generate new model or kinetic structures as prerequisite for a model discrimination step. The potential of model-based experimental design was shown. An experiment specifically targeted to identify dependencies of the reaction kinetics on protein and denaturant concentration was designed, resulting in a new set of reaction kinetics to better describe the refolding process, specifically the concentration of native protein. The example also showed difficulties with co-linear model parameters. Furthermore, due to the low protein concentration resulting from the high dilutions in phase II and III, the yield of native protein was increased by almost two-fold compared to a fed-batch experiment with only one feed.

#### 5. Acknowledgements

The authors acknowledge financial support through the COMET Centre CHASE, funded within the COMET Competence Centers for Excellent Technologies program by the BMK, the BMDW and the Federal Provinces of Upper Austria and Vienna. The COMET program is managed by the Austrian Research Promotion Agency (FFG). The authors also acknowledge insights and support provided by the project partner Bilfinger Life Science.

#### References

- Brik Ternbach, M., Bollman, C., Wandrey, C., Takors, R., 2005. Application of model discriminating experimental design for modeling and development of a fermentative fed-batch L-valine production process. *Biotechnol. Bioeng.* 91, 356–368.
- Brun, R., Reichert, P., Künsch, H.R., 2001. Practical identifiability analysis of large environmental simulation models. *Water Resour. Res.* 37, 1015–1030.
- Daume, S., Kager, J., Herwig, C., 2019. Time Resolved Sensitivity & Identifiability Analysis for Directed Parametrization of Highly Dynamic Models, in: *Computer Aided Chemical Engineering*. Elsevier, pp. 1111–1116.
- Daume, S., Kofler, S., Kager, J., Kroll, P., Herwig, C., 2020. Generic Workflow for the Setup of Mechanistic Process Models, in: Pörtner, R. (Ed.), *Animal Cell Biotechnology: Methods and Protocols*. Springer US, New York, NY, pp. 189–211.
- FDA, 2004. *Guidance for Industry PAT - A Framework for Innovative Pharmaceutical Development, manufacturing, and Quality Assurance*.
- Franceschini, G., Macchietto, S., 2008. Model-based design of experiments for parameter precision: State of the art. *Chem. Eng. Sci.* 63, 4846–4872.
- Kager, J., Herwig, C., Stelzer, I.V., 2018. State estimation for a penicillin fed-batch process combining particle filtering methods with online and time delayed offline measurements. *Chem. Eng. Sci.* 177, 234–244.
- Kager, J., Tuveri, A., Ulonska, S., Kroll, P., Herwig, C., 2020. Experimental verification and comparison of model predictive, PID and model inversion control in a Penicillium chrysogenum fed-batch process. *Process Biochem.* 90, 1–11.
- Pauk, J.N., Raju Palanisamy, J., Kager, J., Koczka, K., Berghammer, G., Herwig, C., Veiter, L., 2021. Advances in monitoring and control of refolding kinetics combining PAT and modeling. *Appl. Microbiol. Biotechnol.* 105, 2243–2260.
- Sinner, P., Kager, J., Daume, S., Herwig, C., 2019. Model-based Analysis and Optimisation of a Continuous Corynebacterium glutamicum Bioprocess Utilizing Lignocellulosic Waste. *IFAC-Pap.* 52, 181–186.
- Ulonska, S., Waldschitz, D., Kager, J., Herwig, C., 2018. Model predictive control in comparison to elemental balance control in an E. coli fed-batch. *Chem Eng Sci* 191, 459–467.
- Wu, N., Kenway, G., Mader, C.A., Jasa, J., Martins, J.R., 2020. pyOptSparse: A Python framework for large-scale constrained nonlinear optimization of sparse systems. *J. Open Source Softw.* 5, 2564.

# Mechanistic modeling of product formation in recombinant *Escherichia coli* cultivations

Don Fabian Müller<sup>a</sup>, Philipp Pably<sup>a</sup>, Daniel Wibbing<sup>c</sup>, Julian Kager<sup>d</sup> and Christoph Herwig<sup>\*a,b</sup>

<sup>a</sup>*Institute of Chemical, Environmental and Bioscience Engineering, TU Wien, Gumpendorfer Straße 1a, 1060 Vienna, Austria*

<sup>b</sup>*Competence Center CHASE GmbH, 4040 Linz, Austria*

<sup>c</sup>*Festo SE & Co. KG, 73770 Denkendorf, Germany*

<sup>d</sup>*Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 228A, 2800 Kgs. Lyngby, Denmark*  
[christoph.herwig@tuwien.ac.at](mailto:christoph.herwig@tuwien.ac.at)

## Abstract

In this work a nonlinear mechanistic model was developed predicting the product formation of recombinant lactate dehydrogenase (LDH) using fed-batch cultivations of *Escherichia coli*. The derived product formation kinetics consider the inclusion body (IB) fraction as well as the soluble protein fraction. After the structural identifiability of the model parameters was analyzed, five fed-batch cultivations were conducted in total in order to identify the model parameters and to validate the model subsequently. The accuracy of the final model was overall good with NRMSE values of around 0.1 to 0.2 for both product fractions. The derived model can be seen as an enabler for advanced monitoring and control strategies like model predictive control and other advances technologies defined in the process analytical technology (PAT) initiative.

**Keywords:** *Escherichia coli*, bioprocess control, product formation, mechanistic model, process analytical technology

## 1. Introduction

In contrast to the chemical process industry, industrial bioprocess control strategies are still mainly based on linear closed-loop control schemes like PID control or simple predefined open-loop trajectories (Narayanan et al. 2020). So far in bioprocess development, feed rate optimization is often realized by design of experiments (DoE) approaches. Instead of optimizing the feed rate directly, modern approaches aim to optimize the biomass specific substrate uptake rate ( $q_s$ ) as a physiological parameter resulting in a more generic approach independent of the bioreactor scale (Wechselberger et al. 2012). However, these solutions are assuming a direct linear relationship between  $q_s$  and the specific product formation rate ( $q_p$ ), which might not be valid throughout the course of the process. Furthermore, batch-to-batch variability is an inherent nature of bioprocesses and therefore open-loop control strategies come to their limits as they cannot react on process disturbances. Advantages of using more advanced model-based control algorithms were discussed previously (Ulonska et al. 2018; Wechselberger, Sagmeister, and Herwig 2013; Kager et al. 2022). Especially due to the complex multivariate and nonlinear relationships in bioprocesses, model predictive control (MPC), which is taking the interrelations between process variables into account, offers higher flexibility than decoupling the control task into several single setpoint tracking control loops (Mears et

al. 2017). Kager et al. (2022) developed a two-degree of freedom control strategy using nonlinear feedback linearization in order to directly control the specific product formation rate  $q_p$  in cultivations of recombinant *Escherichia coli* and stated increased biomass specific titers with this model-based controller compared to optimal constant feed-rate setpoints. The prerequisite of such advanced control strategies is a process model representing the dynamics of the system with sufficient accuracy. In this work we developed a simple mechanistic model for the product formation of a recombinant *Escherichia coli* strain producing LDH. In five fed-batch strain characterization experiments we analyzed the soluble product content as well as the aggregated inclusion body (IB) fraction of the product and included both as separated states in our model, which to our knowledge has not been done before.

## 2. Methods

### 2.1. Cultivation of *Escherichia coli*

Five cultivations were conducted in total. Four out of five experiments (1, 2, 4, 5) were used as a training dataset for parameter estimation and one experiment (3) was used as validation dataset. Recombinant *Escherichia coli* BL21(DE3) was cultivated in fed-batch mode on DeLisa (DeLisa et al. 1999) minimal medium. The target product was the enzyme L-lactate dehydrogenase (LDH). The bioreactor system Labfors 5 (Infors, Bottmingen, Switzerland) with a total volume of 3.3 L was used for cultivation. 1 L medium was supplemented with 100 mM Ampicillin and 15 gL<sup>-1</sup> glucose and inoculated with 100 mL of preculture (1 L shake flask, 37 °C, 230 rpm, 12 h) The controlled value for pH was 6.75 (by addition of NH<sub>4</sub><sup>+</sup> and H<sub>3</sub>PO<sub>4</sub>), for temperature 37 °C before induction and 31.5 °C after induction, and for pO<sub>2</sub> 30% via stirrer speed (400 rpm to 1200 rpm). After the batch phase was finished a fed-batch phase was initiated by controlling the specific substrate uptake rate  $q_s$  to 0.25 gg<sup>-1</sup>h<sup>-1</sup>. After 24 h of process time the product formation was induced with 1 mM isopropyl  $\beta$ -d-1-thiogalactopyranoside (IPTG). Then, different values for  $q_s$  were applied. Table 1 shows the mean values of  $q_s$  after induction of the product formation.

Table 1: Conducted experiments

Name	$q_{s,mean}$
Process 1	0.25 gg <sup>-1</sup> h <sup>-1</sup>
Process 2	0.14 gg <sup>-1</sup> h <sup>-1</sup>
Process 3	0.35 gg <sup>-1</sup> h <sup>-1</sup>
Process 4	0.23 gg <sup>-1</sup> h <sup>-1</sup>
Process 5	0.14 gg <sup>-1</sup> h <sup>-1</sup>

### 2.2. Analytical methods

Online measurements of pH, dissolved oxygen tension (pO<sub>2</sub>), liquid temperature ( $\theta_L$ ), substrate feed rate ( $F_R$ ), airflow rate ( $F_{AIR}$ ), oxygen flow rate ( $F_{O_2}$ ) and offgas analysis (BlueInOne, BlueSens, Herten, Germany) were recorded and used for real-time calculations. For offline reference analytics samples of the culture broth were taken every 2 h during the bioprocess. For biomass quantification the dry cell weight (DCW) was analyzed. After centrifugation of 1 mL sample (20000 g, 10 min), the pellet was dried at 105 °C and gravimetrically quantified in 1.5 mL tubes. The optical density (OD) was determined at 600 nm and the substrate content was obtained through an automated photometric assay using a hexokinase kit (BioHT, Roche, Germany). The live-dead analysis was carried out via a Flow Cytometry (CyFlow® Cube 6, Sysmex Partec GmbH, Germany) with the staining protocol described in (Langemann et al. 2016). The recorded forward-scatter and fluorescence signals were correlated with the amount of dead cells. In order to measure the intracellularly produced LDH the cell pellets were resuspended and disrupted by a combination of chemical lysis and a sonication rod. After centrifugation the supernatant was analyzed for its soluble content of the target enzyme

and the pellet containing the IBs was further processed. The latter was washed and then solubilized using a solution of Guanidinium chloride, denaturing and dissolving the aggregated protein chains. By diluting the resulting protein solution the enzyme refolded into its active form. The enzymatic activity of the soluble and the refolded LDH was determined using an adaptation of the enzyme assay procedure described in (Ponsoda et al. 1991), utilizing the catalytic nature of LDH for the reduction of pyruvate to lactate. The change in adsorption was measured at 340 nm with the automatic Spark® multimode microplate reader (Tecan Trading AG, Switzerland). Additionally, several samples were analyzed with a RP-HPLC (Thermo Fisher Scientific Inc., USA) using the method of (Kopp et al. 2020) together with a BSA-standard curve. The activities of the same samples were correlated to their protein content and a conversion factor of 51400U g<sup>-1</sup> was determined. This was further used to translate all activity measurements to protein concentrations.

### 2.3. Bioprocess model

The model structure which was used is an unstructured model of six states, namely liquid volume (V), living biomass (X<sub>v</sub>), dead biomass (X<sub>d</sub>), substrate (S), total product (P) and inclusion body product (P<sub>IB</sub>).

$$\begin{aligned} \frac{dV}{dt} &= F_{in} - F_{out} & \frac{dX_v}{dt} &= Y_{X/S} \cdot q_S \cdot X_v - k_d \cdot X_v - F_{out} \cdot \frac{X_v}{V} \\ \frac{dX_d}{dt} &= k_d \cdot X_v - F_{out} \cdot \frac{X_d}{V} & \frac{dS}{dt} &= F_R \cdot c_{SR} - q_S \cdot X_v - F_{out} \cdot \frac{S}{V} \\ \frac{dP}{dt} &= q_P \cdot X_v - k_d \cdot P - F_{out} \cdot \frac{P}{V} & \frac{dP_{IB}}{dt} &= q_{P_{IB}} \cdot X_v - k_d \cdot P_{IB} - F_{out} \cdot \frac{P_{IB}}{V} \end{aligned} \quad (1)$$

Three kinetic equations were used to describe the cell specific rates. Monod kinetics were used to describe specific substrate uptake  $q_S$ , a combination of Monod and Haldane kinetics were used to describe total product formation  $q_P$  and Moser kinetics were used to describe IB formation  $q_{P_{IB}}$ .  $q_P$  is modeled dependent on  $q_S$  and on the cumulated metabolized substrate  $S_{met}$ , which is obtained as the integral of  $q_S \cdot X_v$  since the induction.

$$q_S = q_{Smax} \cdot \frac{c_S}{k_S + c_S} \quad q_{P_{IB}} = q_{P_{IB},max} \cdot \frac{q_P^N}{q_P^N + K_{P,q_{P_{IB}}^N}} \quad (2)$$

$$q_P = q_{P,max} \cdot \frac{q_S}{q_S + K_{S,q_S}} \cdot \frac{S_{met}/X}{\frac{(S_{met}/X)^k}{K_{I,q_P}} + S_{met}/X + K_{S,q_P}}$$

### 2.4. Model analysis and parameterization

Before parameter estimation a structural identifiability analysis was performed. Therefore, the sensitivity matrix

$$S = \frac{\partial(x_1, \dots, x_m)}{\partial(\theta_1, \dots, \theta_n)}, \quad S_{i,p} = \frac{\partial x_i}{\partial \theta_p} \quad (3)$$

of the model was derived and the rank was calculated. If S has full rank, the model is structurally identifiable, which means that all model parameters  $\theta$  can be uniquely identified from the measured states  $x$ .

### 2.5. Real-time implementation

Data analysis and model development was executed with MATLAB® and Simulink® (Version R2022a, MathWorks, Natick, Massachusetts, USA). Real-time process control and communication to the equipment was realized through the Lucullus Process

Information Management System PIMS (Securecell, Urdorf, Switzerland). Lucullus was connected via OPC to a computer running MATLAB<sup>®</sup>, where the data preprocessing, the feed control and the state estimation model were executed for the different experiments. The numerical derivative of the feed weight signal was computed and smoothed with a Savitzky-Golay-Filter approach (Savitzky and Golay 1964) to obtain the feed rate  $F_R$  used as model input.

### 3. Results & Discussion

The exploration of the collected data yielded a correlation between the productivity  $q_P$ , the specific substrate uptake  $q_S$  and the metabolized substrate  $S_{met}$  as an inhibitor, which acts as an indicator for cell age. This hypothesis is supported by the data as seen in fig. 1 for the biomass specific product titer  $P_{specific}$ .

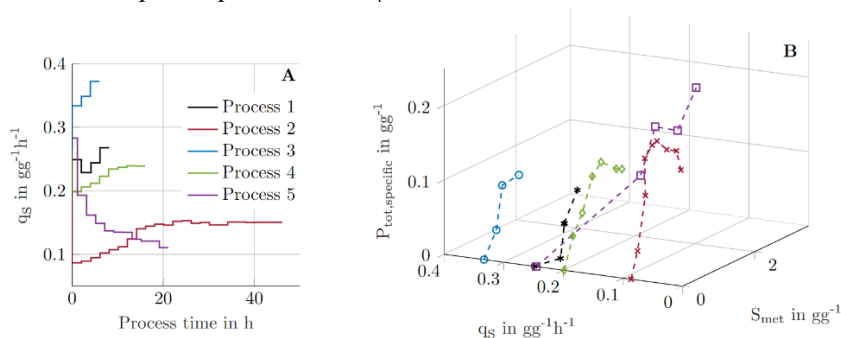


Figure 1: Cell-specific total product titer in  $gg^{-1}$  dependent on substrate uptake rate  $q_S$  and metabolized substrate  $S_{met}$  for all six processes.

These states were further used to find a kinetic equation modeling the cell productivity after induction. Several kinetic formulations, combining Monod and other substrate uptake equations with inhibition terms were tested for their suitability and performance. Additionally, the structural identifiability of the resulting models were analyzed to proof that the selected structure can be uniquely parameterized when no measurement error is assumed. The same was done for the subsequent inclusion body formation, expressed through  $q_{PIB}$ . Here, a linear connection to the overall productivity  $q_P$  was observed in the data and was therefore chosen as the variable for the kinetic equations. The structural analysis and parameterization with the data sets yielded two optimal kinetic equations by comparing their scores for the  $R^2_{adjusted}$  and the NRMSE values. The specific productivity  $q_P$  of choice was a combination of a Monod term with a Haldane inhibition term triggered by  $S_{met}$ . For the product localization the best specific inclusion body formation  $q_{PIB}$  among the tested kinetics was one of Moser type (eq. 3). The performance of the final model against the given data sets together with the respective scores can be seen in fig. 2. The identifiability analysis yielded that the chosen kinetics with 5 parameters for the total productivity and 3 parameters for the IB formation yield a structural identifiable model. A practical analysis of the identifiability also revealed a sensitivity towards the process noise, which can create problems for the parameterization and lead to unstable solutions for the parameter vector or leave some positions as unidentifiable.

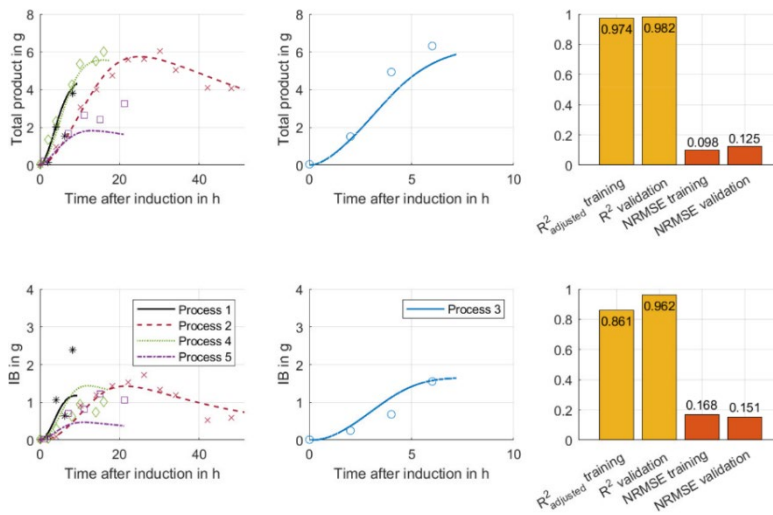


Figure 2: Results of the parameterization with the final model kinetics for the total product and the IBs. Model parameters:  $q_{p,max} = 0.0396 \text{ gg}^{-1}\text{h}^{-1}$ ,  $K_{s,q_s} = 0.0727 \text{ gg}^{-1}\text{h}^{-1}$ ,  $k = 5.56$ ,  $K_{s,q_p} = 0.4669 \text{ gg}^{-1}$ ,  $K_{L,q_p} = 3.6960$ ,  $q_{PIB,max} = 0.0195 \text{ gg}^{-1}\text{h}^{-1}$ ,  $K_{p,qPIB} = 0.0282 \text{ gg}^{-1}\text{h}^{-1}$ ,  $N = 1.9627$

The simulation shows that the final model structure and choice of variables allows a good tracking of the productivity, both in the strong surge right after the induction, as well as the decline over the progress of the process. The NRMSE values for the parameterization datasets as well as for the validation datasets range between 0.1 and 0.2 indicating an overall satisfactory model calibration. The model validity spans in a high range of feed rate inputs ( $q_s$  from  $0.14 \text{ gg}^{-1}\text{h}^{-1}$  to  $0.35 \text{ gg}^{-1}\text{h}^{-1}$ ) and the model can probably be easily transferred to other feeding strategies. Overall, the impact of the chosen  $q_s$ -setpoints on the product and IB formation is well represented enabling the usage of this model in future prediction and control algorithms.

#### 4. Conclusion

We developed a mechanistic model for the product formation of *Escherichia coli*. The model is based on macro-kinetic principles describing product formation dependent on the specific rate of substrate uptake  $q_s$  and the cumulative metabolized substrate  $S_{met}$ , which were considered as indicators for the metabolic load of the cell. In contrast to some other unstructured models described in literature (Neubauer, Lin, and Mathiszik 2003; Kim et al. 2022; Kager et al. 2022) here we describe the formation of soluble product as well as the aggregation into inclusion bodies which opens up the possibility to separately control the two states of the product and possibly steer the production in a specific direction. Model predictive control (MPC) can serve as a framework to implement this strategy and control directly on the product formation and its status. Further, transferability of the identified mechanism to other products and organism needs to be proven to see the developed model as a platform.

#### 5. Acknowledgements

The authors acknowledge the TU Wien Bibliothek for financial support through its Open Access Funding Program. The authors acknowledge financial support through the COMET Centre CHASE, funded within the COMET – Competence Centers for Excellent Technologies programme by the BMK, the BMDW and the Federal Provinces of Upper



Austria and Vienna. The COMET programme is managed by the Austrian Research Promotion Agency (FFG).

## References

- DeLisa, M. P., J. Li, G. Rao, W. A. Weigand, and W. E. Bentley. 1999. "Monitoring GFP-Operon Fusion Protein Expression during High Cell Density Cultivation of Escherichia Coli Using an on-Line Optical Sensor." *Biotechnology and Bioengineering* 65 (1): 54–64.
- Kager, Julian, Johanna Bartlechner, Christoph Herwig, and Stefan Jakubek. 2022. "Direct Control of Recombinant Protein Production Rates in E. Coli Fed-Batch Processes by Nonlinear Feedback Linearization." *Chemical Engineering Research and Design* 182: 290–304.
- Kim, Jong Woo, Niels Krausch, Judit Aizpuru, Tilman Barz, Sergio Lucia, Peter Neubauer, and Mariano Nicolas Cruz Bournazou. 2022. "Model Predictive Control and Moving Horizon Estimation for Adaptive Optimal Bolus Feeding in High-Throughput Cultivation of E. Coli." arXiv. <http://arxiv.org/abs/2203.07211>.
- Kopp, Julian, Florian B. Zauner, Andreas Pell, Johanna Hausjell, Diana Humer, Julian Ebner, Christoph Herwig, Oliver Spadiut, Christoph Slouka, and Reinhard Pell. 2020. "Development of a Generic Reversed-Phase Liquid Chromatography Method for Protein Quantification Using Analytical Quality-by-Design Principles." *Journal of Pharmaceutical and Biomedical Analysis* 188 (September): 113412.
- Langemann, Timo, Ulrike Beate Mayr, Andrea Meitz, Werner Lubitz, and Christoph Herwig. 2016. "Multi-Parameter Flow Cytometry as a Process Analytical Technology (PAT) Approach for the Assessment of Bacterial Ghost Production." *Applied Microbiology and Biotechnology* 100 (1): 409–18. <https://doi.org/10.1007/s00253-015-7089-9>.
- Mears, Lisa, Stuart M. Stocks, Mads O. Albaek, Gürkan Sin, and Krist V Gernaey. 2017. "Mechanistic Fermentation Models for Process Design, Monitoring, and Control." *Trends in Biotechnology* 35 (10): 11.
- Narayanan, Harini, Martin F. Luna, Moritz Stosch, Mariano Nicolas Cruz Bournazou, Gianmarco Polotti, Massimo Morbidelli, Alessandro Butté, and Michael Sokolov. 2020. "Bioprocessing in the Digital Age: The Role of Process Models." *Biotechnology Journal* 15 (1): 1900172.
- Neubauer, P., H. Y. Lin, and B. Mathiszik. 2003. "Metabolic Load of Recombinant Protein Production: Inhibition of Cellular Capacities for Glucose Uptake and Respiration after Induction of a Heterologous Gene InEscherichia Coli." *Biotechnology and Bioengineering* 83 (1): 53–64.
- Ponsoda, Xavier, Ramiro Jover, José Vicente Castell, and María José Gómez-Lechón. 1991. "Measurement of Intracellular LDH Activity in 96-Well Cultures: A Rapid and Automated Assay for Cytotoxicity Studies." *Journal of Tissue Culture Methods* 13 (1): 21–24.
- Savitzky, Abraham., and M. J. E. Golay. 1964. "Smoothing and Differentiation of Data by Simplified Least Squares Procedures." *Analytical Chemistry* 36 (8): 1627–39.
- Ulonska, Sophia, Daniel Waldschitz, Julian Kager, and Christoph Herwig. 2018. "Model Predictive Control in Comparison to Elemental Balance Control in an E. Coli Fed-Batch." *Chemical Engineering Science* 191 (December): 459–67. <https://doi.org/10.1016/j.ces.2018.06.074>.
- Wechselberger, Patrick, Patrick Sagmeister, Helge Engelking, Torsten Schmidt, Jana Wenger, and Christoph Herwig. 2012. "Efficient Feeding Profile Optimization for Recombinant Protein Production Using Physiological Information." *Bioprocess and Biosystems Engineering* 35 (9): 1637–49. <https://doi.org/10.1007/s00449-012-0754-9>.
- Wechselberger, Patrick, Patrick Sagmeister, and Christoph Herwig. 2013. "Model-based Analysis on the Extractability of Information from Data in Dynamic Fed-batch Experiments." *Biotechnology Progress* 29 (1): 285–96. <https://doi.org/10.1002/btpr.1649>.

# Physics-informed neural networks for optimization of polymer reactor design

Yubin Ryu<sup>a,b</sup>, Sunkyu Shin<sup>c</sup>, Jay J. Liu<sup>d</sup>, Wonbo Lee<sup>c</sup>, Jonggeol Na<sup>a,b,\*</sup>

<sup>a</sup> Department of Chemical Engineering and Materials Science, Ewha Womans University, Ewhayeodae-gil 52, Seodaemun-gu, Seoul 03760, Republic of Korea

<sup>b</sup> Graduate Program in System Health Science and Engineering, Ewha Womans University, Ewhayeodae-gil 52, Seodaemun-gu, Seoul 03760, Republic of Korea

<sup>c</sup> School of Chemical and Biological Engineering, Seoul National University, Gwanak-ro 1, Gwanak-gu, Seoul 08826, Republic of Korea

<sup>d</sup> Department of Chemical Engineering, Pukyong National University, Yongso-ro 45, Nam-gu, Busan 48513, Republic of Korea

[jgna@ewha.ac.kr](mailto:jgna@ewha.ac.kr)

## Abstract

Computational fluid dynamics (CFD) is utilized for chemical reactor design/analysis and enables accurate predictions by solving numerical methods. However, real-time prediction is currently infeasible due to its high computational cost. As a result, physics-informed neural networks (PINNs), which integrate data-based machine learning and physical principles, have recently emerged as a surrogate model of CFD (Raissi, M., et al. (2019)). But previous PINNs can only predict properties in terms of location or time. Here, we offer novel PINNs that can also handle operating conditions as input variables to evaluate the impact of various design variables for reactor design optimization (Raissi, M., et al. (2020)). The proposed model is applied to an autoclave reactor for free radical polymerization of ethylene and demonstrated to be capable of interpolating/extrapolating solutions accurately under various settings.

**Keywords:** Computational fluid dynamics; deep learning; optimization; reactor design; artificial intelligence

## 1. Introduction

Computational fluid dynamics (CFD) solves fundamental nonlinear differential equations using numerical methods to analyze flow characteristics. Especially in chemical process engineering, CFD optimizes operating conditions and considerably reduces costs. However, for the physics residual of the governing equations to converge, a huge computational cost is required. As a result, a data-based model called physics-informed neural networks (PINNs) emerged as a surrogate model of CFD.

To integrate the data-driven model and physics-based model, PINNs solve partial differential equations (PDEs) by transforming the issue of finding the solution into a loss function optimization problem. The physical residual terms are added to the loss function and act like a penalty that limits the space of the solution (Cuomo, S., et al. (2022)). Therefore, PINNs can consider governing principles, rather than inferring solutions based only on training data. In addition, PINNs are based on a mesh-free approach and can be applied to complex geometries satisfying physical details with enhanced generalization,

training speed, and accuracy (Karniadakis, G.E., et al. (2021)). Initial PINNs models have been developed to solve mathematical models or inverse problems, solving nonlinear PDEs such as Schrödinger, Burgers, and Allen-Kahn equations (Raissi, M., et al. (2017)). Moreover, PINNs have also been applied to fluid dynamics to solve high-dimensional problems dominated by the Navier-Stokes equation. However, since previous PINNs models could only predict properties in terms of space and time, we intend to develop a novel PINNs model that can even handle operating conditions as input variables.

This study uses PINNs to analyze the flow and reaction in an autoclave reactor where ethylene is polymerized. Unlike previous models, the proposed model was constructed by adding two operating conditions to search for optimal conditions. The potential of PINNs as a surrogate model of CFD was demonstrated by comparing the accuracy of prediction according to location and two driving conditions (inlet temperature and concentration of initiator).

## 2. Methodology

### 2.1. Polymerization reactor modeling via CFD

Figure 1 depicts the geometry of the 3D autoclave reactor, with a pitched blade impeller rotating at a speed of 100 rpm. In this reactor, the monomer flows in from the inlet at the top, the initiator flows in from the side, and generated polymer flows out through the outlet at the bottom. The reactor model is divided into two domains: the inner domain and the outer domain. The governing equations depend on the domain. In the case of the inner domain, the governing equation is determined by the moving reference frame (MRF), since it is a region directly affected by the impeller. In comparison, the outer domain is the part excluding the inner domain, so the governing equation is determined by the stationary frame. The governing equations used in the reactor model are shown in Table 1. Details of variables used in partial differential equations are listed in the appendix.

To generate data, polymerization reactor modeling was performed using ANSYS-Fluent v21.2. Each component's velocity, pressure, temperature, and concentration were extracted according to location and operating conditions. A total of 357 combinations of operating conditions were used as training data, 17 for inlet temperature and 21 for inlet concentration.

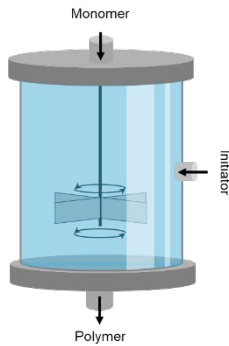


Figure 1. Schematic of autoclave reactor

Table 1. Governing Equations of Inner Domain

Continuity	$u_x + v_y + w_z$
Momentum	$\rho(u_t + u_x u + u_y v + u_z w - \omega(u_x y - u_y x - v)) + p_x - \mu(u_{xx} + u_{yy} + u_{zz})$
	$\rho(v_t + v_x u + v_y v + v_z w - \omega(v_x y - v_y x + v)) + p_y - \mu(v_{xx} + v_{yy} + v_{zz})$
	$\rho(w_t + w_x u + w_y v + w_z w) + p_z - \mu(w_{xx} + w_{yy} + w_{zz})$
Species Transport	$(u l_x + u_x l + v l_y + v_y l + w l_z + w_z l) - k_i e^{-\frac{E_a}{T}}$

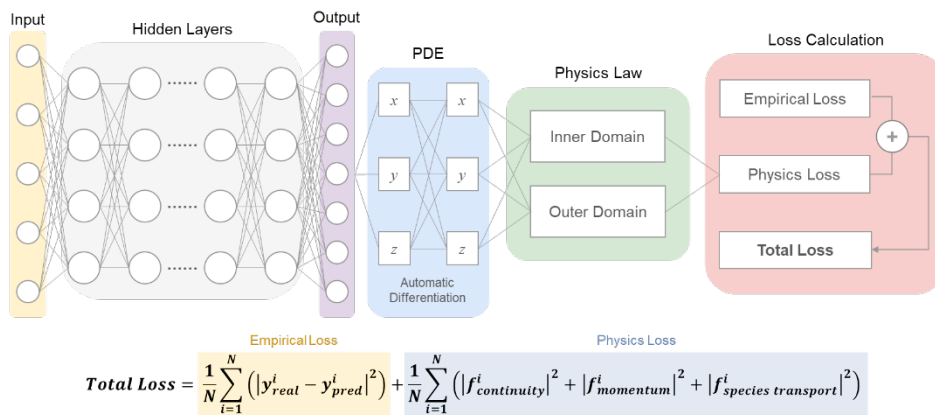
Table 2. Governing Equations of Outer Domain

Continuity	$u_x + v_y + w_z$
Momentum	$\rho(u_t + u_x u + u_y v + u_z w) + p_x - \mu(u_{xx} + u_{yy} + u_{zz})$
	$\rho(v_t + v_x u + v_y v + v_z w) + p_y - \mu(v_{xx} + v_{yy} + v_{zz})$
	$\rho(w_t + w_x u + w_y v + w_z w) + p_z - \mu(w_{xx} + w_{yy} + w_{zz})$
Species Transport	$(u l_x + u_x l + v l_y + v_y l + w l_z + w_z l) - k_i e^{-\frac{E_a}{T}}$

## 2.2. Physics-informed neural networks for reaction systems

### 2.2.1. Model Description

The architecture of the PINNs model is shown in Figure 2. Input variables are coordinates ( $x$ ,  $y$ ,  $z$ ) and inlet temperature, inlet concentration of the initiator. Output variables are velocity, temperature, pressure, and concentration of initiator and monomer. The loss function of PINNs is summed up of empirical loss and physics loss. Empirical loss refers to the mean square error (MSE) between prediction and target data. On the other hand, the physics loss is the residual of the governing equations obtained by automatic differentiation. This loss term forces PINNs to obey the laws of physics. To consider the scale between the governing equations (continuity, momentum, and species transport), physics loss is calculated by assigning a weight corresponding to each loss. PINNs are trained by updating the neural network parameters to minimize total loss.

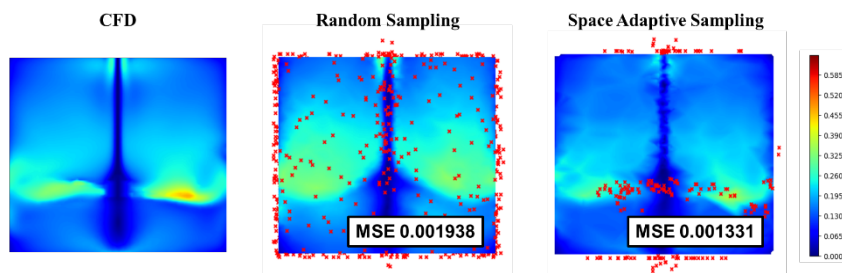


**Figure 2.** Schematic of PINNs: The architecture of PINNs consists of 13 layers including 100 neurons for each layer. Empirical loss measures how well the neural network reflects the training data. Physics loss ensures that the solution is consistent with the known physics.

### 2.2.2. Learning Strategy

For training data, a total of 200000 data with 4000 collocation points were used. We split the CFD simulation data into 3 sets: a train set used to train the model, a validation set used to limit overfitting, and a test set used to test the performance of our model. Learning rate reduction and early stopping by monitoring validation loss were also applied to our model to prevent overfitting. Additionally, we normalized our dataset to consider the scale between input and output variables and improve model stability. To reduce learning time, the proposed model is trained using distributed data-parallel (DDP) utilizing Pytorch v1.7.1 and GPU of NVIDIA GeForce RTX 3090.

In the process of obtaining the derivative for PDE calculation, the computational cost of PINNs increases exponentially. Therefore, we implemented space adaptive sampling to increase the efficiency of training. For every 1000 epochs, we added the top 200 data with the greatest error to the training dataset. Space adaptive sampling was compared with random sampling, which continuously adds 200 random data. The experiment in figure 3 verified that space adaptive sampling had a 30% accuracy improvement compared to random sampling.

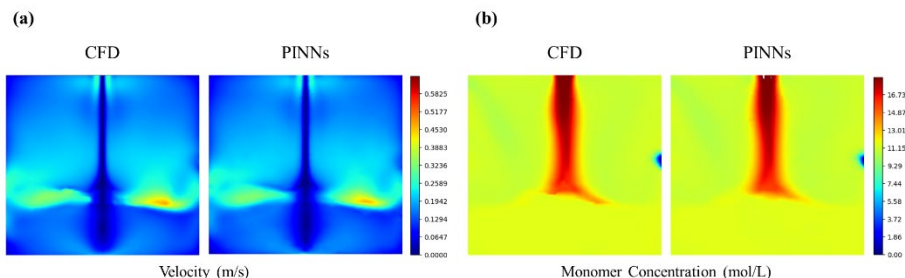


**Figure 3.** Comparison of CFD, random sampling, and space adaptive sampling model for velocity: Red points are the top 200 points with the largest empirical loss and are added to the training data.

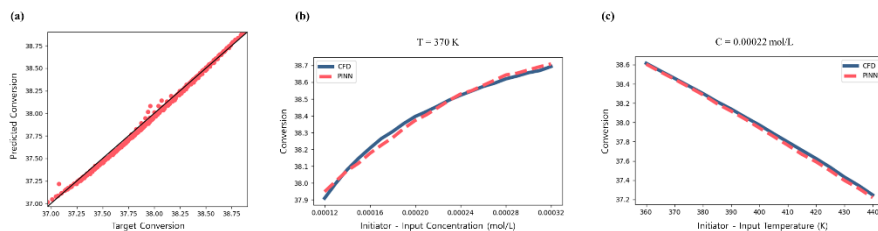
### 3. Results

#### 3.1. Performance of physics-informed neural networks

To check the validity of the proposed model, the target data and prediction of PINNs were compared by contour plot of velocity and monomer concentration. Figure 4 shows that PINNs prediction has high accuracy for both velocity and monomer concentration, and MSE between the target and predicted value was within 0.01%. In addition, to demonstrate that PINNs can reflect changes in driving conditions, we compared the conversion value between PINNs and CFD simulation data (Figure 5). The predicted conversion and target value were comparable, and the conversion graph form based on operating conditions was equivalent to CFD. The higher the inlet concentration of the initiator and the lower the inlet temperature of the initiator, the higher the conversion. To summarize, PINNs can accurately predict fluid flow and reaction.



**Figure 4.** Comparison of CFD and PINNs for (a) velocity and (b) monomer concentration



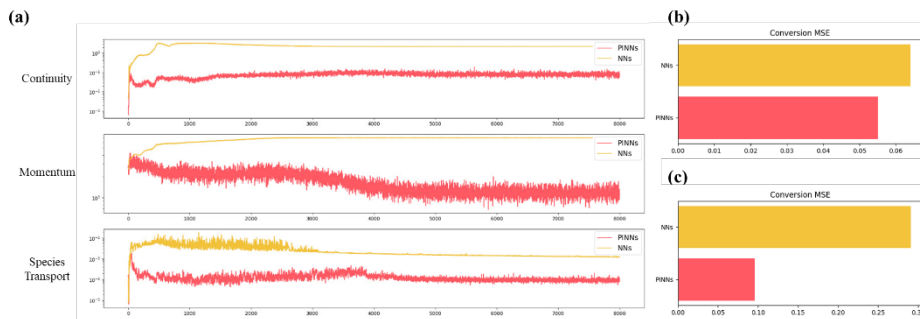
**Figure 5.** Comparison of CFD and PINNs for conversion: (a) Parity plot of the predictions for conversion with the proposed PINNs over the corresponding CFD results, (b) Conversion corresponding to input concentration of the initiator, (c) Conversion corresponding to input temperature of the initiator.

### 3.2. Operating condition interpolation & extrapolation

As the results show above, the suggested model is possible to search for optimal operating conditions, when applied to chemical reactors. However, there was no substantial difference between PINNs and neural networks (NNs). Therefore, to obtain the distinction of PINNs through additional experiments, the physical residual, interpolation, and extrapolation performance of PINNs were compared with NNs.

Figure 6 (a) compares the physical residuals of PINNs and NNs. The fact that PINNs have smaller physical residuals in governing equations indicates that it is trained to obey physical rules more than NNs. The bar plot in figure 6 (b) compares the interpolation and extrapolation performances of PINNs and NNs. Interpolation was performed inside the training data range, and extrapolation was derived using 264 simulated data outside of the training data range.

The accuracy of conversion prediction was tested, and PINNs outperformed NNs in both interpolation and extrapolation. The performance difference was greater in extrapolation than in interpolation. In the case of interpolation, the area to be predicted existed inside the training data range, hence the training data had a significant influence on interpolation results. Extrapolation, on the other hand, amplified the effect of the physics law, resulting in a considerable difference between PINNs and NNs.



**Figure 6.** Comparison of PINNs and NNs: (a) Physics loss of continuity, momentum, species transport equation, (b) Conversion MSE based on interpolation results, (c) Conversion MSE based on extrapolation results

## Conclusion

In this study, we proposed a novel PINNs model that can predict changes based on driving conditions. According to experimental results, the suggested method can explore optimal operating conditions by resolving nonlinear problems and interpolating solutions under various settings. Overall, we emphasize the significance of PINNs when it comes to data-driven system optimization and control. Our future work will be focused on geometric optimization to complete the PINNs-based reactor design framework. Following that, the PINNs-based framework is expected to be a key component of digital twin or multi-scale simulations.

## Notation

$x$	x coordinate
$y$	y coordinate
$z$	z coordinate
$u$	x-velocity
$v$	y-velocity
$w$	z-velocity
$T$	Temperature
$P$	Pressure
$I$	Initiator Concentration
$M$	Monomer Concentration
$E_a$	Activation Energy
$\rho$	Density
$k_i$	Reaction constant
$\omega$	Viscosity

## References

- Raissi, M., Perdikaris, P., & Karniadakis, G. E., 2019, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *Journal of Computational Physics*, 378, 686-707.
- Raissi, M., Yazdani, A., & Karniadakis, G. E., 2020, Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations, *Science*, 367(6481), 1026-1030.
- Cuomo, S., Di Cola, V. S., Giampaolo, F., Rozza, G., Raissi, M., & Piccialli, F., 2022, Scientific Machine Learning through Physics-Informed Neural Networks: Where we are and What's next, *arXiv preprint arXiv:2201.05624*.
- Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L., 2021, Physics-informed machine learning, *Nature Reviews Physics*, 3, 422-440.
- Raissi, M., Perdikaris, P., & Karniadakis, G. E., 2017, Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations, *arXiv preprint arXiv:1711.10561*.

# Extending the SI Decomposition to Continuous-Time Two-Stage Scheduling Problems

Daniel Montes<sup>a\*</sup>, José Luis Pitarch<sup>b</sup>, César de Prada,<sup>a,c</sup>

<sup>a</sup>*Department of Systems Engineering and Automatic Control, Universidad de Valladolid, C/ Prado de la Mergelina 3-5, Valladolid (47011), Spain.*

<sup>b</sup>*Instituto Univ. de Automática e Informática Industrial (ai2), Universitat Politècnica de Valencia, Camino de Vera s/n, Valencia (46022), Spain.*

<sup>c</sup>*Institute of Sustainable Processes, Universidad de Valladolid, C/ Dr. Mergelina s/n, Valladolid (47011), Spain.*

\*[danielalberto.montes.lopez@uva.es](mailto:danielalberto.montes.lopez@uva.es)

## Abstract

Scheduling often involves making decisions in presence of uncertainty, which governs the pricing of raw materials, energy, resource availability, demands, etc. A common approach to incorporate uncertainty in the decision-making process is using two-stage stochastic formulations. Unfortunately, the mathematical complexity of the resulting problems grows exponentially with the number of uncertainty scenarios, which is further complicated by the presence of binary variables. The authors have recently proposed a method using the so-called Similarity Index for discrete-time two-stage scheduling problems that enable scenario-based decomposition. This paper extends this method for scheduling problems formulated on a continuous-time basis. The fundamental idea is to use the Similarity Index to meet non-anticipation in the binary variables and Progressive Hedging on the continuous ones. The proposal is tested on a literature case study that consists of a multiproduct plant with a single processing unit. The combined SI-PH decomposition managed to solve the problem much faster than its monolithic counterpart.

**Keywords:** Similarity Index, Progressive Hedging, Optimization under uncertainty.

## 1. Introduction

The authors recently presented a decomposition method based on the idea of the Similarity Index (SI) to efficiently solve two-stage stochastic scheduling problems (TSSP) formulated using a discrete-time basis (Montes et al., 2022a, 2022b). TSSP handles uncertainty by discretizing the underlying probability distribution of uncertain parameters in a set of scenarios with associated probabilities. Then, the decision variables of the problem are grouped into two sets: first-stage and second-stage ones. In the first stage, the actual value that uncertain parameters will have is still unknown, but some decisions must be made regardless. Hence, the schedule within this first stage cannot be tailored to each scenario, as the uncertainty is not revealed yet. This is known as non-anticipation, i.e., decisions in the first stage must be unique. In the second stage, the values of the uncertain parameters are realized and assumed known, so the decisions can be adjusted to each realization accordingly.

Two-stage formulations are challenging to solve, as the problem size grows exponentially with the number of scenarios considered. This fact is further complicated by the presence of many binary variables, usual in scheduling formulations. One approach to tackle this



issue is to solve each scenario as an independent subproblem of reduced size, but the non-anticipation criterion for the first stage must be satisfied. The Progressive Hedging Algorithm (PHA) was originally proposed for convex problems involving only continuous variables and has convergence and optimality guarantees (Rockafellar & Wets, 1991). However, although the assumptions and guarantees vanish in presence of discrete variables, this method has been used successfully as a heuristics for mixed-integer problems (Bashiri et al., 2021; Peng et al., 2019). In the PHA, the non-anticipation constraints are relaxed, which enables scenario decomposition. Non-anticipation is then progressively enforced by penalizing the deviation from the average schedule in the first stage, i.e., the decision-variables average values computed from the subproblems solutions in a previous iteration.

Likewise, the Similarity-Index decomposition also enables scenario decomposition by approximating, or estimating, the degree of non-anticipation in the discrete decisions of the first stage. The scenario subproblems are solved independently and the similarity among their solutions is then measured using the SI. A weighing parameter is used to progressively increase the importance of maximizing the similarity. Eventually, all scenario solutions are equal in the first stage so non-anticipation is met. However, the SI decomposition was originally devised for discrete-time scheduling formulations. In this work, the authors aim to combine the SI decomposition with Progressive Hedging to decompose TSSP based on continuous-time formulations. The SI handles the binary variables of the problem, while Progressive Hedging handles the continuous ones. The proposal is tested successfully on a literature case study.

The paper is organized as follows: Section 2 presents a summary of the SI concept and its extension to continuous-time scheduling formulations. Section 3 describes the proposed SI-PH decomposition algorithm. The case study is presented in Section 4. A comparison between the proposed approach and the standard monolithic approach is summarized in Section 5. Conclusions and open issues are outlined in Section 6.

## 2. Similarity Index in Continuous-Time Formulations

In discrete-time basis, the Similarity Index relies on the idea of fuzzifying the binary decision variables along the surrounding time periods. The generated areas from the fuzzification for each scenario are compared and their overlap is used as a measure of similarity. Consequently, even if the decisions are not made in the same period among scenarios, one could quantify their similarity by temporal proximity. As the non-anticipation constraints (NAC) require the scenario schedules to be equal in the first stage, the SI can be used to remove the NACs from the formulation, enabling scenario decomposition. The SI can be pushed up through adding it to the objective function.

Two main difficulties arise for extending the Similarity-Index Decomposition to continuous-time TSSP: Time synchronization of decisions among scenarios; and weighing the slot duration as part of the SI computation. To get around these limitations, we propose: 1) not to consider the slot duration in the fuzzification process, and 2) to fuzzify in a slots basis instead of time periods. Then, the SI formula becomes:

$$SI: = \sum_{l \in \mathcal{L}_1} \frac{\min_{e \in \mathcal{E}} \|y_{le} + 0.5y_{(l+1)e} + 0.5y_{(l-1)e}\|_1}{2|\mathcal{L}| - 1} \quad (1)$$

Where the time slots  $l$  belong to an ordered set  $\mathcal{L}$ . The subset  $\mathcal{L}_1 \subset \mathcal{L}$  implies that the Similarity Index is only computed for the decisions made in the first stage of the problem.  $y_{le}$  are the decision variables indexed by the slots and the scenarios  $e \in \mathcal{E}$ . In essence, the Similarity Index is computed as the intersection of the generated areas from fuzzifying the decisions divided by the maximum intersection possible. Formula (1) implies that the similarity is only weighed in the two surrounding slots. As in the time-basis case (Montes et al., 2022b), this formula can be generalized for different fuzzification lengths.

Note that (1) does not cover the slot duration and other continuous variables that might be present in the problem (such as inventory levels, production levels, etc.). We propose to use Progressive Hedging (PH) to handle these variables, which results in a hybrid SI-PH method. Incorporating both the SI and the PH in a general TSSP, each subproblem formulation is given by:

$$\begin{aligned} \min_{x,y} J_e - \lambda SI_e + \sum_{l \in \mathcal{L}_1} \left( \omega_{le} x_{le} + \frac{\rho}{2} \left( \frac{x_{le} - \bar{x}_l}{\bar{x}_l} \right)^2 \right) \\ \text{s. t. :} \\ \text{Specific process constraints} \\ SI_e = \frac{\sum_{l \in \mathcal{L}_1} |s_{le}|_1}{2|\mathcal{L}| - 1} \\ s_{le} \leq y_{le} + 0.5y_{(l+1)e} + 0.5y_{(l-1)e} \quad \forall l \in \mathcal{L}_1 \\ s_{le} \leq \bar{y}_l + 0.5\bar{y}_{l+1} + 0.5\bar{y}_{l-1} \quad \forall l \in \mathcal{L}_1 \\ x \in \mathbb{R}, y \in \{0,1\} \end{aligned} \tag{2}$$

Note that the exact/global SI cannot be computed inside the subproblems to be solved independently, and that the element-wise operator  $\mathbf{min}\{\cdot\}$  in (1) is non-linear, so it cannot be used in the usual MILP scheduling formulations. Thus, the SI is estimated locally in each subproblem (2) by replacing the intersection of the fuzzified variables with a set of slack variables that are upper-bounded. The slack variables  $s_{le}$  are maximized in the objective function so this bound is tight. The  $SI_e$  is then an approximatin of the local SI by scenario, computed as the similarity of the scenario schedule with a reference schedule defined by fixed values  $\bar{y}_l$ . For details about this estimation refer to (Montes et al., 2022b).  $\lambda$  and  $\omega_{le}$  are the weighing parameters for the SI and the PH. They are updated in each iteration  $k$  from the scenario solutions, as follows, where  $\bar{x}_l$  is the expected value of  $x$ , using probabilities  $p_e$ , in the first stage:

$$\begin{aligned} \lambda^{(k+1)} = \lambda^{(k)} - \alpha(SI - 1); \quad \omega_{le}^{(k+1)} = \omega_{le}^{(k)} + \rho \left( \frac{x_{le} - \bar{x}_l}{\bar{x}_l} \right) \quad \forall l \in \mathcal{L}_1, e \in \mathcal{E} \\ \bar{x}_l = \sum_{e \in \mathcal{E}} p_e x_{le} \quad \forall l \in \mathcal{L}_1 \end{aligned} \tag{3}$$

The parameters  $\alpha$  and  $\rho$  are the step-size for  $\lambda$  and  $\omega$ . They greatly influence the algorithm behavior, both in terms of optimality and convergence, as discussed later.

### 3. SI-PH Decomposition

Non-anticipation can be enforced progressively using the SI and PH ideas. Hence, the original problem can be decomposed into scenario subproblems (2) by dropping the non-anticipation constraints. Each subproblem is solved independently and in parallel. The subproblems solutions are later used to update the required parameters for the SI and PH

with (3). The procedure is repeated until either the convergence criteria are met, or the maximum number of iterations is reached. Algorithm 1 below shows the pseudo-code for the SI-PH decomposition.

---

**Algorithm 1**


---

**Require**  $\alpha, \rho, k_{\max}, \Delta, \text{tol}$

1:  $k \leftarrow 0, \lambda^{(0)} \leftarrow 0, \omega_{le}^{(0)} \leftarrow 0, \bar{y}_l \leftarrow 0, \bar{x}_l \leftarrow 0$

2: **repeat**

3:   **for**  $e \in \mathcal{E}$  **do**

4:      $y_{le}^*, x_{le}^* \leftarrow \arg \min_{y_{le}, x_{le}, s_{le}} J_e - \lambda^{(k)} SI_e + \sum_{l \in \mathcal{L}_1} \left( \omega_{le}^{(k)} x_{le} + \frac{\rho}{2} \|x_{le} - \bar{x}_l^{(k)}\|_2^2 \right)$

5:      $SI_e^* \leftarrow SI_e(s_{le}^*)$

6:   **end for**

7:    $SI \leftarrow SI(y_{le}^*)$  using (1)

8:    $\bar{y}_l \leftarrow \arg \min_{y_{le}^*} SI_e^*$

9:    $\lambda^{(k+1)} \leftarrow \lambda^{(k)} - \alpha(SI^{(k)} - 1)$

10:    $\bar{x}_l^{(k+1)} \leftarrow \sum_{e \in \mathcal{E}} p_e x_{le}^{(k)} \quad \forall l \in \mathcal{L}_1$

11:    $\omega_{le}^{(k+1)} \leftarrow \omega_{le}^{(k)} + \rho \left( x_{le}^{(k)} - \bar{x}_l^{(k+1)} \right) \quad \forall l \in \mathcal{L}_1, e \in \mathcal{E}$

12:    $k \leftarrow k + 1$

13: **until**  $(SI = 1 \wedge \|x_{le}^{(k)} - \bar{x}_l^{(k+1)}\|_2 \leq \text{tol}) \vee k = k_{\max}$

14: **return**  $y_{le}^*, x_{le}^*$

---

The convergence criteria need to check that non-anticipation is met. For binary variables, that is easy to verify, as the SI needs to reach the value of 1. For continuous variables, non-anticipation can be confirmed if the 2-norm of the difference between the scenario solutions and the reference value is beyond a small enough tolerance.

#### 4. Case Study

The SI-PH decomposition was tested on a case study originally presented by Dogan & Grossmann (2006). The model uses a continuous-time representation for simultaneously integrating both the scheduling and planning of a continuous multiproduct plant with a single processing unit. The planning horizon is divided into fixed-duration time periods  $t \in \mathcal{T}$  (weeks), which are further subdivided into variable-duration time slots  $l \in \mathcal{L}$ . Products  $i \in \mathcal{I}$  are committed to each slot using binary variables  $W_{ilt}$ . Only the modifications to the original formulations are listed in this work for brevity and not the original equations.

The original formulation of Dogan & Grossmann (2006) is here extended to consider the production rates  $r_i$  as uncertain. Consequently, an additional index  $e \in \mathcal{E}$ , corresponding to the set of uncertainty realization scenarios, is added to all variables and equations of their model. As the time horizon is divided into both weeks and slots, additional variables  $W_{ije}^{\text{aux}}$ , with an additional auxiliary slot index  $j$ , are introduced to compute the Similarity Index as in (1). This variable is equal to the product assignment  $W_{ilt}$  but indexed by  $j$ . The index  $j$  enumerates the slots from 1 to  $N \cdot |\mathcal{T}_1|$ . Where  $N$  is the number of slots per time period and  $\mathcal{T}_1$  is the subset of time periods belonging to the first stage. For instance,

given time periods made up of three slots, the first slot of the second time period corresponds  $j = 4$ .

$$W_{ije}^{\text{aux}} = W_{ilte} \quad \forall j = l + |\mathcal{L}| \cdot (t - 1), l, t, e \quad (4)$$

The original formulation of the problem did not explicitly consider the slot duration  $\phi_{lte}$ . As the non-anticipation criterion of the two-stage formulation requires the slot duration variables to be equal among the scenarios in the first stage, some additional constraints are added to compute them from the starting time of the slots  $TS_{lte}$ :

$$\phi_{lte} = TS_{l+1,te} - TS_{lte} \quad \forall l \neq l_N, t \in \mathcal{T}_1, e \quad (5)$$

$$\phi_{l_N te} = HT_t - TS_{l_N te} \quad \forall t \in \mathcal{T}_1, e \quad (6)$$

Where  $HT_t$  is the elapsed time (in hours) from the beginning of the time horizon, at the end of each time period.

In addition to the economic function of the problem  $z^P$ , the objective function shall include the terms corresponding to the SI and the PH. The SI is computed for the product assignment variables  $W_{ije}^{\text{aux}}$  while the PH is applied to the slot duration  $\phi_{lte}$ , and the inventory levels  $INV_{ite}$ . The objective function to minimize in each subproblem is:

$$J_e := z_e^P - \lambda SI_e + \sum_{i,t \in \mathcal{T}_1, e} \omega_{ite} \frac{INV_{ite}}{\overline{INV}_{it}} + \sum_{l,t \in \mathcal{T}_1, e} \mu_{lte} \frac{\phi_{lte}}{\overline{\phi}_{lt}} + \frac{\rho_1}{2} \sum_{i,t \in \mathcal{T}_1, e} \left( \frac{INV_{ite} - \overline{INV}_{it}}{\overline{INV}_{it}} \right)^2 + \frac{\rho_2}{2} \sum_{l,t \in \mathcal{T}_1, e} \left( \frac{\phi_{lte} - \overline{\phi}_{lt}}{\overline{\phi}_{lt}} \right)^2 \bar{x}_l \quad (7)$$

After solving each subproblem, the multipliers  $\lambda$ ,  $\omega_{ite}$  and  $\mu_{lte}$ , and the expected values  $\overline{INV}_{it}$  and  $\overline{\phi}_{lt}$  need to be updated, as follows:

$$\begin{aligned} \lambda^{(k+1)} &= \lambda^{(k)} - \alpha(SI - 1); & \overline{INV}_{it} &= \sum_{e \in \mathcal{E}} p_e INV_{ite} \quad \forall i, t \in \mathcal{T}_1; \\ \omega_{ite}^{(k+1)} &= \omega_{ite}^{(k)} + \rho_1 \frac{INV_{ite} - \overline{INV}_{it}}{\overline{INV}_{it}} \quad \forall i, t \in \mathcal{T}_1, e; & \overline{\phi}_{lt} &= \sum_{e \in \mathcal{E}} \pi_e \phi_{lte} \quad \forall l, t \in \mathcal{T}_1; \\ \mu_{lte}^{(k+1)} &= \mu_{lte}^{(k)} + \rho_2 \frac{\phi_{lte} - \overline{\phi}_{lt}}{\overline{\phi}_{lt}} \quad \forall l, t \in \mathcal{T}_1, e \end{aligned} \quad (8)$$

Values  $\alpha$ ,  $\rho_1$ , and  $\rho_2$  are the tuning parameters of the algorithm.

## 5. Preliminary results

The combined SI-PH decomposition algorithm was tested on a case study instance of five products (A, B, C, D, E), eight weeks, and five slots. A total of eight uncertainty realization scenarios were considered. Problem 2a parameters from (Dogan & Grossmann, 2006) were used as a basis for the instance. The resulting problem had 17601 variables, 2240 of which were binary, and 17827 equations. GAMS 40.2.0 was used to code the model, and Gurobi 10.0 to solve it. All calculations were performed on a 2-CPU Xeon Gold 6130 computer. The optimality gap for Gurobi was set to 0.5% in all cases.

The monolithic instance was assigned 32 threads. The solver failed to provide an optimal solution after 10 hours of computation. The reported objective value of the best feasible solution found till that moment was  $z^P = 57643$ , reporting a 6.06% optimality gap.

Each of the subproblems in the decomposed instance was assigned 8 threads. Note that although the machine only has 32 cores, some scenarios are solved much faster than others, which frees cores for other threads. Note that, if too few threads are assigned per subproblem, the overall CPU capacity may be underutilized.

The tuning parameters greatly affected the quality of the solution. The values  $\rho_1 = 95$ ,  $\rho_2 = 120$ , and  $\alpha = 190$  provided the highest quality solution. The decomposition approach arrived at an objective value of  $z^P = 54682$  in 1552 seconds. Compared to the objective value of the monolithic approach, this solution is around 5% worse. However, just setting  $\rho_2 = 115$  reaches a much suboptimal  $z^P = 48717$ , and with  $\rho_2 = 125$  the algorithm did not converge after 2000 iterations.

## 6. Summary

This paper presented a hybrid decomposition method for TSSPs formulated on a continuous-time basis. The non-anticipation constraints are replaced by a combination of the Similarity Index and Progressive Hedging. This allows decomposing the original problem into smaller subproblems that are easier to solve. Their solutions are combined to iteratively build a high-quality feasible solution to the original problem.

Unfortunately, tuning the algorithm is not smooth due to the PH part, and small parameter variations lead to very different objective values. Some reported solutions are even far from optimality. Future work will focus on rethinking the PH decomposition strategy.

## Acknowledgments

These results are funded by the Spanish MCIN/AEI as part of the a-CIDiT (PID2021-123654OB-C31, PID2021-123654OB-C32) and LOCPU (PID2020-116585GB-I00) research projects. The first author has received financial support from the 2020 call for pre-doctoral contracts of the University of Valladolid and Banco Santander.

## References

- Bashiri, M., Nikzad, E., Eberhard, A., Hearne, J., & Oliveira, F. (2021). A two stage stochastic programming for asset protection routing and a solution algorithm based on the Progressive Hedging algorithm. *Omega (United Kingdom)*, 104.
- Dogan, M. E., & Grossmann, I. E. (2006). A decomposition method for the simultaneous planning and scheduling of single-stage continuous multiproduct plants. *Industrial and Engineering Chemistry Research*, 45(1), 299–315.
- Montes, D., Pitarch, J. L., & de Prada, C. (2022a). Decomposition of Two-stage Stochastic Scheduling Problems via Similarity Index. *Computer Aided Chemical Engineering*, 51, 985–990.
- Montes, D., Pitarch, J. L., & de Prada, C. (2022b). The Similarity Index to Decompose Two-Stage Stochastic Scheduling Problems. *IFAC-PapersOnLine*, 55(7), 821–826.
- Peng, Z., Zhang, Y., Feng, Y., Rong, G., & Su, H. (2019). A Progressive Hedging-Based Solution Approach for Integrated Planning and Scheduling Problems under Demand Uncertainty. *Industrial and Engineering Chemistry Research*, 58(32), 14880–14896.
- Rockafellar, R. T., & Wets, R. J. B. (1991). Scenarios and Policy Aggregation in Optimization Under Uncertainty. *Mathematics of Operations Research*, 16(1), 119–147.

# Modelling of water vapor adsorption/desorption on hopcalite in an industrial medical air dryer

Cristian Cardenas,<sup>a,b,\*</sup> & David Farrusseng<sup>a</sup>

<sup>a</sup>*Catalysis and Environment Research Institute Lyon, 2 Avenue Albert Einstein, 69626 Villeurbanne, France.*

<sup>b</sup>*MIL'S Society, 15 rue de Genève, 69746 Genas, France.*

\**cristian.cardenas@ircelyon.univ-lyon1.fr*

## Abstract

A dynamic multiscale model was developed to simulate the adsorption and desorption process in a fixed bed for air drying. This model is described by equations of mass balance, thermodynamics, and adsorption/desorption kinetics. The temperature-dependent Toth-Aranovich-Donohue equation was applied to predict the water vapor amount adsorbed on hopcalite. Experimental adsorption and desorption breakthrough curves at different water vapor concentrations and gas flow rates are measured and used to identify the different kinetic resistances, particularly the axial dispersion, external film mass transfer, pore diffusion, and internal mass transport. The model was implemented and solved within COMSOL Multiphysics<sup>®</sup>. The identified model is validated using different water vapor adsorption and desorption breakthrough curves from those used for parameter identification. The model predictions and the measurements showed a good agreement, quantified using performance indices and confirmed by a Kolmogorov-Smirnov statistic test.

**Keywords:** Water vapor adsorption, Hopcalite, Breakthrough curves, Modeling and simulation, Multiscale model.

## 1. Introduction

The air treatment system is an essential part of medical air installations: it ensures the drying and purification of the compressed air produced, to provide healthcare facilities with medical air that complies with the European Pharmacopeia n°1238-medical air in NF EN ISO 7396-1. A medical compressed air dryer is produced by pressure swing adsorption (PSA) using water-selective adsorbents. The dew point of the air treatment system reaches a minimum of -45°C which allows for high air quality. The production of dry medical air is ensured by air compressors and a column operating in PSA mode with a counter-current regeneration mode. The common industrial column is a fixed bed with various moisture adsorbents such as alumina, hopcalite and activated carbon in series in the form of granules. The state of the art is based on the choice of the nature of the adsorbents, their arrangement in the column and their respective volume (Cardenas et al, 2022). Qualitatively, the adsorbents placed at the bottom of the column are the least hydrophilic and adsorb the most moisture, while the adsorbents placed at the top of the column are the most hydrophilic and allow moisture adsorption at low concentrations. In the literature some models to describe the dynamic adsorption of water vapor on alumina are found, contrary to hopcalite where there is limited. Therefore, this study proposes to develop a fickian multiscale model to predict the adsorption and desorption breakthrough

fronts on hopcalite as a function of gas flow rate, concentration, and temperature. This model is described by mass balance, thermodynamic, and adsorption/desorption kinetics equations. Several parameters are needed to implement the model, thus the hopcalite is characterized, and water vapor adsorption isotherm measurements are carried out. Experimental adsorption and desorption breakthrough curves at different water vapor concentrations and gas flow rates are measured and used to estimate the axial dispersion, external film mass transfer, mixture diffusion and internal mass transport coefficients involved in the model equations, which are implemented, and solved within COMSOL Multiphysics®. The identified model is validated using different water vapor adsorption and desorption breakthrough curves from those used for parameter identification. The model predictions and the measurements showed a good agreement, quantified using performance indices and confirmed by a Kolmogorov-Smirnov statistic test. The validated model can be used as a predictive tool for the design and optimization of an industrial air dryer.

## 2. Adsorption isotherms

In the model development of the water vapor adsorption/desorption process on hopcalite, one of the main issues is to select the adsorption equilibrium isotherm that best describes the process. Experimental adsorption isotherms of water vapor on hopcalite in a temperature range of 288-313 K and pressures up to the saturated concentration were measured using a manometric BelsorpMax 2 instrument.

### *Toth-Aranovich-Donohue (Toth-AD) equation*

The temperature-dependent Toth-Aranovich-Donohue equation was applied to approximate the equilibrium relationship of water vapor on hopcalite. The model equation is as follows:

$$q_e = \frac{q_m \cdot b \cdot c}{\left[1 + (b \cdot c)^{t_0}\right]^{\frac{1}{t_0}} \left[1 - b_2 \cdot c\right]^{n_2}} \quad (1)$$

where  $q_e$  (mol.kg<sup>-1</sup>) is the equilibrium amount adsorbed,  $c$  (mol.m<sup>-3</sup>) is the water vapor concentration,  $q_m$  (mol.kg<sup>-1</sup>) is the maximum amount adsorbed,  $b$  and  $b_2$  (m<sup>3</sup>.mol<sup>-1</sup>) are the equilibrium constants at low and high concentrations, respectively,  $t_0$  is the Toth's heterogeneity factor and  $n_2$  is an adjustable dimensionless factor. The temperature-dependent parameters  $q_m$ ,  $b$ ,  $b_0$  and  $t_0$  are detailed in the study by Cardenas et al. (2022a).

## 3. Dynamic adsorption/desorption modeling

A comprehensive one-dimensional fickian multiscale model is developed to simulate the dynamic process of adsorption and desorption of water vapor. The model is based on mass balance, thermodynamic, and kinetic equations.

### *3.1. Model assumptions*

The assumptions of the model are as follows: (i) the gas mixture obeys the ideal gas law, (ii) water vapor is the only adsorbed molecule, (iii) the pressure drop along the bed is negligible, (iv) the resistance to mass transfer through the boundary layer surrounding the solid particles is characterized by an external film mass transfer coefficient, (v) the mass transfer within the particles is due to Fick's diffusion and characterized by an effective diffusion coefficient in the macropores, (vi) the mass transport of water vapor in the mesopores is due to mesopore diffusion, (vii) the Toth-Aranovich-Donohue equation describes the equilibrium isotherm of the gas phase with the adsorbent, (viii) the adsorbent particles are assumed to be spherical and homogeneous in size and density, (ix)

the temperature and physical properties of the adsorbent are assumed to be constant, (x) only the axial dimension is considered (one-dimensional model), and (xi) the adsorption and desorption process is reversible.

### 3.2. Fickian multiscale model

The multiscale model considers the bed scale and the particle scale. The particle is divided into two phases: a gas and a solid phase.

(i) In the bed column, the equation of the water vapor mass balance in the gas phase is expressed as follows:

$$\frac{\partial c}{\partial t} - D_{ax} \frac{\partial^2 c}{\partial z^2} + \frac{\partial(vc)}{\partial z} = -\frac{1-\varepsilon_b}{\varepsilon_b} \frac{3}{r_{pe}} k_f \left( c - c_{pe}|_{r_p=r_{pe}} \right) \left( \frac{1}{1+B_i} \right) \quad (2)$$

(ii) In the particle, assuming the absence of lateral and azimuthal variations of the concentration, the mass balance in the radial direction can be written as:

$$\varepsilon_{pe} \frac{\partial c_{pe}}{\partial t} + (1-\varepsilon_{pe}) \frac{\partial q}{\partial t} = \frac{D_e}{r_p^2} \frac{\partial}{\partial r_p} \left( r_p^2 \frac{\partial c_{pe}}{\partial r_p} \right) \quad (3)$$

(iii) In the particle, the water vapor adsorbed on the solid surface is approximated by the linear driving force LDF model and is written as:

$$\frac{\partial q}{\partial t} \approx \frac{\partial \bar{q}}{\partial t} = k_l (q_e^* - \bar{q}) \quad (4)$$

where  $c$  (mol.m<sup>-3</sup>) is the water vapor concentration in the gas phase,  $v = u/\varepsilon_b$  (m.s<sup>-1</sup>) is the interstitial gas velocity,  $D_{ax}$  (m<sup>2</sup>.s<sup>-1</sup>) is the axial dispersion coefficient,  $\varepsilon_b$  (m<sub>gas</sub><sup>3</sup>.m<sub>col</sub><sup>-3</sup>) is the bed porosity,  $r_{pe}$  (m) is the mean radius of the adsorbent particle,  $k_f$  (m.s<sup>-1</sup>) is the external film mass transfer coefficient,  $c_{p}|_{r_p=r_{pe}}$  (mol.m<sup>-3</sup>) is the water vapor concentration at the surface of the particle,  $B_i = r_{pe}k_f/(5D_e)$  is the material Biot number,  $D_e = \varepsilon_{pe}^{4/3}D_M$  (m<sup>2</sup>.s<sup>-1</sup>) is the effective diffusion coefficient, which depends on the porosity, and the mixture diffusion  $D_M$  (m<sup>2</sup>.s<sup>-1</sup>) in the particle,  $\bar{q}$  (mol.m<sup>-3</sup>) is the average water vapor amount adsorbed on the solid surface,  $q_e^* = q_e \rho_s$  (mol.m<sup>-3</sup>) is the equilibrium water vapor amount adsorbed,  $\rho_s$  (kg.m<sup>-3</sup>) is the solid density, and  $k_l$  (s<sup>-1</sup>) is the internal linear driving force factor. This equation is completed by the following initial and boundary conditions:

- For  $0 \leq z \leq L$ , at  $t = 0$ :  $c = 0$ ,  $c_{pe} = 0$ ,  $\bar{q} = 0$

- For  $t$  and  $r_p > 0$ , at  $z = 0$ :  $-D_{ax} \frac{\partial c}{\partial z} = v(c_0 - c)$ ,  $\bar{q} = 0$ , and at  $z = L$ :  $\frac{\partial c}{\partial z} = 0$

- For  $t$  and  $z > 0$ , at  $r_p = 0$ :  $\frac{\partial c_{pe}}{\partial r_p} = 0$ , and  $r_p = r_{pe}$ :  $-D_e \left( \frac{\partial c_{pe}}{\partial r_p} \right)_{r_p} = k_f \left( c - c_{pe}|_{r_p=r_{pe}} \right)$

Because of this reversibility, we can propose to treat the desorption breakthrough curve as if it were an adsorption breakthrough curve. Only the initial conditions and the boundary conditions are different. Therefore, the same model as described above is used along with the following modified initial and boundary conditions:

- For  $0 \leq z \leq L$ , at  $t = 0$ :  $c = c_0$ ,  $c_{pe} = c_{pe}|_{r_p=r_{pe}}$ ,  $\bar{q} = q_e$

- For  $t$  and  $r_p > 0$ , at  $z = 0$ :  $\frac{\partial c}{\partial z} = 0$

### 3.3. Implementation and simulation of the model

The external mass transfer coefficient is estimated using the Wakao and Funazkri (1978) correlation:

$$k_f = \frac{D_{AB}Sh}{2r_{pe}} \text{ with } Sh = c_{m1} + c_{m2}Re_p^{0.6}Sc^{1/3} \text{ and } c_{m1} = 2 \quad (5)$$



For the axial dispersion coefficient, the correlation of Rastegar and Gu (2017) was used:

$$D_{ax} = \frac{2r_{pe}v}{Pe'} \text{ with } \frac{1}{Pe'} = \frac{0.7D_{AB}}{2r_{pe}v} + \frac{\varepsilon_b}{c_{m3} + c_{m4}(Re_p)^{0.59}} \quad (6)$$

where  $Sh$ ,  $Re_p$ ,  $Sc$ , and  $Pe'$  are the dimensionless Sherwood, Reynolds, Schmidt and Peclet numbers respectively,  $D_{AB}$  ( $\text{m}^2 \cdot \text{s}^{-1}$ ) is the molecular diffusion coefficient. It is important to point out that  $D_{ax}$  and  $k_f$  increase with increasing gas flow rate. To keep this property,  $c_{m2}$ ,  $c_{m3}$  and  $c_{m4}$  are also considered as unknown parameters to be identified from the measurements (Esposito et al, 2022). A total number of five estimable parameters is therefore identified, i.e.,  $k_f$ ,  $D_M$ ,  $c_{m2}$ ,  $c_{m3}$  and  $c_{m4}$ . Comsol Multiphysics<sup>®</sup> version 6 is used to implement and solve the set of partial differential equations of the model. The estimability analysis algorithm using the orthogonalization algorithm described by Yao et al. (2003) is applied to the multiscale model and the value of the estimability threshold was set equal to 0.04 as suggested by Bouchkira et al., (2021). The following estimability order:  $k_f > D_M > c_{m2} > c_{m3} > c_{m4}$ . It should be noted that all parameters are estimable based on available measurements.

### 3.4. Model prediction performance

The prediction performance of the model is quantitatively assessed by means of criteria which compares model predictions with the experimental measurements. In this study, it is evaluated according to two different criteria: The root mean-square error function (*RMSE*) and the Pearson correlation coefficient (*r*).

## 4. Results and discussion

### 4.1. Parameters identification

The experimental data used consist of breakthrough curve measurements carried out at a room temperature of 20 °C, at gas flow rates of 90, 120, and 150  $\text{mL} \cdot \text{min}^{-1}$  with a water vapor concentration of 0.67  $\text{mol} \cdot \text{m}^{-3}$ , and at a gas flow rate of 135  $\text{mL} \cdot \text{min}^{-1}$  with four water vapor concentrations of 0.19, 0.38, 0.48, 0.72, and 0.77  $\text{mol} \cdot \text{m}^{-3}$  i.e., eighth adsorption and desorption experimental measurements to identified parameters. The optimal parameters identified and estimated during adsorption are used in the desorption simulation, except for the parameter  $k_f$ , which is estimated simultaneously from available desorption experimental measurements. Figures 1 compare the experimental measurements to the model predictions obtained using the optimized parameters in the adsorption and desorption process. The water vapor/hopcalite isotherm model used is the Toth-Aranovich-Donohue.

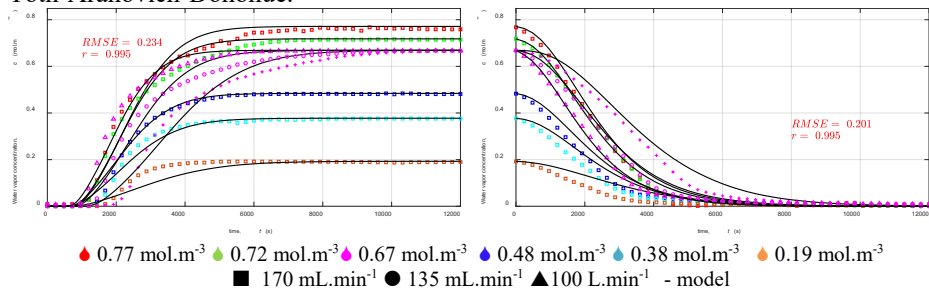


Figure 1: Comparison of water vapor adsorption (left) and desorption (right) breakthrough curves predicted with the measurements using the multiscale model.

The figure show that the predictions of the multiscale model are consistent with the experimental measurements because the Pearson correlation coefficient is close to 1

( $r=0.99$ ) and the root mean square error ( $RMSE<0.23$ ) is low. The optimized parameter values as well as their 95% confidence intervals ( $CI$ ), and the dimensionless material Biot number ( $Bi$ ) of is reported in the Table 1.

Table 1. Optimized parameter values for water vapor adsorption on hopcalite

Parameters	$B_i=4.63$					
	Adsorption			Desorption		
$c_{m2}$	$c_{m3}$	$c_{m4} \times 10^{-5}$	$D_M, (m^2 \cdot s^{-1})$	$k_{l,ads} (s^{-1})$	$k_{l,des} (s^{-1})$	
Value	5.93	0.005	1.01	$4.3 \times 10^{-5}$	0.065	0.0018
CI (95%)	$\pm 1.26$	$\pm 0.002$	$\pm 0.09$	$\pm 8.5 \times 10^{-6}$	$\pm 0.005$	$\pm 0.0006$

The fast saturation of the adsorbent bed at higher flow rates is since as the flow rate increases, the thickness of the film surrounding the particles decreases, reducing the resistivity value of the film for mass transfer. The value of the material Biot number (see Table 1) is greater than 3. Therefore, the adsorption mechanism is controlled by internal pore diffusion. Moreover, the breakthrough time, defined as the time required to reach 10% of the water vapor concentration in the bed column, decreases with increasing water vapor concentration at the inlet. It can also be observed that an increase in water vapor concentration at the inlet results in bed saturation being reached earlier due to relatively faster transport. Since the confidence interval values are lower than the parameter values reported in Table 1, the parameter identification can be considered accurate given the excellent fit indices. The coefficients for external film mass transfer and axial dispersion could be estimated with a different correlation than the one used, for example, by Delgado (2006) and Valdés-Solis et al., (2004) to improve the agreement between measurements and predictions. Despite this discrepancy, the values of the estimated coefficients at different gas flow rates reported in Table 1 are consistent with the magnitudes reported in the literature (Cardenas et al. 2021).

#### 4.2. Model validation

The identified model is then validated with additional measurements of water vapor adsorption and desorption breakthrough curves obtained on hopcalite at different temperatures, i.e., 15, 30, 40, and 50 °C, corresponding to the water vapor concentrations of 0.48, 0.47, 0.46, 0.45 mol.m<sup>-3</sup> and a gas flow rate of 150 mL.min<sup>-1</sup>. The results are shown in Figure 2. It is important to note that these operating conditions (concentration, gas flow rate, and temperature) are different from those used in parameter identification. The Kolmogorov-Smirnov statistical test is used and is based on the calculation of the divergence ( $D_n$ ) between two distributions. For more details, see the study by Cardenas et al. (2022b).

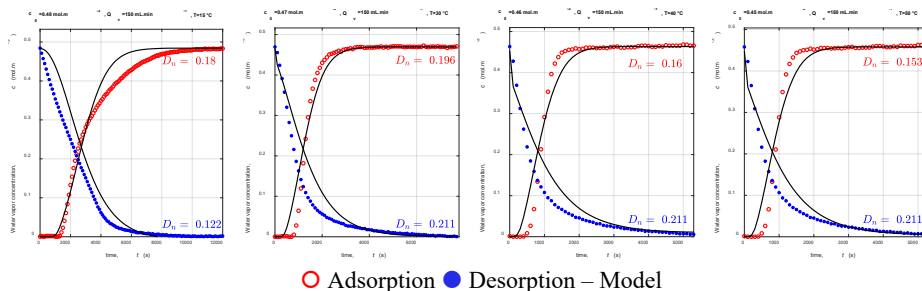


Figure 2: Comparison between predicted and measured breakthrough curves on hopcalite at four temperatures of 15, 30, 40, and 50 °C corresponding to water vapor concentrations of 0.48, 0.47, 0.46, 0.45, mol.m<sup>-3</sup> and a gas flow rate of 150 mL.min<sup>-1</sup>.

The calculated divergence values that are below the critical value ( $D_n^*=0.2$ ) allow us to accept the null hypothesis ( $H_0$ ), i.e., we can claim with 95% probability that the predicted and measured breakthrough curves have the same distribution. As shown in Figure 2, the predictions of the water vapor adsorption model guarantee this null hypothesis in almost all cases, so we can fully validate the multiscale model prediction of water vapor adsorption on hopcalite. On the other hand, the predictions of the water vapor desorption model, despite being very close to the critical value, cannot be guaranteed to be predicted.

## 5. Conclusions

The accuracy of the one-dimensional dynamic adsorption and desorption model using the temperature-dependent Toth-Aranovich-Donohue equation developed, identified, and validated for the simulation of a packed bed column in which water vapor is physically adsorbed on hopcalite was demonstrated. The fickian multiscale model allowed the finding that the adsorption process is controlled by the diffusion of water vapor in the macropores by a value of the material Biot number higher than 3. The statistical validation and the performance of the model performance indices ensure the quality of the model predictions, which is not always the case in modeling studies. This model coupled with an energy balance to improve the design and optimization of commercial adsorption columns using multilayer beds.

## References

- I. Bouchkira, A. M. Latifi, L. Khamar, S. Benjelloun, 2021. Global sensitivity based estimability analysis for the parameter identification of pitzer's thermodynamic model. *Reliability Engineering & System Safety*, 207, 107263.
- C. Cardenas, A. M. Latifi, C. Vallières, S. Marsteau, L. Sigot, 2021. Analysis of an industrial adsorption process based on ammonia chemisorption: Modeling and simulation. *Computers & Chemical Engineering*. 154, 107474.
- C. Cardenas, D. Farrusseng, C. Daniel, R. Aubry, 2022. Modelling of Water Vapor Adsorption and Desorption in an Industrial Medical Air Dryer. Available at SSRN 4259051.
- C. Cardenas, D. Farrusseng, C. Daniel, R. Aubry, 2022a. Modeling of equilibrium water vapor adsorption isotherms on activated carbon, alumina and hopcalite. *Fluid Phase Equilibria*, 561, 113520.
- C. Cardenas, A. M. Latifi, C. Vallières, S. Marsteau, 2022b. Analysis of an industrial adsorption process based on ammonia chemisorption: model validation. *Computer Aided Chemical Engineering*. 51, 457-462.
- J. Delgado, 2006. A critical review of dispersion in packed beds. *Heat Mass Transf.* 42, 279-310.
- F. Esposito, C. Cardenas, A. M. Latifi, S. Marsteau, 2022. CFD modeling and simulation of an ammonia adsorption process. *Computer Aided Chemical Engineering*. 51, 463-468.
- S. O. Rastegar, T. Gu, 2017. Empirical correlations for axial dispersion coefficient and peclet number in fixed-bed columns. *Journal of Chromatography*. 1490, 133-137.
- T. Valdés-Solís, M. Linders, F. Kapteijn, F. G. Marbán, & A. Fuertes, 2004. Adsorption and breakthrough performance of carbon-coated ceramic monoliths at low concentration of n-butane. *Chemical Engineering*. 59, 2791-2800.
- N. Wakao, T. Funazkri, 1978. Effect of fluid dispersion coefficients on particle-to-fluid mass transfer coefficients in packed beds. *Chemical Engineering Science*. 33, 10, 1375-1384.
- K. Z. Yao, B. M. Shaw, B. Kou, K. B. McAuley, D. W. Bacon, 2003. Modeling ethylene/butene copolymerization with multi-site catalysts: Parameter estimability and experimental design. *Polymer Reaction Engineering*. 11, 3, 563-588.

# Multi-fidelity Bayesian Optimisation of Reactor Simulations using Deep Gaussian Processes

Tom Savage,<sup>a</sup> Nausheen Basha,<sup>a</sup> Omar K. Matar,<sup>a</sup> Antonio del Rio Chanona,<sup>a</sup>

<sup>a</sup>*Imperial College London, London SW7 2BX, United Kingdom.*

## Abstract

Coiled tube reactors under pulsed-flow conditions have been shown to provide promising mixing characteristics. In order to validate performance in an industrial setting, and investigate the underlying physics of successful mixing, coiled tube reactors must be optimised. In this work, we apply a novel framework to locate optimal solutions to this nonlinear, computationally expensive, and derivative-free problem. Our optimisation framework takes advantage of deep Gaussian processes to learn a multi-fidelity surrogate model. We apply this model within a novel Bayesian optimisation algorithm, using faster, less accurate, and potentially biased lower-fidelity simulations to enable faster reactor optimisation. We subsequently investigate the physical insights into the swirling flows of these optimal configurations, directly informing the design of future coiled-tube reactors under pulsed-flow conditions. We demonstrate our design framework to be extensible to a broad variety of expensive simulation-based optimisation problems, supporting the design of the next-generation of highly parameterised chemical reactors.

**Keywords:** Optimisation, Computational Fluid Dynamics, Multi-fidelity Modelling

## 1. Introduction

### 1.1. Background

Achieving plug flow performance in chemical reactors is critical for processes involving the manufacture of chemicals, pharmaceuticals, biodiesel production, and waste treatment to minimise energy consumption and costs, and maximise product quality. Plug-flow conditions are dependent upon the reactor geometry, with product distributions of low variance being desired. Here, we study coiled-tube reactors, which have been shown to demonstrate promising plug flow performance in previous computational and experimental studies (Mansour, Liu et al. 2017, McDonough, Ahmed et al. 2019, McDonough, Murta et al. 2019, de Brito Lira, Riella et al. 2022).

Chemical reactors including coiled-tube reactors have been extensively explored through computational fluid dynamics (CFD) simulations, where numerical solutions are solved iteratively for a system of partial differential equations (PDEs) with large degrees of freedom resulting in large computational cost. In addition to being expensive, gradient information is practically unavailable. Derivative-free optimisation has found significant application in domains where mathematical expressions or gradients are unavailable. Examples include the optimisation of proprietary chemical process software (Caballero and Grossmann 2008, Savage, Basha et al. 2022), chemical reaction optimisation (Felton, Rittig et al. 2021), and topology optimisation of two-dimensional chemical reactor

channels (Cai, Guo et al. 2021). With the advent of new technologies in reactor design, reactor geometries are becoming more highly parameterised, resulting in higher-dimensional, more complex derivative-free optimisation problems.

In certain situations, computational expense can be traded off with accuracy via one or multiple fidelity parameters. Examples include training epochs in the context of hyper-parameter optimisation (Wu, Toscano-Palmerin et al. 2019, Schmucker, Donini et al. 2020), mesh fidelities in the context of finite element analysis (Huang, Allen et al. 2006), or combining real-time measurements and predictions in industrial processes (Petsagkourakis, Chachuat et al. 2021). Including fidelity control within a Bayesian optimisation framework enables optimisation with fewer computational resources whilst gaining a ‘high fidelity’ solution (Lindauer, Eggenesperger et al. 2019). Equation 1 describes this multi-fidelity problem

$$x^* = \arg \max_{x \in X} f(x, z^*). \quad (1)$$

where potentially M different fidelities,  $z \in \mathbb{R}^M$  become controllable parameters, and  $z^*$  indicates the highest-fidelity evaluation that is the one of interest.

**Contribution:** In this work, we present the real-world problem of optimising the geometry of a coiled-tube reactor to maximise its plug-flow performance. We apply a state-of-the-art deep GP (DGP)-based, multi-fidelity Bayesian optimisation algorithm to identify novel reactor configurations using an amalgam of different fidelity simulations, modelled using a DGP. Figure 1 demonstrates how our approach takes advantage of lower-fidelity simulations. Our approach contains no additional hyper-parameters when compared to standard Bayesian optimisation. Having identified an optimal geometry, we investigate the physical insights to inform future design of pulsed flow coiled tube reactors.

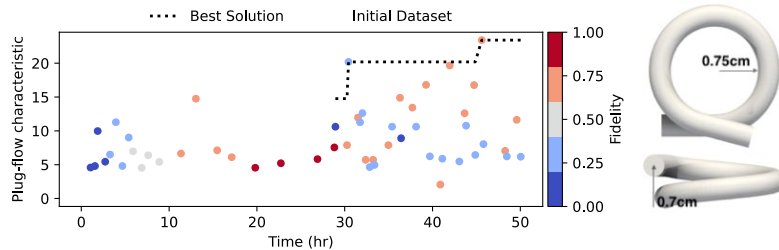


Figure 1: Left: The progression of multi-fidelity Bayesian optimisation using DGPs as a multi-fidelity model. Multiple different fidelity levels are selected for evaluation throughout optimisation, reducing optimisation time. Right: the best coil geometry which has a relatively large coil radius and low pitch.

## 2. Method

By applying DGPs within a Bayesian optimisation framework, we enable an end-to-end model of all fidelities. A more accurate model of higher fidelities should enable the optimisation procedure to make more evaluations at lower, less expensive fidelities.

### 2.1. Parameterisation

A coiled tube reactor is parameterised by a coil radius, and a coil pitch. Inversions within coiled tube reactors have been shown to provide effective mixing properties (Singh and

Nigam 2016, Rossi, Gargiulo et al. 2017, McDonough, Murta et al. 2019). Therefore, in addition to coil radius and coil pitch, we include an additional inversion parameter  $\delta$  to



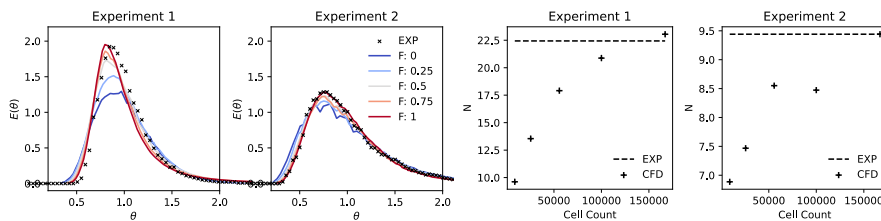
investigate geometries which involve a change of coil direction.  $\delta$  takes values between 0 and 1, and specifies where along the coil the inversion takes place. Figure 2 demonstrates the effect of  $\delta$  on helical tube geometry.

(a): The effect of  $\delta$  for a helical tube with fixed length, viewed from above.



(b): The effect of  $\delta$  for a helical tube with fixed length, viewed from side. Figure 2: From left to right:  $\delta$  is evaluated at 0, 0.15, 0.3, 0.6, and 0.75 with coil radius and pitch remaining constant.

Two fidelity aspects, axial and radial, can be varied when meshing is performed given a tube geometry. (Savage, Basha et al. 2022) demonstrates how these two fidelities affect the final mesh. Five discrete fidelities were simulated and compared with experimental data, in order to test assumptions regarding the accuracy of a simulation. Figure 3 validates the tracer concentration profile of simulations at five fidelities against two sets of experimental data. The objective of the optimisation is to maximise the number of equivalent tanks-in-series (McDonough, Murta et al. 2019), where high values of  $N$  are an indicator for good radial mixing and poor axial mixing. Figure 3 also demonstrates how increasing fidelity (and therefore cell count) results in a closer approximation to the experimental value of  $N$ , derived from each concentration profile. The accuracy of prediction when increasing cell-count in experiment 2 drops slightly.



*Left:* The concentration profile of a tracer injection at five fidelity levels between 0 and 1 against experimentally obtained data.  $E(\theta)$  represents dimensionless concentration as a function of dimensionless time  $\theta$

*Right:* The value of  $N$  corresponding to the concentration profile from each fidelity, which has been converted to cell count.  $N$  represents plug-flow characteristic which is the quantity optimised for.

Figure 3: Validation of five discrete mesh fidelities corresponding to different cell counts, across two sets of experimental data under different conditions.

## 2.2. Model Specification

To model each fidelity level we apply Deep Gaussian processes as demonstrated by (Cutajar, Pullin et al. 2019). DGPs provide a natural extension to sequentially trained multi-fidelity models, where a generating function  $f$  at a given discrete fidelity  $z$  is modeled as a linear or nonlinear function of lower fidelities plus a mismatch term. For example, in the nonlinear case,  $f_z(x)$  is given by

$$f_z(x) = \rho_{z-1}(f_{z-1}(x), x) + \phi_{z-1}. \quad (2.1)$$

Multi-fidelity DGPs combine  $\rho_{z-1}$  and  $\phi_{z-1}$  into a single term  $g_{z-1}$  which is modelled using a GP and trained end-to-end, resulting in a composition of GPs. For example, for  $Z$  different fidelity levels, and observations  $x$  at each fidelity, the highest fidelity is modeled as

$$f_z(x_z, \dots, x_1) = g_z(\dots g_{z-1}(g_1(x_1), x_{z-1}), x_z). \quad (2.2)$$

In this work, we assume five discrete fidelity levels. Thus the resulting DGP has five layers ( $Z = 5$ ), corresponding to simulations at fidelities equal to 0, 0.25, 0.5, 0.75, and 1. MF-DGPs are implemented in Python using Emukit (Paley, Pullin et al. 2019).

## 2.3. Discrete Multi-fidelity Bayesian Optimisation using Deep Gaussian Processes

Whilst continuous approximations have been applied for multi-fidelity Bayesian optimisation (Kandasamy, Dasarathy et al. 2017), we present an algorithm that takes advantage of the discrete fidelity predictions of a deep Gaussian process. Our approach is similar in spirit to that of (He, Tuo et al. 2017) and the experimental design work of (Thodoroff, Kaiser et al.). We perform a standard Bayesian optimisation step on the layer corresponding to the highest fidelity. Highlighting our interest only in the high-fidelity model. We then subsequently evaluated the variance of this solution with respect to the variance of other fidelities, providing insight into potential information gain. These variances are weighted by simulation cost enabling us to trade off information-gain with computational expense. Algorithm 1 highlights this approach.

---

### Algorithm 1 Deep GP-based Multi-fidelity Bayesian Optimisation

---

**Require:**  $f_1(x) \dots f_{z^*}(x)$ ,  $\mathcal{X}$ ,  $n$   
**for**  $z$  in  $1, \dots, z^*$  **do**  
    Generate  $n$  samples,  $\mathbf{x}_z$ , and evaluate  $f_t(\mathbf{x})$  resulting in  $\mathbf{y}_z$ .  
     $\tau_z \leftarrow$  average simulation time  
**end for**  
**while** Budget not exhausted **do**  
    Train DGP using  $\mathbf{x}_1, \dots, \mathbf{x}_{z^*}$  and  $\mathbf{y}_1, \dots, \mathbf{y}_{z^*}$   
    Solve UCB for highest-fidelity:  $x^* \leftarrow \arg \max_x \{ \mu_{z^*}(x) + \beta^{1/2} \sigma_{z^*}(x) | x \in \mathcal{X} \}$   
    Choose fidelity based on variance of DGP and simulation cost:  $z^* \leftarrow \arg \max_z \{ \gamma_z \beta^{1/2} \sigma_z(x^*) \}$  where  $\gamma_z = \max(\tau) / \tau_z$   
    Evaluate  $f_{z^*}(x^*)$ , add  $x^*$  to  $\mathbf{x}_{z^*}$  and  $f_{z^*}(x^*)$  to  $\mathbf{y}_{z^*}$   
**end while**

---

We note that in choosing  $\gamma_z$  based on the points sampled to construct the initial DGP, the algorithm contains no additional hyper-parameters than standard UCB Bayesian optimisation. In addition,  $\gamma_z$  may be updated after a simulation has been performed, more accurately reflecting the simulation time at that fidelity.

## **Experimental Results**

Figure 1 demonstrates the MF-DGP based optimisation, which shows the selection of multiple simulation fidelities throughout optimisation. We note that initialising the DGPs for multi-fidelity modelling requires simulations at each fidelity. This presents a disadvantage in discrete over both single-fidelity and continuous fidelity settings, as the number of fidelities grows so does the initial data required. The number of initial data points and at what respective fidelities is the subject of future work.

### *2.4. Recommendations*

Overall, we find that it may not be beneficial in situations where a large number of discrete fidelities are available, to apply all fidelities within an optimisation framework. A large number of fidelities results in a large number of hyper-parameters, more difficult to train multi-fidelity models such as DGPs with more layers resulting in longer inference times, and potentially slower exploration. If a discrete multi-fidelity Bayesian optimisation method is to be used, we make the recommendation to apply 2 or 3 discrete fidelities in systems with no prior knowledge, despite more being available. Alternatively, if more fidelities are required, modelling multi-fidelity data in a continuous space and applying an appropriate algorithm such as the one presented in (Kandasamy, Dasarathy et al. 2017) may be necessary. Future work will apply multi-fidelity Bayesian optimisation approaches in continuous domains, as well as apply similar methods on more industrial case studies for engineering discovery. Further experiments will be performed to correlate dimensionality of the problem with the number of discrete fidelities.

## **4. Conclusions**

The optimisation of coiled tube reactor geometry is critical to maximise plug-flow behaviour and investigate industrial viability. The optimisation problem is formulated as an expensive black-box problem. In this paper, we propose multi-fidelity Bayesian optimisation using Deep Gaussian processes to find good reactor configurations, taking advantage of less accurate but faster simulations. We demonstrate experimental validation of five discrete fidelities and present a modified multi-fidelity Bayesian optimisation algorithm which relies on fewer hyper-parameters than existing approaches. A multi-fidelity DGP provides correct quantification and propagation of uncertainty, which we use to select not only the next experimental design, but also the fidelity of the evaluation within the algorithm. Our approach can be extended to other problems involving parameterised CFD simulations. This work demonstrates an industrially relevant use case of multi-fidelity deep Gaussian processes for the optimisation of expensive black-box functions. We hope it provides insight and inspiration for people in the chemical engineering community to apply multi-fidelity methods across a variety of potential problems.

## **Acknowledgements**

The authors acknowledge funding by the Engineering & Physical Sciences Research Council, United Kingdom through the PREMIERE (EP/T000414/1), and thank Ilya Orson Sandoval for providing insights on this work. The authors would also like to thank Dr Jonathan McDonough, Newcastle University. Tom Savage acknowledges the support of the Imperial College President's scholarship.



## References

- Caballero, J. A. and I. E. Grossmann (2008). "An algorithm for the use of surrogate models in modular flowsheet optimization." *AIChE journal* **54**(10): 2633-2650.
- Cai, H., K. Guo, H. Liu, C. Liu and A. Feng (2021). "Derivative-free level-set-based multi-objective topology optimization of flow channel designs using lattice Boltzmann method." *Chemical Engineering Science* **231**: 116323.
- Cutajar, K., M. Pullin, A. Damianou, N. Lawrence and J. González (2019). Deep Gaussian Processes for Multi-fidelity Modeling.
- de Brito Lira, J. O., H. G. Riella, N. Padoin and C. Soares (2022). "Fluid dynamics and mass transfer in curved reactors: A CFD study on Dean flow effects." *Journal of Environmental Chemical Engineering* **10**(5): 108304.
- Felton, K. C., J. G. Rittig and A. A. Lapkin (2021). "Summit: benchmarking machine learning methods for reaction optimisation." *Chemistry-Methods* **1**(2): 116-122.
- He, X., R. Tuo and C. J. Wu (2017). "Optimization of multi-fidelity computer experiments via the EQIE criterion." *Technometrics* **59**(1): 58-68.
- Huang, D., T. T. Allen, W. I. Notz and R. A. Miller (2006). "Sequential kriging optimization using multiple-fidelity evaluations." *Structural and Multidisciplinary Optimization* **32**(5): 369-382.
- Kandasamy, K., G. Dasarathy, J. Schneider and B. Póczos (2017). *Multi-fidelity bayesian optimisation with continuous approximations*. International Conference on Machine Learning, PMLR.
- Lindauer, M., K. Eggensperger, M. Feurer, A. Biedenkapp, J. Marben, P. Müller and F. Hutter (2019). BOAH: A Tool Suite for Multi-Fidelity Bayesian Optimization; Analysis of Hyperparameters.
- Mansour, M., Z. Liu, G. Janiga, K. D. P. Nigam, K. Sundmacher, D. Thvenin and K. Zhringer (2017). "Numerical study of liquid-liquid mixing in helical pipes." *Chemical Engineering Science* **172**: 250-261.
- McDonough, J. R., S. M. R. Ahmed, A. N. Phan and A. P. Harvey (2019). "The development of helical vortex pairs in oscillatory flows } A numerical and experimental study." *Chemical Engineering and Processing - Process Intensification* **143**: 107588.
- McDonough, J. R., S. Murta, R. Law and A. P. Harvey (2019). "Oscillatory fluid motion unlocks plug flow operation in helical tube reactors at lower Reynolds numbers (Re  $\leq 10$ )." *Chemical Engineering Journal* **358**: 643-657.
- Paley, A., M. Pullin, M. Mahserici, C. McCollum, N. Lawrence and J. González (2019). *Emulation of physical processes with Emukit*. Second Workshop on Machine Learning and the Physical Sciences, NeurIPS.
- Petsagkourakis, P., B. Chachuat and E. A. del Rio-Chanona (2021). Safe Real-Time Optimization using Multi-Fidelity Gaussian Processes.

Rossi, D., L. Gargiulo, G. Valitov, A. Gavriilidis and L. Mazzei (2017). "Experimental characterization of axial dispersion in coiled flow inverters." Chemical Engineering Research and Design **120**: 159-170.

Savage, T., N. Basha, O. M. Ehecattl and A. D.-R. Chanona (2022). "Deep Gaussian Process-based Multi-fidelity Bayesian Optimization for Simulated Chemical Reactors." arXiv preprint arXiv:2210.17213.

Schmucker, R., M. Donini, V. Perrone and C. Archambeau (2020). Multi-objective multi-fidelity hyperparameter optimization with application to fairness. NeurIPS 2020 Workshop on Meta-learning.

Singh, J. and K. Nigam (2016). "Pilot plant study for effective heat transfer area of coiled flow inverter." Chemical Engineering and Processing: Process Intensification **102**: 219-228.

Thodoroff, P., M. Kaiser, R. Williams, R. Arthern, S. Hosking, N. Lawrence, J. Byrne and I. Kazlauskaitė "Multi-fidelity experimental design for ice-sheet simulation."

Wu, J., S. Toscano-Palmerin, P. I. Frazier and A. G. Wilson (2019). Practical Multi-fidelity Bayesian Optimization for Hyperparameter Tuning.



# Design of a sustainable biodiesel production process by a multi-objective optimization

Barcia-Quimi A.F.<sup>a,\*</sup>, Risco-Bravo A.<sup>a</sup>, Alcivar-Espinoza K.<sup>a</sup>, Tinoco-Caicedo D.L.<sup>a,b</sup>

<sup>a</sup> *Facultad de Ciencias Naturales y Matemáticas (FCNM), Escuela Superior Politécnica del Litoral Ecuador, 090903 Guayaquil, Ecuador.*

<sup>b</sup> *Centro de Energías Renovables y Alternativas (CERA), Escuela Superior Politécnica del Litoral Ecuador, 090903 Guayaquil, Ecuador.*

\*[abarcia@espol.edu.ec](mailto:abarcia@espol.edu.ec)

## Abstract

A multi-objective optimization is proposed for determining the optimal operating conditions for biodiesel production. The objective functions considered were the production cost, the carbon footprint, and the yield of palm oil transesterification. The kinetic model proposed by Narváez et al. (2015) was validated with state-of-the-art experimental data. The optimization was conducted using Python 3.8.8 and openLCA 1.11.0 to determine environmental parameters. Four-decision variables were optimized by using 3 methodologies: Pareto Set, R-Method, and a Techno-economic analysis. The optimal solution resulted in a maximum yield of 80.29% with a minimum production cost of 0.44\$/kg, and a minimum carbon footprint of 0.144 kg CO<sub>2</sub>-eq/kg.

**Keywords:** Multi-objective Optimization, Kinetic model, Carbon footprint, Biodiesel

## 1. Introduction

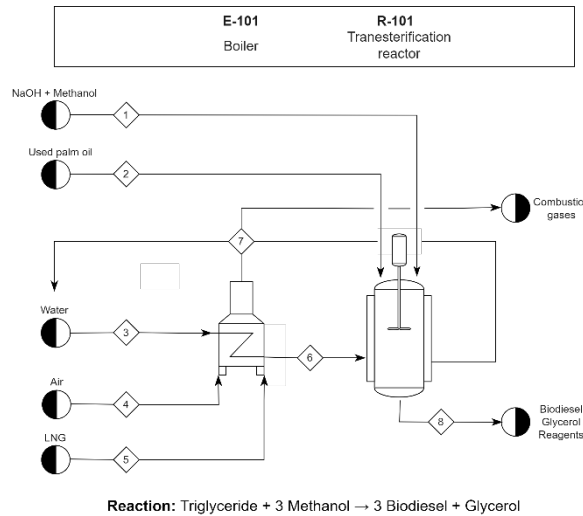
Nowadays, the predominant energy source is fossil fuels. However, research on alternative fuels has begun due to increased energy demand and limited fossil resources. One major alternative is biodiesel production which is now implemented in many countries. Nevertheless, the drawback of the industrialization of this process is the high volume of reactants used such as methanol and conventional catalysts, which leads to a high production cost and a high carbon footprint. For instance, an Ecuadorian company produces 144 thousand tons per year (Roca Meza, 2014), and the production costs exceed \$1000/ton. This causes the price of biodiesel to be 5.56 \$/gal (Clickgasoil, 2021) higher than diesel 4.33 \$/gal (EP PETROECUADOR, 2022). Several experimental studies have focused on finding the optimal conditions to maximize performance. Castillo González et al. (2020) shows that, for the formation of alkyl esters, a molar ratio of alcohol:oil 6:1 result in a high yield of biodiesel. Narváez et al. (2007) performed experimental tryouts on the kinetics of palm oil transesterification, obtaining high-yield results with elevated temperatures and molar ratios. Suzihaque et al. (2022) state that the catalyst's concentration is an important variable to consider since low concentrations imply a

smaller catalytic surface area to conduct the transesterification reaction, while high concentrations could lead to saponification. Either way, yield is affected negatively. Research on biodiesel optimization (Pugazhendhi et al., 2020) is often focused on maximizing yield. However, an optimization also including production cost and carbon footprint would be a suitable approach toward sustainable production. Therefore, the scope of this study is to perform a multi-objective optimization that minimizes the production cost and the carbon footprint and maximizes the yield of the transesterification reaction of palm oil.

## 2. Methodology

### 2.1. System description

Figure 1 shows the process flow for the biodiesel batch production proposed. A NaOH-methanol mixture (Stream 1) and palm oil (Stream 2) enter the batch-type reactor (R-101) for the transesterification reaction. R-101 uses steam supplied (Stream 6) by the boiler (E-101). E-101 works with liquefied natural gas (Stream 5) as fuel and air intake (Stream 4) at ambient conditions for combustion.



**Figure 1.** Process flow diagram to produce biodiesel from waste oil

### 2.2. Kinetic model and Validation

The kinetic model for the palm oil's transesterification reaction was implemented in Python 3.8.8. The model was based on a previous study from Narváez et al. (2015) which assumes that palm oil's transesterification reaction involves three reversible chain reactions that follow an elementary rate law. Equations 1 to 6 present the kinetic model,

$$\frac{d[TG]}{dt} = -k_1 e^{\frac{E_1}{RT}} \left( \frac{C_{cat}}{k_{1cat} + C_{cat}} \right) [TG][M] + k_{-1} e^{\frac{E_{-1}}{RT}} \left( \frac{C_{cat}}{k_{-1cat} + C_{cat}} \right) [DG][EM] \quad (1)$$

$$\frac{d[DG]}{dt} = k_1 e^{\frac{E_1}{RT}} \left( \frac{C_{cat}}{k_{1cat} + C_{cat}} \right) [TG][M] - k_{-1} e^{\frac{E_{-1}}{RT}} \left( \frac{C_{cat}}{k_{-1cat} + C_{cat}} \right) [DG][EM] - k_2 e^{\frac{E_2}{RT}} \left( \frac{C_{cat}}{k_{2cat} + C_{cat}} \right) [DG][M] + k_{-2} e^{\frac{E_{-2}}{RT}} \left( \frac{C_{cat}}{k_{-2cat} + C_{cat}} \right) [MG][EM] \quad (2)$$

$$\frac{d[MG]}{dt} = k_2 e^{\frac{E_2}{RT}} \left( \frac{C_{cat}}{k_{2cat} + C_{cat}} \right) [DG][M] - k_{-2} e^{\frac{E_{-2}}{RT}} \left( \frac{C_{cat}}{k_{-2cat} + C_{cat}} \right) [MG][EM] - k_3 e^{\frac{E_3}{RT}} \left( \frac{C_{cat}}{k_{3cat} + C_{cat}} \right) [MG][M] + k_{-3} e^{\frac{E_{-3}}{RT}} \left( \frac{C_{cat}}{k_{-3cat} + C_{cat}} \right) [G][EM] \quad (3)$$

$$\frac{d[EM]}{dt} = k_1 e^{\frac{E_1}{RT}} \left( \frac{C_{cat}}{k_{1cat} + C_{cat}} \right) [TG][M] - k_{-1} e^{\frac{E_{-1}}{RT}} \left( \frac{C_{cat}}{k_{-1cat} + C_{cat}} \right) [DG][EM] + k_2 e^{\frac{E_2}{RT}} \left( \frac{C_{cat}}{k_{2cat} + C_{cat}} \right) [DG][M] - k_{-2} e^{\frac{E_{-2}}{RT}} \left( \frac{C_{cat}}{k_{-2cat} + C_{cat}} \right) [MG][EM] + k_3 e^{\frac{E_3}{RT}} \left( \frac{C_{cat}}{k_{3cat} + C_{cat}} \right) [MG][M] - k_{-3} e^{\frac{E_{-3}}{RT}} \left( \frac{C_{cat}}{k_{-3cat} + C_{cat}} \right) [G][EM] \quad (4)$$

$$\frac{d[G]}{dt} = k_3 e^{\frac{E_3}{RT}} \left( \frac{C_{cat}}{k_{3cat} + C_{cat}} \right) [MG][M] - k_{-3} e^{\frac{E_{-3}}{RT}} \left( \frac{C_{cat}}{k_{-3cat} + C_{cat}} \right) [G][EM] \quad (5)$$

$$\frac{d[M]}{dt} = -\frac{d[EM]}{dt} \quad (6)$$

where *TG* represents triglycerides, *DG* diglycerides, *MG* monoglycerides, *G* glycerol, *EM* methyl esters, and *M* methanol. This model considers the influence of the reaction temperature and catalyst concentration on the kinetics of the reaction.

To validate the kinetic model, the degree of fit of the experimental kinetic data of Narváez et al. (2007) and Leevijit et al. (2004) were compared with the curve predicted by the model. Then, experimental data from Narváez et al. (2007), Cheng et al. (2004), Silva et al. (2021) and Wongjaikham et al. (2021) were used to compare their reported yield (*Y*%) or conversion (*C*%) results with those predicted by the kinetic model according to Equations 7 and 8 respectively. The relative errors between the experimental and model-predicted results were determined.

$$C \% = \frac{[TG]_0 - [TG]_f}{[TG]_0} \times 100\% \quad (7) \quad Y \% = \frac{[EM]_f}{3 [TG]_0} \times 100\% \quad (8)$$

### 2.3. Multi-objective optimization

For the multi-objective optimization, an algorithm was developed and implemented in Python 3.8.8. The independent variables considered were *reaction temperature* (*T*), *reaction time* (*t*), *NaOH catalyst concentration* (*C<sub>cat</sub>*), and *methanol: oil molar ratio* (*RM*). Therefore, the vector with *e* decision variables is presented in Equation 9:

$$x = (T, t, C_{cat}, RM) \quad (9)$$

The optimization algorithm started by evaluating each of the independent variables within the objective functions included in *Table 1*. The openLCA 1.11.0 program along with the *ef\_secondarydata\_202202* database was used to propose the carbon footprint objective function.

Three different strategies were used to discriminate the results of the solution set. First, the Pareto set was determined using the Python package *pareto* 1.2.0. This package is based on the skyline operator methodology that identifies the non-dominated solutions of the set, i.e., those that are the best in all or at least one of the objectives. Then, the R-Method proposed by Rao & Lakshmi (2021) was used. This is a ranking algorithm to select the *best* solution based on the level of importance of each of the objectives. In this case, yield and production costs were set as objectives with the same degree of interest. Finally, the ten solutions with the highest score from the R-Method analysis were selected to perform a techno-economic analysis. The methodology proposed by Turton et al.

(2018) was used to determine the cash flow, Internal Rate of Return (IRR), Net Present Value (NPV), and Payback Period for each solution. The solution with the highest IRR, NPV, and shortest payback period was selected as the optimal solution. This solution would not only maximize yield, minimize production costs, and carbon footprint but also would have the best economic potential.

**Table 1.** Objective functions

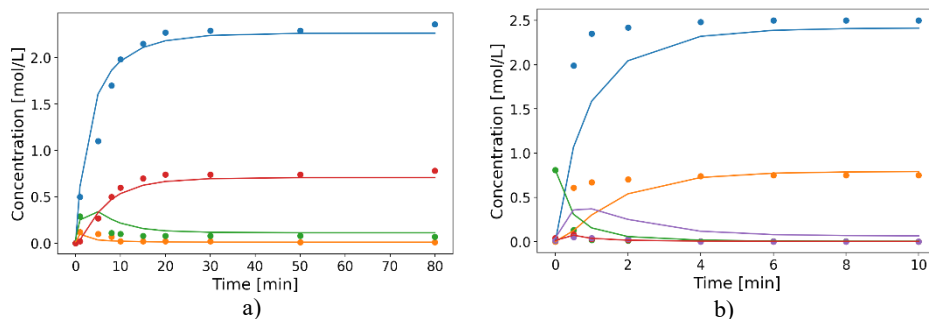
Yield* (%)	Production Cost** (\$/kg)	Carbon footprint** (kg CO <sub>2</sub> -eq/kg)
$\max_x \frac{[EM]_f}{3 * [TG]_0} \times 100\%$	$\min_x \frac{C_M + C_{Cat} + C_E + C_C}{M_{bio}}$	$\min_x \frac{CF_M + CF_{Cat} + CF_E + CF_F}{M_{bio}}$
Restrictions		
$45 \leq T (^{\circ}C) \leq 60$	$30 \leq t (min) \leq 120$	$3:1 \leq RM \leq 10:1$
		$0.5 \leq C_{cat} \left(\frac{\%w}{w}\right) \leq 1.5$

\* $[EM]_f$  the final concentration of methyl ester,  $[TG]_0$  the initial concentration of triglycerides.

\*\* $C$  and  $CF$  represent the cost and carbon footprint of methanol ( $M$ ), catalyst ( $Cat$ ), electricity ( $E$ ), and fuel ( $F$ ).  $M_{bio}$  is the mass of biodiesel produced.

### 3. Results

The kinetic model for the palm oil's transesterification reaction was validated. As shown in *Figure 2*, the theoretical kinetic model in comparison with the data from Narváez et al. (2007) (a) and Leevijit et al. (2004) (b) have a suitable approximation with  $R^2$  values of 0.983 and 0.924 respectively.



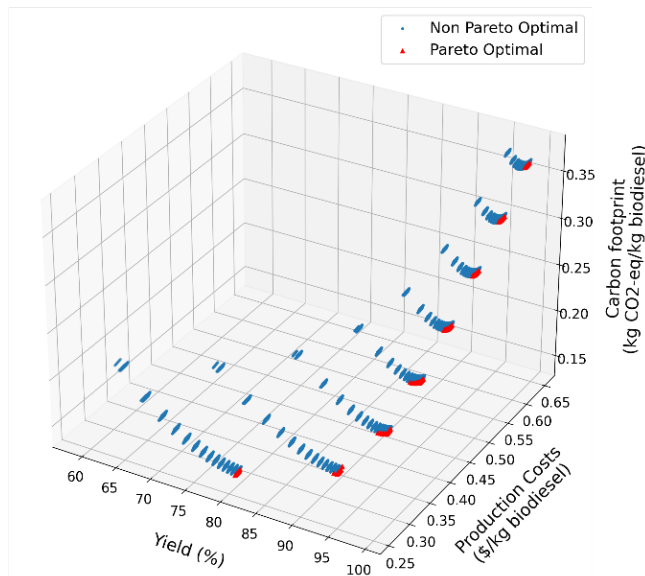
**Figure 2.** Comparison between experimental data (●) and the kinetic model (—) for palm oil transesterification at RM 6:1 and a)  $T = 50^{\circ}C$  and  $C_{cat} = 0.2\%$  and b)  $T = 60^{\circ}C$  and  $C_{cat} = 1.0\%$

*Table 2* presents the experimental and model-predicted results for conversion or yield. The relative errors are below 5%. Based on this, it can be concluded that the model proposed by Narváez et al. (2015) is highly reliable to describe the kinetics of palm oil transesterification. Also, that is possible to extrapolate the model for the selected operating ranges and obtain acceptable results.

The solution set consisted in 12544 results for *Yield (%)*, *Production Cost (\$/kg)* and *Carbon Footprint (kg CO<sub>2</sub>-eq/kg)*. As yield increases, production costs and carbon footprint also increase as shown in *Figure 3*. The Pareto set contained 286 solutions that are on the lower boundary of the solution set showing that these are the non-dominated solutions according to the criteria of this optimization.

**Table 2.** Kinetic model validation results

Reference	T (°C)	t (min)	C <sub>cat</sub> (% w/w)	RM	Parameter	Experimental	Model	% Error
Cheng et al. (2004)	60.5	7	0.5	10:1	C %	99.00	99.77	0.78
Narváez et al. (2007)	60.0	80	1.0	6:1	Y %	97.60	96.94	0.68
Silva et al. (2021)	45.0	60	1.0	4:1	Y %	88.50	90.08	1.75
Wongjaikham et al. (2021)	62.5	40	1.5	3:1	Y %	77.00	80.41	4.39



**Figure 3.** Yield, production cost and carbon footprint results of the solution set

The operating conditions that maximize yield while minimizing production costs and carbon footprint are  $T = 45^{\circ}\text{C}$ ,  $t = 30$  min,  $\text{RM} = 3:1$  and  $C_{\text{cat}} = 1.5\%$ . Under these conditions, a biodiesel production yield of 80.29%, a production cost of 0.26 \$/kg, and a carbon footprint of 0.144 kg CO<sub>2</sub>-eq/kg is achieved. This production cost only includes raw materials and reactor utilities as proposed in its objective function. However, in the techno-economic analysis, a final cost of 0.44 \$/kg was estimated considering other costs such as operating labor and maintenance.

Similar proposals for biodiesel production via vegetable oil transesterification obtained a biodiesel cost between 0.58 \$/kg (El-Gharbawy, 2017) and 1.35 \$/kg (Mohammadshirazi et al., 2014). The product cost can be reduced up to 67.37% when the process operates at the obtained optimal conditions. Therefore, this solution turns out to be the most favorable not only in the studied range but also against other proposals. In addition, the optimal solution makes the proposal attractive and viable, with an IRR of 128.94% and a NPV of \$1.42 million. The invested capital would be recovered in 8 months and 12 days.



#### 4. Conclusions

The multi-objective optimization algorithm determined the optimal temperature, molar ratio, catalyst concentration and reaction time. This solution not only maximizes yield while minimizing production costs and carbon footprint but also has the best economic potential. In addition, this solution reduces production costs by up to 67.37% compared to other similar proposals. The proposed framework could be a potential tool for developing countries to target the biofuel market with sustainable production. In future work, a different optimization algorithm could be used in order to compare and further evaluate the obtained results.

#### References

- Castillo González, J. P., Álvarez Gutiérrez, P. E., Adam Medina, M., López Zapata, B. Y., Ramírez Guerrero, G., & Vela Valdés, L. (2020). Effects on Biodiesel Production Caused by Feed Oil Changes in a Continuous Stirred-Tank Reactor. *Applied Sciences*, *10*(3), 992.
- Cheng, S. F., Choo, Y. M., Ma, A. N., & Chuan, C. H. (2004). Kinetics study on transesterification of palm oil. *Journal of Oil Palm Research*, *16*(2), 19–29.
- Clickgasoil. (2021). *Evolución del precio de Biodiesel*. <https://www.clickgasoil.com/c/evolucion-del-precio-biodiesel>.
- El-Gharbawy, A. (2017). Cost Analysis for Biodiesel Production from Waste Cooking Oil Plant in Egypt. *INTERNATIONAL JOURNAL of SMART GRID*, *1*(1), 16–25.
- EP PETROECUADOR. (2022). *Precios de venta en terminal para las comercializadoras calificadas y autorizadas a nivel nacional*. <https://www.eppetroecuador.ec/wp-content/uploads/downloads/2022/05/estructura-de-precios-mayo-2022.pdf>
- Leevijit, T., Wisutmethangoon, W., Prateepchaikul, G., Tongurai, C., & Allen, M. (2004). A Second Order Kinetics of Palm Oil Transesterification. *The Joint International Conference on "Sustainable Energy and Environment (SEE)"*, *3*, 277–281.
- Mohammadshirazi, A., Akram, A., Rafiee, S., & Bagheri Kalhor, E. (2014). Energy and cost analyses of biodiesel production from waste cooking oil. *Renewable and Sustainable Energy Reviews*, *33*, 44–49.
- Narváez, P. C., Noriega, M. A., & Cadavid, J. G. (2015). Kinetics of palm oil ethanolysis. *Energy*, *83*, 337–342.
- Narváez, P. C., Rincón, S. M., & Sánchez, F. J. (2007). Kinetics of Palm Oil Methanolysis. *Journal of the American Oil Chemists' Society*, *84*(10), 971–977.
- Pugazhendhi, A., Alagumalai, A., Mathimani, T., & Atabani, A. E. (2020). Optimization, kinetic and thermodynamic studies on sustainable biodiesel production from waste cooking oil: An Indian perspective. *Fuel*, *273*, 117725.
- Rao, R. V., & Lakshmi, R. J. (2021). Ranking of Pareto-optimal solutions and selecting the best solution in multi- and many-objective optimization problems using R-method. *Soft Computing Letters*, *3*, 100015.
- Roca Meza, L. G. (2014). *Valoración Financiera de la empresa La Fabril S.A.* Escuela Superior Politécnica del Litoral.
- Silva, F. C. da, Guardiola, J. F. H., Teixeira, L. P., Maria, A. C. L., Souza, L. A. de, & Belém, A. L. (2021). Optimization of palm oil biodiesel production using response surface methodology. *Revista Brasileira de Ciências Ambientais (Online)*, *56*(2), 274–285.
- Suzihaque, M. U. H., Alwi, H., Kalthum Ibrahim, U., Abdullah, S., & Haron, N. (2022). Biodiesel production from waste cooking oil: A brief review. *Materials Today: Proceedings*.
- Turton, R., Shaeiwitz, J. A., Bhattacharyya, D., & Whiting, W. B. (2018). *Analysis, Synthesis, and Design of Chemical Processes* (5th ed.). Prentice Hall.
- Wongjaikham, W., Wongsawaeng, D., Ratnitsai, V., Kamjam, M., Ngaosuwan, K., Kiatkittipong, W., Hosemann, P., & Assabumrungrat, S. (2021). Low-cost alternative biodiesel production apparatus based on household food blender for continuous biodiesel production for small communities. *Scientific Reports*, *11*(1), 13827.

# Modeling of an Industrial Delayed Coker Unit

Anne Firstauthor,<sup>a</sup> Tim B. Secondauthor,<sup>b</sup> James Q. Thirdauthor,<sup>a, b</sup>

<sup>a</sup>*First affiliation, Address, City and Postcode, Country*

<sup>b</sup>*Second affiliation, Address, City and Postcode, Country*

Gizem Kuşoğlu,<sup>a, b</sup> Yaman Arkun,<sup>b</sup>

<sup>a</sup> *Turkish Petroelum Refinery, Körfez, Kocaeli, 41790, Turkey*

<sup>b</sup> *Chemical and Biological Engineering, Koç University, Rumelifeneri, İstanbul, 34450, Turkey*

*gkusoglu16@ku.edu.tr, gizem.kusoglu@tupras.com.tr*

## Abstract

Delayed Coker Unit (DCU) converts the vacuum residual feedstock to lighter and more valuable products such as motor fuels and eliminates the low-order and environment-damaging streams. Thus, optimal operation of this unit provides great economic return. In this direction, we have modeled an industrial DCU which exists in the TUPRAS Refinery. The steady-state model consists of the furnace and coke drums and implemented using MATLAB. Physical properties are determined by Aspen HYSYS. The obtained model was used for predicting the coke level in the coke drums where the reaction takes place and the distribution of the products. Both furnace and coke drum models were verified by comparing obtained results with actual plant data.

**Keywords:** Delayed coking, furnace and coke drum models.

## 1. Introduction

Delayed Coker Unit (DCU) is a critical refinery unit that enables the conversion of heavy materials such as vacuum residue into valuable products. This conversion occurs in coke drums where thermal cracking reactions take place. In the literature, there are a limited number of studies about modeling of the unit. Among these studies, Lima et al. (2016) and Borges et al. (2015) have modeled coke drums as plug flow reactor (PFR) by preparing in-house codes or by using simulation programs such as Aspen HYSYS/PLUS. On the other hand, Bozzano and Dente (2005) and Zhou et al. (2007) focused on the estimation of product yields with kinetic models as well. With this study, separate models of the coke drums and the furnace were created and the content of the coke drum outlet stream was estimated close to the reality.

## 2. Process Description

The main purpose of Delayed Coker Unit is thermal cracking the Vacuum Residue (VR) feed into valuable products such as Coker Fuel Gas, LPG, Stabilized Naphtha, Light Coker Gas Oil (LCGO), Heavy Coker Gas Oil (HCGO), and Coke. A simplified diagram of the delayed coking section is shown in Fig. 1. It consists of a fractionator tower, four

coke drums, and coke heaters. Fresh vacuum residue is preheated and enters the bottom of the main fractionator where it is combined with HCGO recycle. The combined fresh feed and recycle is pumped through the heaters in order to increase the temperature to a point at which coke formation occurs in the coke drums. As can be seen in Fig. 1, coking occurs only at one of the coke drum pairs, while the other drum is being cleaned and prepared for the next coking cycle. The mixture leaving the heaters passes through the coke drums where cracking reactions take place and solid coke accumulates. Product vapors obtained by cracking are returned to the bottom of the main fractionator above the fresh feed entry tray.

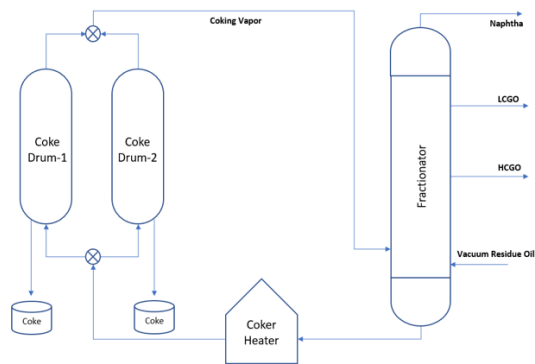


Figure 1. General representation of the Delayed Coker Unit (DCU)

### 3. Delayed Coker Furnace Model

The main role of the furnace is to increase the temperature of the VR feed before sending it to coke drums. VR enters into the furnace as a subcooled liquid and is transformed into a vapor-liquid mixture containing fresh vacuum residue and some products by some cracking reactions in the furnace tubes. While the two-phase stream flows through the furnace tubes, significant pressure drop can develop. In this study, the rigorous furnace model is generated based on the assumptions that there is two-phase flow with endothermic reaction under pressure drop effect, and there is no coke formation. For modeling of the coker furnace, the model structure proposed by Skogestad et al. (1986) has been used with Vapor-Liquid Equilibrium (VLE) for phase split calculations. It is basically based on the mass, energy and momentum conservation equations.

The mass balance is obtained for plug flow reactor (PFR) based on the kinetic model of 5 discrete lumps proposed by Singh et al. (2005). For each discrete lump, kinetic parameters are estimated based on feed and product properties by using the correlations proposed by Castellanos et al. (1991). The reactions only take place in liquid phase. Therefore, the reaction rate expressions have been derived by assuming first-order power-law approach for the PFR as following.

$$\frac{dF_i}{dz} = -r_i A_{liq} \quad (1)$$

where  $i = \text{VR, HCGO, LCGO, Naphtha, Gas lumps}$ ;  $A_{liq}$ , liquid fraction in the tube ( $m^2$ );  $r_i$ , reaction rate of  $i^{\text{th}}$  lump ( $\text{kmol}/\text{sm}^3$ );  $F_i$ , molar flowrate ( $\text{kmol}/\text{s}$ ).

From the perspective of energy, it is assumed that temperature inside the tubes is only changing with convective heat transfer for getting rid of complex equations. Effects of the conduction and radiation are neglected. Resulting enthalpy balance is given in Eq. (2).

$$H(z) = H(0) + \int_0^z Q(z) \cdot \pi \cdot D \cdot dz \quad (2)$$

where  $Q$  is the heat duty ( $\text{kW/m}^2$ );  $H(0)$  enthalpy of the feed at the inlet of the furnace ( $\text{kW}$ ). Pressure drop values along the tubes are being calculated based on three different terms such as frictional, acceleration and gravity gradients. Frictional pressure gradient is calculated with Martinelli correlation with the assumption that two phases can have different properties (Thome and Cioncolini, 2015). Gradient is basically defined as functions of liquid-only or gas-only pressure multiplier and liquid-only or gas-only pressure gradients. Two-phase flow frictional pressure gradient is calculated as in Eq. (3).

$$\left(\frac{dP}{dz}\right)_{fr} = \left(\frac{dP}{dz}\right)_{lo} \phi_{lo}^2 = \left(\frac{dP}{dz}\right)_{go} \phi_{go}^2 \quad (3)$$

where  $\left(\frac{dP}{dz}\right)_{lo}$ , single phase frictional pressure gradients for the liquid-only;  $\left(\frac{dP}{dz}\right)_{go}$ , single phase frictional pressure gradients for the gas-only;  $\phi_{lo}$ , two-phase liquid-only multiplier;  $\phi_{go}$ , two-phase gas-only multiplier.

Acceleration pressure drop gradient is explained on the basis of momentum density  $\rho_m$  ( $\text{kg/m}^3$ ) gradient and mass flux  $G$  ( $\text{kg/m}^2\text{s}$ ). Momentum density term is a function of liquid density  $\rho_L$  ( $\text{kg/m}^3$ ), gas density  $\rho_G$  ( $\text{kg/m}^3$ ), quality ( $x$ ), void fraction ( $\varepsilon$ ). Acceleration pressure drop gradient is computed as follows:

$$\left(\frac{dP}{dz}\right)_{acc} = G^2 \frac{d}{dz} \left(\frac{1}{\rho_m}\right) = G^2 \frac{d}{dz} \left[ \frac{(1-x)^2}{\rho_L(1-\varepsilon)} + \frac{x^2}{\rho_G\varepsilon} \right] \quad (4)$$

where void fraction is calculated based on Lockhart and Martinelli correlation. The last term left for gradient pressure calculation is gravitational pressure drop which is expressed in terms of density and void fraction, such that (Thome and Cioncolini, 2015):

$$\frac{\Delta P_g}{\Delta z} = [\rho_L(1-\varepsilon) + \rho_G\varepsilon] g \sin\theta \quad (5)$$

where  $\frac{\Delta P_g}{\Delta z}$ , gravitational pressure drop;  $\theta$ , the angle of channel inclination from horizontal;  $g$ , acceleration gravity ( $9,81 \text{ m/s}^2$ ).

#### 4. Coker Drum Model

The VR feed enters into the coker drums as two phase composed of vapor and liquid and it exposes to the cracking reactions after a while. Cracking reactions result in vapor products and solid coke which means that there are three phases as coke, residue and distillate in the reactor. From this point of view, it can be modeled similar to the trickle bed reactor. Liquid vacuum residue phase is the primary phase and the gas distillate phase is secondary phase. The system can be modeled in the Eulerian-Eulerian approach. The coke phase is assumed as homogeneous porous medium and accumulates with time.

#### 4.1.1. Material Balance

Vacuum residue is decomposed into gas, distillate products (HCGO, LCGO, Naphtha) and coke. Coke amount changes with time but does not have any positional dependency. Therefore, vacuum residue is assumed to crack only to gas and distillate for steady-state model and coke level is associated with void fraction. Reaction kinetic expression and parameters are calculated based on the correlation proposed by Esfahani et al. (2019). General balance and component-based mass balances are given in Eq. (6a-c).

$$\frac{d(\gamma\alpha_R\rho_R v_{z,R})}{dz} = \gamma\alpha_R\rho_R\Gamma_R \quad (6-a)$$

$$\frac{d(\gamma\alpha_D\rho_D v_{z,D})}{dz} = \gamma\alpha_D\rho_D\Gamma_D \quad (6-b)$$

$$\frac{d(\gamma\alpha_R\rho_R v_{z,R} Y_{i,R})}{dz} = \gamma\alpha_R m_{i,R} \quad (6-c)$$

where  $i$  represents Gas, Distillate, Vacuum Residue lumps;  $\gamma$ , coke void fraction;  $\alpha_R$ , residue volume fraction;  $\alpha_D$ , distillate volume fraction;  $v_{z,R}$ , velocity of the residue phase (m/s);  $v_{z,D}$ , velocity of distillate phase(m/s);  $\Gamma$ , reaction velocity(1/s);  $m_{i,R}$ , mass transfer of component  $i$  from another phase (kg/m<sup>3</sup>s);  $Y_{i,R}$ , weight fraction of  $i^{\text{th}}$  lump.

#### 4.1.2. Momentum Balance

Even though the assumption of no mass transfer between phases, the phase interactions encountered in the trickle beds are involved. The momentum balance of each phase is provided in Eq. (7a-b).

$$\frac{d(\gamma\alpha_R\rho_R v_{z,R}^2)}{dz} = \gamma F_G - \gamma\alpha_R \frac{dP}{dz} + \gamma\alpha_R\rho_R g \quad (7-a)$$

$$\frac{d(\gamma\alpha_D\rho_D v_{z,D}^2)}{dz} = \gamma F_L - \gamma\alpha_D \frac{dP}{dz} + \gamma\alpha_D\rho_D g \quad (7-b)$$

where  $F_G, F_L$ , phase interaction terms (kg/m<sup>2</sup>s<sup>2</sup>);  $g$ , acceleration gravity (9,81 m/s<sup>2</sup>);  $\frac{dP}{dz}$ , pressure drop (kg/m<sup>2</sup>s<sup>2</sup>);  $\gamma$ , coke void fraction. Phase interactions are calculated by the equations provided in the Alopaus et al. (2006)'s study.

#### 4.1.3. Energy Balance

Assuming that the coke, vacuum residue and product vapors are in close contact, the temperatures of the three phases are assumed to be the same throughout the reactor. Therefore, the energy balance is given as following.

$$\frac{d(\gamma\alpha_R\rho_R v_{z,R} C_{P,R} T + \gamma\alpha_D\rho_D v_{z,D} C_{P,D} T)}{dz} = \gamma\alpha_R\rho_R\Gamma_R\Delta\dot{H}_R + U \frac{A}{V_r} (T_C - T) \quad (8)$$

where  $C_{P,R}, C_{P,D}$  are heat capacity of residue and distillate phases, respectively (J/kgK);  $\Delta\dot{H}_R$ , heat of reaction (J/kg);  $A$ , surface area(m<sup>2</sup>);  $V_r$ , volume of the reactor(m<sup>3</sup>);  $U$ , overall heat transfer coefficient(W/Km<sup>2</sup>);  $T_C$ , temperature of surrounding(K);  $T$ , temperature of the fluid(K).

## 5. Results

The models are run for vacuum residue inlet stream of three different days that differed from each other in terms of density (API) and concarbon (CCR). API and CCR values are given in Table 1. The CCR content of the VR affects the yield of the products and so the accuracy of the model predictions was evaluated against the variation of the content.

Table 1. Properties of vacuum residue stream for three different days

Date	API	CCR
1 <sup>st</sup> Day	4,35	23,55
2 <sup>nd</sup> Day	6,50	22,00
3 <sup>rd</sup> Day	4,80	18,38

Models were coded over MATLAB for only one dimension (z-direction). The balance equations of furnace model are solved sequentially until values of variables converge to a point. The coke model is computed by ode15s solver. All required properties (density, phase splits, heat capacity, enthalpy) of each lump are calculated by using Aspen HYSYS. Since the vacuum residue (VR) was first fed to the furnace, the furnace model was run first and then the coke drum model was triggered. The furnace model was run with the VR stream of the 1<sup>st</sup> day and the results in Fig. 2 were obtained. After that, the results of the heater model are used as the input of the coke drum model. The obtained coke drum model results for  $\gamma = 0,7$  are plotted as in Fig. 3.

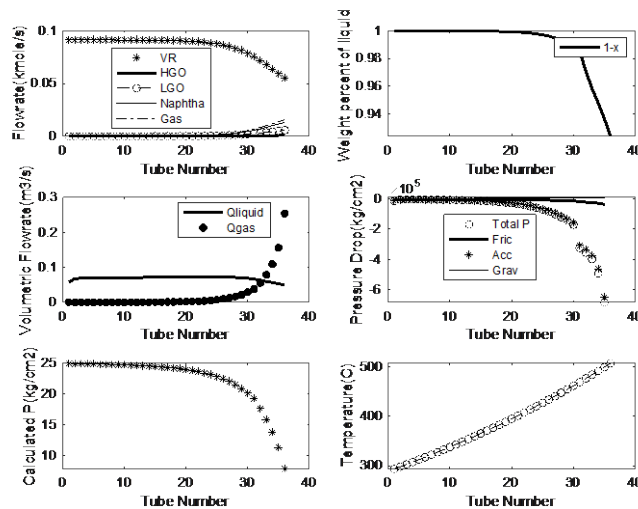


Figure 2. Results of the coker heater model for 1<sup>st</sup> Day

According to Fig. 2 and Fig. 3, it is seen that the estimated values of outlet pressures are quite close to the real outlet values. Real outlet values are represented as Actual T and Actual P in the following graph. The relative error values with respect to temperature, pressure and distillation weight fraction in Table 2 also support this. Despite varying CCR contents of VR, the maximum relative error is 15,77% with steady-state models.

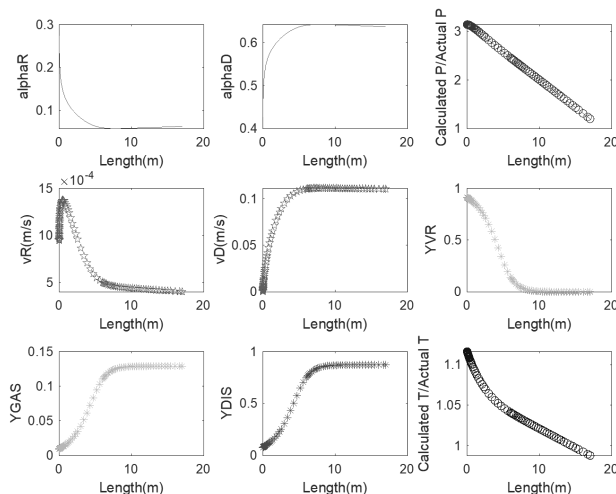


Figure 3. Results of the coke drum model for 1<sup>st</sup> Day (for  $\gamma = 0.7$ )

Table 2. Relative Error with respect to Coke Drum Outlet Conditions

Date	Abs Relative Error (%) Based on Temperature	Abs Relative Error (%) Based on Pressure	Abs Relative Error (%) Based on $Y_{DIS}$
1 <sup>st</sup> Day	0,88	12	6,33
2 <sup>nd</sup> Day	0,91	11	5,79
3 <sup>rd</sup> Day	0,92	15,77	2,79

To obtain these results, correlations proposed in the literature for reaction kinetic parameters were used. On the other hand, only diameter of the coke particle was adjusted to capture the real results in the coke drum model. As shown in Fig. 4, the coke diameter increases with the coalescence of the coke particles while coke level in the drum increases.

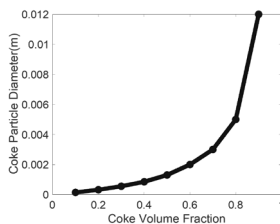


Figure 4. Coke diameter

## 6. Conclusion

Models of both coker heater and coke drums were developed to predict the performance of DCU and verified with the real data sets from Tupras Refinery. The coker heater model is designed as two-phase plug-flow reactor based on 5 discrete lumps (VR, HCGO, LCGO, Naphtha, Gas). The coke drums are assumed as the three-phase trickle-bed reactor with 3 discrete lumps (VR, Distillate, Gas). It has been observed that models running sequentially give realistic results regardless of both density and CCR content of the VR.

**References**

- N.M.N. Lima, P. Sbaite, C. B. Batistella, R. M. Filho, M. R. W. Maciel, F. Manenti, M. Embiruçu, 2016, Simulation, Optimization and Control of Heavy Oil Upgrading Processes: Application to a Delayed Coking Unit, *Computer Aided Chemical Engineering*, Elsevier, 38, 1671-1676.
- C. N. Borges, M. A. Mendes, R. M.B. Alves, 2015, Mathematical Modeling of an Industrial Delayed Coking Unit, *Computer Aided Chemical Engineering*, Elsevier, 37, 515-520.
- G. Bozzano, M. Dente, 2005, A mechanistic approach to delayed coking modelling, *Computer Aided Chemical Engineering*, Elsevier, 20, 2005, 529-534.
- X.L. Zhou, S.Z. Chen, C. L. Li, 2007, A Predictive Kinetic Model for Delayed Coking, *Petroleum Science and Technology*, 25, 1539-1548.
- S. Skogestad, T. Gundersen, O. Johnsen, 1986, Compositional simulation of a refinery coker furnace-An industrial example of two-pase with chemical reaction, *Modeling, Identification and Control*, 7, 25-44.
- J. Singh, M.M. Kumara, A.K. Saxena, S. Kumar, 2005, Reaction pathways and product yields in mild thermal cracking of vacuum residues: A multi-lump kinetic model, *Chemical Engineering Journal*, 108,239-248.
- J. Castellanos, J.L. Cano, R. Del Rosal, V.M. Briones, R.L. Mancilla, 1991, Kinetic model predicts visbreaker yields, *Oil & Gas Journal*, 89, 76-82.
- J. R. Thome and A. Cioncolini, 2015, *Encyclopedia of two-phase heat transfer and flow*, Set 1: Fundamentals and Methods, Volume 3: Flow Boiling in Macro and Microchannels, World Scientific Publishing, Chapter 6: Two-Phase Pressure Drop.
- F. T. Esfahani, M.R. Ehsani, J. Ivakpour, 2019, A Kinetic Model for Delayed Coking Process of Iranian Vacuum Residues, *Petroleum Science and Technology*, 37, 2049-2057.
- V. Alopaeous, K. Hynynen, J. Aittamaa, M. Manninen, 2006, Modeling of Gas-Liquid Packed-Bed Reactors with Momentum Equations and Local Interaction Closures, *Industrial Engineering Chemistry Research*, 2006, 45, 8189-8198.





# **A sustainable energy supply system in smart cities: an integrated natural gas and electricity network operation**

Ali Ahmadian,<sup>a</sup> Ali Almansoori,<sup>b</sup> Ali Elkamel,<sup>c</sup>

<sup>a</sup>*Department of Electrical Engineering, University of Bonab, Bonab, Iran.*

<sup>b</sup>*Department of Chemical Engineering, Khalifa University, Abu Dhabi 127788, United Arab Emirates*

<sup>c</sup>*Department of Chemical Engineering, University of Waterloo, 200 University Avenue West, Waterloo, ON, N2L 3G1, Canada*

## **Abstract**

In this paper, a stochastic approach is presented for optimal planning of an integrated natural gas and electricity network. The planned network can play an important role in the smart cities in order to reduce the emission pollution and operation cost. In addition, various resource and elements of a sustainable green energy system including renewable energy units and energy storage are modeled. A comprehensive investigation is carried out based on different scenarios and the extracted simulation results are discussed. In order to guarantee the global optimum results, the modeling equations of the studied problem have been linearized and the problem became to a mixed-integer linear programming (MILP). The linearized problem optimized using a mathematical based optimization algorithm.

**Keywords:** Sustainable green energy, integrated energy networks, smart cities.

## **1. Introduction**

Today's, the sustainable green energy systems play an important role in the smart cities. Reducing the emission pollution and the operation cost of energy systems is one of the main purpose of the sustainable cities' planner. Several ideas have been proposed in the literature in order to optimal management of energy systems. Optimal energy management in a micro-grid [1], optimal management of electric vehicles charging demand in terms of smart charging [2], optimal energy storage planning and operation in the energy systems [3], etc. are the most famous approaches that are presented in the literature. Although the optimal energy management in the electric network could reduce the operation cost, it can be more beneficial if the electric grid and the natural gas are integrated together [4]. Hence, the integrated energy networks are introduced to optimal operation and planning of both natural gas and electricity networks, simultaneously. In order to study the integrated network, the energy systems' elements including the power plants, gas fired units, gas storage, energy storage, renewable energy resources, etc. should be modeled mathematically. In addition, the power flow in the electricity network and the gas flow in the natural gas network should be modeled and their constraints should be considered.

In this paper, an approach is presented for optimal planning of integrated electricity and gas networks. In addition, the stochastic parameters are modeled based on the Copula method which can model the stochastic parameters more efficiently than other stochastic methods such as Monte Carlo. Furthermore, to decrease the computational burden of the scenario-based method the Kantorovich method is utilized. A comprehensive study by

modeling different resources such as renewable energy resources, and different storages are presented. Based on this method, the system operator can get the most out of these resources. The optimization problem turns into a mixed-integer linear programming (MILP) approach by using appropriate methods which can guarantee the global optimal solution.

<b>Nomenclature</b>	
<b>Indices</b>	
$c$	CHP units index
$e$	Emission type index
$n$	Index of electric and gas storages
$g$	Gas supplier index
$u$	Unit Index
$gs$	Gas storages index
$p, q$	Gas network nodes indices
$s$	Scenarios index
$t$	Time index
<b>Parameters</b>	
$a_u, b_u, c_u$	Unit's cost coefficients
$EC_e$	Externality cost of each emission type (\$/lb)
$EF_{u,e}$	Emission factor of each unit (lb/MWh)
$ES_n^{CH,MAX}, ES_n^{DIS,MAX}$	Maximum rate of charging and discharging in storages
$H_c^A, H_c^B, H_c^C, H_c^D$	Coefficients related to the amount of produced heat by CHP units (MW)
$HD_{s,t}$	Total heat demand (MW)
$P_c^A, P_c^B, P_c^C, P_c^D$	Coefficients related to the amount of produced power by CHP units (MW)
$P_u^{MIN}, P_u^{MAX}$	Minimum and maximum generated active power by units (MW)
$GP_p^{MIN}, GP_p^{MAX}$	Minimum and maximum pressure limit of nodes (Psig)
$Q_u^{MIN}, Q_u^{MAX}$	Minimum and maximum generated reactive power by units (MVar)
$SU_u, SD_u$	Startup and shutdown cost of each unit (\$)
$\eta_n^{CH}, \eta_n^{DIS}$	The efficiency of charging and discharging in storages
$\lambda_g^C$	Gas price (\$/kcf)
$\pi_s$	Probability of each scenario

Variables	
$GF_{p,q,s,t}$	Gas flow in pipelines (kcf)
$G_{g,s,t}^G$	Produced gas by gas suppliers (kcf)
$P_{u,s,t}, Q_{u,s,t}$	Generated active (MW) and reactive (MVar) power by each unit
$ES_{n,s,t}^{CH}, ES_{n,s,t}^{DIS}$	Amount of storages charge and discharge
$GP_{p,s,t}$	Pressure in nodes of the gas network (Psig)
$SOC_{n,s,t}$	Stored power (MWh) or gas (kcf) in storages
$U_{u,s,t}$	Binary variable for on/off of units
$U_{n,s,t}^{CH}, U_{n,s,t}^{DIS}$	Binary variables to control charging and discharging states

## 2. Problem Formulation

An integrated electricity and gas network as shown in Fig. 1 is considered in this paper.

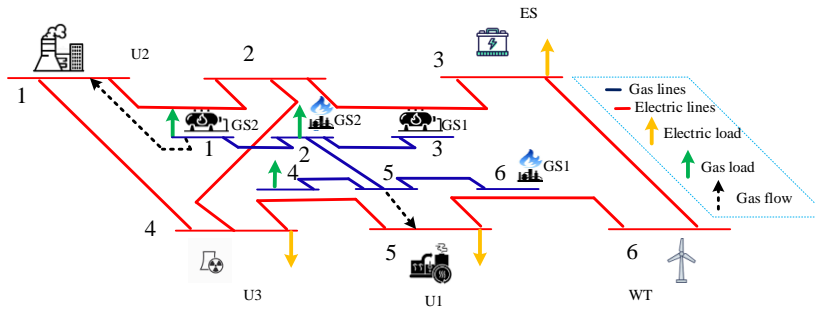


Fig. 1: overall structure of the proposed method.

This study aims to minimize the total cost of both the electricity and gas networks together as expressed in equation (1). This cost consists of three terms: the first term is the emission cost of fossil-fueled based units, the second term explains the operation cost of units to generate electricity, and the third term is related to the cost of natural gas.

$$\min \sum_s \pi_s \sum_t (\underbrace{\sum_e \sum_u EC_e EF_{u,e} P_{u,s,t}}_{\text{Emission Cost}} + \underbrace{\sum_u (a_u P_{u,s,t}^2 + b_u P_{u,s,t} + c_u + S U_u U_{u,s,t} (1 - U_{u,s,t-1}) + SD_u U_{u,s,t-1} (1 - U_{u,s,t}))}_{\text{Operation Cost of Generation units}} + \underbrace{\sum_g \lambda_g^G G_{g,s,t}^G}_{\text{Gas Cost}}) \quad (1)$$

Constraints related to the generation units are described in equations (2) and (3); these constraints limit the total active and reactive power generated by units. The minimum up and down time limits and ramp rate of units are modeled based on reference [5].

$$P_u^{MIN} U_{u,s,t} \leq P_{u,s,t} \leq P_u^{MAX} U_{u,s,t} \quad (2)$$

$$Q_u^{MIN} U_{u,s,t} \leq Q_{u,s,t} \leq Q_u^{MAX} U_{u,s,t} \quad (3)$$

For the CHP units which are responsible for satisfying heat demand besides the above constraints, constraints are investigated to control the amount of generated heat by these units [6]. Moreover, in order to model the electric and gas storages equations are used. Linear power flow equations are used to check the electric network constraints based on reference as presented in [7]. To meet the limitation of the gas network, the presented equations in [8] are utilized.

### 2.1. Equation linearization

In the equations of the system, there are some equations that make the problem a mixed-integer nonlinear programming (MINLP). Solving MINLP problems has a high computational cost. In addition, it is possible to stick to local optimal points. To overcome these issues, the appropriate linearization methods are utilized. For linearization, the multiplication of binary variables, and continuous problems the methods presented in [9] are used. For the linearization of quadratic operation cost of generation units, the piecewise linear method is utilized [10]. Finally, the Weymouth equation is linearized.

## 3. Numerical results and discussion

### 3.1. Input data

To show the effectiveness of the proposed method the electric-gas network which is shown in Fig. 1 is implemented.  $AP^{DR}$  is considered as 10%.  $\eta_n^{CH}$  and  $\eta_n^{DIS}$  are 0.8. Emission data are considered based on [7]. Generation units constraints are presented in references [4] that show the parameters of the gas network.

### 3.2. Uncertainty Modeling

In this paper, the Copula method is investigated for modeling uncertainties. The advantage of copula against other generative methods such as Monte Carlo is its ability to consider the relationship between different features.

### 3.3. Simulation Results

The results after the implementation of the proposed method are shown in figures 2-5 and Table 1. Fig. 2 shows the generated power by three units. The first unit which is responsible for heat demand too generated its maximum power during the day. However, unit 3 is not operated at all hours because the emission cost is implemented for units in this study; in this regard, unit 3 is not economical. Unit 2 produces its maximum power in the middle of the day when we have the maximum power and gas demands.

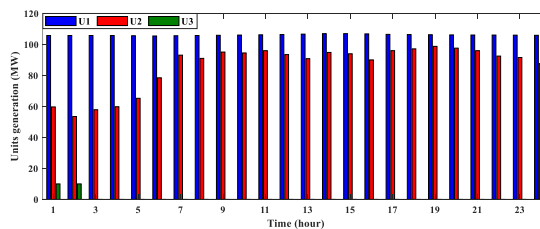


Fig. 2: Generated power by different units.

Fig. 3 demonstrates the total gas demand in the network which is the sum of gas demand and consumed gas by units. This gas demand is supplied by two gas suppliers as shown in Fig. 1, the amount of generated gas by two units are presented in Fig. 4.

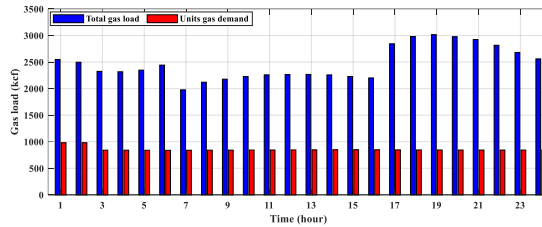


Fig. 3: Total gas demand.

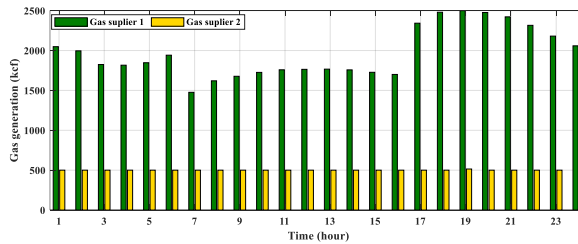


Fig. 4: The amount of generated gas by gas suppliers.

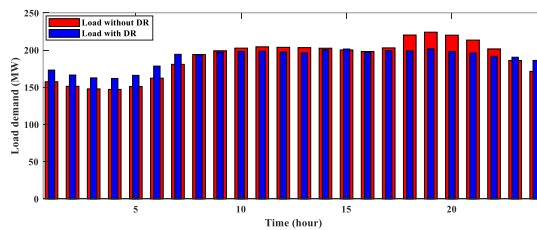


Fig. 5: Electric demand before and after implementation of DRP.

To better clarify the results the detail of the results is presented in Table 1 for each source.

Table 1: Different costs of integrated electricity and gas networks.

Parameters	Value of Cost (\$)
Unit 1	20483
Unit 2	21519
Unit 3	286
Gas cost	154106
Emission cost	4821
<b>Total cost</b>	<b>201215</b>

#### 4. Conclusion

In this paper, an integrated natural gas and electricity network was investigated. The elements and resources of the both networks were modeled mathematically and the power flow in the electricity network and the gas flow in the natural gas network were modeled. A mixed integer linear programming (MILP) approach is presented to optimal planning and operation of the studied integrated network. The simulation results verified the effectiveness and robustness of the proposed methodology in reducing of the emission pollution and operation cost.

#### References

- [1] Thirunavukkarasu, Gokul Sidarth, Mehdi Seyedmahmoudian, Elmira Jamei, Ben Horan, Saad Mekhilef, and Alex Stojcevski. "Role of optimization techniques in microgrid energy management systems—A review." *Energy Strategy Reviews* 43 (2022): 100899.
- [2] Sadeghian, Omid, Arman Oshnoei, Behnam Mohammadi-Ivatloo, Vahid Vahidinasab, and Amjad Anvari-Moghaddam. "A comprehensive review on electric vehicles smart charging: Solutions, strategies, technologies, and challenges." *Journal of Energy Storage* 54 (2022): 105241.
- [3] Sedghi, Mahdi, Ali Ahmadian, and Masoud Aliakbar-Golkar. "Optimal storage planning in active distribution network considering uncertainty of wind power distributed generation." *IEEE Transactions on Power Systems* 31, no. 1 (2015): 304-316.
- [4] AlHajri I, Ahmadian A, Elkamel A. Techno-economic-environmental assessment of an integrated electricity and gas network in the presence of electric and hydrogen vehicles: A mixed-integer linear programming approach. *J Clean Prod* 2021:128578.
- [5] Sadeghi S, Jahangir H, Vatandoust B, Golkar MA, Ahmadian A, Elkamel A. Optimal bidding strategy of a virtual power plant in day-ahead energy and frequency regulation markets: A deep learning-based approach. *Int J Electr Power Energy Syst* 2021;127:106646. <https://doi.org/10.1016/j.ijepes.2020.106646>.
- [6] Gougheri SS, Dehghani M, Nikoofard A, Jahangir H, Golkar MA. Economic assessment of multi-operator virtual power plants in electricity market: A game theory-based approach. *Sustain Energy Technol Assessments* 2022;53:102733.
- [7] Gougheri SS, Jahangir H, Golkar MA, Ahmadian A, Aliakbar Golkar M. Optimal participation of a virtual power plant in electricity market considering renewable energy: A deep learning-based approach. *Sustain Energy, Grids Networks* 2021;26:100448. <https://doi.org/10.1016/J.SEGAN.2021.100448>.
- [8] Jiang Y, Xu J, Sun Y, Wei C, Wang J, Liao S, et al. Coordinated operation of gas-electricity integrated distribution system with multi-CCHP and distributed renewable energy sources. *Appl Energy* 2018;211:237–48.
- [9] Gougheri SS, Jahangir H, Golkar MA, Moshari A. Unit Commitment with Price Demand Response based on Game Theory Approach. 2019 *Int. Power Syst. Conf., IEEE*; 2019, p. 234–40.
- [10] Luo S, Yang L, Zhang X, Chen W, Wang K, Xu Y. A fully linear-constrained optimal electricity-gas flow in an integrated energy system. 2018 *2nd IEEE Conf. Energy Internet Energy Syst. Integr., IEEE*; 2018, p. 1–6.

# A hybrid forecasting framework for the energy sector

Stefanos G. Baratsas,<sup>a,b</sup> Funda Iseri,<sup>a,b</sup> Efstratios N. Pistikopoulos<sup>a,b</sup>

<sup>a</sup>*Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77843, United States*

<sup>b</sup>*Texas A&M Energy Institute, Texas A&M University, College Station, TX 77843, United States*

## Abstract

Energy prices are nonlinear, noisy, non-stationary time-series, which are sensitive and volatile to supply and demand imbalances, changes in monetary and fiscal policies, new environmental regulations, major political and social events, as well as technological breakthroughs. Despite the fact that numerous methods have been proposed in the literature to forecast the prices of the several types and sources of energy, there is no single forecasting method that systematically performs better over the other methods for any given data set. To this respect, Energy Price Index (EPIC) extended forecasting capabilities are presented here. Statistical and machine learning forecasting methods of different nature have been incorporated into the framework, enabling the forecasting of energy prices up to 14 months in the future. The historical prices of each energy product are studied independently, so as to identify seasonality, patterns, trends and outliers. The results reveal the superiority of the deep neural networks, especially those with feedback loops, over statistical forecasting methods.

**Keywords:** forecasting framework, energy prices, machine learning, hyperparameters tuning.

## 1. Introduction

A time-series is a sequence of observations made over a period of time at equal time intervals. An inherent characteristic of time-series is that adjacent observations are often dependent. Determining the probability law that governs a time-series is a crucial task towards unraveling the underlying dynamics, and eventually forecasting future events.

Traditional methods such as autoregressive integrated moving average (ARIMA) (Box et al., 2015) and exponential smoothing (Gardner et al., 1985; Winters, 1960), focus mainly on parametric models which are approximations of the real stochastic dynamics that govern a data set. The increase in data availability and computing power over the last decades led to a new class of methods that are capable of identifying nonlinear patterns, learning by trial-and-error complex data structures, and improving their performance over time, without requiring the underlying knowledge or an analytical model (Montgomery et al., 2015). As such, various machine learning (ML) methodologies have been proposed for a variety of applications (Lim et al., 2021). However, the prevalence of the more complicated ML methodologies in fitting and learning the temporal dynamics of the available data, does not necessarily lead to more accurate forecasts (Makridakis et. al.,



2018). To this respect, hybrid models that combine both statistical and ML methodologies have illustrated superior performance against their individual components in a range of applications (Smyl, 2020).

In this work, we extend the forecasting capabilities of the Energy Price Index (EPIC) (Baratsas et al., 2021, 2022) with the incorporation of 35 statistical and machine learning forecasting methods of different natures, enabling the forecasting of the energy prices up to 14 months in the future. Since EPIC consists of 56 different energy products with 33 unique time-series data, the extended forecasting framework is applied to each one of these univariate time-series individually and the forecasting model that demonstrates the most accurate forecasting results over a testing period is then utilized for the future forecasts.

## 2. Hybrid Forecasting Framework for the Energy Sector

The large number of unique time-series that demonstrate unique patterns, trends, and cycles and need to be predicted for different forecasting horizons, introduce major challenges in the forecasting process. Figure 1 highlights these unique characteristics on a subset of the time-series.

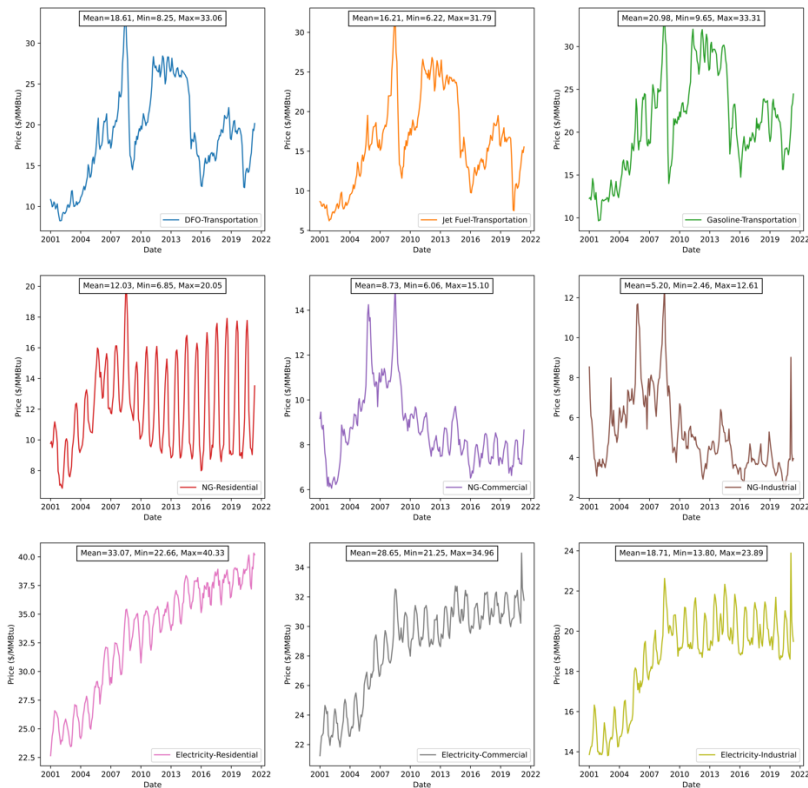


Figure 1: 20 years of monthly prices of the 9 energy products with the highest energy demand in 2017-2021

The 33 time-series are monthly indexed with data sets of 161 to 425 months. The forecasting framework incorporates 35 statistical and ML forecasting methods into 7 distinct groups, which are summarized in Table 1. The first forecasting group "Benchmark Forecasting" consists of five rather simple but quite effective methods (Hyndman, 2018). They are used to provide a benchmark on forecasting performance against the rest of the forecasting groups which are more advanced and computationally expensive. The second forecasting group "Forecasting with Decomposition" utilizes STL decomposition (Gardner et al., 1985) and then the forecasts of the STL objects are obtained by applying a non-seasonal forecasting method to the seasonally adjusted data and re-seasonalizing using the last year of the seasonal component. The third forecasting group "Exponential Smoothing" (Winters, 1960; Bergmeir et. al., 2016) consists of nine forecasting methods, where forecasts are weighted averages of past observations with the weights though decaying exponentially for the older observations. A complementary approach to exponential smoothing is the fourth forecasting group, namely "ARIMA", which combines autoregressive and moving average models, while allowing differencing of the data series. Seasonal and non-seasonal ARIMA models ( $ARIMA_{(p,d,q)(P,D,Q)m}$ ), are used. Due to the large number of hyperparameters to be determined, grid search is conducted in the last two methods (fc4d, fc4e) and the configuration with the smallest AICc and AIC value (Cleveland et. al, 1990) respectively is selected. The fifth forecasting group consists of advanced statistical forecasting methods, including dynamic regression and TBATS models (Cleveland et. al, 1990; Crone et. al., 2010), along with bootstrapped time series that use the Box-Cox and Loess-based decomposition bootstrap (De Livera et. al, 2011).

Table 1: Hybrid forecasting framework for energy prices

Group	Forecasting Method	Description of Forecasts
Benchmark Forecasting	fc1a	Average based on the historical mean
	fc1b	Random Walk (RW) – Naïve naïve w/ RW model
	fc1c	Seasonal Naïve ARIMA(0,0,0)(0,1,0) model
	fc1d	RW w/ Drift RW w/ drift
	fc1e	RW w/ Drift Transformed & Bias Adjusted fc1d adj Box-Cox transformation
Forecasting with Decomposition	fc2a	Automated STL automated STL decomposition
	fc2b	ARIMA Automated STL ARIMA seasonally adjusted data
	fc2c	Naïve Automated STL naïve seasonally adjusted data
	fc2d	RW w/ Drift Automated STL RW w/ drift seasonally adjusted data
	fc2e	ETS Automated STL STL decomposition, w/ ETS seasonally adjusted back Box-Cox transformed data
Exponential Smoothing	fc3a	Simple Exponential Smoothing
	fc3b	Holt’s Damped trend
	fc3c	Holt-Winters’ (HW) Additive
	fc3d	HW Multiplicative
	fc3e	HW Damped Additive
	fc3f	HW Damped Multiplicative
	fc3g	Error, Trend, Seasonal (ETS) exponential smoothing state space

	fc3h	ETS Transformed Bias Adjusted	fc3g w/ adj Box-Cox transform
	fc3i	Bagged ETS	ETS to all bootstrapped series
<b>ARIMA</b>	fc4a	Manual ARIMA	manually selected terms
	fc4b	Auto. Arima Non-Seasonal	automated non-seasonal ARIMA
	fc4c	Auto. Arima Seasonal	automated seasonal ARIMA
	fc4d	Grid Search Arima	grid searching of hyperparameters
	fc4e	Grid Search Arima	grid searching of hyperparameters
<b>Advanced Forecasting Methods</b>	fc5a	Dynamic Regression model	w/ Fourier terms
	fc5b	TBATS	
	fc5c	Bootstrapping ARIMA	Box-Cox & Loess-based decomposition bootstrap
<b>Neural Networks (NN)</b>	fc6a	Feed-forward NN	single hidden layer (hl), lagged inputs
	fc6b	MLP_1	MLP w/ 5 hl & 20 training reps
	fc6c	MLP_2	MLP w/ 5 hl & 50 training reps
	fc6d	MLP_3	fc6c w/ 12 autoregressive lags
<b>Grid Search NN</b>	fc7	MLP	MLP model w/ 2 hl, followed by corresponding dropout layers, and a final output layer
	fc8	RNN	RNN model w/ 2 hl layers and a final output layer
	fc9	LSTM	LSTM model w/ a single LSTM hl, followed by a dropout layer, a dense fully connected layer followed by a dropout layer, and a final output layer
	fc10	CNN-LSTM-DNN	Hybrid CNN-LSTM model w/ 2 convolutional layers for 1D inputs, followed by a pooling & a flattened layer, 2 LSTM hl, and a final output layer

The last two forecasting groups consist of ML methods. In the sixth forecasting group "NN", the hyperparameters of the models are pre-selected, while in the seventh forecasting group "Grid Search NN", a grid search is applied for tuning the hyperparameters. More details about the grid searches are given in Table 2. In the last forecasting group, each configuration is evaluated 3 times. The average of these values for each accuracy measure is considered as the final one.

Table 2: Hyperparameters tuning through grid search

#	Details	Hyperparameters
fc4d	Grid Search Arima	p=[1...5], d=1, q=[1...3], P=[1...5], D=[0...1], Q=[1...3], m=18
fc4e	Grid Search Arima	
fc7	Dense class of Keras API v2.4.3 for Python 3.9 with TensorFlow v2.5.0	inp=[h, 2h], nod=[32, 64], Dropout rate = [0.1], Learning rate = [1e-4], epoch=[250], batch=[8, 16], Differences = [0, 1, h], Standardization = [T, F]
fc8	RNN and Dense class of Keras API v2.4.3 for Python 3.9 with TensorFlow v2.5.0	
fc9	LSTM and Dense class of Keras API v2.4.3	

	for Python 3.9 with TensorFlow v2.5.0	
<b>fc10</b>	Dense class of Keras API v2.4.3	nodes=[64, 128], Learning rate = [1e-4, 1e-6], epochs=[500, 1000], batches=[64], seq=[3]
	for Python 3.9 with TensorFlow v2.5.0	Steps: # of timesteps within each subsequence, steps=[12], filters=[128, 256], ker=[3, 6]

### 3. Data Preprocessing and Evaluation Metrics

Each time-series is preprocessed to achieve stationarity in its mean and variance, and is decomposed so as to extract time-series patterns such as trend, seasonality and cycles. This is done using Seasonal and Trend decomposition using Loess (STL), Box-Cox transformation, first order, and seasonal differencing so as to remove trend and seasonality. The residuals between the historical data and the corresponding fitted values, should be uncorrelated and have zero mean. The data sets are split into 80%-20% sets, with the first 80% of the data used for training/validating the models, and the last 20% used for testing their forecasting accuracy. Three accuracy measures are used in this paper: Root Mean Squared Error (RMSE), symmetric Mean Absolute Percentage Error (sMAPE) and Mean Absolute Error (MAE), and are defined as follows:

$$RMSE = \sqrt{\frac{1}{h} \sum_{t=1}^h (y_t - \hat{y}_t)^2} \quad sMAPE = \frac{2}{h} \sum_{t=1}^h \frac{|y_t - \hat{y}_t|}{|y_t| + |\hat{y}_t|} * 100\%$$

$$MAE = \frac{1}{h} \sum_{t=1}^h |y_t - \hat{y}_t|$$

where  $h$  is the forecasting horizon,  $y_t$  are the historical data, and  $\hat{y}_t$  are the forecasts produced by the model at point  $i$ .

### 4. Results

The forecasting framework is applied to each one of the 33 time-series. In each case, the forecasting model with the highest forecasting accuracy based on the three accuracy measures is selected for the future forecasts. If different models result from the accuracy measures, then the following process is followed: i) if the same model results from two accuracy measures, this one is selected, ii) if a different model results from each accuracy measure, then the model that has the highest overall ranking among all three accuracy measures is selected e.g. fc3a: 1<sup>st</sup> RMSE, 2<sup>nd</sup> sMAPE, 3<sup>rd</sup> MAE. In case there is still no clear winner, the model with the lowest RMSE is selected. Table 3 summarizes the forecasting results for the 33 time-series.

Table 3: Best forecasting models and accuracy measures

Energy Product	Selected Method	Average RMSE (\$/MMBtu)	Average sMAPE (%)	Average MAE (\$/MMBtu)
Distillate Fuel Oil (DFO) [R]	fc7	1.436	4.760	1.002
Kerosene [R, C, I]	fc9	3.403	8.982	1.962
Hydrocarbon Gas Liquids (HGL) [R]	fc3	1.430	5.003	1.111
DFO [C, I]	fc7	1.025	4.262	0.775
HGL [C, I, T]	fc7	1.014	4.241	0.767
Gasoline [C, I]	fc8	1.050	4.130	0.824
Petroleum Coke [C, I]	fc9	0.232	9.792	0.194
Residual Fuel Oil (RFO) [C, I, T]	fc8	0.708	5.856	0.545

Asphalt and Road Oil [I]	fc9	1.069	7.492	0.834
Lubricants [I, T]	fc9	1.044	7.443	0.829
Other Petroleum Products [I]	fc7	1.085	6.601	0.817
Aviation Gasoline [T]	fc9	1.668	4.982	1.229
DFO [T]	fc8	0.836	3.857	0.663
Jet Fuel [T]	fc9	1.071	5.765	0.771
Motor Gasoline [T]	fc9	1.019	3.950	0.813
Geothermal Energy [R, C, I]	fc9	0.527	0.510	0.125
Solar [R]	fc9	0.484	0.280	0.157
Biomass Energy [R, C, I]	fc9	0.176	0.299	0.082
Hydroelectric Power [C, I]	fc7	0.468	1.601	0.206
Solar [C, I]	fc7	1.035	0.997	0.349
Wind Energy [C, I]	fc8	0.088	0.365	0.043
Biomass (Fuel-Ethanol), [C, I, T]	fc9	1.094	3.719	0.856
Biomass (Bio-Diesel) [T]	fc9	1.444	5.064	1.165
Natural Gas, [R]	fc8	0.596	4.183	0.467
Natural Gas, [C]	fc8	0.263	2.685	0.204
Natural Gas, [I]	fc8	0.891	8.846	0.393
Natural Gas, [T]	fc9	0.611	4.223	0.401
Electricity [R]	fc10	0.551	1.125	0.430
Electricity [C]	fc4a	0.651	1.115	0.354
Electricity [I]	fc8	0.810	1.915	0.389
Electricity [T]	fc2d	0.664	1.855	0.532
Coal, [C]	fc9	0.059	0.953	0.042
Coal, [I]	fc9	0.037	0.704	0.027

*R: Residential, C: Commercial, I: Industrial, T: Transportation*

The above results clearly demonstrate the superiority of the last forecasting group "Grid Search NN" over the other forecasting groups, since 31 out of the 33 energy products are modeled more accurately using one of its four forecasting methods. In particular, LSTM is the most accurate method since it is used for 15 energy products, followed by RNN, MLP, and CNN-LSTM-DNN which are used for 9, 6, and 1 energy products respectively. Just two models are modeled using forecasting with decomposition and ARIMA groups. It is worth mentioning, that even without considering the last forecasting group, the NNs of the sixth forecasting group would only be used for 5 out of 33 energy products. Thus, NNs require special attention and proper tuning for forecasting time-series.

## 5. Conclusions

The historical prices of each energy product are analyzed independently considering seasonality, patterns, trends, cycles and outliers. Each time-series is pre-processed to establish stationarity in the mean and variance before being split into training and testing datasets. Then, the forecasting framework that consists of 35 statistical and ML forecasting methods is applied to each time-series individually, and the forecasting model that produces the best accurate results throughout a testing period is used for predictions up to 14 months in the future. The results highlight the critical importance of accurately tuning a NN's hyperparameters and demonstrate that deep NNs, especially those with feedback loops, perform better than statistical forecasting techniques.

## 6. Acknowledgments

Financial support from Texas A&M Energy Institute is also gratefully acknowledged.

## References

- Baratsas, S. G., Niziolek, A. M., Onel, O., Matthews, L. R., Floudas, C. A., Hallermann, D. R., et al., 2021. A framework to predict the price of energy for the end-users with applications to monetary and energy policies. *Nature Communications* 12, 1–12.
- Baratsas, S. G., Niziolek, A. M., Onel, O., Matthews, L. R., Floudas, C. A., Hallermann, D. R., et al., 2022. A novel quantitative forecasting framework in energy with applications in designing energy-intelligent tax policies. *Applied Energy* 305, 117790.
- Bergmeir, C., Hyndman, R. J. and Benítez, J. M., 2016. Bagging exponential smoothing methods using STL decomposition and box–cox transformation, *International journal of forecasting*, vol. 32, no. 2, pp. 303–312.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M., 2015. *Time series analysis: forecasting and control*. John Wiley & Sons.
- Cleveland, R. B., Cleveland, W. S., McRae, J. E. and Terpenning, I., 1990. Stl: A seasonal-trend decomposition, *J. Off. Stat.*, vol. 6, no. 1, pp. 3–73.
- Crone, S. F., and Kourentzes, N., 2010. Feature selection for time series prediction—a combined filter and wrapper approach for NN, *Neurocomputing*, vol. 73, no. 10-12, pp. 1923–1936.
- De Livera, A. M., Hyndman, R. J., and Snyder, R. D., 2011. Forecasting time series with complex seasonal patterns using exponential smoothing, *Journal of the American statistical association*, vol. 106, no. 496, pp. 1513–1527.
- Gardner Jr, Everette S., 1985. Exponential smoothing: The state of the art. *J. Forecasting* 4, 1–28
- Géron, A., 2019. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media.
- Goodfellow, I., Bengio, Y., and Courville, A., 2016. *Deep learning*. MIT press.
- Hyndman, R. J., and Athanasopoulos, G., 2018. *Forecasting: principles and practice*. OTexts.
- Lim, B., & Zohren, S., 2021. Time-series forecasting with deep learning: a survey. *Philosophical Transactions of the Royal Society A* 379, 20200209.
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V., 2018. Statistical and machine learning forecasting methods: Concerns and ways forward. *PloS one* 13, e0194889.
- Montgomery, D. C., Jennings, C. L., & Kulahci, M., 2015. *Introduction to time series analysis and forecasting*. John Wiley & Sons.
- Smyl, S., 2020. A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. *International Journal of Forecasting* 36, 75–85.
- Winters, P. R., 1960. Forecasting sales by exponentially weighted moving averages. *Management Science* 6, 324–342



# A Chance-Constrained Programming Approach to Optimal Planning of Low-carbon Transition of a Natural Gas Supply System: A Case Study of China

Jiaqi Zhang <sup>a</sup>, Pei Liu <sup>\*</sup>, Zheng Li

<sup>a</sup> *Tsinghua BP Clean Energy Research and Education Centre, Department of Energy and Power Engineering, Tsinghua University, Beijing 100084, China*  
*liu\_pei@tsinghua.edu.cn*

## Abstract

Low-carbon transition of energy systems is an inevitable trend to address climate change challenges, featuring more natural gas consumption to replace coal and oil. Planning of a natural gas supply system with multiple supply sources, end-consumers, and large infrastructures are challenging tasks. Uncertainty and spatial and seasonal mismatch of natural gas supply and demand makes the natural gas supply infrastructure planning problem more complex. In this study, a chance-constrained programming approach to optimal planning of low-carbon transition of a natural gas supply system is presented. This approach considers the uncertainties of end-consumer demand, pipe aging and import fluctuation to ensure natural gas supply reliability. China is taken as a case study. Results show that with consideration of uncertainty, an additional 40.75 % of LNG ports capacity and an additional 25.59 % of gas storage capacity need to be installed per year on average, with a 38.10 % increase in the transition cost. Finally, sensitivity analysis results are provided, showing that a rapid increase of transition costs occurs when confidence level exceeds 99.9 %. The results in this study provide references for decision makers to plan the low-carbon transition of natural gas supply systems as well as weighing transition costs against natural gas supply stability.

**Keywords:** chance-constrained programming, natural gas supply system, optimization, decarbonization

## 1. Introduction

Decarbonisation is an inevitable trend to address climate change challenges. In recent years, global natural gas consumption has grown rapidly to replace coal and oil, and the share of natural gas in primary energy consumption has reached a high level of 24% (BP, 2022). Natural gas supply highly depends on specialized and high-cost infrastructure, including pipeline networks, liquid natural gas (LNG) ports, and storage facilities. Spatial mismatch between natural gas resources and natural gas demand requires a large amount of infrastructure. China surpassed Japan as the world's largest LNG importer and accounted for close to 60% of global LNG demand growth in 2021 (BP, 2022).

For the low-carbon transition in developing regions, more natural gas is needed at the current stage. Infrastructure costs are a major part of the cost of the gas supply system, and if infrastructure costs are too high, the low-carbon transition will slow, disrupt or even reverse. In addition, natural gas probably be replaced by zero-carbon energy sources in the future. Therefore, developing an appropriate planning framework to obtain the optimal planning by minimizing the overall cost would facilitate the transition and avoiding investment waste.



Modelling and optimization of a natural gas supply system has been studied in many researches. Li et al. (2019) propose a monthly-scale, multi-period and multi-regional modelling and optimization framework for planning and operation of a natural gas supply system at a transient stage. Huang et al. (2022) proposes a method based on Bayesian networks (BNs) to optimize the reliability of gas supply in natural gas pipeline networks. Dieckhoner et al. (2013) and Chaudry et al. (2014) developed a pipeline network model and obtained an expansion plan to address increasing demand of natural gas based on existing infrastructure. An interconnected natural gas and electrical transmission model was developed for the natural gas system in Illinois, considering spatiotemporal changes in energy demand (Chiang, 2016). Zhang et al. (2016) established a multi-regional natural gas distribution model of China, and optimized gas flow and infrastructure deployment by minimizing the total supply cost.

However, previous studies focus less on the impact of uncertainty in natural gas supply system optimization planning. In recent years, international turmoil led to huge fluctuations in natural gas import. There also have uncertainties about pipeline aging and natural gas demand in a natural gas supply system, which need to be considered in planning.

In this study, a chance-constrained programming approach to optimal planning of a natural gas supply system is established. It can consider the natural gas infrastructure condition of a developing region, as well as using chance-constrained programming (Charnes and Cooper, 1959) to consider the uncertainty of import volatility, aging pipelines and natural gas demand. Under the optimization goal of minimum transition cost, the optimal planning of low-carbon transition of a natural gas supply system is given. In addition, this study also quantitatively analyses the cost increase caused by considering the degree of uncertainty, so as to provide reference for policy makers.

The structure of this paper is as follows. Section 2 introduces the methodology, section 3 introduces the case study and its results, and section 4 is the conclusion.

## 2. Methodology

Regional natural gas supply system optimization model is used to quantitatively optimize the natural gas supply system transition path with the consideration of uncertainty. Input parameters include the information of existing natural gas infrastructure, monthly natural gas demand, upper limit of infrastructure construction ability, and infrastructure investment price. Decision variables of model are used to describe the capacity size and construction location of liquefied natural gas (LNG) ports, natural gas storage system and natural gas pipeline as well as how they change from year to year. The structure of the model is shown in Figure 1.

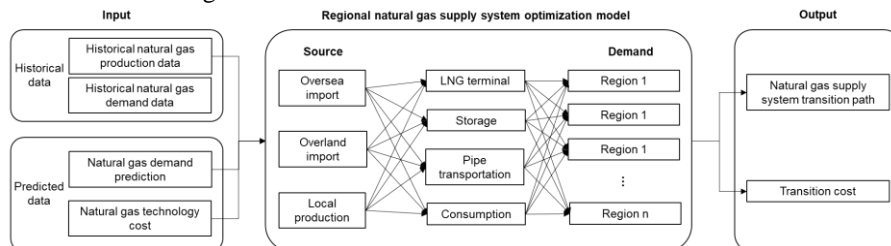


Figure 1. The structure of regional natural gas supply system optimization model

### 2.1. Mathematical formulation

In Regional natural gas supply system optimization model, the core is to meet the supply of natural gas. In each region and in each month, total supply equals total demand. Total

supply comprises of domestic production, net import, net transmission from other regions and net storage reduction, as shown in Eq. (1).

$$qua_{dem,r,t,m} = qua_{pro,r,t,m} + qua_{sto,r,t,m} + qua_{imp,r,t,m} + qua_{tra,r,t,m} \quad (1)$$

Domestic production and net import make up total natural gas supply. Natural gas import is constrained by import capacity as shown in Eq. (2). Natural gas can only be transmitted between two adjacent regions, and the transmission quantity cannot exceed the upper bound constrained by infrastructure, as shown in Eq. (3). A binary variable  $y_{r,rr}$  is introduced to represent geographic constraints, where  $y_{r,rr}$  equals to zero representing two regions are not adjacent. Natural gas storage is constrained with capacity and period, where storage quantity cannot exceed its maximum capacity as shown in Eq. (4), and ending stock of this period equals the beginning of the next period, as shown in Eq. (5).

$$qua_{imp,r,t,m} \leq capa_{imp,r,t}/12 \quad (2)$$

$$qua_{tra,r,rr,t,m} \leq capa_{tra,r,rr,t} \times y_{r,rr}/12 \quad (3)$$

$$qua_{sto0,r,t,m} \leq capa_{sto,r,t} \quad (4)$$

$$qua_{sto0,r,t,m} = qua_{sto1,r,t,m-1} \quad (5)$$

Along with the increase of the natural gas demand, more pipelines, LNG ports and storage facilities would be constructed. Expansions of natural gas storage facilities are continuous design variables, whilst expansions of pipelines and ports are discontinuous variables. The new capacity is an integer multiple of the standard capacity, as shown in Eqs. (6~7). Annual capacities of them are calculated by Eq. (8).

$$inc_{imp,r,t} = int_{imp,r,t} \times std_{imp,r,t} \leq 3 \times 41.4 \quad (6)$$

$$inc_{tra,r,t} = int_{tra,r,t} \times std_{tra,r,t} \leq 5 \times 300 \quad (7)$$

$$capa_{i,r,t} = capa_{i,r,t-1} + inc_{i,r,t} - dec_{i,r,t} \quad (8)$$

A chance-constrained programming method is established to consider the uncertainty of import fluctuation, pipeline aging and natural gas demand. At the time of natural supply, the actual values of import, pipeline capacity and natural gas demand can be seen as the sum of the predicted value and random fluctuation, as shown in Eq. (9).

$$\begin{cases} qua_{imp,t,m} = qua_{imp0,t,m} + \delta_{imp,t,m} \\ capa_{tra,t,m} = capa_{tra0,t,m} + \delta_{tra,t,m} \\ qua_{dem,t,m} = qua_{dem0,t,m} + \delta_{dem,t,m} \end{cases} \quad (9)$$

Considering chance-constraint, Eq. (1) can be expressed as a probabilistic constraint. In Eq. (10),  $\alpha$  is the confidence level, which shows the extent of considering uncertainty.

$$\Pr \{ qua_{pro,r,t,m} + qua_{imp,r,t,m} + qua_{tra,r,t,m} + qua_{sto,r,t,m} - qua_{dem,r,t,m} \geq 0 \} \geq \alpha \quad (10)$$

This study assumes these variables with uncertainty follow a normal distribution, Eq. (10) can be converted into Eq. (11).

$$qua_{pro,r,t,m} + qua_{sto,r,t,m} \geq \mu qua_{imp,r,t,m} + \mu qua_{tra,r,t,m} - \mu qua_{dem,r,t,m} + z_\alpha \times \sqrt{\sigma_{imp,r,t,m}^2 + \sigma_{tra,r,t,m}^2 + \sigma_{dem,r,t,m}^2} \quad (11)$$

The objective function of regional natural gas supply system optimization model describes the total cost of transition path, including the cost of infrastructure investment as well as import, gasification, transportation and storage of natural gas, as shown in Eq. (12).

$$cost_t = cost_{inv,t} + cost_{imp,t} + cost_{ify,t} + cost_{tra,t} + cost_{sto,t} \tag{12}$$

The mixed integer nonlinear programming optimization problem is implemented in General Algebraic Modelling System (GAMS).

### 3. Case Study

China is the largest developing country in the world. It aims to peak carbon emissions by 2030, and strive to achieve carbon neutral goal by 2060. In this study, China is taken as a case. The above approach is used to optimize the transition path of natural gas supply system of China. At the same time, the impact of consideration and non-consideration of uncertainty on planning is compared. Finally, Sensitivity of confidence level is analyzed. This model takes 2020 as base year and the transition path is optimized until 2060.

#### 3.1. Supply structure of natural gas

Figure 2 shows the changes in natural gas supply structure of China. For China, its natural gas supply will mainly rely on imports until 2050, which proportion will exceed 40% by 2050. Natural gas is mainly supplied from Central Asia import and oversea import with other import methods also account for a proportion. Multi-source imports improve energy security.

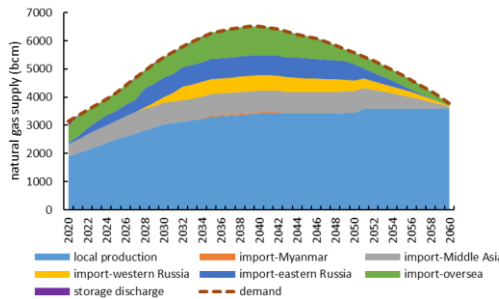


Figure 2 Natural gas supply structure of China

#### 3.2. Comparing results between considering and non-considering uncertainty

The confidence level used in the scenario of considering uncertainty is 99%. In both scenarios of considering and not considering uncertainty, the capacity of LNG ports will peak in 2040. However, after considering uncertainty, an additional 40.75% of LNG ports capacity need to be installed per year on average. Figure 3 shows the distribution of LNG ports in 2040 and 2060 with consideration of uncertainty. LNG ports will mainly be built in Shandong, Zhejiang and Guangdong. The first two will also remain the main provinces importing natural gas in 2060.

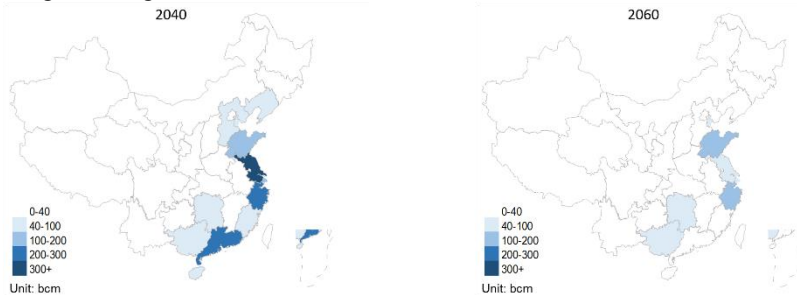


Figure 3 Distribution of LNG ports of China considering uncertainty

For nature gas storage facilities, its capacity will peak in 2040 in both scenarios. There is an additional 25.59% of storage facilities need to be installed per year on average after considering uncertainty, the distributions of which in 2040 and 2060 are shown in Figure 4. Natural gas storage is mainly built in Xinjiang, Heilongjiang and Shaanxi.

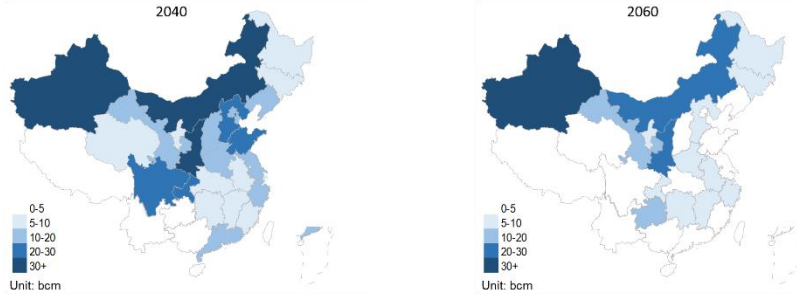


Figure 4 Distribution of storage facilities of China considering uncertainty

For natural gas networks, considering uncertainty hardly changes the total capacity of the network, but makes the capacity of each pipeline more uniform. Figure 5 shows the distribution of natural gas pipeline networks in 2040 under both scenarios. The orange lines represent the main pipelines with an annual transmission capacity of more than 100bcm, which is less than 100bcm in the grey lines. It shows that the distribution of the main pipe network is wider after considering the uncertainty.



Figure 5 Distribution of pipeline of China in 2040

### 3.3. Sensitivity Analysis

When considering uncertainty, different confidence levels may lead to different energy system optimization results. The results of sensitivity analysis are shown in Figure 6. when the confidence level is 90 % ~ 99 %, considering the uncertainty will not bring great changes to the natural gas supply system structure and transition cost. When the confidence level is higher than 99.9 %, the impact of capacity of LNG ports and the cost of transition begin to show an exponential rise. The results can provide a reference for policy makers to build what size of natural gas infrastructure in response to the uncertainty while taking the cost of transition into account.

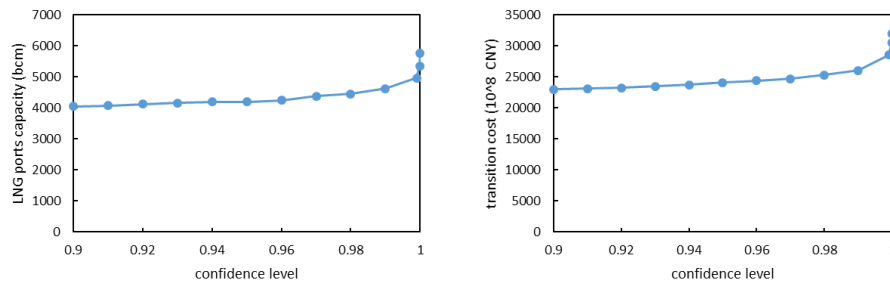


Figure 6 Sensitivity analysis of confidence levels

#### 4. Conclusions

This study proposes a regional natural gas supply system optimization model, which can be used to quantitatively analyze the optimal transition path for a natural gas supply system with minimum transition costs. A chance-constrained programming approach is adopted to consider the impact of uncertainty of import fluctuation, pipeline aging and natural gas demand. Finally, sensitivity analysis of confidence level is carried out to show the impacts of it on the capacity planning of natural gas infrastructure as well as the transition costs.

The results show that increasing capacity of LNG ports and natural gas facilities can enhance the ability to cope with uncertainty. With consideration of uncertainty, under confidence level as 99 %, an additional 40.75 % of LNG ports capacity and an additional 25.59 % of gas storage capacity need to be installed per year on average. In addition, Sensitivity analysis of confidence level shows that rapid growth in transition cost occurs when the confidence level exceeds 99.9 %, which provides a reference for decision makers to make energy system planning while weighing transition costs against natural gas supply stability.

For future research, more uncertainties can be taken into account, such as price uncertainty. In addition, the coupling research of natural gas system and hydrogen system is meaningful. Natural gas is a transition period energy, and hydrogen will be promising in the future. Researching on the coupling of them that natural gas infrastructure can be converted into hydrogen infrastructure can reduce investment costs.

#### References

- BP, 2022, Statistical Review of World Energy.
- T. Li, P. Liu, and Z. Li, 2019, Modelling and optimization of a natural gas supply system at a transient stage: a case study of China. *BMC Energy*. 1: 5.
- Huang W., Li Y., Yu W., Yu H., Shan X., Wang H. and Gong J., 2021, An evaluation index system of the user satisfaction for the natural gas pipeline network. *Journal of Pipeline Science and Engineering*. 1(4): 452-458.
- Dieckhoener C, Lochner S, Lindenberger D, 2013, European natural gas infrastructure: the impact of market developments on gas flows and physical market integration. *Appl Energy*. 102: 994-1003.
- Chaudry M, Jenkins N, Qadrdan M, Wu JZ, 2014, Combined gas and electricity network expansion planning. *Appl Energy*. 113:1171-87.
- Chiang N-Y, Zavala VM, 2016, Large-scale optimal control of interconnected natural gas and electrical transmission systems. *Appl Energy*. 168:226-35.
- Zhang Q, Li Z, Wang G, Li H., 2016, Study on the impacts of natural gas supply cost on gas flow and infrastructure deployment in China. *Appl Energy*, 162, 1385-98.
- Charnes A., Cooper W. W., 1959, Chance-Constrained Programming, *Management Science*, 6(1), 73-79.

# Efficient 1D modelling of hot melt extrusion process for pharmaceutical applications

Michaela Vasilaki,<sup>a</sup> Umair Zafar,<sup>b</sup> Ioannis .S. Fragkopoulos,<sup>c</sup> Ioannis K. Kookos,<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, University of Patras, Rio, 26504, Patras Greece*

<sup>b</sup>*Oral Formulation Research, Novo Nordisk A/S, Måløv, Denmark*

<sup>c</sup>*Oral Drug Product Process Development, Novo Nordisk A/S, Måløv, Denmark*

## Abstract

Hot melt extrusion (HME) involves the melting of materials and mixing and pushing of various solids under elevated temperature and pressure. It is a mature technology that has been extensively used for a broad spectrum of applications in the polymers, food, and pharmaceutical industries. HME can be a promising downstream processing method due to its high efficiency and economic value, however mathematical modelling of HME processes has always been challenging due to the limited understanding and the demanding experimental techniques. In this work, an 1D model is developed for the efficient simulation of the steady state and dynamic operation of HME processes. It is shown that the steady state operation can be described as a classical complementarity problem. The solution of the resulting complementarity problem can be performed efficiently in available software. The determination of the consistent initial conditions facilitates the solution of the system of differential and algebraic equations (DAEs).

**Keywords:** Hot melt extrusion, Pharmaceutical industry, Simulation, Complementarity problem.

## 1. Introduction

HME is a mature technology, that has been used extensively in the polymers industry, and involves the melting of materials and/or mixing and pushing of various solids under elevated temperature and pressure. There has been an active interest in extending the application of HME technology in the polymers, food, and pharmaceutical industry (Crowley et al., 2007). Mathematical modelling of HME processes has always been challenging due to the limited understanding of the material behavior, its complex geometry, the complex interactions between the material processed and the equipment, the existence of multiple zones (solids zone and melt zone) and the demanding experimental techniques necessary to validate theoretical models.

There has been extensive research in developing efficient mathematical models which can be used for process synthesis. Most of the existing modelling approaches have been reviewed recently (Grimard et al., 2016; Kohlgruber, 2020). The objective of this study is to propose a novel approach to the solution of 1D models for HME processes. The proposed methodology is based on transforming the algebraic model that describes the steady state operation using material, energy, and momentum balances as well as logical conditions, into an equivalent complementarity problem (a problem that has been extensively studied in the mathematical programming literature). The proposed novel approach is proven to be computationally efficient and is extended into the dynamic model of the HME process.

## 2. 1-D model of the HME process

### 2.1. Steady state model

Figure 1 shows the general pressure-throughput relationship for extruders that is discussed extensively by Kohlgrüber (2020). The following model is proposed for creeping flows of Newtonian fluids:

$$\dot{V}^* = \dot{V}_{max}^* - \frac{\dot{V}_{max}^*}{\Delta\dot{p}_{max}^*} \Delta p^* \quad (2.1)$$

where the dimensionless pressure drop  $\Delta p^*$  and flowrate  $\dot{V}^*$  are defined as follows:

$$\Delta p^* = \frac{D}{\eta \mathcal{N}} \cdot \frac{\Delta p}{\Delta z}, \quad \dot{V}^* = \frac{\dot{V}}{\mathcal{N} \cdot D^3} \quad (2.2)$$

where  $D$  is the diameter of the barrel,  $\eta$  is the viscosity,  $\mathcal{N}$  is the rotational frequency,  $\Delta p$  is the pressure drop observed in an element with axial length  $\Delta z$  and  $\dot{V}$  is the volumetric flowrate. After substituting Equations (2.2) into Equation (2.1) we obtain:

$$\dot{m} = m_{fb} + m_p = \dot{V}_{max}^* \cdot \rho (\mathcal{N} \cdot D^3) - \frac{\dot{V}_{max}^*}{\Delta\dot{p}_{max}^*} D^4 \frac{\rho}{\eta} \cdot \frac{\Delta p}{\Delta z} \quad (2.3)$$

where  $\rho$  is the density,  $m_{fb}$  denotes the mass flowrate for forward ( $m_f$ ) or backward ( $m_b$ ) moving elements and  $m_p$  the pressure induced mass flowrate. Equation (2.3) holds true for completely filled elements. For partially filled elements with filling ratio  $f$ :

$$\dot{m} = m_{fb} + m_p \quad (2.4)$$

$$m_{fb} = \dot{V}_{max}^* \cdot f \cdot \rho (\mathcal{N} \cdot D^3) \quad (2.5)$$

$$m_p = - \frac{\dot{V}_{max}^*}{\Delta\dot{p}_{max}^*} D^4 \frac{\rho}{\eta} \cdot \frac{dp}{dz} \quad (2.6)$$

In order to develop a mathematical description of the twin-extruder, the extruder is divided into  $i=\{1, 2, \dots, n\}$  discrete elements and material balances are applied in each element. The discrete representation of the solution space is shown in Figure 2. A material balance can be written for each discrete element as follows:

$$m_f(i-1) + m_b(i+1) + m_p(i-1) = m_f(i) + m_b(i) + m_p(i) \quad (2.7)$$

$$m_{fb}(i) = \dot{V}_{max}^*(i) \cdot (1 - \varepsilon(i)) \cdot \rho(i) \cdot (\mathcal{N} \cdot D^3) \quad (2.8)$$

$$m_p(i) = - \frac{\dot{V}_{max}^*(i)}{\Delta\dot{p}_{max}^*(i)} D^4 \frac{\rho(i)}{\eta(i)} \cdot \left( \frac{p(i+1) - p(i)}{z(i+1) - z(i)} \right) \quad (2.9)$$

where we have introduced the void fraction  $\varepsilon(i) = 1 - f(i)$ . The pressure-induced mass flowrate implied by Equation (2.9) involves a conditional statement: pressure buildup or pressure generation can only be developed in completely filled elements ( $f=1$  or  $\varepsilon=0 \Rightarrow p(i)>0$ ). For partially filled elements the pressure is equal to zero ( $f<1$  or  $\varepsilon>0 \Rightarrow p(i)=0$ , apparently the pressure is defined relative to the atmospheric pressure).

Equation (2.7), using Equations (2.8) and (2.9), can be written at steady state in the following form where the unknowns ( $\varepsilon(i)$ ,  $p(i)$ ) appear linearly:

$$\mathbf{A}(\dot{V}_{max}^*(i), \Delta\dot{p}_{max}^*(i), D, \mathcal{N}; \rho, \eta) \cdot \mathbf{x} = \mathbf{b}(\dot{V}_{max}^*(i), D, \mathcal{N}; \rho) \quad (2.10)$$

$$\mathbf{x} = \begin{bmatrix} \boldsymbol{\varepsilon} \\ \mathbf{p} \end{bmatrix} = [\varepsilon(1) \quad \varepsilon(2) \quad \dots \quad \varepsilon(n) \quad p(1) \quad p(2) \quad \dots \quad p(n)]^T \quad (2.11)$$

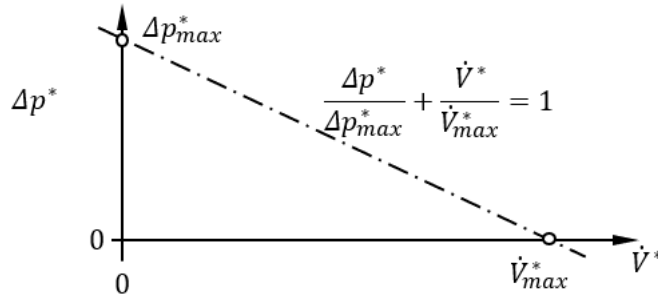


Figure 1. Pressure versus volumetric flowrate for a typical extruder element.

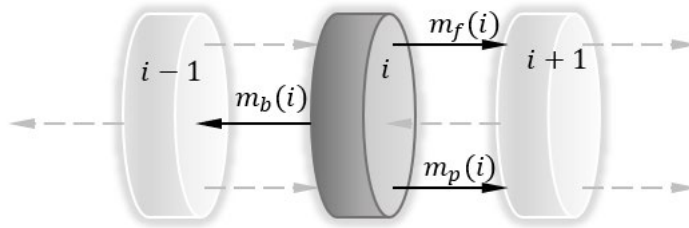


Figure 2. Building block of the 1D model for a twin-screw extruder.

For constant thermophysical properties ( $\rho, \eta$ ) and given flow characteristic and operating parameters ( $\dot{V}_{max}^*(i), \Delta p_{max}^*(i), D, \mathcal{N}$ ), Equation (2.10) is a linear underdetermined system of  $n$  equations in  $2n$  unknowns (see also Choulak et al., 2004 and Eitzlmayr et al., 2014). The system of linear equations can be solved by adding  $n$  specifications to the mathematical description. This has been done in previous studies by first establishing the state of each element (partially filled/completely filled) and then calculating the resulting pressure profile. Numerical problems have been reported in previous studies which have been attributed to inefficiencies of the commercial software used. A systematic methodology is presented in what follows that does not suffer from methodological or numerical issues and allows solution of the problem using arbitrarily fine discretization of the solution space.

The important observation towards developing an efficient methodology for the steady state simulation of extruders is that, apart from the system of  $n$  underdetermined Equations (2.10), there is for each element  $i$  an additional logical condition that needs to be satisfied. This condition can be stated as: either the element is partially filled and the pressure is zero or the element is completely filled, and the void fraction is zero. This can be expressed mathematically as follows:

$$\varepsilon(i) \cdot p(i) = 0, i = 1, 2, \dots, n \quad (2.12)$$

this is a set of  $n$  additional nonlinear equations. When Equation (2.12) is written together with Equation (2.10), then the system of  $n$  equations in  $n$  unknowns is obtained:



$$\mathbf{A} \cdot \begin{bmatrix} \boldsymbol{\varepsilon} \\ \mathbf{p} \end{bmatrix} = \mathbf{b}, \quad \boldsymbol{\varepsilon}^T \cdot \mathbf{p} = \mathbf{0}, \quad \mathbf{0} \leq \boldsymbol{\varepsilon} \leq \mathbf{1}_n, \quad \mathbf{0} \leq \mathbf{p} \leq \mathbf{1}_n P_{max} \quad (2.13)$$

Formulation (2.13) is well known in the mathematical programming literature as the linear complementarity problem (LCP), for which efficient solution methodologies are now available (Kookos, 2022, Cottle, 1992). One of the alternative methodologies can be based on the definition of binary variables  $\delta(i)$ :

$$\delta(i) = \begin{cases} 1, & \text{if } \varepsilon(i) = 0 \\ 0, & \text{if } p(i) = 0 \end{cases} \quad i = 1, 2, \dots, n \quad (2.14)$$

Using the definition (2.14) the LCP problem can be postulated as a mixed-integer, linear programming problem (MILP):

$$\begin{aligned} \min_{\boldsymbol{\varepsilon}, \mathbf{p}, \boldsymbol{\delta}} \quad & \mathbf{1}_n^T \boldsymbol{\delta} \\ \text{s. t.} \quad & \end{aligned} \quad (2.15)$$

$$\mathbf{A} \cdot \begin{bmatrix} \boldsymbol{\varepsilon} \\ \mathbf{p} \end{bmatrix} = \mathbf{b}, \quad \mathbf{0} \leq \boldsymbol{\varepsilon} \leq \mathbf{1}_n - \boldsymbol{\delta}, \quad \mathbf{0} \leq \mathbf{p} \leq \boldsymbol{\delta} \cdot P_{max}, \quad \boldsymbol{\delta} \in \{0, 1\}^n$$

## 2.2. Dynamic model of an extruder

The dynamic model of an extruder consists of the dynamic material balances:

$$\frac{dM(i)}{dt} = m_f(i-1) + m_b(i+1) + m_p(i-1) - (m_f(i) + m_b(i) + m_p(i)) \quad (2.16)$$

$$\begin{aligned} \frac{dm(i)}{dt} = & (m_f(i-1) + m_p(i-1)) \cdot \omega(i-1) - (m_f(i) + m_b(i) + m_p(i)) \cdot \omega(i) \\ & + m_b(i+1) \cdot \omega(i+1) \end{aligned} \quad (2.17)$$

where  $M(i) = \rho(i) A_{cr} \delta z f(i)$  is the total mass holdup in element  $i$ ,  $m(i) = M(i)\omega(i)$  is the mass holdup of component in element  $i$ ,  $\omega(i)$  is the mass fraction of the component of interest,  $A_{cr}$  is the cross-sectional area available for materials flow and  $\delta z$  the axial length of the discrete element. The dynamic model also includes the algebraic equations (2.8) and (2.9). There are as many differential equations as differential variables but there are  $n$  more algebraic variables than algebraic equations. To avoid using the complementarity constraint in the Differential-Algebraic Equations (DAE) model, that will give rise to a high index problem, we are assuming that the pressure dynamics is significantly faster than the flow dynamics. If the solution of the DAE model is available at time  $t$ , then the filling ratio is known, and the pressure profile can be established. With the pressure profile known the DAE model can be solved and obtain the dynamics of the material flow through the extruder. In this way the DAE model is, in effect, simplified to an ordinary differential equation (ODE) model.

## 3. Case study and discussion

We consider the twin-screw extruder studied experimentally by Puaux et al., (2000) (screw configuration B) and theoretically by Eitzlmayr et al., (2014). The experimental extruder is a twin-screw, corotating extruder with a centerline distance of  $C_L=0.021$  m, an external screw diameter of  $D=0.025$  m and a screw core diameter of  $D_C=0.016$  m. The extruder consists of seven screw type elements arranged in series with lengths (from hopper to die) 50/50/100/50/150/200/300 mm and pitches ( $T_S$ ) 16/25/33/25/16/25/33 mm. All elements are right-handed elements (forward pushing and/or pressure generating elements) apart from the fourth one which is left-handed (backward pushing or pressure

consuming element). Several combinations of operating conditions with mass flowrates in the range  $\dot{m}_0 \in [3.8 \text{ kg/h}, 10.8 \text{ kg/h}]$  and screw speeds of  $\mathcal{N} \in [200 \text{ rpm}, 500 \text{ rpm}]$  are considered in the experimental equipment and the results of residence time distribution determination experiments are presented in Puaux et al., (2000). The parameters of the pressure/volumetric flowrate characteristics ( $\dot{V}_{max}^*$ ,  $\Delta p_{max}^*$ ) of the various screw elements need to be estimated. For the maximum dimensionless volumetric flowrate Kohlgrüber (2020) proposes that  $\dot{V}_{max}^* = T_s/2D$ . Parameter  $\Delta p_{max}^*$ , which needs to be determined experimentally, was assumed to be equal to 2000, a value that appears to be reasonable for the screw diameter and pitch of the elements in the experimental extruder (according to experimental curves reported by Kohlgrüber (2020)).

The LCP formulation (2.15) is solved first for the case of  $\mathcal{N} = 400 \text{ rpm}$  and  $\dot{m}_0 = 6.5 \text{ kg/h}$  using CPLEX/GAMS with an increasing number of discretization elements. The overall length of the extruder is 900 mm and 90, 900, 9000 and 90000 discrete elements are selected corresponding to discrete element lengths of 10 mm, 1 mm, 0.1 mm, and 0.01 mm, correspondingly. In Figure 3.a the profile for the filling ratio for 90 and 90000 discrete elements, are presented and compared (the results for 900 and 9000 discrete elements coincide with that of the 90000 elements). The necessary computational time was 0.07 s for the 90 elements and 20.1 s for the 90000 elements (Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz, 6GB memory computer). The computational time is found to scale with  $n^2$  for the intermediate discretizations. As the formulation as an LCP problem was proven to be extremely efficient, we were able to solve repeatedly the formulation by varying parameters such as the feed mass flowrate or the speed of rotation. In Figure 3.b the results obtained by varying systematically the mass flowrate of the feed for a constant speed of rotation is shown. Similar plots can be obtained by varying the speed of rotation for a constant mass flowrate of the feed stream.

To demonstrate the ability of the proposed formulation to predict the dynamic response of the experimental extruder, the dynamic model was solved for a unit pulse experiment in MATLAB using the ode15s DAE solver (the pressure dynamics is assumed faster than the flow dynamics). The volume of each element of the discrete formulation was determined based on the geometry, by calculating the cross-sectional area based on the work by Booy (1980). The experimental results are compared with the simulation results in Figure 4. The agreement is more than satisfactory and proves that the proposed formulation is sufficient for simulating the dynamics of complex twin-screw extruders.

#### 4. Summary and Conclusions

Preliminary results on a novel approach for solving the steady state and dynamic model for complex twin-screw extruders has been presented and validated with experimental results. The proposed formulation that is based on the linear complementarity problem has been proven to be a powerful approach for solving the steady state determination problem in twin-screw extruders. The model, at its present form, has been developed for constant thermophysical properties that are independent from pressure and temperature (density and viscosity). The model needs to be extended to incorporate the temperature dynamics that involves the energy balance for the material, barrel, and screw. It also involves the use of constitutive equations for the density and viscosity to link material flow dynamics with energy dynamics. The preliminary results are promising enough to guarantee that further investigation will improve the predictive capabilities of the model and will allow us to obtain a better and more detailed understanding of the complex and multiscale dynamics of twin-screw extruders. Finally, discrete element simulations can

be used to determine the characteristic parameters of the pressure/throughput curve so as to reduce our reliance on experimental results for developing high fidelity models.

### Acknowledgement

M. Vasilaki & I.K. Kookos were supported for this work by the project “INVALOR: Research Infrastructure for Waste Valorization and Sustainable Management” (MIS 5002495) which is implemented under the Action “Reinforcement of the Research and Innovation Infrastructure”, funded by the Operational Program “Competitiveness, Entrepreneurship and Innovation” (NSRF2014–20) and co-financed by Greece and the EU (EU Regional Development Fund).

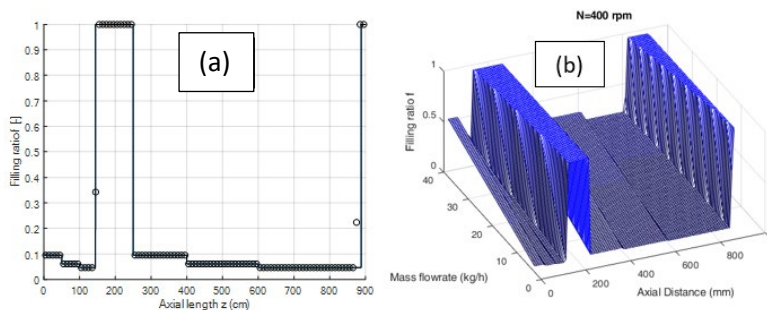


Figure 3. Filling ratio for the experimental extruder for (a) different number of discrete elements (continuous line 90000 elements, o 90 elements) and (b) several values of the mass flowrate and fixed frequency of rotation ( $N = 400$  rpm).

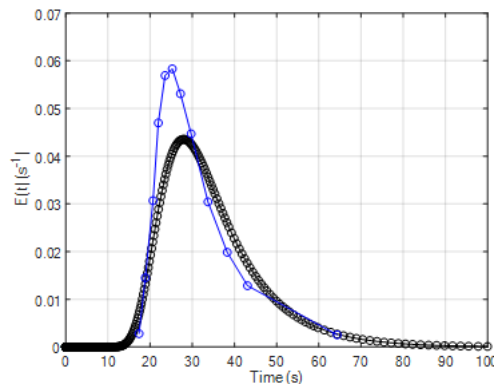


Figure 4. Comparison of the experimental results and simulation results for the RTD experiment ( $N = 400$  rpm,  $\dot{m}_0 = 6.5$  kg/h).

### References

- M.L. Booy, M.L., 1980, Isothermal Flow of Viscous Liquids in Corotating Twin Screw Devices, *Polymer Engineering and Science*, Vol 20(18), pp. 1220-1228.
- S. Choulak, Couenne, F., Gorrec, Y.L., Jallut, C., Cassagnau, P., Michel, A., 2004, Generic Dynamic Model for Simulation and Control of Reactive Extrusion, *Industrial Engineering Chemistry Research*, 43, pp. 7373-7382.
- R.W. Cottle, Pang, J.S., Stone, R.E., 2009, *The linear Complementarity Problem*, Siam Philadelphia.

- M.M., Crowley, Zhang, F., Repka, M.A., Thumma, S., Upadhye, S.B., Kumar Battu, S., McGinity, J.W. and Martin, C., 2007, Pharmaceutical Applications of Hot-Melt Extrusion: Part I., Drug Development and Industrial Pharmacy, 33(9), pp.909–926.
- A. Eitzlmayr, Koscher, G., Reynolds, G., Huang, Z., Booth, J., Shering, P., Kninast J., 2014, Mechanistic modelling of modular co-rotating extruders, International Journal of Pharmaceutics, 474, pp. 157-176.
- J. Grimard, Dewasme, L., Vande Wouwer, A., 2016. A Review of Dynamic Models of Hot-Melt Extrusion. Processes, 4(2), p.19.
- K. Kohlgruber, 2020, Co-Rotating Twin-Screw Extruders: Fundamentals, Carl Hanser, Munich.
- I.K. Kookos, 2022, Practical Chemical Process Optimization, Springer-Nature.
- J.P. Puaux, Bozga, G., Aisner, A., 2000, Residence Time Distribution in a Co-rotating Twin-Screw Extruder, Chemical Engineering Science, 55, pp. 1641-1651.



# Towards Machine Learning of Power-2-Methanol Processes

Carl Julius Martensen,<sup>a</sup> Christoph Plate,<sup>a</sup> Tobias Keßler,<sup>b</sup> Christian Kunde,<sup>c</sup> Lothar Kaps,<sup>c</sup> Achim Kienle,<sup>b,c</sup> Andreas Seidel-Morgenstern,<sup>c,d</sup> Sebastian Sager<sup>a</sup>

<sup>a</sup>*Institute of Mathematical Optimization, Otto-von-Guericke University, Universitätsplatz 2, 39106 Magdeburg, Germany*

<sup>b</sup>*Institute for Automation Engineering, Otto-von-Guericke University, Universitätsplatz 2, 39106 Magdeburg, Germany*

<sup>c</sup>*Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany*

<sup>d</sup>*Institute of Process Engineering, Otto-von-Guericke University, Universitätsplatz 2, 39106 Magdeburg, Germany*

## Abstract

Many dynamic models in process engineering rely on uncertain or even largely unknown mechanism. These parts are often modeled using best-practice knowledge or heuristics, which may result in 1) overly complex models, 2) models which do not reproduce the expected predictive outcome or 3) ill posed optimization problems. A promising approach is the use of machine learning surrogate models inside a partially known mechanistic model. In this study the impact of measurement noise and sample rate on the predictive performance of such a hybrid model is investigated. As an example, process methanol synthesis is used. Additionally, it is shown that the surrogate model can learn even unobservable states from the data.

**Keywords:** Scientific Machine Learning, Methanol Synthesis, Automated Modeling, System Identification

## 1. Introduction

Methanol is an important basic chemical. In the frame of the upcoming energy revolution, power-2-methanol processes become an interesting option for methanol production. These processes rely on green hydrogen, which is produced from regenerative electrical energy by water electrolysis and are therefore usually operated under dynamic conditions due to inevitable fluctuations in the regenerative energy supply. One approach is to generate methanol through the hydrogenation of CO and CO<sub>2</sub> (Olah et al., 2009). Although methanol synthesis is a commonly employed process, its kinetic modeling is still a field of ongoing research. The influence of the dynamic catalyst state on the reaction rates is an important part of the model, which is, however, hard to model from first principles. A promising approach to overcome this challenge is to replace this sub-model with an artificial neural network (ANN) and combine it with the remaining first principle model to obtain a hybrid model. It is not clear a priori how much data is required to successfully train the resulting model. This study investigates the influence of data quality on such a data-driven model. Artificial datasets of reactor output trajectories with

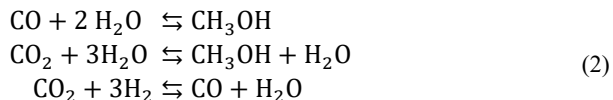
different measurement frequencies and amounts of added noise are generated from a rigorous model (Seidel et. al., 2018, 2021). A hybrid model is then derived by replacing the sub-model that describes the influence of the catalyst state on the reaction rates with an ANN. Hyperparameter optimization of different surrogate architectures, objectives, and penalty functions is conducted, and the best-performing model structure is identified. The hybrid model is trained on the different datasets and the results are evaluated by comparing the state trajectories with those of the rigorous model. Additionally, the trained ANNs are compared with the original sub-model to evaluate the ability to reconstruct the original sub-model behaviour from the different synthetic datasets. The contributions of this analysis are twofold: First, it leads to a better understanding of the necessary data quality for training hybrid models of heterogeneously catalysed methanol synthesis. Second, it generates insights into the parsimony of available rigorous models depending on the operating conditions.

## 2. Materials and Methods

The overall dynamics of the continuous stirred tank reactor (CSTR) for methanol synthesis can be described as a system of differential algebraic equations (DAE) of the form (Seidel et. al., 2018, 2021)

$$\begin{aligned} f(\dot{x}, x, p, t, u, c) &= 0 \\ h(x, p, t, u, c) &= y \end{aligned} \quad (1)$$

where  $x \in \mathbb{R}^{n_x}$  denotes the states of the system and  $\dot{x} \in \mathbb{R}^{n_x}$  its time derivative subject to the set of DAE given by  $f$ . The states of the system consist of the mole fractions of the individual species  $x_i \in (0,1)$  with the additional constraint  $\sum_i x_i = 1$ ,  $i \in (\text{CH}_3\text{OH}, \text{CO}_2, \text{CO}, \text{H}_2, \text{H}_2\text{O}, \text{N}_2)$  and the percentage of reduced centres of the catalyst  $x_\phi \in (0,1)$  (Seidel et. al., 2018).  $p \in \mathbb{R}^{n_p}$  is a vector of parameters,  $t \in \mathbb{R}$  the time,  $u \in \mathbb{R}^{n_u}$  the external inputs to the system, representing an incoming feed of species and a specific volumetric flow rate. The function  $h$  maps the states onto the observed variables  $y \in \mathbb{R}^{n_y}$ . The constant operating conditions  $c = (T, P)$  collect the information about temperature and pressure  $T \in \mathbb{R}$  and  $P \in \mathbb{R}^+$ , respectively. Three reaction pathways are used to describe the synthesis of methanol from CO, CO<sub>2</sub>, and H<sub>2</sub> over CuO/ ZnO/ Al<sub>2</sub>O<sub>3</sub> (unreduced state):



where the first two reactions describe the methanol synthesis via hydrogenation and the last reaction is the reverse water-gas shift reaction (RWGS). The dynamical system following from [Eq. \(2\)](#) is still not fully understood. Especially the influence of the catalyst

on the reaction rate is not clear. Following (Seidel et. al., 2018, 2021) a scaled reaction rate is used to take this effect into account:

$$\bar{r} = [\gamma_{CO} r_{CO} \quad \gamma_{CO_2} r_{CO_2} \quad \gamma_{RWGS} r_{RWGS}]^T \quad (2)$$

where  $\bar{r}_i \in \mathbb{R}$  represents the effective reaction rate of the reaction  $i \in$  as the product of the influence of the catalyst  $\gamma_i \in \mathbb{R}^+$  and the reaction rates based upon a Langmuir–Hinshelwood mechanism  $r_i \in \mathbb{R}$ .

The influence of the catalyst state on the effective reaction rates is of special interest, given that it is a challenging reaction over a distributed parameter space. The current implementation assumes a dependence on the catalyst state and is modeled based on practical heuristics rather known first principles

$$\gamma = \left[ 1 - x_\phi \quad x_\phi^2 \quad \frac{x_\phi}{1 - x_\phi} \right]^T \quad (4)$$

Within our study an existing model of an isochoric, isobaric, isothermal CSTR for methanol synthesis is used, which is based on the dynamics of a steady-state experiment. Its goal is to investigate the influence of data quality and sample distance on the approximative performance of an embedded ANN. While the idea of using ANNs itself is not new, see e.g., (Ljung et. al., 2020), data-driven inference of dynamical systems has gained much attention during the past decade, e.g., extracting (physical) signals (Cranmer et. al., 2020), operator learning (Kovachki et. al., 2021), and approximating solutions to PDE, ODE and DAE systems (Raissi et. al., 2021).

The rigorous model (Seidel et. al., 2018, 2021) has been implemented using Julia (Benzanson et. al., 2015). To assess the use of ANN as a tool to extract and learn unknown parts of our process model, the approach described in (Rackauckas et. al., 2021) is followed. Here, an ANN is embedded inside the known model, effectively extending the previously introduced model with a universal approximator  $\hat{\gamma} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_\theta} \mapsto \mathbb{R}^3$  containing trainable parameter  $\theta \in \mathbb{R}^{n_\theta}$ . The hybrid model takes the form

$$\begin{aligned} \hat{f}(\dot{x}, x, p, t, u, c, \hat{\gamma}(x, \theta)) &= 0 \\ h(x, p, t, u, c) &= \hat{y} \end{aligned} \quad (3)$$



where  $\hat{f}$  denotes the known parts of the model in composition with the ANN  $\hat{\gamma}$  replacing  $\gamma$  within the effective reaction rates  $\tilde{r}$ .

### 3. Numerical Experiments

The fully known ground-truth model was used to generate a dataset  $\mathcal{D} = (d_1, \dots, d_{30})$  of 30 trajectories  $d_i = (X, t_X)$ . Here,  $X \in \mathbb{R}^{n_x \times n_m}$  denotes the collected state trajectory over  $m$  measurements in a matrix form and  $t_X \in \mathbb{R}^{n_m}$  the corresponding time points. The initial conditions of the different species mole fractions were sampled randomly from a uniform distribution and normalized. Additionally, the condition that  $x_{H_2} \in [0.35, 0.8]$  has been imposed by rejecting invalid samples. The initial conditions of the catalyst state  $x_\phi$  are sampled uniformly within the interval  $[0.2, 0.85]$ . The operating conditions are sampled from intervals  $[460.15, 490.15]$  K for the temperature and  $[55.0, 65.0]$  bar for the total pressure. Over all datasets, the incoming species and volumetric flow rate is kept constant. Using the forward solution of the underlying DAE system, dense solutions were created which can be subsampled. Each trajectory was simulated within the time interval  $t \in [0, 3000]$  s and a measurement sample rate  $t_M \in \mathbb{R}$  of 1 s. Synthetic noise has been added to each observable state<sup>1</sup>  $y = [x_{CH_3OH} \ x_{CO_2} \ x_{CO} \ x_{H_2} \ x_{H_2O} \ x_{N_2}]^T$  with a multiplicative error  $\tilde{y}_i = (1 + \epsilon_i)y_i$ . The error is sampled from a truncated normal distribution  $\epsilon_i \sim \mathcal{N}(0.0, \sigma_M)$ , such that  $\epsilon_i \in [-0.1, 0.1]$ . To investigate the impact of the noise, the variance  $\sigma_M \in \mathbb{R}^+$  is varied. To optimize the parameters  $\theta$  of the ANN under varying noise and sampling rates, a maximum likelihood approach is used

$$\min_{\theta, \sigma, \sigma_\theta} - \sum_{d \in \tilde{\mathcal{D}}} \ln \bar{\mathcal{L}}(d, \theta, \sigma, \sigma_\theta), \tilde{\mathcal{D}} \subseteq \mathcal{D}, s.t. \ \sigma_i \geq 0, i = 1, \dots, n_y \quad (4)$$

Here  $\bar{\mathcal{L}} : \mathbb{R}^{n_y \times n_m} \times \mathbb{R}^{n_\theta} \times \mathbb{R}^{n_y} \times \mathbb{R}^+ \mapsto \mathbb{R}^+$  denotes the likelihood function for the overall problem for a subset  $\tilde{\mathcal{D}}$  of the training data

$$\begin{aligned} \ln \bar{\mathcal{L}}(d, \theta, \sigma, \sigma_\theta) = & \overbrace{\sum_{(y_i, t_i) \in d} \ln \mathcal{L}_\Sigma(y_i, \hat{y}_i, \sigma)}^{\text{Data fitting}} + \overbrace{\sum_{\theta_i \in \theta} \ln \mathcal{L}_\theta(\theta_i, \sigma_\theta)}^{\text{Parameter penalty}} \\ & + \underbrace{\sum_{\sigma_i \in \sigma} \ln \mathcal{L}_\sigma(\sigma_i)}_{\text{Prior of the scales}} \end{aligned} \quad (7)$$

where the error distributions likelihood is given as  $\mathcal{L}_\Sigma : \mathbb{R}^{n_y} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_y} \mapsto \mathbb{R}^+$ .  $\mathcal{L}_\theta : \mathbb{R}^{n_\theta} \times \mathbb{R}^+ \mapsto [0, 1]$  represents a prior on the parameters of the ANN. The parameter  $\sigma_\theta \in \mathbb{R}^+$  represents the scaling of the prior distribution of the tuneable parameters  $\theta$ ,

<sup>1</sup> Strictly speaking  $x_{H_2O}$  is not observable. Given that  $\sum_i y_i = 1$ , we can compute it if all other mole fractions are known.

which may be interpreted as a parameter penalty to avoid overfitting<sup>2</sup>. As it is common,  $\sigma_\theta$  is introduced as a hyperparameter rather than a true optimization variable. Finally,  $\mathcal{L}_\sigma : \mathbb{R}^+ \mapsto [0,1]$  denotes a prior on the scale parameter of the error distribution  $\sigma \in \mathbb{R}^{n_y}$ , assuming  $\sigma_i \sim \text{Lognormal}(0,1)$ .

#### 4. Results

By common practice, hyperparameter optimization was performed to find a suitable architecture for the ANN, the right distribution for the prediction error, and for the prior distribution of the parameters  $\theta$  and their scale  $\sigma_\theta$ . Furthermore, it is assumed that all initial conditions are known, given that only the fitting capabilities of the ANN are investigated. The activation of the output layer has been fixed to be softplus, enforcing positive definite scaling  $\hat{y}$ . Using a Monte Carlo approach, different candidate configuration are sampled and trained using [Eq. \(6\)](#) with ADAM (Kingma and Ba, 2014). As an evaluation metric the Bayesian Information Criterion (BIC) has been chosen. Hyperparameter optimization on the generated dataset was performed using a sample distance  $t_M = 30.0$  s and noise variance  $\sigma_M = 5e^{-2}$  with a train-test-split of 0.8. In addition, the training dataset has been batched to reduce training time while keeping a relatively good performance. In total, 300 samples were evaluated. The best architecture consists of the input layer with  $n_x$  neurons followed by a single hidden layer with 2 neurons, both using a selu activation. The error and prior distribution have been chosen to be Laplacian distributed. [Figure 1](#) shows the result of the hyperparameter optimization for trajectories from both train and test set. Since the ground truth [Eq. \(4\)](#) is known, we can directly compare the learned model. A statistical evaluation of the approximation error for the data can be seen in [Table 1](#), a plot of the approximation of the missing equations can be seen in [Figure 2 \(A\)](#). The ANN learns the scaling for the hydrogenation pathways over CO<sub>2</sub> close to ideal but fails generalize on CO and RWGS. This is in good agreement with available a priori knowledge that for the conditions employed here, RWGS is close to chemical equilibrium and CO<sub>2</sub> is the dominating path for methanol synthesis.

---

<sup>2</sup> E.g., for a normal distributed variable  $x$  with mean  $\mu$  and variance  $\sigma$  we can write the loglikelihood (up to a constant) as  $-\frac{1}{2\sigma^2}(x - \mu)^2 = -\lambda(x - \mu)^2$  which can be thought of as a L<sub>2</sub> penalty with a penalty parameter  $\lambda = \frac{1}{2\sigma^2}$ .

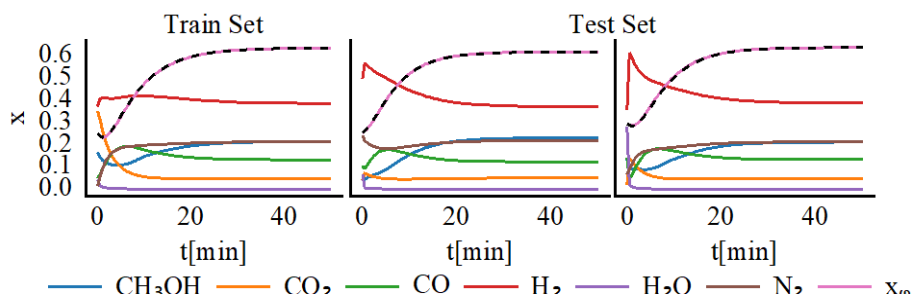


Figure 1 Example trajectories from train and test set. The trained ANNs prediction (full line) from the noisy data (scattered) match the ground truth trajectory of the unobserved state (dashed black).

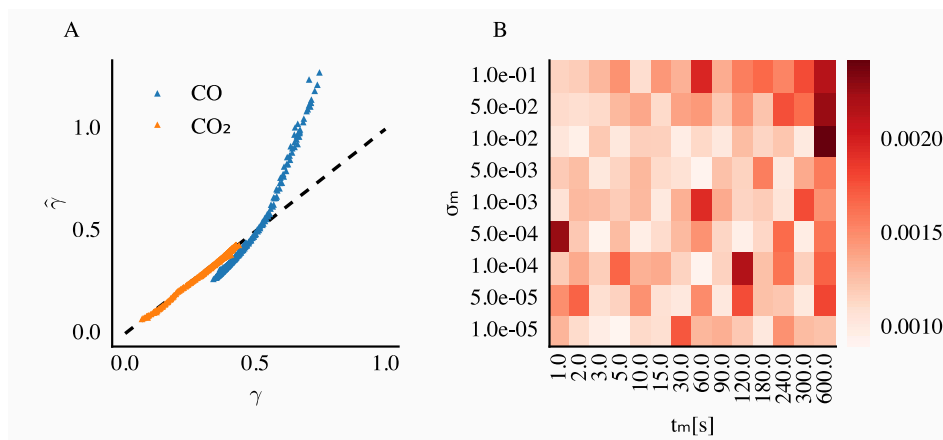


Figure 2 (A) Parity of the ground truth and approximation of the scaling of the hydrogenation pathways evaluated on the test set. RWGS is not shown. (B) Minimum  $L_1$  norm error of the unknown state over varying noise levels and sampling distances evaluated on the test set using the minimum sample distance.

Table 1 : Summary of the  $L_1$  error between the ground truth and approximation of the scaling  $\gamma$

	Train set				Test set			
	Min	Max	Mean	Std.	Min	Max	Mean	Std.
CO - Hydrogenation	1.3e-4	7.2e-1	7.9e-2	7.0e-2	6.8e-4	5.3e-1	8.0e-2	6.8e-2
CO <sub>2</sub> - Hydrogenation	8.5e-7	1.5e-1	5.0e-3	6.3e-3	5.1e-5	2.5e-1	5.5e-3	3.8e-3
RWGS	3.0e-4	2.0e0	7.9e-2	8.0e-2	4.4e-3	1.8e-1	7.4e-2	4.1e-2

For the optimal architecture, the sample distance and noise variance are varied to study the impact of the data quality on the predictive performance. For each pair  $(\sigma_M, t_M)$  the optimal structure is retrained five times and evaluated on the ground truth data with no noise and fully observable measurements. [Figure 2 \(B\)](#) shows the results of the noise

variance study. Overall, the ANN performs well with an overall median error 0.0017 on the test set. Only high noise and high measurement rates perform bad.

## 5. Conclusion

It has been shown that an embedded ANN is able to learn missing components of a physics-based model efficiently using a loglikelihood driven approach. The influence of measurement noise and rate on the predictive performance of the ANN has been

investigated and the resulting model compared to the known ground truth, showing good performance over a wide range of measurement rates and noise levels for the state trajectories. Even if some components cannot be approximated sufficiently, a suitable output is achieved. Several promising future directions open for investigation, e.g., using (optimal) design of experiments, active learning approaches, and learning with additional constraints and unknown parameters like initial conditions. In addition, further testing of our approach on real data collected via an experimental setup will be conducted.

## Acknowledgments

The work was funded by the German Research Foundation DFG within the priority program 2331 'Machine Learning in Chemical Engineering' under grants KI 417/9-1, SA 2016/3-1, SE 586/25-1. Carl Julius Martensen was supported by the European Regional Development Fund (ERDF) within the Research Center of Dynamic Systems (CDS).

## References

- Bezanson, J., Edelman, A., Karpinski, S., Shah, V.B., 2015. Julia: A Fresh Approach to Numerical Computing arXiv:1411.1607.
- Cranmer, M., Sanchez Gonzalez, A., Battaglia, P., Xu, R., Cranmer, K., Spergel, D., Ho, S., 2020. Discovering symbolic models from deep learning with inductive biases. *Adv Neur In* 33, 17429–17442.
- Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. doi:10.48550/ARXIV.1412.6980.
- Kovachki, N., Li, Z., Liu, B., Azizzadenesheli, K., Bhattacharya, K., Stuart, A., Anandkumar, A., 2021. Neural Operator: Learning Maps Between Function Spaces arXiv:2108.08481.
- Ljung, L., Andersson, C., Tiels, K., Schön, T.B., 2020. Deep learning and system identification. *IFAC-PapersOnLine* 53, 1175–1181. doi:10.1016/j.ifacol.2020.12.1329.
- Olah, G.A., Goepfert, A., Prakash, G.K.S., 2009. *Beyond Oil and Gas: The Methanol Economy*. Wiley. doi:10.1002/9783527627806.
- Rackauckas, C., Ma, Y., Martensen, J., Warner, C., Zubov, K., Supekar, R., Skinner, D., Ramadhan, A., Edelman, A., 2021. Universal Differential Equations for Scientific Machine Learning arXiv:2001.04385.

- Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J Comput Phys* 378, 686–707. doi:10.1016/j.jcp.2018.10.045.
- Seidel, C., Jörke, A., Vollbrecht, B., Seidel-Morgenstern, A., Kienle, A., 2018. Kinetic modeling of methanol synthesis from renewable resources. *Chemical Engineering Science* 175, 130–138. doi:10.1016/j.ces.2017.09.043.
- Seidel, C., Nikolić, D., Felischak, M., Petkovska, M., Seidel-Morgenstern, A., Kienle, A., 2021. Optimization of methanol synthesis under forced periodic operation doi:10.3390/PR9050872.

# Water resource portfolio design for optimum water productivity in agriculture

Fatima-Zahra Lahlou,<sup>a</sup> Rajesh Govindan,<sup>a</sup> Bilal Hazrat,<sup>a</sup> Hamish R. Mackey,<sup>b</sup>  
Tareq Al-Ansari<sup>a\*</sup>

<sup>a</sup>*Division of Sustainable Development, College of Science and Engineering, Hamad Bin Khalifa University, Qatar foundation, Doha, Qatar*

<sup>b</sup>*Department of Civil and Natural Resources Engineering, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand*

\*Corresponding Author: [talansari@hbku.edu.qa](mailto:talansari@hbku.edu.qa)

## Abstract

As the demand for food production increases, it is crucial to reassess current water management practices to improve and optimize water use. In this study, the potential of treated sewage effluent in replacing part of the conventional water resources in the agricultural sector is assessed using a mixed integer linear programming optimization model. The proposed model is applied to a case scenario in the state of Qatar, where water from different sources is allocated to four different agricultural subsectors to minimize the carbon footprint while limiting the water stress to a safe global average level which will improve the water productivity of the agricultural sector. Results demonstrate that the proposed model emits twice as much carbon footprint as compared to the base scenario; however, the water stress level is reduced from 323% to 25%, and the water productivity of the agricultural sector increases from  $0.0027 \text{ m}^3$  to  $0.036 \text{ m}^3$ .

**Keywords:** Treated Wastewater; Agriculture; Optimization; Water Productivity; Carbon Footprint.

## 1. Introduction

The global water demand has increased by 600% in the last century and it is now estimated at  $4,600 \text{ km}^3$  (Ritchie and Roser, 2018). With the ever-increasing population and living standards, the demand is expected to increase further by at least 20% in the next 30 years (Wada et al., 2016). While there are enough freshwater resources in the world to meet the global demand, the resources are not equally distributed which leads to water scarcity in some parts of the world. Heading the list of water-scarce regions is the middle east, where countries have been relying on different solutions to reduce water stress including desalination and wastewater reuse. The state of Qatar happens to be the most water-scarce country in the world. It has an annual renewable freshwater resources rate as low as  $21 \text{ m}^3 \cdot \text{capita}^{-1} \cdot \text{year}^{-1}$ ; nevertheless, it has one of the highest water consumption rates globally with over  $182 \text{ m}^3 \cdot \text{capita}^{-1} \cdot \text{year}^{-1}$  (Planning and Statistics Authority, 2019). Whether it is groundwater (GW) abstraction, desalination or wastewater reuse, Qatar must take a holistic approach to find the most suitable solutions. Given its status as one of the most water-scarce countries globally, Qatar must continue to prioritize policies and strategies aimed at reducing its reliance on finite freshwater. The two main direct consequences of the high-water usage in Qatar are the over-abstraction of GW resources and the heavy reliance on energy-intensive desalination to

meet the population's needs. The overuse of GW has resulted in an increase in well salinity to different levels across the country, especially near the shore, such that Reverse Osmosis (RO) is required to treat GW before its safe usage. The state of Qatar also produces a large amount of treated sewage effluent (TSE) of which only a small percentage is reused. That said, TSE represents an opportunity to alleviate the water scarcity in Qatar. Currently, around 50% of the produced TSE is reused and is directed for green space irrigation and fodder production (Planning and Statistics Authority, 2019). If public perception about TSE can be overcome, the reuse of TSE represents a promising solution for Qatar as it can reduce the energy consumed in desalination, reduce water stress and increase water productivity (Lahlou et al., 2022, 2021). That said, around 95 million m<sup>3</sup> of TSE is available and can be injected into the agricultural sector which heavily relies on GW resources. Introducing TSE in the rest of the agricultural subsectors would help reduce water stress and increase the level of water productivity.

This study suggests a water planning framework that aims to optimally allocate the available water resources in Qatar which consist of GW resources, desalinated water (DW) and TSE, to four agricultural sub-sectors which currently rely 100% on wells to meet their needs. The objectives are to do so at a reduced carbon footprint and water stress, for an enhanced water productivity. The developed model is generic and can be applied to other regions.

## 2. Methodology

### 2.1. Case Study

The agricultural subsectors considered in this study are dairy, red meat, egg and poultry production. The annual GDP from these subsectors is 1,86 million Qatari Ryal (Q.R.) (Planning and Statistics Authority, 2020). The water consumption of these sectors is estimated using the annual production information provided by the local authorities and the global average water required to produce the commodities (Mekonnen and Hoekstra, 2010). The water resources considered are GW, DW, and TSE. Groundwater resources have an annual safe yield of 57 million m<sup>3</sup>.year<sup>-1</sup> and have salinity levels ranging between 1000 mg.L<sup>-1</sup> up to 12000 mg.L<sup>-1</sup> based on the well distance from the shore (Baalousha and Ouda, 2017). It is assumed that the energy required for GW pumping is 0.84 kWh.m<sup>-3</sup> and for RO to get rid of the excess Total Dissolved Solids (TDS) level is between 2.6-5.5 kWh.m<sup>-3</sup>, and that the reject water rate ranges between 30-50% (Chen et al., 2019; Enikeeva and Khadra, 2020; Philippe, 2021). The DW is produced by 13 different plants across the country. The three technologies used are Multi-Stage Flash (MSF), Multi-Effect Distillation (MED) and RO. The least energy-intensive technology is RO as it requires as little as 5kWh.m<sup>-3</sup> as opposed to MSF and MED which require 21 kWh.m<sup>-3</sup> (Darwish et al., 2016). The wastewater treatment plants considered in this study are the six plants that receive and treat at least 1% of the total wastewater produced in the country. All six plants use tertiary treatment in addition to Nitrogen & Phosphorus removal (Planning and Statistics Authority, 2019). The supply of DW is carried out using pipeline distribution. As for TSE, two types of distribution are considered, pipeline and truck tankers. Pipeline water distribution requires 0.0055 Kwh.m<sup>-3</sup>.km<sup>-1</sup> (Tow et al., 2021). As for truck tankers distribution, it required an average of 0.4 kWh.km<sup>-1</sup> for every truck tanker with a capacity of 38 m<sup>3</sup> (Lahlou et al., 2020).

### 2.2. Problem Formulation

The objective of the model is to reduce the carbon footprint associated with the supply and treatment of the three water resources considered in this study and it is described in eq (1) while limiting the water stress level of the GW resources to a safe average of 25% which would enhance the water productivity of the agricultural sector. The water stress level and water productivity are calculated using eq (2) and (3).

$$CF_t = (e^p \sum_{i=1}^{i=4} \sum_{j=1}^{j=6} W_i \cdot x_{p-i,j}^{TSE} \cdot d_{p-i,j}^{TSE} + \sum_{i=1}^{i=4} e^{TSE} \cdot W_i \cdot \sum_{j=1}^{j=6} x_{p-i,j}^{TSE} + x_{r-i,j}^{TSE} + e^p \sum_{i=1}^{i=4} \sum_{j=1}^{j=13} W_i \cdot x_{p-i,j}^{DW} \cdot d_{p-i,j}^{DW} + \sum_{i=1}^{i=4} W_i \cdot \sum_{j=1}^{j=13} e_j^{DW} \cdot x_{p-i,j}^{DW} ) \cdot CF_u + e^{GWP} \sum_{i=1}^{i=15} W_i \cdot x_i^{GW} \cdot (1 + \frac{GW_{RO-i}^{\%} \cdot R_i^{\%}}{1-R_i^{\%}}) + \sum_{i=1}^{i=15} W_i \cdot e_i^{GWT} \cdot x_i^{GW} \cdot CF_u + (\sum_{i=1}^{i=4} \sum_{j=1}^{j=6} \text{ceil}(\frac{W_i \cdot x_{r-i,j}^{TSE}}{V_c}) \cdot e^r \cdot d_{r-i,j}^{TSE}) \cdot \frac{CF_d}{E_d} \quad (1)$$

$$WS = \frac{WWD_t}{WA_t - WWI} \quad (2)$$

$$WP = \frac{GDP}{WWD_t} \quad (3)$$

$$WWD_t = \sum_{i=1}^{i=15} W_i \cdot x_i^{GW} \cdot (1 + \frac{GW_{RO-i}^{\%} \cdot R_i^{\%}}{1-R_i^{\%}}) \quad (4)$$

Such that:

- $CF_t$  is the total carbon footprint which needs to be minimized [ton-CO<sub>2</sub>-eq.month<sup>-1</sup>],
- $x_{p-i,j}^{DW}$  is the percentage of the water requirement of sinks  $i$  that is sourced from the desalination plant  $j$  [%],
- $x_{p-i,j}^{TSE}$  represents the percentage of the water requirement of sink  $i$  that is sourced from WWTP  $j$  via pipeline [%],
- $x_{r-i,j}^{TSE}$  represents the percentage of the water requirement of sink  $i$  that is sourced from WWTP  $j$  via road [%],
- $d_{p-i,j}^{DW}$  pipeline distance between the desalination plant  $j$  and the sink  $i$  [km],
- $d_{p-i,j}^{TSE}$  &  $d_{r-i,j}^{TSE}$  distances between sink  $i$  and WWTP  $j$  via pipeline and via road respectively [km],
- $x_i^{GW}$  is the percentage of the water requirement of sink  $i$  that is extracted from GW [%],
- $CF_u$  is the carbon footprint of 1 kWh depending on the country energy mix [g-CO<sub>2</sub>-eq.kWh<sup>-1</sup>],
- $E_d$  is the Energy in 1 litre of diesel [kWh.L<sup>-1</sup>],
- $CF_d$  is the Carbon footprint associated with burning 1 litre of diesel [g-CO<sub>2</sub>-eq.L<sup>-1</sup>],
- $e^p$  is the energy required to transport 1 m<sup>3</sup> of water through 1 km via pipe [kWh.m<sup>3</sup>.km<sup>-1</sup>],
- $e^p$  is the energy required for truck-tanker to travel 1 km [kWh.km<sup>-1</sup>],
- $e^{TSE}$  is the energy required to treat 1 m<sup>3</sup> of TSE [kWh.m<sup>-3</sup>],
- $e^{DW}$  is the energy required to produce 1 m<sup>3</sup> of DW [kWh.m<sup>-3</sup>],
- $e^{GWP}$  is the energy required to pump 1 m<sup>3</sup> of GW [kWh.m<sup>-3</sup>],
- $R_i^{\%}$  is the % of rejected water as brine from the RO unit for sink  $i$  [%],
- $e_i^{GWT}$  is the energy required to produce 1 m<sup>3</sup> of freshwater from GW using RO in sink  $i$  [kWh.m<sup>-3</sup>],
- $GW_{pure-i}^{\%}$  is the percentage of untreated GW required with salinity level  $TDS_{in-i}$  [%],
- $GW_{RO-i}^{\%}$  is the percentage of treated GW required with salinity level  $TDS_f$  [%],
- $R_i^{\%}$  is the percentage of GW that is rejected as brine after reverse osmosis [%],
- $GDP$  is the total gross domestic product [QAR],
- $WWD_t$  is the total water withdrawals by all economic sectors [million m<sup>3</sup>],
- $WA_t$  is the total renewable water availability [million m<sup>3</sup>],
- $WWI$  is the water withdrawal intensity [million m<sup>3</sup>],



The model is constrained by the capacities of the desalination and wastewater treatment plants, the water requirements of the subsectors, in addition to logical constraints as described by eq (5-11).

$$\sum_{i=1}^{15} W_i \cdot (x_{p,i,j}^{TSE} + x_{r,i,j}^{TSE}) \leq C_j^{TSE}, \forall j \quad (5)$$

$$\sum_{i=1}^{15} W_i \cdot x_{p,i,j}^{DW} \leq C_j^{DW}, \forall j \quad (6)$$

$$x_i^{GW} + \sum_{j=1}^{j=6} x_{p,i,j}^{TSE} + x_{r,i,j}^{TSE} + \sum_{j=1}^{j=13} x_{p,i,j}^{DW} = 1, \forall i \quad (7)$$

$$0 \leq x_i^{GW} \leq 1, \forall i \quad (8)$$

$$0 \leq x_{p,i,j}^{TSE} \leq 1, \forall i \quad (9)$$

$$0 \leq x_{p,i,j}^{TSE} \leq 1, \forall i \quad (10)$$

$$0 \leq x_{p,i,j}^{DW} \leq 1, \forall i \quad (11)$$

### 3. Results and Discussion

The total GW usage of the base scenario is estimated at over 184 million m<sup>3</sup>.year<sup>-1</sup>. Assuming that the considered agricultural subsectors are the only users of GW resources, the excessive usage puts the water stress level at over 300%, while the global water stress, which is considered to be safe, is around 18.6% (United Nations, 2022). The carbon footprint of the base scenario is found to be 121.3 thousand ton-CO<sub>2-eq</sub>. As for the water productivity it is estimated to be as little as 0.01 Q.R.m<sup>-3</sup> which is equivalent to \$0.0027 m<sup>-3</sup>.

For limited water stress that is set below or equal to 25%, and enhanced water productivity, the minimum carbon footprint that can be achieved is 253.2 thousand ton-CO<sub>2-eq</sub> which is more than twice that of the base scenario. This is due to the limited GW and TSE resources and the need to meet the rest of the demand via energy-intensive desalination. In fact, more than 70% of the calculated carbon footprint comes from the sea water desalination process (Figure 1). Although the carbon footprint of the base scenario is lower than what can be achieved using the proposed scenario, the water stress level is reduced from 323% to 25%, and the water productivity level increases from 0.010 Q.R.m<sup>-3</sup> to as much as 0.13 Q.R.m<sup>-3</sup>, equivalent to \$0.036 m<sup>-3</sup>. That said, TSE has the potential to alleviate the water stress as well as increase water productivity in the state of Qatar. The water allocation required to achieve these results is described in Figure 2.

Given that all wastewater produced in the State of Qatar has to be treated before being reused or disposed of, it can be argued that the carbon footprint associated with the process of treatment should not be borne by the agricultural sector. This would result in a 19% reduction in the carbon footprint. In addition to that, sludge produced from WWTPs can be used to offset part or all of the carbon footprint associated with the practice, if converted to power.

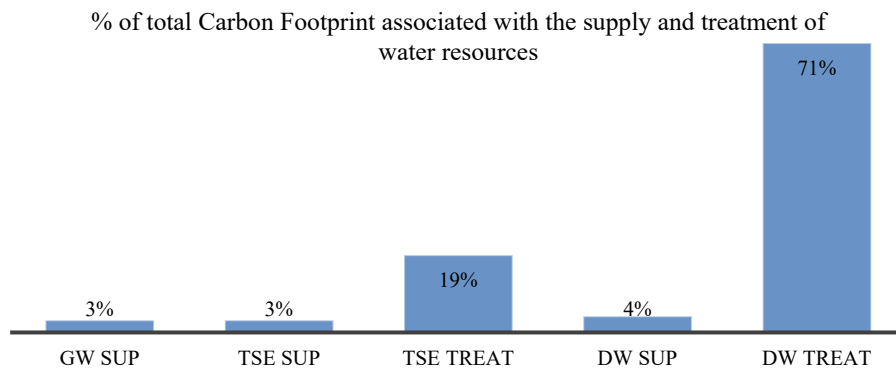


Figure 1: Percentage of the carbon footprint associated with the supply (SUP) and treatment (TREAT) of groundwater (GW) resources, TSE, and desalinated water (DW).

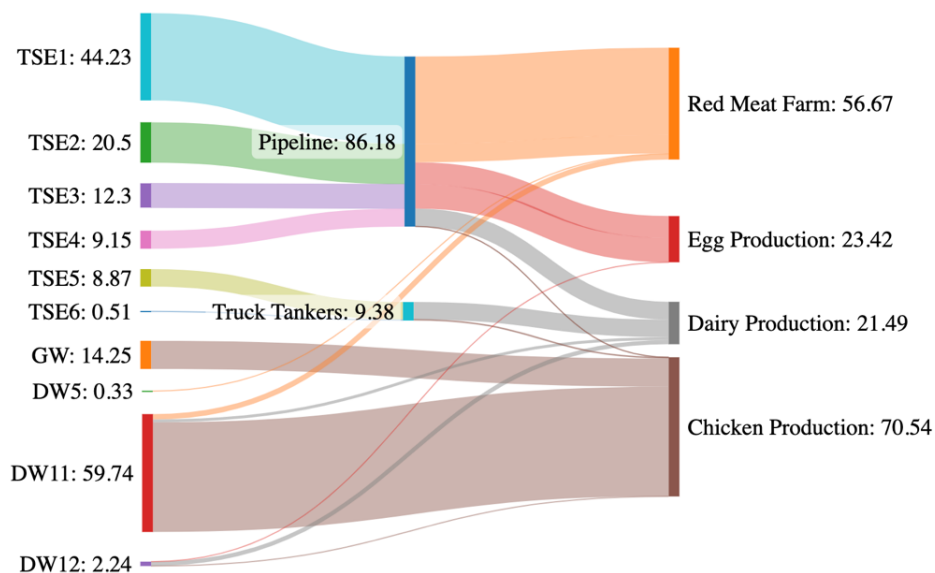


Figure 2: Water resources allocation for the four agricultural subsectors in million  $m^3 \cdot year^{-1}$ .

#### 4. Conclusion

The state of Qatar is one of the most water-scarce countries in the world but also one of the highest users of water. For this reason, there is a need for sustainable water budget planning in the country. One of the possible solutions that can alleviate the water stress in the State of Qatar is the reuse of TSE. This work investigates the optimum water planning framework for GW, DW and TSE water use in four agricultural subsectors for a reduced carbon footprint, limited water stress and improved water productivity. Results suggest that limiting the use of GW resources will increase the carbon footprint associated with the supply and treatment of the agricultural sector’s water requirements, as the needs will need to be met through desalinating seawater. However, this scenario is still considered an improved one since the water stress is reduced from 323% to 25% and the resulting water productivity level is 13 times higher than the base scenario. For future

work, it would be interesting to include the other subsectors of the agricultural sector such as outdoor and indoor farming. In addition to that, it would be interesting to look at how the use of sludge, another important byproduct of WWTPs, can further improve the environmental aspects of the agricultural sector.

### 5. Acknowledgement

This publication was made possible by grant reference MME01-0922-190049 from the Ministry of Municipality and Environment in the State of Qatar. The findings herein reflect the work and are solely the responsibility of the authors.

### References

- Baalousha, H.M., Ouda, O.K.M., 2017. Domestic water demand challenges in Qatar. *Arab. J. Geosci.* 10, 537. <https://doi.org/10.1007/s12517-017-3330-4>
- Chen, X., Thorp, K.R., Ouyang, Z., Hou, Y., Zhou, B., Li, Y., 2019. Energy consumption due to groundwater pumping for irrigation in the North China Plain. *Sci. Total Environ.* 669, 1033–1042. <https://doi.org/10.1016/j.scitotenv.2019.03.179>
- Darwish, M.A., Abdulrahim, H.K., Hassan, A.S., 2016. Realistic power and desalted water production costs in Qatar. *Desalin. Water Treat.* 57, 4296–4302. <https://doi.org/10.1080/19443994.2014.992977>
- Enikeeva, K., Khadra, C., 2020. Water saving greenhouse, trial and demonstration center. Doha.
- Lahlou, F.-Z., Mackey, H.R., Al-Ansari, T., 2022. Role of wastewater in achieving carbon and water neutral agricultural production. *J. Clean. Prod.* 339, 130706. <https://doi.org/10.1016/j.jclepro.2022.130706>
- Lahlou, F.-Z., Mackey, H.R., McKay, G., Onwusogh, U., Al-Ansari, T., 2020. Water planning framework for alfalfa fields using treated wastewater fertigation in Qatar: An energy-water-food nexus approach. *Comput. Chem. Eng.* 141, 106999. <https://doi.org/10.1016/j.compchemeng.2020.106999>
- Lahlou, F.Z., Mackey, H.R., McKay, G., Al-Ansari, T., 2021. Reuse of treated industrial wastewater and bio-solids from oil and gas industries: Exploring new factors of public acceptance. *Water Resour. Ind.* 26, 100159. <https://doi.org/10.1016/j.wri.2021.100159>
- Mekonnen, M.M., Hoekstra, A.Y., 2010. The green, blue and grey water footprint of farm animals and animal products.
- Philippe, B., 2021. Desalination of seawater and brackish water. *Encycl. l'énergie.*
- Planning and Statistics Authority, 2020. *Agricultural Statistics.* Doha.
- Planning and Statistics Authority, 2019. *Water Statistics in the State of Qatar.* Doha.
- Ritchie, H., Roser, M., 2018. *Water use and stress.*
- Tow, E.W., Hartman, A.L., Jaworowski, A., Zucker, I., Kum, S., AzadiAghdam, M., Blatchley, E.R., Achilli, A., Gu, H., Urper, G.M., Warsinger, D.M., 2021. Modeling the energy consumption of potable water reuse schemes. *Water Res.* X 13, 100126. <https://doi.org/10.1016/j.wroa.2021.100126>
- United Nations, 2022. *Ensure availability and sustainable management of water and sanitation for all.*
- Wada, Y., Flörke, M., Hanasaki, N., Eisner, S., Fischer, G., Tramberend, S., Satoh, Y., van Vliet, M.T.H., Yillia, P., Ringler, C., Burek, P., Wiberg, D., 2016. Modeling global water use for the 21st century: the Water Futures and Solutions (WFaS) initiative and its approaches. *Geosci. Model Dev.* 9, 175–222. <https://doi.org/10.5194/gmd-9-175-2016>

# S-GNN: State-Dependent Graph Neural Networks for Functional Molecular Properties

Adem R.N. Aouichaoui, Alessandro Cogliati, Jens Abildskov, Gürkan Sin \*

*Process and Systems Engineering Centre (PROSYS), Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søtofts Plads, Building 228A, 2800 Kgs. Lyngby, Denmark*

[\\*gsi@kt.dtu.dk](mailto:*gsi@kt.dtu.dk)

## Abstract

Property models are an integral part of many chemical engineering applications and have been the subject of a lot of interest, especially with recent advancements in deep learning such as graph neural networks. Despite being of major importance, little effort has been dedicated to functional properties where the property dependency goes beyond the molecular structural information and depends on the state variables such as temperature and pressure. In this work, we demonstrate a flexible framework to extend graph neural networks to account for such use cases. A total of 13 different temperature-dependent properties were modeled covering enthalpy of vaporization, various heat capacities, densities, and thermal conductivities as well as surface tension and vapor pressure. While many were successfully modeled with some reaching an average absolute relative error below 6%, some still require further attention such as surface tension and thermal conductivity to achieve good accuracy.

**Keywords:** Graph neural networks, Thermophysical properties, Molecules, QSPR.

## 1. Introduction

In-silico evaluation of the various thermal, physical, safety, and transport properties of molecules is an integral element in many process simulation and product design software. These properties are usually the product or derived from experimental measurements and as such predictive models are necessary to perform in-silico evaluation of the properties. Quantitative structure-property relationships (QSPRs) are mathematical models that can provide such a service by leveraging the structural information of the molecule in inferring its property (Aouichaoui et al., 2023, 2022). Molecular properties of interest in the chemical engineering field can generally be divided into primary, secondary, and functional properties (Kontogeorgis and Gani, 2004). The primary properties are single-value properties that can be determined solely from the structural information, while secondary properties cannot be explicitly calculated uniquely from the structural information but rather as a function of other properties. An example of primary properties is the critical point properties as well as the normal boiling point, while the secondary properties could be the acentric factor or the solubility parameters. Both types of properties have long been the subject of various QSPR modeling through e.g. the use of group-contribution methods (Hukkerikar et al., 2012) as well as more advanced deep learning models (Aouichaoui et al., 2023, 2022). While much focus has been on taking advantage of the recent development in the field of deep learning (DL) to develop increasingly better QSPR models for primary properties and secondary properties, little

attention has been given to functional properties. Functional properties on the structural information as well as a set of intensive variables such as temperature, pressure, or composition. Such properties include saturated liquid density, vapor pressure, heat capacities, and thermal conductivity. Predictive models that can infer these properties are an integral part of any process simulation software as many unit operations are operated at varying temperatures and pressures. These properties are usually predicted through correlations that include compound-specific constants and as a function of other properties of the compounds. An example of the correlation of saturated vapor pressure ( $P^{sat}$ ) using the Antoine equation as seen in Eq.(1), where T is the temperature and A, B, and C are compound-specific constants determined through regression. A further example is a correlation for the liquid density ( $\rho^{sat}$ ) using the modified Racket equation as seen in Eq.(2), where R is the universal gas constant,  $T_c$  and  $P_c$  are the critical temperature and pressure respectively,  $T_r$  is the reduced temperature and  $Z_{RA}$  is a compound specific constant determined through regression.

$$\log_{10}P^{sat} = A - \frac{B}{T + C} \quad (1)$$

$$\frac{1}{\rho^{sat}} = \frac{R T_c}{P_c} Z_{RA}^{(1+(1-T_r)^{2/7})} \quad (2)$$

The two examples show that the property is dependent on other inputs that are either determined experimentally or through other correlations (e.g. the use of other property models to evaluate the critical properties). One major drawback of this is the fact that uncertainties become untraceable due to many input parameters. A further issue is when the property evaluation requires other QSPR models. As such, we advocate the leveraging end-to-end learning framework from the field of deep learning and extending it to build property models for functional properties that solely rely on the molecular structure and the intensive state variable such as the temperature. The developed framework is denoted state-dependent graph neural networks (S-GNN) and is elaborated on in the following.

## 2. S-GNN

Graph neural networks (GNNs) provide a unified approach that combines the feature extraction process and the regression process (Aouichaoui et al., 2022). GNNs takes as input a graph representation of the molecule where nodes correspond to atoms and edges correspond to the bonds. The node and edge are associated with feature vectors with information about the atoms and bond e.g. atom type, number of hydrogen attached, and bond type. Using a series of operations (also called convolutional layers) involving matrix multiplication, the graph representation is updated and converted into a numerical vector that is used as the molecular descriptor. This process is denoted message-passing and governs the feature extraction process and distinguishes GNN models from each other.

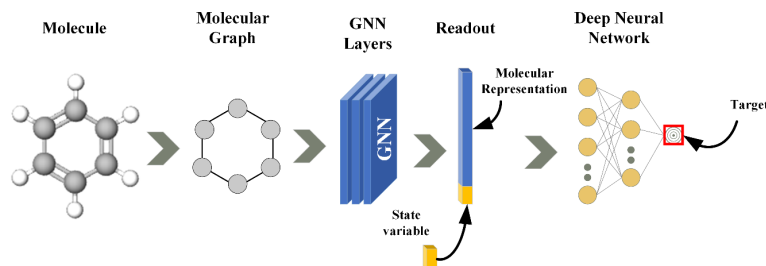


Figure 1: Schematic of the workflow of S-GNN

The S-GNN extends on the classical GNN models by considering the state variables as global variables that are concatenated to the learned molecular representation resulting from the GNN feature extraction step as shown in Figure 1. In this work, we demonstrate the concept using a GNN model known as *attentiveFP*, which in recent studies has demonstrated state-of-the-art performance across many properties (Aouichouai et al., 2022; Xiong et al., 2020). The model combines the attention mechanism and recurrent neural networks on two levels: node-level like any other GNN model and as part of the readout by operating on a graph where the molecule is considered a super node, while all atoms are linked to it. The node operations can be seen in Eq.(3) to Eq.(8). The initialization step aims to dimension the atom features ( $x_v^{atom}$ ) of node  $v$  to the desired hidden dimension through a learnable weight matrix  $W$  and a non-linear activation function. The edge features  $x_w^{atom}$  of neighboring atoms are combined with the node features to produce the latent representation of the neighboring atoms. The alignment aims to combine the node representation with that of its neighbors, which is used as input to a softmax to produce a weight coefficient. The coefficient is used as a weight for the importance/influence the hidden representation of the neighboring node should have when updating the hidden representation of a given node as seen in the context.

$$\text{Initialization} \quad h_v^0 = \text{ReLU}(W \cdot x_v^{atom}) \quad \text{for } t = 0 \quad (3)$$

$$n \quad h_w^0 = \text{ReLU}(W \cdot [x_w^{atom}, x_{v,w}^{bond}]) \quad \text{for } t = 0 \quad (4)$$

$$\text{Alignment} \quad \epsilon_{v,w}^{t+1} = \text{leakyReLU}(W \cdot [h_v^t, h_w^t]) \quad \text{for } t + 1 \leq T \quad (5)$$

$$\text{Weighting} \quad \alpha_{v,w}^{t+1} = \text{softmax}(\epsilon_{v,w}^{t+1}) \quad \text{for } t + 1 \leq T \quad (6)$$

$$\text{Context:} \quad C_v^{t+1} = \text{elu} \left( \sum_{w \in \text{Neighbor}(v)} \alpha_{v,w} \cdot W \cdot h_w^t \right) \quad (7)$$

$$\text{Update} \quad h_v^{t+1} = \text{GRU}(h_v^t, C_v^{t+1}) \quad \text{for } t + 1 \leq T \quad (8)$$

Similar reasoning is applied for the graph level embedding seen in Eq.(9) to Eq.(13), where  $L$  is the number of graph embedding layers and  $s$  denotes the supernode.

$$\text{Initialization} \quad h_s^0 = \sum_{v \in G} h_v^T \text{ for } l = 0 \quad (9)$$

$$\text{Alignment} \quad \epsilon_{s,v}^{l+1} = \text{leakyReLU}(W \cdot [h_s^l, h_v^T]) \text{ for } l + 1 \leq L \quad (10)$$

$$\text{Weighting} \quad \alpha_{s,v}^{l+1} = \text{softmax}(\epsilon_{s,v}^{l+1}) \text{ for } l + 1 \leq L \quad (11)$$

$$\text{Context} \quad C_s^{l+1} = \text{elu} \left( \sum_{v \in \text{Neighbor}(s)} \alpha_{s,v} \cdot W \cdot h_v^t \text{ for } l + 1 \leq L \right) \quad (12)$$

$$\text{Update} \quad h_s^l = \text{GRU}(h_s^l, C_s^{l+1}) \text{ for } l + 1 \leq L \quad (13)$$

### 3. Case study: modeling of 13 temperature-dependent properties

#### 3.1. Data

A total of 13 temperature-dependent properties were investigated: ideal gas heat capacity (ICP), liquid heat capacity (LCP), solid heat capacity (SCP), liquid and solid density (LDN, SDN), enthalpy of vaporization (HVP), surface tension (ST), liquid-solid and vapor thermal conductivity (LTC, STC, VTC), vapor pressure (VP) and liquid-vapor viscosity (LVS, VVS). All data were retrieved from the AIChE DIPPR 801 Database and only considered experimental data, smoothed or derived (Rowley et al., 2019). Chemical diversity analysis showing the different species (hydrocarbons, oxygenated, nitrogenated, chlorinated, fluorinated, brominated, iodinated, sulfonated, phosphorated, silicon-containing, or multi-functional) present can be seen in Table 1 where the number of total distinct species ( $N_{\text{spec}}$ ) and the total amount of observations ( $N_{\text{obs}}$ ) are also reported.

Table 1: Chemical diversity analysis of the dataset

	CH	O	N	Cl	F	Br	I	P	S	Si	Mul	$N_{\text{spec}}$	$N_{\text{obs}}$
<b>ICP</b>	30	23	4	3	9	1	1	0	7	0	10	88	1,529
<b>LCP</b>	219	285	59	26	16	13	6	0	44	0	73	741	13,831
<b>SCP</b>	167	163	33	17	12	5	0	0	38	0	52	487	16,567
<b>LDN</b>	412	626	107	63	43	17	7	0	70	0	194	1,539	25,485
<b>SDN</b>	57	129	18	11	2	4	3	0	5	0	50	279	718
<b>HVP</b>	397	601	105	61	36	18	7	0	66	0	183	1,474	24,279
<b>ST</b>	104	337	61	42	25	14	7	0	42	0	104	736	6,814
<b>LTC</b>	89	157	26	28	16	10	5	0	2	0	42	375	5,104
<b>STC</b>	2	9	1	0	0	0	0	0	0	0	4	16	132
<b>VTC</b>	47	34	9	6	18	2	1	0	1	0	17	135	1,550
<b>VP</b>	411	629	105	65	39	18	7	0	71	0	195	1,540	50,232
<b>LVS</b>	200	392	71	45	23	15	7	0	35	0	117	905	12,192
<b>VVS</b>	49	26	4	7	15	2	1	0	3	0	13	120	1,734

### 3.2. Model training

For the model training, the data were split randomly with 70% allocated for training and the remaining evenly split between validation and testing. No data scaling was performed. The model was trained over a maximum of 500 epochs using the ADAM optimizer, with mean absolute error as a loss function and the early stopping policy as a contingency against overfitting. The training-specific hyperparameters and their bounds consisted of: the batch size ([16, 1024]), the initial learning rate ( $[10^{-6}, 10^{-1}]$ ), the weight decay ([0, 0.25]), and the dropout ([0, 0.25]). The model-specific hyperparameters covered: the number of node embedding layers ([1,4]), the number of graph embedding layers ([1,4]), the length of the graph feature ([50,512]), and the number of DNN layers ([1,3]). The hyperparameter optimization was conducted using Bayesian optimization with a Gaussian process (GP) model with Matern32 kernel as covariance function, with 100 random exploration steps and 100 BO steps while the acquisition strategy chosen was as the upper confidence bound (UCB). The BO was performed using the “*Bayes\_Opt*” toolbox in python (Jiménez and Ginebra, 2017). The node and edge features used in this work are similar to those in previous studies (Aouichaoui et al., 2022; Xiong et al., 2020). The model can be accessed via: <https://github.com/gsi-lab/s-afp>. Since the data are proprietary, they are not made available, but the workings of the model are illustrated using another freely available dataset.

### 3.3. Results

The performance of the models in terms of the  $R^2$  and mean absolute percentage error and the mean absolute error and vis á vis the test set as well as the complete dataset can be seen in Table 2, an example of the parity plot for the liquid density and the saturated vapor pressure can be seen in Figure 2.

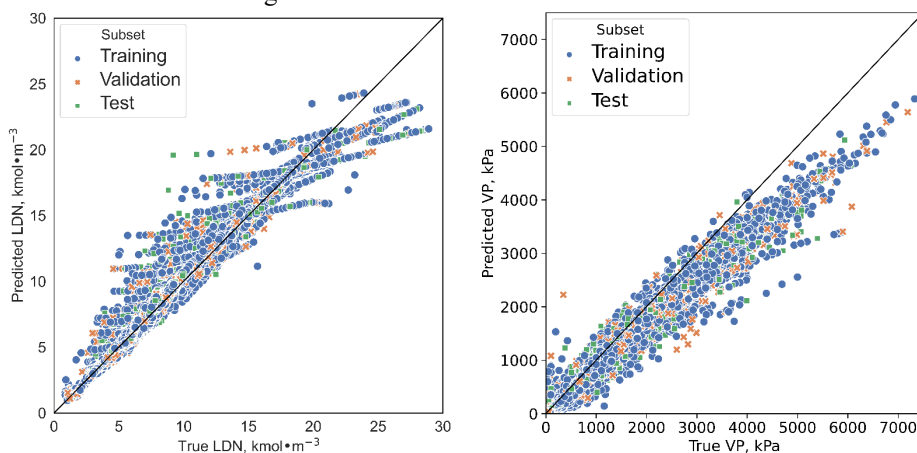


Figure 2: Parity plots for liquid density (LDN-left) and vapor pressure (VP-right)

Table 2: Model performance considering the test set and the overall data

Property	Unit	$R^2$		MAPE		MAE	
		test	all	test	all	test	all
ICP	[J·kmol <sup>-1</sup> ·k <sup>-1</sup> ]	0.99	0.99	1.2	1.1	1,958	1,678
LCP	[J·kmol <sup>-1</sup> ·k <sup>-1</sup> ]	0.99	0.99	3.6	3.3	7,810	7,260
SCP	[J·kmol <sup>-1</sup> ·k <sup>-1</sup> ]	0.98	0.98	110.4	67.9	10,360	10,370
LDN	[kmol·m <sup>-3</sup> ]	0.95	0.96	5.9	5.9	0.50	0.50
SDN	[kmol·m <sup>-3</sup> ]	0.98	0.98	4.3	4.7	0.49	0.46



HVP	[J·kmol <sup>-1</sup> ]	0.96	0.96	5.6	8.6	2.3e6	2.2e6
ST	[N·m <sup>-1</sup> ]	0.63	0.63	16.3*	15.8*	5e-3	5e-3
LTC	[W·m <sup>-1</sup> ·k <sup>-1</sup> ]	0.96	0.97	3.7	3.4	0.005	0.004
STC	[W·m <sup>-1</sup> ·k <sup>-1</sup> ]	0.12	0.44	30.1	50.4	0.06	0.10
VTC	[W·m <sup>-1</sup> ·k <sup>-1</sup> ]	0.98	0.88	12.1	12.4	3e-3	3e-3
VP	[kPa]	0.95	0.95	35.6	35.9*	58.6	60.31
LVS	[Pa·s]	N.A	0.35	204e2*	203e2*	1.2	2.7
VVS	[Pa·s]	0.90	0.91	9.9	5.7	1.1e-6	9.4e-7
*: Median absolute percentage error was used instead							

Many properties achieve good accuracy with 7 datasets reaching below 10% MAPE. In some cases, the use of the MAPE is not suitable as some datapoint are very small and close to zero and as such even small errors lead to a very large relative error which is misleading instead we suggest using the median absolute percentage error as it is robust to such outliers in the sample data. Especially the results obtained for the LDN, suggest the model results are similar to the state-of-the-art reported in previous studies using the modified Racket equation (Spencer and Adler, 1978). However, the absence of overall performance metrics makes it difficult to perform a comparison. However, noteworthy is also the fact that some datasets present a challenge to achieve a good accuracy performance, especially considering STC, LVS, and to some extent ST. One possible explanation for the STC dataset is the fact that the chemical diversity is very low with only 16 distinct molecules, which makes it difficult to learn any suitable representation. For the LVS, we observed some experimental data that were 7 times larger in order of magnitude, which might have been data that were tagged as derived. This highlights the requirement of addressing the data quality used during the model training as these are detrimental in allowing the model to extract knowledge it can correlate to the target property. For the ST, we observed a wide range of data with a notable bias (an intercept on the y-axis) when constructing the parity plot. This could suggest that the learning of the model has not achieved convergence enough to be able to reduce the variance observed. Another piece of evidence is the observation on the magnitude of the data points was between 0 and 0.06, which would require an even smaller learning rate than the one used during BO. One possible solution would be to perform data scaling as this will limit the spread of the data in addition to performing outlier detection to remove odd data present in the dataset.

#### 4. Conclusions

In this work, we provide an extension to graph neural networks that can generalize to any graph neural network and allow it to evaluate properties that depend on system state variables such as temperature and pressure in addition to the structural information of the molecule. Out of 13 properties evaluated, the majority of the properties were modeled with high accuracy ( $R^2 > 0.9$ ), while surface tension and thermal conductivity properties need further focus as they prove challenging to predictively model.

#### References

- Aouichaoui, A.R.N., Fan, F., Mansouri, S.S., Abildskov, J., Sin, G., 2023. Combining Group-Contribution Concept and Graph Neural Networks Toward Interpretable Molecular Property Models. *J Chem Inf Model*.

- Aouichaoui, A.R.N., Mansouri, S.S., Abildskov, J., Sin, G., 2022. Uncertainty estimation in deep learning-based property models: Graph neural networks applied to the critical properties. *AIChE Journal* 68.
- Hukkerikar, A.S., Sarup, B., ten Kate, A., Abildskov, J., Sin, G., Gani, R., 2012. Group-contribution + (GC +) based estimation of properties of pure components: Improved property estimation and uncertainty analysis. *Fluid Phase Equilib* 321, 25–43.
- Jiménez, J., Ginebra, J., 2017. pyGPGO: Bayesian Optimization for Python. *The Journal of Open Source Software* 2, 431–433.
- Kontogeorgis, G.M., Gani, R., 2004. *Computer Aided Property Estimation for Process and Product Design: Computers Aided Chemical Engineering*. Elsevier.
- Rowley, R.I., Wilding, W.V., Oscarson, J.L., Giles, N.F., 2019. *DIPPR Data Compilation of Pure Chemical Properties*.
- Spencer, C.F., Adler, S.B., 1978. A critical review of equations for predicting saturated liquid density. *J Chem Eng Data* 23, 82–89.
- Xiong, Z., Wang, D., Liu, X., Zhong, F., Wan, X., Li, X., Li, Z., Luo, X., Chen, K., Jiang, H., Zheng, M., 2020. Pushing the boundaries of molecular representation for drug discovery with the graph attention mechanism. *J Med Chem* 63, 8749–8760.



# Switching Multi-Objective Dynamic Optimization (MODO) for the Production of Value-Added Products

Juan C. Acosta-Pavas,<sup>a</sup> Carlos. E Robles-Rodríguez,<sup>a</sup> Jérôme Morchain,<sup>a</sup> David Camilo Corrales,<sup>b</sup> Claire Dumas,<sup>a</sup> Arnaud Cockx,<sup>a</sup> César A. Aceves-Lara,<sup>a</sup>

<sup>a</sup>*TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France*

<sup>b</sup>*Toulouse White Biotechnology (TWB), INRAE, 135 Avenue de Rangueil, 31077 Toulouse, France*

## Abstract

A Multi-Objective Dynamic Optimization (MODO) applied to the biomethanation process is developed. The MODO strategy was designed to find the trade-off between the maxima yield and productivity of methane and acetate modifying the inlet liquid and gas flow rates. First, a multi-objective optimization was applied to find the Pareto optimal set between productivity and yield of methane and acetate independently. Then, Pareto optimal points (POPs) were selected to develop five cases in dynamic optimization, which used a Model Predictive Control (MPC) approach. Cases 1-2 corresponded to the model simulation and the use of the POPs directly in simulation. Cases 3-4 addressed the POPs to maximize the Euclidean length for methane and acetate. Case 5 was performed to demonstrate a switch between these objectives. The ability of the MODO strategy to perform the switch was demonstrated. Additionally, the dynamic optimization allowed the reduction of the inlet gas flow rate to  $1.5m^3/d$ .

**Keywords:** Biomethanation, Multi-objective Optimization, Pareto Optimal Set, Dynamic Optimization, Process Control

## 1. Introduction

In Anaerobic Digestion (AD) organic matter is transformed by the synergistic work of different microorganisms into CH<sub>4</sub>, and CO<sub>2</sub> through four steps: hydrolysis, acidogenesis, acetogenesis, and methanogenesis. The produced biogas contains 50-75% of CH<sub>4</sub>, and 25-50 % of CO<sub>2</sub> (Iglesias et al., 2021). Biomethanation uses microorganisms to convert the CO<sub>2</sub> in the biogas from AD and syngas (a mix of H<sub>2</sub>, CO, and CO<sub>2</sub>) into CH<sub>4</sub>. This process, however, could also be used to produce value-added products such as acetate (Chaikitkaew et al., 2021). Acetate serves as a chemical platform in the textile, polymer, pharmaceutical, and food industries (Martín-Espejo et al., 2022). Nevertheless, controlling this type of process is an arduous task due to the multiple reactions and microorganisms involved. As a result, obtaining desired performances of yields or productivities at an industrial scale remains difficult, particularly when you need to optimize several of them simultaneously. The use of dynamic models plays a crucial role in the design of control strategies. For instance, Model Predictive Control (MPC) (Morales-Rodelo et al., 2020) is implemented to maintain or optimize several variables. MPC refers to control actions that optimize a criterion in a future behavior of the system, which is determined by the dynamic model (Camacho and Bordons, 2007). Multi-

Objective Optimization (MOO) implies optimizing problems where there is more than one objective to be optimized simultaneously. These objectives are usually conflictive (Vertovec et al., 2021).

This work aims at proposing a Multi-objective Dynamic Optimization (MODO) to maximize yield and productivity of the biomethanation process regarding two of the potential products: CH<sub>4</sub> and acetate. The proposed dynamic optimization approach used a MPC with two control variables corresponding to the inlets of liquid and gas flow rates. MPC uses the Pareto optimal sets, where each solution is considered as a Pareto Optimal Point (POP). The robustness of the MODO strategy is analyzed by performing a switch between the optimized values for CH<sub>4</sub> and acetate.

## 2. Multi-Objective Dynamic Optimization as Control Strategy

A MODO strategy was proposed in a previous work (Acosta-Pavas et al., 2022a) to determine the optimal values of the objectives. This strategy entails the following steps: **Step 1-Model definition.** Proposition of the dynamic model to represent the biological process.

**Step 2-Definition of the objective functions  $Y_1^*, \dots, Y_m^*$**  to be maximized/minimized in the MOO. A MOO can be formulated as a minimization problem, Eq (1).

$$\min_{x,u,p,t} \{Y_1^*(x, u, p), \dots, Y_m^*(x, u, p)\} \quad (1)$$

$$\text{Subject to } \begin{cases} dy/dt = \xi(x, u, p, t) \\ f_i(x, u, p, t) \leq 0 & i = 1, 2, \dots, n \\ g_i(x, u, p, t) = 0 & i = 1, 2, \dots, k \\ u^L \leq u \leq u^U \end{cases} \quad (2)$$

where  $Y_1^*, \dots, Y_m^*$  are the  $m$  objective functions,  $x$  the state variables,  $u$  the control variables and  $p$  the control variables. The dynamic model is represented by  $dy/dt$ ;  $f_i$  and  $g_i$  indicate inequality and equality constraints.  $u^L$  and  $u^U$  corresponds to the lower and upper bounds of the control variables.

**Step 3-Selection of the POP.** Determination of the Pareto optimal set  $Y_1^*, \dots, Y_m^*$  and selection of the POP to be used as the reference trajectory in the dynamic optimization.

**Step 4-Definition of the dynamic problem with a single weighted objective.** Formulation of an objective function considering the previously identified POP in terms of a MPC problem. The dynamic optimization determines the input variables that minimize the following objective function,

$$\min_u \left( \sum_{j=t}^{t+H_p} \left( \frac{|Y^* - y(t+j|t)|}{Y^*} \right)^2 + \sum_{j=t}^{t+H_c} W_u \Delta u(t+j|t)^2 \right) \quad (3)$$

where  $u$  is the vector of the control variables;  $H_p$  and  $H_c$  are the prediction and control horizons;  $y(t+j|t)$  is the output prediction calculated at time instant  $t+j$  using the information available at time instant  $t$ .  $Y^*$  is a reference trajectory that enables to reach the set point and is determined by the MOO. The term  $\Delta u(t+j|t)$  is the control move at time instant  $t+j$  calculated using information available at time instant  $t$ . The problem in Eq. (3) is also subject to Eq. (2).

### 3. Multi-objective Dynamic Optimization in Biomethanation Process

The main goal was to optimize yields and productivities of two value-added products, CH<sub>4</sub> and acetate, in order to demonstrate that a control strategy could attained larger values than the ones obtained from the model simulation. Two manipulated variables were proposed for the optimization: the inlet gas ( $q_{gas}^{in}$ ) and liquid ( $q_{liq}^{in}$ ) flow rates.

**Step 1.** The model used for simulation was proposed in a previous study by Acosta-Pavas et al., (2022b). This model was built upon experimental data from (Sun et al., 2021) and it considered the conversion of H<sub>2</sub> and CO to enhance CH<sub>4</sub> production. The experimental data from Sun et al. (2021) was carried out in a bubble column reactor with a working volume of  $3.75 \times 10^{-2} \text{ m}^3$  and a hydraulic retention time (*HRT*) of 20 days. The reactor operated at 37°C for 207 days. The organic loading rate (*OLR*) corresponded to  $0.53 \text{ kg}_{COD}/\text{m}^3 \text{ d}$  with  $q_{liq}^{in}$  of  $1.9 \times 10^{-3} \text{ m}^3/\text{d}$ . Gas addition was carried out in five stages (I - V) in which  $q_{gas}^{in}$  and the gas loading rate (*GLR*) were varied in time. The model (Acosta-Pavas et al., 2022b) can be rewritten as,

$$\frac{dS_{gas,i}}{dt} = \frac{q_{gas}^{in}}{V_{gas}} S_{gas,i}^{in} + N_i \left( \frac{V_{liq}}{V_{gas}} \right) - \frac{q_{gas}}{V_{gas}} S_{gas,i} \quad (4)$$

$$\frac{dS_{liq,j}}{dt} = \frac{q_{liq}^{in}}{V_{liq}} (S_{liq,j}^{in} - S_{liq,j}) + \sum_k Y_k f_{j,k} \mu_k - N_i \quad (5)$$

$$\frac{dX_k}{dt} = \frac{q_{liq}^{in}}{V_{liq}} (X_k^{in} - X_k) + Y_k \mu_k - \mu_{k,dec} \quad (6)$$

Sub-indices  $j = 1-8$  denote glucose, butyrate, propionate, acetate, H<sub>2</sub>, CH<sub>4</sub>, CO, and CO<sub>2</sub> in liquid phase; sub-indices  $i = \text{H}_2, \text{CH}_4, \text{CO}, \text{and CO}_2$  in the gas phase; while sub-indices  $k =$  biomass that degraded glucose, butyrate, propionate, acetate, H<sub>2</sub>, and CO in liquid phase.  $q_{gas}$  is the outlet gas flow rate,  $V_{liq}$  and  $V_{gas}$  are the liquid and gas volume, respectively.  $S_{liq,j}^{in}$ ,  $S_{gas,i}^{in}$ , and  $X_k^{in}$  are the inlet concentrations of components  $j$  in the liquid phase, the inlet concentration of components  $i$  in the gas phase, and the inlet concentration of biomass  $k$  in the liquid phase.  $Y_k$  is the yield of biomass  $k$ ,  $f_{j,k}$  are the stoichiometric coefficients,  $\mu_k$  and  $K_{k,dec}$  are the growth rate and the decay constant for biomass  $k$ . Finally,  $N_i$  is the mass transfer rate of component  $i$ .

**Step 2.** Two objectives were considered to optimize: the yields  $Y_{CH_4}$ ,  $Y_{ac}$  and productivities  $P_{CH_4}$ ,  $P_{ac}$ , for CH<sub>4</sub> and acetate, respectively. These are defined as,

$$Y_{CH_4} = \frac{q_{gas,CH_4}}{kgCOD_{added}}; P_{CH_4} = \frac{q_{gas,CH_4}}{V_{liq}} \quad (7)$$

$$Y_{ac} = \frac{S_{liq,ac}}{kgCOD_{added}}; P_{ac} = \frac{S_{liq,ac}}{V_{liq}} \quad (8)$$

The MOO for the maximization of  $Y_h$  and  $P_{hi}$  was proposed as,

$$\max_{\{q_{gas}^{in}, q_{liq}^{in}\}} (P_h, Y_h) \quad (9)$$

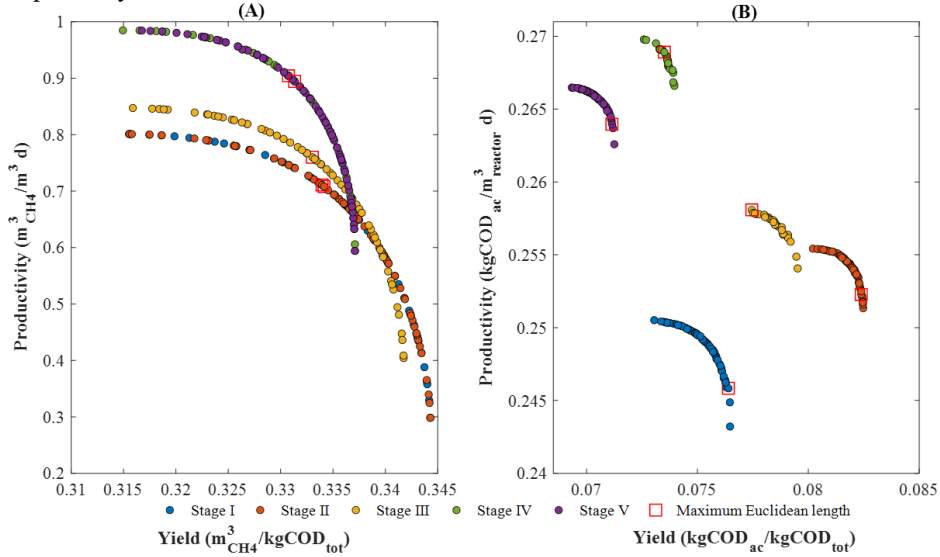
$$\text{Subject to } \begin{cases} \text{Equations 4 - 6} \\ Y_{CH_4} \leq Y_{CH_4}^{max} = 0.39 \text{ m}_{CH_4}^3/\text{kgCOD}_{added} \\ 1 \times 10^{-3} \leq q_{gas}^{in} \leq 10 \text{ m}^3/\text{d} \\ 1 \times 10^{-3} \leq q_{liq}^{in} \leq 0.1 \text{ m}^3/\text{d} \end{cases} \quad (10)$$

where  $h = \text{CH}_4, \text{acetate}$ ,  $Y_{CH_4}^{max}$  corresponds to  $0.39 \text{ m}_{CH_4}^3/\text{kgCOD}_{added}$  which represents the theoretical yield of CH<sub>4</sub> that could be obtained by kg of COD at 37°C. This

corresponds to  $0.35 m_{CH_4}^3 / kgCOD_{added}$  at standard temperature and pressure conditions (only for  $P_{CH_4}$ , and  $Y_{CH_4}$  maximization).

In this study, the simulations were run using an Intel® Core i7 8665U 2.11 GHz, 16 GB RAM computer. The *paretosearch* function from MATLAB® was used to obtain the Pareto optimal set for each stage.

**Step 3.** The Pareto fronts were computed for each stage. Concerning  $CH_4$ , 60 Pareto points were calculated for all stages. Regarding acetate, 60 Pareto points were calculated for stages I, II, and V, while 29 and 35 Pareto points were obtained for stages III and IV, respectively.



**Figure 1.** Pareto optimal sets for  $CH_4$  (A), and acetate (B) at each stage.

Figure 1 shows the Pareto optimal sets for  $CH_4$  and acetate at each stage. A progressive increase in  $P_{CH_4}$  was observed, while  $Y_{CH_4}$  decreased slightly.

$Y_{ac}$  and  $P_{ac}$  increased from stage I to II, then  $Y_{ac}$  decreased for the subsequent stages. However,  $P_{ac}$  continued to increase until stage IV to encounter a decrease in stage V.

At each stage, the POP were selected to maximize the Euclidean length ( $d_{max}$ ) for  $CH_4$  and for acetate (red squares in Fig. 1). For both cases,  $d_{max}$  was calculated as the distance from the origin, using a normalization as in Eq. (11).

$$d_{max} = \max \left( \sqrt{\left( \frac{Y_{CH_4}^* - \min(Y_{CH_4}^*)}{\max(Y_{CH_4}^*) - \min(Y_{CH_4}^*)} \right)^2} + \left( \frac{P_{CH_4}^* - \min(P_{CH_4}^*)}{\max(P_{CH_4}^*) - \min(P_{CH_4}^*)} \right)^2} \right) \quad (11)$$

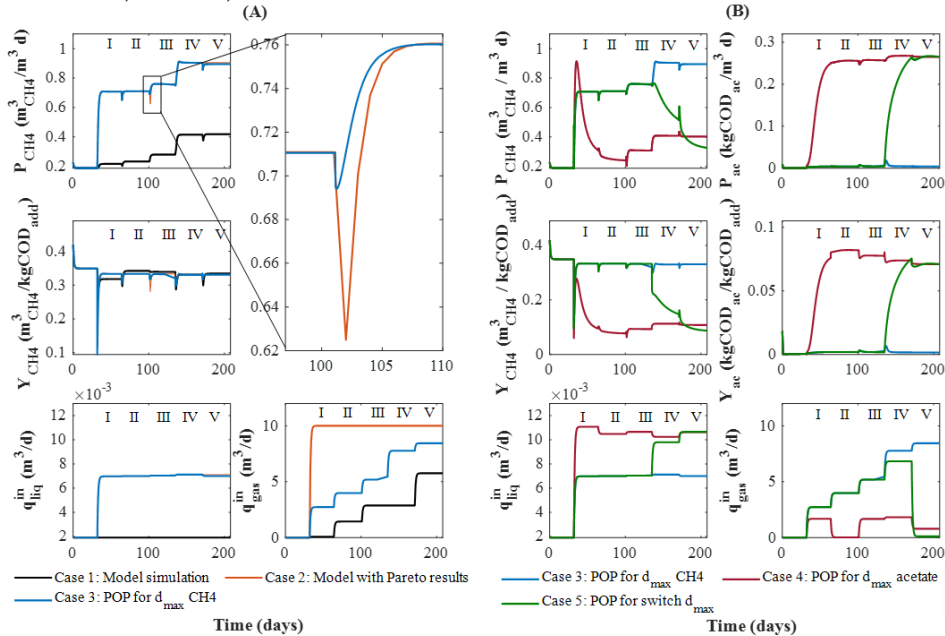
**Step 4.** The dynamic optimization for  $Y_h$  and  $P_h$  was proposed as follows ( $h = CH_4$ , acetate),

$$\min_{\{q_{gas}^in, q_{liq}^in\}} \left( \sum_{j=t}^{t+H_p} \left( \frac{|Y_h^* - Y_h(t)|}{Y_h^*} \right)^2 + \left( \frac{|P_h^* - P_h(t)|}{P_h^*} \right)^2 + \sum_{j=t}^{t+H_c} W_{u,1} \Delta q_{gas}^{in}(t)^2 + W_{u,2} \Delta q_{liq}^{in}(t)^2 \right) \quad (12)$$

Eq. (12) is subject to the constraints in Eq. (10).  $Y_i^*$  and  $P_i^*$  denote the POP values for yield and productivity computed by the MOO,  $\Delta q_{gas}^{in}(t)^2$  and  $\Delta q_{liq}^{in}(t)^2$  are the differences between the inlet gas and liquid flow rates, respectively, before and after each step of the dynamic optimization.  $W_{u,1}$  and  $W_{u,2}$  are the parameters that weight the importance of the control effort term in the optimization. In all cases, the initial values for both manipulated

variables were  $1 \times 10^{-3} \text{ m}^3/\text{d}$ .  $H_p$  and  $H_c$  were equal and equivalent to the time length of each stage. Optimization was performed with the *patternsearch* algorithm in MATLAB®.

**Step 5.** Five cases were studied to assess the dynamic optimization. Case 1 regarded model simulation (data without control). The other cases correspond to: the direct use of the POP in simulation (case 2); the inclusion of POP into a MODO for the maximization for  $d_{\max} \text{ CH}_4$  (case 3) and  $d_{\max} \text{ acetate}$  (case 4), and a switch strategy aiming at maximizing  $\text{CH}_4$  in stages I-III and then maximizing acetate in stages IV-V (case 5). The weights  $W_{u,1}$  and  $W_{u,2}$  (for cases 2-5) were manually adjusted to values of  $1 \times 10^{-6}$ .



**Figure 2.** Yields, productivities, injected and inlet gas flow rates in the MODO. Cases 1-5 (A); Case 6 (B).

Figure 2-A presents the results of cases 1 to 3. It is observed that  $Y_{\text{CH}_4}$  was maximized in the experimental part. On the other side,  $P_{\text{CH}_4}$  can be enhanced through the utilization of MODO.

The advantages to use the dynamic optimization were observed when cases 2 and 3 are compared. Both of them achieved similar results for  $P_{\text{CH}_4}$ . However, the transition between stages was smoother and faster in case 3, due to the dynamic part of the MODO (zoom in Figure 2-A). For instance, between stages II and III the time to reach 95% of the steady state decreased from 104 days in case 2 to 102 days in case 3, which could translate into less substrate consumption and thus a higher profit. Regarding  $Y_{\text{CH}_4}$ , it was similar in all cases, while  $q_{\text{liq}}^{\text{in}}$  varied from  $7.1 \times 10^{-3}$  to  $6.9 \times 10^{-3} \text{ m}^3/\text{d}$  at stage V. For  $q_{\text{gas}}^{\text{in}}$ , a value of  $10 \text{ m}^3/\text{d}$  was obtained for all stages in case 2. Nonetheless, this value was reduced in case 3 where  $q_{\text{gas}}^{\text{in}}$  ranged between 2.73 and  $8.44 \text{ m}^3/\text{d}$ .

Figure 2-B displays the cases 3 to 5. In stage V of case 3, a value of  $0.89 \text{ m}^3_{\text{CH}_4}/\text{m}^3 \text{ d}$  was obtained for  $P_{\text{CH}_4}$ , while  $0.33 \text{ m}^3_{\text{CH}_4}/\text{kgCOD}_{\text{added}}$  was obtained for  $Y_{\text{CH}_4}$ . In stage V of case 4, values of  $0.27 \text{ kgCOD}_{\text{ac}}/\text{m}^3 \text{ d}$  and  $0.07 \text{ kgCOD}_{\text{ac}}/\text{kgCOD}_{\text{added}}$  were achieved for  $P_{\text{ac}}$  and  $Y_{\text{ac}}$ , respectively.



In case 5, a switch was applied at the beginning of stage IV, where the reference values for  $P_{ac}$  and  $Y_{ac}$  were achieved at the end of the stage. Then, in stage V, values of  $0.27 \text{ kgCOD}_{ac}/\text{m}^3 \text{ d}$  and  $0.07 \text{ kgCOD}_{ac}/\text{kgCOD}_{added}$  were obtained for  $P_{ac}$  and  $Y_{ac}$ , respectively, while the control variables were calculated to maintain the values from cases 3 and 4. This case shows that the MODO strategy is robust when switching multi-objectives and attains the optima founded by the MOO. When comparing the performance of the MODO for CH<sub>4</sub> with the model simulation, (case 3 vs case 1), it is observed that  $P_{CH_4}$  increased between 3.26 to 2.13 times from stage I to V concerning case 1, while  $Y_{CH_4}$  was maintained similar.

#### 4. Conclusion

A MODO strategy was successfully applied over a biomethanation process based on the dynamic model ADM1\_ME. The feasibility of using Paretos to find the trade-off between objective functions, such as yields and productivities of CH<sub>4</sub> and acetate was demonstrated. The application of a dynamic optimization allowed an improvement in the response by reducing approximately 2 days the time in which the steady state was reached in the stage changes. Additionally, a reduction in gas flow rates of  $1.5 \text{ m}^3/\text{d}$  was achieved with the dynamic optimization. The robustness of the MODO strategy was demonstrated with a successful switch between the products of interest. These results showed that both control variables ( $q_{liq}^{in}$  and  $q_{gas}^{in}$ ) have a key role at each case and stage. Although these results show the feasibility of the proposed strategy, it is important to note that these are simulation results, and the adaptation of the microorganisms to the propose switching conditions might be different. However, this work showed the possibility to optimize in a smoother way the production of these two products and the change of objectives.

#### References

- J.C. Acosta-Pavas, C.E. Robles-Rodríguez, C.A. Méndez Suarez, J. Morchain, C. Dumas, A. Cockx, C.A. Aceves-Lara, 2022a. Dynamic Multi-Objective Optimization Applied to Biomethanation Process. *Chemical Engineering Transactions* 96, 319–324.
- J. C. Acosta-Pavas, J. Morchain, C. Dumas, V. Ngu, A. Cockx, C.A. Aceves-Lara, 2022b. Towards Anaerobic Digestion (ADM No. 1) Model's Extensions and Reductions with In-situ Gas Injection for Biomethane Production. *IFAC-PapersOnLine* 55, 635–640.
- E.F. Camacho, C. Bordons, 2007. *Model Predictive control*, Advanced Textbooks in Control and Signal Processing. Springer London, London.
- S. Chaikitkaew, J. Seengenyong, C. Mamimin, N.-K. Birkeland, A. Reungsang, S. O-Thong, 2021. Simultaneous biogas upgrading and acetic acid production by homoacetogens consortium enriched from peatland soil. *Bioresource Technology Reports* 15, 100701.
- R. Iglesias, R. Muñoz, M. Polanco, I. Díaz, A. Susmozas, A.D. Moreno, M. Guirado, N. Carreras, M. Ballesteros, 2021. Biogas from Anaerobic Digestion as an Energy Vector: Current Upgrading Development. *Energies* 14, 2742.
- J.L. Martín-Espejo, J. Gandara-Loe, J.A. Odriozola, T.R. Reina, L. Pastor-Pérez, 2022. Sustainable routes for acetic acid production: Traditional processes vs a low-carbon, biogas-based strategy. *Science of The Total Environment* 840, 156663.
- K. Morales-Rodelo, M. Francisco, H. Alvarez, P. Vega, S. Revollar, 2020. Collaborative control applied to bsm1 for wastewater treatment plants. *Processes* 8, 1–22.
- H. Sun, Z. Yang, Q. Zhao, M. Kurbonova, R. Zhang, G. Liu, W. Wang, 2021. Modification and extension of anaerobic digestion model No.1 (ADM1) for syngas biomethanation simulation: From lab-scale to pilot-scale. *Chemical Engineering Journal* 403, 126177.
- N. Vertovec, S. Ober-Blöbaum, K. Margellos, 2021. Multi-objective minimum time optimal control for low-thrust trajectory design, in: 2021 European Control Conference (ECC), Delft, Netherlands, 1975-1980.

# Dynamic modeling of a waste incinerator furnace in view of supervision

Lionel Sergent,<sup>a,b</sup> François Lesage,<sup>a</sup> Abderrazak M. Latifi,<sup>a</sup> Jean-Pierre Corriou,<sup>a</sup>  
Alexandre Grizeau<sup>b</sup>

<sup>a</sup>Laboratoire Réactions et Génie des Procédés, CNRS-ENSIC, Université de Lorraine, 1  
rue de Grandville, 54000 Nancy, France

<sup>b</sup>SUEZ RV France, 4 rue du Clausenhof, 67590 Schweighouse-sur-Moder, France  
[lionel.sergent@suez.com](mailto:lionel.sergent@suez.com)

## Abstract

Roughly 30% of municipal waste is incinerated in the European Union. Because of the heterogeneity of the waste and lack of local measurements, the industry relies on traditional control strategies. However, advanced modeling strategies have been used to gain insights on the design of such facilities. In this paper, a first-principles model of the waste combustion based on the simplification of past literature works is developed. Simulations are carried out using the walking column approach, reproducing the major trends observed in other works and on site. The dynamic responses resulting from the model will serve as the centrepiece of a global model to be used for supervision assistance.

**Keywords:** Waste Incineration, First-principles Model, CFD Simulation, Supervision.

## 1. Introduction

The first municipal incinerator appeared in London in 1870 as a response to fast urbanization and sanitation issues (Bertolini, 2000). To minimize pollution and cost, and to maximize the recovery of energy, the process was improved over time.

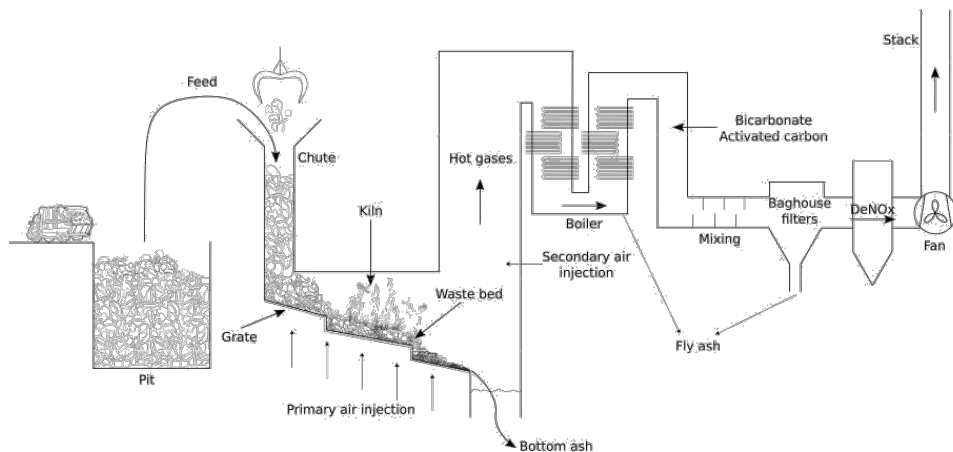


Figure 1: Schematic view of a typical incineration process.

Nowadays, the process can be roughly separated in three main stages: the feeding stage, hot stage, and flue-gas treatment stage, each one containing several units. A typical incineration process is given in Figure 1. For municipal waste, the most common type of kiln is the reciprocating grate, as it provides good mixing and a steady movement of the solids while being robust.

Regarding chemical engineering, the difficulties inherent to the incineration process are twofold. First, the waste heterogeneity and lack of control of the feeding imply strong uncertainties right after the process inlet. Second, there are few sensors in or near the kiln due to harsh mechanical and thermal conditions. There are many more sensors towards the process outlet. Therefore, it is relevant to develop a model that considers the process from inlet to outlet. Nevertheless, as the furnace is both the heart of the process and the host of the most complicated phenomena, it deserves specific attention. This paper presents a model of the behavior of a burning bed of waste to be integrated in a complete model of a real plant. Although greatly inspired by the literature, it differs by providing significant simplifications to a family of models only dedicated to the kiln.

## 2. Literature survey

### 2.1. Experiments

Because in industrial facilities only global fluid flows and temperatures can be measured, it is useful to rapidly present the 3 main experiments conducted in the literature as they give information on the local behavior.

1. The thermogravimetric analysis gives a classification of the fractions of waste known as the immediate analysis (humidity, volatile matter, char and ash). The ultimate analysis may be obtained to get a breakdown by elements. Thermal analysis also provides data for pyrolysis kinetics. Despite a high number of experiments, it must be noted that the resulting kinetics laws are quite unreliable (White et al., 2011).
2. The so-called pipe-bowl experiments (Menard, 2003) are similar to cylindrical wood stoves. They feature an array of thermocouples and gas analyzers that allow to follow locally the combustion of medium size samples. This experiment unravels the propagation of reaction fronts through the combustible.
3. Cold grate experiments study how solid particles travel and mix along an industrial grate. In the absence of combustion, it is much easier to track objects by means of visual hints or RFID tags (Džiugys et al., 2006).

### 2.2. Modeling approaches

Few works present models geared towards the control of incinerators. Van Kessel (2003) uses a 0D CSTR model and Magnanelli et al. (2020) present a 1D horizontal model for that purpose. However, these models cannot explain the behavior observed during the 3 mentioned experiments. Therefore, other models were considered, generally developed to gain insight on the design and pollutants behavior. Lagrangian approaches offer detailed simulations at the cost of tedious mechanical descriptions and high computation time (Brosch et al., 2014). Eulerian approaches are more common, generally solved in 2D (Menard, 2003). Neglecting horizontal gradients leads to a method introduced by Goh et al. (2001) as the "walking column" approach, which is mostly retained for the present work in order to alleviate computation time.

### 3. Proposed model

#### 3.1. Model features

The waste present on the grate is subdivided in  $N$  columns which are modeled as a porous medium. It evolves dynamically while moving towards the bottom-ash pit at a constant speed. The porous medium is represented by a reduced set of 13 vertically distributed state variables: 4 solid mass fractions (humidity, volatile matter, residual carbon, inert matter), 7 gas mass fractions ( $N_2, O_2, H_2O, CH_4, CO, CO_2, H_2$ ), the solid and gas temperatures. It can be noted that in the literature, generally more state variables are necessary.

The evolution of state variables is determined by partial differential equations derived from the general transport equation, expressed as

$$\partial_t(\rho\phi) + \partial_z(\rho v\phi) = \partial_z(\Gamma\partial_z\phi) + S \quad (1)$$

where  $\rho$  is the apparent density,  $\phi$  a conservative variable corresponding to a state variable,  $\Gamma$  the diffusion coefficient,  $v$  the vertical velocity, and  $S$  the source term.

The bottom boundary condition is imposed by the primary air injection, which varies along the grate, while the top boundary condition is given by the pressure generated by the induced draft fan.

The gas phase properties are estimated by correlations (Poling et al., 2001). The properties of the solid phase are generally constant or linear with temperature, and the values are the standard estimates used in the industry.

Equation (1) written for the total mass of gas is rearranged into a pressure equation by assuming Darcy's and ideal gas laws. Its dynamics are fast and neglecting the accumulation term had no influence on the resulting pressure field. Removing the accumulation term allows to analytically solve the flow field, offering a significant simplification compared to the more general momentum equation widely used. It also makes the numerical solution more stable.

#### 3.2. Main phenomena involved

The source terms are constituted by a combination of various rates describing the main phenomena involved in the process:

- The drying rate is calculated through a classical Sherwood correlation. The driving force is the difference between the partial and saturation pressures of water.
- The pyrolysis rate is assumed to be a first order reaction. The pre-exponential factor and energy activation have been fitted with thermogravimetric experiments.
- The residual carbon burning rate is modeled by a first-order reaction and transport of oxygen through the particle boundary layer.
- The rates of light gases ( $CH_4, CO, H_2$ ) combustion are calculated with kinetic laws, which are all monomial in the gas concentrations, with mixing and channelling limitations.
- The radiative exchange rate with the flame and the kiln structure is modeled by Bouguer's law, while the radiative exchange between particles is treated using the radiative diffusion approximation.
- The convective heat exchange rate between the solid and gas phases is calculated with a classical Nusselt correlation.

The proportions of gases produced by pyrolysis are precalculated by a routine which ensures that the measured lower heating value, immediate analysis and ultimate analysis

of the waste are all respected. The atomic and energy balances used motivate the choice of exactly 7 gas mass fractions, as a smaller number would not allow the respect of the mentioned constraints.

### 3.3. Numerical solution

Simulations were carried out using MATLAB environment. Discretization is based on the finite volume method and the FVTool library (Eftekhari and Schüller, 2022). The advection term is discretized with the upwind scheme. The used grid is denser near the boundaries. The time integration is done with an adaptive implicit Euler scheme. A dynamic simulation with fixed input parameters to reach a steady state typically takes a few minutes on a personal computer, although the calculation time is highly sensitive to several parameters.

Because the vertical velocity of the solid is neglected, the volume reduction effect does not appear in transport equation (1). Instead, the volume of each cell is modified at each time step. This calculation reflects the fact that drying and devolatilization mostly leave intact pores in the particles, while char burning consumes the remaining structure and rapidly decreases the size of particles, resulting in a collapse of the solid bed.

## 4. Results and discussion

### 4.1. 1D simulation of a single column

Figure 2 presents time profiles of the gas outputs of a single column using a one-dimensional model. To make this simulation comparable with most pipe-bowl experiments found in literature, the air supply is kept constant at under-stoichiometric conditions, while the height of the waste bed is 30 cm.

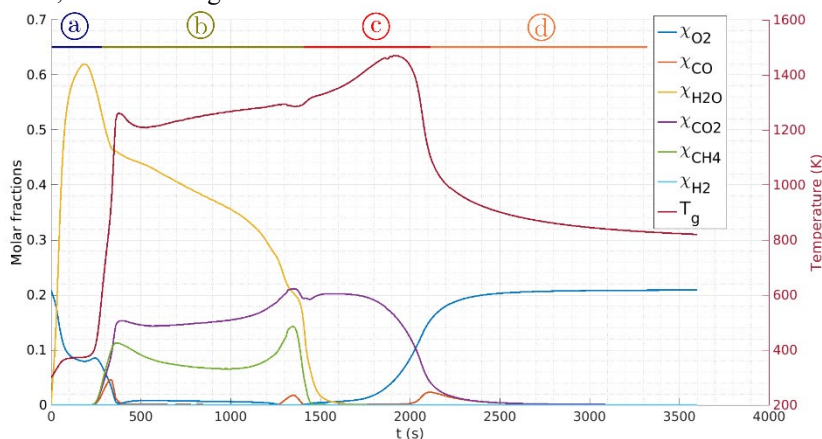


Figure 2: Evolution of the gas output of a 1D column model with constant air supply.

In phase (a) (Figure 2), the radiative flux heats up and dries the top of the bed, characterized by a short temperature plateau around 373 K. In phase b, the pyrolysis starts. Light gases are emitted and quickly burned, resulting in a stiff temperature increase and a near-depletion of oxygen. Some of the heat generated is used to dry the lower layers that will in turn pyrolyze and burn. When this reaction front reaches the bottom, the remaining carbon is burned out in phase c which results in a temperature peak. In phase d, the combustion is completed, and a thermal equilibrium is reached between radiative heating and convective cooling by the primary air injection.

#### 4.2. 2D simulation of the entire bed

For a waste bed traveling on a grate, the state variables can be represented in maps such as the one given in Figure 3. The phases that were observed against time on Figure 2 can now be roughly observed with respect to position, despite large differences in primary air and bed height. After half a meter of pure drying, the pyrolysis starts early in zone 1, but not enough air is present to burn the volatile gases yet. Because the conditions are over-stoichiometric in zone 2, pyrolysis, light gas combustion and carbon combustion happen mostly simultaneously.

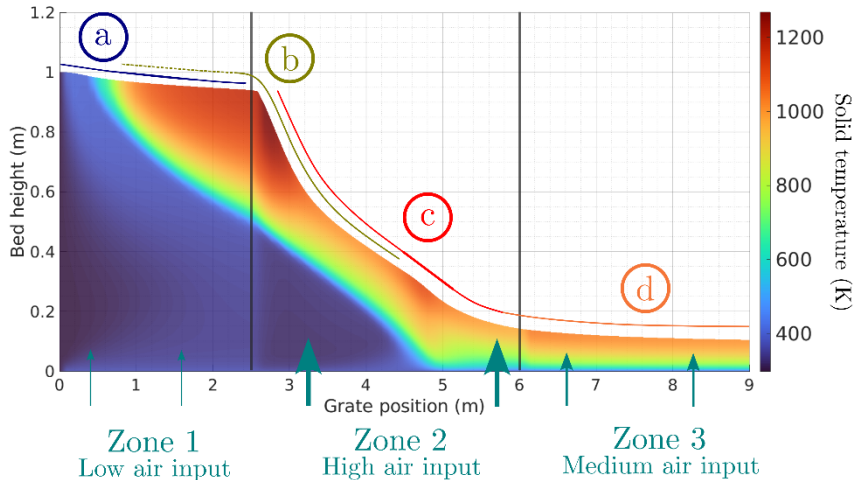


Figure 3: Solid temperature (K) distribution of a simulated waste bed. The grate and primary air injection parameters are taken from the incinerator of Schweighouse-sur-Moder (France) operated by the SUEZ company. The computation grid is 100 by 100.

#### 4.3. Waste bed response to a step change

Figure 4 displays the waste bed response to a step decrease of primary air injection.

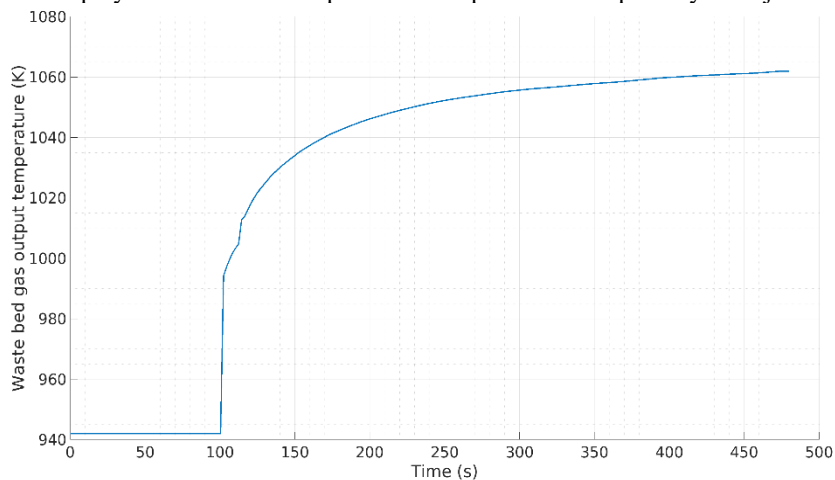


Figure 4: Response of the kiln to a primary air step decrease of 30% at time  $t=100$ s.

The waste bed gas output temperature is computed by integrating the gas enthalpy and mass flows at the bed surface before any secondary combustion can happen. Note that because the walking columns algorithm produces nearly cyclic computation artefacts, a moving average has been applied to obtain a smoother curve (Figure 4). Since the conditions are over-stoichiometric, a decrease in primary air results in an immediate increase in temperature as less dilution occurs. Then, the temperature continues to increase closely to a 1st order response. This is mainly due to the solid thermal inertia. If the conditions were under-stoichiometric, a decrease in primary air would result in a temperature decrease due to incomplete combustion.

## 5. Conclusion and perspectives

A simplified first-principles model of the behavior of a burning bed of waste in an incinerator furnace has been developed and used to carry out simulations. Major trends observed on site and in the literature were successfully reproduced. For the first time, this type of model is used to study the dynamic responses of a burning waste bed. However, the results appear to be very sensitive to the model parameters, including the solid representation (particles, porosity), kinetic rates and heat transfer. Improvements could be made on the numerical resolution, such as an adaptive grid to follow the reaction front, the use of Total Variation Diminishing advection schemes and higher order integrators.

The next step consists in modeling the freeboard which will allow partial validation of the kiln combustion model with plant data, thanks to the secondary combustion chamber temperature and O<sub>2</sub> sensors. Then, the boiler, with its thermal inertia, will be modeled and offer steam production responses. Once the flue gas treatment stage is modeled, comparison with online gas analyzers will be possible. The final process model will be used as a basis for the prototype of a supervision tool whose exact form will depend on the characteristics of the global dynamic model, including its ability to treat uncertainty.

## References

- G. Bertolini, 2000, Aperçu historique (jusqu'en 1950) du développement de l'incinération des ordures ménagères dans le monde, *Déchets, sciences et techniques*, 20, 3–8.
- B. Brosch, V. Scherer, S. Wirtz, 2014, Simulation of municipal solid waste incineration in grate firing systems with a particle based novel Discrete Element Method, *VGB Powertech*, 75–83.
- A. Džiugys, B. Peters, H. Hunsinger, L. Krebs, 2006, Evaluation of the Residence Time of a Moving Fuel Bed on a Forward Acting Grate, *Granular Matter*, 8, 125–135.
- A.A. Eftekhari, K. Schüller, 2022, FVTool: Finite volume toolbox for Matlab.
- Y.R. Goh, Y.B. Yang, R. Zakaria, R.G. Siddall, V. Nasserzadeh, J. Swithenbank, 2001, Development of an Incinerator Bed Model for Municipal Solid Waste Incineration, *Combustion Science and Technology*, 162, 1, 37–58.
- E. Magnanelli, O.L. Tranãs, P. Carlsson, J. Mosby, M. Becidan, 2020, Dynamic modeling of municipal solid waste incineration, *Energy*, 209, 118426.
- Y. Menard, 2003, Modélisation de l'incinération sur grille d'ordures ménagères et approche thermodynamique du comportement des métaux lourds, PhD Thesis, Institut National Polytechnique de Lorraine, Nancy, France.
- B.E. Poling, J.M. Prausnitz, J.P. O'Connell, 2001, *The properties of gases and liquids*, 5th ed, McGraw-Hill, New York.
- R. van Kessel, 2003, Stochastic disturbances and dynamics of thermal processes, PhD Thesis, Technische Universiteit Eindhoven, Netherlands.
- J.E. White, W.J. Catallo, B.L. Legendre, 2011, Biomass pyrolysis kinetics: A comparative critical review with relevant agricultural residue case studies, *Journal of Analytical and Applied Pyrolysis*, 91, 1, 1–33.

# Reinforcement learning combined with model predictive control to optimally operate a flash separation unit

Dean Brandner<sup>a\*</sup>, Torben Talis<sup>b\*</sup>, Erik Esche<sup>b</sup>, Jens-Uwe Repke<sup>b</sup>, Sergio Lucia<sup>a</sup>

<sup>a</sup>*Technische Universität Dortmund, Laboratory of Process Automation Systems, Emil-Figge-Straße 70, Building G2, Dortmund 44227, Germany*

<sup>b</sup>*Technische Universität Berlin, Process Dynamics and Operations Group, Str. des 17. Juni 135, Berlin 10623, Germany*

\**dean.brandner@tu-dortmund.de, t.talis@tu-dortmund.de*

## Abstract

Model predictive control (MPC) and reinforcement learning (RL) are two powerful optimal control methods. However, the performance of MPC depends mainly on the accuracy of the underlying model and the prediction horizon. Classic RL needs an excessive amount of data and cannot consider constraints explicitly. This work combines both approaches and uses Q-learning to improve the closed-loop performance of a parameterized MPC structure with a surrogate model and a short prediction horizon. The parameterized MPC structure provides a suitable starting point for RL training, which keeps the required data in a reasonable amount. Moreover, constraints are considered explicitly. The solution can be obtained in real-time due to the surrogate model and the short prediction horizon. The method is applied for control of a flash separation unit and compared to a MPC structure that uses a rigorous model and a large prediction horizon.

**Keywords:** Reinforcement Learning, Model Predictive Control, Flash Separation Unit

## 1. Introduction

Model predictive control (MPC) is a powerful control concept as it can deal with process constraints and nonlinear multivariable systems. However, the control performance mainly depends on the accuracy of the system model as well as its prediction horizon. Using a rigorous first principles model and a long prediction horizon may lead to limited real-time capability due to large computational effort. One possibility to reduce the computational effort is to use surrogate models that are less complex than the original first principles ones. In addition, the prediction horizon can be kept short to reduce the number of decision variables of the optimal control problem. However, both strategies will lead to worse closed-loop performance.

Due to the digitalization of the process industry, an increasing amount of data becomes available. Data can improve process operation in different manners, for example with data-driven surrogate models or via reinforcement learning (RL). However, traditional RL approaches require very large amounts of data, and they cannot deal easily with process constraints.

In this work, we follow the success of recent approaches and utilize a parameterized MPC structure within an RL scheme to improve the closed loop performance of an MPC controller with an imperfect surrogate model and to consider process constraints



explicitly. We first show how the parameterized MPC controller can be used to approximate the action-value function of the optimal control problem and illustrate afterwards the results of the application on a flash unit.

## 2. Background on surrogate modelling for dynamic systems

Constrained dynamic systems are usually described by differential algebraic equation (DAE) systems. They consist of both ordinary differential equations and algebraic equations and can be written in semi-explicit form with state variables  $x = (y, z)$ , the control variables  $u$ , the parameters  $p$ , and time  $t$ :

$$\dot{y} = f(y, z, u, p, t) \quad (1)$$

$$0 = g(y, z, u, p, t). \quad (2)$$

High index problems frequently arise in chemical engineering, e.g., by assuming negligible vapor holdup. These systems cannot be integrated by explicit methods and are usually rather large DAE systems, which are slow and difficult to solve. This is limiting the applicability of optimization-based operation and control, such as MPC. Therefore, data-driven surrogate models are increasingly introduced.

There are many different approaches to model dynamic systems using data, e.g., Gaussian processes (Kocijan, 2016), different recurrent neural network architectures (Esche et al., 2022), and many more. In case a system is Markovian, a feedforward neural network is sufficient and can predict the state of the system at the next timestep, based on the current state and controls (Himmelblau, 2008).

## 3. Q-Learning for parameterized MPC structures

Reinforcement learning (RL) is a tool to solve Markov decision process (MDP) problems via the interaction of an agent (controller) and an environment (plant) (Sutton and Barto, 2018). The agent computes an action  $\mathbf{a}_k$  based on the current state of the environment  $\mathbf{s}_k$ . The action is applied to the system, which replies with the next state  $\mathbf{s}'_k$  and a reward  $r_k$  (stage cost). The combination gives an element of a Markov chain  $\mathcal{M} = \{(\mathbf{s}_k, \mathbf{a}_k, r_k, \mathbf{s}'_k)\}_{k=0}^N$ . The state of the environment  $\mathbf{s}_k$  does not have to be the system state  $\mathbf{x}_k$  only, for example the state of the environment can also be made up of past control actions  $\mathbf{u}_{k-1}$ . The agent's action is computed according to its policy  $\pi(\mathbf{s})$ . The policy can be learned implicitly from the optimization of the action-value function.

Q-learning focusses on finding the optimal action-value function  $Q^*(\mathbf{s}, \mathbf{a})$  of a MDP. The action-value function is a measure of how good an action  $\mathbf{a}$  in a given state  $\mathbf{s}$  is, when following a certain policy  $\pi$

$$Q(\mathbf{s}, \mathbf{a}) = r(\mathbf{s}, \mathbf{a}) + \gamma \cdot Q^\pi(\mathbf{s}', \pi(\mathbf{s}')) \quad (3)$$

where  $\gamma \in [0; 1]$  is a so-called discount factor. The optimal control policy can then be derived by optimization over the optimal action-value function. For control, the optimal policy  $\pi^*$  minimizes the expected cost. In the solution process for continuous control tasks, the action-value function is approximated using a function approximator  $Q_\theta$ .

A common choice to approximate that function is to choose deep neural networks (Mnih et al., 2015). However, they may suffer from lack of structure, therefore requiring an excessive amount of data. One way to reduce the amount of data is to use a more

structured function approximator like a parameterized MPC structure (Gros and Zanon, 2020). MPC takes an underlying system model  $f_d$  and predicts future states with it to derive an optimal control input. A parameterized MPC formulation can be written as:

$$\begin{aligned}
 Q_{\theta}(\mathbf{s}, \mathbf{a}) = \min_{\substack{\mathbf{x}_{[0:N]} \\ \mathbf{u}_{[0:N-1]} \\ \boldsymbol{\sigma}_{[0:N]}}} & \gamma^N (V_{f,\theta}(\mathbf{x}_N) + \mathbf{w}_f^T \boldsymbol{\sigma}_N) + \sum_{i=0}^{N-1} \gamma^i (\ell_{\theta}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{w}^T \boldsymbol{\sigma}_k) \\
 \text{s. t.} & \quad \mathbf{x}_{k+1} = f_{d,\theta}(\mathbf{x}_k, \mathbf{u}_k), \quad \mathbf{x}_0 = \mathbf{s}, \quad \mathbf{u}_0 = \mathbf{a}. \\
 & \quad g(\mathbf{u}_k) \leq 0, \quad \boldsymbol{\sigma}_k \geq 0 \\
 & \quad h_{\theta}(\mathbf{x}_k, \mathbf{u}_k) \leq \boldsymbol{\sigma}_k, \quad h_{f,\theta}(\mathbf{x}_N) \leq \boldsymbol{\sigma}_N.
 \end{aligned} \tag{4}$$

The parameterization is performed using the parameter vector  $\boldsymbol{\theta} \in \mathbb{R}^{n_{\theta}}$ , which can parameterize the objective function, the system model, and the constraints. For classic MPC, the discount factor is chosen as  $\gamma = 1$ . The parameterized MPC structure consists of the objective function, which is compound of the stage cost  $\ell_{\theta}$ , the terminal cost  $V_{f,\theta}$ , and penalties for the slack variables  $\boldsymbol{\sigma}$ , but also input and state constraints. The state constraints are considered as soft constraints.

Classic Q-learning updates may cause instabilities (Sutton and Barto, 2018). Mnih et al. introduce robustness via a replay buffer and target function  $Q_{\hat{\theta}}$  (Mnih et al., 2015). The replay buffer stores data from the past. In each update step, past samples are taken to perform a batchwise update. The target function is updated slower than the actual function. The target parameters are updated as  $\hat{\boldsymbol{\theta}} \leftarrow (1 - \tau)\hat{\boldsymbol{\theta}} + \tau\boldsymbol{\theta}$  with  $0 \leq \tau \leq 1$ . The Q-learning update requires the gradient of the Q-function. For the parametric MPC, the gradient of Q-function equals the gradient of the Lagrangian  $\mathcal{L}$  of the underlying optimization problem evaluated at the optimal points (Büsken and Maurer, 2001):

$$\nabla_{\boldsymbol{\theta}} Q_{\theta}(\mathbf{s}, \mathbf{a}) = \nabla_{\boldsymbol{\theta}} \mathcal{L}(\mathbf{x}_{[0:N]}^*, \mathbf{u}_{[0:N-1]}^*, \boldsymbol{\lambda}^*, \mathbf{v}^* | \mathbf{s}, \mathbf{a}). \tag{5}$$

The main idea of our work is to adapt the MPC structure (4) via the parameters  $\boldsymbol{\theta}$ , using Q-learning to counteract the imperfect surrogate model  $f_{d,\theta}$  and short prediction horizon.

## 4. MPC parameter optimization with RL on a flash simulation

### 4.1. Dynamic pressure-driven flash model

In order to build a realistic environment, a first principles model of a dynamic flash unit is created, see Figure 1. It describes the binary mixture of ethanol and water. The vapor liquid equilibrium is described assuming an ideal gas, but non-ideal liquid phase using Wilson's activity coefficient. Pressure and level are controlled via distillate and bottom stream respectively.

A surrogate model of the first principles model will be used within the parameterized MPC. For this, the rigorous model is sampled to generate data, which is used for training of a dynamic surrogate model. The control variables and their bounds for the model are the supplied heat flux  $\dot{Q}$  (0-30 kW), feed flow  $F^F$  (1-2 mol s<sup>-1</sup>), and feed composition  $x^F$  (0.01-0.99). The feed temperature is constant at 365 K.

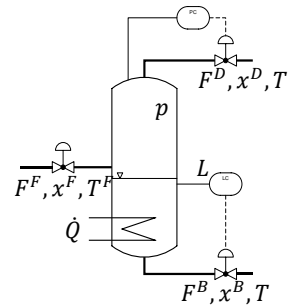


Figure 1. Flowsheet of the flash unit with 4 inputs:  $F^F$ ,  $x^F$ ,  $T^F$ ,  $\dot{Q}$  and 7 outputs  $p$ ,  $T$ ,  $L$ ,  $F^D$ ,  $x^D$ ,  $F^B$ ,  $x^B$ .

To create training data, the former two are excited with an amplitude modulated pseudo-random binary sequence (APRBS) - which can be interpreted as a sequence of step signals with different holding times with a frequency of  $1 \text{ s}^{-1}$ . The feed composition is held constant for 50 s and then increased by 0.01. A second trajectory is created to later test the surrogate model, which uses a lower frequency for the APRBS while allowing the same number of steps.

The 50'000 samples of the first trajectory are shuffled and split into (50 %, 75 %, or 95 %) training data, while the rest is used for validation. These points are used to train three similar feed forward neural network with 10 nodes in one hidden layer for 10 epochs with the Adam optimizer (Kingma and Ba, 2014). The hyperbolic tangent was used for the activation function. The MSE on the validation data is  $5.7 \cdot 10^{-5}$ ,  $5.0 \cdot 10^{-5}$  and  $1.6 \cdot 10^{-5}$  for the three different training scenarios. The second trajectory is then used for testing and gets predicted with an MSE of  $3.1 \cdot 10^{-4}$ ,  $5.8 \cdot 10^{-5}$ , and  $1.5 \cdot 10^{-5}$ , respectively. In the following section the first model (model 1) is used to show that the presented approach works well, even for surrogate models with a rather large plant model mismatch.

#### 4.2. Design of the RL-environment based on the rigorous dynamic flash model

The RL environment consists of the rigorous first principles model. The current state of the environment  $\mathbf{s}_k$  is composite of all system states  $\mathbf{x}_k = (p, T, L, F^D, y^D, F^B, x^B)^T_k$  at time  $k$  and the previously applied control action  $\mathbf{u}_{k-1} = (F^F, Q)^T_{k-1}$  at time  $k-1$ .  $x^F$  and  $T^F$  are kept constant. Hence, it is built as  $\mathbf{s}_k = (\mathbf{x}_k^T, \mathbf{u}_{k-1}^T)^T$ . The control action applied to the environment is  $\mathbf{a}_k = \mathbf{u}_k$ . The stage cost  $\ell(\mathbf{s}_k, \mathbf{a}_k)$  of the environment is a composite of three subterms:

$$\ell(\mathbf{s}_k, \mathbf{a}_k) = L(\mathbf{x}_k, \mathbf{u}_k) + r(\mathbf{u}_k, \mathbf{u}_{k-1}) + \epsilon(\mathbf{x}_k, \mathbf{u}_k). \quad (6)$$

$L(\mathbf{x}_k, \mathbf{u}_k)$  is a quadratic cost term, which penalizes deviations of the system states to the setpoints.  $r(\mathbf{u}_k, \mathbf{u}_{k-1})$  penalizes changes in the control inputs, while  $\epsilon(\mathbf{x}_k, \mathbf{u}_k)$  penalizes constraint violations. The terms are defined as follows with the values from Table 1

$$L(\mathbf{x}_k, \mathbf{u}_k) = 10 \cdot (y_k^D - y_{SP}^D)^2 + 1 \cdot (F^D - F_{SP}^D)^2 \quad (7)$$

$$r(\mathbf{u}_k, \mathbf{u}_{k-1}) = \Delta \mathbf{u}_k^T \begin{pmatrix} 1 & 0 \\ 0 & 30 \end{pmatrix} \Delta \mathbf{u}_k \quad (8)$$

$$\epsilon(\mathbf{x}_k, \mathbf{u}_k) = \mathbf{w}_x^T \boldsymbol{\sigma}_{x,k} + \mathbf{w}_u^T \boldsymbol{\sigma}_{u,k}. \quad (9)$$

Table 1: Setpoints (SP), lower bounds (lb), upper bounds (ub) and constraint penalties ( $w_i$ ) for the states and control actions within the environment.

	$p$	$T$	$F^D$	$y^D$	$F^B$	$x^B$	$\dot{Q}$	$F^F$
	bar	°C	mol s <sup>-1</sup>	—	mol s <sup>-1</sup>	—	kW	mol s <sup>-1</sup>
SP	—	—	0.2	0.45	—	—	—	—
lb	0	20	0	0.4	0	0	0	1
ub	1	90	0.5	1	2	1	30	2
$w_i$	20	20	20	20	20	20	250	250

#### 4.3. Parameterized MPC structure for optimal control of the flash separation unit

The structure from (4) is used to learn the optimal control policy. Model 1 from section 4.1 is used to predict future states. The model is kept unchanged. The parameterization occurs solely in the terminal cost  $V_{f,\theta}$ , which is defined as

$$V_{f,\theta}(x_N) = 10^2 \cdot \theta_{11} \cdot (F_N^D - F_{SP}^D)^2 + 10^4 \cdot \theta_{22} \cdot (y_N^D - y_{SP}^D)^2 \quad (10)$$

with initial values of  $\theta_{11} = \theta_{22} = 0$ . The MPC acts subject to the constraints given in Table 1. The considered stage cost is the same as given in equation (6). The prediction horizon is chosen to be  $N = 10$ . A Q-learning update is then conducted at each time instance using a minibatch of  $m = 16$  samples taken from the replay buffer of size 4'096 with a learning rate of  $\alpha_\theta = 10^{-4}$ . The target MPC is directly updated using  $\tau = 1$ .

#### 4.4. Control performance of parameterized MPC

Figure 2 shows the control performance of a benchmark MPC (rigorous model,  $N = 20$ , solid line) compared to the parameterized MPC (section 4.3, dashed line). The figure shows the controlled states in the left half of each subplot and the control actions on the right half. The dotted lines demonstrate the respective setpoints. The left plot compares the control performance of the benchmark MPC and the parameterized MPC before any training took place. This is similar to applying just MPC with the surrogate model. It can be seen that  $y^D$  is not tracked properly by the parameterized MPC. Also, the control input trajectories differ from the benchmark especially for  $F^F$ . However, it provides a very suitable starting point for RL training in contrast to conventional RL with neural networks, for which initialization is critical for performance but very difficult to choose properly. The right plot shows the performance of the parameterized MPC controller after 100 episodes of training. One episode is a closed loop trajectory of 5 s, each consisting of 50 samples. An update is performed after each time step. The setpoints are now both tracked better. Furthermore, also the control trajectories are closer to those obtained with the benchmark controller. The benchmark achieved a closed-loop cost after 5 s of 0.21367, while the parameterized MPC reduces from initially 0.22063 (untrained) to 0.21749 (100 episodes).

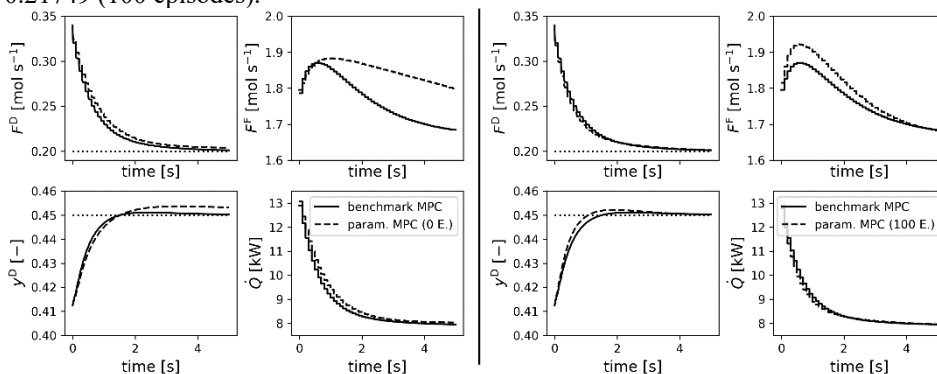


Figure 2. Left: Application of the MPC controller without RL improvements. Right: Application of the RL improved MPC controller. The dotted lines specify the setpoints. 0 E. and 100 E. specify the number of trained episodes.

Both models are also compared with respect to their computation time. The computations were conducted on an intel core i9-10900K CPU, using Python 3.9.13 and the toolbox

do-mpc with CasADi and ipopt. The computation of a single control input for the benchmark MPC took  $0.414 \pm 0.081$  s on average, which is approximately four times the sampling interval of 0.1 s. For the parameterized MPC with the surrogate model, the computation only took  $0.015 \pm 0.003$  s, which is approximately 0.15 times the sampling interval and clearly illustrates the computational advantages of the proposed approach.

## 5. Conclusion

We showed that MPC and RL can be combined to improve controllers in chemical engineering applications. Just by using a simple parameterization of the MPC structure within the terminal cost and a small amount of data, the performance of the controller could be improved to resemble the MPC with a full detailed model, while reducing the computation time by over an order of magnitude.

Future work will include the study of more detailed parametrizations, comparisons with other RL techniques and the extension to more complex control tasks.

## Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 466380688 – within the Priority Program “SPP 2331: Machine Learning in Chemical Engineering”.

## References

- Büskens, C. and Maurer, H. (2001) ‘Sensitivity Analysis and Real-Time Optimization of Parametric Nonlinear Programming Problems’, in Grötschel, M., Krumke, S. O. and Rambau, J. (eds) *Online Optimization of Large Scale Systems*, Berlin, Heidelberg, Springer Berlin Heidelberg, pp. 3–16.
- Esche, E., Weigert, J., Brand Rihm, G., Göbel, J. and Repke, J.-U. (2022) ‘Architectures for neural networks as surrogates for dynamic systems in chemical engineering’, *Chemical Engineering Research and Design*, vol. 177, pp. 184–199.
- Gros, S. and Zanon, M. (2020) ‘Data-Driven Economic NMPC Using Reinforcement Learning’, *IEEE Transactions on Automatic Control*, vol. 65, no. 2, pp. 636–648.
- Himmelblau, D. M. (2008) ‘Accounts of Experiences in the Application of Artificial Neural Networks in Chemical Engineering’, *Industrial & Engineering Chemistry Research*, vol. 47, no. 16, pp. 5782–5796.
- Kingma, D. P. and Ba, J. (2014) *Adam: A Method for Stochastic Optimization*.
- Kocijan, J. (2016) *Modelling and Control of Dynamic Systems Using Gaussian Process Models*, Cham, Springer International Publishing.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S. and Hassabis, D. (2015) ‘Human-level control through deep reinforcement learning’, *Nature*, vol. 518, no. 7540, pp. 529–533.
- Sutton, R. S. and Barto, A. (2018) *Reinforcement learning: An introduction*, Cambridge, Massachusetts, London, England, The MIT Press.

# Architecture and Design of a Modern Commercial Process Simulator

Ian Boys,<sup>a</sup> Jochen Steimel<sup>b</sup>

<sup>a</sup>*AVEVA, 26561 Rancho Parkway South, Lake Forest, CA 92630, USA*

<sup>b</sup>*AVEVA, Mainzer Landstraße 178-190, 60327 Frankfurt, Germany*

*Ian.Boys@aveva.com*

## Abstract

Over the past 10 years AVEVA (formerly SIMSCI, Invensys) has been developing a next generation process simulator, with a clean design from the ground up. In this article we will share the experience developing and marketing this software and reflect on how early design ideas have evolved and matured into a commercially available software application. The software is based on a novel concept for setting up, parametrizing, initializing, and solving equation-based simulation models targeting the next generation of process engineers. We will discuss some of the successes and challenges in realizing this next-generation concept. We will furthermore discuss customer reaction and feedback, and how the software has been received in the marketplace.

**Keywords:** Simulation, Software, Numerical Methods, Modeling

## 1. Introduction

In recent years, two contributions were published, one from industry Broecker (2021), and one from academia Pistikopoulos (2021), which both highlight the need for process simulation software to consider advances in computing technology as well as modern algorithms and user experience expectations. Both publications highlight the need for open interfaces, enabling cross-domain applications, which combine real-world data with first principles and data-driven models. Broecker points out that future innovation will arise at the intersection and combination of tools, as complex problems demand cross-domain expertise and no tool can fill all needs. This vision of course requires open tools that allow interaction and integration into larger, more complex workflows. There seems to be a consensus that established tools have so far been lacking in this regard.

In their review article, Mitsos et. al (2018) spelled out the following conflict: “[in the petrochemical industry] many of the developed chemical processes are prototypical applications [...]. Often, there is no good business case to justify the implementation of a special solution to a very specific problem class in commercial simulation software”. Further in the article, they lament the limited interaction between software vendors and the industry and the slow adoption of novel academic techniques into commercial software. As a possible solution, they conjecture that third parties could fill the gap of capabilities lacking in commercial tools, but also admit that the limited accessibility of those tools often prohibits the implementation of new process systems engineering methods into an integrated offering acceptable for industrial use. The authors want to address this gap, by discussing a tool that is open for extension and research by academia, yet fit for every-day use by industrial practitioners.

We observe: the world of process engineering is changing fast, and next-generation tools must have an outstanding degree of flexibility and customizability to serve the needs of a future generation of process engineers, to empower them to tackle the various challenges induced by the energy transition and the need to switch to sustainable feedstocks. With this contribution we want to discuss the salient features of one commercial simulation software that was developed from the ground up with extensibility and future use-cases in mind, made possible by strong customer interaction and short development cycles.

## 2. Properties of a Next-Generation Simulator

### 2.1. Overall architecture

To address the needs discussed above AVEVA started work on developing a next-generation process simulator about 10 years ago. The main goal of this project was to create a platform (called AVEVA Process Simulation, further referred to as the software) that would be powerful enough to support all kinds of process engineering, simulation, and design workloads, while also being in line with modern standards of user experience, scalability, and responsiveness. The software is designed to drive the digital transformation in process engineering through lifecycle simulation and a holistic approach, as demonstrated in de Beer (2021).

A central pillar is our modern client/server architecture. In the software, client interactivity is separated from simulation analysis, model instantiation and solving. This design liberates the software from the requirement that both parts are being run on the same machine, enabling scalability and cloud computing.

For the end-user the simulation platform is flexible and can be adapted to changing use cases. This flexibility is mainly provided by an open-equation custom modeling environment powered by an equation-oriented solver. Instead of implementing process equations in subroutines, as sequential-modular process simulators might do, in the software, (nearly) every equation, variable, model and library can be inspected and modified by the end-user. The only exceptions to this are external equation sets, which allow external code to be connected to the platform (see sec. 2.4.). The availability of all equations and variables as first-class objects enables the dynamic restructuring and re-interpretation of the equation system depending on the problem statement. For example, a seamless transition from steady-state to dynamic simulation is supported by the platform, as only the time derivatives need to be handled differently between the solver and the dynamic integrator. Thus, it is possible to go back-and-forth between steady-state and dynamic mode for the same model, without the need to manually adjust equations or specifications.

When developing new modeling capabilities, the development team is using the same tools available to the end-user. Everything in our libraries can be reconstructed by the user as there are no hidden authoring tools, making the software an open and transparent modeling tool. A screenshot illustrating the model editor, for a model describing the heat transfer through a pipe wall, making use of steady-state and dynamic equations side-by-side is given in fig. 1. An example for this extensibility can be found in Bishop (2020), where the author developed a model for a membrane reactor using our software and performed operability studies.

In line with the platform design, thermodynamic calculations and databases are kept separate from the client/server. AVEVA Process Simulation offers a wide set of industry standard thermodynamic routines for equation of states or activity coefficient models, as well as thermodynamic codes for specialized systems (e.g., electrolytes, cryogenic hydrogen, or supercritical CO<sub>2</sub>).

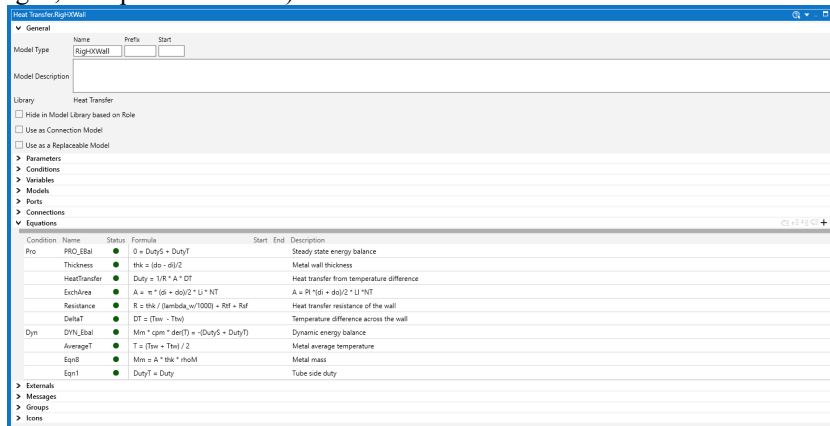


Figure 1: Model equation editor showing steady-state and dynamic equations

## 2.2. Numerics

The numerical engine powering the platform is a general-purpose equation solving system consists of an instantiation algorithm that translates a model into a flat equation system, a block decomposition algorithm that reduces a large equation system into smaller equation systems that can be solved independently, and a solver factory that applies tailored solvers to each of the equation systems. The block decomposition system is based on the graph algorithms of Tarjan (1972) and Dulmage (1958). Each well-specified block is analyzed, and if the convergence criteria are not met, a solver suitable for the block is selected to converge the residuals. If the block decomposition determined that any blocks are independent from each other, as often happens with phase equilibrium equation sets, they can be solved in parallel to reduce execution time.

## 2.3. Interactivity

Over the last years, expectations on interactivity in computer programs have shifted dramatically. Apps and smartphones have reached widespread adoption, which led to an increased demand in application ergonomics. In the past, simulators often followed a four-step procedure: define flowsheet topology, provide specifications, trigger solve and inspect output. Many times, users needed to backtrack when it became apparent that a certain topology would not allow the desired set of specifications, but that fact only became apparent during the inspection of the log files. Today users expect instant response and an exploration-driven user interface where making mistakes is cheap and easy to remedy and information is not hidden deep in nested dialogs. To address this shift in expectations for industrial software, AVEVA Process Simulation supports three paradigms that were quite uncommon in simulation tools so far: solve-as-you-go, undo/redo and a simulation repository.

Empowered by the state-of-the-art algorithms, the simulation building core of the software can identify parts of a flowsheet that have changed after a user interaction and need an update. This results in smaller equation systems to be rebuilt and solved and a very fast response and presentation of results. Likewise, the undo/redo system can undo



changes to parameter and variable values and even impactful operations like optimization runs. In addition to reverting an action, the system uses implicit snapshots of the “state before” to forego the need for an additional convergence pass after undo, resulting in an even faster response. The simulation repository is an abstraction on the file system. Once a simulation is imported from the file system into the repository, it is managed by the software. Any change to the flowsheet is persisted instantaneously (“auto-save”). Combined with undo/redo, this leads to a seamless modeling experience, without the fear of losing data (e.g., caused by a crash of the software).

Together, these three features empower a special way of approaching equation-oriented modeling and simulation that is fundamental to our software: incremental initialization. It is well-known that finding good initial guesses for an equation-oriented simulation is a non-trivial task and crucial in simulators relying on a general equation solver. In the software, the flowsheet is built up one unit operation at a time. By following a few basic rules (fully specify each unit before moving to the next, always have a square simulation with the same number of unknown variables and equations, undo as soon as the solver cannot solve a problem) it is very easy, also for inexperienced users, to bootstrap large-scale process flowsheet models, even with recycle loops. At the time when a recycle needs to be closed, the returning stream already is at a converged open-loop solution, so convergence is often quick and robust. In addition, a snapshot of the current simulation can be taken, which allows the user to re-initialize to a good solution.

#### *2.4. Extensibility*

As mentioned before the software is designed as an open platform, ready to be extended to fit the user’s need in process simulation. To support those use cases, the platform offers two main entry points for extension beyond the already flexible open-equation modeling framework.

As mentioned briefly in section 2.1 the software offers an external equation interface which is part of a software development kit (SDK). With the SDK users can provide custom calculations that cannot be expressed in algebraic form (e.g., thermodynamic algorithms, machine-learning models). Users need to provide callback methods that return equation residuals and partial derivatives to fit into our equation-oriented solver framework, but how they calculate those values is fully under their control. In this way it becomes possible to extend modeling to non-smooth equations, conditional logic, or algorithmic structures. To exemplify the power of the SDK, the default thermodynamic methods provided by the software are implemented using the external equation toolkit, keeping in line with our philosophy that any platform user should be able to use the same tools that our developers are using.

The second extension point is the scripting interface. As a side effect of the client/server architecture, any action that can be performed in the GUI client is accessible as a function call for a scripting client. The scripting interface is bundled with the product as a Python module and a .NET library. Any clients, such as Excel/VBA, Python scripts or full native desktop applications can automate our software to implement complex computing workloads such as structural optimization, training of surrogate models and uncertainty analysis.

### **3. Experiences in Developing a Next-Generation Simulator**

As part of the go-to-market strategy, the software benefitted greatly from a customer-centric approach. Instead of developing against an internally defined roadmap, customers were engaged early on and asked to provide feedback on the priority of development activities. Instead of our treating us as a regular software vendor, formal partnerships allowed both parties to synchronize migration and development activities, so that important features were made available when they were needed. We have mapped our target audience against archetypal personas, representing a wide range of users, including students, junior/senior design engineers, thermodynamicists, modelling experts and control engineers. When new features are proposed, we make sure to consult our list of personas to make sure the functionality fits with their needs.

The agile software development process underlying the development activities made it possible to release incremental improvements early and rapidly. This in turn allowed customers to give feedback quickly so that features could be fine-tuned while still in development resulting in a shorter turnover time for new features.

The users of the software are very engaged and are given ample opportunities to be heard, from user meetings over advisory boards to direct meetings with the development team in voice-of-customer sessions. The level of trust resulting from this close contact, allowed developers to get first-hand feedback from real users, a crucial requirement for our agile process. Some of that feedback was expected, some was surprising. In the following, we want to share the aspects that were most important to us.

Not surprisingly, adapting to the equation-oriented paradigm was a challenge to users accustomed to sequential-modular simulators. The benefits coming from it, e.g., near instant results and being able to select any variable to be specified or calculated were highly regarded, but especially finding the root cause of an unconverged simulation proved to be very hard for many users. Interestingly, we found out that younger users, who were not as used to the sequential-modular paradigm, found it easy and intuitive to learn our software, whereas experienced engineers struggled more often.

To address this feedback, we greatly expanded the information about the solution that was reported to the user, in the form of log files and small badges shown over flowsheet icons. Whenever possible this information was translated from the mathematical domain into the process engineering domain. For example, instead of reporting a singular solution, the software reports that for a certain variable no unique value can be determined and that a different specification set should be tried. In another example, instead of just reporting degrees of under- or overspecification, the system visually marks potential variable candidates that are part of the under- or over-specified equation sets. Our goal is to provide a list of small, actionable suggestions that could lead to a solution of the problem instead of a general “solve failed” message.

A well-liked feature is the customizability of the flowsheet canvas. The software supports variable references to be added next to unit operations, which allow variables to be displayed and changed without the need to open a new dialog. Similarly, equations can be added to the flowsheet, so that custom design specifications can be shown next to the process equipment. This feature enables customers to document their flowsheets and

simulations within the software itself and reduces the need for external explanations, improving reuse and collaboration.

#### 4. Conclusion

In this article, we have shared the fundamental architecture and design philosophy for our next-generation process simulation tool. Our conclusion is that the next generation concepts envisaged at the outset have proven viable in practice and are able to sustain an industrially relevant and capable process simulator. We base this conclusion upon customer feedback, both from industry and academia (research and teaching). Interactivity is a huge factor in attaining a better understanding of a process and is appreciated by all customers.

AVEVA Process Simulation is reaching wider adoption in the chemical industry and a feature maturity that will allow the software to expand into modern usage scenarios, such as design optimization or uncertainty quantification. We therefore think that the next generation of process engineers can look forward to new ways of working with process simulation, where the workflow is driven more by the goals of the engineer and less by the algorithms of the software.

#### References

- S. Broecker et. al., 2021, Process Simulation – Fit for the future?, Position Paper of the Process Net Working Committee, DECHEMA e.V.
- E.N. Pistikopoulos et. al., 2021, Process system engineering – The generation next?, *Computers & Chemical Engineering*, 147, 107252
- A. Mitsos, 2018, Challenges in process optimization for new feedstocks and energy sources, *Computers and Chemical Engineering*, 113, 209-221
- J. de Beer et. al., 2021, The Role of Process Engineering in the Digital Transformation, *Computers & Chemical Engineering*, 154, 107423
- B. Bishop et. al., 2020, Modeling, Simulation, and Operability Analysis of a Nonisothermal, Countercurrent, Polymer Membrane Reactor, *Processes*, 8, 78
- R. Tarjan, 1972, Depth-First Search and Linear Graph Algorithms, *SIAM Journal on Computing*, 1, 14-160
- A.L. Dulmage et. al., 1958, Coverings of Bipartite Graphs, *Canadian Journal of Mathematics*, 10, 517-534

# ***High and Robust Fault Detection via Polynomial Approximated Isomap Embeddings***

Burak Alakent<sup>a</sup>

<sup>a</sup>: *Department of Chemical Engineering, Bogazici University, Istanbul, 34342, Turkiye*  
*burak.alakent@boun.edu.tr*

## **Abstract**

In Multivariate Statistical Process Monitoring (MSPM), a large number of measured variables is monitored online usually in a latent variable space. To this end, various linear/nonlinear, Gaussian/non-Gaussian, and static/dynamic methods have been proposed. While implementations of deep learning methods in MSPM are frequently seen nowadays, less emphasis is given to unsupervised nonlinear manifold learning (ML) methods, such as Isomap. It is important to note that ML takes the geometry of the data into consideration while reducing dimensions, hence may have an advantage, particularly, over Kernelized methods. However, requirement of approximations for out-of-sample (test) points renders ML methods less practical. To remedy this issue, we have recently proposed independent component analysis of polynomial approximation to Isomap embeddings coupled with principal component analysis (ICA<sub>plso</sub>-PCA) method, and showed that fault detection and isolation performances are drastically improved compared to traditional methods. In the current study, we include lagged process measurements in ICA<sub>plso</sub>-PCA, i.e. dynamic ICA<sub>plso</sub>-PCA (dICA<sub>plso</sub>-PCA), and show that fault detection rate is further increased, and more robust with respect to number of selected components compared to various methods on Tennessee Eastman plant.

**Keywords:** Fault detection, dynamic PCA, ICA, Manifold Learning, Tennessee Eastman.

## **1. Introduction**

Occupational and process safety issues are nowadays more of concern than ever, hence fault detection, isolation, identification and prognosis have recently gained utmost importance. To achieve this aim, Multivariate Statistical Process Monitoring (MSPM) is one of the most frequently employed methods mainly due to i) difficulties in obtaining accurate mechanistic mathematical models for chemical industries, and ii) developments in statistical and machine learning methods making it possible to extract information from high dimensional, nonlinear and collinear process variables (Ji and Sun; 2022). Though Principal Components Analysis (PCA) and Projection to Latent Structures (PLS) are historically the benchmark methods for MSPM (Kourti and MacGregor; 1995), both methods are known to suffer from various limitations, such as extracting only static relations, and relatively poor performance under nonlinear process dynamics and non-Gaussian variable distributions. Various combinations of Kernel methods (Lee et al., 2004) and Independent Components Analysis (ICA) (Kano et al., 2003; Lee and Qin, 2007) are commonly employed to overcome these limitations. Furthermore, applications

of deep Artificial Neural Network (ANN) models are shown to increase the fault detection performance in MSPM applications (Qian et al., 2022).

With the new millennium, there has been a surge of research activity in manifold learning, such as Isomap (Tenenbaum, de Silva, and Langford, 2000), and locality preserving projections (LPP) (He and P. Niyogi, 2003), and various applications have been implemented for MSPM purposes. While manifold learning methods, unlike Kernel and ANN methods, offer the capability of determining the geometry of a low-dimensional manifold, on which the process data is deemed to lie, nonlinear manifold learning methods lack explicit mapping function to transform the future query points into embedding space. A number of solutions have been proposed to tackle this issue. One solution method is to approximate the nonlinear embedding using a linear transformation in the optimization objective function (Luo, 2014; Miao, Song, and Shen, 2015). Another method is to approximate the nonlinear transformation using various learners, such as ANN (van der Maaten, 2009). Recently, we have proposed ICA of polynomial approximation to Isomap embeddings coupled with PCA (ICA<sub>pIso</sub>-PCA) method, in which embedding approximations are performed using a high order (regularized) polynomial, rendering both high fault detection rates, and contribution plots with highly reduced smearing effect (Alakent, 2022). In the current study, we aim to investigate whether fault detection rate can further be improved including lagged process measurements; the resulting method is named dynamic ICA<sub>pIso</sub>-PCA (dICA<sub>pIso</sub>-PCA). Applications on Tennessee Eastman simulation plant show that fault detection rates of dICA<sub>pIso</sub>-PCA are generally higher compared to those of traditional and state-of-the-art MSPM methods, and ICA<sub>pIso</sub>-PCA. The most significant advantage of the suggested dICA<sub>pIso</sub>-PCA method is shown to be its high fault detection accuracy over a wide range of number of components selected, making it easier to construct a reliable monitoring model without using a very large dataset.

## 2. Dimensional Reduction Tools used in SPM

### 2.1. PCA

The aim in PCA is to determine linear projections of the original coordinates with the highest variance. Given the mean-centered (and scaled) data matrix  $\mathbf{X} \in \mathbb{R}^{N \times D}$ , where  $N$  and  $D$  correspond to number of samples and process variables, or features, respectively, eigenvalue decomposition of the covariance matrix  $\mathbf{C}$  yields the loading matrix  $\mathbf{P} \in \mathbb{R}^{p \times p}$ , on which  $\mathbf{X}$  is projected to obtain  $\mathbf{T} \in \mathbb{R}^{N \times D}$  scores matrix:

$$\mathbf{C} = \mathbf{X}^T \mathbf{X} / (N - 1) = \mathbf{P} \mathbf{\Lambda} \mathbf{P}^T \quad (1)$$

$$\mathbf{T} = \mathbf{X} \mathbf{P} \quad (2)$$

Usually, the first  $d$  ( $< D$ ) principal components (PCs) with the highest eigenvalues (variances) are used to construct a lower dimensional subspace, in which essential process variability is deemed to be captured:

$$\mathbf{T}_d = [\mathbf{t}_1 \ \mathbf{t}_2 \ \dots \ \mathbf{t}_d] = \mathbf{\Psi} \mathbf{P}_d = \mathbf{\Psi} [\mathbf{p}_1 \ \mathbf{p}_2 \ \dots \ \mathbf{p}_d] \quad (3)$$

### 2.2. ICA

The generative model of ICA is  $\mathbf{x} = \mathbf{A} \mathbf{s}$ , in which  $\mathbf{s} \in \mathbb{R}^m$ , consisting of  $m$  independent components each with unit variance, and  $\mathbf{x} \in \mathbb{R}^D$  is a zero-mean random variable vector, and  $\mathbf{A} \in \mathbb{R}^{D \times m}$  is the mixing matrix. A demixing matrix  $\mathbf{W} \in \mathbb{R}^{m \times D}$  is to be determined:

$$\mathbf{Y} = \hat{\mathbf{S}} = \mathbf{W} \mathbf{X} \quad (4)$$

The problem stated above may be solved via Fast-ICA procedure, in which each vector in  $\mathbf{W}$  and  $\mathbf{Y}$  is determined iteratively (Hyvärinen, Karhunen and Oja, 2001). Similar to PCA, a reduced  $d$  ( $< D$ ) number of ICs ( $\mathbf{W}_d$ ) may be used. However, ranking ICs is not straightforward as that in PCA, since variance of each IC is constrained to be equal to unity. In the current study, ranking ICs is based on the negentropy of  $\mathbf{y}$  values.

### 2.3. Isomap

In Isomap, sample points are assumed to lie on a smooth manifold, and approximate geodesic distances, i.e. distances measured via shortest tracks on the manifold, are used to represent pairwise sample distances. The K-Isomap method involves connecting  $K$  nearest neighbors ( $K=5$  used currently) to each sample point, and storing the squares of the geodesic distances, determined via summing the distances along the shortest paths, in a  $\mathbf{S} \in \mathbb{R}^{N \times N}$  matrix. Then, eigenvalue decomposition is employed on  $\mathbf{S}$  mean-centered via  $\mathbf{H} = \mathbf{I} - \mathbf{1}\mathbf{1}^T/N$ , in which  $\mathbf{I}$  is the identity matrix, and  $\mathbf{1} = [1 \ 1 \ \dots \ 1]^T \in \mathbb{R}^N$ :

$$-\frac{1}{2}\mathbf{H}\mathbf{S}\mathbf{H} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T \quad (5)$$

Similar to PCA, the eigenvalues are ranked in decreasing order, and the largest  $d$  eigenvectors of  $\mathbf{V}$  are used to construct the Isomap coordinates:

$$\mathbf{Y} = \mathbf{\Lambda}_d^{1/2}\mathbf{V}_d^T \quad (6)$$

Unlike  $\mathbf{P}_d$  and  $\mathbf{W}_d$  in PCA and ICA, respectively, Isomap does not yield an explicit transformation matrix that may be used for embedding novel samples; hence, approximations are required to determine the Isomap coordinates of future points.

### 2.4. Dynamic PCA (dPCA)

PCA is based on the assumption of time-independent samples, that is easily violated in industrial processes. In order to include the autocorrelation in process measurements in constructing the PC space, dPCA was proposed (Ku, Storer and Georgakis, 1995). The dPCA involves augmenting the features matrix  $\mathbf{\Psi}$  using the current (sample  $n$ ) and  $\tau$  time lagged (samples  $n-1, n-2, \dots n-\tau$ ) measurements as follows:

$$\mathbf{\Psi} = [\mathbf{X}(n) \ \mathbf{X}(n-1) \ \dots \ \mathbf{X}(n-\tau)] \in \mathbb{R}^{(N-\tau) \times D(\tau+1)} \quad (7)$$

The rest of dPCA involves applying eigenvalue decomposition on the covariance matrix of  $\mathbf{\Psi}$  and employ traditional PCA, as in Eqns. (1-3). The dPCA method was shown to improve fault detection performance since deterioration of the time-dependence of the variables could also be monitored. Including time lagged measurements in other dimension reduction tools, such as dynamic ICA (dICA), have been widely used.

## 3. Principal Fault Detection Indices in SPM

Based on the traditional application of PCA,  $T^2$  and  $SPE$  are the most commonly used surrogate variables (indices), which measure the proximity of the test samples to origin of normal operation in the reduced subspace, and the magnitude of the residual component of the test sample, unaccounted by the reduced space, respectively.

$$T^2 = \sum_{i=1}^d \frac{t_i^2}{\lambda_i} = \sum_{i=1}^d \frac{(\mathbf{x}^T \mathbf{p}_i)^2}{\lambda_i} \quad (8)$$

$$SPE = \|\mathbf{x} - \mathbf{P}_d^T \mathbf{t}\|_2^2 \quad (9)$$

In Eqn. 8, the coordinates of the query sample in the original (mean-centered) and projected spaces are shown with  $\mathbf{x}$  and  $\mathbf{t} = [t_1 \ t_2 \ \dots \ t_d]^T$ , respectively, while  $\lambda_i$  denotes the variance of scores of normal training data. In Eqn. 9, the squared norm of the residuals of the query point is determined. Phase I of MSPC consists of projecting historical normal operation data to a lower dimensional space, followed by determining the control limits

(CLs) of surrogate variables, i.e.  $T_{lim}^2$  and  $SPE_{lim}^2$ , at a specific Type-I error rate. Phase II consist of checking whether the surrogate indices of novel measurements ( $T^2$  and  $SPE^2$ ) fall within CLs.

#### 4. The Proposed Method: dICA<sub>pIso</sub>-PCA

We recently proposed ICA<sub>pIso</sub>-PCA, in which Isomap embeddings are approximated using ridge regression of fourth to sixth ordered polynomials, followed by ICA and PCA (Alakent, 2022), and, in the current study, we aim to improve the fault detection capability of ICA<sub>pIso</sub>-PCA via addition of lagged variables. The dICA<sub>pIso</sub>-PCA starts with employing PCA and Isomap separately on  $\Psi$ , which is the features matrix consisting of time lagged measurements (see Eqn. 7), and all the PC scores and the first  $I_p$  ( $I_p=8$  used currently) embedding coordinates are stored. Then, the following regressor matrix  $\Psi^*$  is formed:

$$\Psi^* = [\Psi \quad (\Psi)^{o2} \quad \dots \quad (\Psi)^{or}] \in \mathbb{R}^{(N-\tau) \times D\tau(\tau+1)} \quad (10)$$

Here,  $(\cdot)^{oa}$  is used to raise the power of each element in the matrix to  $a = 1, 2, \dots, r$ , and  $r$  is taken to be equal to six. Taking  $\mathbf{F} \in \mathbb{R}^{N \times I_p}$  as the matrix of Isomap embeddings, ridge regression ( $\alpha = 10^{-4}$ ) is employed to determine the approximated Isomap embeddings:

$$[\hat{f}_1 \hat{f}_2 \dots \hat{f}_{I_p}] = \Psi^* (\Psi^{*T} \Psi^* + \alpha \mathbf{I})^{-1} \Psi^{*T} \mathbf{F} \quad (11)$$

Highest linearly correlated PC scores with Isomap approximations are replaced by each of  $[\hat{f}_1 \hat{f}_2 \dots \hat{f}_{I_p}]$ , i.e. PC scores, most similar to high variance approximated Isomap embeddings, are replaced by these approximations, yielding a “nonlinear scores” matrix of reduced dimensions of the training data,  $\mathbf{T}_{nl} \in \mathbb{R}^{(N-\tau) \times D'}$ ,  $D' \geq D(\tau + 1)$ . Then, ICA and PCA are sequentially employed on  $\mathbf{T}_{nl}$ , i.e. application of ICA yields the following.

$$\hat{\mathbf{S}} = [\hat{s}_1 \hat{s}_2 \dots \hat{s}_N] = \mathbf{W} \mathbf{T}_{nl}^T \quad (12)$$

Here,  $\mathbf{W}$  is the demixing matrix in Eqn. (4), and used to transform the nonlinear (correlated) scores matrix to independent scores  $\hat{\mathbf{S}}$ . The IC scores are computed via  $I^2 = \hat{\mathbf{S}}^T \hat{\mathbf{S}}$ , and 99% CL for  $I^2$  indices ( $I_{lim}^2$ ) is computed using kernel density estimator. The  $I^2$  value of a new query point is compared with  $I_{lim}^2$ , and out-of-control alarm is given if  $I^2 > I_{lim}^2$ . The same procedure is repeated with  $T^2$  and  $SPE$  indices, which are obtained via employing PCA on the residuals of the nonlinear scores, unexplained by the ICA. For a more detailed explanation of the procedure, please refer to Alakent (2022).

#### 5. Results and Discussion

We employed dICA<sub>pIso</sub>-PCA on the simulated Tennessee Eastman plant (TEP) (Downs and Vogel, 1993). The main products in TEP are G and H with the purged byproduct F from a feed consisting of reactants A, C, D, E and inert component B. In our study, data from 33 process variables (excluding concentration measurements) sampled at 3 min intervals were used. Normal operation data is comprised of 960 samples, while test data consists of one normal operation, and 21 different faults of 960 samples (Chiang, Russell, and Braatz, 2000). The process diagram, variables and fault descriptions of TEP, omitted here for space considerations, can be found in the sources given in the References. We compared Missed Detection Rate (MDR) values, i.e. percent of faulty samples deemed to be normal, of dICA<sub>pIso</sub>-PCA with the best results obtained from dynamic and/or Kernel methods, such as KICA, dPCA, dICA, dICA<sub>PCA</sub> and dKPCA. CLs of the monitoring indices in all methods were adjusted to yield a false detection rate of 0.01 (rate of false signals for normal query points) in order to render an unbiased comparison of MDR values (Mutlu and Alakent, 2019). We used two time lags in all dynamic dimension

reduction methods. To save space, we here show only the minimum MDR obtained from all monitoring indices in Table 1. We included a state-of-the-art method, namely Global Local Preserving Projection (GLPP), which utilizes the NPP to capture both the local and global structure of data (Miao, Song and Shen, 2015).

In Table 1,  $dICA_{plso}$ -PCA gives the smallest average MDR, significantly smaller than  $ICA_{plso}$ -PCA, indicating that including time-lagged measurements improves fault detection rate of our method. Looking in more detail,  $dICA_{plso}$ -PCA yields the top and second best performances in 14 and the remaining 7 faults, respectively. On the other hand,  $dGKPP$  has the best and second best performance in 9 and 6 of the faults, respectively, while exhibiting a moderate performance in the remaining 6 faults, showing a poorer fault detection capability compared to that of  $dICA_{plso}$ -PCA.

Table 1. MDRs of various MSPC models against 21 faults in TEP.

Fault no.	KICA <sup>i</sup>	GKPP <sup>ii</sup>	$ICA_{plso}$ -PCA <sup>iii</sup>	dPCA <sup>iv</sup>	dKICA <sup>v</sup>	dKPCA <sup>vi</sup>	dGKPP <sup>vii</sup>	$dICA_{plso}$ -PCA <sup>viii</sup>
1	<b>0</b> <sup>ix</sup>	0.25	<b>0</b>	0.500	0.125	0.125	0.125	0.125
2	1.5	1.5	1.5	1.75	1.75	1.625	1.5	<b>1.125</b>
3	96.9	97	97.1	98.4	96.1	96.3	95.5	<b>95.1</b>
4	<b>0</b>	0.625	<b>0</b>	0.249	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>
5	<b>0</b>	<b>0</b>	<b>0</b>	72.1	<b>0</b>	0.125	<b>0</b>	<b>0</b>
6	<b>0</b>	<b>0</b>	<b>0</b>	0.249	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>
7	<b>0</b>	<b>0</b>	<b>0</b>	0.249	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>
8	2	2.13	2.13	2.63	1.88	1.63	1.63	<b>1.5</b>
9	96.6	97.5	96	98.1	97	96	96.5	<b>95.9</b>
10	32.6	11.1	11	63.6	7.875	13.75	<b>5.75</b>	6.38
11	19.8	28.4	17.5	12.5	11.9	6.75	19.6	<b>4.88</b>
12	0.25	0.125	0.125	0.998	<b>0</b>	0.25	0.125	<b>0</b>
13	4.75	5.38	4.75	4.99	4.25	4.38	<b>3.63</b>	3.88
14	<b>0</b>	<b>0</b>	<b>0</b>	0.249	<b>0</b>	<b>0</b>	0.125	<b>0</b>
15	95.6	94.1	94	96.6	83.4	92.1	<b>60.9</b>	65
16	38.6	8	8.63	70.1	6.25	8.375	<b>3.13</b>	4.00
17	3.75	7.50	2.75	3.74	2.38	2.13	2.63	<b>2.00</b>
18	10	10.1	9.50	10.1	9.375	<b>9.25</b>	9.50	9.38
19	46.6	13.9	5.13	28.8	0.75	1.88	0.375	<b>0.125</b>
20	33	9.13	8.75	46.9	11.1	28.8	<b>8</b>	8.38
21	43.8	47.8	43.6	49.1	54.4	51.3	49.9	<b>43.5</b>
Avg <sup>ix</sup>	25.0	20.7	19.2	31.5	18.5	19.7	17.1	<b>16.3</b>

<sup>i</sup> 40 components, <sup>ii</sup> 19 components with  $K=10$  and  $\sigma = 2$ , <sup>iii</sup> 1 IC and 4 PCs, <sup>iv</sup> 21 components, <sup>v</sup> 10 ICs, <sup>vi</sup> 65 PCs,  $\sigma = 300$ , <sup>vii</sup> 29 components with  $K=10$  and  $\sigma = 10$ , <sup>viii</sup> 6 ICs and 4 PCs.

<sup>ix</sup> Bold values correspond to smallest MDR values obtained for each fault type.

<sup>x</sup> Average MDR values for each model shown in the last row.

Additionally, Fig. 1 shows that average MDR of  $dICA_{plso}$ -PCA does not significantly depend on the number of selected components, unlike dKPCA and dGKPP, which, in practical applications, would require separate datasets comprising faulty operating conditions to adjust the number of components. Hence, selecting the “right” number of components, that is still an unsolved issue in MSPM applications, seems to be less of an issue for  $dICA_{plso}$ -PCA.



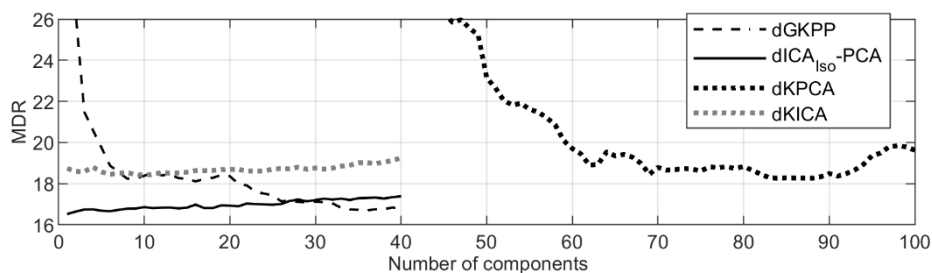


Figure 1. Change of average MDR with respect to number of PCs or ICs.

## References

- C. Ji and W. Sun (2022), A Review on Data-Driven Process Monitoring Methods: Characterization and Mining of Industrial Data, *Processes*, 10, 335.
- T. Kourti and J.F. MacGregor (1995), Process analysis, monitoring and diagnosis, using multivariate projection methods, *Chemom. Intell. Lab. Sys.*, 28, 3-21.
- J.-M. Lee, C.K. Yoo, S.W. Choi, P.A. Vanrolleghem, and I.-B. Lee (2004), Nonlinear process monitoring using kernel principal component analysis, *Chem. Eng. Sci.*, 59, 223-234.
- M. Kano, S. Tanaka, S. Hasebe, I. Hashimoto, and H. Ohno (2003), Monitoring independent components for fault detection, *AIChE J.*, 49, 969-976.
- J.-M. Lee and S.J. Qin (2007), Fault Detection of Non-Linear Processes Using Kernel Independent Component Analysis, *Can. J. Chem. Eng.*, 85, 526-536.
- J. Qian, Z. Song, Y. Yao, Z. Zhu, X. Zhang (2022), A review on autoencoder based representation learning for fault detection and diagnosis in industrial processes, *Chemom. Intell. Lab. Syst.*, 231, 104711.
- B. Tenenbaum, V. de Silva, and J.C. Langford (2000), A global geometric framework for nonlinear dimensionality reduction, *Science*, 290, 2319-23.
- X.F. He and P. Niyogi (2003), Locality preserving projections. *In Proceedings of the Conference on Advances in Neural Information Processing Systems*, Dec 8–13, Vancouver, Canada; MIT Press: Cambridge, MA.
- L. Luo (2014), Process Monitoring with Global-Local Preserving Projections, *Ind. Eng. Chem. Res.*, 53, 7696–7705.
- E.C. Mutlu and B. Alakent (2019), Revisiting reweighted robust standard deviation estimators for univariate Shewhart S-charts, *Qual. Reliab. Eng.*, 35, 995-1009.
- A. Miao, Z. Ge, Z. Song and F. Shen (2015), Nonlocal structure constrained neighborhood preserving embedding model and its application for fault detection. *Chemom. Intell. Lab. Syst.*, 142, 184-196.
- L. van der Maaten (2009), Learning a parametric embedding by preserving local structure, *AISTATS*, volume 5 of JMLR Proceedings, 384-391.
- B. Alakent (2022), Reducing Smearing Effect in Contribution Plots and Improving Fault Detection via Polynomial Approximated Isomap Embeddings, *Can. J. Chem. Eng.*, DOI: 10.1002/cjce.24689.
- A. Hyvärinen, J. Karhunen and E. Oja (2001), *Independent Component Analysis*, John Wiley & Sons, Inc..
- W. Ku, R.H. Storer and C. Georgakis (1995), Disturbance detection and isolation by dynamic principal component analysis, *Chemom. Intell. Lab. Syst.*, 30, 179-196.
- J.J. Downs and E.F. Vogel (1993), A plant-wide industrial process control problem, *Comput. Chem. Eng.*, 17, 245-255.
- L.H. Chiang, E.L. Russell, and R.D. Braatz (2000), Fault diagnosis in chemical processes using Fisher discriminant analysis, discriminant partial least squares, and principal component analysis, *Chemom. Intell. Lab. Syst.*, 50, 243–252.

# A comparative study of distributed feedback-optimizing control strategies

Vegard Aas,<sup>a</sup> Risvan Dirza,<sup>a</sup> Dinesh Krishnamoorthy,<sup>a</sup> Sigurd Skogestad,<sup>a\*</sup>

<sup>a</sup> *Department of Chemical Engineering, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway*

*\*sigurd.skogestad@ntnu.no*

## Abstract

In industry, the processes often consist of several subsystems with a common constraint, for example, a shared resource. This paper considers the problem of steady-state real-time optimization (RTO) for a subsea gas-lifted oil production network with multiple wells and constrained access to shared gas-lift supply. Such problems can be solved by a centralized numerical optimization, which can be computationally expensive. To avoid the use of numerical optimization, one can utilize either online primal or dual decomposition instead, where the problem is converted into a feedback-based problem. The main benefit of primal method is that it distributes local setpoints which complies with primal feasibility, however, the dual method is more general. Both the primal and dual methods allow for a distributed implementations. The dual method is more general in terms for allowing for many constraints, but as shown for the simulations with uncertainty and measurement noise, the primal method may give better dynamic constraint satisfaction.

**Keywords:** Distributed optimization, Feedback control, Production optimization.

## 1. Introduction

Industrial process often consists of several subsystems with a common constraint, for example, a shared resource. Problems like this can be decomposed and solved using distributed optimization. However, this can be computationally expensive as it solves several rounds of numerical optimization problems online at each sample time. This can be addressed by indirectly moving the optimization problem into the control layer (Morari et al. (1980)). Such problems are known as feedback-optimizing control, which can be implemented using simple tools such as Proportional-Integral-Derivative (PID) controllers.

In our previous work (Dirza et al., 2022(b)), we have experimentally validated a recently developed method of feedback-optimizing control called distributed feedback-based RTO. This method is developed based on dual decomposition and optimally handles steady-state changes in active constraints. However, the constraints are controlled in a slower time scale by updating the dual variables. This leads to the need for significant “back-off” strategy, which could lead to profit loss in the long run. To eliminate or reduce the “back-off,” Dirza et al. (2022(a)) introduces an alternative distributed feedback-optimizing control based on online primal decomposition using feedback and constraint controller(s) which distribute local setpoints without violating the common constraint to avoid or minimize use of a “back-off” strategy.

In this paper, we provide a comparative analysis of the two distributed feedback-optimizing control approaches. The model we consider represents our lab-scale experimental rig that consider uncertainty and measurement noise. The rig emulates a subsea oil production network with gas-lift rate as the manipulated variables. The experimental results will be provided as the future work.

## 2. Feedback-optimizing control

Consider the optimization problem for the entire system built by a network of  $N$  subsystems. We assume each subsystem is optimized locally and that we always have active constraint. Thus, the steady-state optimization problem can be expressed as follows,

$$\min_{u_i, \forall i \in N} J_N = \sum_{i \in N} J_{N,i} \quad (1a)$$

$$s. t. \quad \mathbf{f}_i(\mathbf{x}_i, \mathbf{u}_i, \mathbf{d}_i) = 0 \quad i \in [1, N] \quad (1b)$$

$$\sum_{i \in N} g_i(\mathbf{x}_i, \mathbf{u}_i, \mathbf{d}_i) - g^{max} = 0 \quad (1c)$$

where  $\mathbf{x}_i$ ,  $\mathbf{u}_i$  and  $\mathbf{d}_i$  denote the vector of states, inputs and disturbances respectively. Constraint (1b) is related to the entire system model, and (1c) is an equality constraint.

### 2.1. Distributed Feedback-Optimizing Control using Online Primal Decomposition

To solve problem (1) using simple feedback controller, it is possible to construct a method based on online primal decomposition (Dirza et. al (2022(a))). The main motivation for this method is that we want to achieve optimal steady-state operation in a distributed manner, with minimal dynamic constraint violation and without demand of solving numerical optimization problems online. By introducing an initial value of local constraint  $g_i^{sp}$ , (1c) can be written as

$$g_i(\mathbf{x}_i, \mathbf{u}_i, \mathbf{d}_i) - g_i^{sp} = 0 \quad i \in [1, N] \quad (1d)$$

$$\sum_{i \in N} g_i^{sp} = g^{max} \quad (1e)$$

instead. As long as (1e) is satisfied, primal feasibility is guaranteed. Each subproblem estimates local Lagrange multipliers, which is used in central constraint controllers. These controllers update the setpoints iteratively, where the goal is to provide setpoints that satisfy the primal feasibility. In this paper we assume that the constraint is always active, therefore we have two types of approaches of updating the local setpoints. To ensure that (1e) is satisfied, one of the local setpoints is updated as follows,

$$g_N^{sp,k+1} = g^{max} - (g_1^{sp,k+1} + \dots + g_{N-1}^{sp,k+1}) \quad (2)$$

This is called the compensator subsystem, and the objective is to ensure primal feasibility. For the remaining subsystems,  $i = \{1, \dots, N-1\}$ , the local setpoints is updated by

$$g_i^{sp,k+1} = g_i^{sp,k} + K_{i,i}(-\lambda_1^k + \lambda_N^k) \quad (3)$$

We may use integrating controllers with integral gain  $K_{i,i} = \frac{1}{K_i(\tau_{c,i})}$ , where  $K_i$  is the step response gain and  $\tau_{c,i}$  is the closed-loop time constant. However, a proportional integral (PI) controller can also be used. The local Lagrange multipliers,  $\lambda_i$ , can be estimated

$$\lambda_i = -\nabla_{u_i} \hat{J}_{N,i} \left( \nabla_{u_i} \hat{g}_i(\mathbf{x}_i, \mathbf{u}_i, \mathbf{d}_i) \right)^{-1} \quad (4)$$

Where  $\nabla_{u_i} \hat{J}_{N,i}$  and  $\nabla_{u_i} \hat{g}_i$  are the estimated gradient of local cost and local setpoints respectively. In figure 2.1 the online primal decomposition framework is illustrated. The central constraint controllers, which contains both normal and compensator subsystems,

provide new set points for the local constraints. As the constraints always are active, these set points are considered as inputs to the subsystems. When there is a presence of disturbance, we can use the plants current information to estimate the current state and parameters by implementing a local dynamic estimator, for instance Extended Kalman Filter (EKF). By applying the inputs, estimated states and parameters, we can estimate cost gradient as well as constraint gradient to compute the local Lagrange multipliers. The multipliers are then implemented in the central constraint controller to calculate the new set points.

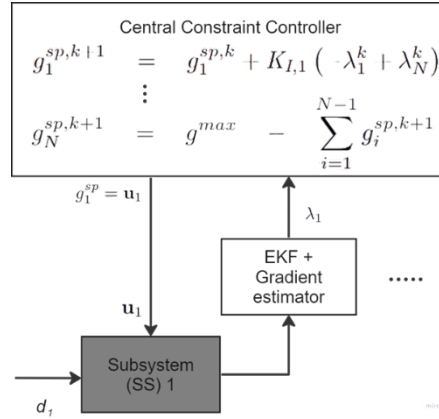


Figure 2.1 The online primal decomposition control structure.

### 3. Simulation descriptions

#### 3.1. Simulation model setup

To emulate the subsea gas-lifted oil production system, we use a model created in MATLAB R2021a based on a lab-scale experimental rig from a paper by Matias et al. (2022) and has been tested and used in earlier papers related to the lab-rig. It is implemented with noise similar to what is expected from the lab-rig. The reservoir section of the model is implemented by valve openings  $\mathbf{p} = [\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3]$  which determine the liquid flow. These valves represent the disturbances from the reservoir. In the MATLAB model the gas-lift flow controllers are given  $\mathbf{u}_i$  as setpoints. They are implemented as a 5 second delay plus input noise to simulate controller action instead of coding the actual controllers. The setpoint controllers in the central constraint controller are integral controllers tuned using SIMC tuning rules (Skogestad (2003)).

#### 3.2. Optimization problem setup

The objective of the optimization problem in this model setup is to maximize the liquid flow rate, which equals the sum of the liquid production of the three wells, with a limited amount of gas-lift injection, which is input shared constraint. Considering problem (1). the economic objective can be expressed as below,

$$J(\mathbf{u}, \mathbf{p}) = \sum_{i=1}^3 f_i(u_i, p_i) = -20Q_{l,1}(u_1, p_1) - 25Q_{l,2}(u_2, p_2) - 30Q_{l,3}(u_3, p_3) \quad (5)$$

where  $Q_{l,1}$ ,  $Q_{l,2}$  and  $Q_{l,3}$  are the produced liquid flow rates of well 1, 2 and 3 respectively. We assume different values of the hydrocarbon flows as shown in eq. 6, to illustrate how different values affect the behavior of the subsystems. The input vector is defined as  $\mathbf{u} = [Q_{gl,1}^{sp} \quad Q_{gl,2}^{sp} \quad Q_{gl,3}^{sp}]^T$  where  $Q_{gl,1}^{sp}$ ,  $Q_{gl,2}^{sp}$  and  $Q_{gl,3}^{sp}$  are the gas-lift set points from the central controller of well 1, 2 and 3 respectively. In addition, the reservoir valve  $\mathbf{p}$  is time varying.

#### 3.3. Comparative method

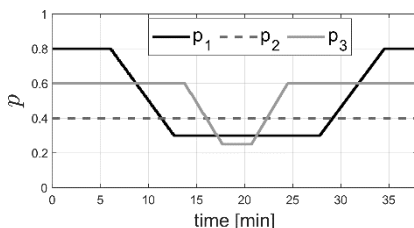
We consider our previous work, distributed feedback-based optimization with dual decomposition, to compare with primal decomposition. The implemented dual decomposition method is based on Dirza et al. (2022(b)) and Krishnamoorthy et al. (2021). In figure 3.1 the implemented dual decomposition framework is shown.

#### 4. Results and Discussion

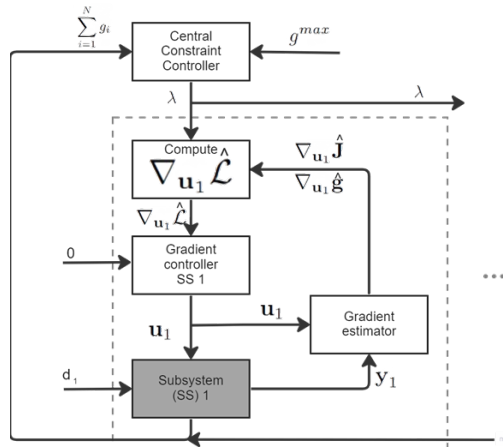
Figure 4.1 shows the disturbance in this simulation, which corresponds to the reservoir valve openings  $\mathbf{p}$  in the lab-rig model. The first disturbance occurs when  $p_1$  gradually decreases from  $t = 6$  to  $t = 12.5$  min. During this interval, we expect the gas-lift injection in well 1 to decrease, and redistribution of the gas to the other wells. The second disturbance occurs when also  $p_3$  gradually decreases from  $t = 14$  to  $t = 18$  min. As before, we expect that the gas-lift injection rate in well 3 will go down. At the same time, we expect the other wells will gain a higher gas-lift injection. The third and fourth disturbance occur from  $t = 21$  to  $t = 24.5$  and  $t = 28$  to  $t = 34.5$  respectively. During this time  $p_3$  and  $p_1$  are gradually increased back up to the initial values. This is because we want to see how the controllers behave for both decrease and increase in the disturbance.

In this paper subsystem 3 has been used as the compensator for the primal decomposition. This subsystem was chosen because it has the highest gain magnitude, which results in most profit for cases with active shared constraint.

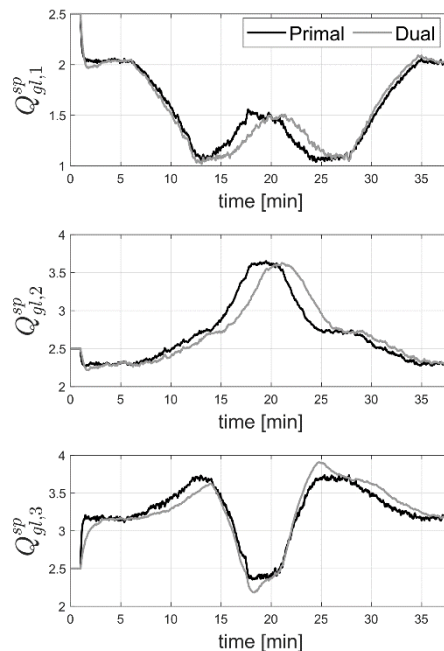
In figures 4.2 – 4.7 we compare the simulation results from primal decomposition and dual decomposition. The calculated input setpoints are shown in figure 4.2, the actual gas-lift flow rate will deviate slightly from the setpoints due to the implemented measurement noise in the model. Another cause is how the gas flow rate controllers are implemented in the model, which is described in section 3.1.



**Figure 4.1** Disturbance profile during the simulation.

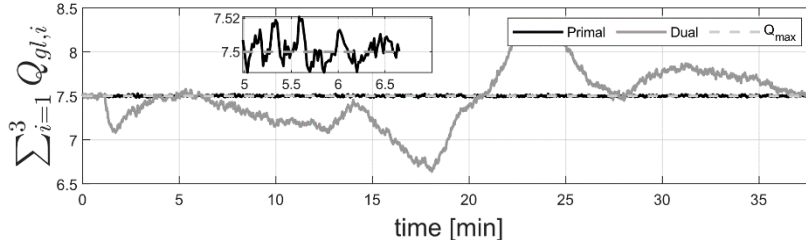


**Figure 3.1** Block diagram of dual decomposition control structure for one well. The area within the dashed gray box is duplicated  $N$  times. For a more detailed version the reader is referred to Dirza et al. (2022b).



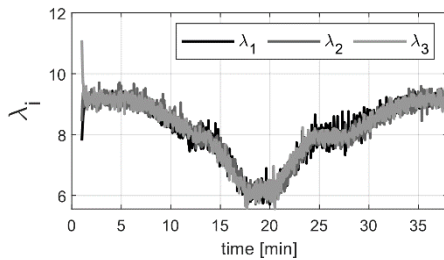
**Figure 4.2** The gas-lift flow rate setpoint  $\mathbf{u}^{sp} = Q_{gl,i}^{sp}$  of every well due to disturbance from the simulation.

In figure 4.2 and 4.3 we see that the dual control is responding slower to the disturbances than the primal control for correcting the gas-lift setpoints. This is a consequence of the time-scale separation between the central controller and the gradient controllers needed in the dual structure (Dirza et al (2022b)).

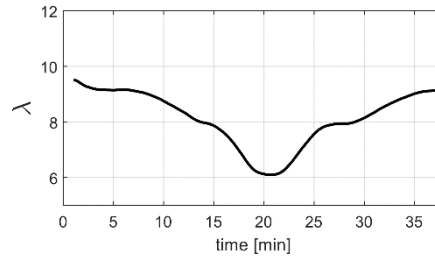


**Figure 4.3** The constraint satisfaction of both primal and dual. There is a magnifying plot in time window 5 to 6.5 min showing only constraint satisfaction for primal.

Figure 4.3 shows the constraint satisfaction. As expected, the Primal decomposition performs much better than the dual decomposition here. This is because the compensator system, subsystem 3, "absorbs" the deviations from the constraint. The absorbing ability comes from how the compensator set point is calculated, see eq. 2, and is the reason for the primal decomposition's capability to maintain the active constraint. In terms of what this means for the operation, with the use of primal decomposition we can run the system without any significant back-off and remain feasible at all times. On the other and, for the dual decomposition we must implement back-off, especially from  $t = 21$  when the disturbances start to increase again.



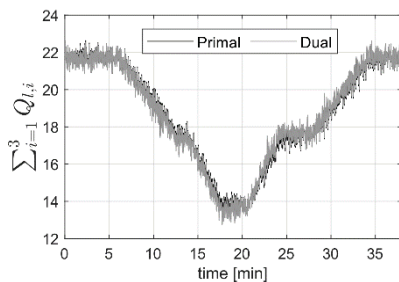
**Figure 4.4** Local Lagrange multipliers for primal decomposition.



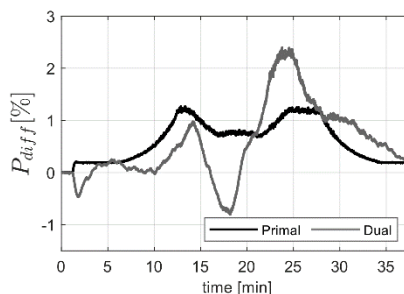
**Figure 4.5** Lagrange multiplier for dual decomposition

Figure 4.4 depicts that the local Lagrange multipliers for Primal Decomposition converge to the same value. Figure 4.5 show the Lagrange multiplier for the dual decomposition. We can observe that the multiplier converges slower and is smoother for this method. This is because the central controller for dual decomposition operates on a slower timescale. When the Lagrange multiplier converges around  $t = 3.5$  we can see that the active constraint is controlled almost as good as for the Primal case. It is also worth to notice that the Lagrange multipliers converges to the same value for both Primal and Dual.

To analyze the optimization performance of primal and dual decomposition, we compare the profit obtained by the two methods with a naive approach, where it is considered fixed inputs,  $\mathbf{u}_i = Q_{gl}^{max} / 3$ . The naive approach represents the case where no information is available, therefore, the best approach is to divide the available gas equally among the



**Figure 4.6** Total liquid flow of the system



**Figure 4.7** Instantaneous profit compared to naive approach.

wells. This additional approach is used as benchmark to show how primal and dual decomposition compares to not do any optimization at all. To display the performance, we plot the difference in percentage between the instantaneous profit of primal/dual and the naive approach. The difference is calculated as  $J_{diff} = \frac{J - J_{naive}}{J_{naive}} \cdot 100$ , where  $J$  is the profit of the method of interest and  $J_{naive}$  is the profit of the naive approach. Figure 4.7 shows that primal is more profitable than dual until  $t = 21$  min, after this dual appears to be favorable. However, in figure 4.3 we see that after  $t = 21$  dual does not achieve primal feasibility and therefore the profit here is not viable.

## 5. Conclusion

In this work, we have done simulation with uncertainty and measurement noise to compare feedback-based real-time optimization with online primal and dual decomposition. Based on the results we can conclude that primal decomposition is able to effectively ensure primal feasibility for separable systems with an active shared input constraint. It performs better than dual decomposition in this aspect, mainly because of the timescale difference of the central constraint controllers. This significantly reduces the need for any significant back off, which results in more profitable operations. While primal ensures constant feasibility, it is less general than dual. As the continuation of this work, we consider obtaining the experimental result, implementing dual with override and develop a primal structure able to switch between active constraints.

## References

- Dirza, R., Rizwan, M., Skogestad, S., & Krishnamoorthy, D. (2022a). Real-time Optimal Resource Allocation using Online Primal Decomposition. *IFAC-PapersOnLine*, 55(21), 31-36.
- Matias, J., Oliveira, J. P., Le Roux, G. A., & Jäschke, J. (2022). Steady-state real-time optimization using transient measurements on an experimental rig. *Journal of Process Control*, 115, 181-196.
- Dirza, R., Matias, J., Skogestad, S., & Krishnamoorthy, D. (2022b). Experimental validation of distributed feedback-based real-time optimization in a gas-lifted oil well rig. *Control Engineering Practice*, 126, 105253.
- Krishnamoorthy, D. (2021). A distributed feedback-based online process optimization framework for optimal resource sharing. *Journal of Process Control*, 97, 72-83.
- M. Morari, Y. Arkun, G. Stephanopoulos, 1980. Studies in the synthesis of control structures for chemical processes: Part i: Formulation of the problem. process decomposition and the classification of the control tasks analysis of the optimizing control structures. *AIChE Journal* 26, 220 – 232.
- Skogestad, S. (2003). Simple analytic rules for model reduction and PID controller tuning. *Journal of Process Control*, 13(4), 291–309.

# Superstructure configuration and optimization of the biofuel production perspectives in Austria

Safdar Abbas,<sup>a</sup> Ada Josefina Robinson<sup>a</sup>, Stavros Papadokonstantakis<sup>a</sup>

<sup>a</sup>*Institute of Chemical, Environmental and Bioscience Engineering  
TU Wien, 1060 Wien, Austria  
[Safdar.abbas@tuwien.ac.at](mailto:Safdar.abbas@tuwien.ac.at)*

## Abstract

This paper presents a superstructure model of mid to long-term biofuel production systems in Austria and determines the optimal configurations that simultaneously satisfy the profit maximization and environmental impact minimizations. A Linear Programming (LP) mathematical model formulation is developed based on data collected from different studies on the biofuel production process. The superstructure-based model considers the availability of Austrian second-generation feedstock, conversion technologies, and final products and by-products. The economic objective function considers the total net profit. The environmental objective function considers the overall environmental impact measured through the Prospective life cycle assessment. The multi-objective model is solved in GAMS software using the  $\epsilon$ -constraint method. The results demonstrate the solution for minimum Environmental Impact is zero and profit also zero which means nothing is produced then the model solved for maximum profit which is 284 M€/Year and an overall Environmental Impact is 0.67 Mt of CO<sub>2</sub>-Eq.

**Keywords:** Biomass, Biofuels, Prospective Life Cycle Assessment (PLCA), Superstructure, Optimization.

## 1. Introduction

Renewable energy resources are emerging as an attractive option for ensuring future energy security around the world. Biomass has been demonstrated to have the potential for fulfilling energy and chemical needs as a good alternative to petroleum because of its rising prices and uncertain availability. There has been a lot of concern about the sustainability of several first-generation biofuels, which has led to discussions about the possibility of producing biofuels from non-food crops, known as second-generation biofuels (Sharma et al., 2013). Second-generation biofuels offer many potential benefits, such as converting waste oils into fuel, reclaiming abandoned land and reducing food crop competition. Also, second-generation feedstock made from locally available sources can be very beneficial to rural and emerging regions of a country's economy. A high yielding feedstock ensures a reasonable price for commercial scale biofuel production. This is a very important aspect of feedstock selection. The benefits of biomass over other renewable energy sources are that it can be stored for long periods of time until needed and that it is an excellent alternative to liquid fuels in today's transportation system. Fuels obtained can be blended to replace classic petroleum-based fuels, or used as (without blending) in existing vehicles (Kowalski et al., 2022).

Biofuels are rapidly gaining interest in society and policies for secure, sustainable, and circular economy concepts, and thus designing and optimizing the entire biorefinery value chains from various biomass feedstock types to biofuel production and end-use has become a significant research challenge (Yue et al., 2014). Superstructure-based modelling and optimization is a systematic approach when considering the commercial feasibility and perspectives of biofuel production, diverse technological options, supply chain, market scenarios, and environmental objectives. Superstructure based approach also helps to setting the boundaries for the scope of the analysis at a country level (e.g., for Austria in this study) or for a union of countries (e.g., for EU, if extending the scope) to lead the results of higher practical importance for decision makers and industrial stakeholders that can influence policies and investments, respectively. This approach is typically influence by spatial aspects with respect to feedstock supply chains and market conditions. To this end, currently the first-generation feedstock is used to produce biofuel in Austria but their overall carbon footprint, land-use change, food availability, price, and biodiversity have all been criticized (The European Parliament and the Council of the European Union, 2015).



## 2. Problem Statement

The problem addressed in this work concerns the superstructure-based model of mid to long-term second-generation biofuel production system in Austria while optimizing two objectives: the profit and the environmental impact. Figure 1 shows the superstructure-based model accounting the feedstock available in Austria according to RED II Annex IX, conversion pathways, final products, and by-products. The conversion technologies have different technology module because these are designed using a different catalyst (Morales et al., 2021). The environmental objective function considers the overall environmental impact in terms of a global warming potential including feedstock extraction, processing and final disposal of products. The economic aspects are considered in terms of total net profit. The goal of this study is to determine the optimal configuration of second-generation biofuel production system in Austria that simultaneously maximize the profit and minimize the environmental impact.

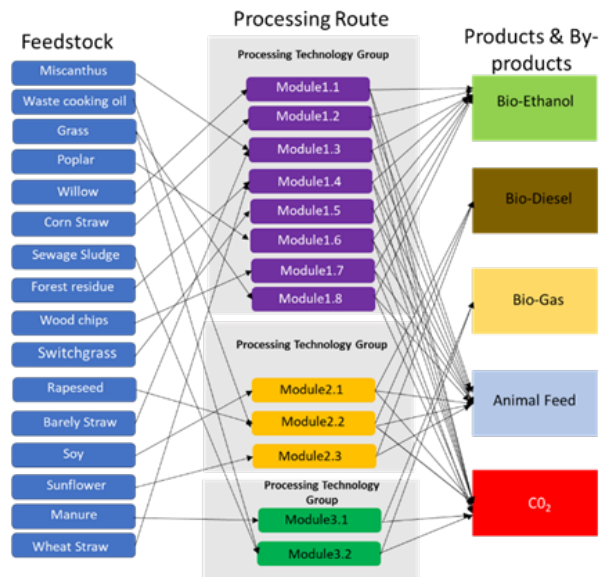


Figure 1: Superstructure for the second-generation biofuel production system

## 3. Model Formulation

The mathematical formulation is developed based on the work performed by Santibañez-Aguilar et al., 2011 to determine the optimal configuration as presented in Table 1. The mathematical model is coded in the GAMS software and solved using the solver CPLEX. The  $\epsilon$ -constraint method (Soroudi, 2017) is used to determine the set of optimal solutions that satisfy both objectives and that are used to build the Pareto front. In order to solve the multi-objective problem by  $\epsilon$ -constraint method, first the maximum and minimum values of the profit are calculated. These border values are determined by maximizing and minimizing profit separately as the objective function. Then, the profit is assumed as a constraint, and the environmental impact is calculated as the sole objective function. A Pareto optimal front is then derived for both objective functions.

### 3.1. Economic Analysis

The economic analysis considered in terms of profit is expressed in Table 1. The  $C_p^{value}$  is the selling price of products and  $C_b^{value}$  is the selling price by-products in €/Ton (IEA, 2022). The annual operating cost involve mainly the feedstock cost in Austria (Ruiz et al., 2019) and utilities consumed during the processing (including energy, catalyst, solvents, etc.). The total capital investment is calculated by process blocks build up method. This method is adopted from Tsagkari et al., 2020 and it consists of a modular approach to estimate the processing cost of different blocks inside the battery limit. The cost of each process block is calculated by the

power law equation (Tsagkari et al., 2020). The sum of the process block costs is the total capital investment. The equation, which follows the power law functions, shows a decrease in unit production costs with an increase in plant capacity. Therefore, nonlinear cost functions are linearized over different intervals between given processing limits to obtain a linear function.

Table1: Mathematical formulation of superstructure configuration and optimization of the biofuel production.

<b>Sets, Positive Variables, Decisions Variables, Parameters, Objective Function, and Constraints</b>	
<p><b>Sets</b>  q: Feedstock  p: Products  b: by products  r: Processing routes</p> <p><b>Positive Variables</b>  <math>u_{pqr}</math>: Flowrate of products  <math>F_{pqr}</math>: Flowrate of feedstock</p> <p><b>Decision Variable</b>  Profit: Net annual gain  EI: Environmental Impact</p>	<p><b>Parameters</b>  Cost(p): Product Cost  Cost(q): Feedstock Cost  Cost(b): By-product Cost  Processing Cost  Max(q): Maximum available resources  Max(p): Maximum demand of products  EI(q): Environmental Impact of feedstock  EI(p): Environmental Impact of products  EI(r): Environmental Impact of processing routes  <math>\omega_{pqr}</math>: Conversion factor to product  <math>\gamma_{pqr}</math>: Conversion factor to by product</p>
<b>Objective Function</b>	
Objective Function = [max profit; min EI]	
Where,	
<b>Profit</b> = (Annual Sales – Annual Operating cost – Total Capital investment)	
<b>Annual Sales</b> = ( $\sum_p \sum_q \sum_r u_{pqr} C_p^{value} + \sum_p \sum_q \sum_r \sum_b \gamma_{pqr} C_b^{value}$ )	
<b>Total Capital Investment</b> = sum of the total capital investment cost of all process block	
<b>Total Capital cost of one process block</b> = $I_{rk}^{ref} \left[ \frac{\text{Output capacity of a new plant}}{\text{Output capacity of a reference plant}} \right]^{n_{r,k}}$	
<b>EI</b> = (LCA of feedstock + LCA of processing route + LCA of final product)	
<b>Constraints</b>	
$\sum_p \sum_r f_{pqr} \leq f_q^{max} \quad q \in Q$	A first set of inequality constraints are the maximum availability of the feedstock. it is restricted to use more than the existing quantity of feedstock for its processing to the corresponding final products.
$\sum_q \sum_r u_{pqr} \leq u_p^{max} \quad p \in P$	A second constraints is the product demand.This constraint helps to prevent the higher production rate than demand and also avoid to waste of sources.
$u_{pqr} \leq u_{pqr}^{max} \quad q \in Q, r \in R, p \in P$	As a final constraint, the maximum processing limit for feedstocks also considered. This is only for upper limit which are the maximum amount of feedstock for each processing route.

### 3.2. Environmental Analysis

The Environmental analysis includes the overall environmental impact measured through the prospective Life Cycle Assessment (PLCA) methodology. Prospective Life Cycle Assessment is useful tool to determine the environmental impact of current and emerging technologies in the future. The PLCA requires an inventory database that reflects the anticipated changes in technology and the environment at a particular point in time, based on a specific socio-techno-economic path. In this study PLCA contains an Ecoinvent 3.8 database along

with Integrated assessment model (IAM)(Sacchi et al., 2022). IAM describes transformation pathways of the interlinked energy-economy-land-climate system. IAM is based on Shared Socioeconomic Pathways (SSPs) which create a framework for the study of climate-related scenario outcomes. The background provides for quantification of SSP2 as a reference or 'marker' implementation (Fricko et al., 2017).

The SSP2 narrative describes a middle-of-the-road development in the mitigation and adaptation challenges space. There has been an extension of the historical experience in many dimensions of the SSP2 marker implementation, especially in terms of improving the baseline carbon and energy intensity. SSP2 can serve as a starting point to further explore integrated solutions for achieving multiple societal objectives in light of the climate adaptation and mitigation challenges that society could face over the 21st century (Fricko et al., 2017). An open source software Activity browser is used to conduct the overall environmental impact in this study. The PLCA is carried out using an attributional approach because the background scenario does not change. The goal of this LCA is to identify the environmental improvement of different technologies. The results of this analysis are used to build a superstructure optimization model. The results of Environmental impact analysis are presented in terms of GWP (kg CO<sub>2</sub> Eq) and are shown in Table 2 and 3, respectively.

Table 1: Environmental Impact Analysis for Each Feedstock using PLCA

Feedstock	GWP (Kg CO <sub>2</sub> Eq)	Feedstock	GWP (Kg CO <sub>2</sub> Eq)
Miscanthus	0.08	Switchgrass	0.08
Willow	0.051	Rapeseedoil	0.76
Poplar	0.072	Waste cooking oil	0.083
Wheatstraw	0.038	Soy	0.3
Cornstraw	0.086	Sunflower	0.55
Barleystraw	0.030	Manure	0
Forest residue	0.034	Sewage sludge	0.0099
Wood chips	0.039	Grass	0.1

Table 2: Environmental Impact Analysis for Each Processing Route PLCA

Feedstock	Processing Routes	GWP (Kg CO <sub>2</sub> Eq)	Feedstock	Processing Routes	GWP (Kg CO <sub>2</sub> Eq)
Miscanthus	Fermentation and saccharification	0.74	Switchgrass	Fermentation and saccharification	0.74
Willow	Fermentation and saccharification	0.32	Rapeseedoil	Transesterification	2.00
Poplar	Fermentation and saccharification	0.39	Waste cooking oil	Transesterification	0.38
Wheatstraw	Fermentation and saccharification	0.37	Soy	Transesterification	0.42
Cornstraw	Fermentation and saccharification	0.42	Sunflower	Transesterification	0.40
Barleystraw	Fermentation and saccharification	0.37	Manure	Biodigestion	0.67
Forest residue	Fermentation and saccharification	0.23	Sewage sludge	Biodigestion	0.40
Wood chips	Fermentation and saccharification	0.57	Grass	Biodigestion	0.16

#### 4. Results & Discussions

The proposed methodology is tested through a case study to establish a biorefinery system in Austria. This study involves the 16 types of biomass, mainly agriculture and forest residue available in Austria (Ruiz et al., 2019) used as a feedstock to obtained bioethanol, biodiesel and biogas as a product and cattle feed and carbon dioxide as by product. The conversion efficiency of processing route that converts the feedstock into products and byproducts were obtained by previous results (Morales et al., 2021). Additional required data such as selling price of products byproducts and demand of the product were taken from governmental institutions (IEA,2022; Biofuels,2022.).The solution obtained by solving this multi-objective problem is presented in Figure 2. First, the solution for the minimum profit is obtained when nothing is produced, so the profit is equal to zero which is shown at Point A in Figure 2. The solutions for point B and C in Pareto curve are shown in Figure 3 and 4 respectively. Regarding the production of bioethanol at point B 76% demand is fulfil while the biodiesel and biogas is only 20% and 2% of the total demand respectively. The

maximum profit at point B is 173.3 M€/Year and EI is 0.39 Mt of CO<sub>2</sub>-Eq. Figure 4 shows that 100% bioethanol demand is fulfilled with only 60% feedstock of woodchips available in Austria. The Production of biodiesel and biogas remains the same as point B. The overall Profit at point C is 284 M€/Year and EI is 0.67 Mt of CO<sub>2</sub>-Eq.

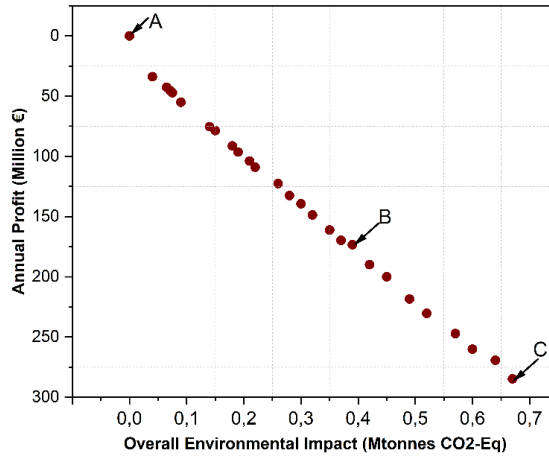


Figure 2: Pareto front of optimal solutions

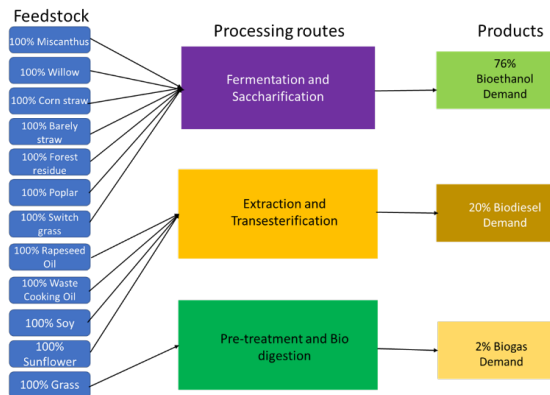


Figure 3: Solution for point B of Pareto front

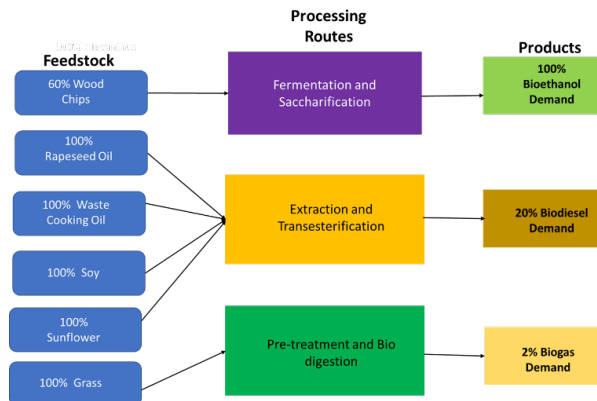


Figure 4: Solution for point C of Pareto front

## 5. Conclusions

This paper presents a systematic approach to the design of superstructure for optimal configuration of biofuel considering simultaneously maximum profit and minimum environmental impact. An Austrian second-generation feedstock mainly agriculture residue and forest residue considered. In this study we considered different constraints the maximum feedstock available, maximum demand of biofuel in Austria prospective. The multi-objective optimization is solved using the  $\epsilon$ -constraints method to establish the tradeoff among the economic and environmental objectives. Results shows that model can help to solve the problem of biofuel production from agriculture residue or forest residue.

## 6. References

- 2022, I. (2022) <https://www.iea.org/countries/austria>  
 Biofuels, I. (2022) <https://www.iea.org/fuels-and-technologies/bioenergy>
- Fricko, O., Havlik, P., Rogelj, J., Klimont, Z., Gusti, M., Johnson, N., Kolp, P., Strubegger, M., Valin, H., Amann, M., Ermolieva, T., Forsell, N., Herrero, M., Heyes, C., Kindermann, G., Krey, V., McCollum, D. L., Obersteiner, M., Pachauri, S., ... Riahi, K. (2017). The marker quantification of the Shared Socioeconomic Pathway 2: A middle-of-the-road scenario for the 21st century. *Global Environmental Change*, 42, 251–267. <https://doi.org/10.1016/j.gloenvcha.2016.06.004>
- Kowalski, Z., Kulczycka, J., Verhé, R., Desender, L., de Clercq, G., Makara, A., Generowicz, N., & Harazin, P. (2022). Second-generation biofuel production from the organic fraction of municipal solid waste. *Frontiers in Energy Research*, 10(August), 1–15. <https://doi.org/10.3389/fenrg.2022.919415>
- Morales, M., Arvesen, A., & Cherubini, F. (2021). Integrated process simulation for bioethanol production: Effects of varying lignocellulosic feedstocks on technical performance. *Bioresource Technology*, 328(December 2020), 124833. <https://doi.org/10.1016/j.biortech.2021.124833>
- Ruiz, P., Nijs, W., Tarvydas, D., Sgobbi, A., Zucker, A., Pilli, R., Jonsson, R., Camia, A., Thiel, C., Hoyer-Klick, C., Dalla Longa, F., Kober, T., Badger, J., Volker, P., Elbersen, B. S., Brosowski, A., & Thrän, D. (2019). ENSPRESO - an open, EU-28 wide, transparent and coherent database of wind, solar and biomass energy potentials. *Energy Strategy Reviews*, 26(June 2019), 100379. <https://doi.org/10.1016/j.esr.2019.100379>
- Sacchi, R., Terlouw, T., Siala, K., Dirnaichner, A., Bauer, C., Cox, B., Mutel, C., Daioglou, V., & Luderer, G. (2022). PRospective EnvironMental Impact asSEment (premise): A streamlined approach to producing databases for prospective life cycle assessment using integrated assessment models. *Renewable and Sustainable Energy Reviews*, 160(April 2021), 112311. <https://doi.org/10.1016/j.rser.2022.112311>
- Sharma, B., Ingalls, R. G., Jones, C. L., & Khanchi, A. (2013). Biomass supply chain design and analysis: Basis, overview, modeling, challenges, and future. In *Renewable and Sustainable Energy Reviews* (Vol. 24, pp. 608–627). Elsevier Ltd. <https://doi.org/10.1016/j.rser.2013.03.049>
- The European Parliament and the Council of the European Union. (2015). Amending Directive 98/70/EC relating to the quality of petrol and diesel fuels and amending Directive 2009/28/EC on the promotion of the use of energy from renewable sources. *Official Journal of The European Union*, 2014(September), 20–30.
- Tsagkari, M., Kokossis, A., & Dubois, J. L. (2020). A method for quick capital cost estimation of biorefineries beyond the state of the art. *Biofuels, Bioproducts and Biorefining*, 14(5), 1061–1088. <https://doi.org/10.1002/bbb.2114>
- Yue, D., You, F., & Snyder, S. W. (2014). Biomass-to-bioenergy and biofuel supply chain optimization: Overview, key issues and challenges. *Computers and Chemical Engineering*, 66, 36–56. <https://doi.org/10.1016/j.compchemeng.2013.11.016>

# Optimal operation of an industrial low pressure carburizing furnace

Fatima Matamoros<sup>a,b</sup>, Pierre-Alexandre Glaude<sup>a</sup>, Roda Bounaceur<sup>a</sup>, Hubert Monnier<sup>b</sup>, Abderrazak Latifi<sup>a,\*</sup>

<sup>a</sup> *Université de Lorraine, CNRS, LRGP, 54000, Nancy, France*

<sup>b</sup> *Institut National de Recherche et de Sécurité, 54500, Vandœuvre-lès-Nancy, France*  
[abderrazak.latifi@univ-lorraine.fr](mailto:abderrazak.latifi@univ-lorraine.fr)

## Abstract

This study presents the dynamic optimization of an industrial low-pressure gas carburizing (LPC) furnace. The objective is to determine the optimal operating conditions that minimize the production of harmful substances, i.e., polycyclic aromatic hydrocarbons (PAH) and soot, while meeting the industrial and mechanical requirements of the treated steel parts. A dynamic optimization problem is proposed and solved by means of a control vector parametrization approach and a gradient-based method using the software MATLAB. Optimized operating conditions are proposed and they are experimentally tested on an industrial LPC furnace. The results confirm that the proper treatment of steels parts is ensured and show that the PAH generated are reduced.

**Keywords:** Low-pressure gas carburizing, polycyclic aromatic hydrocarbons, dynamic hybrid optimization, control vector parameterization.

## 1. Introduction

Low-pressure gas carburizing (LPC) is a process that aims to improve the mechanical properties of steels by increasing the carbon content on their superficial layers. Since its introduction in the 20th century, LPC has become quite widespread in the heat treatment industry as it is an environmentally friendly process that reduces carburizing time and causes no intergranular oxidation of steel (Zajusz et al., 2014). This is achieved by carrying out the process at high temperatures and low pressures while supplying a cyclic feed of carbon-bearing gases like acetylene. Like other carburizing processes however, LPC has proven to be a significant source of occupational exposure to toxic substances like soot and polycyclic aromatic hydrocarbons (PAH) (Champmartin et al., 2017). The challenge is therefore to determine the best sequence of acetylene feed that minimizes the production of toxic compounds without compromising the carbon content constraints on the steel parts. The present paper addresses this challenge by formulating a hybrid dynamic optimization problem involving two modes (acetylene feed on and off), where each mode is characterized by model equations, a transition condition and transition functions. A control vector parametrization (CVP) approach is used to transform the problem into a nonlinear programming (NLP) problem that is solved by means of a gradient-based method. The results of the optimization and the experimental studies carried out to verify the validity of the findings are presented.

## 2. Materials and methods

The experimental installation used for the carburizing of gear box pinions (shown in

Figure 1) is presented in Figure 2. It consists of an industrial furnace made up of two independent chambers, i.e., a quenching chamber and a carburizing chamber, an acetylene and nitrogen feeding system, pressure and temperature control and regulation systems and a gas collection system consisting of a sampling rod and a cold trap connected to a pump. Nitrogen is used to inert the gas collection system to keep oxygen out of the setup.



Figure 1: gear box pinion  
(23MnCrMo5 steel)

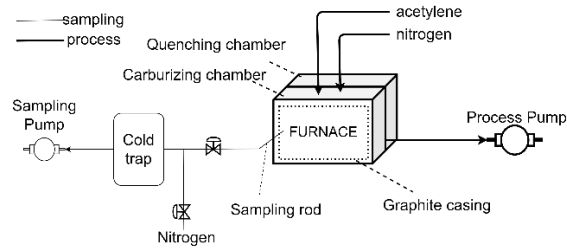


Figure 2 : experimental setup

The experiments consist in programming the operating conditions (the “recipe”) used for carburizing, i.e., temperature, pressure, gases inlet flowrates and feeding durations, then in loading the steel parts into the furnace. The gas collection system is used to collect the gases generated all throughout the process. At the end of carburizing, the metal parts undergo a 15-minute nitrogen quench under 15 bars of pressure. The contents of the cold trap are analyzed offline by gas chromatography coupled to a mass spectrometer (GC-MS). The results of the Vickers hardness tests conducted on the carburized steel parts to determine their micro hardness profiles are presented in this paper. Two recipes are used for the carburizing of the steel parts: an industrial recipe and a recipe obtained by solution of the optimization problem presented in the following section.

### 3. Formulation of the optimization problem

#### 3.1. Process model

The process model considers the formation of gaseous and solid species due to the thermal decomposition of acetylene, as well as the interactions between the fluid and solid phases. The pyrolysis of acetylene is modeled by means of a 140 species and 444 reactions kinetic mechanism obtained by reduction of the detailed mechanism proposed by Bensabath (2017). It is completed by the modeling of soot by means of the method of moments (Frenklach, 2002) and the modeling of pyrocarbon by assuming pseudo-homogenous deposition kinetics (Ziegler, 2004). The external transfer of acetylene to the surface of the steel and the diffusion of carbon in the steel are modeled by means of Fick’s second law. The interaction between the fluid and the steel is described by a Langmuir-Hinshelwood Hougen Watson model. Temperature and pressure are assumed to be constant throughout the reaction zone. The gas mixture is assumed to obey the ideal gas law and the reactor is assumed to be a perfectly mixed stirred tank reactor (PSTR). Under these assumptions, the mass balances on gas species and the differential equations for the moments of soot particles are given by equations (1) and (2) respectively:

$$\rho_g V \frac{dY_k}{dt} = \dot{m}_{in}(Y_{k,in} - Y_k) + \dot{m}_{prod,k} - Y_k(\dot{m}_{surf} - \dot{m}_{soot}) \quad (1)$$

$$\frac{dM_r}{dt} = R_r + G_r + W_r - \frac{M_r Q_{out}}{V} \quad (2)$$

where  $\rho_g$  is the mass density of the gas mixture,  $V$  is the volume of the reactor,  $Y_{kin}$  and  $Y_k$  are respectively the inlet mass fraction and the mass fraction of gaseous species  $k$  in

the reactor,  $\dot{m}_{in}$  is the inlet mass flowrate,  $\dot{m}_{prod,k}$  is the mass production rate of species  $k$  by all reactions,  $\dot{m}_{surf}$  is the total mass production rate of gaseous species by the reaction at the surface of the steel,  $\dot{m}_{soot}$  is the mass production rate of soot,  $M_r$  is the moment  $r$ ,  $R_r$ ,  $G_r$  and  $W_r$  are soot source terms due to the phenomena of nucleation, coagulation and surface growth respectively, and  $Q_{out}$  is the outlet volumetric flowrate. The mass balances in the fluid/solid boundary layer and in the steel lead to set of equations (3) and (4) respectively:

$$\begin{aligned} \frac{\partial [C_2H_2]}{\partial t} &= D_{C_2H_2} \frac{\partial^2 [C_2H_2]}{\partial z^2} \\ [C_2H_2](z, t_0) &= 0 \\ [C_2H_2](z = 0, t) &= [C_2H_2(PSTR)] \\ -D_{C_2H_2} \frac{\partial [C_2H_2]}{\partial z} \Big|_{z=z_L} &= \frac{V_p}{A_p} r_{surf} \end{aligned} \quad (3)$$

$$\begin{aligned} \frac{\partial \omega_C}{\partial t} &= D_C \frac{\partial^2 \omega_C}{\partial y^2} \\ \omega_C(y, t_0) &= \omega_{C,0} \\ -D_C \frac{\partial \omega_C}{\partial y} \Big|_{y=0} &= r_{surf} \frac{V_p}{A_p} \frac{100W_C}{\rho_p} \\ D_C \frac{\partial \omega_C}{\partial y} \Big|_{y=y_L} &= 0 \end{aligned} \quad (4)$$

where  $[C_2H_2]$  is the concentration of acetylene in the boundary layer,  $[C_2H_{2PSTR}]$  is the concentration of acetylene in the PSTR,  $\omega_C$  is carbon weight percentage in the steel,  $D_{C_2H_2}$  and  $D_C$  are the diffusion coefficients of acetylene in the gas phase and of carbon in the steel respectively,  $z$  and  $y$  are acetylene and carbon diffusion directions,  $W_C$  is carbon molar weight,  $V_p$ ,  $A_p$  and  $\rho_p$  are the volume, surface and density of the steel respectively, and  $r_{surf}$  is the volumetric rate of the reaction at the surface of the steel given by:

$$r_{surf} = k_{app} K_1 \frac{[C_2H_{2(surf)}]}{(1 + K_1 [C_2H_{2(surf)}] + K_2 [C_{(fe)}])^2} \quad (5)$$

where  $[C_2H_{2(surf)}]$  is the concentration of acetylene in the gas phase at the surface of the steel and  $[C_{(fe)}]$  is the concentration of carbon at the surface of the steel. Parameters  $K_1$  and  $K_2$  are Langmuir equilibrium constants of the adsorption of acetylene and carbon respectively, and  $k_{app}$  is an apparent kinetic constant. A hybrid model is proposed to take into account the phases where acetylene feeding is on (boost stages, designed by mode 1) and off (diffusion stages, designed by mode 2). The model equations on state variables  $x$  are given by:

$$\dot{x} = f^{(1)}(x) \quad (6)$$

$$\dot{x} = f^{(2)}(x) \quad (7)$$

where  $f$  are the relations described above. The transition conditions  $L$  and transition times  $t_{boost}$  and  $t_{diff}$  are given by:

$$L^{(2k)} = t_{boost}^{(k+1)} - t = 0, \quad \text{where } t_{boost}^{(k)} = \sum_{i=1}^{k-1} (a_i + b_i), \quad k = 1, \dots, N_C - 1, \quad (8)$$



$$L^{(2k-1)} = t_{\text{diff}}^{(k)} - t = 0, \quad \text{where } t_{\text{diff}}^{(k)} = a_k + \sum_{i=1}^{k-1} (a_i + b_i), \quad k = 1, \dots, N_C \quad (9)$$

where  $a_k$  and  $b_k$  are the lengths of boost and diffusion stage  $i$  respectively and  $N_C$  is the number of cycles, where each cycle corresponds to one boost stage followed by a diffusion stage. The transition functions are given by:

$$x_{\text{boost}}^{(k+1)}(t_0^{(k+1)}) - x_{\text{diff}}^{(k)}(t_f^{(k)}) = 0, \quad k = 1, \dots, N_C - 1 \quad (10)$$

$$x_{\text{diff}}^{(k)}(t_0^{(k)}) - x_{\text{boost}}^{(k)}(t_f^{(k)}) = 0, \quad k = 1, \dots, N_C \quad (11)$$

### 3.2. Performance index

The performance index corresponds to the total mass of PAH and soot generated:

$$J = \int_{t_0}^{t_f} (\dot{m}_{\text{HAP}} + \dot{m}_{\text{soot}}) dt = \int_{t_0}^{t_f} \dot{x}_{\text{obj}} dt = x_{\text{obj}}(t_f), \quad \text{where } x_{\text{obj}}(t_0) = 0$$

### 3.3. Decision variables and constraints

The decision variables considered are the operating conditions, i.e., the number of cycles  $N_C$ , the lengths of the boost and diffusion stages  $a_1 \dots a_{N_C}$  and  $b_1 \dots b_{N_C}$ , and the inlet flowrate of acetylene  $Q$ . The lengths of boost and diffusion stages vary in the [1-60] min range and acetylene flowrate in the [800-3500] NL/h range.

### 3.4. Constraints on the state variables

Constraints on the state variables related to the industrial requirements on the steel parts are considered and are given by the following relations:

$$\omega_c^{\text{surf}}(t_f) = 0.6 \% \quad (12)$$

$$\omega_c^{0.4 \text{ mm}}(t_f) = 0.34 \% \quad (13)$$

$$\omega_c^{\text{surf}}(t) \leq 1.2 \%, \quad t_0 \leq t \leq t_f \quad (14)$$

where  $\omega_c^{\text{surf}}$  and  $\omega_c^{0.4 \text{ mm}}$  are the carbon weight percentage at the surface of the steel and a depth of 0.4 mm respectively,  $t_f$  is the duration of the process and the value of 1.2 % corresponds to the saturation of carbon in the steel at the current operating conditions.

### 3.5. Optimization algorithm

To solve the optimization problem, it is necessary to deal with some difficulties that arise due to the nature of the decision variables. Firstly, the number of cycles  $N_C$  is an integer variable, making the optimization problem a mixed integer dynamic optimization problem. To overcome the ensuing difficulty in trying to solve such a problem, a decomposition of the original problem is proposed. Various sub-problems are thus established and for each one, the number of cycles is fixed and is thus not considered as a decision variable. Several sub-problems are then solved, each with a different number of cycles. The optimal number of cycles is then determined by selecting the sub-problem that leads to the minimum of the optimization criterion. Another difficulty that arises is the fact that the inlet flowrate of acetylene  $Q$  is a time-dependent variable of infinite dimension. A CVP approach is adopted to transform the problem into a finite dimension problem. The time interval is therefore divided into  $2N_C$  segments and  $Q$  is approximated by a piecewise constant function.

$$Q(t) = Q_k, \quad t \in [t_{\text{boost}}^{(k)}, t_{\text{diff}}^{(k)}], \quad k = 1, 2, \dots, N_C \quad (15)$$

$$= 0 \quad \text{otherwise}$$

The resulting NLP problem is solved by means of *fmincon* optimizer in the MATLAB environment. The hybrid model is integrated by means of *odes15s* solver.

## 4. Results and discussion

### 4.1. Optimization results

Figure 3 shows the results of the optimization : Figure 3(a) shows the performance index (PI), the quantity of acetylene consumed and the carbon integrated into the steel parts; Figure 3(b) shows the total durations of boost and diffusion stages. Six different sub-problems were considered and  $N_c$  varies between 5 and 10. For each  $N_c$ , the results are normalized by the results obtained with the industrial recipe R0.

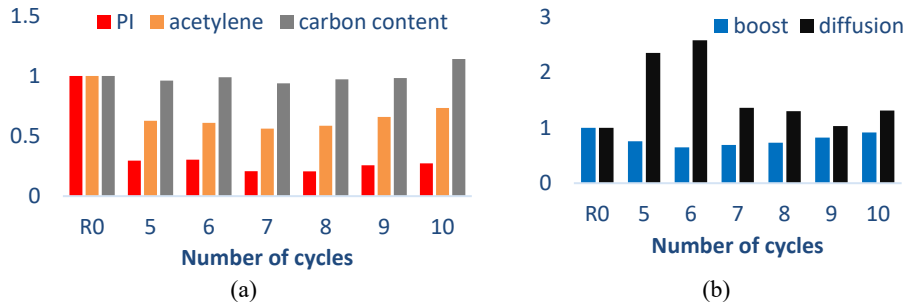


Figure 3 : Results of the optimization of the different sub-problems and predictions with the industrial recipe R0

The results show that the optimization of the different sub-problems leads to a considerable reduction of the optimization criterion when compared to the industrial recipe (Figure 3 (a)). This can be directly related to the reduction of the quantity of acetylene that is fed to the reactor, as the toxic substances are generated from its thermal decomposition. However, even if less acetylene is used, the carbon added to the steel parts is of the same order of magnitude for all sub-problems and for R0. This suggests that acetylene is in excess with R0 and thus decomposes to produce PAH and soot. Figure 3(b) shows that all sub-problems lead to an increase in the total duration of the diffusion stages and a decrease in the total duration of boost stages when compared to R0. Recipes with shorter boost times and longer diffusion times thus lead to a decrease of toxic compounds.

### 4.2. Experimental results

Among the different sub-problems, the operating conditions of sub-problem  $N_c = 7$  were chosen as the optimal conditions. Compared to the industrial recipe, the total process duration is 33% longer but leads to a 44% decrease in acetylene consumption. The results of the hardness tests conducted on a pinion carburized with the industrial recipe and one carburized with the optimized recipe are shown in Figure 4(a) and Figure 4(b) respectively. The tests were performed according to different directions along the teeth of the pinions to study the homogeneity of carburizing in space. The results show that both recipes lead to homogeneous carburizing and to very similar hardness profiles. The results obtained for PAH production with both recipes are shown in Figure 4(c). They show a clear decrease in the production of PAH with the optimized recipe as compared to the industrial one. The experimental results obtained thus show that it is possible to significantly reduce toxic compounds without compromising the industrial quality of steel parts.

## 5. Conclusions

In this study, a first-principles modeling and optimization methodology was proposed to optimize an industrial low-pressure gas carburizing furnace. Experiments were performed on an industrial scale to quantify PAH formation and to evaluate the quality of the steel parts carburized with an optimized recipe. The results show that it is possible to ensure

industrial quality while reducing toxic compounds by carefully selecting the acetylene feed profile. In future works, although the objective will remain the minimization of the process toxicity, the methodology adopted here will be used to consider other aspects to ensure the feasibility of the approach from an industrial viewpoint. Multi-objective optimization will thus be applied to deal with different case studies, e.g., problems that consider the total process duration or even the deposition of pyrocarbon on the furnace walls. Indeed, the former entails higher energy demands, and the latter leads to an increased number of maintenance operations, which not only expose the operators to toxic compounds, but also represent higher operating expenses.

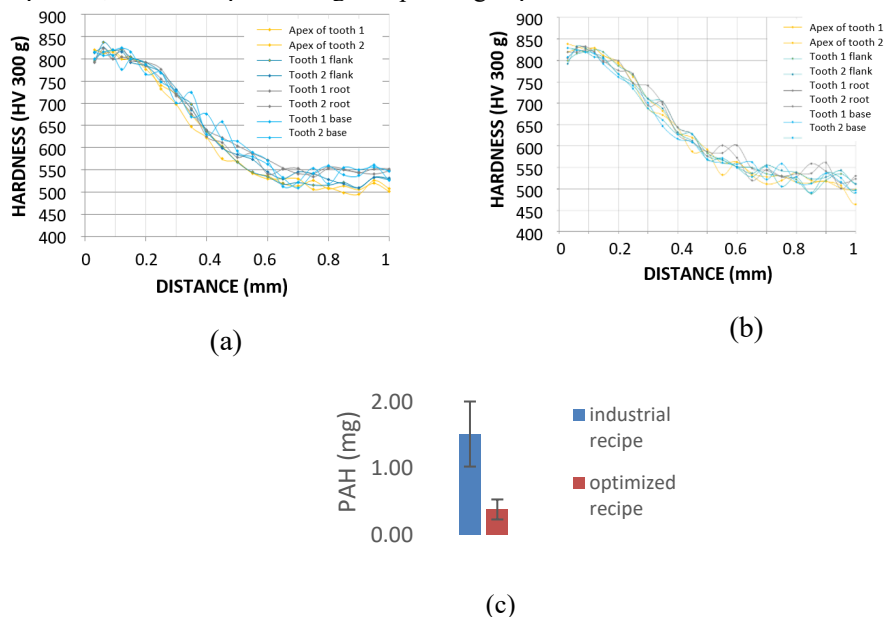


Figure 4: Experimental results (a) hardness profile of a pinion carburized with the industrial recipe (b) hardness profile of a pinion carburized with the optimized recipe and (c) PAH generated with the industrial and optimized recipes

## References

- T. Bensabath, 2017, Approche préventive pour une réduction des Hydrocarbures Aromatiques Polycycliques (HAP) dans les fours à pyrolyse, PhD thesis, Université de Lorraine, France.
- C. Champmartin, F. Jeandel, H. Monnier, 2017, Maintenance of Low-Pressure Carburising Furnaces: A Source of PAH Exposure. *Annals of Work Exposures and Health*, 61(3), 321–332, <https://doi.org/10.1093/annweh/wxw024>.
- M. Frenklach, 2002. Method of moments with interpolative closure. *Chemical Engineering Science*, 57(12), 2229–2239, [https://doi.org/10.1016/S0009-2509\(02\)00113-6](https://doi.org/10.1016/S0009-2509(02)00113-6).
- M. Zajusz, K. Tkacz-Śmiech, M. Danielewski, 2014, Modeling of vacuum pulse carburizing of steel. *Surface and Coatings Technology*, <https://doi.org/10.1016/j.surfcoat.2014.08.023>.
- Ziegler, 2004. Modélisation cinétique des dépôts de pyrocarbone obtenus par pyrolyse d'hydrocarbures (Issue 2004INPL045N) [Université de Lorraine]. <https://hal.univ-lorraine.fr/tel-01750078>.

# Development of a holistic Python package for optimal selection of experimental design criteria in kinetic model discrimination

Maerthe Theresa Tillmann<sup>a,b</sup>, Federico Galvanin<sup>a\*</sup>

<sup>a</sup> *Department of Chemical Engineering, University College London,  
Torrington Place, London WC1E 7JE, UK*

<sup>b</sup> *Faculty of Mechanical Engineering, RWTH Aachen University,  
Eilfschornsteinstraße 18, 52062 Aachen, Germany*

\* *f.galvanin@ucl.ac.uk*

## Abstract

Starting with a candidate set of kinetic models for a reaction, model-based design of experiment (MBDoE) techniques can be used to determine experimental conditions for fast model identification of reaction kinetics using the minimum number of experimental runs to specify both model structure and corresponding parameters. However, practically, determining optimal settings, including criteria and selection methods for model discrimination for efficient model identification under consideration of parametric uncertainty in the whole identification procedure is still an open and challenging task.

In this work, a holistic Python package is presented which comprises MBDoE for model discrimination and, subsequently, MBDoE for parameter precision. The new package is tested on in-silico experiments for the identification of a Baker's Yeast model to evaluate and compare: *i*) the total number of experiments required for kinetic model identification considering different experimental design criteria; *ii*) the rate of correct model selections using different model selection methods.

**Keywords:** model-based design of experiments, model discrimination, experimental design criteria

## 1. Introduction

In chemical engineering, kinetic models are used to mathematically describe and predict the time-dependent behavior of reaction systems whereby mechanistic models, in contrast to data-driven models, can provide easy interpretability and reliable extrapolations. Starting with several candidate model formulations and assuming insufficient information to statistically determine the best formulation among the proposed ones, model-based design of experiment (MBDoE) can be a promising strategy to select further experimental conditions to facilitate model identification (Asprey and Macchietto, 2000). In sequential MBDoE strategies, the optimal experimental design problem is formulated as an iterative optimization problem through the maximization of a design criterion related to model identification based on the current state of knowledge from collected data. The MBDoE strategy reflects the motivation of pursuing model identification with the minimum number of experiments to avoid experimental costs, resources and time.

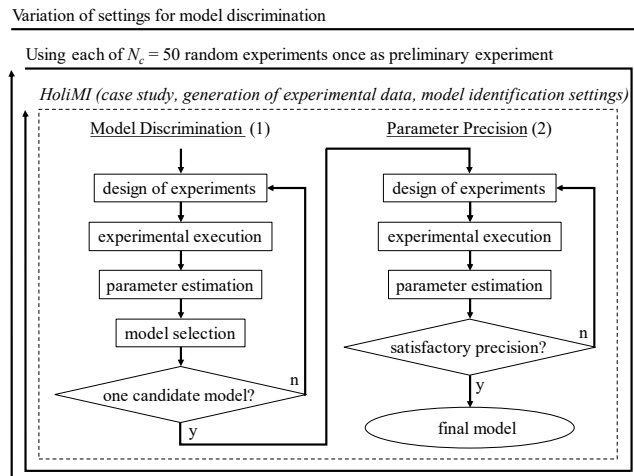
The classical procedure for model identification using MBDoE according to Asprey and Macchietto (2000) comprises three stages: *i*) a theoretical preliminary identifiability analysis of the set of proposed model candidates, *ii*) MBDoE for model discrimination

and *iii*) parameter precision for the selected model determined in *ii*). Several MBDoE design criteria for model discrimination, used in stage *ii*) as design objective function, have been proposed in literature in terms of statistical approaches of discrimination and complexity of the formulation (Michalik et al., 2010, Olofsson et al., 2019). Multiple MBDoE design criteria also exist for improving parameter precision, but the practical differences have been discussed in literature (Franceschini and Macchietto, 2008). So far, comparative studies of different design criteria for model discrimination (Burke et al., 1997, Olofsson et al., 2019) are limited to the stage of model discrimination only, so that the impact of parametric uncertainty on the whole identification procedure, and thus the effect on the number of experiments required in stage *iii*), is neglected. Practically, selection of settings for model identification is necessary, including experimental design criteria and probabilistic model selection settings. However, it is generally complex to determine which settings for the stage of model discrimination would provide the best prerequisites for model identification with the minimum number of experiments.

In order to address the question of optimal settings for model discrimination, a Python package is presented in this paper where different MBDoE criteria for model discrimination and for parameter precision are implemented as well as different model selection strategies from literature (cf. Olofsson et al., 2019). Applying this tool to the case study of yeast fermentation demonstrates a new holistic approach for comparing MBDoE design criteria and model selection settings to achieve fast model identification with the minimum number of experimental runs.

## 2. Model Identification using Model-Based Design of Experiments

The proposed framework of comparative study using the holistic model identification function (*HoliMI*) implemented in the Python package *HoliMI* is illustrated in Figure 1. For each combination of settings, the model identification function (*HoliMI*) is called repeatedly considering different scenarios for preliminary experiments.



**Figure 1** Proposed framework of comparative study and model identification function (*HoliMI*) implemented in the Python package including variants of settings for model discrimination.

The sequence of operations implemented as main function (*HoliMI*) in the Python package follows the MBDoE procedure for holistic model identification presented in Asprey and Macchietto (2000). Function inputs are: *i*) specifications of the case study, *ii*) configurations defining in-silico experiments and *iii*) settings of the implemented

sequence of operations for model identification. In the comparative study, settings for *iii*) related to model discrimination (criteria, model selection settings) are investigated.

### 2.1. Bioprocess Case Study

The fermentation of Baker's Yeast is considered as in-silico system for the comparative study (Chen and Asprey, 2003). The best approximating model structure, which describe the biomass ( $X$ ) and substrate ( $S$ ) concentration in a fed-batch reactor, i.e.  $\hat{\mathbf{y}} = [\hat{S}, \hat{X}]$ , should be selected among four proposed candidates. An experimental budget of  $N_{\text{exp max}} = 10$  experiments limits the model identification task. Each experiment is defined by a design vector  $\boldsymbol{\varphi} = [D, S_f, X_0]$  including dilution rate  $D$  [ $\text{h}^{-1}$ ], substrate concentration in the feed  $S_f$  [g/L] and initial concentration of biomass  $X_0$  [g/L] while samples are taken at fixed time points  $\mathbf{t}_{\text{sp}} := [10 \text{ h}, 20 \text{ h}, 30 \text{ h}, 40 \text{ h}]$ . Candidate model formulations, initial conditions of the system, initial parameter guesses and experimental design space are chosen in accordance with Chen and Asprey (2003).

### 2.2. In-Silico Experiments

For the generation of in-silico experiments, Monod kinetics including a term for biomass mortality are assumed as true model of the system. In addition to the mathematical solution of the model, a gaussian-distributed measurement error affects the true model predictions, i.e. the in-silico generated experimental data is defined by  $\mathbf{y} = N(\hat{\mathbf{y}}(\boldsymbol{\varphi}), \boldsymbol{\Sigma})$ , where  $\boldsymbol{\Sigma} := \begin{pmatrix} 0.12 & 0 \\ 0 & 0.12 \end{pmatrix}$  is the time-constant variance-covariance matrix used.

### 2.3. Sequence of Operations for Model Identification

The holistic model identification function (*HoliMI*), see Figure 1, comprises the iterative procedure for model discrimination (step 1) and the subsequent iterative procedure for parameter precision of the selected model (step 2) which approximates the experimental data best among the proposed candidates. Using MBD<sub>oE</sub>, experiments are designed such that the design criterion  $\Psi$  for model discrimination is maximized in each iteration and the specific formulation of the design criterion for model discrimination is defined by the settings as function input. Based on the set of experimental data, the parameters of each model are then estimated so that the weighted sum of squared residuals  $\chi^2$  as measure for the lack of fit is minimized (Chen and Asprey, 2003). The lack of fit  $\chi^2$  is then used to assign the models a probability  $p$  of being the best approximating model structure. These relative probabilities are used by Schwaab et al. (2006) to reject model candidates with low probabilities. If one of the model probabilities reaches 99.9%, model discrimination is terminated and parameter precision for the selected model is pursued until the confidence interval of all parameter estimates fall below 0.05. D-optimality is used as design criterion for the subsequent MBD<sub>oE</sub> for parameter precision.

## 3. Comparative Study

Besides following a holistic approach for comparing the number of experiments required for correct model identification including a precise parameter estimation, the comparative study is characterized by taking different parametric initializations for the first MBD<sub>oE</sub> iteration into account. Results are averaged over multiple runs considering accurate and inaccurate parameter estimates scenarios based on preliminary system knowledge.

### 3.1. Characterization of Preliminary Knowledge on Model Parameters

In order to provide the most general comparison of settings for model identification, numerous scenarios to describe the preliminary knowledge on model parameters are considered. A set of 50 randomly selected experimental conditions and corresponding measurements from in-silico experiments is generated. For each combination of settings,

the model identification function (*HoliMI*) is called  $N_c = 50$  times to consider a different preliminary experiment from the generated set in each realization so that different scenarios of parametric uncertainty in the subsequent MBDoE steps are considered.

### 3.1.1. Settings Used in the Comparative Study

Investigated settings in the comparative study are *i*) the MBDoE design criterion used for model discrimination as objective function of optimization problem; *ii*) the probability calculation method for model selection; *iii*) the probability-based model rejection method. The design criterion of Hunter and Reiner (1965), Eq. (1), is a summation of differences between model predictions over all pairs of  $M$  candidate models. In contrast to this basic formulation, the criterion of Box and Hill (1967), Eq. (2), takes model probabilities  $p$  and the variance of model predictions  $\hat{\mathbf{V}}$  into account. A statistically different approach has been proposed by Buzzi-Ferraris and Forzatti (1990), Eq. (3). Michalik et al. (2010) have derived the Akaike's weights design criterion, Eq. (4), from the Akaike Information Criterion of the candidate models with  $N_\theta$  as number of estimated model parameters. Recently, Olofsson et al. (2019) presented the Jensen-Rény divergence as design criterion, Eq. (5), for  $N_y$  measurement variables. Furthermore, the comparison of different MBDoE design criteria is set in relation to a random design of experimental conditions.

$$\Psi_{\text{HR}} = \sum_{i=1}^{M-1} \sum_{j=i+1}^M (\hat{\mathbf{y}}_i - \hat{\mathbf{y}}_j)^T (\hat{\mathbf{y}}_i - \hat{\mathbf{y}}_j) \quad (1)$$

$$\Psi_{\text{BH}} = \sum_{i=1}^{M-1} \sum_{j=i+1}^M p_i p_j \left( \text{tr} \left\{ \frac{\hat{\mathbf{V}}_i}{\hat{v}_j} + \frac{\hat{\mathbf{V}}_j}{\hat{v}_i} - 2\mathbf{I} + (\hat{\mathbf{y}}_i - \hat{\mathbf{y}}_j)^T \left( \frac{1}{\hat{v}_i} + \frac{1}{\hat{v}_j} \right) (\hat{\mathbf{y}}_i - \hat{\mathbf{y}}_j) \right\} \right) \quad (2)$$

$$\Psi_{\text{BF}} = \sum_{i=1}^{M-1} \sum_{j=i+1}^M \frac{(\hat{\mathbf{y}}_i - \hat{\mathbf{y}}_j)^T (\hat{\mathbf{y}}_i - \hat{\mathbf{y}}_j)}{\hat{v}_i + \hat{v}_j} \text{tr} \left\{ \frac{2\boldsymbol{\Sigma}}{\hat{\mathbf{V}}_i + \hat{\mathbf{V}}_j} \right\} \quad (3)$$

$$\Psi_{\text{AWDC}} = \sum_{i=1}^M \frac{1}{\sum_{j=i}^M \exp \left\{ -\frac{1}{2} (\hat{\mathbf{y}}_i - \hat{\mathbf{y}}_j)^T \hat{\mathbf{V}}_i^{-1} (\hat{\mathbf{y}}_i - \hat{\mathbf{y}}_j) + N_{\theta_i} - N_{\theta_j} \right\}} \quad (4)$$

$$\Psi_{\text{JRD}} = H_2(\sum_{i=1}^M p_i g_i) - \sum_{i=1}^M p_i H_2(g_i) \text{ with} \quad (5)$$

$$H_2(\sum_{i=1}^M p_i g_i) = \frac{N_y}{2} \log\{2\pi\} - \log \left\{ \sum_{i=1}^M \frac{p_i^2}{2^{\frac{N_y}{2}} |\hat{\mathbf{V}}_i|^{\frac{1}{2}}} + 2 \sum_{j=1}^{i-1} p_i p_j \exp \left\{ -\frac{1}{2} \phi_{ij} \right\} \right\},$$

$$\phi_{ij} = \frac{\hat{\mathbf{y}}_i^T \hat{\mathbf{y}}_i}{\hat{v}_i} + \frac{\hat{\mathbf{y}}_j^T \hat{\mathbf{y}}_j}{\hat{v}_j} - \frac{(\hat{\mathbf{y}}_i + \hat{\mathbf{y}}_j)^T (\hat{\mathbf{y}}_i + \hat{\mathbf{y}}_j)}{\frac{1}{\hat{v}_i} + \frac{1}{\hat{v}_j}} + \log|\hat{\mathbf{V}}_i| + \log|\hat{\mathbf{V}}_j| + \log \left| \frac{1}{\hat{v}_i} + \frac{1}{\hat{v}_j} \right| \text{ and}$$

$$H_2(g_i) = \frac{N_y}{2} \log\{4\pi\} + \frac{1}{2} \log|\hat{\mathbf{V}}_i|$$

Four different calculation methods of probability weights  $w$  with  $p_i = w_i / \sum_{k=1}^M w_k$  are compared: M1 as inverse of lack of fit; M2 resulting from comparing the lack of fit with the  $\chi^2$  distribution including the degrees of freedom of the model; M3 being based on the Akaike Information Criterion and M4 being similarly based on the likelihood function  $L_i = \prod_{q=1}^{N_{\text{exp}}} \prod_{r=1}^{N_{\text{sp}}} \prod_{s=1}^{N_y} (1/2\pi \boldsymbol{\Sigma}_{s,s})^{1/2} \exp \left\{ - (y_{q,r,s} - \hat{y}_{i,q,r,s})^2 / 2\boldsymbol{\Sigma}_{s,s} \right\}$  (Michalik et al., 2010) as M3 but without penalty term to include the degrees of freedom of the model.

$$\text{M1: } w_i = \frac{1}{\chi_i^2} \quad (6) \quad \text{M2: } w_i = 1 - F_{\text{DoFi}}(\chi_i^2) \quad (7)$$

$$\text{M3: } w_i = \frac{1}{\sum_{k=1}^M \exp\{AIC_i - AIC_k\}} \quad (8) \quad \text{M4: } w_i = \frac{1}{\sum_{k=1}^M \exp\{\pi_i - \pi_k\}} \quad (9)$$

with  $AIC_i = -2\log\{L_i\} + 2\text{DoFi}$  with  $\pi_i = -2\log\{L_i\}$

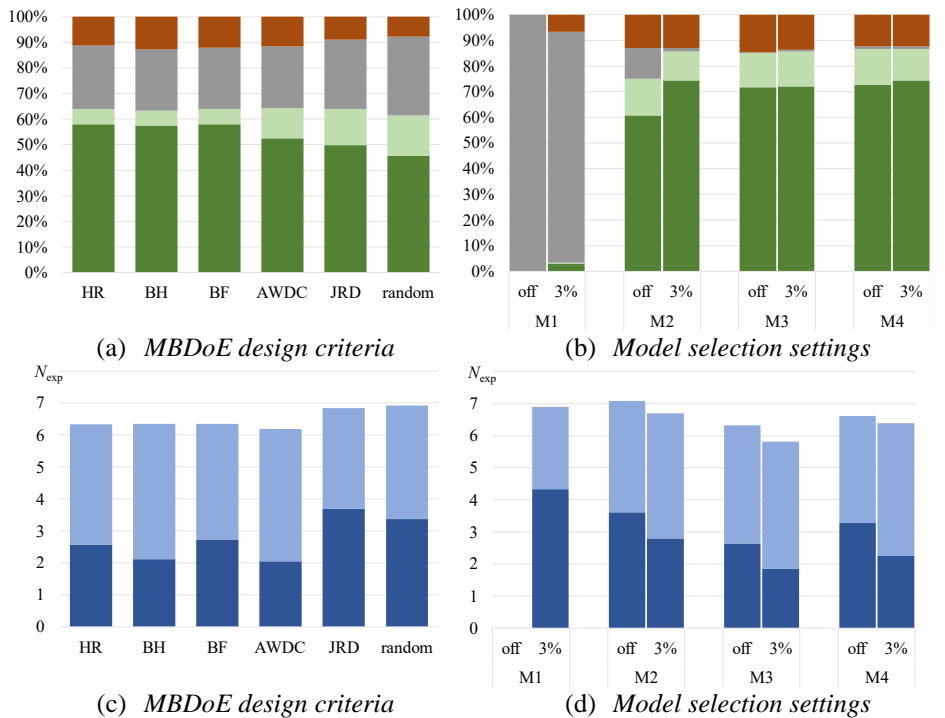
The probability calculation method and preferences for model rejection define the settings used for model selection: either probability-based model rejection is deactivated (“off”) or candidate models with  $p \leq 3\%$  are rejected as indicated by Schwaab et al. (2006).

3.2. Evaluation Metrics

The model identification function (*HoliMI*) is not deterministic since the measurement error generated for the in-silico experiments is randomly distributed. Possible return values are: *i*) correctly selected model with precise parameter estimation; *ii*) correctly selected model with interruption (i.e. maximum number of experiments reached) of parameter precision; *iii*) interruption of model discrimination and *iv*) false model selection. Thus, the different setting variants are compared by means of effectiveness  $EFN \propto \frac{\text{number of correct model selections}}{\text{number of model identifications}}$  and efficiency  $EFC \propto \frac{1}{\text{total number of experiments}}$  of cases with identification of correct model structure and precise parameters, return *i*).

3.3. Results and Discussion

Figure 2 shows the distribution of returns from the model identification function (*HoliMI*) averaged over all realizations with respective design criterion (2.a) and model selection settings (2.b). The average number of experiments required for correctly selected models and precise parameter estimation with respective settings is also presented (2.c, 2.d). The required number of experiments is split in number of experiments required for model discrimination and those required for parameter precision.



**Figure 2** Performance of MBDoe design criteria and a random design strategy (a, b). Distribution of experiments required for model identification (c, d).

(a, b) wrong (orange), undecided (grey), correct: imprecise parameters (light green), correct: precise parameters (dark green)  
 (c, d) experiments for parameter precision (light blue), experiments for model discrimination (dark blue)



In general, MBD<sub>oE</sub> in contrast to a random design is the more efficient strategy to design experiments for the identification of both model structure and parameters. However, MBD<sub>oE</sub> effectiveness is affected by parametric uncertainty as can be seen by the small improvement compared to a random design (2.a, 2.c) when averaging over multiple parametric scenarios. In contrast, the probability calculation method M1 is significantly less effective than the other methods (2.b), with M3 performing most efficiently (2.d). Furthermore, rejecting models at a probability of 3% resulted in faster model identification than in probability-free model rejection methods and in lower rates of false model selections due to the risk of rejecting the true model.

#### 4. Conclusion

The proposed [Python package HoliMI](#) allows a holistic comparison of settings for model identification using MBD<sub>oE</sub> techniques for both model discrimination and improvement of parameter precision. Results on the application of the package to a benchmark case study related to the identification of kinetics in a bioreactor system showed that the selection of model selection settings can affect the model identification performance substantially. Depending on the number of candidate model, a threshold for probability-based model rejections should be considered to enhance model identification with minimum number of experiments. Furthermore, it is recommended to use probability weights based on the probability likelihood function (M3 or M4) for efficient correct model selections. General guidelines on optimal MBD<sub>oE</sub> design criteria for model discrimination can be derived from the application to further study cases and the implementation of the proposed package in autonomous reactor platforms.

#### References

- Asprey, S.P., Macchietto, S. (2000). Statistical tools for optimal dynamic model building. *Computers & Chemical Engineering*, 24, 1261–1267
- Box, G.E.P., Hill, W.J. (1967). Discrimination among mechanistic models. *Technometrics*, 9 (1), 57–71
- Buzzi-Ferraris, G., Forzatti, P. (1990). An improved version of a sequential design criterion for discriminating among rival multiresponse models. *Chemical Engineering Science*, 45 (2), 477–481
- Burke, A.L., Duever, T.A., Penlidis, A. (1997). Discriminating between the terminal and penultimate models using designed experiments: An overview. *Industrial & Engineering Chemistry Research*, 36 (4), 1016–1035
- Chen, B.H., Asprey, S.P. (2003). On the design of optimally informative dynamic experiments for model discrimination in multiresponse nonlinear situations. *Industrial & Engineering Chemistry Research*, 42 (7), 1379–1390
- Franceschini, G., Macchietto, S. (2008). Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science*, 63 (19), 4846–4872
- Hunter, W.G., Reiner, A.M. (1965). Designs for discriminating between two rival models. *Technometrics*, 7 (3), 307
- Michalik, C., Stuckert, M., Marquardt, W. (2010). Optimal experimental design for discriminating numerous model candidates: The awdc criterion. *Industrial & Engineering Chemistry Research*, 49 (2), 913–919
- Olofsson, S., Hebing, L., Niedenführ, S., Deisenroth, M.P., Misener, R. (2019). Gpdoemd: A python package for design of experiments for model discrimination. *Computers & Chemical Engineering*, 125, 54–70
- Schwaab, M., Silva, F.M., Queipo, C.A., Barreto, A.G., Nele, M., Pinto, J.C. (2006). A new approach for sequential experimental design for model discrimination. *Chemical Engineering Science*, 61 (17), 5791–5806

# A Model-based Computational Approach to Optimal Design of Nose-to-Brain Drug Delivery Systems

Athina Vasileiadou,<sup>a</sup> Costas Kiparissides,<sup>a,b</sup>

<sup>a</sup>*Department of Chemical Engineering, Aristotle University of Thessaloniki, 54124, Thessaloniki, Greece*

<sup>b</sup>*Chemical Process & Energy Resources Institute, 6th km Harilaou-Thermi rd., P.O. Box 60361, 57001, Thessaloniki, Greece*

## Abstract

In the present study, a dynamic polymer degradation-drug diffusion model is developed to calculate the drug release rate from a population of size-distributed drug-loaded biodegradable polymer particles and analyze the effects of particle size distribution (PSD), drug loading distribution (DLD), and polymer molecular properties on degradation kinetics and drug release rate. The derived model, comprising a system of partial differential and algebraic equations, is numerically solved using the method of lines and validated against published experimental data on the drug release rate from drug-loaded PLGA microspheres. By controlling the PSD, DLD, molecular and morphological properties of the polymeric carriers, a desirable drug release profile can be realized. Toward this objective, a model-based dynamic optimization approach is developed to determine the optimal PSD and DLD of the biodegradable polymer carriers to achieve a drug release rate of zero-order over an administration period of three weeks.

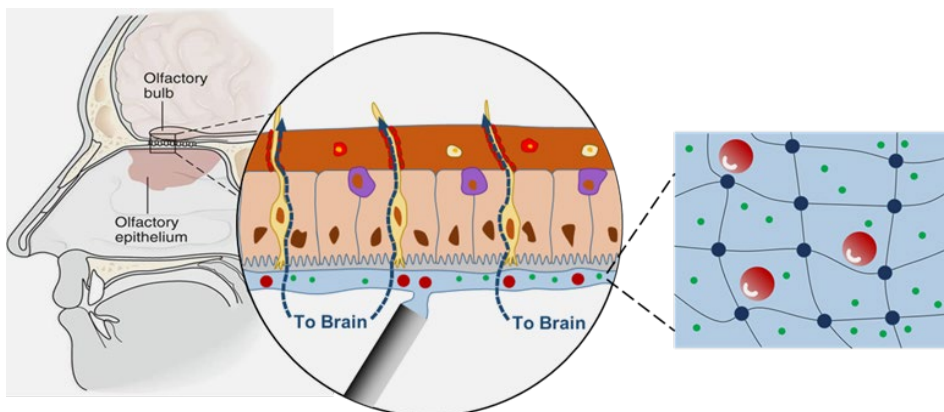
**Keywords:** Controlled drug delivery; polymer degradation kinetics; drug diffusion; model-based optimal design of DDS.

## 1. Introduction

Neurological diseases encompass a broad spectrum of brain pathological conditions, i.e. Parkinson's and Alzheimer's diseases, multiple sclerosis, meningitis, etc., which can cause alterations to neural functions and progressive loss of neural tissue. Despite the recent advances in both drug delivery research and improved understanding of the pathogenesis of neurological diseases, effective treatment options are still missing due to the complexity of the central nervous system (CNS) and the alleged multifactorial pathogenic mechanisms. Thus, currently available therapeutic formulations have been generally proven ineffective because they are unable to reverse or/and completely restore normal neural functions. Moreover, the delivery of therapeutics to the brain by conventional systemic approaches is hindered by several barriers, including the Blood brain barrier (BBB) and the Blood-Cerebrospinal Fluid Barrier (BCSFB). In all cases, the delivery of active pharmaceutical ingredients (APIs) to the brain in sufficient quantities and in a highly controlled and sustainable way (e.g., over a time period of three weeks) is required in order to achieve a desired therapeutic effect.

In a recent EU Horizon collaborative research program, [www.n2b-patch.eu](http://www.n2b-patch.eu), a novel formulation was developed for nose-to-brain controlled delivery of biopharmaceuticals for the chronic treatment of CNS disorders. The novel drug formulation consisted of

biodegradable PLGA particles loaded with a cognition enhancing drug, embedded into a biodegradable hydrogel matrix. The hydrogel patch was deposited, via a nasal endoscopic applicator as a thin liquid-gelling film, onto the olfactory region (Kiparissides et al., 2019). The research goal was to obtain a desirable drug release profile over a long period of time (e.g., 3-4 weeks) by controlling the PSD, drug loading, molecular and morphological properties of the embedded PLGA drug carriers into a hydrogel patch. By controlling the molecular and morphological properties of the biodegradable hydrogel and embedded PLGA carriers, a desirable drug release profile from the embedded particles into the hydrogel matrix to the mucus layer and then to the olfactory bulb can be established.

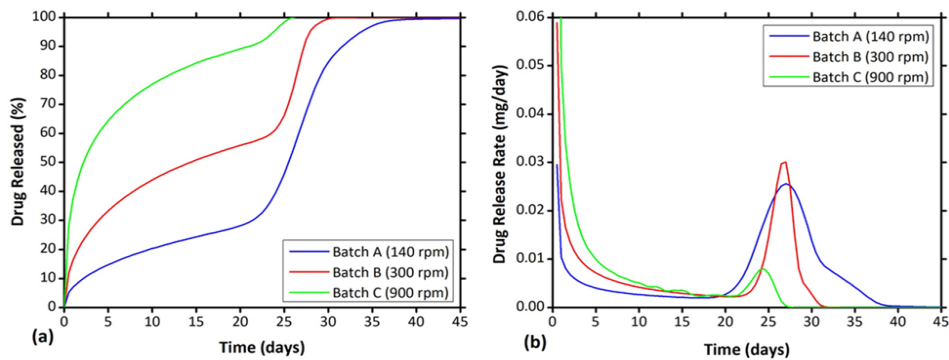


**Figure 1.** The basic concept of the proposed nose-to brain drug delivery system. Illustration of local hydrogel formation at the olfactory cleft via the aid of an endoscopic applicator. Drug-loaded nanoparticles are embedded into the deposited hydrogel patch. Blue arrows show the potential intracellular and extracellular pathways through which the API is transferred to the brain ([www.n2b-patch.eu](http://www.n2b-patch.eu)).

In the present contribution, a process systems computational approach, comprising models at different time and length scales, is developed to elucidate the fundamental physical, transport and biological processes in relation to the nose-to-brain delivery of selected APIs from drug-loaded biodegradable PLGA carriers, embedded into a hydrogel film deposited at the olfactory mucosa, to the epithelium and then to the olfactory bulb (Kiparissides et al. 2020). A polymer degradation-drug diffusion model is derived to calculate the drug release rate from a population of size-distributed drug-loaded particles and analyze the effects of particle size distribution (PSD), drug loading distribution (DLD), polymer molecular and morphological properties on polymer degradation kinetics on drug release profile. The model takes into account the spatial-temporal variation of the drug and water diffusion coefficients due to polymer degradation that results in a decrease of the molecular weight of the polymer. The unknown model parameters are estimated by fitting model predictions to the experimental data of Berchane et al. (2007). Finally, a multi-parametric optimization problem is formulated to calculate the optimal PSD and DLD of drug-loaded polymer carriers to obtain a desired drug release profile over a drug administration period of several weeks.

## 2. Model-based optimal design of drug-loaded polymeric carriers

Frequently, a zero-order release profile is required for sustainable, controlled delivery of a specific drug dose to a patient in order to achieve a maximum therapeutic outcome. As shown in Figure 2a, the experimentally prepared drug delivery formulations of Berchane et al. (2007) exhibit a tri-phasic drug release profile (see blue, red and green lines). The corresponding drug release rates of the three formulations are depicted in Figure 2b. As can be seen the respective drug release rate exhibits an initial burst followed by an almost zero drug release rate and at approximately 20-30 days, the release rate shows a second drug release burst. This means that the prepared formulations cannot maintain a sustainable drug release rate of zero-order over a long period and thus cannot yield an optimal therapeutic effect. In the present study, a model-based dynamic optimization approach was developed to determine the optimal particle size and drug-loading distributions to obtain a sustainable drug release rate profile of zero-order over a specified period of time.



**Figure 1.** (a) Experimental profiles of cumulative drug amount released from three size-distributed populations of piroxicam-loaded PLGA particles. (b) Respective drug release rates for the three experimental formulations (Berchane et al. 2007).

### *Selection of the optimal particle size distribution*

In this study, it was assumed that the optimal drug-loaded PLGA particles followed a log-normal volume probability distribution,  $N_V(d)$ , given by

$$N_V(d) = \frac{1}{d\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln d - \mu}{\sigma}\right)^2\right); \quad 0 < d < \infty \quad (1)$$

Where  $\mu$  is the mean value and  $\sigma$  is the standard deviation of the log-normal distribution. Thus, the values of  $\mu$  and  $\sigma$  were selected as control variables in the optimization problem.

### *Selection of the optimal drug loading distribution*

Regarding the drug loading distribution (i.e., drug mass per unit of polymer mass), it was assumed to follow a size-dependent logarithmic function with respect to particle size. Moreover, the initial particle concentration at  $t=0$  was assumed to be independent of particle radius.

$$f_{DL}(d) = -\alpha \cdot \ln(d) + \beta \quad (2)$$

Thus, the two drug loading parameters  $\alpha$  and  $\beta$  were treated as two additional control variables in the optimization problem.

*Drug release from a size distributed population of drug-loaded particles*

To calculate the drug release rate from a monodisperse population of biodegradable drug-loaded PLGA microspheres, two partial differential equations were formulated and numerically solved to calculate the spatial-temporal variation of drug and water concentrations in the PLGA microspheres. Accordingly, the following partial differential equation and associated initial and boundary conditions were derived to describe the drug diffusion and release from the PLGA microspheres.

$$\begin{aligned}\frac{\partial C_d(r, t)}{\partial t} &= \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 D_{d,eff}(t, r) \frac{\partial C_d(r, t)}{\partial r} \right) \\ &= D_{d,eff} \left( \frac{\partial^2 C_d}{\partial r^2} + \frac{2}{r} \frac{\partial C_d}{\partial r} \right) + \frac{\partial C_d}{\partial r} \frac{\partial D_{d,eff}}{\partial r},\end{aligned}\quad (3)$$

$$\left. \frac{\partial C_d(r, t)}{\partial r} \right|_{r=0} = 0 \quad (4)$$

$$D_{d,eff}(t, r) \left. \frac{\partial C_d(r, t)}{\partial r} \right|_{r=R} = -k_{c,d}(C_d(R, t) - C_{out}(t)) \quad (5)$$

$$C_d(r, 0) = C_{in} \quad (6)$$

$$V_{out} \frac{dC_{out}(t)}{dt} = N_p k_{c,d} A_p (C_d(R, t) - C_{out}(t)) \quad (7)$$

The symbols  $C_d(r, t)$  and  $C_{out}(t)$  denote the drug concentration in the particles and release medium, respectively.  $C_{in}$  is the initial drug concentration in the particles at  $t=0$ , independent of particle radius.  $D_{d,eff}(r, t)$  and  $k_{c,d}$  are the effective intra-particle drug diffusion coefficient and drug mass transfer coefficient from the particle's external surface ( $A_p$ ) to the continuous release medium, respectively.  $R$  is the particle radius,  $V_{out}$  is the volume of the release medium,  $N_p$ , is the total number of drug-loaded particles. Note Eq. (7) is a dynamic mass balance on the drug concentration in the continuous release medium (sink).

Similarly, one can derive a partial differential equation describing the temporal-spatial variation of water in the PLGA particles

$$\begin{aligned}\frac{\partial C_w(r, t)}{\partial t} &= \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 D_{w,eff}(t, r) \frac{\partial C_w(r, t)}{\partial r} \right) - k_{a,1} C_w(r, t) \mu_0(r, t) \\ &\quad + k_{\alpha,2} C_{H^+}(r, t) \lambda_0(r, t) \\ &\quad - k_d C_{H^+}(r, t) C_w(r, t) (\mu_1(r, t) - \mu_0(r, t)) \\ &= D_{w,eff} \left( \frac{\partial^2 C_w}{\partial r^2} + \frac{2}{r} \frac{\partial C_w}{\partial r} \right) + \frac{\partial C_w}{\partial r} \frac{\partial D_{w,eff}}{\partial r} - k_{a,1} C_w \mu_0 \\ &\quad + k_{\alpha,2} C_{H^+} \lambda_0 - k_d C_{H^+} C_w (\mu_1 - \mu_0)\end{aligned}\quad (8)$$

$$\left. \frac{\partial C_w(r, t)}{\partial r} \right|_{r=0} = 0 \quad (9)$$

$$D_{w,eff}(t, r) \left. \frac{\partial C_w(r, t)}{\partial r} \right|_{r=R} = -k_{c,w}(C_w(R, t) - C_{w,eq}) \quad (10)$$

$$C_w(r, 0) = 0 \quad (11)$$

Where  $C_w(r, t)$  and  $C_{w,eq}$  are the water concentration in the particles and at equilibrium, respectively.  $D_{w,eff}(r, t)$  is the effective intra-particle water diffusion coefficient and  $k_{c,w}$  is the mass transfer coefficient of water from the release medium to the particle's external surface.

Note that in order to reduce the computational effort (i.e., the number of differential equations to be integrated), the particle size distribution was discretized into four particle classes,  $N_c$ . The final model comprised a system of 2549 non-linear ordinary differential equations (i.e., 9 variables for each discrete particle class times 71 or 70 discretization points in the radial coordinate plus the  $C_{out}$  variable) and 28 algebraic equations.

*Definition of the objective function and numerical solution*

Based on the above stated optimization objective (i.e., of zero-order drug release rate), the following dynamic optimization problem was formulated.

$$\min J_{obj,2} = \|\mathbf{R}_p - \mathbf{R}_d\|_2, \quad w. r. t. (\sigma, \mu, \alpha, \text{ and } \beta) \quad (12)$$

Where the vectors  $\mathbf{R}_d = (R_{d,1}, R_{d,1}, \dots, R_{d,max})$  and  $\mathbf{R}_p = (R_{p,1}, R_{p,1}, \dots, R_{p,max})$  denote the desired and model-based optimal values of the drug release rate, respectively. The maximum number of discrete drug release rate values was set equal to 50 over a drug release horizon of 25 days, which corresponds to one drug release rate value per 1/2 day.

It is important to note that the above optimization problem was solved together with the following equality constrain regarding the conservation of the total drug mass in the system.

$$\left( \sum_{j=1}^{N_c} m_{d,j}(t = 0) \right) = m_{d,total} \quad (13)$$

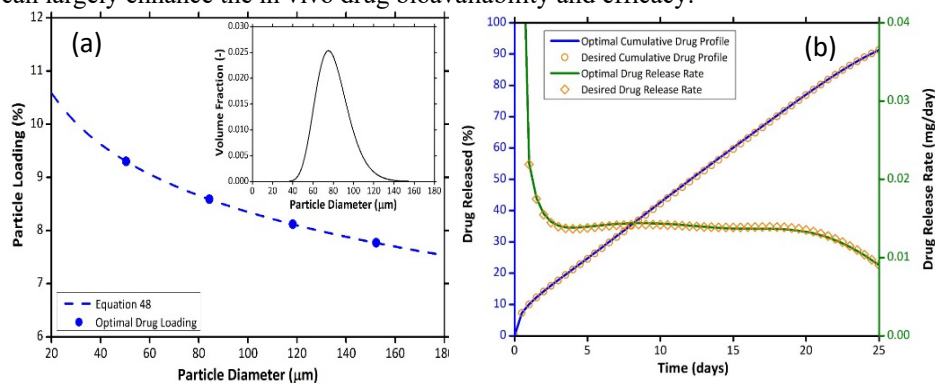
where  $m_{d,total}$  is the total drug dose (i.e., 0.4 mg ) and  $m_{dj}$  is the drug mass in the “j<sup>th</sup>” particle class at t=0.

Thus, our optimization objective was to deliver a total therapeutic drug dose of 0.4 mg over a time period of 25 days by maintaining an almost zero-order drug release rate, 0.0128 mg/day. The above stated optimization problem was numerically solved in MATLAB®, employing the ode15s solver for the integration of the differential-algebraic equations and the “MultiStart” global optimizer (a MATLAB® built-in function) for the estimation of the four control variables. From the numerical solution of the optimization problem the parameters  $\mu$  and  $\sigma$  of the log-normal PSD (Eq. (1)) and the parameters  $\alpha$  and  $\beta$  of the drug loading function (Eq. (2)) were estimated so that a desired zero-order drug release rate profile could be realized. The optimal estimates of the four control variables are reported in Table 1 as well as the optimal dug loading per discrete class of particles.

**Table 1.** Estimated optimal values of the four control variables.

Control Variables				Drug Loading Per Particle Class			
$\mu_x$ ( $\mu\text{m}$ )	$\sigma_x$ ( $\mu\text{m}$ )	$\alpha$	$\beta$	Class 1	Class 2	Class 3	Class 4
80.15	16.60	0.0139	0.1475	9.30%	8.59%	8.12%	7.77%

In Figure 3a, the calculated optimal PSD and optimal drug loading distribution are depicted. As can be seen the calculated particle size distribution varies from 40  $\mu\text{m}$  to 140  $\mu\text{m}$  while the drug loading per particle class (i.e., % of drug mass per polymer mass of a class) decreases as the particle size increases. In Figure 3b, the optimal and desired percentage of cumulative drug released and respective drug release rate are plotted with respect to the drug release time. As can be seen the optimal formulation of the drug-loaded PLGA carriers (see optimal values of  $\sigma$ ,  $\mu$ ,  $\alpha$ , and  $\beta$  in Table 1) closely follows the desired drug release rate of zero-order from the second to the twenty third day of the total release period. Note that the observed deviations in the calculated drug release rate from the desired zero-order profile are primarily due to the inherent initial burst in drug release rate while the observed deviation after the 23<sup>rd</sup> day is due to the depletion of drug concentration in the PLGA carriers. The results of the present investigation do show that the proposed model-based optimal design of drug-loaded polymer carriers can aid the development of new drug delivery systems exhibiting a desired drug release profile that can largely enhance the *in vivo* drug bioavailability and efficacy.



**Figure 3.** (a) Calculated optimal drug loading values per particle class. In the inset graph, the calculated optimal particle volume fraction distribution is depicted. (b) Total drug released and drug release rate from a size-distributed population of particles.

## References

- N.S. Berchane, K.H. Carson, A.C. Rice-Ficht, M.J. Andrews, 2007, Effect of Mean Diameter and Polydispersity of PLG Microspheres on Drug Release: Experiment and Theory, *International Journal of Pharmaceutics*, 337, 1-2, 118-126.
- C. Kiparissides, A. Vasileiadou, F. Karageorgos, S. Serpetsi, 2020, A Computational Systems Approach to Rational Design of Nose-to-Brain Delivery of Biopharmaceutics, *Industrial & Engineering Chemistry Research*, 59, 6, 2548–2565.
- T.D. Knab, S.R. Little, R.S. Parker, 2015, A Systems Approach to Modeling Drug Release from Polymer Microspheres to Accelerate *in vitro* to *in vivo* Translation, *Journal of Control Release*, 211, 74-84.
- S.H.S Koshari, D.P. Chang, N.B. Wang, I.E. Zarraga, K. Rajagopal, A.M. Lenhoff, N.J. Wagner, 2019, Data-Driven Development of Predictive Models for Sustained Drug Release, *Journal of Pharmaceutical Sciences*, 108, 11, 3582-3591.

# Global estimability analysis and model-based design of experiments in surrogate modeling

Ilias Bouchkira,<sup>a,b</sup> Sanae Elmisaoui,<sup>b,c</sup> Saad Benjelloun,<sup>c</sup> Abderrazak M. Latifi,<sup>b,\*</sup>

<sup>a</sup> *Department of Chemical Engineering, Loughborough University, Epinal Way, Loughborough, LE11 3TU, United Kingdom.*

<sup>b</sup> *Laboratoire Réactions et Génie des Procédés, CNRS-ENSIC, Université de Lorraine, 1 rue Grandville, BP20451, Nancy Cedex, 54001, France.*

<sup>c</sup> *Mohammed VI Polytechnic University, Hay Moulay Rachid, Ben Guerir 43150, Morocco.*

[abderrazak.latifi@univ-lorraine.fr](mailto:abderrazak.latifi@univ-lorraine.fr)

## Abstract

In this work, a novel approach for the development of surrogate models is presented. It aims to combine traditional model-based design of experiments and global estimability analysis. The approach is based on the Fisher Information Matrix to set two objective functions to be maximized. The first is given by D-optimal design of experiments criterion that maximizes the determinant of FIM, thus minimizing the volume of the confidence region and resulting in more precise parameter estimate. The second is given by the sum of Euclidian norms of FIM's columns to improve the estimability of all the unknown parameters. The resulting multi-objective optimization problem is solved to determine the Pareto front of the optimal solutions. The Multi-Attribute Utility Theory is used as a decision making aid method to select the best solution, needed for the development of the surrogate model. For demonstration purposes, a polynomial Response Surface Model (RSM) based surrogate is developed. It mimics a computationally expensive first-principles model that predicts the conversion rate of a phosphate ore digestion by phosphoric acid. The results are very promising, showing how the approach allows to select the most estimable model parameters from intelligently designed data points. High performance of the developed surrogate is demonstrated by comparing its predictions and computation time to those obtained by first-principles model.

**Keywords:** Surrogate modeling, Model-based design of experiments, Sensitivity and Estimability analysis, Multi-objective optimization, Decision making aid.

## 1. Introduction

The development of reliable first-principles models is of utmost importance in process engineering. Recently, there is an increasing demand for high fidelity, large and multi-scale mathematical models. However, in most cases, the process models usually involve sets of ODEs, PDEs etc., that take time to integrate making the developed models very slow/expensive to converge. To overcome these problems, high-fidelity and accurate surrogate models that allow to quickly compute the required model outputs from the inputs are needed. Indeed, surrogate models have been increasingly investigated recently (Rabhi et al., 2018; Elmisaoui et al., 2022). They are developed to mimic the behavior of a given simulation model or an outcome of interest that cannot be easily measured or computed, as closely as possible, while being computationally cheap to evaluate.

Surrogate modeling is usually based on a limited number of intelligently chosen data points (black-box modeling), and on choosing an appropriate mathematical expression of the surrogate, which involves in most cases many unknown parameters to be identified.



For the development of accurate surrogates, it is very important to design the most optimal data points that maximize the estimability of the surrogate model parameters in order to determine the most estimable parameters to be identified.

These two objectives are generally overlooked in most of the current published papers. The main purpose of the present work is to develop a novel approach that enables to (i) firstly, select intelligently the dataset of points to mimic by the surrogate model, and (ii) secondly to determine with accuracy, the unknown surrogate model parameters to be identified from the selected dataset of points.

This approach is based on a constrained multi-objective optimization problem that accounts for two objectives, one from Model-based design of experiments and one from parameter estimability, while the constraints are mainly given by the range of variation of the operating conditions. The solution of the formulated problem will be a set of optimal profiles, i.e., a Pareto front, from which, the best solution can be chosen using a decision making aid method. For demonstration purposes, the approach is applied to the case of polynomial Response Surface Model (RSM)-based surrogate, to mimic the behavior of a computationally expensive first-principals model that predicts the conversion rate of the digestion reaction of phosphate ore by phosphoric acid.

## 2. Models and methods

### 2.1. First-principals model of phosphate ore digestion

The case study in this work deals with a phosphate ore digestion model previously developed by Elmisaoui et al. (2021). It describes the dissolution behavior of a pure tri-calcium phosphate (TCP) in phosphoric acid to produce mono-calcium phosphate (MCP). The dissolution reaction is given by:



The conversion rate of this reaction depends on several operating conditions such as, the reaction temperature, the initial phosphoric acid concentration, the tri-calcium phosphate particle size distribution and porosity, the hydrodynamic conditions, the solid/liquid ratio, the residence time etc. In this work, four sensitive parameters are considered. They include the initial TCP particle size ( $R_0$ ), the initial phosphoric acid concentration ( $C_0$ ), the tank stirring speed through a hydrodynamic parameter ( $\alpha$ ), and the residence time ( $\tau$ ) inside the reactor. Moreover, to predict the dissolution rate, the shrinking core model is used considering three main phases: (i) the liquid bulk, (ii) the liquid film that surrounds the solid particles, (iii) and the solid phase. Some assumptions are made for the development of the first-principals model and can be found in Elmisaoui et al. (2021).

The resulting model is given by the following system of mass balance equations:

$$\frac{dX}{dt} = \frac{3M_{TCP}D_{MCP}}{x_{TCP}R_0\delta\rho_S} (1 - X)^{\frac{2}{3}} (C_{MCP}|_{r=R} - C'_{MCP}) \quad (2)$$

$$\frac{\partial C_i}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left( D_i r^2 \frac{\partial C_i}{\partial r} \right) \quad i = H_3PO_4, Ca(H_2PO_4)_2 \quad (3)$$

$$r = R: -D_i \frac{\partial C_i}{\partial r} \Big|_{r=R} = v_i k_r C_{H_3PO_4} \Big|_{r=R} \quad (4)$$

$$r = R + \delta: C_i|_{r=R+\delta} = C'_i \quad (5)$$

$$\delta = R \left[ 1 + \alpha \left( \frac{R}{R_0} \right)^{8/9} D_{MCP}^{-1/3} \right]^{-1} \quad (6)$$

$$C'_{MCP} = \frac{3n_{TCP}^0}{V_L} \text{ and } C'_{\Pi_3PO_4} = C_{H_3O_4}^0 - \frac{4n_{TCP}^0}{V_L} \quad (7)$$

In the system of equations above, (2) describes the solid phase, (3) describes the liquid film surrounding the solid particles, (4 and 5) are the two main boundary conditions, (6) is used to predict the film thickness  $\delta$ , and (7) describes the liquid bulk.  $m_s$  and  $R$  are respectively the mass and radius of the solid phosphate ore particles,  $x_{TCP}$  and  $M_{TCP}$  are the mass fraction of TCP and its molecular weight.  $C_{MCP}|_{r=R}$  and  $C'_{MCP}$  are the concentrations of MCP at the solid surface and in the liquid bulk.  $k_r$  is a rate constant,  $D_i$  and  $v_i$  are the diffusion coefficient and stoichiometric coefficient of component  $i$ , respectively.  $n_{TCP}^0$  is the initial number of moles of TCP in the solid,  $V_L$  is the liquid volume in the tank,  $\alpha$  is a hydrodynamic parameter calculated from the stirring speed of the tank. The solution of (2-6) allows to predict the conversion rate  $X(\tau, C_0, R_0, \alpha)$  to be mimicked by the surrogate model.

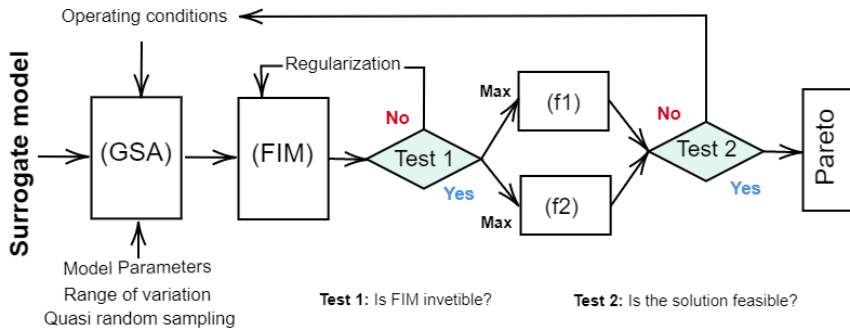
### 2.2. Polynomial Response Surface Model-based surrogate

Surface Response model-based surrogates have been extensively used recently. They proved to be accurate, flexible, and available in many commercial data analysis software. The third degree polynomial RSM-based surrogate model used in this work writes as:

$$\hat{Y} = y_0 + \sum_i a_i X_i + \sum_i b_i X_i X_i + \sum_{i \neq j} d_{i,j} X_i X_j + \sum_i c_i X_i X_i X_i + \sum_{i \neq j \neq k} e_{i,j,k} X_i X_j X_k \quad (7)$$

*Cst*      *Linear terms*      *Binary interaction terms*      *Ternary interaction terms*

Considering four main sensitive operating parameters, i.e.,  $u = \{\tau, C_0, R_0, \alpha\}$ , the surrogate model above involves 23 unknown parameters (table 1). The main challenges that the proposed approach addresses are (i) which data points (from the first-principals model) to mimic, and (ii) which unknown model parameters to estimate.



**Figure 1:** Framework of the multi-objective model-based design of experiments. (GSA): Global Sensitivity Analysis. (FIM): Fisher Information Matrix.

### 2.3. Multi-objective model-based design of experiments

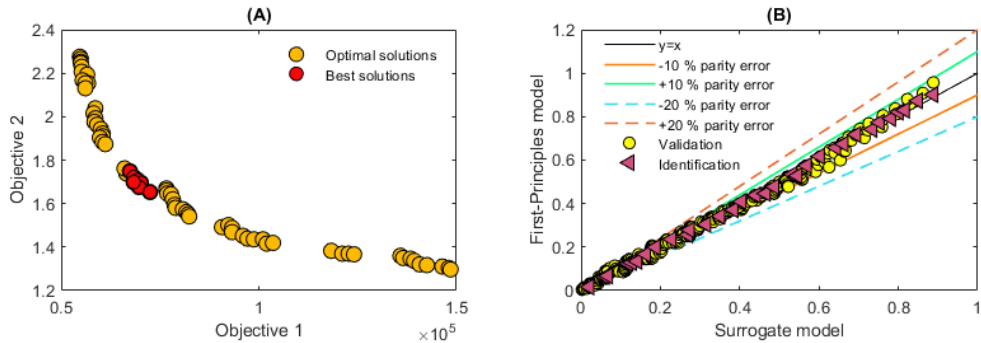
Figure 1 shows the proposed approach for the development of the surrogate models. It is based on the Fisher Information Matrix (FIM), which is computed from the global sensitivity indices of the model's unknown parameters as:

$$Z = \begin{pmatrix} S_{1,1|m_1} & \cdots & S_{1,d|m_1} \\ \vdots & \ddots & \vdots \\ S_{n,1|m_n} & \cdots & S_{n,d|m_n} \end{pmatrix}, \quad \mathbf{FIM} = Z.Z^T \quad (8)$$

where  $m_i$  is the  $i^{th}$  sampling point,  $d$  is the number of model parameters and  $n$  is the number of measurements.  $S_{i,j}$  are the individual sensitivity indices of the  $i^{th}$  output with respect to the  $j^{th}$  parameter. They are calculated by means of the method developed by Saltelli (2012). FIM is then used to compute the two objective functions  $f_1$  and  $f_2$ . They represent the D-optimal model-based design of optimal experiments (MBDoe) criterion and the Euclidian norm of FIM's columns, respectively (Bouchkira et al., 2021; 2022; 2023). A regularization test is performed to ensure that the FIM is invertible (Test 1). Then, multi-objective optimization is performed to maximize the two objective functions and to determine the Pareto front of the optimal solutions (Test 2). The multi-objective optimization problem is formulated as:

$$\begin{aligned} \max_{\varphi} \mathbf{F} &= [f_1, f_2]^T \\ f_1 &= \det(\mathbf{FIM}) \\ f_2 &= \sum_{i,j} \|S_{i,j}\| \\ \mathbf{FIM} &= Z(X, \theta).Z(X, \theta)^T \\ \varphi &= [C_{0,1}, \dots, C_{0,n}, R_{0,1}, \dots, R_{0,m}, \alpha_1, \dots, \alpha_k, t_1, \dots, t_w] \\ \text{s.t.} \quad &0 < C_{0,j-1} \leq C_{0,j}, j \in [1; n]; \quad 0 < R_{0,r-1} \leq R_{0,r}, r \in [1; m]; \\ &0 < \alpha_{s-1} \leq \alpha_s, s \in [1; k]; \quad 0 < t_{p-1} \leq t_p, p \in [1; w] \\ &C_{0,j-1} - C_{0,j} \leq \omega_C; \quad R_{0,r-1} - R_{0,r} \leq \omega_R \\ &\alpha_{s-1} - \alpha_s \leq \omega_\alpha; \quad t_{p-1} - t_p \leq \omega_t \\ &Cond(\mathbf{FIM}) \leq \theta \end{aligned} \quad (\text{S.1})$$

where  $\omega_C$ ,  $\omega_R$ ,  $\omega_\alpha$  and  $\omega_t$  are used to avoid duplicated experiments.  $\theta$  is a cut-off criterion to guarantee that FIM is revertible,  $\theta = 10$  is considered as in (Cenci et al., 2022).



**Figure 2:** Optimization and identification results. (A): Pareto front. (B): Parity diagram.

**Table 1:** Surrogate modeling and estimability results. Parameters with (\*) are not estimable

Parameter	Terms	Value	Precision ( $\pm$ )	Estimability magnitude	Rank
$x_0$	$cst$	-0.076	0.0361	0.156	3
$a_1$	$C_0$	-9.187	$7,72. 10^{-3}$	0.051	11
$a_2$	$R_0$	$-5.88. 10^5$	149.1	0.061	8
$a_3^*$	$\alpha$	-	-	0.015	13
$a_4$	$t$	$0.15. 10^{-2}$	$7.19. 10^{-4}$	0.487	2
$b_1$	$C_0, C_0$	$-2.02. 10^{-2}$	$8.23. 10^{-6}$	0.094	4
$b_2$	$R_0, R_0$	$7.22. 10^5$	$2.03. 10^4$	0.054	10
$b_3^*$	$\alpha, \alpha$	-	-	0.003	20
$b_4$	$t, t$	$-1.79. 10^{-1}$	$2.05. 10^{-4}$	0.705	1
$c_1^*$	$C_0, C_0, C_0$	-	-	0.002	21
$c_2$	$R_0, R_0, R_0$	$-2.20. 10^{13}$	$9.03. 10^{10}$	0.066	7
$c_3^*$	$\alpha, \alpha, \alpha$	-	-	0.012	14
$c_4$	$t, t, t$	$0.01. 10^{-7}$	$1.78. 10^{-9}$	0.081	5
$d_1^*$	$C_0, R_0$	-	-	0.008	15
$d_2^*$	$C_0, \alpha$	-	-	0.008	16
$d_3$	$C_0, t$	$2.90. 10^{-1}$	$1.88. 10^{-4}$	0.045	12
$d_4^*$	$R_0, \alpha$	-	-	0.001	23
$d_5$	$R_0, t$	-0.509	0.0138	0.061	9
$d_6^*$	$\alpha, t$	-	-	0.004	19
$e_1^*$	$C_0, R_0, \alpha$	-	-	0.001	22
$e_2$	$C_0, R_0, t$	-0.111	$2.239. 10^{-2}$	0.073	6
$e_3^*$	$C_0, \alpha, t$	-	-	0.004	18
$e_4^*$	$R_0, \alpha, t$	-	-	0.006	17

### 3. Results and discussion

Global sensitivity analysis is first performed based on the method of Saltelli (2002), and a Quasi-Monte Carlo sampling of size 10.000 was sufficient to achieve sensitivity indices convergence. The sensitivity indices were then used to build the FIM which is used in the multi-objective optimization problem. The latter is then implemented and solved under MATLAB, and a Genetic Algorithms solver was used to achieve global optimality.

Resolution of the multi-objective optimization problem allowed to generate the Pareto front (Figure 2A) of the optimal compromises to be mimicked by the surrogate model. It is worth mentioning that all the points on the Pareto front are optimal. However, for further development of the surrogate model, the best compromises of the Pareto front should be selected to be implemented. To this end, the Multi-attribute Utility Theory method was used as a decision making aid method as in our previous works (Bouchkira et al. 2022). It allowed to classify the solutions and to select the best one. Five best solutions are highlighted in red color in figure 2A. The best solution was selected to be implemented and suggests to mimic 32 outputs of the first-principles model predictions, based on the combination of two initial phosphoric acid concentrations i.e.,  $C_0 = 115$  and  $C_0 = 453$ ; two initial particle radius, i.e.,  $R_0 = 6.9 10^{-4}$  and  $R_0 = 5.1 10^{-4}$ ; two hydrodynamic parameters, i.e.,  $\alpha = 2.73$  and  $\alpha = 3.18$ ; and four residence times, i.e.,  $\tau = 155$ ,  $\tau = 215$ ,  $\tau = 403$ , and  $\tau = 600$ .

Table 1 summarizes the results obtained in the development of the polynomial RSM-based surrogate model using the best solution. It shows first the ranking of the surrogate unknown parameters according to their estimability from the selected best solution. Moreover, the results show that only 12 unknown parameters out from 23 are estimable, and this is deduced from the estimability magnitudes of the parameters that were computed within the optimization approach, and their comparison to an estimability cut-off criterion proposed in (Yao et al. 2011). The values of the estimable parameters are provided along with their confidence intervals.

It should be noted here that the proposed approach has the advantage of identifying only the estimable model parameters, which avoids the model overfitting and improves the model precision by ignoring uncertain parameters.

Figure 2B shows a comparison between the predictions of the developed surrogate model and those of the expensive first-principals model. The results show the accuracy and good performance of the developed surrogate model, and this can be shown in the parity errors that do not exceed -10% and +10% for all the data points used for the identification of the model parameters (red triangle), while the errors on other data points that were not used during the identification process (validation: yellow circles) lie between -20% and +20%. It is finally worth mentioning that the developed surrogate model proved to be more than three times faster than the first-principals model, and this was deduced by running both models several times on an ASUS workstation with 12th Gen Intel(R) Core (TM) i9-12900K 3.20 GHz 128GB Ram processor.

#### 4. Conclusion

A novel multi-objective optimization-based approach for the development of accurate surrogate models is presented. It has the advantage of designing intelligently the data points from the computationally expensive model to be mimicked by the surrogate model, while selecting only its estimable unknown parameters. To demonstrate the effectiveness of the proposed approach, it is used in this work for the development of a surrogate model to mimic the predictions of a computationally expensive first-principals model that predicts the conversion rate of a phosphate ore digestion using phosphoric acid. High performance of the developed surrogate model was finally demonstrated by comparing its predictions and computation time to those obtained by the first-principals model.

#### References

- I. Bouchkira, A. Latifi, L. Khamar, S. Benjelloun, 2021, Global sensitivity based estimability analysis for the parameter identification of Pitzer's thermodynamic model. *Reliability Engineering & System Safety*, 207, 107263.
- I. Bouchkira, S. Benjelloun, L. Khamar, A. Latifi, 2023. Thermodynamic modeling and parameter estimability analysis of a wet phosphoric acid process with impurities. *Fluid Phase Equilibria*, 564, 113594.
- I. Bouchkira, A. Latifi, L. Khamar, S. Benjelloun, 2022. Modeling and multi-objective optimization of the digestion tank of an industrial process for manufacturing phosphoric acid by wet process. *Computers & Chemical Engineering*, 156, 107536.
- F. Cenci, G. Bano, C. Christodoulou, Y. Vueva, S. Zomer, M. Barolo, F. Bezzo, P. Facco, (2022). Streamlining tablet lubrication design via model-based design of experiments, *International Journal of Pharmaceutics*, 614, 121435.
- S. Elmisaoui, A. M. Latifi, L. Khamar, M. Salouhi, (2021), Analysis of the dissolution mechanism in the phosphoric acid manufacturing process: modelling and simulation. *Computer Aided Chemical Engineering* 50, 891–897.
- S. Elmisaoui, S. Benjelloun, A. Chkifa, A.M. Latifi, (2022), A sparse polynomial surrogate model for the shrinking core model in phosphate ore digestion, *Computer Aided Chemical Engineering*, 51, 1291-1296.
- A. Rabhi, A. Chkifa, S. Benjelloun, A. Latifi, (2018), Surrogate-based modeling in flotation processes, *Computer Aided Chemical Engineering*, 43, 229-234.
- A. Saltelli, (2002), Making best use of model evaluations to compute sensitivity indices. *Computer physics communications*, 145(2), 280–297.
- K.Z. Yao, M.B. Shaw, B. Kou, K.B. McAuley, D.W. Bacon, (2003), Modeling ethylene/butene copolymerization with multi-site catalysts: parameter estimability and experimental design, *Polymer Reaction Engineering*, 11(3), 563-588.

# Tractable Data-driven Solutions to Hierarchical Planning-scheduling-control

Damien van de Berg<sup>a</sup>, Roberto Xavier Jimenez Jimbo<sup>a</sup>, Nilay Shah<sup>a</sup>, Ehecatl Antonio del Rio-Chanona<sup>a,\*</sup>

<sup>a</sup>*Sargent Centre for Process Systems Engineering, Roderic Hill Building South Kensington Campus. London, SW7 2AZ, United Kingdom*

## Abstract

Using numerical optimization for the hierarchical integration of decision-making units is crucial to provide feasibility and optimality of all levels. However, realistically modelling hierarchical decision-making calls for multilevel formulations, which are numerically intractable and mathematically difficult. In this work, we show how to leverage two data-driven techniques – derivative-free optimization and optimality surrogates – to decrease the computational burden of multilevel problems. We reformulate a tri-level planning-scheduling-control problem into a single-level black-box problem wherein each evaluation calls a scheduling instance with embedded optimal control surrogates. We show that solving this integrated problem instead of the single-level instance leads to changes in the optimal production planning and scheduling sequence, and discuss trade-offs associated with both techniques.

**Keywords:** Derivative-free optimization; Optimization with embedded surrogates; Enterprise-wide optimization; Hierarchical decision-making

## 1. Introduction

Chemical enterprises can capture significant value by using mathematical optimization for the integration of hierarchical levels of operations from supply chain and process design down to control. Traditionally, in *sequential* decision-making, upper-level decisions are taken while disregarding lower-level considerations, and then fed as setpoints to the lower levels. In the *monolithic* approach, lower-level feasibility considerations are included into upper-level optimization problems accompanied by a significant drop in computational tractability. In the *hierarchical* approach, upper-level decision-makers consider lower-level optimality, accounting for how lower-level optimization problems affect upper-level objectives. This results in multi-level formulations, which are often numerically intractable and mathematically difficult (Chu and You, 2015). In this work, we show how we can leverage derivative-free optimization (DFO) (van de Berg et al., 2022) and surrogate modelling techniques to efficiently approach the tri-level optimum of an integrated planning-scheduling-control optimization problem.

To solve hierarchically integrated optimization problems, we reformulate multiple nested levels of decision-making into a single-level formulation. To this end, we can leverage data-driven techniques in a bottom-up or top-down approach. In the bottom-up approach, we first construct ‘optimality surrogates’ that map the optimal response variables of the lower level appearing in the upper levels as a function of the upper-level variables

appearing in lower levels. Similar techniques are known under various names across different fields, such as multi-parametric model predictive control, approximate explicit optimal control laws, amortized learning in Machine Learning, or policy-based methods in Reinforcement Learning (Avraamidou and Pistikopoulos, 2019; Sacchio et al., 2021). The top-down approach relies on derivative-free optimization (DFO), which is closely related to black-box and simulation-based optimization. The reader is referred to van de Berg et al. (2022) for a review on DFO for process systems engineering applications and to Beykal et al. (2022) for an exposition of how DFO could be used to solve multilevel problems.

Figure 1 shows our approach to the planning-scheduling-control problem: The bottom-up technique constructs optimality surrogates for the lower-level control problem. This surrogate is created offline and included as a constraint in the middle scheduling problem. As such, we collapse the two lower levels of scheduling and control into a single level. We then use DFO to find the optimal upper-level planning variables: DFO iteratively proposes a set of planning-level variables which are fixed and fed as setpoints to the lower-level integrated scheduling-control problem. The lower-level is then solved via traditional optimization software (Pyomo) which returns the complicating quantities that inform the upper-level objective to be optimized via DFO. As such, we manage to find a tractable solution to an otherwise intractable trilevel integrated planning-scheduling-control problem. It is worth noting that there will always be an approximation error when using surrogates, therefore, these must be trained and cross-validated appropriately.

## 2. Methodology

### 2.1. Problem Statement

We consider canonical planning, scheduling, and planning problems as described in Chu and You (2015). Using the hierarchical approach to integrate these decision levels yields the following tri-level optimization problem:

$$\begin{aligned}
 & \min_{\substack{\mathbf{x}_p, \mathbf{x}_{\text{prod}}, \mathbf{x}_s^*, \mathbf{c}_s^*, \\ \mathbf{x}_b^*, \mathbf{c}_c^*, t^*, \mathbf{x}_c^*}} f_p(\mathbf{x}_p, \mathbf{x}_{\text{prod}}) + \mathbf{c}_s^* + \mathbf{c}_c^* & (1) \\
 & \text{s. t.} \quad \mathbf{h}_p(\mathbf{x}_p, \mathbf{x}_{\text{prod}}) = \mathbf{0}, \\
 & \quad \mathbf{g}_p(\mathbf{x}_p, \mathbf{x}_{\text{prod}}) \leq \mathbf{0}, \\
 & \quad \substack{\mathbf{x}_s^*, \mathbf{c}_s^*, \mathbf{c}_c^* \\ \mathbf{x}_b^*, t^*, \mathbf{x}_c^*} \in \arg \min_{\substack{\mathbf{x}_s, \mathbf{x}_{\text{batch}}, \mathbf{c}_s, \\ \mathbf{x}_{\text{ctrl}}^*, t_{\text{ctrl}}^*, \mathbf{c}_{\text{ctrl}}^*}} f_s(\mathbf{x}_s, \mathbf{x}_{\text{batch}}, t_{\text{ctrl}}^*; \mathbf{x}_{\text{prod}}) \\
 & \quad \text{s. t.} \quad \mathbf{h}_s(\mathbf{x}_s, \mathbf{x}_{\text{batch}}, \mathbf{c}_s, t_{\text{ctrl}}^*; \mathbf{x}_{\text{prod}}) = \mathbf{0}, \\
 & \quad \mathbf{g}_s(\mathbf{x}_s, \mathbf{x}_{\text{batch}}, t_{\text{ctrl}}^*; \mathbf{x}_{\text{prod}}) \leq \mathbf{0}, \\
 & \quad \mathbf{x}_{\text{ctrl}}^*, t_{\text{ctrl}}^*, \mathbf{c}_{\text{ctrl}}^* \in \arg \min_{\substack{\mathbf{x}_c, t_{\text{ctrl}}, \mathbf{c}_{\text{ctrl}}} f_c(\mathbf{x}_c, t_{\text{ctrl}}, \mathbf{x}_{\text{batch}}) \\ \text{s. t.} \quad \mathbf{h}_c(\mathbf{x}_c, t_{\text{ctrl}}, \mathbf{c}_{\text{ctrl}}; \mathbf{x}_{\text{batch}}) = \mathbf{0} \\
 & \quad \mathbf{g}_c(\mathbf{x}_c, t_{\text{ctrl}}, \mathbf{x}_{\text{batch}}) \leq \mathbf{0}
 \end{aligned}$$

Eq. (1) denotes a tri-level optimization problem where  $f_p(\cdot), f_s(\cdot), f_c(\cdot)$  denote the objective functions,  $\mathbf{h}_p(\cdot), \mathbf{h}_s(\cdot), \mathbf{h}_c(\cdot)$  denote the equality constraints, and  $\mathbf{g}_p(\cdot), \mathbf{g}_s(\cdot), \mathbf{g}_c(\cdot)$  denote the inequality constraints of the planning, scheduling, and control problems respectively.  $\mathbf{x}_p, \mathbf{x}_{\text{prod}}$  denote the planning-specific and the production targets. The highest-level objective consists not only of  $f_p(\cdot)$  but also the optimal scheduling and operating costs appearing as  $\mathbf{c}_s^*, \mathbf{c}_c^*$  in the planning,  $\mathbf{c}_s, \mathbf{c}_{\text{ctrl}}^*$  in the scheduling, and  $\mathbf{c}_c^*$  in the

control (no scheduling cost in the control problem) that are obtained as the result of the scheduling optimization and its nested control optimization in the constraints. The mid-level scheduling problem takes the planning variables as setpoints  $\mathbf{x}_{\text{prod}}$  and optimizes the scheduling and batch variables  $\mathbf{x}_s, \mathbf{x}_{\text{batch}}$  which appear as optimal response variables in the planning level as  $\mathbf{x}_s^*, \mathbf{x}_b^*$ . The scheduling problem is also influenced by the optimal batch processing time as defined in the lower-level control problem. The control problem takes as setpoint the batch variables from the scheduling  $\mathbf{x}_{\text{batch}}$  and determines the optimal control variables, batch processing times, and operating cost  $\mathbf{x}_c, \mathbf{t}_f, c_{\text{ctrl}}$  (planning),  $\mathbf{x}_{\text{ctrl}}, \mathbf{t}_f^*, c_{\text{ctrl}}^*$  (scheduling), and  $\mathbf{x}_c, \mathbf{t}_f, c_{\text{ctrl}}$  (control) for said batch. It is worth noting that our considered planning, scheduling, and control problems would, in isolation, represent linear, mixed-integer linear, and optimal control problems, respectively.

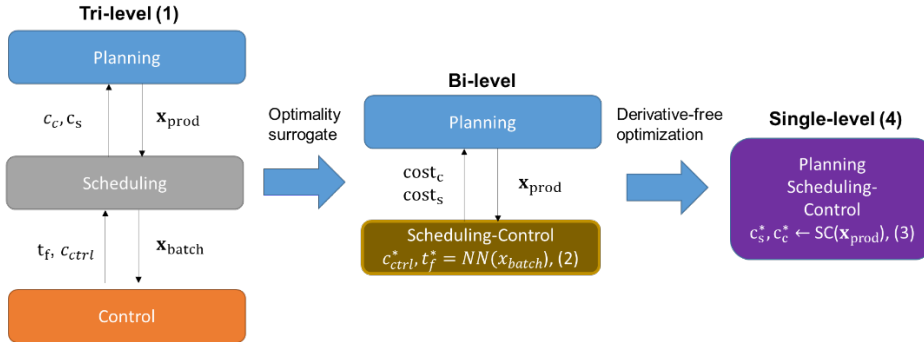


Figure 1. We reformulate tri-level planning-scheduling-control problem (1) by collapsing the control problem as an optimality surrogate into the scheduling problem (2), and then using derivative-free optimization to solve the remaining bi-level problem as a single-level black-box optimization problem (4), where each evaluation calls the scheduling model with the embedded optimal control surrogate (3).

As illustrated in Figure 1, we will start by reformulating the bi-level scheduling and control problems into a single-level formulation using the bottom-up approach, before collapsing the integrated scheduling-control level into the planning level.

## 2.2. Collapsing scheduling and control levels using optimality surrogates

By inspection of Problem (1), we notice that the optimal operating cost and batch processing time  $c_{\text{ctrl}}^*$  and  $\mathbf{t}_f^*$  are the only control-level variables appearing in upper levels. Consequently, we can formulate surrogates that map the optimal control response variables  $c_{\text{ctrl}}^*$  and  $\mathbf{t}_f^*$  as a function of the batch production targets  $\mathbf{x}_{\text{batch}}$  as the only setpoints feeding into the control level.

$$c_{\text{ctrl}}^*, \mathbf{t}_f^* = NN(\mathbf{x}_{\text{batch}}) \quad (2)$$

While there are many potential surrogates we could use (Dias and Ierapetritou, 2020), we use neural networks for their general approximation qualities, their potential for MILP reformulation and their ease of incorporation into optimization pipelines (Cecon et al., 2022). We uniformly sample batch sizes and solve the corresponding optimal control problem, logging the optimal response variables. Then, we train a small neural network of two hidden layers with five nodes each using Pytorch to predict the optimal response variables given the setpoints. We can then reformulate the neural network surrogate into mixed-integer linear constraints that can be incorporated as a Pyomo block into the planning Pyomo model using OMLT (Cecon et al., 2022). Since we are using rectified linear unit (ReLU) activation functions, we can use the full-space formulation to ensure



that the bi-level scheduling-control model can be reformulated into a single-level mixed-integer linear program (3), which is crucial for computational tractability. This allows us to use specialized solvers (Gurobi). Following the monolithic approach or using other formulations would require less efficient mixed-integer *nonlinear* solvers.

$$\begin{aligned} \mathbf{x}_s^*, \mathbf{c}_s^*, \mathbf{c}_c^*, \mathbf{x}_b^*, \mathbf{t}^*, \mathbf{x}_c^* \in \text{SC}(\mathbf{x}_{\text{prod}}) = \arg \min_{\substack{\mathbf{x}_s, \mathbf{x}_{\text{batch}}, \mathbf{c}_s, \\ \mathbf{t}_f, \mathbf{c}_{\text{ctrl}}}} f_s(\mathbf{x}_s, \mathbf{x}_{\text{batch}}, \mathbf{t}_f^*; \mathbf{x}_{\text{prod}}) \\ \text{s. t.} \quad \mathbf{h}_s(\mathbf{x}_s, \mathbf{x}_{\text{batch}}, \mathbf{c}_s, \mathbf{t}_f^*; \mathbf{x}_{\text{prod}}) = \mathbf{0}, \\ \mathbf{g}_s(\mathbf{x}_s, \mathbf{x}_{\text{batch}}, \mathbf{t}_f^*; \mathbf{x}_{\text{prod}}) \leq \mathbf{0}, \\ \mathbf{c}_{\text{ctrl}}^*, \mathbf{t}_f^* = \text{NN}(\mathbf{x}_{\text{batch}}) \end{aligned} \quad (3)$$

### 2.2.1. Single-level black-box optimization given optimal scheduling-control

After collapsing the bi-level scheduling-control problem into an integrated single-level formulation, we need to incorporate the integrated model with the higher-level planning, resulting in bi-level Formulation (4):

$$\begin{aligned} \min_{\substack{\mathbf{x}_p, \mathbf{x}_{\text{prod}}, \mathbf{x}_s^*, \mathbf{c}_s^*, \\ \mathbf{x}_b^*, \mathbf{c}_c^*, \mathbf{t}^*, \mathbf{x}_c^*}} f_p(\mathbf{x}_p, \mathbf{x}_{\text{prod}}) + \mathbf{c}_s^* + \mathbf{c}_c^* \\ \text{s. t.} \quad \mathbf{h}_p(\mathbf{x}_p, \mathbf{x}_{\text{prod}}) = \mathbf{0}, \\ \mathbf{g}_p(\mathbf{x}_p, \mathbf{x}_{\text{prod}}) \leq \mathbf{0}, \\ \mathbf{x}_s^*, \mathbf{c}_s^*, \mathbf{c}_c^* \in \text{SC}(\mathbf{x}_{\text{prod}}) \\ \mathbf{x}_b^*, \mathbf{t}^*, \mathbf{x}_c^* \end{aligned} \quad (4)$$

Inspired by Beykal et al. (2022), we intend to use derivative-free optimization (DFO) to find the planning variables that optimize the planning-level objective. DFO is closely related to black-box and simulation-based optimization. Its aim is to find input variables that optimize a system whose gradient expressions cannot be cheaply obtained. DFO solvers can be loosely categorized into direct methods, which handle function evaluations directly, and model-based methods, which rely on the intermediate construction and optimization of surrogates. As opposed to traditional numerical optimization where all variables are considered ‘decision variables’, in DFO, we only optimize the degrees of freedom, a subset of the variables that when fixed allow for the remaining quantities to be determined using the remaining equality constraints. To this end, we first decompose the planning-specific variables  $\mathbf{x}_p$  into  $\mathbf{x}_{\text{store}}$ ,  $\mathbf{x}_{\text{transp}}$ ,  $\mathbf{x}_{\text{sales}}$ , the storage, transportation, and sales variables respectively. We notice that when fixing  $\mathbf{x}_{\text{prod}}$ ,  $\mathbf{x}_{\text{transp}}$ ,  $\mathbf{x}_{\text{sales}}$ , we can obtain the storage variables  $\mathbf{x}_{\text{store}}$  by leveraging  $\mathbf{h}_p(\mathbf{x}_{\text{store}}, \mathbf{x}_{\text{transp}}, \mathbf{x}_{\text{sales}}, \mathbf{x}_{\text{prod}}) = \mathbf{0}$ . We can then obtain all quantities of the planning-level objective:  $\mathbf{c}_s^*, \mathbf{c}_c^* \in \text{SC}(\cdot)$  and  $f_p(\cdot)$ . In the DFO, we augment the planning-level objective with a penalty term on the inequalities

$$\begin{aligned} g_{\text{viol}} = \max(\mathbf{0}, \mathbf{g}_p(\cdot))^2 \\ \min_{\mathbf{x}_{\text{prod}}, \mathbf{x}_{\text{transp}}, \mathbf{x}_{\text{sales}}} f_p(\mathbf{x}_{\text{store}}, \mathbf{x}_{\text{transp}}, \mathbf{x}_{\text{sales}}, \mathbf{x}_{\text{prod}}) + \mathbf{c}_s^* + \mathbf{c}_c^* + g_{\text{viol}} \\ \text{where} \quad \mathbf{x}_{\text{store}} \leftarrow \mathbf{h}_p(\mathbf{x}_{\text{store}}, \mathbf{x}_{\text{transp}}, \mathbf{x}_{\text{sales}}, \mathbf{x}_{\text{prod}}) = \mathbf{0}, \\ g_{\text{viol}} \leftarrow \mathbf{g}_p(\mathbf{x}_{\text{store}}, \mathbf{x}_{\text{transp}}, \mathbf{x}_{\text{sales}}, \mathbf{x}_{\text{prod}}) \leq \mathbf{0}, \\ \mathbf{c}_s^*, \mathbf{c}_c^* \leftarrow \text{SC}(\mathbf{x}_{\text{prod}}) \end{aligned} \quad (4)$$

As such, we have reformulated the tri-level Problem (1) into a single-level black-box optimization problem (4), where in each evaluation of the DFO input we evaluate the remaining planning-level variables necessary to compute the planning-specific objective

and inequality penalty term. We obtain the lower-level optimal costs by calling the integrated planning-scheduling optimization instance.

2.3. Notes on the derivative-free optimization

Since the black-box evaluations are expensive, we only have a restricted budget of 400 evaluations to optimize the 178-dimensional problem. This makes the DFO problem very challenging to be solved without *a priori* insights. As such, we hot-start the solution to the problem by extracting the relevant DFO variables from the solution of the following planning-only problem:

$$\begin{aligned} \min_{\mathbf{x}_p, \mathbf{x}_{\text{prod}}} \quad & f_p(\mathbf{x}_p, \mathbf{x}_{\text{prod}}) \\ \text{s. t.} \quad & \mathbf{h}_p(\mathbf{x}_p, \mathbf{x}_{\text{prod}}) = \mathbf{0}, \\ & \mathbf{g}_p(\mathbf{x}_p, \mathbf{x}_{\text{prod}}) \leq \mathbf{0}, \end{aligned} \tag{5}$$

Given the expensive nature of the evaluations, we only have enough budget to perform a ‘local search’ around the optimum obtained from (5) rather than perform extensive exploration of the solution space. Informed by van de Berg et al. (2022), we use the trust-region method Py-BOBYQA given its highly exploitative nature.

3. Results

The alternative to solving the tri-level optimization problem is solving the planning-only problem in (5) and sending its optimal planning target as a setpoint to the scheduling problem which is optimized and finally fed to control (sequential approach). We use the solution of the planning problem of the sequential approach as an initial guess for the DFO and decrease the tri-level objective from a cost of 4.762 to -0.132 million, turning the tri-level objective profitable. The tri-level solution also translates into tangible production and scheduling changes: Figures 2A and 2B demonstrate that going from the initial guess to the DFO solution finetunes the production targets of some products (in red case) given more accurate accounting of storages costs from the scheduling level. Figures 2C and 2D illustrate that the production assignment of products P1 and P2 changes (in red case) due to a more accurate estimate of the processing time from the control problem.

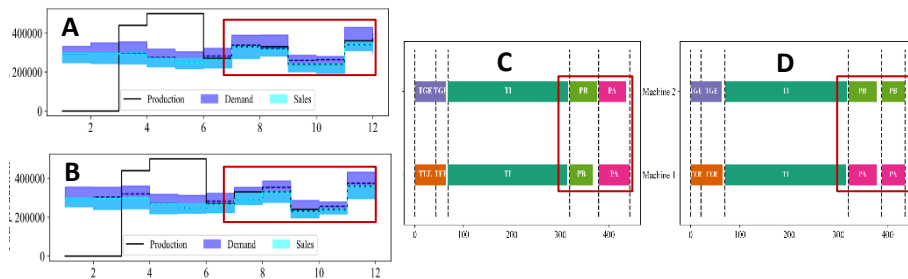


Figure 2. Figures A and B represent the production target of one product in kg over 12 weeks for the planning-only initial guess (A) and the tri-level solution (B) respectively. Figures C and D show the scheduling job assignments for all products in a particular week for the planning-only initial guess (C) and the tri-level solution (D) respectively. Changes highlighted in red.

4. Discussion

Our proposed methodology lends itself well to the decision-making frameworks of the process industries. Given the expensive nature of monolithic or hierarchical decision-making, planning, scheduling, and control problems are usually solved sequentially. Our proposed integrated approach leverages an initial guess to the DFO problem using the

single-level planning problem. As such, we can always fall back on the sequential approach if the DFO fails to make significant progress. This way, when applying exploitative DFO solvers, our proposed methodology can be interpreted as ‘fine-tuning’ the single-level planning optimum towards a tri-level objective in the available evaluation budget available to decision-makers. Our methodology begs the question if the two techniques (DFO and optimality surrogates) can be used interchangeably. There are clear trade-offs associated with both techniques. DFO for bi-level optimization has the advantage that it does not have to sacrifice solution accuracy through approximation errors. However, DFO performs best in lower dimensions. To use DFO for bi-level optimization requires all degrees of freedom in the upper-level optimization problem to be determined, which scales rapidly with the problem size, especially given our tight evaluation budget due to the expensive scheduling-control solution at each iteration. On the other hand, while optimality surrogates present the more tractable alternative, they inevitably introduce approximation error. As such, the choice of optimality surrogates should be followed by strategies that mitigate approximation error, such as cross-validation of the surrogate’s performance. Future work will apply this methodology to different case studies, investigate the potential of the surrogate approach to the planning-scheduling integration and of decision trees to be used as optimality surrogates.

## 5. Conclusion

We leverage DFO and optimality surrogates to reformulate a tri-level optimization problem into a single-level black-box optimization problem that we solve using the derivative-free optimizer Py-BOBYQA where each evaluation calls a scheduling instance with an embedded optimal control surrogate. We suggest hot-starting the solution to our problem using the solution to the planning-only optimization problem. As such, we use the remaining computational budget to approach the tri-level optimum from the hierarchical approach, which most accurately reflects decision-making. Nonetheless, we can always fall back on the ‘status quo’, i.e. the solution of the planning-only optimization problem from the sequential method. Our approach leads to tangible changes in the optimal production planning and scheduling sequence from the ‘status quo’.

## References

- S. Avraamidou, E. N. Pistikopoulos, 2019, A Bi-Level Formulation And Solution Method For The Integration Of Process Design And Scheduling. *Proceedings of the 9th International Conference on Foundations of Computer-Aided Process Design*, 47, 17-22
- B. Beykal, S. Avraamidou, E. N. Pistikopoulos, 2022, Data-driven optimization of mixed-integer bi-level multi-follower integrated planning and scheduling problems under demand uncertainty, *Computers & Chemical Engineering* 156, 107551
- F. Ceccon, & J. Jalving, J. Haddad, A. Thebelt, C. Tsay, C. Laird, R. Misener, 2022, OMLT: Optimization & Machine Learning Toolkit, arXiv preprint arXiv:2202.02414
- Y. Chu, F. You, 2015, Model-based integration of control and operations: Overview, challenges, advances, and opportunities, *Computers & Chemical Engineering*, 83, 2–20
- L.S. Dias, M.G. Ierapetritou, 2020, Integration of planning, scheduling and control problems using data-driven feasibility analysis and surrogate models. *Computers & Chemical Engineering*, 134, 106714
- S. Sachio, M. Mowbray, M. Papanthasiou, E.A. del Rio-Chanona, P. Petsagkourakis, 2021, Integrating process design and control using reinforcement learning. *Chemical Engineering Research and Design*, 183, 160-169
- D. van de Berg, T. Savage, P. Petsagkourakis, D. Zhang, N. Shah, E. A. del Rio-Chanona, 2022, Data-driven optimization for process systems engineering applications. *Chemical Engineering Science*, 248, 117135

# A graph classification approach to determine when to decompose optimization problems

Ilias Mitrai,<sup>a</sup> Prodrimos Daoutidis<sup>a\*</sup>

<sup>a</sup>*Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, 55455 MN, USA*

\* *Corresponding author, e-mail: daout001@umn.edu*

## Abstract

Decomposition-based optimization algorithms are widely used for the solution of optimization problems. However, the choice of whether a decomposition-based solution approach should be selected over a monolithic one is not apparent in general. In this work, we propose a graph classification approach for determining a-priori when to use a decomposition-based solution approach for the solution of convex mixed integer nonlinear optimization problems. We apply the proposed approach to benchmark optimization problems and analyze the predictive performance of the classifier.

**Keywords:** Graph Classification, Decomposition-based solution algorithms, convex MINLP

## 1. Introduction

Decomposition-based solution algorithms have been widely used to solve optimization problems by exploiting the underlying structure of the problem. Although these algorithms are usually employed for the solution of large-scale problems (Conejo, 2006) for some cases, such as convex Mixed Integer Nonlinear Programs (MINLPs), they are more efficient than monolithic solution approaches even for problems with few variables and constraints (Duran and Grossmann, 1986). However, the effectiveness of a decomposition-based solution approach can only be determined after an optimization problem is solved with both approaches. Although this is a reasonable task for a simple problem, it is computationally expensive for large-scale problems. Hence, a method to determine a-priori when to apply a decomposition-based solution algorithm is necessary. In this work, we will pose this as an algorithm selection problem (Rice, 1976) in which given an optimization problem  $\mathcal{P}$  and a set of algorithms  $\mathcal{A} = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ , one must find the algorithm  $\alpha^*$  that solves problem  $\mathcal{P}$  in the minimum computational time. Mathematically this problem can be stated as follows:

$$\alpha^* \in \arg \min_{a \in \mathcal{A}} m(\mathcal{P}, a), \quad (1)$$

where  $m : \mathcal{P} \times \mathcal{A} \rightarrow R$  is the solution time of problem  $\mathcal{P}$  using algorithm  $\alpha$ . This is a black box optimization problem since the function  $m$  is not known explicitly. Machine learning techniques have been used to approximate either the solution time or the solution of the problem itself with a surrogate model based on some features of the problem (Kerschke et al. (2019), Xu et al. (2011), Bonami et al. (2022), Kruger et al. (2017)). These approaches considered the selection of monolithic solution algorithms for Mixed

Integer Linear (MILPs) and Mixed Integer Quadratic Programs (MIQPs) and determined if a MILP should be solved using the Dantzig-Wolfe (DW) reformulation or Branch and Cut. In all these cases, the regression or classification is performed based on problem features obtained from the problem formulation and the preprocessing step. However, the features aggregate information about the problem and do not explicitly consider the structural coupling between the variables, which is the basis for the application of decomposition-based solution algorithms (Conejo (2006), Allman et al. (2019), Mitrai and Daoutidis (2021), Mitrai et al. (2022)).

In this paper, we propose a graph classification approach to determine if a convex MINLP should be solved using a monolithic or a decomposition-based solution approach. Specifically, we use graph convolutional neural networks (Kipf and Welling, 2016) to capture the structural and functional interactions among the variables of the problem. Given this representation, a graph classifier is trained to determine if a convex MINLP should be solved using Branch and Bound (B&B) (Gupta and Ravindram, 1985), a monolithic solution approach, or the Outer Approximation (OA) algorithm (Duran and Grossmann, 1986) a commonly used decomposition-based solution approach. The rest of the paper is organized as follows: In Section 2 we present the combined structural and functional graph representation of an optimization problem. In Section 3 we present the graph classification-based algorithm selection solution approach and in Section 4 we analyze the classifier's performance on benchmark convex MINLPs.

## 2. Graph representation of an optimization problem

We assume that the following convex MINLP must be solved:

$$\begin{aligned} \mathcal{P} : \min_{x,y} f(x,y) \\ \text{s. t. } g(x,y) \leq 0 \\ x \in \mathbb{Z}^{n_x}, y \in \mathbb{R}^{n_y} \end{aligned} \quad (2)$$

where  $n_x + n_y = n$ , and  $g, f$  are convex in the variables  $y$  and in the relaxed integer variables  $x$ . Based on our previous work (Mitrai et al. (2022), Allman et al. (2019)), we will represent problem  $\mathcal{P}$  as a graph.

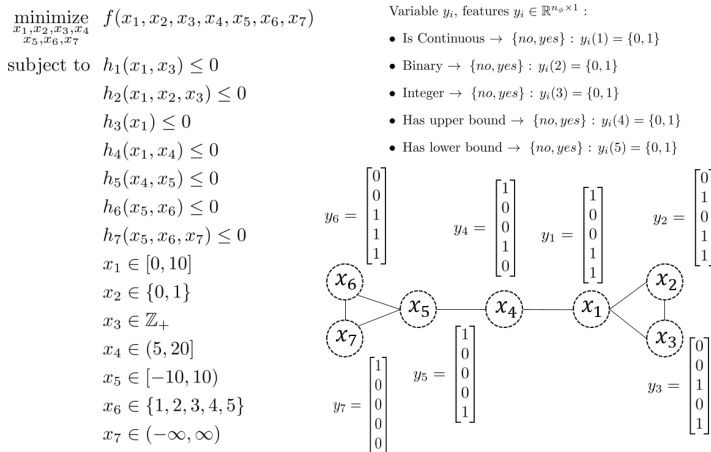


Figure 1: Graph and feature representation of an optimization problem

Three graphs can be generated to capture the structural coupling among the variables and constraints, the bipartite variable-constraint graph, the constraint graph, and the variable graph. For the rest of the paper, we will use the variable graph  $\mathcal{G}_n(V_n, E_n)$  ( $|V_n| = n$ ), where every node represents a variable and the edges represent constraints that couple two variables. In this work, we extend the aforementioned representation to a combined structural and functional graph representation of an optimization problem. Specifically, for every node in a graph, we add a set of features that capture different characteristics of the node, such as the domain of the variable and the upper and lower bounds (see Fig. 1). These features form the feature matrix  $F$ . Given this representation, a graph is represented by the adjacency matrix  $A$  and the feature matrix  $F$ , i.e.,  $\mathcal{G}_n(A, F, E)$ .

### 3. Algorithm selection via graph classification

#### 3.1. Graph classification approach

Given an optimization problem  $\mathcal{P}$  and the variable graph  $\mathcal{G}_n(A, F, E)$  the goal is to develop a classifier  $\mathcal{C}: F(\mathcal{P}) \times A(\mathcal{P}) \mapsto p \in \mathbb{R}^{N_a}$  to determine if problem  $\mathcal{P}$  should be solved using Branch and Bound (B&B) or the Outer Approximation (OA) algorithm ( $\mathcal{A} = \{B\&B, OA\}, |\mathcal{A}| = N_a = 2$ ). The inputs in the classifier are the adjacency  $A(\mathcal{P})$  and feature  $F(\mathcal{P})$  matrix (which depend on the problem  $\mathcal{P}$ ), and the output is a vector  $p \in \mathbb{R}^{N_a}$ , where  $p_i$  is equal to the probability that algorithm  $\alpha_i$  solves the problem in the minimum computational time. Under this setting, the algorithm selection problem is transformed into a graph classification problem, where a given problem  $\mathcal{P}$  is solved using the algorithm with the maximum probability  $a^* = \arg \max p$  with  $p = \mathcal{C}(F(\mathcal{P}), A(\mathcal{P}))$ . The overall framework is presented in Fig. 2.

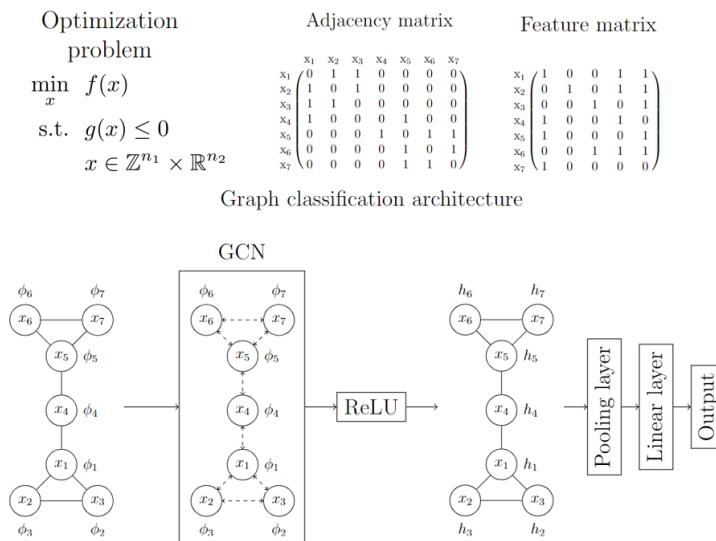


Figure 2: Learning when to decompose framework

#### 3.2. Graph classifier architecture

The graph classification task is performed using graph convolutional neural (GCN) networks (Kipf and Welling, 2016). Given an optimization problem  $\mathcal{P}_i$  (see Eq. 2) with  $N_v$  variables and  $N_m$  constraints, we can generate the variable graph  $\mathcal{G}_v(V_v, E_v)$  ( $|V_v| = N_v$ ) and obtain the adjacency  $A \in \mathbb{R}^{N_v \times N_v}$  and feature  $F \in \mathbb{R}^{N_v \times N_\phi}$  matrices where  $N_\phi$

is the number of features per node. We define as  $\phi_i$  the features of node  $i$  ( $\phi_i$  is the  $i^{th}$  row of matrix  $F$ ). We will assume that  $L$  GCN layers are stacked sequentially and the features of the nodes are updated using a GCN layer as follows:

$$H^{(1)} = \sigma \left( \widehat{D}^{-\frac{1}{2}} \widehat{A} \widehat{D}^{-\frac{1}{2}} F W^{(1)} \right)$$

$$H^{(l)} = \sigma \left( \widehat{D}^{-\frac{1}{2}} \widehat{A} \widehat{D}^{-\frac{1}{2}} H^{(l-1)} W^{(l)} \right) \forall l \in \{2, L\},$$

where  $l \in \{1, L\}$  denotes the layer number,  $\sigma$  is the activation function,  $\widehat{A} = A + I$  ( $I$  is the identity matrix of appropriate dimension),  $\widehat{D}_{ii} = \sum_{j=0}^{N_v} \widehat{A}_{ij}$  is the degree of node  $i$  and  $\widehat{D} = \text{diag}(\{\widehat{D}_{ii}\}_{i=1}^{N_v})$ .  $W^{(1)} \in \mathbb{R}^{N_\phi \times N_h}$  are the weights for the first layer,  $W^{(l)} \in \mathbb{R}^{N_h \times N_h}$  are the weights for layer  $l$ , and  $N_h$  is equal to the number of hidden features for each node. The output of the last layer  $L$  is passed through a pooling layer which creates a graph-level feature  $r \in \mathbb{R}^{N_h \times 1}$  which is equal to

$$r_i = \frac{1}{N_v} \sum_{j=1}^{N_v} H_{ij}^L \quad \forall i = 1, \dots, N_h.$$

Finally, the graph level feature is an input to a linear transformation layer, where the output is  $y \in \mathbb{R}^{N_A \times 1}$  ( $y_i$  is the probability that algorithm  $i$  solves the problem in the minimum computational time) and is equal to  $y = \Theta r + b$ , where  $\Theta \in \mathbb{R}^{N_A \times N_h}$ ,  $b \in \mathbb{R}^{N_A \times 1}$  are parameters.

This classifier can classify problems with different number of variables and constraints since the pooling layer after the last convolution layer creates the vector  $r$  which has  $N_h \times 1$  dimension. Therefore, for the final classification step, every problem is represented in the hidden features dimension space  $N_h$ .

### 3.3. Training of the graph classifier

We assume that a set of optimization problems  $\{\mathcal{P}_i\}_{i=1}^{N_{data}}$  is available and for these problems we can obtain the variable graphs  $\{\mathcal{G}_i\}_{i=1}^{N_{data}}$ , the adjacency and feature matrix  $\{A_i, F_i\}_{i=1}^{N_{data}}$ , and labels  $\{z_i\}_{i=1}^{N_{data}}$  where  $z_i$  is the algorithm that solves problem  $\mathcal{P}_i$  in the minimum computational time. Given these data, the task is to learn the classifier's parameters that will maximize its accuracy by optimizing some loss function, such as the cross entropy.

## 4. Deciding when to decompose convex MINLPs

In this section, we apply the proposed approach to benchmark convex MINLPs from the MINLPLib (2022) to determine if a given convex MINLP should be solved using Branch and Bound (Gupta and Ravindram, 1985) or the Outer Approximation algorithm (Duran and Grossmann, 1986) as implemented in the BONMIN solver (Bonami et al., 2008).

### 4.1. Feature representation of the problem and data generation for classification

We will consider the following features for every node in the variable graph: variable domain (continuous, binary, integer) and bound (upper, lower). We use one-hot encoding for the representation of these features as follows:

$$\phi(1) \in \{0,1\}: 1 \text{ if the variable is continuous and } 0 \text{ otherwise}$$

$$\phi(2) \in \{0,1\}: 1 \text{ if the variable is binary and } 0 \text{ otherwise}$$

- $\phi(3) \in \{0,1\}$ : 1 if the variable is integer and 0 otherwise
- $\phi(4) \in \{0,1\}$ : 1 if the variable has an upper bound and 0 otherwise
- $\phi(5) \in \{0,1\}$ : 1 if the variable has a lower bound and 0 otherwise

Under this representation, every node has five features ( $N_\phi = 5$ ).

#### 4.2. Data generation for classification

The MINLPlib library contains 295 convex optimization problems which are solved with both algorithms (B&B and OA) using BONMIN 1.8.8 with a maximum computational time of 3000 seconds. From the 295 problems, 227 are solved with at least one solver (151 problems are solved faster with OA and 57 with B&B), and for each problem we obtain the adjacency and feature matrix  $(A_i, F_i)$ . Overall, we obtain the dataset  $\mathcal{D} = \{(A_i, F_i), y_i\}_{i=1}^{N_{data}}$ , where 66% of the data points have label OA and 34% B&B.

#### 4.3. Graph classification architecture and implementation

We split the dataset  $\mathcal{D}$  at random into a training set and a testing set. The training set has 197 data points and the testing set has 30 data points picked at random (15 random data points have label OA and 15 have label B&B). The graph classifier has four convolution layers ( $L = 4$ ), a global mean pooling layer, a linear layer (with dropout probability equal to 0.5), the number of hidden features is 12 ( $N_h = 12$ ), and ReLU is used as the activation function. For the training, we use the Adam algorithm (Kingma and Ba, 2014) for 50 iterations with random initialization, the learning rate equal to 0.005, and batch size equal to 10. The loss function is the cross entropy and different weights are assigned to the two classes to account for the imbalance in the training dataset. The weights are computed as follows:  $\omega_i = N_{data} N_\alpha^{-1} \left( \sum_{j=1}^{N_{data}} 1_i(y_j) \right)^{-1}$ , where  $1_i$  is the indicator function, i.e.,  $1_i(y_j) = 1$  if  $y_j = i$  and 0 otherwise. The weight for class B&B is 1.6148 and for class OA is 0.7243. The GCN is implemented in PyTorch Geometric (Fey and Lenssen, 2019) and the training is done using PyTorch (Paszke et al., 2015).

#### 4.4. Classification results

The performance metrics of the classifier on the testing data set are presented in Table 1.

Table 1: Performance metrics for the graph classifier on the testing data set

Accuracy	0.9	F1 score	0.9	Confusion matrix		
				Predicted label		
Label	Precision	Recall	F1 score	True label	OA	B&B
OA	0.83	1.00	0.91	OA	15	0
BB	1.00	0.80	0.89	B&B	3	12

From the results, we observe that the classifier has 90% accuracy and the F1 score is 0.9. A few data points with label B&B are misclassified as OA. This misclassification could be caused by the small size and the imbalance in the training set. Finally, we observe that the classifier can accurately predict all the problems that are solved faster with the OA algorithm.

## 5. Conclusions

The selection of the appropriate solution method for a general optimization problem can be a challenging and time-consuming step. In this work, we propose a graph classification approach to determine whether a convex MINLP should be solved using a decomposition-based or a monolithic-based solution approach by considering simultaneously the



structural and functional coupling between the variables and the constraints of a problem. Application of this approach to benchmark problems shows that the graph classifier can achieve 90% accuracy. However, more data are necessary to reduce the imbalance in the dataset and improve the performance of the classifier.

## 6. Acknowledgment

This work was supported by the National Science Foundation (NSF-CBET, award number 1926303) and a Doctoral Dissertation Fellowship (DDF) from University of Minnesota

## References

- A. Allman, W. Tang, P. Daoutidis, DeCODE: a community-based algorithm for generating high-quality decompositions of optimization problems, *Optimization and Engineering* 20 (4) (2019) 1067–1084.
- A. J. Conejo, E. Castillo, R. Minguez, R. Garcia-Bertrand, *Decomposition techniques in mathematical programming: engineering and science applications*, Springer Science & Business Media, 2006.
- A. Paszke, et al., Pytorch: An imperative style, high-performance deep learning library, in: H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, R. Garnett (Eds.), *Advances in Neural Information Processing Systems* 32, Curran Associates, Inc., 2019, pp. 8024–8035.
- D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv, preprint arXiv:1412.6980 (2014).
- I. Mitrai, W. Tang, P. Daoutidis, Stochastic blockmodeling for learning the structure of optimization problems, *AIChE Journal* 68 (6) (2022) e17415.
- I. Mitrai, and P. Daoutidis, 2021. Efficient solution of enterprise-wide optimization problems using nested stochastic blockmodeling. *Industrial & Engineering Chemistry Research*, 60(40), pp.14476-14494.
- J. R. Rice, The algorithm selection problem, in: *Advances in computers*, Vol. 15, Elsevier, 1976, pp. 65–118.
- Kerschke, P., Hoos, H.H., Neumann, F. and Trautmann, H., 2019. Automated algorithm selection: Survey and perspectives. *Evolutionary computation*, 27(1), pp.3-45.
- L. Xu, F. Hutter, H. H. Hoos, K. Leyton-Brown, Hydra-mip: Automated algorithm configuration and selection for mixed integer programming, in: RCRA workshop on experimental evaluation of algorithms for solving problems with combinatorial explosion at the international joint conference on artificial intelligence (IJCAI), 2011, pp. 16–30.
- M. A. Duran, I. E. Grossmann, An outer-approximation algorithm for a class of mixed-integer nonlinear programs, *Mathematical programming* 36 (3) (1986) 307–339.
- M. Fey, J. E. Lenssen, Fast graph representation learning with pytorch geometric, arXiv preprint arXiv:1903.02428 (2019).
- M. Kruber, M. E. Lübbecke, A. Parmentier, Learning when to use a decomposition, in: *International conference on AI and OR techniques in constraint programming for combinatorial optimization problems*, Springer, 2017, pp. 202–210.
- MINLP Lib (2022) Mixed-integer nonlinear programming library. <http://www.minlp lib.org/>. Accessed 27 May 2022
- O. K. Gupta, A. Ravindran, Branch and bound experiments in convex nonlinear integer programming, *Management science* 31 (12) (1985) 1533–1546.
- P. Bonami, A. Lodi, G. Zarpellon, A classifier to decide on the linearization of mixed-integer quadratic problems in cplex, *Operations Research* (2022).
- P. Bonami, L. T. Biegler, A. R. Conn, G. Cornuéjols, I. E. Grossmann, C. D. Laird, J. Lee, A. Lodi, F. Margot, N. Sawaya, et al., An algorithmic framework for convex mixed integer nonlinear programs, *Discrete optimization* 5 (2) (2008) 186–204.
- T. N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, arXiv preprint arXiv:1609.02907 (2016).

# An Enhanced Group-Interaction Contribution Method for the Prediction of Freezing Point and Enthalpy of Fusion for Pure Organic Compounds

Khaled Mansour<sup>a,\*</sup>, Mourad Korichi<sup>a</sup>

<sup>a</sup>*Univ. Ouargla, Fac. des Sciences Appliquées, Lab. Dynamique, Interactions et Réactivité des Systèmes, BP 511 Ouargla, Algeria.*  
*khaledmansour47@yahoo.fr*

## Abstract

In this work, we developed new models using an enhanced group-interaction contribution method for the prediction of freezing point and enthalpy of fusion of pure organic compounds. These enhancements proved their efficiency by results presented on an earlier published paper, (Mansour and Korichi (2016)), which includes the extension of the training dataset by new experimental data, add 85 new structural groups and additional correction groups to distinct between isomers including Cis or Trans as well as Ortho, Meta or Para structures. 1124 and 501 experimental data points used for the new ( $T_{fp}$ ) and ( $\Delta H_f$ ) models development and testing respectively.

**Keywords:** Thermophysical Properties Prediction, Group-Interaction Contribution, Freezing point, Enthalpy of Fusion, Pure Organic Compounds.

## 1. Introduction

Freezing point is of high importance property for pure organic compounds characterization; it is influenced by the presence of other substances (impurity) made it an important criterion for purity.

The enthalpy of fusion also known as the latent heat of fusion ( $\Delta H_f$ ) is an important physical property of the solid state, and can be useful in correcting thermo-chemical data to a standard state when combined with other thermodynamic properties, Wang et al., (2010).

The knowledge of these two properties for pure organic compounds is a great challenge for any new process and/or products design, especially when taking in consideration the enormous number of molecules, and not forgetting time and costs. An alternative solution available to solve the lack of experimental data and the request for these properties, prediction methods. Numerous based group-contribution methods and models developed for the prediction of freezing point and enthalpy of fusion in order to fill the gap between new available molecules and the properties values, Boethling et al., (2000); Poling et al., (2004).

Group contribution methods divide the molecules in to groups (fragments; atoms, bonds or group of atoms, etc) each group has a partial value contribute in the total value of the calculated property. These contributions calculated from known experimental data. Property of a compound obtained by a summing up the values of all contributions presented in the molecule via the developed mathematical model. Kolská, et al., (2012).

Marrero and Pardillo (1999) proposed, the group-interaction contribution (GIC) which defined a fragment as the interaction between two simple groups e.g the interaction between  $-\text{CH}_3$  and  $>\text{CH}_2$  make a group, this approach was not applied for the development of models allowing the prediction of freezing point and enthalpy of fusion.

The appearance of new experimental data helps the development of new models, an enhanced group interaction contribution method published, Mansour and Korichi (2016) presenting new models using the (GIC) approach. In this paper, the enhanced approach applied for the development of new models for the prediction of freezing point and enthalpy of fusion at the freezing point.

## 2. Models development

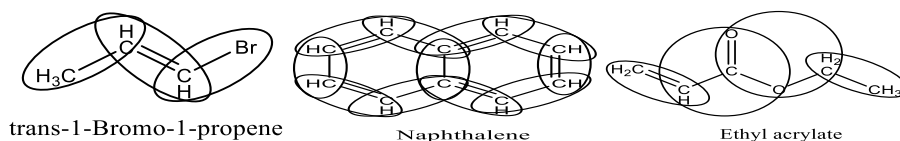
### 2.1. Approach

The mean benefits of the new developed models are accuracy and isomers distinction (due to new available data, which conduct to new added groups). In this work, the (GIC) approach adopted for the prediction of these properties with some modifications in order to enhance the models prediction efficiency:

- Initially, after studying the utility of the groups, some of them were eliminated because they are useless, e.g. ( $=\text{CH}_2$  and  $=\text{CH}_2$ ), which appears only in ethene, this group has no use because it will not be used in the prediction of new compounds properties, and we have the properties of the mentioned compound.
- Secondly, modify some structural groups, such as ( $-\text{CH}_3$  and  $-\text{CN}$ ) and ( $-\text{CH}_2$  and  $-\text{CN}$ ), which they will be replaced by only one group ( $-\text{C}\equiv$  and  $\equiv\text{N}$ ) to reduce the groups number in order to simplify the method.
- Then, add new groups according to the new available data in our database taking in consideration the approach philosophy in generating groups, for example ( $>\text{CH}-$  and  $-\text{Br}$ ), ( $>\text{CH}-$  and  $-\text{C}\equiv$ ) are new added groups.
- Also, add some groups based on the functional chemical groups e.g.  $-\text{OH}$  and  $-\text{OR}$  for the distinction between alcohols and ethers,  $\text{O}=\text{C}-\text{OH}$  and  $\text{O}=\text{C}-\text{O}-\text{R}$  for the distinction between acids carboxylic and esters...etc, to enhance the method isomers distinction ability.
- Finally, introduce the differentiation between the aromatic and double bonds, which was not the case for the method of Marrero and Pardillo (1999).

### 2.2. Groups selection

Selection of contributions is based mainly on the interaction between two groups and the type of this interaction (bonds type; simple, double, triple or aromatic), these groups are similar to the groups used by Marrero and Pardillo (1999) with new added groups according to the modifications mentioned above, Fig. (1) shows the group's selection methodology. The groups used for the generation of contributions illustrated in table (1).



**Fig. 1:** Groups selection methodology.

**Table 1:** Groups used for the generation of group-interaction ((a), aromatic; (s), single; (d), double; (r), rig. For the correction groups, we have proposed cis, trans, ortho, meta and para.)

<b>Groups with single bond</b>					
-CH <sub>3</sub>	>CH <sub>2</sub>	>CH-	>C<		
<b>Groups with double bond</b>					
=CH <sub>2</sub>	=CH- (s)	-CH= (d)	=C< (s)	>C= (d)	=C=
<b>Groups with triple bond</b>					
	≡CH			≡C-	
<b>Groups within rig with single bond</b>					
	>CH <sub>2</sub> (r)	>CH-(r)	>C<(r)		
<b>Groups within rig with double bond</b>					
	=CH-(s) (r)	-CH= (d) (r)	>C= (d) (r)		
<b>Groups within aromatic rig</b>					
	=CH- (a)		=C< (a)		
<b>Oxygen group with single bond</b>			<b>Amine group with single bond</b>		
	-O-			>NH (s)	
<b>Sulfur group with single bond</b>			<b>Amine group within rig with double bond</b>		
	-S- (s)			-N= (r)	
<b>Acetone group within rig</b>			<b>Acetone groups with single bond</b>		
	>CO(r)			>CO	
<b>Isocyanate group with double bond</b>					
	=N- (s)			-N= (d)	

### 2.3. Database

Special care was taking while the collection of the experimental data by choosing the right sources, the book of Yaws and Narasimhan (2008) was the main sources of our data base, beside, other sources were also used as verification for the database; table (2) resumes the available data points for each property, for every class.

**Table 2:** Properties and their experimental data points in the database

Class*	CH	O	N	S	H	multi	Total
T <sub>fp</sub>	422	299	123	60	192	29	1124
ΔH <sub>f</sub>	265	95	37	33	62	9	501

\*CH, hydrocarbons; O, oxygenated; N, nitrogenated; S, sulfur containing; H, halogenated; multi, multifunctional compounds.

### 2.4. Proposed models

Two types of regression used for the obtaining of the models. The first is a nonlinear regression using 100% of the data points, which allows as getting the adjusted parameters a<sub>1</sub> and b<sub>1</sub>. The second one is a linear regression, the data points was divided in two parts, 90% was used as training set for the obtaining of the contributions values and the adjusted parameters c<sub>1</sub> and c<sub>2</sub>, and 10% reserved for the testing of models performances.

Several models with different types of functions passed through the processes of models development and testing as mentioned above, and the adopted models for the prediction of ( $T_{fp}$ ) and ( $\Delta H_f$ ) are respectively equations (1) and (2):

$$\frac{T_{fp}}{(K)} = a_1 * \left( \sum_i N_i C_i + c_1 \right)^{b_1} \quad (1) \quad \text{Where } a_1=10.944, b_1= 1.522 \text{ and } c_1=3.6155$$

$$\frac{\Delta H_f}{(kJ/mol)} = T_{fp} * \left( \sum_i N_i C_i + c_2 \right)^2 \quad (2) \quad \text{Where } c_2=0.14327$$

### 3. Results and discussion

The mean tools for the validation of our models are: MAPE; Mean Absolute Percentage Error, MAE; Mean Absolute Error and SD; Standard Deviation Qiang *et al.* (2009) which also will be used for the judgment of the proposed models performances, beside the two determination coefficients  $R^2$  and adjusted  $R^2$ . The values of these parameters calculated for the total data points and for each class, as illustrated in table (3).

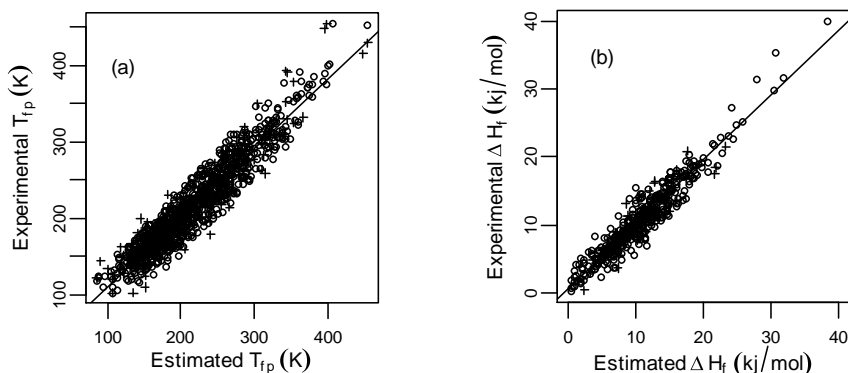
**Table 3:** Performances of the proposed models for the prediction of  $T_{fp}$  and  $\Delta H_f$

Class	$T_{fp}$				$\Delta H_f$			
	Data points	MAPE (%)	MAE (K)	SD (K)	Data points	MAPE (%)	MAE (kJ/mol)	SD (kJ/mol)
CH	421	8.19	14.76	18.48	265	17.32	1.24	1.63
O	299	6.44	14.79	18.77	95	10.25	1.07	1.56
N	123	6.12	14.55	18.74	37	10.11	1.15	1.6
S	60	6.93	11.82	15.90	33	10.7	0.84	1.3
H	192	8.27	16.76	20.30	62	17.76	1.01	1.43
Multi	29	4.60	11.69	15.68	9	11.06	1.31	2.03
Testing set	112	8.35	16.78	20.36	50	16.52	1.55	1.98
Total	1124	7.35	14.85	18.72	501	14.95	1.15	1.58
		$R^2 = 0.901$	$Adj. R^2 = 0.879$			$R^2 = 0.905$	$Adj. R^2 = 0.868$	

The new proposed models present accurate results for all the classes according to the validation parameters presented in the table (3), the hydrocarbon classes for both models offer the most accurate results among all the classes.

The value of the two determination coefficients  $R^2$  and adjusted  $R^2$  for the two models are very high and close to one unit, which represent another indication for the accuracy and reliability of the suggested models.

Beside validation parameters, distributions between experimental values and estimated values for both properties presented in Fig. (2) as validation tool for the chosen models. As illustrated, we can observe the accuracy obtained using the proposed models, these figures testify to an excellent reliability presented by the new models.



**Fig. 2:** Distributions of experimental and estimated values for (a)  $T_{fp}$ ; (b)  $\Delta H_f$  for training set (○) and testing set (+).

**Table 4:** Comparison between our model and earlier published models for the prediction of  $T_{fp}$

Methods	Data Points	Groups number	MAPE (%)	MAE (K)	SD (K)
Constantinou and Gani (1994)	312	66 1 <sup>st</sup> order	8.90	17.39	22.51
		41 2 <sup>sd</sup> order	7.23	14.03	18.28
Marrero and Gani (2001)	1547	166 1 <sup>st</sup> order	9.30	24.90	33.87
		117 2 <sup>sd</sup> order	7.90	21.41	29.52
		62 3 <sup>rd</sup> order	7.60	20.22	27.67
Wang, et al., (2008) CPG	730	93	6.67	14.46	NA
<b>This work</b>	<b>1124</b>	<b>203</b>	<b>7.35</b>	<b>14.85</b>	<b>18.71</b>

**Table 5:** Comparison between our model and earlier published models for the prediction of  $\Delta H_f$

Methods	Data points	Groups number	MAPE (%)	MAE (kJ/mol)	SD (kJ/mol)
Marrero and Gani (2001)	711	131 1 <sup>st</sup> order	18.30	2.58	4.16
		68 2 <sup>sd</sup> order	16.40	2.32	3.88
		40 3 <sup>rd</sup> order	15.70	2.17	3.65
<b>This work</b>	<b>501</b>	<b>143</b>	<b>14.95</b>	<b>1.15</b>	<b>1.58</b>

Tables 4 and 5 presents the comparison between the proposed models and the models available in the literature, using the validation parameters MAPE, MAE and SD. Although, the difference between the accuracy of new models and old ones is not great, but the extension of the application range and isomers distinction gives more power to the new models, beside the simplicity of the method.

#### 4. Conclusion

We presented new models allowing the estimation of ( $T_{fp}$ ) and ( $\Delta H_f$ ); these models developed via an enhanced group-interaction contribution method. Comparing with other methods available in literature, the proposed models showed excellent results concerning reliability; accuracy and application range.

Although, the (GIC) method has no theoretical background on how the groups are selected, but it bases on the bound type as criteria for the groups selection, which enhance its performance concerning the distinction between structural isomers Cis, Trans and Ortho, Meta and Para, which offers more accuracy for the prediction of the mentioned properties.

#### 5. References

- B.E. Poling, J.M. Prausnitz and J.P. O'Connell, 2001, *The Properties of Gases and Liquids*, 5th Edition, Mc GRAW HILL, New York.
- C.L. Yaws and P.K. Narasimhan, 2008, *Critical properties and acentric factor-organic compounds*, *Thermophysical Properties of Chemicals and Hydrocarbons*, Texas, William Andrew, pp 01-809.
- K. Mansour and M. Korichi. 2016, *An Enhanced Group-Interaction Contribution Method for the Prediction of Normal Boiling and Critical points*, 26<sup>th</sup> European Symposium on Computer Aided Process Engineering 1<sup>st</sup> Edition, Vol, 38, pp1237-1242.
- K.G. Joback and R.C. Ried, 1987, *Estimation of Pure-Component Properties from Group-Contributions*, *Chemical Engineering Communications*, Vol.57, pp 233-243.
- Kolská, Z., Záborský, M., & Randova, A. (2012). *Thermodynamics - Fundamentals and Its Application in Science*. IntechOpen. Chapter 6: Group Contribution Methods for Estimation of Selected Physico-Chemical Properties of Organic Compounds. pp. 135-162
- L. Constantinou and R. Gani, 1994, *New Group - Contribution Method for the Estimation of Properties of Pure Organic Compounds*, *AIChE Journal*, Vol 10, pp 1697-1710.
- M.J. Marrero and F.E. Pardillo. 1999. *Estimation of Pure Compounds Properties Using Group-Interaction Contribution*, *AIChE Journal*, Vol 45, pp 615-621.
- R.S. Boethling, D. Mackay and with a foreword by W.J. Lyman, 2000, *Handbook of Property Estimation Methods for Chemicals: Environmental and Health Sciences*, CRC Press LLC. Florida.
- W. Qiang, J. Qingzhu, Z. Zhenyan and D.Yu, 2010, *Position Group Contribution Method for the Prediction of the Fusion Enthalpy of Organic Compounds*, 2010 AIChE Annual Meeting, New York, USA.
- W. Qiang, M. Peisheng and N. Shifeng, 2009, *Position Group Contribution Method for Estimation of Melting Point of Organic Compounds*, *Chinese Journal of Chemical Engineering*, Vol 17, N° 3, pp 468-472.

# Single and Multi-Objective Superstructure Optimization of an Integrated Continuous Multistage Reaction-Crystallization-Filtration Process with Recycles

Jiaxu Liu, Brahim Benyahia\*

*Department of Chemical Engineering, Loughborough University, Loughborough, LE11 3TU, United Kingdom*

*B.Benyahia@lboro.ac.uk*

## Abstract

A superstructure optimization methodology is developed to identify optimal design and operation strategies of a combined continuous multistage reaction and crystallization process. The production of aspirin is used as a case study. The main objective is to maximize yield and minimize wastes while meeting tight critical quality constraints. The synthetic pathway involves a two-step reaction, which runs continuously, followed by a multistage seeded continuous cooling antisolvent crystallization process, and finally a wash filtration stage. To reduce the environmental footprint of the process and enhance its circularity, the filtrate (mother liquor) is purified using a multistage purification process then made available as recycle streams to the reaction and crystallization steps. The superstructure also considers that part of the filtered crystals can be recycled to the crystallization stages. A purge is used to control the level of impurities present in the system. Based on the proposed approach, several optimization scenarios were developed with a large set of decision variables, including jacket temperatures, antisolvent flow rates, recycle flow rates, residence times, and recycled crystal mass flow rate. The objective functions include maximizing the mean crystal size, and yield and minimizing the coefficient of variation (CV). It was shown that a maximum mean crystal size of 628  $\mu\text{m}$ , a maximum yield of 87.91% and a minimum CV of 0.3596 can be achieved using the single objective optimization options. In addition, a multi-objective mixed integer nonlinear optimization problem was proposed to identify the best design and operation compromises captured by the Pareto front. Finally, a multicriteria decision-aiding approach based on the multi-attribute utility theory was used to rank the optimal Pareto trade-offs and help identify the most suitable solution.

**Keywords:** Multistage Reaction and Crystallization, Superstructure Optimization, MINLP, Multi-Objective Optimization, Multicriteria Decision Aiding.

## 1. Introduction

Over the last decade, continuous processing and flow chemistry have gained increasing interest in pharma as the most cost-effective, flexible, and environmentally friendly operation option. Despite the increased adoption of continuous processing, the development of fast and systematic approaches to synthesize integrated continuous processes and the identification of robust operating strategies is still an open topic in the pharmaceutical industry. Despite the high costs of pharma grade raw materials, which includes high purity reagent, intermediates and solvents, the implementation of recycles to rationalize resources consumption and minimize wastes is still not a common practice in pharma, and only a few investigations have been published to date (Benyahia 2018).



However, the presence of impurities and the inherent stringent regulatory requirements make the design and implementation of recycling even more challenging in pharma. As such, the development of more rigorous and systematic design and operation strategies is critical to demonstrate the benefits of the recycles and reduce inherent risks.

Pharmaceutical upstream processing, also referred to as drug substance manufacturing, commonly involves several reaction and purification steps which often involve a crystallization step followed by a wash-filtration. Often, the filtrate (mother liquor), which in the current practice is considered as waste, contains significant amounts of residual API, reagents, and solvents. However, the filtrate may also contain substantial amounts of impurities. Consequently, the design of an effective and environmental process with recycles requires effective purification technologies of the filtrates, and most importantly, optimal and robust design and operating strategies.

To address the optimal design problems from a more general perspective while embedding circularity principles, several operating and design options were introduced to the system along with systematic recycling strategies, which include recycling both solid products (crystals) and purified filtrate (mother liquor). In addition, optimal seeding policies were also considered as a possible option in all crystallization stages. These generic design options provide a solid basis for a systematic and effective process synthesis of multistage reaction and crystallization processes based on a superstructure optimization framework. The objective is to maximize flexibility and enhance the targeted key performance indicators such as the mean crystal size, yield, and coefficient of variation (CV). The generic optimal design problem, which captures all possible design scenarios, will be introduced in the following part. The resulting optimization problem is a mixed integer nonlinear programming problem (MINLP) which will deliver the optimal design options and the most effective recycling strategies. In addition to the single objective MINLP problem, a multi-objective MINLP is developed to identify the best design and operation compromises. Furthermore, the resulting set of optimal Pareto solutions is analyzed and ranked based on a multicriteria decision-aiding method to help determine the best design and operation option.

## **2. Method**

The synthesis and purification of aspirin is used as a case study. This involves two continuous reaction steps (Joiner, Billeter et al. 2014), a multistage continuous crystallization (Liu and Benyahia 2022), and a filtration stage. The mathematical models and experimental data were obtained from the literature and the proposed superstructure is shown in figure 1. The fresh feed and antisolvent are added to each stage separately at a fixed total mass flowrate, used as a constraint in the optimization problems. Crystal seeds and recycles (both liquid and crystals) can be introduced at any crystallization stage. Before recycling, the filtrate or mother liquor, obtained at the filtration stage, goes through a multistage membrane purification process to increase the concentration of API (Espinoza-Gomez, Saucedo-Castillo et al. 2018) but must be delivered at undersaturated level to avoid crystallization in the recycle stream. A purge stream is also used to control the level of impurities in the process.

Combining the reaction and crystallization allows us to investigate the impact of recycling unreacted reagents and even API and solvents on the overall process performance. Here the recycle is connected to the second reactor due to the fact that reaction 1 is a water-free reaction and the addition of water will significantly reduce reaction yield. The potential advantage is to reduce the antisolvent (water) in the recycled mother liquor and

recycle the API in the mother liquor, which will increase the overall yield and decrease the SA waste (all added water will end up in the SA waste stream).

Based on these generic settings, a series of optimization scenarios were developed by manipulating a large set of decision variables, namely: the jacket temperatures, antisolvent flow rates, feed flow rates, seeding policies, liquid recycles, crystals recycles, and a set of binary (integer) variables which determine whether a stream is required or not (0/1). The objective of the superstructure optimization problem is to maximize the mean crystal size, yield or minimize the CV. These criteria are optimized individually using single objective optimization strategies, and then a combination of these criteria is addressed using a multi-objective optimization approach to deliver the best design and operation compromises. Finally, the Pareto optimal solution are classified using the Multi-Attribute Utility Theory (MAUT) (Benyahia et al. 2011).

### 2.1. Formulation of the optimization problems

Firstly, a MINLP is proposed based on the generic flow diagram depicted in figure 1. Each of the recycles, fresh feeds and seeding streams is allocated a real variable, to determine the required flow rate, and an integer variable, to determine whether the stream is needed. The mathematical formulation of the optimization problem is shown in equation 1. The objective function can be the mean crystal size, CV, Yield (for single objective optimization) or a combination of the above (multi-objective optimization).

In the mathematical formulation below (equation 1),  $T_{j,i}$  is the temperature of jacket temperature ( $^{\circ}\text{C}$ ),  $F_{AS,i}$  is the antisolvent flow rate (g/min),  $F_{feed,i}$  is the feed flow rate (g/min), and  $F_{seed,i}$  is the suspended seed flow rate (g/min).  $\mu_{0,seed}$ ,  $d_{seed}$  are the number density of the suspended seed solution and crystal size ( $\mu\text{m}$ ) respectively.  $R_{l,i}$ ,  $R_{p,i}$  are the recycles flow rate of mother liquor (g/min) and crystals (g/min).  $I_i$  is the binary (integer) variable associated with the  $i$ th stage.

The constraints C1 to C9 represent the upper bound and lower bounds of the decision variables. C10 is used to allow cooling only, and C11-C13 defined the final constraints of the jacket temperature and antisolvent ratio. The supersaturation should be maintained above 1 during the process as no dissolution is allowed.

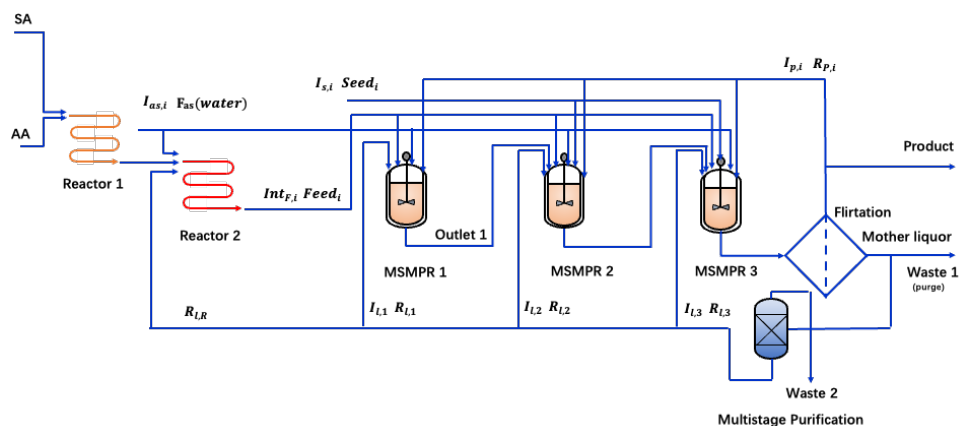


Figure 1 Generic superstructure of the multistage reaction and purification process with recycles.

$$\begin{aligned}
& \min_{T_{j,i}, F_{AS,i}, F_{feed,i}, F_{seed,i}, \mu_{0,seed}, d_{seed}, R_{l,i}, R_{p,i}, I_{S,i}, I_{F,i}, I_{AS,i}, I_{p,i}, I_{l,i}} \{-d, -Yield, CV\} \\
& s.t. \quad \dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{p}, t) \quad \mathbf{x}_{t=0} = \mathbf{x}_0 \\
& \quad 0 = \mathbf{h}(\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{p}, t) \\
& \quad C1: 20 \leq T_{j,i} \leq 40 \\
& \quad C2: 0 \leq F_{AS,i} \leq 20 \\
& \quad C3: 0 \leq F_{feed,i} \leq 30 \\
& \quad C4: 0 \leq F_{seed,i} \leq 20 \\
& \quad C5: 0 \leq \mu_{0,seed} \leq 10^6 \\
& \quad C6: 0 \leq R_{l,j} \leq 30 \\
& \quad C7: 0 \leq R_{p,i} \leq 5 \\
& \quad C8: 50 \leq M_{l,max,k} \leq 500 \\
& \quad C9: 50 \leq M_{l,max,3} \leq 10000 \\
& \quad C10: T_{i+1} \leq T_i \\
& \quad C11: \omega_{as,k} \leq 60\% \\
& \quad C12: \sum F_{feed,i} = 25 \\
& \quad C13: \omega_{as,3} = 60\% \\
& \quad C14: S_i \geq 1 \\
& \quad C15: I_{S,i}, I_{F,i}, I_{AS,i}, I_{p,i}, I_{l,i} = 0 \text{ or } 1 \\
& \quad i = 1,2,3, k = 1,2
\end{aligned} \tag{1}$$

### 3. Results and discussion

#### 3.1. Single objective optimization

The MINLP problem described above was solved using a genetic algorithm (GA) in MATLAB. The maximum mean crystal size of 628  $\mu\text{m}$ , 87.91% yield, and 0.36 CV were obtained based on the three single objective optimization scenarios. Here, for the sake of brevity, only the first scenario which maximizes the mean crystal size, is discussed in detail. The corresponding optimal process design shows that the optimal solution, shown in Table 1, indicates that all fresh feed must be delivered to the first crystallization stage along with the entire liquid recycle stream (purified mother liquor) to promote growth. Compared to the feed, the purified mother liquor has a lower API concentration. Therefore, its introduction into the second and third stage may dilute the crystallization medium and significantly reduce supersaturation, which in turn leads to reduced crystal growth. The optimal design scenario, in the case maximization of the mean crystal size, recommends that larger residence times are required in the second and third crystallization stages to promote growth, while the first crystallization is used as a nucleator. Furthermore, the crystals recycle streams are implemented at all crystallization stages at their maximum allowed flow rates (upper bound) to prioritize the production of very large crystals. However, the implementation of the crystal recycles strongly depend on the targeted CV value and may be removed under different CV considerations.

#### 3.2. Multi-objective optimization

A multi-objective optimization problem is developed to optimize three critical objectives to help identify the best design and operation compromises. Here, the mean crystal size and yield are maximized while the CV is minimized. The optimization problem is subject to the same set of constraints as those in the problem formulated above. The corresponding Pareto solutions are shown in the 2D bubble map, where the grey scale and

Table 1. Optimal solutions (real and integer variables) for scenario 1.

	$T$ (C°)	$F_{as}$ (g/min)		$R_{ml}$ (g/min)		$R_p$ (g/min)		$F_{feed}$ (g/min)		$M_{max}$ (g)	$F_{seed}$ (g/min)	
MSMPR 1	$T_{j,1}$	$I_{as,1}$	$F_{as,1}$	$I_{l,1}$	$R_{ml,1}$	$I_{p,1}$	$R_{p,1}$	$I_{F,1}$	$F_{feed,1}$	$M_{max,1}$	$I_{S,1}$	$F_{seed,1}$
	33.87	1	9.81	1	0.98	1	4.94	1	20.83	354.56	1	1.44
MSMPR 2	$T_{j,2}$	$I_{as,2}$	$F_{as,2}$	$I_{l,2}$	$R_{ml,2}$	$I_{p,2}$	$R_{p,2}$	$I_{F,2}$	$F_{feed,2}$	$M_{max,2}$	$I_{S,2}$	$F_{seed,2}$
	33.45	1	5.90	0	0	1	5.00	0	0	500	0	0
MSMPR 3	$T_{j,3}$	$I_{as,3}$	$F_{as,3}$	$I_{l,3}$	$R_{ml,3}$	$I_{p,3}$	$R_{p,3}$	$I_{F,3}$	$F_{feed,3}$	$M_{max,3}$	$I_{S,3}$	$F_{seed,3}$
	27.00	1	2.81	1	0.01	1	4.97	0	0	1000	1	2.95
Optimal Seeding properties	$\mu_0$					23909						
	size ( $\mu m$ )					24.40						
	Mean crystal size ( $\mu m$ )					628						

bubble size represents the mean crystal size (third objective function). To help the decision maker identify few alternative optimal designs and operation options for the experimental verification, the multi attribute utility theory (MAUT) is used to rank all Pareto optimal solutions. The 210 Pareto solutions were ranked according to the utility function scores. The size of the circle represents the mean crystal size. It can be found in the 2D Pareto solutions distribution shape (figure 2 (largest)) that the mean crystal size and CV performance go in opposite directions. The crystals lie at the right side of the shape, where CV is the worst. However, the yield doesn't vary significantly but overall stays within acceptable range.

### 3.3. Combined multistage reaction and crystallization process

Here the reactions stages are also optimized. By adding the ratio of SA and AA, represented by  $r$ , as one of the decision variables, the conversion of SA can be manipulated. With the optimal operation profile, a mean crystal size of 307  $\mu m$  can be obtained based on the proposed superstructure optimization.

Compared to scenario 1 (maximize mean crystal size), the optimal solution requires higher seeding flowrates, but the required seed concentration/density is lower compared to the previous case. It is because the seed is suspended in a saturated solution, which has a higher solubility than the outlet flow from the reactor. It means that the dissolved API is a priority to generate more growth. Besides, the optimal solution also indicates that recycling the purified mother liquor is not advised to maximize the mean crystal size

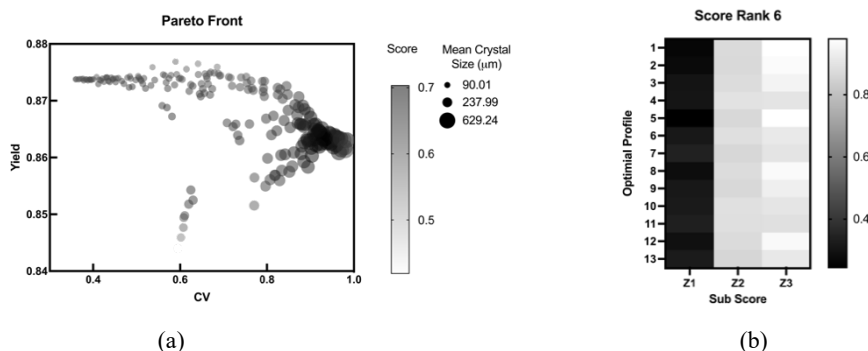


Figure 2 Ranking of the Pareto solutions based on MAUT (a) and sub-scores (b)

when the optimization is based on the combined reaction and crystallization. Because of the conflicting effect between a better reaction conversion and better crystallization yield, the recycle stream may dilute the crystallization system more significantly far beyond what was observed in scenario 1, when the concentration of API is higher than the nominal/initial concentration in the outlet flowrate (fresh feed) from the reactor.

#### 4. Conclusion

A superstructure optimization approach was proposed to design an integrated reaction-crystallization-filtration process with recycles. All possible options for the recycle streams, crystal seeding, and fresh feed additions were considered and allocated binary and real decision variables. Other critical decision variables were also considered which include the design and operating options such as cooling temperatures and crystallizer capacity (residence time). Firstly, single objective optimization problems were solved separately to (1) maximize mean crystal size, (2) minimize the CV, (3) maximize yield. This was followed by a combination of three objectives to address a mixed integer multi-objective optimization problem. The resulting Pareto solutions were analyzed and ranked using MAUT. The exploration of the combination of the reaction and crystallization also indicate that recycling the purified mother liquor to the second and third stage in crystallization are not recommended. And with the recycles to the reactor, the mean crystal size was maximized to 307  $\mu\text{m}$  with less water added in the second reactor. The proposed systematic approach shows that wastes can be significantly reduced, and resources consumption may be rationalized without compromising on product quality.

#### Acknowledgements

The authors acknowledge funding from the UK Engineering and Physical Sciences Research Council (EPSRC), for Made Smarter Innovation – Digital Medicines Manufacturing Research Centre (DM2), EP/V062077/1.

#### References

- Benyahia, B. (2018). Applications of a plant-wide dynamic model of an integrated continuous pharmaceutical plant: Design of the recycle in the case of multiple impurities. Computer Aided Chemical Engineering, Elsevier B.V. **41**: 141-157.
- Benyahia, B., M. A. Latifi, C. Fonteix and F. Pla (2011). "Multicriteria dynamic optimization of an emulsion copolymerization reactor." Computers & Chemical Engineering **35**(12): 2886-2895.
- Espinoza-Gomez, H., E. Saucedo-Castillo, L. Z. Flores-López, E. Rogel-Hernandez, M. Martínez and F. T. Wakida (2018). "Ethanol:water blends separation using ultrafiltration membranes of poly(acrylamide-co-acrylic acid) partial sodium salt and polyacrylamide." The Canadian Journal of Chemical Engineering **96**(3): 763-769.
- Joiner, D. E., J. Billeter, M. E. P. McNally, R. M. Hoffman and P. J. Gemperline (2014). "Comprehensive kinetic model for the dissolution, reaction, and crystallization processes involved in the synthesis of aspirin." Journal of Chemometrics **28**(5): 420-428.
- Liu, J. and B. Benyahia (2022). "Optimal start-up strategies of a combined cooling and antisolvent multistage continuous crystallization process." Computers and Chemical Engineering **159**.
- Mencarelli, L., Q. Chen, A. Pagot and I. E. Grossmann (2020). "A review on superstructure optimization approaches in process system engineering." Computers and Chemical Engineering **136**.

# Prediction of product distribution using machine learning techniques

Praveen Kumar Bommineni,<sup>a</sup> Manohar Kakunuri,<sup>a</sup> Sarat Babu Anne,<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, National Institute of Technology Warangal, Warangal – 506004, India*

## Abstract

The product distribution of multi-product processes involving series-parallel reactions depends on several variables. Hence, traditional experimental and computational approaches are laborious, time consuming and the results may not be satisfactory. Data driven approaches serve as an alternative to the traditional approaches in predicting and optimizing the product distribution. Machine learning (ML) provides reliable and accurate predictions. In this work, application of ML techniques to predict product distribution of one such complex process is demonstrated. Chlorination of methane has been chosen as case study for this purpose. Data required for the ML models is generated by simulating the process flow sheet in aspenONE. Of the several ML models used, k-nearest neighbors (kNN) method provides most accurate product distribution with high R-square value and low mean absolute error (MAE).

**Keywords:** Series-parallel reactions, product distribution, aspenONE, Machine Learning.

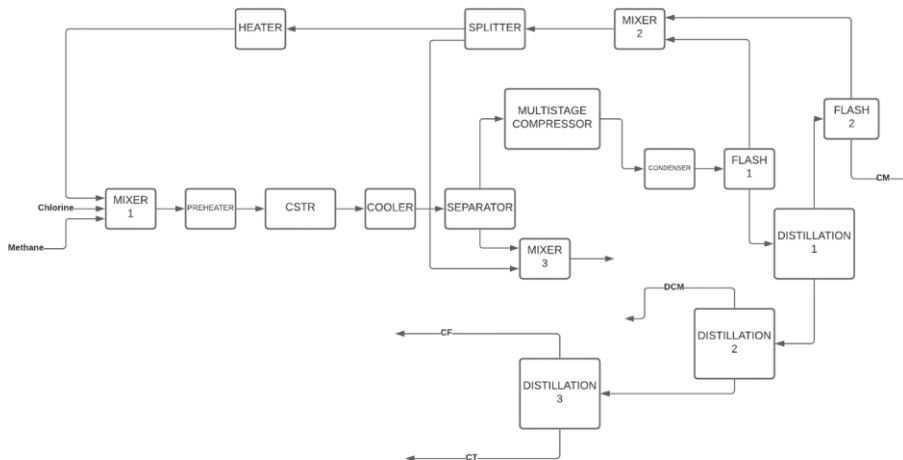
## 1. Introduction

Reactions and processes are typically desired to operate at optimum conditions of process variables such as temperature, pressure, flow rates, compositions etc. (Tabor 2018, Gupta 2021). Therefore, the chief goal is to find the optimal conditions so as to maximize performance in terms of yields, purity etc. However, it is not a trivial task to find optimum conditions even at lab scale as many variables/parameters effect the performance and it requires iterative costly procedures. Data-driven approaches serve as an alternative and reliable way to achieve the same (Shields 2021, Marcus 2021, Hein 2021, Joung 2021).

Machine learning is proving to be useful tool for applications in various fields including materials discovery from failed experiments, optimizing reaction conditions, reaction classification, synthesis of small molecules (Park 2022, Raccuglia 2016, Zunyun 2020, Schwaller 2021, Probst 2022, Segler 2018). In this article we demonstrate prediction of product distribution using ML techniques. First we present the process simulation followed by application of machine learning techniques to predict product flow rates and compositions.

## 2. Methane chlorination process simulation

### 2.1. Steady state process simulation

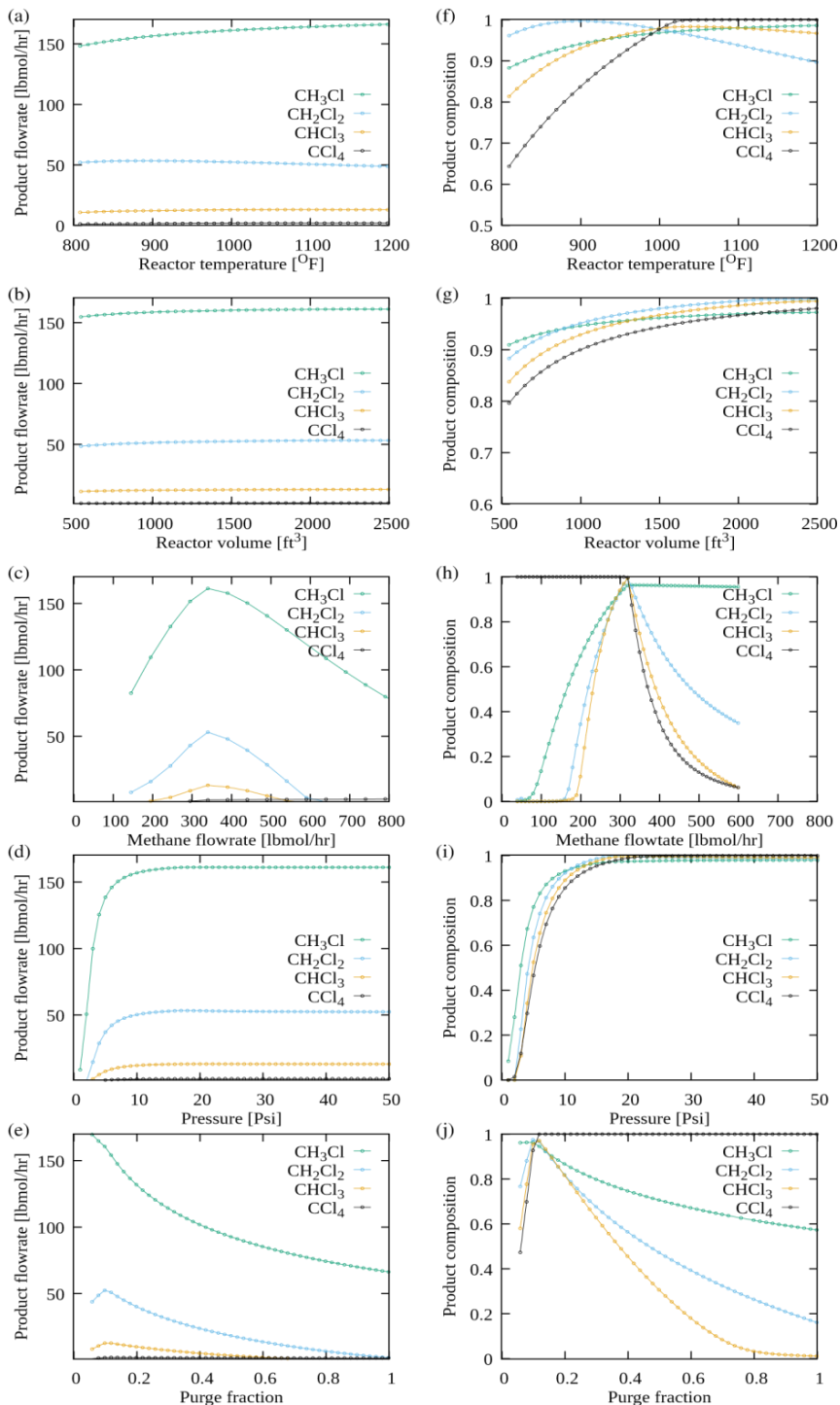


**Figure 1.** Methane chlorination process block diagram

Methane chloride can be produced by hydrochlorination of methanol (Thyagarajan 1966, Habata 1976) and chlorination of methane (Dantus 1996, Dantus 1999) processes. In this work, chlorination of methane which is one of the widely studied systems since 1930's (Robert 1931, Dantus 1996, Dantus 1999, Rozanov 2010) has been simulated. For industrial application, thermal chlorination is generally employed which is highly exothermic in nature. Methane undergoes following chlorination reactions to produce chloromethane ( $\text{CH}_3\text{Cl}$ ), dichloromethane ( $\text{CH}_2\text{Cl}_2$ ), trichloromethane ( $\text{CHCl}_3$ ) and tetrachloromethane ( $\text{CCl}_4$ ).



Of these four reactions, the main reaction that takes place is methane and chlorine reacting to form methyl chloride. Based on the operating conditions, the product distribution varies by a wide margin. The data required for ML techniques is generated by simulating the process flow sheet in aspenONE (Dantus 1996, Dantus 1999). Methane chlorination process block diagram is shown in Figure 1.



**Figure 2.** Effect of all five variables on (a-e) Product flow rates, (f-j) Product compositions.



## 2.2 Sensitivity analysis

Once the basic flow sheet is simulated, sensitivity analysis has been carried out in order to identify the parameters on which product flow rates and compositions are highly dependent. From the sensitivity studies, it is found out that the 5 parameters, i.e., reactor temperature, reactor volume, methane feed flow rate, reactor pressure, and purge fraction are the key parameters. Varying the above key parameters within the effective range for each of these variables, simulations are carried out to obtain product flow rates and compositions. The parameter ranges for which data generated is listed in Table 1. All the four product flow rates and compositions as a function of the five variables are shown in Figure 2(a-e) & (f-j) respectively.

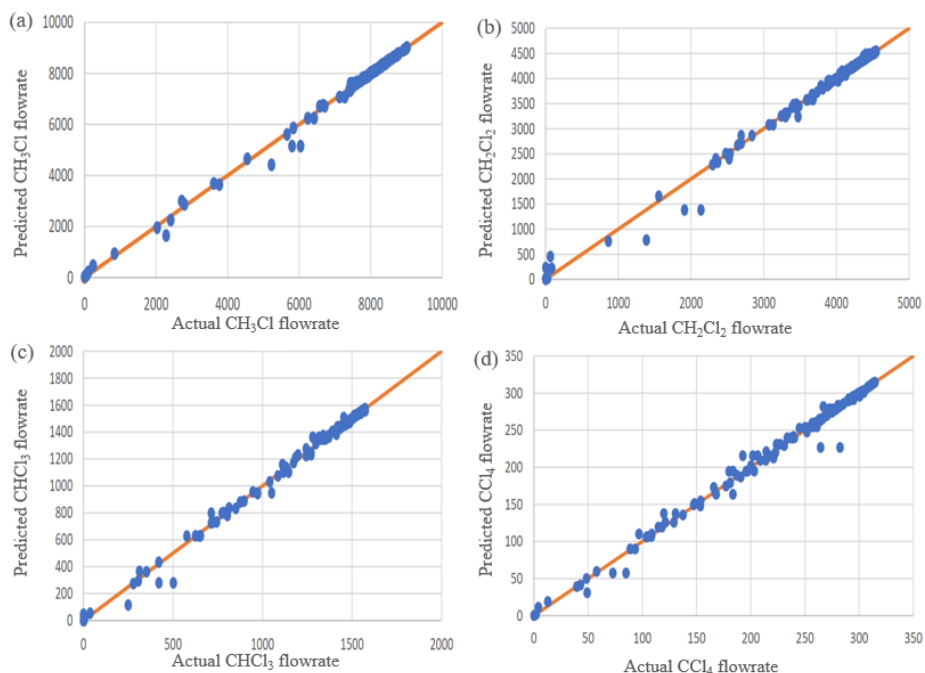
**Table 1.** Effective parameter ranges (in fps units) in which sensitivity analysis is carried out.

Variable	Range
1. Reactor temperature	800 – 1200 °F
2. Reactor volume	500 – 2500 ft <sup>3</sup>
3. Methane feed flow rate	30 – 600 lbmol/hr
4. Reactor pressure	0.1 – 25 psia
5. Purge fraction	0.04 – 0.8

## 3. Machine learning models

The data generated using sensitivity analysis is used to train five different ML models. These models include multi-output linear regression (MLR), support vector regression (SVR), decision trees, kNN and artificial neural networks (ANN). Data set from sensitivity analysis generated consisting of 2500 data points for each variable is divided in the ratio of 80:20 for training and testing ML models respectively. These models train by mapping the function relating input variables to the output variables in the given training data set. The trained model is then evaluated by predicting the output for the inputs given in the test set and comparing the predicted output to actual output. The predicted flow rates and compositions from five ML models are compared against actual values. However, for one ML model i.e., kNN, product flow rates obtained from simulations and values predicted by the model are shown in Figure 3 for four products. Two metrics i.e., R-square and mean absolute error (MAE) are used to measure the goodness of fit and accuracy of the models. These values are given in Table 2. The results indicate that the kNN model is found to be more accurate compared to the other models with high R-square value of 0.995 and low MAE of 23.38 for the process chosen to study. The next best model to kNN is decision trees. MLR has a low R-square and SVR has high MAE compared to other

models. However, it is observed that increasing number of data points to train SVR and ANN results in low MAE.



**Figure 3.** Parity charts of flowrates obtained from simulation and kNN model for all four products (a)  $\text{CH}_3\text{Cl}$  (b)  $\text{CH}_2\text{Cl}_2$  (c)  $\text{CHCl}_3$  and (d)  $\text{CCl}_4$

**Table2:** Comparison of R-square and MAE for different ML models.

ML Model	R-Square	MAE
1. Multi output linear regression	0.57	396.58
2. SVR	-	2777.61
3. Decision trees	0.98	39.27
4. kNN	0.995	23.38
5. ANN	-	345.40

#### 4. Conclusions

In conclusion, we have shown the application of ML techniques to predict the product distribution for multi-product series-parallel reactions. Data generated from simulations is used to train and test different models. R-square and MAE are used as model performance metrics. For the methyl chloride process considered in this study, kNN model performs better than other techniques considered. Better performance of the kNN model may be attributed to the large data set used to train the model. Our work demonstrates that ML models can be successfully used to predict the product distribution our multi product series-parallel reactions with good accuracy and without the involvement of cumbersome experiments or simulations.

## References

- 1 Tabor, D.P. et al (2018), Accelerating the discovery of materials for clean energy in the era of smart automation. *Nat Rev Mater* 3, 5–20.
- 2 Jason E. Hein, (2021), Machine learning made easy for optimizing chemical reactions, *Nature* 590, 40-41.
- 3 Markus Meuwly, 2021, Machine Learning for Chemical Reactions. *Chemical Reviews* 121 (16) , 10218-10239
- 4 Udit Gupta, Dionisios G. Vlachos, 2021, Learning Chemistry of Complex Reaction Systems via a Python First-Principles Reaction Rule Stencil (pReSt) Generator. *Journal of Chemical Information and Modeling* 61 (7) , 3431-3441.
- 5 Joonyoung F. Joung, Minhi Han, Jinhyo Hwang, Minseok Jeong, Dong Hoon Choi, Sungnam Park., 2021, Deep Learning Optical Spectroscopy Based on Experimental Database: Potential Applications to Molecular Design. *JACS Au* , 1 (4) , 427-438.
- 6 Andrzej M. Żurański, Jesus I. Martinez Alvarado, Benjamin J. Shields, Abigail G. Doyle, 2021, Predicting Reaction Yields via Supervised Learning. *Accounts of Chemical Research* , 54 (8) , 1856-1865.
- 7 Hanyu Gao, Jean Pauphilet, Thomas J. Struble, Connor W. Coley, Klavs F. Jensen, 2021, Direct Optimization across Computer-Generated Reaction Networks Balances Materials Use and Feasibility of Synthesis Plans for Molecule Libraries. *Journal of Chemical Information and Modeling* 2021, 61 (1) , 493-504.
- 8 Shields B. J. et al., (2021), Bayesian reaction optimization as a tool for chemical synthesis, *Nature* 590, 89–96.
- 9 S. Park, H. Han, H. Kim, S. Choi, (2022) Machine learning applications for chemical reactions, *Chem. Asian J.* , 17, e202200203.
- 10 Raccuglia, P., Elbert, K., Adler, P. et al. (2016) Machine-learning-assisted materials discovery using failed experiments. *Nature* 533, 73–76.
- 11 Zunyun et al., (2020), Optimizing chemical reaction conditions using deep learning: a case study for the Suzuki-Miyaura Cross coupling reaction, *Org. Chem. Front.* 7, 2269-2277.
- 12 Schwaller, P., Vaucher, A. C., Laino, T., & Reymond, J. L. (2021). Prediction of chemical reaction yields using deep learning. *Machine learning: science and technology*, 2(1), 015016.
- 13 Daniel Probst, Jean-Louis Reymond et al. (2022) Reaction classification and yield prediction using the differential reaction fingerprint DRFP, *Digital Discovery*, 1, 91.
- 14 Segler, M., Preuss, M. & Waller, M. (2018), Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* 555, 604–610.
- 15 Thyagarajan, M. S., Kumar, R., and Kuloor, N. R. (1966), Hydro chlorination of Methanol to Methyl Chloride in Fixed Catalyst Beds, *I & E C Process Design and Development* 5(3): 209-213.
- 16 Habata, K., Tanaka, S., Araki, H., (1976), Shin-Etsu Chemical Co. Ltd. "Process for Preparing Methyl Chloride", USA, US 3,983,180A.
- 17 Mauricio M. Dantus and Karen A. High., (1996), An Economic Tool for the Retrofit of Chemical Processes Through Waste Minimization and Process Integration, *Industrial and Engineering Chemistry* 35(12): 4566-4578.
- 18 Mauricio M. Dantus and Karen A. High, (1999), Evaluation of waste minimization alternatives under uncertainty: a multiobjective optimization approach. *Computers & Chemical Engineering*, 23(10), 1493-1508.

# Architectural Design of Chemical and Biological Pathways through Parameter Sensitivity Oriented Mixed Integer Formulations

Emir Topac,<sup>a</sup> Emrullah Erturk,<sup>a</sup> Ozgun Deliismail,<sup>b</sup> Sahin Sarrafi,<sup>b</sup> Hasan Sildir,<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, Gebze Technical University, Kocaeli, 41400, Türkiye*

<sup>b</sup>*SOCAR Turkey R&D and Innovation Co., Izmir, 35800, Türkiye*

*hasansildir@gtu.edu.tr*

## Abstract

A mixed integer nonlinear programming (MINLP) formulation is developed for the automatized design of chemical and biological pathways. The proposed formulation is solved using rigorous algorithms and enables the selection of reactions whose rate constants have the highest sensitivity. Unlike the traditional parameter estimation approach, the training performance is introduced as constraint which is pre-defined to balance the architectural simplicity and fitting performance. Adjoint sensitivity expressions are additionally solved simultaneously in the formulation and the resulting parameter sensitivities are included in the objective. Binary variables are introduced in the linking constraints for the selection of a reaction subset from a more complex pathway, also tightening the search space for the optimization algorithm. The approach is implemented on a complex chemical/biological pathway which processes glucose as a primary feedstock to produce hydroxymethylfurfural (HMF). The reduced pathway prediction bounds are narrower ensuring a similar mean predictions, despite 25% reduction in the superstructure.

**Keywords:** Parameter identifiability; Mixed integer programming; Sensitivity; Uncertainty.

## 1. Introduction

Chemical and biological pathway models include a high number of rate expressions with tunable parameters to be estimated from the experimental data. The lack of spatial and temporal measurements might result in statistically ill-defined inverse problems with multiple solutions in parameter estimation (McClean and McAuley, 2012). Furthermore, the computational load and identifiability problems increase once the disturbances and measurement noises appear especially when the model architecture is large.

Usually, trial and error procedures or sequential methods are applied for the architectural synthesis and parameter identifiability tasks, which require significant manual effort in addition to decentralized computations on theoretical level. The most sensitive parameter selection based on fisher information matrix is compared to LASSO, which is least absolute shrinkage and selection operator, and applied to a biomechanical system in (Ramadan et al., 2018). Shapiro et al. developed a heuristic algorithm to eliminate some

parameters and re-calculate the sensitivities until desired condition value is obtained (Shapiro et al., 2014).

A simultaneous approach using rigorous formulations to obtain the pathway architecture and corresponding parameters based on the parameter sensitivity values and their contribution to uncertainty is scarce and would have the potential to exploit available statistical knowledge from experimental data for a better model development. This study focuses on the development of a mixed-integer formulation to perform architectural synthesis and parameter estimation tasks based on sensitivities in addition to the model training performance.

## 2. Methodology

A typical reaction pathway is represented by ODEs and given by:

$$\frac{dy}{dt} = f(t, y, p) \quad (1)$$

where  $y$  is the vector of outputs which are typically concentrations;  $p$  is the vector of time-invariant parameters;  $t$  is the time;  $f$  is the vector of mathematical rate expressions.

The derivative of the terms in Eq. 1 results in forward sensitivity ODE for a particular parameter and output:

$$\frac{ds_{ij}}{dt} = \frac{df_j}{dy_j} s_{ij} + \frac{df_j}{dp_i} \quad (2)$$

where  $s_{ij}$  is the sensitivity of  $j^{\text{th}}$  output,  $y_j$ , to  $i^{\text{th}}$  parameter,  $p_i$ . A simultaneous solution of Eq. 1 and Eq. 2 is required to address the specific considerations on parameter selection process to account for the model performance, additionally.

A modified mixed integer formulation is developed for the architectural design and estimation of its parameters who have the highest sensitivity, eliminating the parameters with relatively smaller impact on the outputs. The formulation after the ODEs are discretized based on orthogonal collocation enables the simultaneous solution based on rigorous optimization algorithms and is given by:

$$\begin{aligned} & \text{Min}_{p,b} \quad - \sum s_{ij,k}^2 \\ & \text{s.t.} \\ & \frac{dy}{dt} = f(t, y, p) \\ & \frac{ds_{ij}}{dt} = \frac{df_j}{dy_j} s_{ij} + \frac{df_j}{dp_i} \quad i = 1, 2, \dots, M \quad j = 1, 2, \dots, P \\ & s_{ij}(0) = 0 \quad i = 1, 2, \dots, M \quad j = 1, 2, \dots, P \\ & y_j(0) = (0.25, 0, 0, 0, 0, 0) \quad j = 1, 2, \dots, 6 \\ & \sum \|y_k - y_{m,k}\| \leq \varepsilon \\ & b_i p_{i,\min} \leq p_i \leq b_i p_{i,\max} \quad i = 1, 2, \dots, P \\ & b_{\min} \leq \sum b_i \leq b_{\max} \quad b_i \in \{0, 1\} \quad i = 1, 2, \dots, P \end{aligned} \quad (3)$$

where  $s_{ij,k}$  is the sensitivity at sample time  $k$ ;  $y_k$  is the model prediction at  $k$ ;  $y_{m,k}$  is the measurement at  $k$ ;  $\varepsilon$  is the desired maximum training error;  $p_{i,min}$  and  $p_{i,max}$  are lower and upper bounds of  $i^{th}$  parameter,  $p_i$ , respectively;  $b_i$  is the binary variable to account for the existence of  $p_i$ ;  $b_{min}$  and  $b_{max}$  are lower and upper bounds of number of parameters desired in the ultimate architecture.

Solution of Eq. 3 requires an explicit formulation for sensitivity ODEs for the rigorous approach, which calls for an automatic or symbolic differentiation. There are various solver packages which deliver such expressions (Andersson et al., 2019), to avoid numerical approximation errors. On the other hand, sensitivity equations are mostly nonlinear and further mathematical reformulations and decompositions, in addition to dense discretization, might be essential for pathways with high complexity. In contrast to nonlinear optimization which applies to parameter estimation with fixed architectures, the proposed mixed integer formulation includes binary variables to account for the existence of particular parameter values which represent the reaction rate constant once they have a nonzero value. Linking constraints are also introduced to tighten the search space by fixing the parameter values while eliminating the corresponding reaction path from the network. The formulation is flexible and can address measurements and sensitivity considerations at desired sample times. Current formulation includes the selection of the parameters which deliver the maximum squared sensitivity at measurement points.

The parameter covariance matrix for a particular architecture ( $c_p$ ) is subsequently calculated using bootstrapping method (Godo-Pla et al., 2019) for further analysis on the identifiability issues. The uncertainty in the outputs due to parameter variations is calculated using (Tellinghuisen, 2001):

$$c_y = S_k c_p S_k^T \quad (4)$$

where  $c_y$  is an under estimator of actual output covariance matrix;  $S_k$  is the sensitivity matrix at  $k$ .

The approach is implemented on a complex chemical/biological pathway and solved using BONMIN solver in PYOMO computation environment (Hart et al., 2011).

### 3. Results

The proposed approach is implemented on 5-Hydroxymethylfurfural (HMF) synthesis from glucose (Tang et al., 2017) which can be obtained through biological processes, providing a promising opportunity for sustainability considerations. The full network superstructure is shown in Fig. 1a, where several different reaction paths enable the production of a particular compound, which in turn contributes to identifiability of these paths. With a coupled and dependent set of ODEs, the unique estimation of the parameters becomes a challenging task and the proposed formulation is implemented to demonstrate the impact, unlike (Erturk et al., 2021) where sensitivity considerations are not included. Thus, in addition to six differential equations to represent the compounds, 72 additional sensitivity ODEs are introduced to account for 12 parameters in the full network. Once the simultaneous architectural synthesis and the parameter estimation are performed based on the selection of the most sensitive parameters, the resulting reduced network is presented in Fig. 1b.

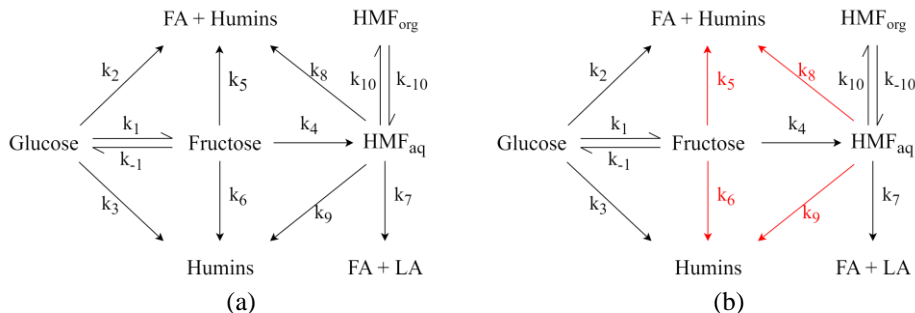


Figure 1. Reaction pathway  
(a) Full network (b) Reduced network

The eliminated reaction paths from Fig. 1a are shown in Fig. 1b using red arrows. Next, the performance of the full network and the reduced network is compared in terms of prediction similarity and uncertainty. The uncertainty for the solutions is evaluated from the diagonal elements of  $c_y$ , which is calculated from Eq. 4 after bootstrapping procedures are performed.

Fig. 2a and Fig. 2b include solutions of pathway ODEs for the full network and the reduced network, respectively, for Fructose and Formic Acid (FA) + Humins. With relatively smaller mean values, Levulinic Acid (LA) is excluded from figures. The mean predictions, which exclude the parameter covariance matrix and related propagation equations, show similar profile due to constraint in Eq. 3 to deliver a satisfactory training performance characterized by  $\varepsilon$  and shown in full lines. The uncertainty ranges are represented by dashed lines. In turn, a similar profile is obtained despite elimination of four reaction paths as shown in Fig. 1b.

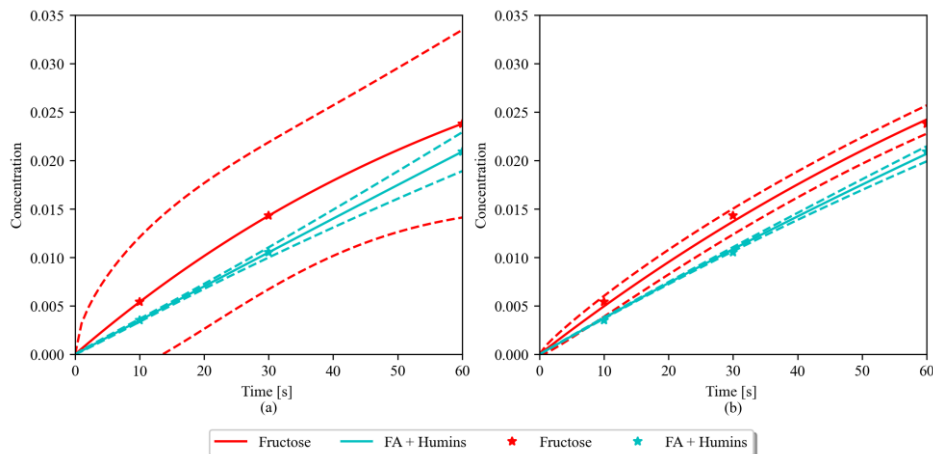


Figure 2. Mean solution and uncertainty interval for Fructose, FA+Humins  
(a) Full network (b) Reduced network

As a result of numerical scale differences among pathway variables, the solution profiles for  $\text{HMF}_{aq}$  and  $\text{HMF}_{org}$ , are presented in Fig. 3b, from which the latter is for the reduced network. Compared to the other components, the impact on  $\text{HMF}_{org}$  uncertainty is less significant as the corresponding component related reactions have not changed (see Fig. 1a and Fig. 1b). However, some decrease in the uncertainty range is still observed due to a more robust parameter estimation which accounts for the identifiability.

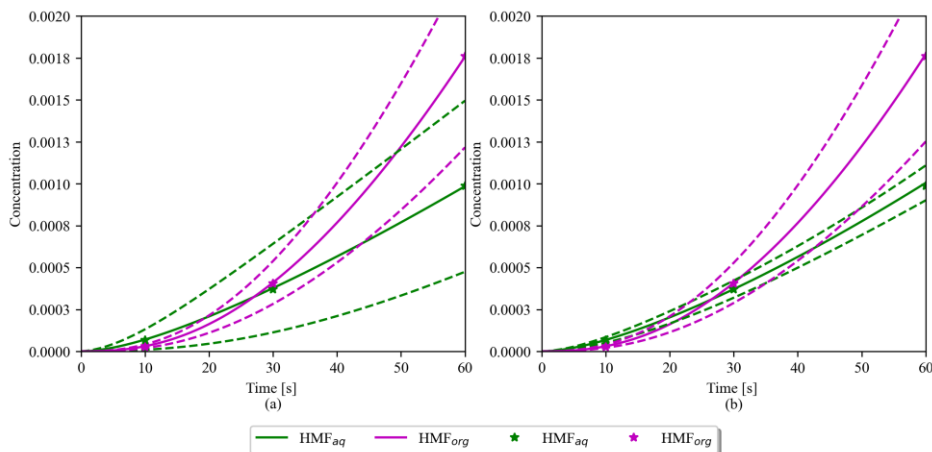


Figure 3. Mean solution and uncertainty interval for  $HMF_{aq}$ ,  $HMF_{org}$   
 (a) Full network (b) Reduced network

Table 1 includes a comparison on the maximum and minimum dependent variable values based on uncertainty at final time. The prediction bounds for the reduced model is significantly more tightened compared to the full network. However, a more drastic reduction in uncertainty is favored by the elimination of the reaction paths associated with a particular component.

Table 1. Impact on output uncertainty

	Full Network [ $\times 100$ ]		Reduced Network [ $\times 100$ ]		Difference [%]
	Min	Max	Min	Max	
Glucose	19.15	21.16	19.96	20.32	82.09
Fructose	1.41	3.35	2.28	2.56	85.57
$HMF_{aq}$	0.05	0.15	0.09	0.11	80.58
$HMF_{org}$	0.12	0.23	0.13	0.23	5.45
FA+Humins	1.89	2.29	1.99	2.14	62.50
FA+LA	0.00	0.10	0.00	0.09	10.00

#### 4. Discussion

Traditional parameter estimation problems for complex pathways usually process a fixed architecture and include the training error in the objective function under physically constrained rate expression parameters. In contrast to traditional approach, the proposed mixed-integer formulation performs simultaneous architectural design and parameter estimation by adding binary variables which consider the existence of a particular path and tighten the search space during the optimization through linking constraints. The model training performance is introduced as a constraint to balance the structural reduction and fitting performance, since elimination of some reaction paths from the superstructure hinders the fitting capability, in theory. Moreover, the simultaneous evaluation and consideration of sensitivity expressions enable the calculations to be performed by rigorous optimization algorithms, which can include advanced mathematical reformulations, decompositions, and approximations to account for various



sensitivity requirements. Such a centralized processing of the complex problem delivers the best configuration and high identifiable parameters, especially when global or sophisticated solvers are utilized.

As a downside, the approach calls for more computational load compared to traditional methods, due to binary variables to represent the existence of parameters, evaluation of higher number of equations during optimization, additional nonlinearity terms, the rigorous solution algorithm delivers a more robust model architecture. Thus, subsequent tasks related to model update or control tasks become computationally more feasible.

A biochemical pathway is considered to demonstrate the impact of the approach. However, the problem formulation is flexible and can be further tailored to various applications. Current focus includes the development of sophisticated mathematical reformulations to address the computational problems when larger networks with more parameters are needed.

## References

- Andersson, J.A.E., Gillis, J., Horn, G., Rawlings, J.B., Diehl, M., 2019. CasADi: a software framework for nonlinear optimization and optimal control. *Math. Program. Comput.* 11, 1–36.
- Erturk, E., Erdal, A., Sildir, H., 2021. Tam Sayili ve Surekli Optimizasyon Problemi ile Reaksiyon Ag Modellerinin Kucultulmesi. *Konya Muhendis. Bilim. Derg.* 9, 142–156.
- Godo-Pla, L., Emiliano, P., Valero, F., Poch, M., Sin, G., Monclús, H., 2019. Predicting the oxidant demand in full-scale drinking water treatment using an artificial neural network: Uncertainty and sensitivity analysis. *Process Saf. Environ. Prot.* 125, 317–327.
- Hart, W.E., Watson, J.P., Woodruff, D.L., 2011. Pyomo: Modeling and solving mathematical programs in Python. *Math. Program. Comput.* 3, 219–260. <https://doi.org/10.1007/s12532-011-0026-8>
- Mclean, K.A.P., Mcauley, K.B., 2012. Mathematical modelling of chemical processes-obtaining the best model predictions and parameter estimates using identifiability and estimability procedures. *Can. J. Chem. Eng.* 90, 351–366. <https://doi.org/10.1002/cjce.20660>
- Ramadan, A., Boss, C., Choi, J., Peter Reeves, N., Cholewicki, J., Popovich, J.M., Radcliffe, C.J., 2018. Selecting sensitive parameter subsets in dynamical models with application to biomechanical system identification. *J. Biomech. Eng.* 140.
- Shapiro, B., Barton, M., Mavalvala, N., Mittleman, R., Youcef-Toumi, K., 2014. Selection of important parameters using uncertainty and sensitivity analysis. *IEEE/ASME Trans. Mechatronics* 20, 13–23.
- Tang, J., Zhu, L., Fu, X., Dai, J., Guo, X., Hu, C., 2017. Insights into the kinetics and reaction network of aluminum chloride-catalyzed conversion of glucose in NaCl–H<sub>2</sub>O/THF biphasic system. *ACS Catal.* 7, 256–266.
- Tellinghuisen, J., 2001. Statistical error propagation. *J. Phys. Chem. A* 105, 3917–3921. <https://doi.org/10.1021/jp003484u>

# Data-driven fouling modeling in crude oil heat exchanger

Pelin Dologlu<sup>a</sup>, Alihan Tabak<sup>a</sup>, A. Eren Vedin<sup>a</sup>, Gizem Kayar<sup>b</sup>

<sup>a</sup>*SOCAR Turkey, Digital Transformation Department, Istanbul 34485, Turkey*

<sup>b</sup>*SOCAR STAR Oil Refinery, Process Department, Aliaga, Izmir 35800, Turkey*

## Abstract

Unwanted materials in fluids deposit on heat transfer surfaces in time, which is called fouling, and it negatively affects heat exchangers regarding operation and maintenance costs. Therefore, fouling rate should be monitored to prevent unexpected costs and possible failures due to deposition limit of heat transfer area. Monitoring also allows process engineers to improve operating conditions to keep fouling under control for longer cleaning cycles and schedule cleanings effectively. In this work, to monitor fouling rate in the shell and tube crude oil preheat exchanger in the real refinery, the data-driven heat transfer rate prediction model is developed from processed and to-be-processed crude oil planning, on-line sensor and laboratory sample data. Results of various machine learning and deep learning techniques are compared, and as Multiple Linear Regression is found to be the best performing predictive model. Developed fouling rate model predictions are actively used for scheduling of crude oil preheat exchanger cleaning and improving operating conditions to reduce costs in the refinery.

**Keywords:** crude oil, pre-heat train exchangers, fouling, machine learning, multiple linear regression

## 1. Introduction

Petrochemical industries, especially oil refinery, include one of the most energy-required processes among other industries. Maintaining sustainability and reducing gas emissions while keeping the process profitable is a significant problem for refinery processes. This challenging problem can be solved with energy saving solutions (Mrayed et al., 2021). Since the crude distillation unit (CDU) is the front-end operation of the refinery, any intervention in CDU can reduce the energy consumption of the entire refinery by 24% (Yang et al., 2020). Crude oil preheat train (PHT) exchangers recover energy, which is necessary for CDU, by up to 70%. However, recovered energy via PHT exchangers decreases over time because of the deposition of unwanted materials on the heat transfer surface area, which is called fouling (Coletti et al., 2011). Therefore, the fouling rate should be monitored to prevent unexpected energy and maintenance costs, gas emissions and possible failures due to the deposition limit of the heat transfer area (Santamaria & Macchietto, 2019). Monitoring also allows process engineers to improve operating conditions to keep fouling under control for longer cleaning cycles and schedule cleanings effectively. Thus, monitoring of fouling in crude oil PHT exchangers has been studied by many researchers.

Fouling modeling and as a further step fouling prediction enable monitoring fouling in heat exchangers effectively. In literature, fouling is modeled in the first place as empirical and theoretical (Wang et al., 2019). Since, crude oil fouling contains many simultaneous

and different mechanisms, empirical or theoretical models are not generalized and not enough to reflect the nonlinear nature of fouling properly (Sundar et al., 2020). Therefore, data-driven approach for fouling in heat exchanger is introduced to the literature in recent years. Neural network based methods for fouling monitoring and prediction of heat exchanger are preferred methods thanks to their natural capacity to exhibit nonlinear relationships easily (Davoudi & Vaferi, 2018). Also, Sundar et al. (Sundar et al., 2020) proposed a robust deep learning fouling prediction approach which learns heat exchanger flow and heat transfer physics. Furthermore, Hosseini et al. (Hosseini et al., 2022) estimated the fouling factor in a comprehensive study using five different algorithms namely Gaussian Process Regression, Decision Trees, Bagged Trees, Support Vector Regression and Linear Regression. Data-driven approaches are appropriate methods in Industry 4.0 where lots of data is generated and they have proved their success in heat exchangers to determine fouling in complex processes.

In this work, to monitor the fouling rate in the shell and tube crude oil preheat exchanger, the data-driven overall heat transfer rate prediction model is developed. Daily overall heat transfer rate predictions for the next 15 days are predicted from some of the processed and to-be-processed crude oil planning data, on-line sensor and laboratory sample data to calculate heat transfer rate based on the theoretical model. To the best of authors' knowledge, there are limited PHT exchanger fouling prediction models for long time horizon, that consider all processed and to-be-processed crude oil planning data, on-line sensor and laboratory sample data. In order to choose the best prediction model in real crude oil PHT exchangers, a comparison of different machine learning and deep learning techniques, which are XGBoost, Multiple Linear Regression, Random Forest, Extra Trees, Gradient Boosting, Decision Tree, LightGBM and Multilayer Perceptron are studied. The aim of this study is efficient scheduling of crude oil PHT exchanger cleaning and improving operating conditions by reducing energy loss to reduce costs in the refinery, while considering safety by preventing unexpected failures in normal operation. Moreover, proposed data-driven modeling approach can be generalized for other heat exchangers just by revising crude oil properties related inputs.

## 2. Materials and Methods

### 2.1. Crude Oil Preheat Train

PHT exchangers have a simple design, but they have an important role in terms of the economical aspect of refinery operations. Before crude oil is heated in the furnaces, crude oil passed through these PHT exchangers, which increases energy efficiency. In this work, the multi-pass shell and tube eight identical parallel heat exchangers in PHT, undergoing fouling deposition in both shell and tube side are modeled (see Figure 1). While crude oil flows from the shell-side, the overhead fluid of the atmospheric distillation column flows from the tube-side in this heat exchanger network. Crude oil is heated with overhead fluid while overhead fluid is cooled. If there is fouling in the heat exchanger, temperature and pressure differences are highly affected and energy efficiency decreases over time. To keep the process efficient and safe, the fouling level in the heat exchanger network in refinery should be conducted properly.

### 2.2. Fouling Model

Fouling can be defined as the deposition of unwanted materials on the heat transfer area, and it negatively affects heat transfer efficiency in operations. Cross section of one of heat exchanger from network and detailed fouling scheme of shell and tube heat exchanger can be seen in Figure 1. Fouling formation can be happen both inside and outside of tubes. Foulants in crude oil cause fouling inside tubes and crude oil flows at

low temperature ( $T_c$ ) which increases formation of fouling. Hot overhead fluid flows at shell side and  $T_h$  temperature, which causes fouling outside of tubes. Since one of the fouling indicators is the heat transfer rate, which can be calculated easily via a theoretical model, with the help of on-line sensors and laboratory data also known real PHT exchanger-specific coefficients. Therefore, in this work, the overall heat transfer rate is modeled to monitor fouling in the PHT exchanger network of real-time crude oil distillation unit operations. Energy efficiency of heat exchangers will be increased, and unexpected failures can be prevented by the help of this model. Also, maintenance of heat exchangers can be scheduled efficiently. Furthermore, data-driven modeling approaches can be generalized for other heat exchangers by revising crude oil related inputs.

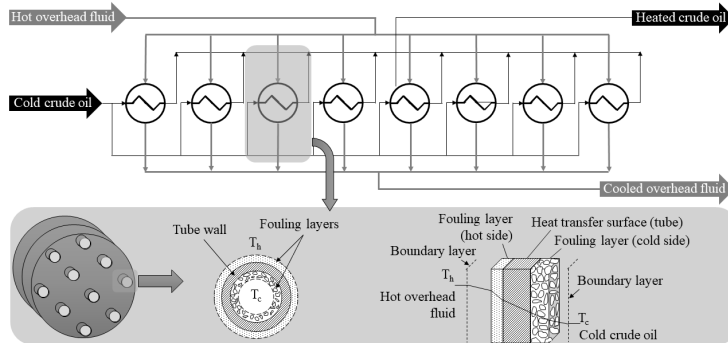


Figure 1. Crude oil PHT exchanger network and detailed fouling scheme of heat exchanger

### 2.3. Industrial Data Set Preparation For Modeling

30 different features related to PHT exchanger network and crude oil assays were used in this study. 16 of them belong to the online sensor data and the rest are laboratory sample data. In addition to online sensor and laboratory sensor data, unique crude oil properties are added to the dataset which is collected from crude oil assays. The planning department decides which crude oil will be used in the future and shares this information with the process units. In this study, to-be processed crude oil blend properties (viscosity, asphaltenes, naphthenes, total acid number, and API) are calculated as weighted averages of unique crude oil properties from crude oil assays according to the planning department's decision. Since different crude oil blend properties can aggregate effects on the process, the cumulative sums of these properties are also added to the dataset. Whole online sensor data, laboratory sample data and crude oil blend properties features have been fetched on a daily basis, and they are aggregated as one dataset. In addition, some new features to calculate the overall heat transfer rate were created by inputs that are already in the dataset. For instance, the shell-side delta temperature was calculated by the subtraction between the shell inlet temperature and the shell outlet temperature. Similarly, tube-side delta temperature was calculated by the subtraction between tube inlet temperature and tube outlet temperature. The delta pressure data for both shell and tube sides were calculated in this way, and they are added to the dataset because of possible indicators of fouling. Since the model predicts the overall heat transfer rate 15 days in advance, 15 days lagged data and the data which represents 1, 3, 7 and 15 days rolling means to see the trends of all inputs in the dataset, except laboratory sample data. Consequently, the final dataset has 947 sample points from real-time crude oil operations and 97 features with one target, which is the overall heat transfer rate.

## 2.4. Machine Learning and Deep Learning Algorithms

### 2.4.1. Multiple Linear Regression

Multiple Linear Regression (MLR) tries to find the linear relationship between a dependent variable (target) and one or more independent variables (predictors). The relationship between the target and predictors is represented as an equation that has an intercept and a coefficient for each predictor. The goal is to find the coefficients that minimize the difference between the predicted and actual values of the target. The obtained model can be used to predict the target value for a new set of predictor values. Linear Regression is a simple but powerful tool for modeling linear relationships and is widely used in various applications (Hosseini et al., 2022). The coefficients are estimated from the training data using normal equation as optimization algorithm in this work.

### 2.4.2. Decision Tree

Decision Tree is a non-parametric supervised learning algorithm. Decision Tree has the capability of predicting the target variable by learning simple decision rules inferred from the data features (Hosseini et al., 2022).

### 2.4.3. Ensemble Methods

Ensemble methods contain set of models, which are integrated to obtain final prediction. Many randomized decision trees on various sub-samples of dataset are used in Random Forest and Extra Trees to improve prediction accuracy. Also, techniques such as Gradient Boosting, XGBoost and LightGBM, that fit boosted decision trees by minimizing an error gradient (Hosseini et al., 2022), are used in this work.

### 2.4.4. Multilayer Perceptron

Multilayer Perceptron (MLP) is widely accepted technique as a promising tool for solving complex engineering problems in real industrial processes. The MLP can include any number of the layer with specified number of neurons in each layer, which transfers information between first and last layers (Davoudi & Vaferi, 2018). Since 1 hidden layer and 30 neurons give the best result, these parameters are used for MLP in this work.

## 3. Results and Discussion

Correlation of features in the final dataset with the target -overall heat transfer rate- is observed, and the most correlated 8 features have been selected as inputs for modeling. Selected inputs are 15 days lagged data of shell-side delta temperature, 15 days lagged data of overall heat transfer rate, 15 days lagged data of viscosity, 1 day rolling mean of viscosity, 3 days rolling mean of viscosity, 7 days rolling mean of viscosity, 15 days rolling mean of viscosity and cumulative sum of asphaltene.

Different machine learning and deep learning techniques are compared their performances to find the best fouling prediction model in real-time operations. While MLR, Decision Tree, Random Forest, Extra Trees, Gradient Boosting, XGBoost and LightGBM were used as machine learning techniques, MLP was used as a deep learning technique. All algorithms are implemented in Python 3.8.

In this study, three different test periods were selected, and each test period contains 15 sample points. The reason why preferred 3 test periods is the business unit wanted to see the results of these periods whether models can catch the expected fouling results. Also, Mean Absolute Error (MAE) was preferred as the model evaluation metric for this work. After the results were tested in 3 different test periods, the best fouling prediction model for real-time operations is the MLR model. Figure 2 shows the comparison of the different model techniques' performances in terms of MAE for a test period.

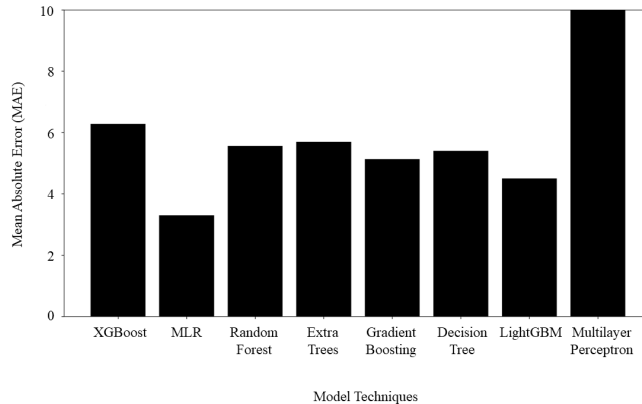


Figure 2. Different machine learning and deep learning techniques comparison

As it can be easily seen from Figure 2, the MLR model has the best performance with 3.29 MAE score. Multilayer Perceptron as a deep learning technique gives the highest MAE score. The reason why Multilayer Perceptron technique fails is deep learning techniques perform well on more sample point dataset than this study dataset. Besides, boosting techniques also failed because of the scarcity of data. MLR technique, which has the best results, is used for real-time PHT fouling modeling within 15 days horizon. One of the test set results for comparison of MLR fouling model predictions and actual overall heat transfer rate real-time operation of PHT exchanger values within 15 days horizon are shown in Figure 3. The results were deemed successful by the business unit. With the help of the fouling prediction model, the business unit can monitor fouling before reaching the critical limit, and can intervene in the process parameters by taking an action to slow down the fouling effect. Also, the business unit can schedule crude oil PHT exchanger cleaning efficiently and can improve operating conditions to reduce costs in the refinery. Furthermore, this modeling approaches can be generalized for other heat exchangers by revising crude oil related inputs.

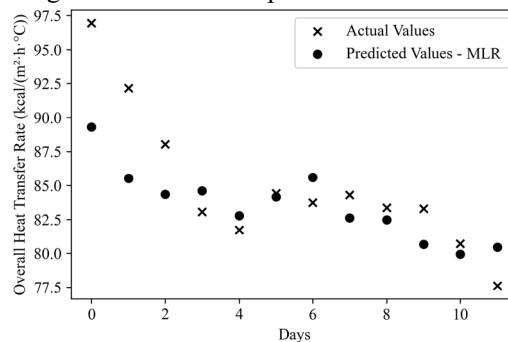


Figure 3. Fouling model predictions and actual values comparison of the test set

#### 4. Conclusions

Oil refineries have the most energy-required processes, thus any energy saving solutions has a great impact on sustainability, gas emissions and operating costs. One of the energy saving solution is that energy efficiency increases while crude oil passed through PHT exchangers before crude oil is heated in the furnaces in CDU. However, fouling causes the deposition of a layer from unwanted materials in these heat exchangers, which

decreases the heat transfer rate. If the fouling rate in PHT exchangers can be known before reaching the critical limit, the business unit can take action that slows down the fouling effect. Therefore, data-driven fouling rate prediction in the shell and tube crude oil PHT exchangers model is developed to real-time monitor fouling rate for a long time horizon. Since one of the fouling indicators is the heat transfer rate, which can be calculated easily via a theoretical model, the overall heat transfer rate is modeled to monitor fouling. Daily overall heat transfer rate predictions for the next 15 days are predicted from some of the processed and to-be-processed crude oil planning data, on-line sensor and laboratory sample data. Combining of all processed and to-be-processed crude oil planning data, on-line sensor and laboratory sample data are increased prediction accuracy and provides predicting capability to fouling modeling for long time horizon. In order to choose the best prediction model in real crude oil PHT exchangers, a comparison of different machine learning and deep learning techniques, which are XGBoost, Multiple Linear Regression, Random Forest, Extra Trees, Gradient Boosting, Decision Tree, LightGBM and Multilayer Perceptron are studied. Proposed data-driven modeling approach can be implemented on other heat exchangers easily just by revising crude oil related parameters. With the help of the fouling prediction model, the business unit can monitor fouling before reaching the critical limit, and can intervene in the process parameters by taking an action to slow down the fouling effect. Also, the business unit can schedule crude oil PHT exchanger cleaning efficiently and can improve operating conditions to reduce costs in the refinery.

## References

- Coletti, F., Macchietto, S., & Polley, G. T. (2011). Effects of fouling on performance of retrofitted heat exchanger networks: A thermo-hydraulic based analysis. *Computers & Chemical Engineering*, 35(5), 907–917. <https://doi.org/10.1016/J.COMPCHEMENG.2011.01.027>
- Davoudi, E., & Vaferi, B. (2018). Applying artificial neural networks for systematic estimation of degree of fouling in heat exchangers. *Chemical Engineering Research and Design*, 130, 138–153. <https://doi.org/10.1016/J.CHERD.2017.12.017>
- Hosseini, S., Khandakar, A., Chowdhury, M. E. H., Ayari, M. A., Rahman, T., Chowdhury, M. H., & Vaferi, B. (2022). Novel and robust machine learning approach for estimating the fouling factor in heat exchangers. *Energy Reports*, 8, 8767–8776. <https://doi.org/10.1016/J.EGYR.2022.06.123>
- Mrayed, S., Shams, M. Bin, Al-Khayyat, M., & Alnoaimi, N. (2021). Application of pinch analysis to improve the heat integration efficiency in a crude distillation unit. *Cleaner Engineering and Technology*, 4, 100168. <https://doi.org/10.1016/j.clet.2021.100168>
- Santamaria, F. L., & Macchietto, S. (2019). Integration of optimal cleaning scheduling and control of heat exchanger networks under fouling: MPCC solution. *Computers & Chemical Engineering*, 126, 128–146. <https://doi.org/10.1016/J.COMPCHEMENG.2019.04.012>
- Sundar, S., Rajagopal, M. C., Zhao, H., Kuntumalla, G., Meng, Y., Chang, H. C., Shao, C., Ferreira, P., Miljkovic, N., Sinha, S., & Salapaka, S. (2020). Fouling modeling and prediction approach for heat exchangers using deep learning. *International Journal of Heat and Mass Transfer*, 159, 120112. <https://doi.org/10.1016/J.IJHEATMASSTRANSFER.2020.120112>
- Wang, F. L., Tang, S. Z., He, Y. L., Kulacki, F. A., & Yu, Y. (2019). Heat transfer and fouling performance of finned tube heat exchangers: Experimentation via on line monitoring. *Fuel*, 236, 949–959. <https://doi.org/10.1016/J.FUEL.2018.09.081>
- Yang, K., Liu, S., He, C., Zhang, B., Chen, Q., & Pan, M. (2020). Improving energy saving of crude oil distillation units with optimal operations. *Journal of Cleaner Production*, 263, 121340. <https://doi.org/10.1016/J.JCLEPRO.2020.121340>

# Explainable formation energy prediction for uncovering the relationship between the electronic structure and stability of the heterogeneous catalyst

Daeun Shin,<sup>a, b</sup> Dong Hyeon Mok,<sup>c</sup> Seoin Back,<sup>c</sup> Jonggeol Na<sup>a, b</sup>

<sup>a</sup>*Department of Chemical Engineering and Materials Science, Ewha Womans University, Seoul 03760, Republic of Korea*

<sup>b</sup>*Graduate Program in System Health Science and Engineering, Ewha Womans University, Seoul 03760, Republic of Korea*

<sup>c</sup>*Department of Chemical and Biomolecular Engineering, Institute of Emergent Materials, Sogang University, Seoul 04107, Republic of Korea*  
[jgna@ewha.ac.kr](mailto:jgna@ewha.ac.kr)

## Abstract

In recent studies, machine learning (ML) applications in the heterogeneous catalysis field for material properties prediction accelerate the catalyst discovery with desired properties. However, due to its high complexity, most ML models suffer from the black-box problem, which cannot provide a basis for prediction. Thus, reliable application and physical insight generation are challenging with conventional black-box models. Here, we developed an ML model that predicts formation energy ( $E_f$ ) from the density of states (DOS). More importantly, by interpreting the model, we also confirmed the possibility of uncovering the relationship between the electronic structure of materials and their stability. Our model achieves successful performance demonstrating its superior capability of DOS featurization.

**Keywords:** Heterogeneous Catalyst, Stability, Artificial Intelligence, Convolutional Neural Network, Explainable AI (XAI)

## 1. Introduction

In the heterogeneous catalysis field, the discovery of new materials with desired properties is challenging because of its large parameter space (Esterhuizen J.A., 2022). In this context, computational material discovery has grown continuously over the past decade thanks to the development of various crystal structure prediction algorithms. Previously, density functional theory (DFT) calculation was a major tool to search chemical space, however, due to its high computational cost, the demand for an efficient tool has been increasing. Recently, machine learning (ML) models, which generate predictions within sub-seconds, have emerged as a powerful approach to bypassing the formidable computational cost and navigating the large chemical space more efficiently (Fung V., 2021). However, to maximize the prediction performance, ML models have become more complex with tremendous trainable parameters, which makes it nearly impossible to interpret the internal logic underlying the prediction (Doosun Hong Jaehoon 2022). Beyond mere prediction tools, in order for an ML model to play a conducive role in scientific advancement in the heterogeneous catalysis field, the model should be interpretable and able to provide meaningful knowledge and testable hypotheses. As an alternative, theory-infused models integrating physical interaction into the algorithm and thus, inherently interpretable were studied (Wang S.H., 2021). However, due to their



restricted architecture and enforced simplicity, it is difficult to achieve high performance. Furthermore, since most of these models are system-specific, there is a limitation in extension to general application for various system. In order to satisfy both predictive performance and interpretability simultaneously, the implementation of the explainable AI (XAI) technique on the fully trained predictive model is in the spotlight. These studies are mainly conducted by utilizing descriptors related to material properties and obtaining the importance of each descriptor in prediction, for ease of interpretation and visualization (OmidvarNoushin 2021). Especially, since it is well-known that the electronic structure is closely related to physical properties of materials, there have been many efforts to utilize electronic features, such as *d*-band center, extracted from the density of state (DOS) (LinicHongliang 2012). Yet, the problem remains that it does not extend well to a wide range of data with only a few existing features identified by human experts. In addition, only constrictive interpretation is possible, which describes intricate physical interaction within the previously defined features. In this perspective, an approach for direct interpretation from the DOS signal itself, without human intervention in feature extraction, is required.

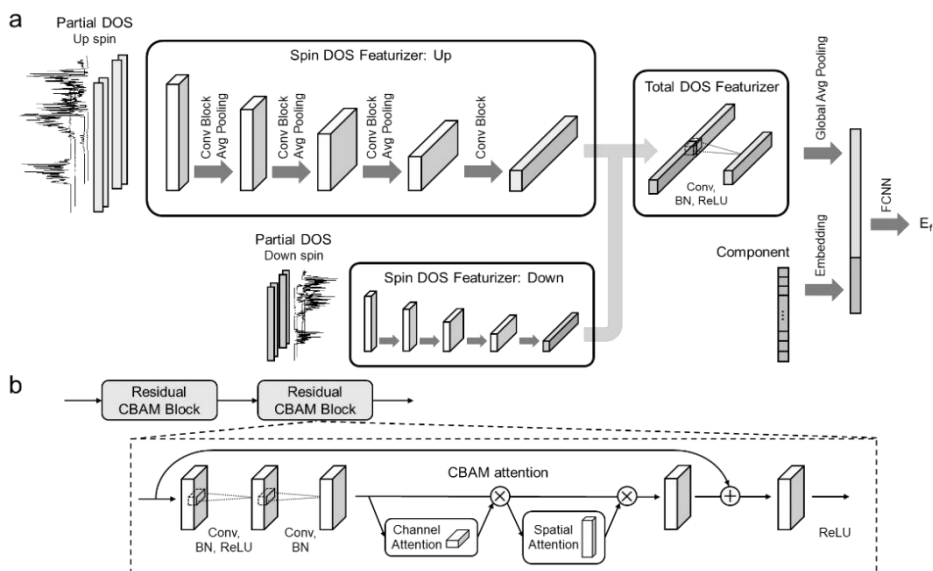
In this work, we developed a convolutional neural network (CNN) based ML model taking DOS directly as input for accurate formation energy prediction. Further, the possibility of translating the relationship between the electronic structure of materials and the crystal's thermodynamic stability via linking the XAI technique to the model is presented. Thanks to the convolution operation, which has strength in recognizing patterns of signal data, electronic features of DOS are self-explored inside the model with high flexibility. Particularly, our model, whose architecture is motivated by chemical background, is specialized to featurize DOS. After training, the SHapley Additive exPlanation (SHAP) (LundbergS.M., 2020) value of features discovered by the model was calculated to determine what factors on each input affected the prediction in which way. By visualizing how SHAP values show the pattern toward formation energy, common characteristics of stable materials can be confirmed.

## 2. Results

### 2.1. ML approach for accurate formation energy prediction

We developed a ML model with a chemically motivated architecture optimized to featurize partially projected DOS for accurate formation energy prediction. The model is built based on a convolutional neural network (CNN), which is well-established for learning high-level features from signal data with multiple channels (KiranyazSerkan 2019). Given that CNN takes input made up of separate channels and featurizes input preserving its spatial information, which indicates spin state according to energy here, it is advantageous to manage DOS input consisting of various projected states. Additionally, the convolution operation gathers input features  $h_k^{n-1}$  multiplying weight matrices  $w_{kj}^n$ , then aggregates them by sum to generate featurized output features represented as  $h_j^n$  (KimYoon 2014). Since each partial DOS represents a relative contribution of a particular state at total DOS, total DOS features can be obtained via aggregating features from partial DOS, which resembles convolution operation. Based on the similarity, the CNN-based featurizer generates abundant features of the total DOS from the input partial DOS.

Meanwhile, the sign of DOS is determined depending on its spin state, that is, positive and negative for up spin and down spin, respectively. Thus, total DOS is generally considered according to its spin state (YeoByung 2019). Inspired by the background, we introduced the CNN-based featurizer for each spin state, named spin



**Figure 1.** (a) Schematic representation of model architecture. Partial DOS in each spin state and component vector are fed into the model as different inputs. Three featurizers are used to generate total DOS feature maps. (b) Convolutional block architecture employing CBAM self-attention and residual connection. Each convolutional block is connected through average pooling.

DOS featurizer. As shown in Figure 1a, there are two spin DOS featurizers for each spin state respectively. Each featurizer takes orbital projected DOS as input according to the spin state and extracts features based on the convolution operation mentioned above. Note that, the featurizers are independent of each other, using separate weight matrices. The spin DOS featurizer is constructed with four convolutional blocks with two residual convolutional block attention module (CBAM) blocks (WooSanghyun 2018) for automatic and adaptive feature learning. Each block is designed with the aim of producing well-learned feature maps which abundantly contain meaningful features, and the detailed structure is represented in Figure 1b. The CBAM self-attention module performs dynamic feature selection that learns the importance of the features differently for each input, rather than imposing the same weight on all extracted features as general CNN does. In addition, the residual connection prevents the loss of these weighted features due to the depth of the network (DuXue 2020). After further learning through the subsequent convolutional blocks, total DOS feature maps of each spin state are obtained at the end of the featurizers.

The maps are concatenated in channel dimension and served as input to the following total DOS featurizer, which finally aggregates the integrated feature maps to form comprehensive feature maps covering all the spin and orbital states. The final feature maps generated through all the CNN-based featurizers described above are flattened and concatenated with component information followed by the embedding layer. Finally, it is taken as the input of fully connected neural network (FCNN) to predict formation energy. The key aspects of our model are optimal operation for aggregation of partially projected DOS and chemically inspired architecture considering DOS based on its spin state. This enables the model to maximize the number of high-quality features for formation energy prediction.

The model trained on DFT-calculated DOS from Materials Project (MP) database (JainAnubhav 2013), which includes 66,411 materials and their DOS data in total. Note that the training data consisted of only *s*, *p*, and *d*-block elements, but it is possible to include *f*-block elements in training/test data by integrating extra channels in the model. We generated input representations, which consist of up and down spin states of 9 orbitals from *s* to *d*. Each one comprises an individual input channel (total 18 channels), where StandardScaler of scikit-learn package was applied to remove the mean and to scale to unit variance. After preprocessing 18 channels, we added one extra channel representing a composition of materials, where one-hot encoding of all the elements in our dataset, 110, was used. Tensorflow 2.4.1 and Keras 2.4.0 are used as the backend for the model implementation. The model was trained for 300 epochs with an early stopping algorithm on an NVIDIA GPU with memory and CUDA 11.0. Additionally, we highlight that our model could be universally used for formation energy prediction tasks as it was trained with data consisting of diverse compositions of elements across more than 80.

## 2.2. Prediction Performance

**Table 1.** Prediction performance with 2e-ORR test set for different algorithms.

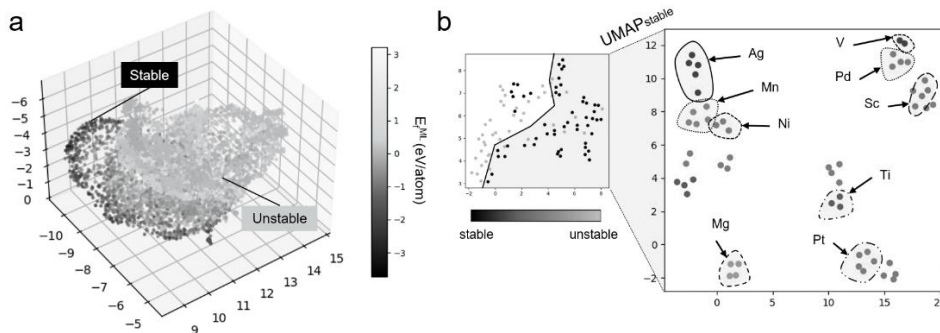
Model	DOS signal-based		DOS feature-based*				
	Proposed	CNN	KRR	SVR	RF	XGBoost	DNN
MAE	<b>0.236</b>	0.273	0.271	0.280	0.356	0.307	0.274
RMSE	<b>0.372</b>	0.397	0.435	0.446	0.562	0.498	0.420

\**d*-band center, Width, Skewness, Kurtosis, Number of fillings, Compositions

We then evaluated the prediction performance of our model and compare it with other ML models based on different input features and algorithms. All models were tested on the same dataset of electrochemical two-electron O<sub>2</sub> reduction reaction (2e-ORR) which contains DOS of 7165 binary alloys, and the overall performances are listed in Table 1. The DOS feature-based models (i.e. kernel ridge regressor (KRR) (SchlkopfBernhard 2018), support vector regressor (SVR) (AwadMarianne 2015), random forest (RF) (BreimanLeo 2001), XGBoost (ChenTianqi 2016), deep neural network (DNN) with two fully connected layers (CourvilleJan 2016)), where 0–4th moment properties of 18 orbitals were used as inputs, underperform compared to the DOS signal-based models (i.e. proposed, CNN (SunDechun 2019)). This result reiterates that only a few existing features are insufficient to properly explain solid materials in the whole dataset. As such, since the excessive simplicity and low flexibility in fitting, these models are difficult to predict as accurately as models that turn all the information on DOS to account. Furthermore, the proposed model here outperforms conventional CNN, highlighting the effectiveness of chemically inspired architecture and CBAM managing input DOS effectively. As a result, our model achieved the highest performance in formation energy prediction with 0.237 eV/atoms of MAE, demonstrating its superior featurization ability.

## 2.3. XAI Interpretation

To unpack the black-box model and confirm the physical interaction between DOS and the formation energy, SHAP analysis was implemented. SHAP is a XAI technique introducing a game theoretic approach to explain the output of the ML model. It decomposes the predicted value into an additive sum of contributions of individual features on input. Due to its guaranteed mathematical background, SHAP values are one of the most-used approaches to understanding ML models. Prior to further analysis, we projected SHAP values obtained from 2e-ORR data into a three-dimensional space using uniform manifold approximation and projection (UMAP) (McInnesLeland 2018) to



**Figure 2.** Visualization of SHAP value distribution for different datasets, generated by UMAP. a) Three-dimensional UMAP for 2eORR dataset. Darker colors indicate low  $E_f^{ML}$ . (b) Two-dimensional UMAP for the single-specie dataset. (left) According to the  $E_f^{ML}$ , less than 0.01 is categorized as stable and others as unstable. (right) Two-dimensional UMAP plotted again with only the data belonging to the stable region.

visualize whether the learned features show a tendency depending on material stability (Figure 2a). The unsupervised analysis displays the continuous change in feature importance on DOS in accordance with the formation energy. Particularly, the darker area, where the data with low formation energy is located close, demonstrates that stable materials share common features on DOS.

Inspired by this result, it is analyzed more specifically what common characteristics appear in SHAP values for materials that are predicted as stable. For ease of analysis, we preliminarily obtained DOS and their SHAP values from various crystal structures, i.e. simple cubic (SC), body-centered cubic (BCC), face-centered cubic (FCC), hexagonal close-packed (HCP), and rhombohedral crystal structure (RHO), for materials consisting of only a single specie, and projected the SHAP values into low dimension. Similar context to the previous result, it is confirmed that the materials predicted as stable and the materials predicted as unstable are distributed in separate regions on UMAP (Figure 2b left). Then, UMAP was performed again targeting only the data that belong to the darker area where 'stable' data are largely located (Figure 2b right). Surprisingly, the data are clustered by the specie, showing that data within the same element share similar features even if having different crystal structures. The similarities are not observed in the data with unstable crystal structures. These results demonstrate the possibility of XAI interpretation that deduces the common properties of stable materials and further uncovers the causal relationship between the electronic structure and the formation energy.

### 3. Conclusion and Future Works

Here we have developed a machine learning model which successfully predicts the formation energy from DOS with superior featurization ability. More importantly, we presented the potentiality of direct XAI interpretation of the DOS signal itself without any prior involvement of human experts in feature restriction. The result shown here is simply that features on DOS have commonality according to their formation energy. Thus, this observation will be developed in future works to reveal the 'which' factor of DOS is decisive in formation energy prediction, and finally, it will shed light on the causal relationship between the electronic structure of materials and the crystal's thermodynamic stability. Ultimately, it will contribute beyond the stability screening by the black-box model to knowledge and hypothesis generation in the field. (Esterhuizen J.A., 2022)

## References

- Awad, Mariette and Khanna, Rahul},. Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers. Apress, 2015.
- Breiman, Leo. "Random Forests." *Machine Learning*, 2001. 5-32.
- Chen, Tianqi and Guestrin, Carlos. "XGBoost: A Scalable Tree Boosting System." *Association for Computing Machinery*, 2016. 785-794.
- Courville, Ian Goodfellow and Yoshua Bengio and Aaron. *Deep Learning*. MIT Press, 2016.
- Doosun Hong, Jaehoon Oh, Kihoon Bang, Soonho Kwon, Se-Young Yun, Hyuck No Lee. "Interpretable Deep learning Model for Analyzing the Relationship between the Electronic Structure and Chemisorption Property." *J. Phys. Chem. Lett.*, 2022. 8628-8634.
- Du, Xue and Liao, Kuo and Shen, Xiaofeng. "Secondary Radar Signal Processing Based on Deep Residual Separable Neural Network." *2020 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS)*, 2020. 12-16.
- Esterhuizen, J.A., Goldsmith, B.R., Linic, S. "Interpretable machine learning for knowledge generation in heterogeneous catalysis." *Nat Catal* 5, 2022. 175-184.
- Fung, V., Hu, G., Ganesh, P. "Machine learned features from density of states for accurate adsorption energy prediction." *Nat Commun* 12, 2021. 88.
- Jain, Anubhav and Ong, Shyue Ping and Hautier, Geoffroy and Chen, Wei and Richards, William Davidson and Dacek, Stephen and Cholia, Shreyas and Gunter, Dan and Skinner, David and Ceder, Gerbrand and others. "Commentary: The Materials Project: A materials genome approach to accelerating materials innovation." *American Institute of Physics AIP*, 2013. 011002.
- Kim, Yoon. "Convolutional Neural Networks for Sentence Classification." *Association for Computational Linguistics*, 2014.
- Kiranyaz, Serkan and Ince, Turker and Abdeljaber, Osama and Avci, Onur and Gabbouj, Moncef. "1-D Convolutional Neural Networks for Signal Processing Applications." *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019. 8360-8364.
- Linic, Hongliang Xin and Adam Holewinski and Neil Schweitzer and Eranda Nikolla and Suljo. "Electronic structure engineering in heterogeneous catalysis: Identifying novel alloy catalysts based on rapid screening for materials with desired electronic properties." *Topics in Catalysis*, 2012. 376-390.
- Lundberg, S.M., Erion, G., Chen, H. "From local explanations to global understanding with explainable AI for trees." *Nat Mach Intell* 2, 2020. 56-67.
- McInnes, Leland and Healy, John and Melville, James. "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction." *arXiv*, 2018.
- Omidvar, Noushin and Pillai, Hemanth S. and Wang, Shih-Han and Mou, Tianyou and Wang, Siwen and Athawale, Andy and Achenie, Luke E. K. and Xin, Hongliang. "Interpretable Machine Learning of Chemical Bonding at Solid Surfaces." *J. Phys. Chem. Lett.*, 2021. 11476-11487.
- Scholkopf, Bernhard and Smola, Alexander J. and Bach, Francis. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. The MIT Press, 2018.
- Sun, Dechun and Chen, Yang and Liu, Jiaao and Li, Yu and Ma, Rui. "Digital Signal Modulation Recognition Algorithm Based on VGGNet Model." *2019 IEEE 5th International Conference on Computer and Communications (ICCC)*, 2019. 1575-1579.
- Wang, SH., Pillai, H.S., Wang, S. "Infusing theory into deep learning for interpretable reactivity prediction." *Nat Commun* 12, 2021. 5288.
- Woo, Sanghyun and Park, Jongchan and Lee, Joon-Young and Kweon, In So. "CBAM: Convolutional Block Attention Module." *Springer International Publishing*, 2018. 3-19.
- Yeo, Byung Chul and Kim, Donghun and Kim, Chansoo and Han, Sang Soo. "Pattern Learning Electronic Density of States." *Scientific Reports*, 2019. 5879.

# Modular stochastic optimization for optimal rainwater harvesting system design

Qiao Yan Soh,<sup>a</sup> Edward O'Dwyer,<sup>a</sup> Salvador Acha,<sup>a</sup> Nilay Shah<sup>z</sup>

*<sup>a</sup>Sargent Center for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, London SW7 2AZ, UK  
qiaoyan.soh13@imperial.ac.uk*

## Abstract

Rainwater Harvesting (RWH) systems can serve a dual functionality as a flood mitigation structure as well as providing local water availability. Optimisation-based design strategies must be transferrable enough to incorporate the influence of the local climate and case-specific catchment area characteristics into the design process, which can be a significant endeavour when required for every individual implementation. To increase the accessibility of optimisation methods in the appropriate sizing of RWH systems, this paper presents a modularised optimisation model, where tank components and dynamics are contained as individual blocks. These blocks can then be pieced together to produce a full system model, allowing optimisation models to be easily built for any combination and design of RWH system. This is implemented with a multi-tank RWH system, where an evaluation of the optimised system configuration showed a good balance between the dual objectives of providing improved flood mitigation and local water reuse, in comparison to an existing system derived through alternative sizing strategies.

**Keywords:** Multi-scale integration, Rainwater harvesting, Stochastic optimization

## 1. Introduction

Large-scale rainwater harvesting (RWH) systems have demonstrated potential as a strategy for simultaneously reducing flood risks by temporarily holding excess water, and minimising water stress by satisfying non-potable water demands with locally harvested water. As the performance of these systems has been demonstrated to be very sensitive to the local climate and catchment area (Semaan, 2020), appropriately sizing these structures has remained a key barrier to further implementations of RWH systems.

Optimisation models have been found to be effective and successful in deriving suitably sized RWH systems (Okoye, 2015; Nnaji, 2017), but with the range of possible system configurations, optimisation-based strategies are not the most scalable of sizing methods since the creation of optimisation models for each individual implementation of a RWH system can require significant time, effort, and expertise. Hence, it is common to find that design guidelines published by public water managers utilise simplified functions, for example of the expected average demand and catchment area to calculate an appropriate tank size. These are normally found to prioritise the flood mitigation objectives, especially in highly developed urban centres as these are likely to have a stable freshwater supply network to satisfy any local water demands. However, this would result in overly sized systems which can reduce the return on investment for the RWH system and incur further costs in the form of utilising the limited space resources available within a city (Matos, 2013).

Hence, to improve the accessibility of and encourage more widespread use of optimisation-model-based strategies for sizing urban RWH systems, this paper proposes

a modularised optimisation model tool. This provides basic optimisation model blocks for a single tank, with a range of possible tank draining configurations which can be further customised through functional constraint blocks. This declarative format would allow individuals with expertise outside of implementing optimisation methods to utilise them as an effective system sizing strategy.

## 2. Methods

The modularised optimisation model is presented in the following sections and applied to a multi-tank RWH system. The tool is used to derive an optimal system size for a multi-tank RWH system, and the output system configuration is evaluated for its ability to achieve the dual objectives of flood mitigation and local water reuse. The computational requirements are measured as a function of the model solution time, as this could be a key factor for increased uptakes of optimisation-model based strategies.

### 2.1. Modular optimization model structure

The proposed modularised optimisation framework consists of function blocks which correspond to the dynamics of a given tank component and build up to a fully stochastic Mixed Integer Linear Program (MILP) model that looks to select an optimal tank height, area, and any applicable orifice heights for the range of rainfall scenarios given to the model.

The component function blocks correspond not only to typical tank dynamics such as the mass balance and flow equations, but also provide design specifications and behaviours including:

- Penalty functions for exceeding threshold overflows or wasted capacities
- Common optimisation objectives including system size reduction, and overflow minimisations.
- Freshwater use for satisfying unmet demand

To provide a turnkey functionality to this framework, further tank configuration blocks were defined which correspond specifically to the different possible combinations of ways a given tank may be drained. These blocks can be initialised immediately and solved when supplied with a desired input parameter set but can also be further supplemented with component function blocks to implement further tank behaviours and or penalty functions for multi-objective systems. A basic implementation workflow for a single tank would hence involve:

- (1) Create a base optimisation model through a tank configuration block
- (2) Add any additional system dynamics and constraints
- (3) Define an objective function, including any penalties
- (4) Populate the model with rainfall, demand, and design parameter options
- (5) Solve the derived model.

Under this framework, multi-tank systems can be implemented and optimised sequentially through a series of much smaller MILPs, which typically would be much more tractable than a single, large optimisation problem.

### 2.2. Case study configuration

The framework was applied to a multi-tank RWH system aimed at simultaneously reducing local flooding risks and providing non-potable water to its service area. This is formed of four specialized tanks, as illustrated in Figure 1. The 'separation' tank serves as an initial receiver for all inflow and acts as a filter that directs flow between the 'harvesting' and 'detention' tanks. These tanks both store water, but separately for the

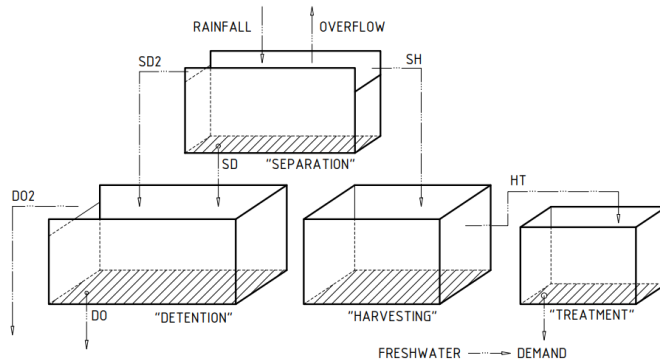


Figure 1: Multi-tank RWH system configuration and flow interactions

purposes of local reuse and temporarily holding excess to prevent downstream inundation respectively. Finally, a ‘treatment’ tank is used to hold treated, ready-to-use water. Under the modularised framework, each individual tank is optimised individually with non-linear flow divisions within the system managed using an external flow division function. The flow of information between the individual tank optimisation models, as well as their corresponding model outputs, are highlighted in Figure 2. The input data comes from a historical timeseries, which has been segregated into 24-hour rolling window samples. Scenarios are then selected randomly from the set of samples, with the set diversity ensured using a minimum threshold value for a variance measure, defined as the sum of the variance of the rainfall values in each timestep across the entire simulation horizon. When the minimum threshold is not reached, the set is reselected.

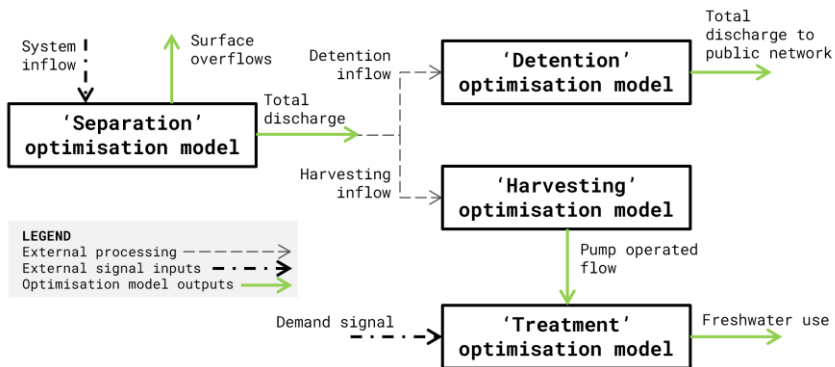


Figure 2: Information flow and interactions between optimisation models in a multi-tank RWH system

Table 1: Implemented objective and penalty functions for each tank optimisation model.

Tank	Objective	Penalty function
Separation	Minimise overflows	Wasted capacity
Detention	Minimise tank capacity	Overflow threshold
Harvesting	Maximise outflow to treatment tank	Overflow threshold
Treatment	Minimise freshwater top-up volumes	Overflow threshold



### 3. Results

The performance of a modularised optimisation model is dependent on the following model parameters:

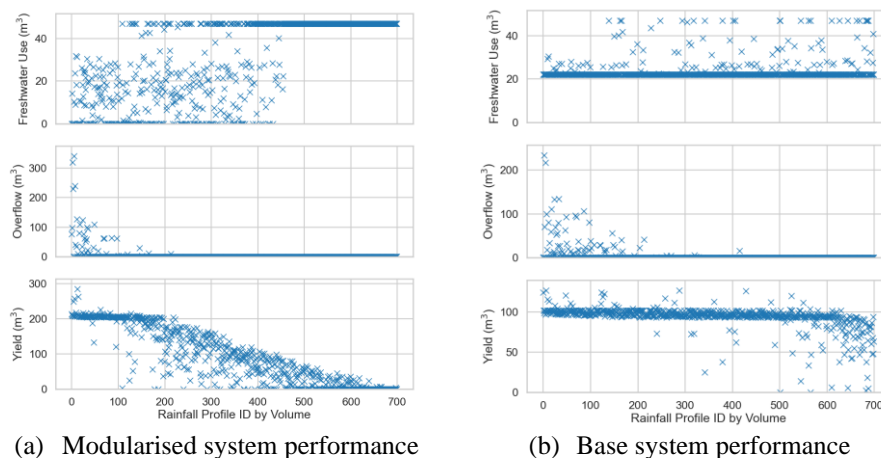
- *System configuration parameters*: The set of possible tank areas, tank heights and orifice heights.
- *Solution time*: The total allowable time for the model to find a solution.
- *Allowable capacity*: The maximum volume of space the derived tank can occupy.
- *Penalty costs*: The penalty values that influence the objective function value, if applicable in the specific model.
- *Stochasticity*: The number of rainfall scenarios used.

The sets of possible parameter values provided for the optimization models are selected as a large possible design space for a system that is to serve a 1.57 ha residential estate. Lower vertical space allowances are used as the underground make-up of a city is typically carefully subdivided between multiple utilities and different layers are hence reserved for distinct, specific purposes. The objective and penalty functions for each of the four tank systems are shown in Table 1, where the overflow threshold penalty assigns a cost to the overflow volume in exceedance of the threshold volume defined for a given timestep, and the wasted capacity penalty function implements a three-tiered cost for any wasted tank volumes above 10% of the total tank capacity.

For effective adoption of a novel sizing strategy, the proposed method needs to be sufficiently easy to implement, whilst providing system configurations that provide adequate levels of flood mitigation and yield collection. Hence, the following subsections discuss the impact of stochasticity on the required solution time and derived solution quality for the modularised optimisation framework.

#### 3.1. Stochasticity and solution time

The model's behaviour as the set of rainfall scenarios increases is characterised by the computational time required to derive an optimal solution. The set of rainfall scenarios is selected separately for each set size to ensure that each set meets the minimum threshold variance value. This was applied to each of the optimisation models in the multi-tank RWH system individually, such that the behaviour of each tank configuration model block can be characterised.



(a) Modularised system performance (b) Base system performance  
**Figure 3:** Scatterplot of key performance indicator values, in order of decreasing input rainfall volume.

The number of scenarios and size of the derived optimisation model have minimal impact on the computational requirement of the individual tank optimisations. On the other hand, the ‘Separation’ tank, which has the highest number of openings and hence constraints and variables, is significantly more computationally intensive than the other tank design types and requires eight-times the amount of computational time to derive an optimal solution.

3.2. Performance of the derived configuration

The output optimal configuration for the full system remained the same despite increasing the number of scenarios used with the modular optimisation model. This optimised output is shown, in comparison to an existing configuration in Table 2. The total time required to derive a full configuration with 25 rainfall scenarios selected with a minimum variance value of 50 was 1.5 hours when implemented on an Intel Core i7-8650U with 16.0 GB ram.

The performance of the derived configuration is determined for its ability to reduce surface overflow risks and provide local water availability using a simulation model. A statistical measure of the key performance indicators is generated by running simulations for 700 days of the highest rainfall volumes extracted from an eight-year historical timeseries, and the KPI values, sorted by descending input rainfall volumes are demonstrated in Figure 3.

Statistical performance metrics are then extracted from these figures to provide quantitative indicators of the system performance. Freshwater use indicates the amount of freshwater required to ensure that local non-potable water demands are fully met. The scatterplots above demonstrate that the existing configuration is limited in its ability to provide water for irrigation purposes in a timely manner and hence has a minimum freshwater requirement for all simulated scenarios. On the other hand, the optimised system is observed to be able to provide for all the required water demand for 16.3% of the simulated scenarios and only fails to provide sufficient water for reuse when the inflow volumes fall below a total of 230 m<sup>3</sup>.

**Table 2:** System configurations for an existing RWH system, and an optimised configuration for the same catchment area.

Tank	Existing configuration		Optimised configuration	
	Height (m)	Area (m <sup>2</sup> )	Height (m)	Area (m <sup>2</sup> )
Separation	1.23	5.02	2.5	80
Detention	2.00	264.52	0.5	10
Harvesting	2.40	38.20	2.5	80
Treatment	2.00	5.00	1.25	100
<b>Total</b>	<b>636.84 m<sup>3</sup></b>		<b>530 m<sup>3</sup></b>	
<i>Orifice</i>	<i>Height (m)</i>		<i>Height (m)</i>	
SH	0.52		1.5	
SD2	0.66		1.75	
DO2	1.8		0.4	

The optimised configuration has a higher peak overflow volume of 285m<sup>3</sup>, but the proportion of scenarios which demonstrated overflow is reduced from 10.86% to 5.86%. Thus, with the optimised configuration, overflow events are less likely to occur, but this comes at the price of having higher overflow volumes if they do. The existing configuration is better at harvesting some amount of water for 99.7% of the simulated scenarios, while the optimal configuration only harvested water in 78.3% of the simulated

scenarios. Hence, the optimal configuration is much better at providing water availability during larger rain events, whilst the existing configuration provides a much more consistent yield performance over the possible range of rainfall volumes.

#### 4. Conclusions

The results demonstrated that the modularised optimisation model can provide a full RWH system configuration for a multi-tank system with low computational time limits of 1.5 minutes. Under this optimisation-based sizing strategy, there is an improved balance between the dual objectives of improving local reuse of harvested water and reducing flood risks, especially in comparison to an existing RWH system configuration which had not been derived using optimisation-based methods.

Thus, the modularised optimisation tool presented in this paper can encourage more implementations of multi-objective RWH systems by improving the accessibility of optimisation-based sizing strategies. Through the tank block framework, optimisation models can easily be built if the desired configuration of discharge openings is known. Further behavioural specifications can also be appended to the model if this is required, and a full optimisation model, specific to the design for the given catchment, can be derived much more quickly in comparison to a full, single optimisation model.

The performance of a stochastic optimisation model, however, remains highly influenced by the data quality used for its input scenarios. Therefore, the performance of the modularised tool continues to be dependent on the availability of high-resolution rainfall data. To mitigate this, there is currently work in progress to integrate such a sizing strategy with a rainfall design model to derive a robust RWH system configuration.

#### Acknowledgements

This research is supported by the Singapore Ministry of National Development and the National Research Foundation, Prime Minister's Office under the Land and Liveability National Innovation Challenge (L2 NIC) Research Programme (L2 NIC Award No. L2NICTDF1-2017-3). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of the Singapore Ministry of National Development and National Research Foundation, Prime Minister's Office, Singapore.

#### References

- M. Semaan, S. D. Day, M. Garvin, N. Ramakrishnan, and A. Pearce, "Optimal sizing of rainwater harvesting systems for domestic water usages: A systematic literature review," *Resour. Conserv. Recycl. X*, vol. 6, no. November 2019, p. 100033, 2020.
- C. O. Okoye, O. Solyali, and B. Akintug, "Optimal sizing of storage tanks in domestic rainwater harvesting systems: A linear programming approach," *Resour. Conserv. Recycl.*, vol. 104, pp. 131–140, 2015.
- C. C. Nnaji, P. G. C. Emenike, and I. T. Tenebe, "An Optimization Approach for Assessing the Reliability of Rainwater Harvesting," *Water Resour. Manag.*, vol. 31, no. 6, pp. 2011–2024, 2017.
- C. Matos, C. Santos, S. Pereira, I. Bentes, and M. Imteaz, "Rainwater storage tank sizing: Case study of a commercial building," *Int. J. Sustain. Built Environ.*, vol. 2, no. 2, pp. 109–118, 2013.
- Q. Y. Soh, E. O'Dwyer, S. Acha, and N. Shah, "Stochastic optimal design for large-scale rainwater harvesting and detention systems," in *ECOS 2022 - the 35th international conference on efficiency, cost, optimization, simulation and environmental impact of energy systems*, 2022.

# Knowledge-based model and simulations to support decision making in wastewater treatment processes

Claire Valentin,<sup>a</sup> Frédéric Lagoutière,<sup>b</sup> Jean-Marc Choubert,<sup>c</sup> Françoise Couenne,<sup>a</sup> Christian Jallut,<sup>a</sup>

<sup>a</sup>*Univ Lyon, Université Claude Bernard Lyon 1, CNRS, LAGEPP UMR 5007, 43 boulevard du 11 novembre 1918, 69100, Villeurbanne, France*

<sup>b</sup>*Univ Lyon, Université Claude Bernard Lyon 1, CNRS, Institut Camille Jordan, UMR 5208, 43 boulevard du 11 novembre 1918, 69622 Villeurbanne Cedex, France*

<sup>c</sup>*INRAE REVERSAAL, 69625 Villeurbanne, France*  
*claire.valentin@univ-lyon1.fr*

## Abstract

In this paper, mass and momentum balances are used to model and simulate the most significant phenomena occurring in a continuous urban sludge clarifier. This model is designed to be the core of a digital twin in the future. This has two consequences: the choice of a one-dimensional model only and a numerical scheme for simulation that allows a reasonable runtime. We propose a different way of writing of the dynamic model of the clarifier. Instead of using volume fractions and velocity of solid particles as state variables, we use volume fractions and flux of solid particles. This approach, which is used for conservation law systems, gives more stable simulation results for hyperbolic systems.

This paper focuses on comparative discussions of these two different versions of the model. The numerical simulation scheme is presented. In addition, the simulation is compared with experimental data obtained in a full-scale wastewater treatment plant.

**Keywords:** Urban sludge continuous settling, Dynamic mass and momentum balances, nonlinear hyperbolic system, Rusanov numerical approximation

## 1. Introduction

The efficiency of wastewater treatment facilities is a worldwide major problem. Urban wastewater treatment plants are regulated by European directives, such as 91/271/EEC. The biological treatment unit of a treatment plant purifies wastewater before it is released into the natural environment by using a clarifier. The quality of the clarified water strongly depends on the instantaneous hydraulic loads arriving upstream, on the design of the equipment and on the operating conditions. The settling of the sludge in the clarifier enables the separation of more concentrated sludge that is pumped down and clarified water that is released up into the environment in an overflow. Thus, the quality of the plant effluent is highly dependent on the performance of the clarifier. In order to optimize its operation, modelling and simulation are the first steps to perform.

In the literature, the modelling of clarifiers is obtained through mass and momentum balance equations for solid particles. The latter is written as a static or dynamic partial differential equation (Chauchat (2013), Garrido (2003)) or replaced by a constitutive relation representing the velocity of solid particles (Li (2014)). In the case where the

momentum balance is represented by a PDE, the global model is weakly hyperbolic and the position of the so-called sludge blanket corresponds to the front of a shock wave. This shock is related to a discontinuity in the solid particle volume fraction and a peak value in the flux.

In this paper, we present two choices of state variables for the sludge settling modelling in a clarifier (Valentin (2022)) and discuss the simulation results with respect to experimental data.

## 2. A 1-D schematic representation of the clarifier

Fig. 1 shows a 1-D schematic representation of the clarifier. It is connected to the wastewater treatment process at three points, one inlet and two outlets:

- one inlet where the sludge, consisting of fluid and particles, flows into the clarifier by gravity from the upstream biological aeration tank at the volume flow rate  $Q_f(t)$  and with a particle concentration  $C_f(t)$ . The sludge feed is situated at a depth of  $z = z_f$ ,
- one top outlet for clarified water at  $z = 0$ , with a volume flow rate  $Q_e(t)$  and particle concentration  $C_e(t)$ ,
- one bottom outlet for compressed sludge at  $z = z_b$ , with a volume flow rate  $Q_u(t)$  and particle concentration  $C_u(t)$ . Some of the compressed sludge is recirculated back into the aeration tank at a volume flow rate  $Q_{ur}(t)$  and some is extracted from the clarifier at a volume flow rate  $Q_{ue}(t)$  such that  $Q_u(t) = Q_{ur}(t) + Q_{ue}(t)$ .

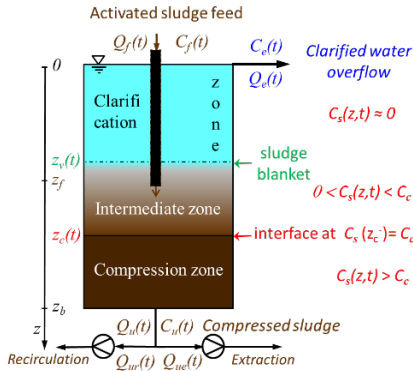


Figure 1: Schematic view of a clarifier

$A$	$1175 \text{ m}^2$
$A_k$	$0.02 \text{ m/s}$
$\epsilon_c$	$4.1 \cdot 10^{-2}$
$n_r$	$2$
$n_s$	$11$
$\rho_l$	$1000 \text{ kg/m}^3$
$\rho_s$	$1030 \text{ kg/m}^3$
$\sigma_0$	$0.5 \text{ kg/ms}^2$
$z_b$	$2.8 \text{ m}$
$z_f$	$1.8 \text{ m}$

Table 1: Model parameter values

The open-air settler content can be divided into two moving interfaces separating three zones:

- The upper interface is the sludge blanket, located at depth  $z_s(t)$ . It separates the clarification zone (which contains no or very few particles) from the intermediate zone,
- The lower interface is defined by the intermediate/compression threshold and is located at depth  $z_c(t)$  where a change in the particles behaviour occurs as the particle concentration  $C_s(z,t)$  exceeds the critical threshold  $C_c$ . Above this threshold, interparticle stress comes into effect. Under  $z_c(t)$ , the liquid phase flows through a porous network of concentrated particles, Toorman (1996). The model parameter values are given in Table 1.

## 3. A 1-D dynamic model of the clarifier

The knowledge-based model includes two dynamic mass and momentum balances of urban sludge added with physical constitutive equations under standard simplifying

assumptions, Valentin (2022). As the liquid and solid phase densities  $\rho_l$  and  $\rho_s$  are constant, the two most natural state variables are the particle volume fraction,  $\varepsilon_s$ , and the particle flux,  $f_s = \varepsilon_s v_s$  with  $v_s$  the particle velocity. The corresponding hyperbolic system of conservation laws is then defined, if  $\varepsilon_s > 0$ , by:

$$\partial_t \varepsilon_s + \partial_z f_s = \frac{f_{1s}}{\rho_s} \delta_f \quad (1)$$

$$\partial_t f_s + \partial_z \left( \frac{f_s^2}{\varepsilon_s} \right) = \varepsilon_s g \left( 1 - \frac{\rho_l}{\rho_s} \right) - \frac{\partial_z \sigma_e(\varepsilon_s)}{\rho_s} + \frac{r(\varepsilon_s)(\varepsilon_s v_m - f_s)}{\rho_s \varepsilon_s (1 - \varepsilon_s)} + \frac{f_{2s}}{\rho_s} \delta_f \quad (2)$$

with  $\sigma_e(\varepsilon_s)$  the interparticle stress between the particles,  $r(\varepsilon_s)$  the liquid/solid drag force and  $v_m$  the average velocity of the liquid/solid mixture (also the total volume flux of the suspension).  $f_{1s}$  and  $f_{2s}$  are the source terms representing the sludge feed inlet in the particle mass and momentum balances equations respectively. They depend on  $Q_f$  and  $C_f$ .  $\delta_f$  is a Dirac function that represents the location of the sludge feed at  $z = z_f$ . The two boundary conditions are  $v_s(0, t) = -\frac{Q_e}{A}$  and  $v_s(z_b, t) = \frac{Q_u}{A}$ .

Various kind of constitutive equations that describe the compression and drag phenomena have been proposed in the literature, Li (2014).  $r(\varepsilon_s) = \frac{\rho_l g}{A_k} \varepsilon_s^{2/(3-n_r)}$  is proposed in Chauchat (2013) and the following  $\sigma_e(\varepsilon_s)$  expression, that depends on the particle volume fraction, in Garrido (2003). If  $\varepsilon_s > \varepsilon_c$  (in the compression zone),  $\sigma_e(\varepsilon_s) = \sigma_0 \frac{\varepsilon_s^{n_s} - \varepsilon_c^{n_s}}{\varepsilon_c^{n_s}}$  else (in the two other zones),  $\sigma_e(\varepsilon_s) = 0$  (see Table 1).  $\sigma_e(\varepsilon_s)$  is a continuous function on  $[0, z_b]$  but its disadvantage is that it is zero over the upper part of the spatial domain which makes the system weakly hyperbolic only.

Most of the papers in the literature present a model based on a dynamic particle mass balance with the state variable,  $\varepsilon_s$ , coupled with a constitutive equation that gives the so-called batch or hindered settling velocity (Garrido, 2003), (Li, 2014). This motivated us and others, Chauchat (2013), Valentin (2022) to use first the state variables  $(\varepsilon_s, v_s)$  according to the following hyperbolic system of two PDEs:

$$\partial_t \varepsilon_s + \partial_z (\varepsilon_s v_s) = \frac{f_{1s}}{\rho_s} \delta_f \quad (3)$$

$$\partial_t v_s + \partial_z \left( \frac{v_s^2}{2} \right) = g \left( 1 - \frac{\rho_l}{\rho_s} \right) - \frac{\partial_z \sigma_e(\varepsilon_s)}{\rho_s \varepsilon_s} + \frac{r(\varepsilon_s)(v_m - v_s)}{\rho_s \varepsilon_s (1 - \varepsilon_s)} + \frac{f_{21s}}{\rho_s} \delta_f \quad (4)$$

This representation is based on the temporal and spatial derivatives of products such as  $\varepsilon_s v_s$ . As shock waves appear on the two state variables, such product derivatives may lead to mathematical problems. Moreover, the difficulty of numerically solving the model based on the state variables  $(\varepsilon_s, v_s)$  and the oscillations obtained on the spatial profiles of the velocity  $v_s$  led us to return to the more "natural" state variables that correspond to those used in the so-called conservative approach, the particle volume fraction,  $\varepsilon_s$ , and the particle flux,  $f_s = \varepsilon_s v_s$  (equations (1) and (2)) and to include the interparticle stress in the flux.

#### 4. Numerical discretization scheme

The simulations were carried out using explicit Euler time-discretization and an efficient numerical scheme adapted to hyperbolic and weakly hyperbolic nonlinear PDE systems: a finite volume method spatial-discretization with the Rusanov approximation of the fluxes (Godlewski (1996), LeVeque (2002)).

The state variable vector  $x$  and the flux vector  $F_s(x)$  are defined if  $\varepsilon_s > 0$  by  $x = \begin{pmatrix} \varepsilon_s \\ f_s \end{pmatrix}$

and  $F_s(x) = \begin{pmatrix} f_s \\ \frac{f_s^2}{\varepsilon_s} + \frac{\sigma_e(\varepsilon_s)}{\rho_s} \end{pmatrix}$ . The state variables are considered as uniform in each volume

$i$  of the mesh and equal to the average values  $\bar{x}_i^k$  at time step  $k$ .

For time step  $k$  and volume  $i$ , the flux at the input interface  $i - \frac{1}{2}$  is approximated by:

$$F_{i-\frac{1}{2}}^k = \frac{1}{2} (F_s^k(\bar{x}_{i-1}) + F_s^k(\bar{x}_i)) - \frac{\omega_s^k}{2} (\bar{x}_i^k - \bar{x}_{i-1}^k)$$

with  $\omega_s^k = \max(\rho_i^k)$ ,  $\rho_i^k$  the spectral radius of the volume  $i$  which depends on  $\bar{x}_i^k$ .

Then:

$$\bar{x}_i^{k+1} = \bar{x}_i^k + \frac{\Delta t}{\Delta z} (F_{i-\frac{1}{2}}(\bar{x}_i, t) - F_{i+\frac{1}{2}}(\bar{x}_i, t)) + \Delta t (S_1(\bar{x}_i) + S_2)$$

where we assume that  $S_1(\bar{x}_i)$  is a good approximation of  $\frac{1}{\Delta z} \int_{i-1/2}^{i+1/2} S_1(x) dz$

and that the variable time-step  $\Delta t$  respects the CFL (Courant-Friedrichs-Lewy) condition:  $\Delta t = \gamma \frac{\Delta z}{\omega_s^k}$  with  $0 < \gamma < 1$ . As  $\sigma_e(\varepsilon_s) = 0$  in the clarification and intermediate zones, the system is only weakly hyperbolic.

#### 5. Simulation of a transient state experiment of continuous settling

The model simulation is compared to experimental data obtained in a full-scale settler operated under the following transient state scenario (Fig.3): the sludge feed rate was abruptly increased by a magnitude of two from a value corresponding to a stationary profile at  $t = 1.40 \text{ am}$  ( $660 \text{ m}^3/\text{h}$ ) and abruptly decreased  $8 \text{ hrs}$  later ( $370 \text{ m}^3/\text{h}$ ). The flow rates and concentrations at the sludge inlet and outlet were measured on-line as well as the sludge blanket position.

Simulations based on this model were performed with various discretization parameters such as spatial mesh size, convergence condition (Courant-Friedrichs-Lewy). They are compared to the experimental data. Measured sludge feed flow rate,  $Q_f(t)$ , recirculation flow rate,  $Q_{ur}(t)$  and extraction flow rate,  $Q_{ue}(t)$  as well their mean values are given in Fig.3.

An N-node spatial mesh was used to run the simulations of the discretized model presented in section 4. with the constitutive equations and boundary conditions given in section 3.

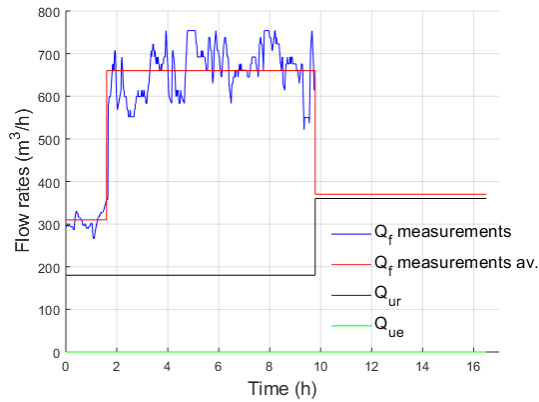


Figure 3: Applied flow rates

The calculated sludge blanket position  $z_v(t)$  corresponds to the location where the maximum gradient of the solid concentration is reached. Fig 4.a shows that the simulated and measured sludge blankets positions are very close.

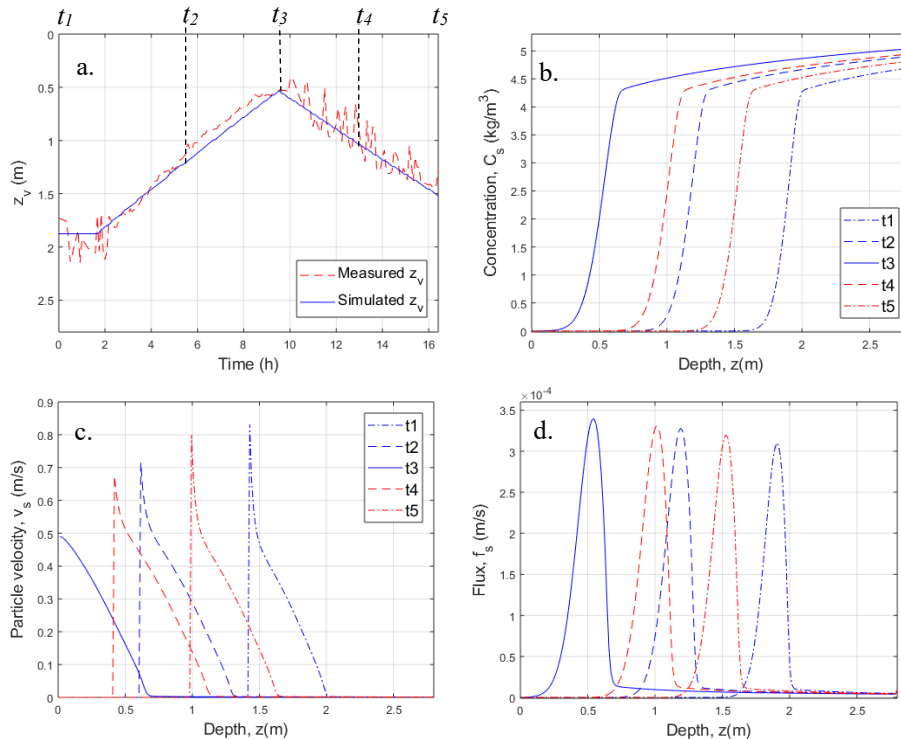


Figure 4: Simulated  $z_v(t)$  and  $C_s(z)$ ,  $v_s(z)$ ,  $f_s(z)$  spatial profiles ( $N=200$ )

Five spatial profiles at  $t_1$ ,  $t_2$ ,  $t_3$ ,  $t_4$  and  $t_5$  are shown in Fig 4.b to Fig 4.d. Fig 4.b presents particle concentration spatial profiles. The particle concentration at the top of the clarifier ( $z=0$ ) is close to zero. It increases sharply at the depth of the sludge blanket and then increases more moderately to the bottom as soon as the critical threshold is exceeded. The effect of the interparticle stress below the critical threshold is also well highlighted.



Spatial profiles  $v_s(z)$  and  $f_s(z)$  are shown in Fig. 4.c and Fig. 4.d. The shock wave is also clearly visible at the "peak" location. The particle velocity at the bottom of the settler  $v_s(z_b)$  is very low compared to the velocity at the sludge blanket position,  $5.3 \cdot 10^{-4} \text{ m/s}$ , and is set by the  $Q_u$  flow rate.

Good results are obtained with this numerical scheme for a mesh size of up to 800 volumes and  $\gamma = 0.99$  in the CFL condition. Choosing a lower  $\gamma$  value gives the same results but with a longer runtime due to a smaller  $\Delta t$ . The runtime of a simulation with 200 volumes on a workstation equipped with an Intel Xeon at 3.8GHz is 8 min 15 s. According to the nature of the numerical scheme, the runtime with  $N = 800$  is 16 times longer that makes 2 hrs 12 min. Although the simulation gives good results with  $N = 800$  with less numerical diffusion, as this model will be the core of a decision support system and then a digital twin, a compromise must be decided. A spatial discretization of  $N = 200$  seems suitable.

## 6. Conclusions

As a conclusion, the well-suited conservative state variables are the volume fraction and the volume flux of the solid particles. The expression of the momentum balance includes the compression stress between the particles based on a nonlinear constitutive law taken from Garrido (2003). A numerical scheme with explicit Euler time-discretization and a finite volume method spatial-discretization with the Rusanov approximation of the fluxes works well. The simulation results are close to experimental results by using a set of well-chosen parameter values. It can be used for prediction and decision support for other scenarios of operation. The runtime with a spatial mesh of 100 volumes is still about 2 min 4 s on a workstation equipped with Intel Xeon at 3.8GHz, which may be a limitation of this approach.

## Acknowledgements

The authors would like to thank the Auvergne-Rhone-Alpes region's PAI 2020 program, along with the C-StaRRE 4.0 project for their financial support, in particular for Hadi ABOU DAHER's internship (Grant number 22PACKR-CVALENTIN-6887).

## References

- J. Chauchat, S. Guillou, D. Pham van Bang and K. Dan Nguyen, 2013, Modeling sedimentation-consolidation in the framework of a one-dimensional two-phase flow model, *Journal of Hydraulic Research*, 51 (3), pp. 293-305.
- P. Garrido, F. Concha, R. Burger, 2003, Settling velocities of particulate systems: 14. Unified model of sedimentation, centrifugation and filtration of flocculated suspensions. *Int. J. Mineral processing*, vol. 72, pp.57-74.
- E. Godlewski, P.A. Raviart, 1996, Numerical Approximation of Hyperbolic Systems of Conservation Laws, *Applied Mathematical Sciences (AMS, volume 118)*, Springer.
- R. LeVeque, 2002, Finite volume methods for hyperbolic problems, Cambridge university press, vol. 31.
- B. Li and M.K. Stenstrom, 2014, Research advances and challenges in one-dimensional modeling of secondary settling Tanks - A critical review, *Water Research*, vol. 65, pp. 40-63.
- E.A. Toorman, 1996, Sedimentation and self-weight consolidation: general unifying theory. *Geotechnique*, vol. 46 (1), pp.103-113.
- C. Valentin, N. Chassin, F. Couenne, J.M. Choubert and C. Jallut, 2022, 1-D Dynamic knowledge based model of urban sludge continuous-flow settling process. Comparison with experimental results, prepublication submitted to a journal, 9 pages (hal-03678231v1).

# Homotopic Method for binodal curves computation in ternary liquid-liquid separation

Nataliya Shcherbakova,<sup>a</sup> Kevin Roger,<sup>a</sup> Joseph Gergaud,<sup>b</sup> Olivier Cots<sup>b</sup>

<sup>a</sup>*Laboratoire de Génie Chimique, Université de Toulouse, CNRS, INP, UPS, Toulouse, France*

<sup>b</sup>*IRIT, Université de Toulouse, CNRS, INP, UPS, Toulouse, France*

## Abstract

This paper analyses the intrinsic geometric structure of binodal curves of phase separation diagrams of ternary mixtures. A new computational method is presented, and the efficiency of the proposed approach is tested for the Flory-Huggins model identification of water – acetone – hexadecane mixture.

**Keywords:** liquid-liquid separation, bitangent plane, binodal curve, differential homotopy

## 1. Introduction

Controlling the phase separation in industrial processes requires detailed knowledge of thermodynamic diagrams of multi-component mixtures (Koningsveld et al., 2001). Many computational problems arise in this context: identification of model parameters from the experimental data, localization of phase separation envelopes, detection of multi-phase regions, etc. Despite the enormous progress in this domain in the past decades, there is still a real need for new fast and reliable algorithms, for instance, in industrial applications using bio-sourced mixtures or in polymer production.

This paper focuses on the analysis of liquid-liquid separation of ternary mixtures at constant temperature and pressure. The topological structure of phase separation diagrams is then defined by binodal and spinodal curves delimiting stable, metastable and unstable domains in the composition space. Knowing these domains is crucial for modeling the demixing process in a given mixture. The experimental detection of spinodals that delimits the material stability domain is a highly non-trivial task. Instead, the various techniques allows measuring the composition of coupled coexisting phases, which can be used to establish a reliable model in order to predict the spinodal location *a-posteriori*.

The standard method of numerical computation of binodal curves uses the Newton-Raphson-like iterative algorithm to solve a system of algebraic equations over the mesh covering the state space of the diagram, while the classical least square method is commonly used for model identification. Often this approach fails to meet at least one of the quality criteria characterizing the “good” model: the binodal curve shape and the bundle of tie-lines indicating the coupled phases composition. Inspired by the ideas of D.J. Korteweg (Levelt Sengers, 2002) applied to the topological analysis of the Gibbs free energy surface, in this paper, we propose a new numerical method for computing binodal curves by a kind of differential homotopy method. The key idea is to consider

binodals as the 2D projections of certain smooth curves in 4D configuration space formed by the pairs of points of the phase diagram. This geometrical point of view is presented in Section 2, where a set of ordinary differential equations (ODE) describing the 4D curve as an integral curve of a smooth vector field in 4D space is derived. A conventional ODE solver can numerically compute this curve at desired precision.

In Section 3, the developed approach is tested on the water-acetone-hexadecane mixture, characterized by two non-miscibility binary gaps among its components. The Flory-Huggins model parameters, including the triple interaction term, are computed by solving a non-trivial minNLP problem associated with a criterion, which accounts for the intrinsic geometry of the binodal curve. The computations performed using Mathematica 9 software yield very promising results that meet both quality criteria cited above.

## 2. Binodal and spinodal curves: thermodynamics vs. geometry

### 2.1 Phase coexistence conditions in multi-component mixtures

The physicochemical properties of an  $N$ -component system can be described by the Gibbs free energy  $G(P, T, n)$ , where  $n = (n_1, \dots, n_N)$ ,  $n_i$  being the number of moles of  $i$ -th component,  $P$  the pressure and  $T$  the temperature of the system. Being a homogeneous function of 1-st order with respect to  $n_i$ ,  $G$  can be expressed in terms of chemical potentials  $\mu_i = \frac{\partial G}{\partial n_i}$  as  $G(P, T, n) = \sum_{i=1}^N n_i \mu_i(P, T, n)$ . The infinitesimal changes in the state of the system obey the fundamental Gibbs equation

$$dG = -S dT + V dP + \sum_{i=1}^N \mu_i dn_i \quad (1)$$

where  $S$  and  $V$  are the entropy and the volume of the system.

All expressions above rest valid for each phase of a closed system maintained at thermodynamic equilibrium with two coexisting phases. Numbering the phases by  $a$  and  $b$ , the equilibrium condition reads  $dG = dG^a + dG^b = 0$ . Since in a closed system  $n_i = n_i^a + n_i^b$  and  $n_{tot} = \sum_{i=1}^N n_i$  are constant, it follows that  $dn_i^a = -dn_i^b$ , and hence Eq.(1) implies that

$$T^a = T^b, \quad P^a = P^b, \quad \mu_i^a(P^a, T^a, n^a) = \mu_i^b(P^b, T^b, n^b), \quad i = 1, \dots, n \quad (2)$$

In this paper, it is assumed that both phases can be described by the same thermodynamic model, though in principle, two different expressions can model be used for chemical potentials of phases  $a$  and  $b$ . In addition, only the isobaric isothermal conditions will be considered, so the first two of Eqs. (2) become trivial.

The physicochemical properties of real mixtures are usually expressed in terms of molar, volume or mass fractions. In particular, the Flory-Huggins model used in Section 3 to describe the mixture of a polymer with water and a solvent, employs the volume fractions. Assuming the total volume of the mixture to be equal to the sum of partial volumes of components, and denoting by  $x_i = v_i/V$  the volume fraction of  $i$ -th component and by  $g(P, T, x) = G(P, T, n(x))/R V$  the free Gibbs energy of mixing per unit of volume, Eqs. (2) can be rewritten in the form

$$\begin{aligned} \frac{\partial g^a(x^a)}{\partial x_i^a} &= \frac{\partial g^b(x^b)}{\partial x_i^b}, \quad i = 1, \dots, N-1, \\ g^a(x^a) - g^b(x^b) - \sum_{i=1}^{N-1} \frac{\partial g^a(x^a)}{\partial x_i^a} (x_i^a - x_i^b) &= 0 \end{aligned} \quad (3)$$

In the binary case ( $N=2$ ) these conditions mean that the graph of function  $g(x)$  admits a bitangent line, as it is shown in Fig. 1a.

### 2.2 Ternary case: binodal curves and bitangent planes of the Gibbs free energy surface

To simplify notations, from now on  $g'_i$  stays for  $\frac{\partial g}{\partial x_i}$ , the symbols  $\nabla_x g$  and  $D_x^2 g$  are used for the gradient and for the Hessian of function  $g$ , and  $(|)$  for the scalar product in  $\mathbb{R}^2$ .

In case of ternary mixtures Eqs.(3) describe the specific properties of the surface, referred to as surface  $W$ , defined by equation  $z = g(x_1, x_2)$  in a 3D Cartesian space endowed with coordinates  $x_1, x_2, z$ , as it shown in Fig. 1b. The material stability condition implies that such a surface has a physical meaning only in the sub-domain of the volume fraction space where the function  $g(x)$  is convex:

$$\Omega = \{x = (x_1, x_2): x_1, x_2 \in [0,1], x_1 + x_2 \leq 1, D_x^2 g \text{ is weakly pos. definite} \}$$

The projection on the  $x$ -plane of the set of singular points of surface  $W$ , i.e. the points where  $D_x^2 g$  is singular, defines the spinodal curve of the phase diagram. The first two of Eqs.(3) guarantee the existence of a pair of points  $P_1$  and  $P_2$  on the surface  $W$  with collinear normals, whereas the latter of Eqs.(3) says that these points belong to the same plane tangent to  $W$ . The projection of the segment  $P_1 P_2$  on  $x$ -plane is called a tie-line. The one-parametric family of conodal points  $P_1, P_2$  define two directrices of a certain ruled surface in 3D space. Their projections on the  $x$ -plane correspond to the two branches of binodal curve on the phase diagram. These branches can meet each other at the critical point of the phase diagram. Since by definition the critical point also belongs to the spinodal, these are the only common point between the binodal and the spinodal curves, and the whole binodal curve lies on the same side with respect to the spinodal curve. Fig.1b. illustrates all these concepts for the mixture having one critical point.

### 2.3 Differential equations of binodal and spinodal curves

As it follows from the above analysis, the binodal curves are formed by the pairs of points. So it is natural to describe these curves in the 4D space  $\Sigma = \Omega \times \Omega$  defined as follows:

$$\Sigma = \{ \mathbf{q} = (\mathbf{q}_1, \mathbf{q}_2) \in \mathbb{R}^4 : \mathbf{q}_1 = \mathbf{x}^a \in \Omega, \mathbf{q}_2 = \mathbf{x}^b \in \Omega \}$$

Consider now the three co-dimension one smooth sub-manifolds in  $\Sigma$  associated to the zero-levels of the functions

$$\begin{aligned} F_1(q) &= g'_1(q_1) - g'_1(q_2), \quad F_2(q) = g'_2(q_1) - g'_2(q_2), \\ F_3(q) &= g(q_2) - g(q_1) + (\nabla_x g(q_1)|q_2 - q_1), \quad q \in \Sigma \end{aligned} \quad (4)$$

These sub-manifolds can intersect each other forming a one-dimensional sub-manifold  $\mathbf{B} = \{ \mathbf{q} \in \Sigma: \mathbf{F}(\mathbf{q}) = \mathbf{0}, \mathbf{i} = \mathbf{1}, \mathbf{2}, \mathbf{3} \}$  whose orthogonal projections on two exemplars of  $\Omega$

define the two branches of the binodal curve. In what follows  $B$  will be referred to as the generalized binodal curve. This curve can be seen as the integral curve of the vector field  $\mathbf{q} \rightarrow \mathbf{V}(\mathbf{q}) \in T_{\mathbf{q}}\mathbf{B}$ . By definition,

$$\nabla_{\mathbf{q}} F_i(\mathbf{q})V(\mathbf{q}) = 0, \quad i = 1,2,3 \quad (5)$$

Define a pair of functions  $\Phi_i$  and  $\Psi_i$  such that  $(\Phi_i, \Psi_i)^T = \mathbf{D}_{\mathbf{x}}^2 g(\mathbf{q}_i)(\mathbf{q}_1 - \mathbf{q}_2)$ , for  $i=1,2$ . Then Eqs. (5) are equivalent to the following conditions:

$$\Phi_1 V_1 + \Psi_1 V_2 = 0 \quad \text{and} \quad \mathbf{D}_{\mathbf{x}}^2 g(\mathbf{q}_1)(V_1, V_2) = \mathbf{D}_{\mathbf{x}}^2 g(\mathbf{q}_2)(V_3, V_4) \quad (6)$$

The straightforward computation shows that  $\mathbf{V}(\mathbf{q})$  can be expressed as

$$\begin{aligned} V_1 &= \Psi_1 \det \mathbf{D}_{\mathbf{x}}^2 g(\mathbf{q}_2), & V_2 &= -\Phi_1 \det \mathbf{D}_{\mathbf{x}}^2 g(\mathbf{q}_2) \\ V_3 &= \Psi_2 \det \mathbf{D}_{\mathbf{x}}^2 g(\mathbf{q}_1), & V_4 &= -\Phi_2 \det \mathbf{D}_{\mathbf{x}}^2 g(\mathbf{q}_1). \end{aligned} \quad (7)$$

Clearly, the critical points of surface  $W$  correspond to the singular points of  $\mathbf{V}(\mathbf{q})$ .

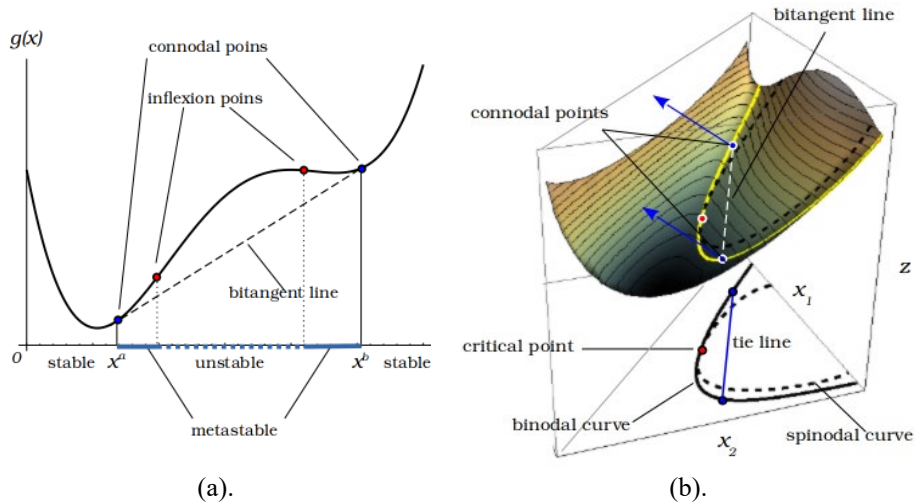


Fig.1. bitangent lines to the free Gibbs energy surface  $W$  in binary (a) and ternary (b) cases, and their projections on the composition space.

**Remark.** The same argument can be used to define the spinodal curve as the integral curve of the vector field  $\nabla_{\mathbf{x}} \mathbf{S}(\mathbf{x})^\perp$ , where  $\mathbf{S}(\mathbf{x}) = \det \mathbf{D}_{\mathbf{x}}^2 g(\mathbf{x})$ .

The described construction is a geometrical interpretation of the differential homotopy method (Allgower and Georg, 2003) for computing smooth curves defined implicitly. Using Eq.(5) instead of Eqs.(3) it is possible to compute binodal curves by a conventional ODE solver without solving the set of algebraic Eqs.(3) or Eqs.(2) by a Newton-Raphson type iterative algorithm, as it is usually done. The only iterative step concerns the computation of the starting point of the generalized binodal (or spinodal) curve, which reduces to a 2D problem. An algorithm based on the same principles was described in detail in N. Shcherbakova et al. 2017 and O. Cots et al. 2021.

### 3. Case study: Flory-Huggins parameters of water-acetone-hexadecane

#### 3.1 The Flory-Huggins model with ternary interaction term

The classical Flory-Huggins model defines the excess free energy of mixing per unit of volume according to the expression

$$g = \varphi_1 \ln \varphi_1 + \frac{1}{N_2} \varphi_2 \ln \varphi_2 + \frac{1}{N_3} \varphi_3 \ln \varphi_3 + \sum_{i,j=1,i<j}^3 \chi_{ij} \varphi_i \varphi_j + \beta \varphi_1 \varphi_2 \varphi_3 \quad (8)$$

Here  $\varphi_i$ ,  $i=1,2,3$  are the volume fractions of water, the solvent and the polymer respectively,  $N_2$  and  $N_3$  are the number of segments in the molecules of the solvent and of the polymer referred to the water segment,  $\chi_{ij}$  and  $\beta$  are the binary and ternary interaction coefficients. In this paper  $N_2$ ,  $N_3$ ,  $\chi_{ij}$  and  $\beta$  are assumed to be constant. Notice that the cross term  $\beta \varphi_1 \varphi_2 \varphi_3$  contains several physical effects. Indeed, it is the simplest possible correction of the Flory-Huggins model that accounts for composition-dependent binary interaction parameters as well as for the failure of the total volume conservation hypothesis when mixing the components. Since the present work aims to propose a generic robust method for binodal identification, it is not necessary to investigate the real physical meaning of the cross term here.

In further computations the volume fractions of acetone and of polymer are chosen as the independent variables so that  $x_1 = \varphi_2$ ,  $x_2 = \varphi_3$  and  $\varphi_1 = 1 - x_1 - x_2$ . Inserting Eq.(8) into Eqs.(4) yields three algebraic conditions to be verified along the binodal curve. Due to the particular form of Eq. (8), these conditions linearly depend on six scalar parameters:  $\chi_{12}$ ,  $\chi_{13}$ ,  $\chi_{23}$ ,  $\beta$ ,  $r_1 = N_2^{-1}$ , and  $r_2 = N_3^{-1}$ .

#### 3.2 Case study: water-acetone-hexadecane

The further analysis is based on the experimental data obtained at  $P = 1 \text{ atm}$  and  $T=23^\circ\text{C}$  using Raman's spectroscopy to measure each species fraction for a given sample. 14 measurements of coupled compositions were acquired defining tie-lines  $q^k \in \Sigma$  (black points in Fig. 2) and 17 non-coupled composition measurements  $x_m \in \Omega$  (white circles in Fig. 2) defining the phase-separation envelope. The latter series contains a pair of points defining the miscibility gap  $x^a$ ,  $x^b$  of the binary mixture acetone-hexadecane. Using Eqs.(3) for this binary mixture,  $\chi_{23}$  and  $N_2$  were computed in terms of  $x^a$ ,  $x^b$ . Analogously,  $\chi_{13}$  and  $N_3$  were found using the miscibility gap limits of water-hexadecane mixture reported in the literature. Due to the significant variation of these data in different sources, the term  $\chi_{13}$  in Eq. (8) was replaced by  $\chi_{13} + \delta$  in the next step of the computation. In this second step the remaining parameters  $\chi_{12}$ ,  $\beta$ ,  $\delta$  were found by solving the following non-linear optimization problem:

$$\min_{\chi_{12}, \beta, \delta} \sum_{k=1}^{14} \frac{F_1^2(q^k) + F_2^2(q^k)}{(1 - F_3^2(q^k))^2} \quad (9)$$

The particular form of criterion Eq.(9) push the minNLP solver toward the solutions which guarantee that each point  $q_k$  corresponds to a pair of points on  $W$  having collinear

normal vectors, whereas the denominator term penalize those pairs of points that do not belong to the same bitangent line.

The described method was implemented using Mathematica 9 package, the result is shown Fig. 2. Grey lines correspond to the measured tie-lines, and the dashed lines denote the computed tie-lines. The thick black curve corresponds to the computed binodal curve, which satisfies both quality criteria discussed above. The right part of Fig.2 shows the acetone vertex zone in high resolution, showing an excellent correspondence between experimental and model identification results.

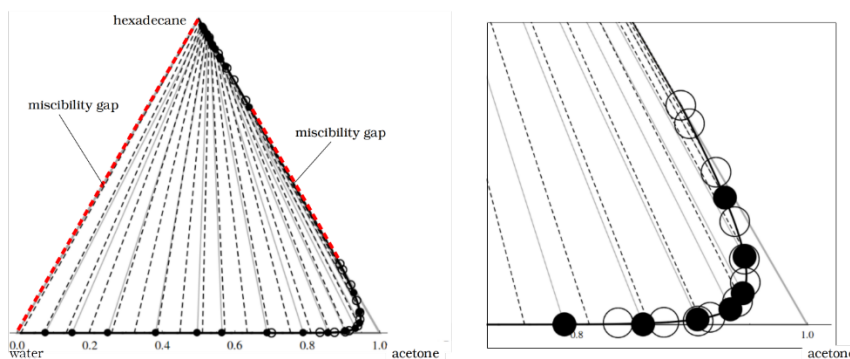


Fig.2: water-acetone-hexadecane mixture, experimental data (black and white points) vs. numerical computation of the binodal curve using Flory-Huggins model (black curve).

#### 4. Conclusion

The geometric description of binodal curves provides a deeper insight into their internal structure. In particular, it allows reducing numerical computation of binodal and spinodal curves to standard ODEs integration, gaining in execution time and accuracy. The numerical method proposed in this paper can be adapted to any thermodynamic model. We plan to implement the resulting differential homotopy algorithm in the next version of the SMITH code (Cots et al, 2021).

#### References

- R. Koningsveld, W.H.Stockmayer, E. Nies. *Polymer Phase Diagrams*. Oxford Univ. Press, 2001
- J. Levelt Sengers, *How fluids unmix*. Royal Netherlands Acad. of Arts and Sc., Amsterdam, 2002
- E. Allgower, K. Georg. *Introduction to numerical continuation methods*. *Classica in Applied Mathematics*, vol.45, SIAM, Philadelphia, USA, 2003
- O. Cots, N. Shcherbakova, J. Gergaud. SMITH: differential homotopy and automatic differentiation for computing thermodynamic diagrams of complex mixtures. *Computer Aided Chem. Eng.*, 50, pp. 1081-1086, 2021
- N. Shcherbakova, I. Rodriguez, J. Abidskov, V. Gerbaud. A Novel Method for Detecting and Computing Univolatility Curves in Ternary Mixtures. *Chem. Eng. Sci.*, 173, pp. 21-36, 2017
- Mathematica 9. <http://www.wolfram.com/mathematica/>

# Modelling of iron oxide reduction with hydrogen

Emiliano Salucci,<sup>a</sup> Snigdha Ghosh,<sup>b</sup> Vincenzo Russo,<sup>c</sup> Henrik Grénman,<sup>a</sup> Henrik Saxén<sup>a</sup>

<sup>a</sup>*Åbo Akademi University, Henrikinkatu 2, Turku 20500, Finland*

<sup>b</sup>*IIT (ISM) Dhanbad, Jharkhand, Dhanbad 826001, India*

<sup>c</sup>*University of Naples Federico II, Via Vicinale Cupa Cintia 26, Napoli 80126, Italy*

## Abstract

The steel industry represents one of the key sectors of our society and it is, therefore, essential to devise new strategies to drastically reduce its carbon dioxide emissions. From this point of view, a potentially beneficial solution would be to replace the carbon sources used in the reduction processes of minerals with hydrogen. This requires a deeper understanding of the kinetics of iron oxide reduction in order to provide crucial elements for selecting the best-operating conditions and designing new technologies. In this work, an mathematical model was developed, based on the assumptions of the shrinking core model, to describe the reaction kinetics and mass transport phenomena occurring during iron ore reduction. Experimental tests were conducted to provide detailed data to support this work. The model results qualitatively agree with what was found in the laboratory tests.

**Keywords:** Hydrogen-based DRI, green-ironmaking, iron oxides, kinetics-modelling.

## 1. Introduction

The aim of reducing greenhouse gas emissions is driving all manufacturing and industrial sectors to seek alternative solutions that are viable in the short term and offer excellent margins for improvement in the years ahead. With around 30% of global direct industrial emissions (2.6 Gt/year), steelmaking represents one of the largest CO<sub>2</sub> emitting sectors (Energy Agency 2020). The classical production process, carried out in a blast furnace (BF), involves the use of iron ores (i.e., iron oxides) which are reduced to metallic iron using carbon monoxide. This gas component is produced at partial oxidation of coal and reacts with the oxygen in the ore to produce iron and CO<sub>2</sub> (Hutny et al. 1991). As a result of the very high operating temperatures of the BF process, which locally exceed 2000°C, iron melts and slag forms in the lower part of the furnace. Developed 40 years ago, the Midrex process represents a possible alternative to conventional iron production. The reduction reaction in this process occurs at lower temperatures than in the BF and it usually operates on reformed natural gas. The gas-solid reactions produce direct reduced iron in solid state (DRI), and thus avoids melting (Atsushi et al. 2010). To achieve the complete decarbonization of this sector, it is necessary to redesign current production facilities and to replace all or most of the carbon source by hydrogen (Patisson et al. 2021). The kinetics of iron oxide reduction by hydrogen has been extensively studied and the literature shows that it is influenced by many factors and parameters. Several studies have focused on the thermodynamics of the reduction. At operating temperatures above 570°C, the reactive network consists of three reversible reactions in series, starting with the reduction of hematite (Fe<sub>2</sub>O<sub>3</sub>) and passing through the intermediates magnetite (Fe<sub>3</sub>O<sub>4</sub>) and wustite (FeO); at lower temperatures, wustite does not form due to its



instability (Zhang et al. 2013). To investigate the kinetic dependence of the reduction on the main operating conditions, several works have examined in depth the effect of temperature (Bernasowski 2014) or the composition of the reducing gas (Wei et al. 2018) on the system. Many authors focused their efforts on the determination of the rate-limiting step of the heterogeneous process and on the importance of diffusion phenomena. For this reason, the study of morphology, geometry, and other properties of the solid is crucial. Several authors have investigated the effect of the size of the iron ore (e.g., particle diameter) (Bonalde et al. 2005; Hou et al. 2012), the initial porosity (Kazemi et al. 2017) or the gas flow rate to assess the resistance to gaseous diffusion around a pellet (Hou et al. 2012). The main mathematical models used to describe the process and evaluate the kinetic parameters are the Avrami Model (Piotrowski et al. 2005), often used in metallurgical work, the Shrinking Core Model (SCM) (Moon et al. 1998) or the Grain Model (Bonalde et al. 2005). As for the assessment of the rate-limiting step, different conclusions have been made, mainly proposing an intraparticle diffusive regime, a kinetic regime, or a variable regime, depending on the reduction stage considered (Lin et al. 2003). Given the variety of experimental conditions and modelling assumptions, a wide range of activation energies, are reported in the literature. The activation energy for individual intermediate reactions or for the whole process can vary between 18 kJ/mol and 246 kJ/mol (Pineau et al. 2006). To accurately describe the chemical kinetics and mass transfer phenomena of the reduction process, an extensive kinetic study is required. This work focuses on the development of mathematical models for the description of iron oxide reduction based on a large set of experiments, which will be useful in the future for the development and design of hydrogen-based DRI processes.

## 2. Materials and methods

### 2.1. Experimental Setup

To support the development of a mathematical model capable of describing the reduction process, several experimental tests have been conducted using an Autochem2910. This device, which is normally used for the characterization of catalysts (e.g., TPR, TPO etc.), makes it possible to study the reduction of iron ores in a small packed-bed reactor. The experimental apparatus consists of a U-tube glass, resistant to high temperatures, with branches of different diameters. Ore fines, loaded in the branch with the largest diameter, is held by a porous support placed above the reduction of the tube diameter. The reactive bed (diameter 8 mm, height approx. 1 mm) consists of hematite powder (purity 97%), with a particle diameter of less than 5  $\mu\text{m}$ . The reduction of hematite was carried out using a gas mixture containing the reducing species ( $\text{H}_2$  purity 99.9%) and an inert gas (Ar, purity 99.9%). The temperature was measured using a thermocouple placed inside the tube. The composition of the gas leaving the bed was analyzed using a thermal conductivity detector (TCD) and a mass spectrometer, from which the degree of iron oxide reduction can be estimated. The tests were conducted isothermally at a specific  $\text{H}_2/\text{Ar}$  ratio and total volumetric gas flow rate.

### 2.2. Mathematical Model

A mathematical model has been developed to describe the reactive and diffusive phenomena present in a bed of iron oxide particles through which a gas mixture containing a reducing gas species flow. The reactive network consists of several equilibrium reactions in series, the number of which varies depending on the operating temperature. Table 1 shows the reactions occurring at temperatures above 570°C:

Table 1. Reaction network,  $T > 570^\circ\text{C}$ , and enthalpies (Zhang et al. 2013)

Reaction fronts ( $j$ )	Reduction reaction	$\Delta H^0$ [kJ/mol]
1	$3 \text{Fe}_2\text{O}_3 + \text{H}_2 \Leftrightarrow 2 \text{Fe}_3\text{O}_4 + \text{H}_2$ Hematite Magnetite	-11.72
2	$\text{Fe}_3\text{O}_4 + \text{H}_2 \Leftrightarrow 3 \text{FeO} + \text{H}_2\text{O}$ Magnetite Wustite	+77.44
3	$\text{FeO} + \text{H}_2 \Leftrightarrow \text{Fe} + \text{H}_2\text{O}$ Wustite Iron	+23.85

To describe the intraparticle reductive process, the mathematical model was developed based on the key assumptions of the SCM, where the reactions only take place on the surface separating two zones within the solid, an unreacted core, consisting of pure solid reactant, and a reacted outer layer, consisting of pure solid product. The reaction interface moves inwards as the core is consumed. By extending the SCM to the series reaction system, three different intraparticle reaction fronts can be described, as shown in Figure 1.

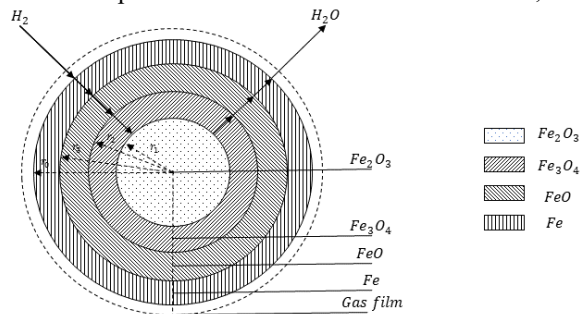


Figure 1. Three-interface SCM, reduction of hematite particle.

### 2.2.1. Mass Balances

The balance equations useful for describing the degrees of reduction of iron oxides over time ( $t$ ) and along the axial coordinate of the reactor ( $z$ ) were expressed as

$$\frac{\partial X_j(t, z)}{\partial t} = \frac{N_j(t, z)}{n_{00} \cdot a_j} \quad (1)$$

where  $X_j$  is the degree of conversion of the  $j$ -th reaction front ( $\text{Fe}_2\text{O}_3$  reduction, etc.),  $n_{00}$  is the initial oxygen content in the system while  $a_j$  is the proportional oxygen content parameter depending on the type of iron oxide considered.  $N_j$  represents the reduction rate of the  $j$ -th reaction front defined under the assumptions of the SCM.  $N_j$  is given by

$$N_j(t, z) = \frac{P}{R_g T} \frac{3}{r_0} \frac{1}{A_j(t, z)|_f} \cdot \left( \frac{R_j(t, z)|_f}{W(t, z)|_f} - y_{eq,j} \right) \quad (2)$$

where  $P$  and  $T$  represent the operating pressure and temperature,  $R_g$  is the gas constant,  $r_0$  is the particle radius of the ore, while  $y_{eq,j}$  represents the fraction of  $\text{H}_2$  at the equilibrium of the  $j$ -th reaction.

In the initial stages of reduction, with the formation of the intermediate species magnetite and wustite, a triple reaction front is established in the particles. Once the reduction of hematite has been completed, at any  $t$  and  $z$  ( $X_1(t, z) = 1$ ), the equations useful for describing the entire reactive process at that specific location must change as the model no longer describes a triple but a double reaction front. Identically, this occurs when  $\text{FeO}$  is the only residual oxide, effectively establishing a single reaction front. The formulae of the coefficients ( $A_j$ ,  $W$ ,  $R_j$ ), which define the reduction rate, are closely related to the

number of available reaction fronts. The value of  $f$  defines the number of reaction fronts present at time  $t$  and the  $z$ -coordinate, according to

$$\begin{cases} \text{if } X_1(t, z) < 1 & f = 3 \quad j = 1, 2, 3 \\ \text{if } X_1(t, z) = 1 \cap X_2(t, z) < 1 & f = 2 \quad j = 2, 3 \\ \text{if } X_1(t, z) = 1 \cap X_2(t, z) = 1 \cap X_3(t, z) < 1 & f = 1 \quad j = 3 \end{cases} \quad (3)$$

As an example, Eqs. (4-5) give the formulas for  $R_1|_{f=3}$  and  $W|_{f=3}$ , which are functions of the terms representing kinetic ( $A_j$ ), intra-particle diffusive ( $B_j$ ) and external diffusive ( $F$ ) resistance, respectively

$$\begin{aligned} R_1(t, z)|_3 &= (A_1 A_2 A_3) \cdot y_{\text{H}_2}(t) + [A_1 A_2 (B_3 + F)] \cdot y_{\text{eq}3} \\ &+ [A_1 A_3 B_2 + A_1 A_3 (B_3 + F) + A_1 B_2 (B_3 + F)] \cdot y_{\text{eq}2} \\ &+ [A_3 B_1 (A_2 + B_2) + A_3 B_1 (A_2 + B_2) + B_1 (A_2 + B_2) (B_3 + F) + A_2 A_3 B_2 \\ &\quad + A_2 A_3 (B_3 + F) + A_2 B_2 (B_3 + F)] \cdot y_{\text{eq}1} \end{aligned} \quad (4)$$

$$\begin{aligned} W(t, z)|_3 &= \{(A_1 + B_1) \cdot [A_3 (A_2 + B_2 + B_3 + F) + (A_2 + B_2) \cdot (B_3 + F)] \\ &\quad + A_2 [A_3 (B_2 + B_3 + F) + B_2 (B_3 + F)]\} \cdot y_{\text{eq}1} \end{aligned} \quad (5)$$

The complex derivation of these equations is reported in the literature (Hara et al. 1974). Examples of typical formulations of the  $A_j$  and  $B_j$  terms are given in Eqs. (6-7) while, in the first instance, due to the small size of the particles, the external transport resistance is considered negligible ( $F=0$ ).

$$A_1(t, z)|_3 = \frac{1}{(1 - X_1(t, z))^{\frac{2}{3}}} \cdot \frac{1}{k_1 \cdot \left(1 + \frac{1}{K_1}\right)} \quad (6)$$

$$B_1(t, z)|_3 = \frac{(1 - X_2(t, z))^{\frac{1}{3}} - (1 - X_1(t, z))^{\frac{1}{3}}}{(1 - X_1(t, z))^{\frac{1}{3}} \cdot (1 - X_2(t, z))^{\frac{1}{3}}} \cdot \frac{r_0}{D_1} \quad (7)$$

where  $k_j$  and  $K_j$  represent the kinetic constant and the equilibrium constant of the  $j$ -th reaction front, while  $D_j$  is the intra-particle diffusion coefficient that depends on the morphological characteristics of the ferrous oxide. The kinetic (Chu et al. 2005), thermodynamic (Zhang et al. 2013), and diffusive (Takahashi et al. 1971) parameters used in the model were taken from the literature.

Under dynamic conditions, two mass balance equations were developed for the gas species involved in the reduction reaction

$$\frac{\partial c_i(t, z)}{\partial t} = \underbrace{-\frac{u}{h} \cdot \frac{\partial c_i(t, z)}{\partial z}}_{\text{Convection}} + \underbrace{\frac{D_B}{h^2} \cdot \frac{\partial^2 c_i(t, z)}{\partial z^2}}_{\text{Diffusion}} + \underbrace{\frac{1}{(1 - \varepsilon_B)} \cdot \sum_{j|f}^3 v_{i,j} \cdot N_j(t, z)}_{\text{Reaction}} \quad (8)$$

where  $c_i$  is the gas concentration in the reactor ( $\text{H}_2$  or  $\text{H}_2\text{O}$ ),  $u$  represents the velocity of the gas,  $h$  is the height of the reactor and  $D_B$  is the axial dispersion coefficient. In the term describing the chemical reaction,  $v$  is the stoichiometric coefficient of the  $i$ -th gas species of the  $j$ -th reaction while  $\varepsilon_B$  is the void degree in the bed. Finally, the sum of the reductive contributions ( $N_j$ ) must consider the number of active reaction fronts.

The boundary conditions of the system of partial differential equations were defined at the reactor inlet ( $z = 0$ ) and at the outlet ( $z = 1$ ):

$$\frac{\partial c_i(t, z)}{\partial z} \Big|_{z=0} = c_{\text{bulk}, i} \quad \frac{\partial c_{\text{H}_2}(t, z)}{\partial z} \Big|_{z=1} = 0 \quad (9)$$

### 2.2.2. Numerical Methods

The mathematical model was developed in MATLAB 2022 by solving the system of partial differential equations, setting 50 grid points for the axial dimensionless coordinates of the reactor. The numerical method used is the finite difference method, which is based on backward differentiation formulas.

### 3. Results

In this work, an intensive simulation study was conducted to assess the responsiveness of the developed mathematical model using kinetic parameters reported in the literature. As an example, Figure 2 shows the main results from a simulation conducted at 600 °C, total gas volume flow rate of 20 mL/min and  $H_2/Ar$ .

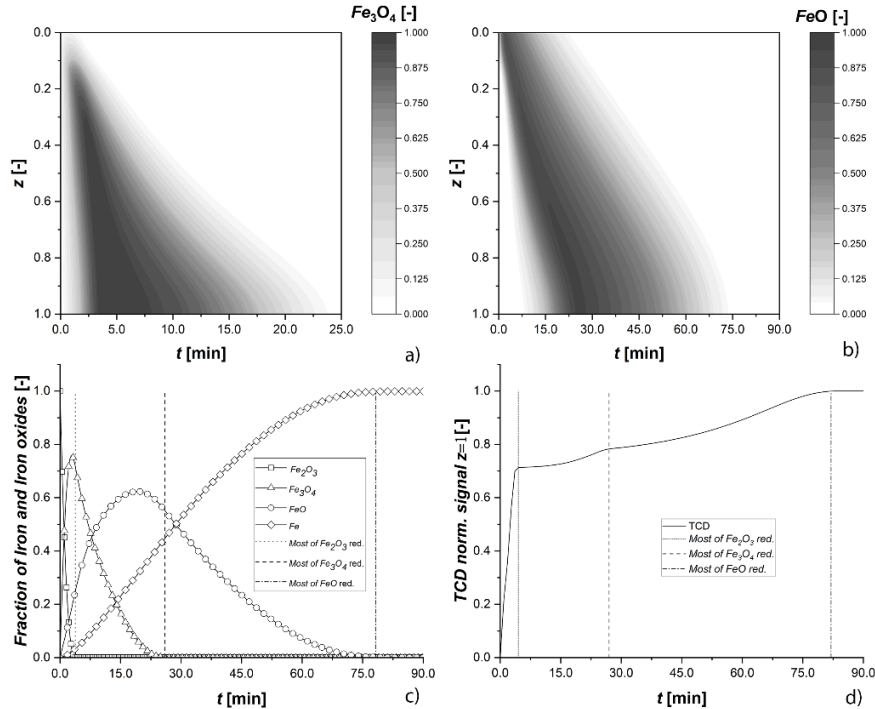


Figure 2. Axial profiles over time of the oxides  $Fe_3O_4$  (a) and  $FeO$  (b). Simulated overall fractions of the oxides and iron over time (c). Simulated TCD signal from off-gas composition (d).

The contour diagrams in Figures 2a and 2b show the axial profiles of the  $Fe_3O_4$  and  $FeO$  fractions in the particles over time, respectively, while Figure 2c shows the overall composition of the system. The reduction of  $Fe_2O_3$  is seen to be completed extremely rapidly, promoting the formation of high amounts of  $Fe_3O_4$  throughout the reactor after only a few minutes. Close to the entrance of the reactor ( $z=0$ ), magnetite does not accumulate in large quantities due to its rapid reduction to wustite, but near the end of the reactor ( $z=1$ ) its accumulation and depletion are considerably slower. The reasons lie in the high consumption of  $H_2$  by the reduction in the overlying areas of the bed and the resistance to intra- and inter-particle transport. Due to hydrogen starvation, the amount of  $H_2$  reaching the lower part of the reactor does not provide the driving force required to reduce the oxides. For this reason, about 25 minutes are required to complete the reduction of  $Fe_3O_4$ . The reaction of  $FeO$  to iron turns out to be the slowest: at the reactor inlet a complete reduction of the particles is observed within about 30 minutes, but close to the outlet the time required is more than doubled. This reveals the impact of the reactor bed height on the overall reaction time. Figure 2d shows a (normalized) TCD signal calculated based on the simulated off-gas composition ( $H_2$ ,  $H_2O$  and  $Ar$ ). This was found to show very many common features with the TCD signals measured in the experiments. Therefore, the objective of forthcoming work is to estimate kinetic parameters of the model by fitting it to the empirical TCD curves. However, the parameter estimation will

not be undertaken until data from a more extensive set of reduction experiments will be available.

#### 4. Conclusions

In this work, an in-depth modeling study was conducted to develop a model capable of describing the entire reductive process of a hematite bed exerted to hydrogen-containing gas. For this purpose, the sensitivity of the model to the variation of several structural parameters was assessed. The acquisition of accurate experimental data using the experimental system make it possible to estimate kinetic parameters of the individual reaction steps, which is the focus of forthcoming work.

#### Acknowledgement

We acknowledge support from Business Finland to the Towards Fossil-free Steel project.

#### References

- M. Atsushi, H. Uemur, & T. Sakaguchi, 2010. MIDREX processes. *Kobelco Technology Review*, 29, 50–57.
- M. Bernasowski, 2014. Theoretical study of the hydrogen influence on iron oxides reduction at the blast furnace process. *Steel Research International*, 85(4), 670–678.
- A. Bonalde, A. Henriquez, & M. Manrique, 2005. Kinetic Analysis of the Iron Oxide Reduction Using Hydrogen-Carbon Monoxide Mixtures as Reducing Agent. *ISIJ Int.*, 45(9), 1255–1260.
- M. Chu, J. Yagi, & F. Shen. (2005). *Modelling on blast furnace process and innovative ironmaking technologies*. Northeastern University Press (Shenyang).
- Energy Agency, I. 2020. *Iron and Steel Technology Roadmap Towards more sustainable steelmaking Part of the Energy Technology Perspectives series*.
- Y. Hara, M. Tsuchiya, & S. Kondo. 1974. Intraparticle Temperature of Iron-Oxide Pellet during the Reduction. *Tetsu to Hagane*. 60, 1261–1270.
- B. Hou, H. Zhang, H. Li, & Q. Zhu, 2012. Study on kinetics of iron oxide reduction by hydrogen. *Chinese Journal of Chemical Engineering*, 20(1), 10–17.
- W. P. Hutny, G. K. Lee, & J. T. Price, 1991. Fundamentals of coal combustion during injection into blast furnace. In *Prog. Energy Combust. Sci*, 17.
- M. Kazemi, M. S. Pour, & D. Sichen, 2017. Experimental and Modeling Study on Reduction of Hematite Pellets by Hydrogen Gas. *Metallurgical and Materials Transactions B: Process Metallurgy and Materials Processing Science*, 48(2), 1114–1122.
- H. Y. Lin, Y. W. Chen, & C. Li, 2003. The mechanism of reduction of iron oxide by hydrogen. *Thermochimica Acta*, 400(1–2), 61–67.
- I. J. Moon, C. H. Rhee, & D. J. Min, 1998. Reduction of hematite compacts by H<sub>2</sub>-CO gas mixtures. *Steel Research*, 69(8), 302–306.
- F. Patisson, O. Mirgaux, & J. P. Birat, 2021. Hydrogen steelmaking. Part 1: Physical chemistry and process metallurgy. *Materiaux et Techniques*, 109(3–4).
- A. Pineau, N. Kanari, & I. Gaballah, 2006. Kinetics of reduction of iron oxides by H<sub>2</sub>. Part I: Low temperature reduction of hematite. *Thermochimica Acta*, 447(1), 89–100.
- K. Piotrowski, K. Mondal, H. Lorethova, L. Stonawski, T. Szymański, & T. Wiltowski, 2005. Effect of gas composition on the kinetics of iron oxide reduction in a hydrogen production process. *International Journal of Hydrogen Energy*, 30(15), 1543–1554.
- R. Takahashi, J. Yagi & Y. Omori, 1971. Reduction Rate of Iron Oxide Pellets with Hydrogen. *Science Reports of the Research Institutes, Tohoku University. Ser. A, Physics, Chemistry and Metallurgy*, 23, 9–30.
- Z. Wei, J. Zhang, B. Qin, Y. Dong, Y. Lu, Y. Li, W. Hao, & Y. Zhang, 2018. Reduction kinetics of hematite ore fines with H<sub>2</sub> in a rotary drum reactor. *Powder Technology*, 332, 18–26.
- W. Zhang, J. Zhang, Q. Li, Y. He, B. Tang, M. Li, Z. Zhang, & Z. Zou, 2013. Thermodynamic Analyses of Iron Oxides Redox Reactions. *The 8th Pacific Rim International Congress on Advanced Materials and Processing*, 777–789.

# Parameter estimation approach applied to microalgae-bacteria photobioreactor

Irina Bausa-Ortiz,<sup>a,b</sup> Raul Muñoz,<sup>a,c</sup> Smaranda P. Cristea,<sup>a,b</sup> Cesar de Prada<sup>a,b</sup>

<sup>a</sup>*Institute of Sustainable Processes, Universidad de Valladolid, Dr. Mergelina s/n, Valladolid 47011, Spain*

<sup>b</sup>*Dpt. of Systems Engineering and Automatic Control, School of Industrial Engineering, Universidad de Valladolid, Prado de la Magdalena 3-5, Valladolid 47011, Spain*

<sup>c</sup>*Dpt. of Chemical Engineering and Environmental Technology, School of Industrial Engineering, Universidad de Valladolid, Dr. Mergelina, s/n, Valladolid 47011, Spain*  
[irina.bausa@uva.es](mailto:irina.bausa@uva.es)

## Abstract

Mechanistic models are commonly used in wastewater treatment systems based on microalgae-bacteria consortia to study microalgae-bacteria interactions and for design and control purposes. Generally, in the proposed mathematical models, the parameters concerning the activity of bacteria and microalgae are obtained through calibration or parameter estimation. Microalgae-bacteria models frequently involve many nonlinear differential equations and parameters, which creates difficulties with the optimization for parameter estimation, resulting in non-optimum convergence in some cases. This work deals with formulating the optimization problem in the parameter estimation context. The goal is to propose an improved approach for parameter estimation when several outputs and parameters are involved in the optimization problem. The proposed approach solves a series of increasing-complexity optimization problems to estimate process parameters gradually, avoiding convergence problems.

**Keywords:** Modeling, Parameter Estimation, Optimization, Wastewater treatment.

## 1. Introduction

Microalgae-based technologies for wastewater treatment have received a renovated interest in the last years in the context of resource recovery and circular economy (Muñoz and Guieysse, 2006). The increasing number of microalgae-based applications has contributed to the development of new models for studying the main processes, factors, and variables affecting microalgae growth in different cultures media, including wastewater (Solimeno, Gómez-Serrano and Ación, 2019), (Casagli *et al.*, 2021). Generally, in microalgae-bacteria models, model parameters are obtained through calibration or parameter estimation (Solimeno, *et al.*, 2017), (Casagli *et al.*, 2021).

The main goal of this study is to propose an improved approach for parameter estimation when several outputs and parameters are involved in the optimization problem. The proposed approach solves a series of increasing-complexity optimization problems to estimate process parameters gradually, avoiding convergence problems. The proposed approach represents an alternative to dealing with large optimization problems in parameter estimation, which are common in these systems. Parameter estimation via dynamic optimization is realized in each optimization step to fit simulated outputs to

experimental data. This methodology has been tested in a photobioreactor for wastewater treatment, in which processes concerning the activity of microalgae and bacteria, the influence of sunlight and nutrients concentration in the growth of microorganisms, chemical equilibrium reactions, and transfer of gases to the atmosphere are involved. Parameter estimation has been performed in a four-step optimization sequence. The first optimization step involves only two system outputs and five parameters. Finally, five model outputs were considered in the optimization problem: Total Suspended Solids (TSS) concentration, dissolved Total Organic Carbon (TOC) concentration, dissolved Inorganic Carbon (IC) concentration, dissolved ammonium concentration, and dissolved oxygen (DO) concentration. Simulation results show that model outputs matched with experimental data. The proposed optimization approach results in a better fit between experimental and simulated data and lower convergence time compared to the whole optimization problem.

## 2. Methodology for parameter estimation

The proposed approach solves a series of increasing-complexity optimization problems to estimate process parameters gradually, avoiding convergence problems. The idea is first to formulate one easier parameter estimation problem involving a subset of system outputs and parameters, replacing the other outputs variables with their experimental values, and then use these estimated values as a starting point for the next step of the optimization problem. The approach is oriented to increase the sub-set dimension until all the system outputs are being included in the optimization problem. The sub-sets selection is performed based on sensitivity analysis. Parameter estimation via dynamic optimization is realized in each optimization step to fit simulated outputs to experimental data.

This methodology has been tested in the photobioreactor for wastewater treatment used in the anoxic–aerobic algal–bacterial photobioreactor configuration with biomass recycling described in (Alcántara *et al.*, 2015). The anoxic reactor model and the settler model of this facility were described in (Bausa *et al.*, 2021, 2022). The data used in this study corresponded with 47 days of experimentation, these data were used for parameter estimation and model validation. In this work, the model BIO\_ALGAE 2 (Solimeno, Gómez-Serrano and Ación, 2019) has been used to describe the photobioreactor. The modeling and simulation of the photobioreactor are coded in dynamic simulation software PROOSIS® (EA International, 2022).

### 2.1. Sensitivity analysis

Previous to parameter estimation, the sensitivity analysis was conducted to identify the parameters with the greatest impact on the model outputs. For this purpose, a subset of the most influential parameters on model outputs was analyzed.

Sensitivity analysis was carried out using PROOSIS®. The results of sensitivity analysis indicated that model outputs are especially sensitive to the maximum specific growth rate of microalgae ( $\mu_{\text{ALG}}$ ) and heterotrophic bacteria ( $\mu_{\text{H}}$ ); the decay rate of microalgae ( $k_{\text{death,ALG}}$ ) and heterotrophic bacteria ( $k_{\text{death,H}}$ ); and the gas-liquid mass transfer coefficients for ammonia ( $K_{\text{la,NH}_3}$ ), oxygen ( $K_{\text{la,O}_2}$ ), and carbon dioxide ( $K_{\text{la,CO}_2}$ ). These decision variables agree with decision variables used in other publications on the topic (Casagli *et al.*, 2021), (Solimeno *et al.*, 2017).

#### 2.1.1. Sensitivity analysis as a guide to sub-sets selection

Results of sensitivity analysis are used here also as a guide to determine the best selection of groups of model outputs to consider. In addition, prior knowledge of system dynamic should be considered.

In the first step of the optimization problem, two model outputs were considered: TSS concentration and DO concentration. The TSS concentration is related to many of the processes of the model and its value depends on the concentration of all particulate components in the photobioreactor. Furthermore, DO concentration is involved in many of the processes concerning to microalgae and bacteria activity and mass transfer processes. Both model outputs depend mainly on the maximum specific growth rate of microalgae ( $\mu_{ALG}$ ), the decay rate of microalgae ( $k_{death,ALG}$ ), and the mass transfer coefficient for oxygen ( $K_{la,O_2}$ ).

Inorganic carbon is included as model output in the second step of the optimization. Microalgae use carbon dioxide as a carbon source for growth. As expected,  $\mu_{ALG}$  is the parameter with the greatest influence on IC concentration. Also, inorganic carbon is highly affected by the gas-liquid mass transfer coefficient for carbon dioxide ( $K_{la,CO_2}$ ).

In the third step, TOC concentration is also considered as model output. In microalgae-bacteria processes, heterotrophic bacteria oxidate the organic matter present in the wastewater, for that reason TOC concentration is mainly affected by the maximum specific growth rate ( $\mu_H$ ) and the decay rate of heterotrophic bacteria ( $k_{death,H}$ ).

Finally, dissolved ammonium concentration is included as model output in the fourth step. Ammonium concentration is mainly affected by  $\mu_{ALG}$ ,  $k_{death,ALG}$ , and the mass transfer coefficient for ammonia ( $K_{la,NH_3}$ ). In order to reduce the size of the vector of parameters, mass transfer coefficients for ammonia and carbon dioxide were expressed as a function of the oxygen one (1) (Casagli *et al.*, 2021).

$$\frac{K_{la,j}}{K_{la,O_2}} = \left(\frac{D_{Sj}}{D_{O_2}}\right)^{0.5} \quad (1)$$

where  $D_{Sj}[\text{m}^2\text{s}^{-1}]$  is the diffusivity coefficient for the gas  $j$ .

## 2.2. Parameter estimation

In this work, the fair function estimator (Huber, 2011) is used as a robust objective function  $J$  in the parameter estimation problem. Unlike Least Square Method, the fair function estimator reduces the effect of outliers, thus making it more robust in nature.

Dynamic optimization problem is represented by (2):

$$\min_{\hat{\theta}} J(\hat{\theta}, \theta) = \sum_{i=1}^n c^2 \left[ \frac{|\varepsilon_j|}{c} - \log \left( 1 + \frac{|\varepsilon_j|}{c} \right) \right] \quad (2)$$

Subject to restrictions (3) and (4)

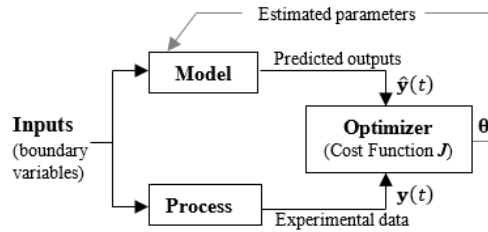
$$\frac{dx(t)}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}, t) \quad (3)$$

$$\hat{\mathbf{y}}(t) = \mathbf{g}(\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}, t) \quad (4)$$

$$\underline{\boldsymbol{\theta}} \leq \hat{\boldsymbol{\theta}} \leq \bar{\boldsymbol{\theta}}$$

Where  $\varepsilon_j = (y_j - \hat{y}_j) / \sigma_j$  represents the error between available process measurements ( $y$ ) and their estimated values ( $\hat{y}$ ) limited between user-defined minimum and maximum values,  $c \in \mathbb{R}^+$  is a user defined fitting parameter to tune the slope for large residues, and  $\sigma$  is the standard deviation of experimental data. The approach to solving a parameter estimation problem in terms of optimization considers that for each value of the vector of parameters  $\boldsymbol{\theta}$  (decision variables) the model provides one prediction of the system output  $\hat{\mathbf{y}}(\boldsymbol{\theta})$  in each experiment. The same sequence of process inputs  $\mathbf{u}(t)$  is applied to the model. For each sample time  $t$ , the prediction error  $\varepsilon_j(t)$  is an indicator of model goodness. Fig. 1 represents the estimation procedure in sequential optimization context.





**Figure 1.** Estimator in sequential optimization

The dynamic optimization problem for the start-up optimization can be converted into a nonlinear programming (NLP) problem via control vector parameterization technique and a proper procedure for computing the cost function. In this work, the SNOPT nonlinear programming algorithm has been used in the PROOSIS® dynamic simulation environment to solve the optimization problem.

Parameter estimation results for each optimization step were provided in Table 1. Initial values for parameter estimation in Step 1 and decision variables ranges for optimization (Table 1) were defined in agreement with previous results reported in the literature.

**Table 1.** Values of estimated parameters in photobioreactor

Parameter	Value				Limits for Optimization
	Step 1	Step 2	Step 3	Step 4	
$\mu_{\text{ALG}}$ [ $\text{d}^{-1}$ ]	1.627	0.990	1.062	1.062	0.4 – 2
$k_{\text{death,ALG}}$ [ $\text{d}^{-1}$ ]	0.101	0.050	0.050	0.050	0.05 – 0.21
$\mu_{\text{H}}$ [ $\text{d}^{-1}$ ]	1.656	1.000	1.210	1.211	1 – 6
$k_{\text{death,H}}$ [ $\text{d}^{-1}$ ]	0.895	0.900	0.900	0.900	0.12 – 0.9
$K_{\text{la, O}_2}$ [ $\text{d}^{-1}$ ]	13.081	4.000	4.000	4.000	4 – 30

### 3. Results

In each step of the optimization problem, data from first 30 days were used for parameter estimation. Validation was performed using data from days 30 to 47. Figures 2 to 6 show simulation results of Step 4.

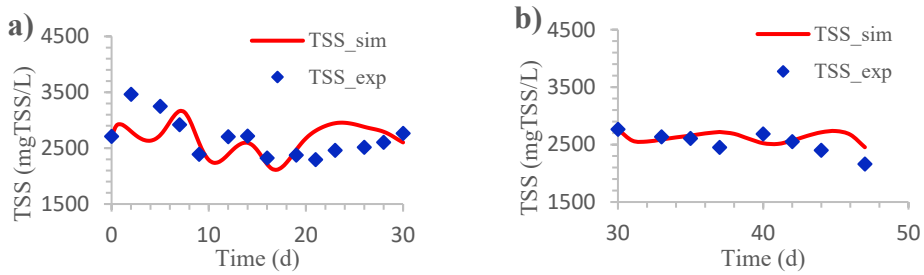
Fig. 2 present simulation results for the concentration of TSS. Fig. 2a) and 2b) represent data set used for parameter estimation and for validation, respectively. Although some discrepancies are observed in transient behavior prediction for TSS concentration, simulated average values are closely with experimental results.

Fig. 3 show simulation results for the dissolved oxygen concentration ( $S_{\text{O}_2}$ ). Fig. 3a) and 3b) represent data set used for parameter estimation and for validation, respectively.

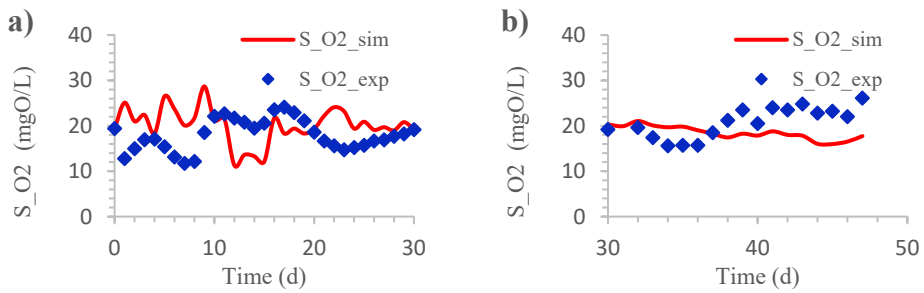
Fig. 4 and Fig. 5 presents simulation results for the Inorganic Carbon concentration and Total Organic Carbon concentration, respectively. Fig. 4a) and 5a) represent data set used for parameter estimation. Fig. 4b) and 5b) represent data used for model validation.

Fig. 6 show simulation results for the dissolved ammonium concentration ( $S_{\text{O}_2}$ ). Fig. 6a) and 6b) represent data set used for parameter estimation and for validation, respectively.

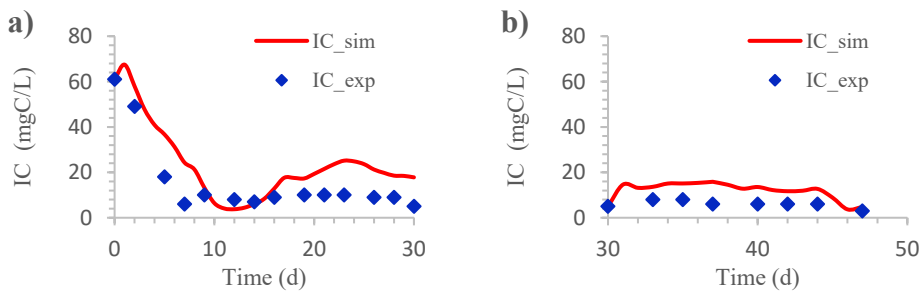
The model was effective in reproducing dynamic behavior of different measured variables.



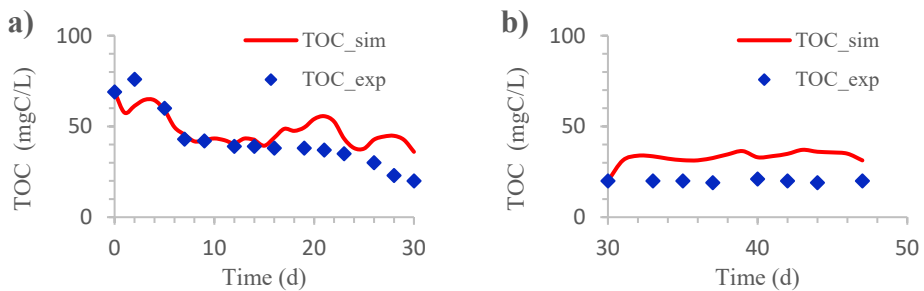
**Fig. 2.** Experimental and simulated Total Suspended Solids concentration



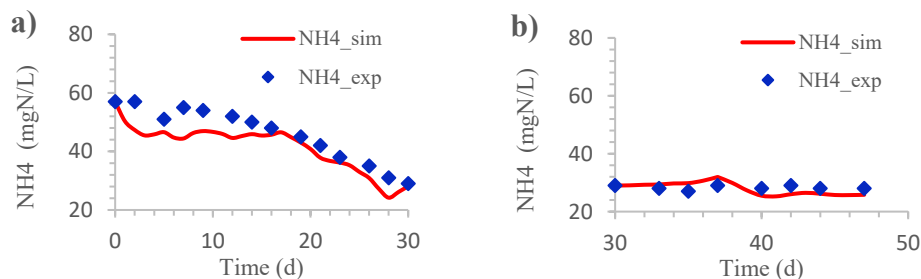
**Fig. 3.** Experimental and simulated dissolved oxygen concentration



**Fig. 4.** Experimental and simulated dissolved Inorganic Carbon concentration



**Fig. 5.** Experimental and simulated dissolved Total Organic Carbon concentration



**Fig. 6.** Experimental and simulated dissolved ammonium concentration

#### 4. Conclusions

Parameter estimation via dynamic optimization was realized in four optimization steps. The proposed optimization approach results in a better fit between the experimental and simulated data, and lower convergence time with respect to the optimization problem statement considering all the system outputs from the beginning. Future work is oriented to design state-observers to estimate non-measurable variables.

#### 5. Acknowledgments

This research was supported by Regional Government of Castilla y León and EU-FEDER program (CLU 2017-09, CL-EI-2021-07, UIC 233, UIC 315), by Regional Government of Castilla y León and the European Social Found (Order EDU/601/2020), and by the Project a-CIDiT (PID2021-123654OB-C31).

#### References

- C. Alcántara *et al.*, 2015, ‘Evaluation of wastewater treatment in a novel anoxic-aerobic algal-bacterial photobioreactor with biomass recycling through carbon and nitrogen mass balances’, *Bioresource Technology*, 191, pp. 173–186. <https://doi.org/10.1016/j.biortech.2015.04.125>.
- I. Bausa *et al.*, 2021, ‘Modelo para la estimación de la concentración de biomasa en una instalación reactor anóxico-fotobiorreactor aerobio de algas y bacterias para el tratamiento de aguas residuales domésticas’, <https://doi.org/10.17979/spudc.9788497498043.427>.
- I. Bausa *et al.*, 2022, ‘Modeling and simulation of anoxic-aerobic algal bacterial photobioreactor for nutrients removal’, *COMPUTER-AIDED CHEMICAL ENGINEERING*, 51. Elsevier, pp. 151–156. <https://doi.org/http://dx.doi.org/10.1016/B978-0-323-95879-0.50026-6>.
- F. Casagli *et al.*, 2021, ‘ALBA: A comprehensive growth model to optimize algae-bacteria wastewater treatment in raceway ponds’, *Water Research*, 190. <https://doi.org/10.1016/j.watres.2020.116734>.
- EA International, 2022, *EcosimPro | PROOSIS - Modelling and Simulation Toolkits and Services*. Available at: <https://www.ecosimpro.com/> (Accessed: 18 July 2022).
- P. J. Huber, 2011, ‘Robust Statistics’, in M. Lovric (ed.) *International Encyclopedia of Statistical Science*. Springer, Berlin, [https://doi.org/https://doi.org/10.1007/978-3-642-04898-2\\_594](https://doi.org/https://doi.org/10.1007/978-3-642-04898-2_594).
- R. Muñoz and B. Guieysse, 2006, ‘Algal-bacterial processes for the treatment of hazardous contaminants: A review’, *Water Research*, 40(15), pp. 2799–2815. <https://doi.org/10.1016/j.watres.2006.06.011>.
- A. Solimeno *et al.*, 2017, ‘Integral microalgae-bacteria model (BIO\_ALGAE): Application to wastewater high-rate algal ponds’, *Science of the Total Environment*, 601–602, pp. 646–657. <https://doi.org/10.1016/j.scitotenv.2017.05.215>.
- A. Solimeno, C. Gómez-Serrano, and F.G. Acién, 2019, ‘BIO\_ALGAE 2: improved model of microalgae and bacteria consortia for wastewater treatment’, *Environmental Science and Pollution Research*, 26(25), pp. 25855–25868. <https://doi.org/10.1007/s11356-019-05824-5>.

# An Event-Based Continuous-Time MILP for Short-Term Crude Oil Scheduling

Irem Martin,<sup>a,b</sup> Sena Kurban,<sup>a</sup> Gizem Kuşoğlu Kaya,<sup>a</sup> Özgür Kabak,<sup>b</sup> Y. İlker Topcu<sup>b</sup>

<sup>a</sup>*Turkish Petroleum Refinery, 41780, Körfez, Kocaeli, Turkey*

<sup>b</sup>*Industrial Engineering Department, Faculty of Management, Istanbul Technical University, 34367, Sarıyer, Istanbul, Turkey*  
[irem.martin@tupras.com.tr](mailto:irem.martin@tupras.com.tr)

## Abstract

In complex real refinery systems, it is essential to find an optimal plan in a short time to maintain operations' efficiency. An effective tool for industrial use in refinery crude oil scheduling is still lacking. The aim of this study is to develop an optimization tool to support efficient decision-making by ensuring the stability of operations and considering economic profit. An event-based continuous-time Mixed Integer Linear Program is developed to handle crude oil operations from vessels to crude distillation units as well as to satisfy product demands. This new formulation describes real-world operational applications of the TUPRAS Refinery.

**Keywords:** Crude oil operations, Short-term scheduling, Mixed-integer linear programming, Optimization.

## 1. Introduction

In the literature, several approaches have been proposed to handle the management of crude oil scheduling. The early approaches proposed by Shah (1996) and Lee et al. (2002) focused on discrete-time representation. However, recently, continuous-time representation has received great attention since it reduces complexity and problem size. Yadav and Shaik (2012) proposed a state-task-network approach using event-based continuous-time representation to manage crude oil operations. In this study, a similar approach is implemented in a complex petroleum refinery. Therefore, the discrete-time approach model, developed by Dologlu et al. (2022), has been modified as a continuous model approach to better reflect the refinery. Unlike the discrete model, many new features have been added in the continuous-time approach model, such as tank rest times, the amount of crude oil remaining in the pipe at the beginning and after the unloading, the unloading of different crude oils from the same vessel, and the decision of crude oil to be selected in case more than one crude oil arrives on the same day in the plan.

## 2. Model Formulation and Problem Definition

Similar to Dologlu et al. (2022), TUPRAS Refinery including one docking station, 12 tanks and 2 Crude Distillation Units (CDUs) whose structure is illustrated in Fig. 1 is studied in this paper. During a given planning horizon, a vessel arrives at the docking station and different types of crude oil are unloaded from vessel into tanks through pipes. Finally, the appropriate crude oil mixtures in the tanks are charged into the CDUs. Before charging operations, the crude oil mixtures in the tanks must meet the quality

requirements. Moreover, about 50 types of crude oil reach and are processed at the refinery. Having too many types of crude oil can lead to operational difficulties for processing in CDUs due to their specific properties. In addition, since the tanks are not dedicated to a single operation and all the unloading, transfer and charging operations can be carried out in all tanks, the scheduling problem becomes more complex.

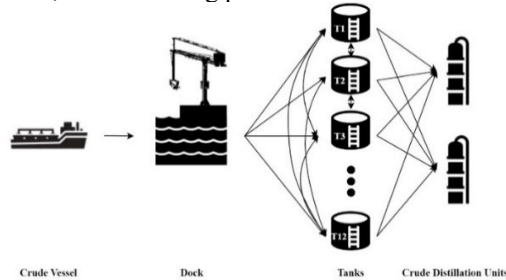


Figure 1. An overview of crude oil operations

The system addressed includes  $i \in I$  crude oil tanks and  $h \in H$  crude oil types. Four main operations of the crude oil system are described here as tasks under  $a$  set, and the four subsets are:  $k(a)$ ,  $ar(a)$  and  $as(a)$  for unloading, transfer and charging tasks, respectively. Crude oils' descriptive properties such as American Petroleum Institute (API) and Sulfur are in the  $p \in P$  set. Over the planning horizon,  $k$  vessels arrive at the system as well as operations take place along a chain of  $n$  events. The set of arriving crude oil types is  $K_h$  and  $T_k$  is the set of tanks to which vessel  $k$  can be unloaded. For simplicity, unloading operations are set to specific events.  $nk'$  and  $nk''$  are the set of unloading events, which is for the main unloading and for the remaining part in the pipe, respectively. The following data is used for scheduling the operations: (1) scheduling time horizon,  $PH$ ; (2) arrival time of vessels,  $KT_{k,t}$  (hr); (3) amount of arriving crude oil vessels,  $KV_k$  ( $m^3$ ); (4) demurrage  $DM_k$  (\$/hr) and lay time of vessels  $LT_k$  (hr); (5) crude oil properties  $PR_{h,p}$ ; (6) Tank properties (min. & max. capacity of tanks,  $LC_{ih}$  and  $UC_{ih}$ ; tank rest times after unloading and transfer,  $R$ ); (7) Initial tank volume and compositions,  $VI_{h,i}$  ( $m^3$ ); (8) bounds on the flowrates of the streams for unloading ( $FU_{ih}$ ), transfer ( $FT_a$ ), and charging ( $FC_{as}$ ). Additional parameters are required for problem formulation.  $B_h$  is the set for crude oils that cannot mix with each other. Besides, since the rest time applies only to the tank that receives a transfer,  $A_{ps}$  and  $A_{ng}$  parameters are used to separate receiving and sending transfer tanks for a transfer operation, respectively. Moreover, there are two other vessel types; the first is the optional crudes that planners must choose one out of two or three different crude oil types. There may be more than one optional crude in a planning period, so they were handled in groups.  $GN_k$  is for the group numbers,  $O_i$  is the set of all optional crudes and  $O_l$  is the different types of crude in group  $l$ . Second is the vessels carrying two or three different crude oils and all of them should be unloaded sequentially.  $S_a$  is the set of these concurrent crudes.  $SF_k$  specifies the ones that are not the last one in the unloading sequence so the remaining pipe volume ( $VP$ ) of them is not going to be planned separately since they will be unloaded one after another and  $by_k$  parameter takes value 1 for them. To introduce the limitations for charging operations, the related parameters are set as: The minimum charging hours for a charging task,  $MH$ . Also, the CDUs have minimum-maximum API and maximum Sulfur limits,  $LA_j$ ,  $UA_j$  and  $US_j$ , respectively. The described problem is formulated in Mixed-Integer Linear Programming (MILP) with event-based continuous time representation with the assumption of perfect mixing in the tanks. The optimum values of the following decision variables are determined to minimize the demurrage cost of vessels: (1) the start and end times of all operations,  $s_{a,n}$  and  $f_{a,n}$ ; (2) the

start and end times of operations in tanks,  $ts_{n,i}$  and  $tf_{n,i}$ ; (3) the start and end times of unloading and transfer operations,  $rs_n$  and  $rf_n$ ; (4) the total volume of tanks,  $vt_{i,n}$ , the volume of components in the tanks,  $v_{i,h,n}$ ; (5) the amount of crude oil that flows from one unit to another in unloading,  $bk_{k,n,i}$ , transfer,  $bh_{ar,n,h}$ , charging,  $bs_{as,n,h,i}$ ; (6) delay time of vessel unloading,  $g_k$ ; (7) total amount of flowing,  $b_{a,n}$ . Binary variables for operational decisions are also defined.  $w_{a,n}$  shows if a task is assigned to an event. It also applies to all of the operations in specific,  $wl_{k,n,i}$ ,  $ws_{as,n,i}$ ,  $wtb_{k,n,i}$  for unloading, charging and unloading of the remaining crude in the pipe, respectively.  $w_{k,n}$  represents whether a vessel is unloaded and  $wb_{k,n}$  is the same for unloading the crude in the pipe. The selection of optional crude oils is represented by  $q_k$ .  $e_{n,i}$  shows whether the volume in the tank is above the minimum capacity, and  $t_{n,i,h}$  shows whether a crude oil type exists in a tank.

The objective function is the minimization of the demurrage cost of vessels, Eq. (1). Tank components balance, total tank volume and capacity limits are given in Eqs. (2-4), respectively. The calculation of processed volumes at each event for transfer and charging tasks are handled in Eqs. (5a-b) and for unloading is handled in two different equations for the pipe part and the main part of the crude, Eqs. (6a-b). Eqs (7a-c) guarantee that if any flow occurs at any event, then the binaries take value 1 where M refers to a big number. In Eqs. (8a-10c), timing and sequencing rules are handled for each of the tasks. For continuous operation of CDUs, it is ensured that charging starts from the beginning and total operation time of each CDU should be equal to the time horizon  $PH$ , Eqs. (11a-b). Besides, since transfer operations are not preferred, they are eliminated in Eq. (12).

$$Cost = \sum_k g_k * DM_k \quad (1)$$

$$v_{i,h,n} = v_{i,h,n-1} + \sum_k bk_{k,n,i} + \sum_k wtb_{k,n,i} * VP + \sum_{ar \in A_{ps}} bh_{ar,n,h} - \sum_{ar \in A_{ng}} bh_{ar,n,h} - \sum_{as} bs_{as,n,h,i} \quad \forall n \geq 2, n'_k, n''_k, h \in K_h, i \in T_k \quad (2)$$

$$vt_{i,n} = \sum_h v_{i,h,n} \quad \forall n, i \quad (3)$$

$$LC_i \leq vt_{i,n} \leq UC_i \quad \forall n, i \quad (4)$$

$$FT_{ar} * (f_{ar,n} - s_{ar,n}) = b_{ar,n} \quad \forall ar, n \quad FC_{as} * (f_{as,n} - s_{as,n}) = b_{as,n} \quad \forall as, n \quad (5a,5b)$$

$$f_{k,n} - s_{k,n} = \sum_i wtb_{k,n,i} * \frac{VP}{FU_i} \quad \forall k, n'_k \quad f_{k,n} - s_{k,n} = \sum_{i \in T_k} \frac{bk_{k,n,i}}{FU_i} \quad \forall k, n'_k \quad (6a,6b)$$

$$b_{ar,n} \leq M * w_{ar,n} \quad \forall ar, n \quad b_{as,n} \leq M * w_{as,n} \quad \forall as, n \quad b_{k,n} \leq M * w_{k,n} \quad \forall k, n'_k, n''_k \quad (7a,7b,7c)$$

$$f_{k,n} = s_{k,n} \quad \forall k, n \quad f_{k,n} \geq s_{k,n} \quad \forall k, n'_k, n''_k \quad s_{k,n} \geq f_{k,n-1} \quad \forall k, n \quad (8a,8b,8c)$$

$$f_{ar,n} \leq s_{ar,n} + 60 * 24 * w_{ar,n} \quad \forall ar, n \quad f_{ar,n} \geq s_{ar,n} \quad \forall ar, n \quad s_{ar,n} \geq f_{ar,n-1} \quad \forall ar, n \quad (9a,9b,9c)$$

$$f_{as,n} \leq s_{as,n} + 60 * 24 * w_{as,n} \quad \forall as, n \quad f_{as,n} \geq s_{as,n} \quad \forall as, n \quad s_{as,n} = f_{as,n-1} \quad \forall n > 1, as \quad (10a,10b,10c)$$

$$\sum_n b_{as,n} \geq FC_{as} * PH * 24 \quad \forall as \quad s_{as,n} = 0 \quad \forall n = 1, as \quad (11a,11b)$$

$$w_{ar,n} = 0 \quad \forall ar, n \quad (12)$$

Eq. (13) ensures that all the volume of vessel should be unloaded except the part in the pipe. The main unloading and pipe unloading for each of the vessels are assigned to an event in Eqs. (14a-b). Eqs. (15a-b) state the same condition for optional crudes for the one that is chosen. If there is any unloading in an event, then the binary related to the tank involving also takes the value 1 (Eqs. (16a-b)). Eq. (17) guarantees that one of the crude oils must be selected for each of the optional crude oil vessel groups. Unloading flow boundaries are in Eqs. (18a-b). The unloading can start only after the arrival of vessel and same for optional crudes (Eqs. (19a-b)). In Eqs. (20a-b), the unloading delay time subject

to demurrage cost is calculated for both types of vessels. Eq. (21) ensures that unloading of remaining crude from the previous vessel should start after the next vessel's arriving. After the pipe crude, the main part of the next vessel should be unloaded immediately (Eq. (22)). Vessels are unloaded sequentially and the concurrent crudes in a vessel are unloaded repeatedly, Eqs. (23-24). Eqs. (25a-b) state that the crude in the pipe should be unloaded to the set of tanks  $T_k$  in which the next vessel can be unloaded. Lastly, Eqs. (26a-28b) are about sequencing unloading and transfer operations.

$$\sum_{i \in T_k} bk_{k,n,i} = (KV_k - (VP * (1 - by_k))) * w_{k,n} \quad \forall k, n'_k \quad (13)$$

$$wb_{k,n} = 1 \quad \forall k \notin O_t, n'_k \quad wk_{k,n} = 1 \quad \forall k \notin O_t, n'_k \quad (14a,14b)$$

$$wb_{k,n} = q_k \quad \forall k \in O_t, n'_k \quad wk_{k,n} = q_k \quad \forall k \in O_t, n'_k \quad (15a,15b)$$

$$\sum_{i \in T_k} wt_{k,n,i} \leq w_{k,n} * I \quad \forall k, n'_k \quad \sum_i wtb_{k,n,i} = wb_{k,n} \quad \forall k, n'_k \quad (16a,16b)$$

$$\sum_{k \in O_l} q_k = 1 \quad \forall l \quad (17)$$

$$bk_{k,n,i} \leq M * wt_{k,n,i} \quad \forall k, n'_k, i \in T_k \quad bk_{k,n,i} \geq VP * wt_{k,n,i} \quad \forall k, n'_k, i \in T_k \quad (18a,18b)$$

$$s_{k,n} \geq KT_k \quad \forall k \notin O_t, n'_k \quad s_{k,n} \geq KT_k * q_k \quad \forall k \in O_t, n'_k \quad (19a,19b)$$

$$g_k \geq f_{k,n} - KT_k - LT_k \quad \forall k \notin O_t, n'_k \quad g_k + 24 * PH * (1 - q_k) \geq f_{k,n} - KT_k - LT_k \quad \forall k \in O_t, n'_k \quad (20a,20b)$$

$$s_{k,n} + GN_k * 24 * PH * (1 - q_k) \geq KT_{k+1} - GN_{k+1} * 24 * PH * (1 - q_{k+1}) \quad \forall k \in K_b, n'_k \quad (21)$$

$$f_{k,n} + GN_k * 24 * PH * (1 - q_k) \geq s_{k+1,n+1} - GN_{k+1} * 24 * PH * (1 - q_{k+1}) \quad \forall k \in K_b, n'_k \quad (22)$$

$$s_{k,n} - GN_k * 24 * PH * (1 - q_k) \leq rf_{n-1} \quad \forall k, n'_k \quad (23)$$

$$s_{k+1,n+1} \leq f_{k,n} \quad \forall k \in S_a, n'_k \quad (24)$$

$$wtb_{k,n,i} = 0 \quad \forall k \notin O_t, n'_k, i \notin T_k \quad wtb_{k,n,i} \leq (1 - q_{k+1}) \quad \forall k \in O_t, n'_k, i \notin T_k \quad (25a,25b)$$

$$rs_n \leq s_{ar,n} \quad \forall ar, n \quad rf_n \geq f_{ar,n} \quad \forall ar, n \quad (26a,26b)$$

$$rs_n \leq s_{k,n} \quad \forall k, n \quad rf_n \geq f_{k,n} \quad \forall k, n \quad (27a,27b)$$

$$rs_n \leq rf_n \quad \forall n \quad rs_n \geq rf_{n-1} \quad \forall n \quad (28a,28b)$$

Eqs. (29a, 30b) relates tank and crude oil-based flows to event-based flow amounts. Tank times for operations should be associated with task times and are handled in Eqs. (31a-b) for charging; in Eqs. (32-36) for unloading; in Eqs. (37-38b) for transfer operations considering the tank rest times. Events in the tanks should be sequenced, Eqs. (39a, 39b). Eq. (40) ensures no simultaneous input-output at any event. Simultaneous unloading and transfer operations are also not allowed due to using the same line, Eq. (41).

$$\sum_{i \in T_k} bk_{k,n,i} = b_{k,n} \quad \forall k, n'_k \quad \sum_i wtb_{k,n,i} * VP = b_{k,n} \quad \forall k, n'_k \quad (29a,29b)$$

$$\sum_h bh_{ar,n,h} = b_{ar,n} \quad \forall ar, n \quad \sum_i \sum_h bs_{as,n,h,i} = b_{as,n} \quad \forall as, n \quad (30a,30b)$$

$$ts_{n,i} \leq s_{as,n} + 24 * 60 * (1 - ws_{as,n,i}) \quad \forall as, n, i \quad tf_{n,i} \geq f_{as,n} - 24 * 60 * (1 - ws_{as,n,i}) \quad \forall as, n, i \quad (31a,31b)$$

$$ts_{n,i} \leq s_{k,n} + 24 * 60 * (1 - wt_{k,n,i}) \quad \forall k, n'_k, i \in T_k \quad (32)$$

$$tf_{n,i} \geq f_{k,n} + R_k * (1 - SF_k) - 24 * PH * (1 - wt_{k,n,i}) \quad \forall k, n'_k, i \in T_k \quad (33)$$

$$ts_{n,i} \leq s_{k,n} + 24 * 60 * (1 - wtb_{k,n,i}) \quad \forall k, n'_k, i \quad (34)$$

$$tf_{n,i} \geq f_{k,n} + R_k * (1 - wt_{k+1,n+1,i}) - 24 * PH * (1 - wtb_{k,n,i}) \quad \forall k \notin O_t, n'_k, i \in T_k \quad (35)$$

$$tf_{n,i} \geq f_{k,n} + R_k * (1 - wt_{k+1,n+1,i}) - 24 * PH * (1 - wtb_{k,n,i}) - 24 * PH * (1 - q_{k+1}) \quad \forall k \in O_t, n'_k, i \in T_k \quad (36)$$

$$ts_{n,i} \leq s_{ar,n} + 24 * 60 * (1 - w_{ar,n}) \quad \forall n, i, ar \quad (37)$$

$$tf_{n,i} \geq f_{ar,n} + R_{ar} - 24 * 60 * (1 - w_{ar,n}) \quad \forall n, i, ar \in A_{ps} \quad tf_{n,i} \geq f_{ar,n} - 24 * 60 * (1 - w_{ar,n}) \quad \forall n, i, ar \in A_{ng} \quad (38a,38b)$$

$$ts_{n,i} \leq tf_{n,i} \forall n, i \qquad \qquad \qquad ts_{n,i} \geq tf_{n-1,i} \forall n, i \qquad (39a,39b)$$

$$\sum_k wt_{k,n,i} + \sum_k wtb_{k,n,i} + \sum_{ar \in A_{ps}} w_{ar,n} + 0.5 * \sum_{ar \in A_{ng}} w_{ar,n} + 0.25 * \sum_{as} ws_{as,n,i} \leq 1 \quad \forall n'_k, n''_k, i \in T_k \qquad (40)$$

$$w_{ar,n} = 0 \quad \forall n'_k, n''_k, ar \qquad (41)$$

If a tank's volume is above the minimum capacity,  $e_{n,i}$  takes value 1 meaning that there is a significant amount of crude that is to be considered for mixing rules, Eq (42). Mixing constraint is handled in Eq. (43). Operational rules for charging are considered as follows: The binding of related variables is given in Eq. (44). Minimum hours of charging for a charging task is handled in Eq. (45). API and Sulfur limitations are given in Eqs. (46-47).

$$e_{n,i} * UC_i \geq vt_{i,n-1} - LC_i \qquad \forall n > 2, i \qquad (42)$$

$$t_{n,i,h} + wt_{k,n,i} \leq (1 - e_{n,i}) + 1 \qquad \forall n'_k \geq 2, i, h \in B_h, k \in K_h \qquad (43)$$

$$\sum_h bs_{as,n,h,i} \leq ws_{as,n,i} * M \quad \forall as, n, i \qquad (44)$$

$$b_{as,n} \geq w_{as,n} * FC_{as} * MH \quad \forall n, as \qquad (45)$$

$$LA_{as} * \sum_i \sum_h bs_{as,n,h,i} \leq \sum_i \sum_h bs_{as,n,h,i} * p_{h,"API"} \leq UA_{as} * \sum_i \sum_h bs_{as,n,h,i} \quad \forall n, as \qquad (46)$$

$$\sum_i \sum_h bs_{as,n,h,i} * p_{h,"Sulfur"} \leq US_{as} * \sum_i \sum_h bs_{as,n,h,i} \quad \forall n, as \qquad (47)$$

### 3. Results and Discussions

The developed continuous-time model is applied for 14-days planning horizon that contains 5 vessels shipped 7 different types of crude oil, 2 CDUs, 12 tanks (2 tanks are in maintenance), including 13 crude oil types. The volume of vessels varies between 40,000 and 175,000 m<sup>3</sup> and the total volume of 5 vessels is about 535,000 m<sup>3</sup>. The MILP model results are shown in Fig. 2 for the first 10 days of the planning horizon.

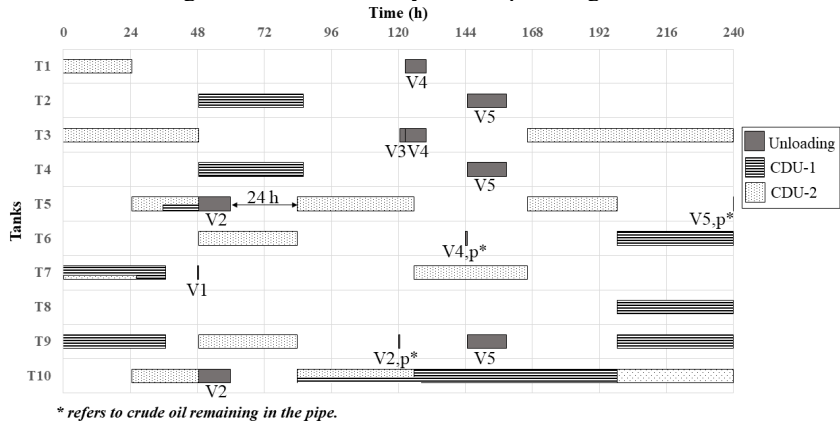


Figure 2. Gantt chart of optimal solution for first 10 days of planning period

Unlike the discrete model developed by Dologlu et al. (2022), in this model, a continuous charge flow is provided to the units. There is a 24-hour rest time in the tanks after unloading. For example, T5 was charged to CDU 24 hours after V2 unload to T5 was completed. The unloading of crude oil remaining in the pipe can be monitored. For example, V1 is the crude oil remaining in the pipe at the start of planning. Immediately after, V2 was unloaded. V3 and V4 are simultaneous incoming vessels and their unloading to the T3 tank appears to be sequentially scheduled. V2 vessel carries 2



different types of optional crude oil. The model made the decision of crude oil in line with the tank occupancy and contents, so as not leading to any demurrage cost.

The main motivation of this study is to compare the computational performance and behavior of models with different time approaches, such as discrete and continuous time, for the real refinery problem. To test the performance of discrete and continuous time models, four cases with different time horizons, 8-days, 14-days, 19-days, and 26-days were studied. All MILP models are executed in GAMS v 28.2.0 on an Intel(R) Xeon(R) Gold 6252 CPU 2.10 GHz machine with 12 GB of RAM using the CPLEX solver. In all cases, the objective function of minimizing the demurrage costs was used. The overall solutions of the four cases are presented in Table 1. When the results of the continuous time model are compared to the discrete time model, the number of constraints, variables, and binary variables is higher, so the model has a longer CPU time. This is because the constraints such as the rest time of the tanks and the evaluation of different unloading states, which were previously ignored in the discrete model, were added to the model to reflect the real refinery case, and the model became more complex. Thus, it is an expected result that the variables, constraints, and CPU time increase in the transition to the continuous model. In addition, as the planning period increased, the number of constraints and variables, and the CPU time increased accordingly for each model time approach.

**Table 1.** Comparison of discrete and continuous time model results

Scenario	The discrete-time approach model				The continuous-time approach model			
	1	2	3	4	5	6	7	8
Planning time horizon (day)	8	14	19	26	8	14	19	26
Number of constraints	781	1594	2382	3633	10977	16008	21101	23690
Number of variables	1440	2796	4046	5920	5178	7541	9960	11191
Number of binary variables	197	350	509	811	926	1306	1696	1888
CPU time (s)	0.28	0.50	2.28	5.98	10.03	149.36	551.80	1567.59

#### 4. Conclusion

In this paper, an event-based continuous time formulation is proposed to address the real-life problem of scheduling of crude-oil operations. Compared to the discrete time formulation, the continuous time formulation provides better production schedules as it reflects the operations of refinery with the added constraints and the continuous time setup. The computation time of the continuous time model increases as more event points are needed to solve refinery-scheduling problems over longer time horizons. Future research will focus on case-specific operational constraints and solving the non-linear problem when the tank contents are uniformly charged to the CDUs as a homogeneous mixture as well as accelerating the model solving time.

#### References

- N. Shah, 1996, Mathematical programming techniques for crude oil scheduling, *Computers & Chemical Engineering*, 20, 1227-1232.
- P. Dologlu, S. Kurban, İ. Marttin, N. Ataman, G.K. Kaya, F. İşeri, Ö. Kabak, Y. I. Topcu, 2022, A new MILP formulation for scheduling of crude oil operations, In *Computer Aided Chemical Engineering*, 51, 991-996.
- S. Yadav, M.A. Shaik, 2012, Short-term scheduling of refinery crude oil operations, *Industrial & Engineering Chemistry Research*, 51(27), 9287-9299.
- Y. H. Lee, S. H. Kim, 2002, Production-distribution planning in supply chain considering capacity constraints, *Computers & Industrial Engineering*, 43(1-2), 169-190.

# Modeling aqueous multi-electrolyte systems for the simulation and optimization of concentrated brine processes

Wajeha Tauqir<sup>a</sup>, Adaeze C. Maduako<sup>a</sup>, Xi Yang<sup>a</sup>, George M. Bollas<sup>a</sup>

<sup>a</sup>*Department of Chemical and Biomolecular Engineering, UTC Institute for Advanced Systems Engineering, University of Connecticut, 159 Discovery Dr, Storrs, CT, 06269, USA. ([george.bollas@uconn.edu](mailto:george.bollas@uconn.edu))*

## Abstract

An advancement of the refined e-NRTL model (Bollas et al., 2008) is used for the modeling of osmotic and activity coefficients associated with aqueous multi-electrolyte systems, with an emphasis on electrolyte systems relevant to concentrated brine solutions. Binary interaction parameters associated with the ions in each electrolyte system are estimated from available activity and osmotic coefficient experimental data. Electrolytes studied are those common in brine solutions (e.g., Na<sup>+</sup>, K<sup>+</sup>, Li<sup>+</sup>, Mg<sup>2+</sup>, Ca<sup>2+</sup>, Cl<sup>-</sup>, SO<sub>4</sub><sup>2-</sup>). Integration of the thermodynamic model in the software package IDAES (Institute for the Design of Advanced Energy Systems), as a thermodynamic property module are presented, and shown to predict various thermodynamic properties (such as activity and osmotic coefficient, and vapor pressure) at higher molalities and temperature ranges. Accurate prediction of solution thermodynamics in brines, then, targets the accurate modeling of brine water treatment systems.

**Keywords:** Concentrated brines, electrolyte thermodynamics, water treatment

## 1. Introduction

Brine solutions are encountered in many natural and industrial processes. Because of their detrimental effects, these highly concentrated salt solutions are rendered dangerous to be disposed of directly to the environment. High salinity in liquid streams is also the major cause of scaling in pipes and equipment. Thus, there is a need to model and develop energy-intensive desalination technologies to increase the reuse of saline waters. Exploration of novel and advanced brine management options is challenging because of the lack of steady state and dynamic simulators that can accurately capture the thermodynamics and kinetics that dominate separation efficiency and processing options. Uncertainty and the high cost associated with the deployment of these systems on a larger scale make modeling a necessity before expensive piloting. The challenge in designing brine wastewater treatment technologies arise because of the high concentration of ions, which when in mixtures are difficult to predict in terms of their impact on the macroscopic properties of the solution. Comprehensive and precise thermodynamic models need to be developed that can accurately predict various thermodynamic properties associated with each ion in the system, over a wide range of molality and temperature ranges.

Electrolyte mixtures are non-ideal even at very low temperature and pressure. To completely model the non-ideality of electrolyte solutions, long-range electrostatic interactions between ions and short-range ion dipole and non-electrostatic interactions should be taken into account. One of the most important thermodynamic properties, which not only quantifies the non-ideality of solutions involving electrolytes but also predicts other properties, is the mean ionic activity coefficient. Thus, it is fundamental to

have models that can accurately predict activity and osmotic coefficients, which can, then, be generalized to excess Gibbs free energy models and predict other solution properties, like enthalpy of mixing and heat capacity. Various activity coefficient models have been used to predict the non-ideal behavior of solutions. The most common of these are the e-NRTL model and its many extensions. The e-NRTL model was first developed for single salt, single solvent solutions. The Pitzer-Debye-Hückel equation was used to model the long-range interactions, while short-range interactions were modeled based on non-random two-liquid, local composition theory. Later, Chen & Evans (1986) modified the model for multi electrolyte systems. The thermodynamics of multi-electrolyte solutions differ largely from single electrolyte systems. Additional terms need to be considered when modeling non-ideality in multi-electrolyte aqueous systems, to cater to increased ions interaction. Besides water treatment, accurate electrolyte models are needed to simulate processes like geothermal energy production, hydraulic fracturing, thermochemical hydrogen production, and pharmaceutical processes (Chen, 2006).

Major ions responsible for salinity and scaling in aqueous solutions include:  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Li}^+$ ,  $\text{Ca}^{2+}$ ,  $\text{Sr}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Ba}^{2+}$ ,  $\text{Cl}^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{HCO}_3^-$ , and  $\text{CO}_3^{2-}$ . Studies have employed extensions and refinements of the e-NRTL model to binary, ternary and quaternary systems involving these ions. (Tanveer and Chen, 2016) developed a thermodynamic model for the aqueous  $\text{Ca}^{2+}$ ,  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Cl}^-$  quaternary system based on the symmetric e-NRTL model. Honarparvar et al. (2017) studied ternary and quaternary systems involving  $\text{Ba}^{2+}$ ,  $\text{Na}^+$ ,  $\text{Cl}^-$ ,  $\text{SO}_4^{2-}$  using an e-NRTL-based model. The e-NRTL model was also studied for the system containing  $\text{Sr}^{2+}$ ,  $\text{Na}^+$ ,  $\text{Cl}^-$ , and  $\text{SO}_4^{2-}$  ions by Honarparvar et al. (2018). Bhattacharia et al. (2015) performed thermodynamic modeling of the  $\text{KCl} + \text{H}_2\text{O}$  binary system, and the  $\text{KCl} + \text{NaCl} + \text{H}_2\text{O}$  ternary system using the e-NRTL model. In this paper, the refined e-NRTL model developed by Bollas et al. (2008) for multi-electrolyte solutions and improved later by Yang et al. (2023) for mixed solvent solutions is applied on 11 aqueous binary systems, with electrolyte concentrations up to their solubility limit, at 298.15 K. The electrolytes studied include three 1-1 electrolytes;  $\text{NaCl}$ ,  $\text{KCl}$ ,  $\text{LiCl}$ , four 2-1 electrolytes;  $\text{MgCl}_2$ ,  $\text{SrCl}_2$ ,  $\text{CaCl}_2$ ,  $\text{BaCl}_2$ ; three 2:1 electrolytes;  $\text{K}_2\text{SO}_4$ ,  $\text{Li}_2\text{SO}_4$ , and  $\text{Na}_2\text{SO}_4$ , and one 2-2 electrolyte,  $\text{MgSO}_4$ . Binary interaction parameters for each aqueous electrolyte were estimated from experimental data on activity and osmotic coefficients. The model predicts accurately the liquid phase non-idealities and through mean ionic activity and osmotic coefficients, providing sound thermodynamic basis for the modeling of process systems that involve brine solutions.

## 2. Thermodynamic Modeling Framework

The refined e-NRTL model used in this work is explained in detail in (Bollas et al. (2008) and Yang et al. (2023)). It expresses the excess Gibbs free energy of the electrolyte solution as a summation of long-range electrostatic contributions, accounting for the ion-ion interactions, and short-range interactions, accounting for local, ion-ion, ion-molecule and molecule-molecule interactions. The model of Yang et al. (2023) describes short-range interactions with a classic NRTL model, but it uses a thermodynamically consistent conversion of the Debye-Hückel equation to model long-range, electrostatic interactions. The short-range model requires two binary interaction parameters  $\tau_{ij}$  and  $\tau_{ji}$ , where  $\tau_{ij} \neq \tau_{ji}$ , and the non-randomness parameter  $\alpha$ . Classically, the value of  $\alpha$  is fixed to 0.2. Binary interaction parameters refer to water-electrolyte and electrolyte-electrolyte pairs, in the case of binary systems. In the case of ternary systems, two electrolytes sharing a common ion in an aqueous system will require six binary interaction parameters; four corresponding to the individual electrolyte-water pairs and two for the electrolyte-electrolyte asymmetric pairs. Other parameters that are treated as adjustable in the model

include the distance of closest approach of the original Debye-Hückel theory, and ionic hydration. Hydration is included to describe its effect on the local composition during the short-range equation formulation. To do this, the mole fractions in the e-NRTL model are replaced by the effective mole fractions of the hydrated solution. To predict the effect of water structure on its neighborhood, hydration numbers are allowed to have negative values. The model with constant hydration numbers can predict mean ionic activity and osmotic coefficients with accuracy at low concentrations. To consider the effect of decreased water activity at increased molality, the stepwise hydration equilibria theory was employed, which accounts for the decreased availability of water molecules as the solution ionic strength increases.

### 3. Methodology

#### 3.1. Thermodynamic Modeling of Binary Systems

The refined e-NRTL model was first implemented to model 11 aqueous binary systems which included two anions ( $\text{Cl}^-$ , and  $\text{SO}_4^{2-}$ ) and seven cations ( $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Li}^+$ ,  $\text{Ca}^{2+}$ ,  $\text{Sr}^{2+}$ ,  $\text{Mg}^{2+}$ , and  $\text{Ba}^{2+}$ ). The model was simultaneously fitted against experimental data from Robinson (1959) of activity and osmotic coefficient with their molality ranging up to the solubility limit of each solution studied, at 298.15 K and 1 atm. The objective function in the non-linear regression algorithm included both the activity and osmotic coefficients to improve model precision and fidelity. Pure component parameters including ionic radii, hydration indices, and effective molar volume were taken at the values estimated with the data of electrolyte solution data reported in Yang et al. (2023). The Average Absolute Relative Deviation (AARD) was calculated for each system.

Table 1: AARD of original e-NRTL, refined e-NRTL with constant hydration and refined e-NRTL model with stepwise hydration equilibria for binary systems relevant to brines.

Electrolyte	Original e-NRTL Model		Refined e-NRTL Constant Hydration Model		Refined e-NRTL Stepwise Hydration Model	
	AARD% $\gamma_{\pm}^m$	AARD% $\Phi$	AARD% $\gamma_{\pm}^m$	AARD% $\Phi$	AARD% $\gamma_{\pm}^m$	AARD% $\Phi$
<b>1-1</b>						
NaCl	1.551	2.088	0.229	0.149	0.047	0.034
KCl	0.808	1.788	0.078	0.071	0.034	0.021
LiCl	18.4931	9.448	4.740	2.109	1.355	0.454
<b>2-1</b>						
MgCl <sub>2</sub>	14.150	7.083	1.565	0.566	0.804	0.343
SrCl <sub>2</sub>	6.598	3.722	0.841	0.559	0.281	0.216
CaCl <sub>2</sub>	18.132	8.421	4.725	2.141	0.806	0.387
BaCl <sub>2</sub>	1.618	1.157	0.177	0.092	0.328	0.316
<b>1-2</b>						
K <sub>2</sub> SO <sub>4</sub>	0.842	1.598	0.53259	0.641	0.53285	0.564
Li <sub>2</sub> SO <sub>4</sub>	2.105	1.955	0.311	0.265	0.208	0.221
Na <sub>2</sub> SO <sub>4</sub>	2.871	3.281	0.419	0.477	0.171	0.209
<b>2-2</b>						
MgSO <sub>4</sub>	3.496	10.836	2.431	5.344	2.714	5.064

#### 3.2. Thermodynamic Modeling of Ternary Systems

For the thermodynamic modeling of ternary systems, involving two electrolytes with a common ion in an aqueous solution, six binary interaction parameters are required, four corresponding to the binary systems involved and two corresponding to the electrolyte-electrolyte pair interactions. The binary interaction parameters corresponding to

electrolyte-water pair are fixed to the values fitted by the binary systems model. The electrolyte-electrolyte interaction parameters are, of course, unknown to the model, as these data have not been used in the fitting of binary systems. Commonly, it is reported that these parameters should be of negligible contribution to the model, especially for dilute solutions. However, this is not the case in highly concentrated brines. The electrolyte-electrolyte relevant binary parameters are, thus, fitted against available experimental data of activity and osmotic coefficients for the ternary system. As a first step, the constant hydration model was deployed. The stepwise hydration model will need to identify the affinity of individual ions of the same charge sign to hydration, which is the topic of on-going work. The hydration indices and the corresponding ionic volumes approximated using the binary system model are fixed (to the values obtain from the study of binary systems) during the fitting of parameters of ternary systems.

#### 4. Results

The AARD values of the original e-NRTL model, the refined e-NRTL model with constant hydration and the refined e-NRTL model with stepwise hydration equilibria for all the binary systems studied are provided in Table 1. The corrected Akaike Information Criterion ( $AIC_c$ ) and the Bayesian Information Criterion (BIC) were calculated for both the refined e-NRTL model and the original e-NRTL model. Both information metrics assess the quality of fit of each model and the potential of overparameterization. Lower values of  $AIC_c$  and BIC in the case of the refined e-NRTL model (5313.74 and 7320.13, respectively) compared to those of the e-NRTL model (20080.80 and 21714.84, respectively) indicate a good balance of goodness of fit. The large F-ratio also signifies a statistically good model. To visualize the accuracy improvement with the refined e-NRTL model, as well as the accuracy improvement of the two hypotheses of the impact and mechanism of hydration, the performance of the models is illustrated for three binary systems shown in Figs. 1-3. These correspond to activity and osmotic coefficients of represented 1-1, 1-2, 2-1 binary electrolyte systems, compared against measurements.

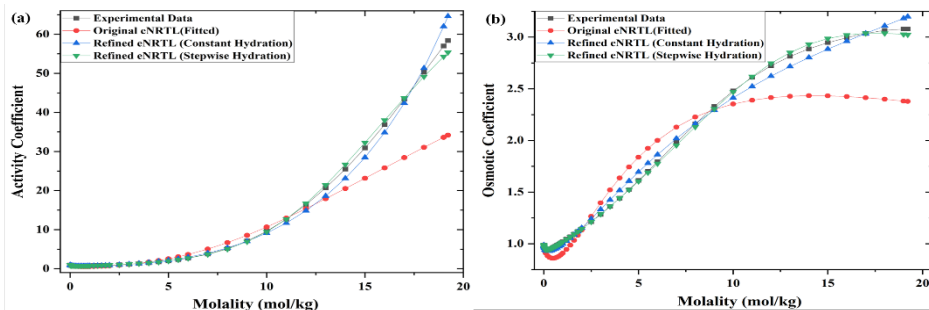


Figure 1: Comparison of the original e-NRTL, refined e-NRTL constant hydration, and refined eNRTL stepwise hydration model for the prediction of (a) activity coefficients of LiCl and (b) osmotic coefficient plot of LiCl.

Fig. 1 presents activity and osmotic coefficients of the LiCl + H<sub>2</sub>O binary system. For these results, the binary parameters of the original e-NRTL model were fitted with the same algorithm used in all other studies, with an objective function that incorporates both activity and osmotic coefficient data. It can be clearly seen that at lower molalities, activity coefficient predictions from all three models are similar and quite accurate. At molalities higher than of 5 mol/kg the stepwise hydration model shows clear superiority.

To further demonstrate the performance of the refined e-NRTL with constant and stepwise hydration assumptions, the Na<sub>2</sub>SO<sub>4</sub> + H<sub>2</sub>O and MgCl<sub>2</sub> + H<sub>2</sub>O binary systems are plotted in Figs. 2 and 3. For the case of Na<sub>2</sub>SO<sub>4</sub> + H<sub>2</sub>O, it can be observed that as the

concentration increases beyond 1 mol/kg, the difference between experimental data and the original e-NRTL model becomes significant, especially in the prediction of osmotic coefficients. However, the results of both the constant and stepwise hydration equilibria assumptions with the refined e-NRTL model are accurate. Activity and osmotic coefficient plots for the  $\text{MgCl}_2 + \text{H}_2\text{O}$  binary system predicted using all the three models are shown in Fig. 3. It can be observed from Fig. 3 that up to molality of 2 mol/kg activity and osmotic coefficients are predicted fairly well by all three models, but beyond that molality range, the stepwise hydration model fits the data best.

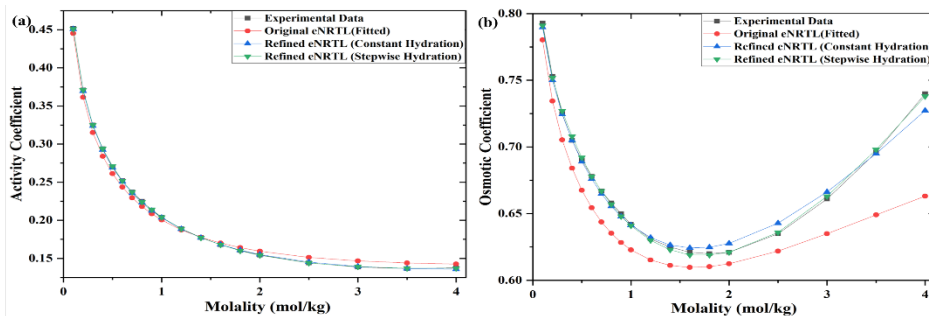


Figure 2: Comparison of the original e-NRTL, refined e-NRTL constant hydration, and refined e-NRTL stepwise hydration model for the prediction of (a) activity coefficient and (b) osmotic coefficient of  $\text{Na}_2\text{SO}_4$ .

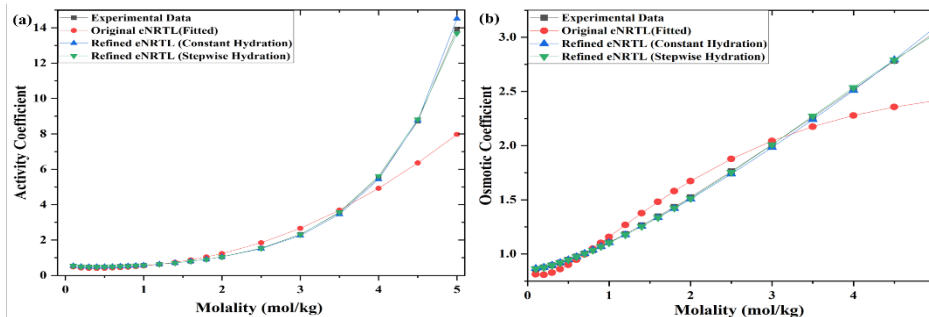


Figure 3: Comparison of the original e-NRTL, refined e-NRTL constant hydration, and refined eNRTL stepwise hydration model for the prediction of (a) activity coefficients and (b) osmotic coefficients of  $\text{MgCl}_2$ .

The ternary  $\text{NaCl-KCl-H}_2\text{O}$  system was studied as an example of ternary systems. Binary interaction parameters for the  $\text{NaCl-H}_2\text{O}$  and  $\text{KCl-H}_2\text{O}$  pairs remain the same for ternary system. The electrolyte-electrolyte binary interaction pairs were regressed against experimental osmotic coefficient data from Dinane et al. (2002). The binary interaction parameters for the  $\text{NaCl-KCl}$  aqueous system are summarized in Table 2. In Fig. 4, the corresponding predicted osmotic coefficient values are plotted against the experimental data. The agreement between model and experiment is very good, by just fitting the electrolyte-electrolyte binary interaction parameters.

## Conclusion

The refined e-NRTL model was used to model the thermodynamics of brine solutions. 11 binary aqueous systems involving the following ion;  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Li}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$ ,  $\text{Cl}^-$ ,  $\text{SO}_4^{2-}$  were studied and one  $\text{NaCl-KCl-H}_2\text{O}$  ternary system was studied for validation of the extrapolation ability of the binary models to mixtures of electrolytes. The binary interaction parameters were fitted against experimental data of activity and osmotic

coefficients of binary systems. The osmotic coefficients of the ternary system were predicted with fair accuracy. The simulation was performed on an Intel(R) Core(TM) i7-7700 processor with CPU of 3.60 GHz and 8.00 GB RAM. This thermodynamic model can be extended to predict other thermodynamic properties beyond the osmotic and activity coefficients presented here and at higher molalities. It is also being implemented in the user-friendly environment of IDEAS to develop open-source libraries of property models for highly concentrated aqueous and heterogeneous electrolyte solutions, embedded in block model libraries for processing and separations. These libraries are then intended to be utilized for process synthesis and intensification of brine treatment options.

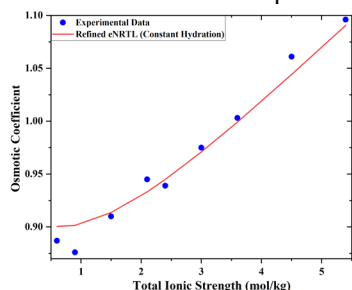


Figure 4: Prediction of osmotic coefficient of NaCl-KCl-H<sub>2</sub>O ternary system.

Table 1: Binary interaction parameters for NaCl-KCl-H<sub>2</sub>O ternary system

	Refined e-NRTL
$\tau_{H_2O,NaCl}$	7.951
$\tau_{H_2O,KCl}$	7.696
$\tau_{NaCl,H_2O}$	-3.984
$\tau_{NaCl,KCl}$	1.3672
$\tau_{KCl,H_2O}$	-3.830
$\tau_{KCl,NaCl}$	1.3672

### Acknowledgments

Funding for this project was provided by the National Alliance for Water Innovation (NAWI) of the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy (EERE), Advanced Manufacturing Office, under Funding Opportunity Announcement Number DE-FOA-0001905.

### References

- Bhattacharia, S.K., Chen, C.-C., 2015. Fluid Phase Equilib 387, 169–77.
- Bollas, G.M., Chen, C.C., Barton, P.I., 2008. AIChE Journal 54, 1608–24.
- Chen, C.-C., 2006. Fluid Phase Equilib 241, 103–12.
- Chen, C.C., Evans, L.B., 1986. AIChE Journal 32, 444–54.
- Dinane, A., Guendouzi, M. el, Mounir, A., 2002. J Chem Thermodyn 34, 423–41.
- Honarparvar, S., Saravi, S.H., Reible, D., Chen, C.C., 2018. Fluid Phase Equilib 470, 221–31.
- Honarparvar, S., Saravi, S.H., Reible, D., Chen, C.C., 2017. Fluid Phase Equilib 447, 29–38.
- Mehrpooya, M., Habibi, R., 2020. J Clean Prod 275, 123836.
- Paukert Vankeuren, A.N., Hakala, J.A., Jarvis, K., Moore, J.E., 2017. Fluid. Environ Sci Technol 51, 9391–9402.
- R. A. Robinson, R.H.S., 1959. Electrolyte Solutions: Second Revised Edition. Dover Publications.
- Tanveer, S., Chen, C.C., 2016. Fluid Phase Equilib 409, 193–206.
- Xi Yang, Paul. I. Barton, George. M. Bollas, 2023. In Preparation.

# Stable two-stage scenario generation via game-theoretic optimisation

Georgios L. Bounitsis, Lazaros G. Papageorgiou, Vassilis M. Charitopoulos

*The Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, University College London, Torrington Place, London WC1E 7JE, UK*  
[v.charitopoulos@ucl.ac.uk](mailto:v.charitopoulos@ucl.ac.uk)

## Abstract

Scenario generation methods constitute an important aspect towards efficient solution of Stochastic Programming (SP) problems and exploitation of big data. The ability of these methods to consistently provide scenario sets which guarantee stability on the solution of the stochastic programs is determinant of their performance. In this context, we present a modification of the existing Distribution and Moment Matching Problem (DMP) which is formulated as Mixed-Integer Linear Programming (MILP) model. The Nash bargaining approach is employed and the different statistical properties of the DMP are considered as players. Through this game-theoretic approach the impact of the user-defined parameters on the scenario generation procedure is investigated. Results from a capacity planning case study highlight the benefits of the proposed approach with respect to in-sample and out-of-sample stability.

**Keywords:** Scenario Generation; Stochastic Programming; Optimisation; Mixed-Integer Linear Programming (MILP); Nash equilibrium.

## 1. Motivation

Optimisation under uncertainty has emerged as a focal topic within the Process Systems Engineering (PSE) research agenda (Li & Grossmann, 2021; Charitopoulos et al., 2018). Existing mathematical frameworks can tackle optimisation problems under uncertainty for various degrees of risk aversion. In particular, risk-neutral Stochastic Programming (SP) takes advantage of scenario-based formulations to optimise the expected value of the problem over an assumed probability distribution. Recently, research interest has been directed towards the development of scenario tree generation (SG) or reduction approaches, which create a smaller and representative set of scenarios to efficiently solve computationally challenging stochastic programs. Among them, the optimisation-based SG techniques form an important category (Bertsimas & Mundru, 2022). Moment Matching Problem (MMP) is part of the latter and can generate a scenario tree by solving a statistical errors' minimisation problem (Høyland & Wallace, 2001). Recently, Bounitsis et al. (2022) proposed a Mixed-Integer Linear Programming (MILP) model for the Distribution and Moment Matching Problem (DMP), which was shown to overcome the so-called under-specification issues of the Nonlinear Programming (NLP) based counterparts (Calfa et al., 2014).

In this work, we present an extension to the work of Bounitsis et al. (2022) which aims to further enhance stability and performance of the SG framework, by mitigating the impact of the model's user defined parameters. A modification of the DMP MILP model is proposed, in which the Nash bargaining approach is used. The problem at hand is modelled as an MILP following a separable programming reformulation (Gjerdrum et al.



2001; Charitopoulos et al., 2020). Finally, the enhanced quality and stability of the optimisation-based SG methods are validated through bias, in-sample and out-of-sample tests (Kaut & Wallace, 2007).

The remainder of the article is organised as follows: in Section 2 a summary of the main theoretical aspects is conducted while the detailed methodology is outlined in Section 3. In Section 4, the proposed model is employed to evaluate its stability on a capacity planning case study. Finally, conclusions are drawn in Section 5.

## 2. Preliminaries and literature review

### 2.1. Stochastic Programming

In Stochastic Programming (SP) problems the uncertainty is modelled via a known discrete probability distribution. Two-stage stochastic programming (TSSP) is the most common version of stochastic programming problems, in which the uncertainty is revealed at one step. Optimisation problems under uncertainty can be reformulated as stochastic programs using discrete realisations for the uncertain parameters. These are also called scenarios ( $k$ ) and are specified by their values for the uncertain parameters ( $\xi_1, \dots, \xi_k$ ,  $k = \{1, \dots, K\}$ ) and corresponding probabilities of occurrence ( $p_1, \dots, p_k$ ,  $k = \{1, \dots, K\}$ ). TSSP optimises simultaneously the objective function of the first-stage decisions and the expected value of the second-stage costs, which is formulated as:

$$\mathbb{E}[Q(x, \xi)] = \sum_{k=1}^K p_k \cdot Q(x, \xi_k) \quad (1a)$$

Thus, the *deterministic equivalent problem* for the TSSP is given by Eq. (1b):

$$\begin{aligned} \min_{x, z_1, \dots, z_k} \quad & c^\top x + \sum_{k=1}^K p_k \cdot q_k^\top \cdot z_k \\ \text{s. t.} \quad & Ax \leq b \\ & T_k x + W_k z_k \leq h_k, \quad k = 1, \dots, K \\ & x, z_k \geq 0, \quad k = 1, \dots, K \end{aligned} \quad (1b)$$

In this formulation,  $k$  is the index of scenarios,  $x \in \mathbb{R}^n$  denotes the “here-and-now” decisions,  $z \in \mathbb{R}^m$  is the set of the “wait-and-see” decisions and the vector  $\xi = (q, T, W, h)$  contains the data (known or uncertain) of the second stage problem.

### 2.2. Scenario generation methods

In cases of either continuous distributions or a very large number of possible realisations of the uncertain parameters, scenario tree generation methods can be used to create a smaller in size set of scenarios for the uncertain parameters, with specific values and probabilities, which replaces the original uncertainty set for the solution of the TSSP.

#### 2.2.1. Distribution and Moment Matching Problem

Moment Matching Problem (MMP) for scenario tree generation was introduced by Høyland and Wallace (2001) and aims at the minimisation of the errors regarding the considered statistical properties between the original uncertain set and the final reduced set. A typical version of MMP accounts for first four statistical moments (i.e., mean, variance, skewness, kurtosis) and covariance between the parameters. The optimisation-based NLP formulation has been extended to simultaneously match the stochastic distribution of the uncertain parameters by minimising the errors regarding the empirical cumulative distribution function (ECDF) (Calfa et al., 2014). This problem is referred to as Distribution Matching Problem (DMP) and remains NLP. However, it may lack of

stability as it suffers from under-specification issues, in which case either a unique scenario is assigned to several nodes, or zero probabilities are assigned to some nodes.

Recently, both MMP and DMP have been reformulated as MILP problems by employing binary variables to indicate the selection of the scenarios from a known original scenario set (Kaut, 2021; Bounitsis et al., 2022). Regarding DMP, Bounitsis et al. (2022) proposed a scenario generation MILP model using  $L^1$  and  $L^\infty$  norms to minimise the errors regarding the statistical properties. The model was integrated into a framework with copula-based sampling and clustering techniques and its unified impact seems to lead to significant mitigation of the under-specification issues of the NLP model.

### 2.2.2. Stability assessment of scenario generation methods

Stability measures assess the performance of a scenario generation or reduction method by testing its capability to generate different trees which lead to consistent expected values. Considering multiple scenario trees,  $T_c$ , which lead to stochastic solutions  $\bar{x}_c$  of the scenario-based problem and omitting the notation for the “wait-and-see” variables, the in-sample and out-of-sample stability are achieved, if Eqs. (2)-(3) are true respectively (Kaut & Wallace, 2007):

$$f(\bar{x}_c; T_c) \approx f(\bar{x}_{c'}; T_{c'}) \quad (2)$$

$$f(\bar{x}_c; \xi) \approx f(\bar{x}_{c'}; \xi) \quad (3)$$

Regarding Eq. (3) a large reference tree  $R$  can be used as an approximation of the true stochastic process  $\xi$  in order to estimate the out-of-sample stability. Finally, bias is also an important measure for the evaluation of the quality of certain stochastic solutions  $\bar{x}$ :

$$B(\bar{x}) = f(\bar{x}; R) - \min_{x \in X} f(x; R) \quad (4)$$

Bias captures the error that fixed here-and-now decisions,  $\bar{x}$ , impose to the so-called full-space problem (FS) which is solved using the reference tree  $R$  (Kaut & Wallace, 2007).

### 2.3. Nash game-theoretic approach

According to Nash bargaining approach a fair solution for all players of a game can be achieved taking into consideration their initial position. Given a game with players  $t \in T$ , their corresponding objective functions and status quo points prior to joining the game (or generally a lower profit requirement point), denoted as  $\pi_t^{SQ}$ , then a fair solution point,  $\pi_t$ , can be achieved. The bargaining solution fairly contributes the payoff to the players obeying the axioms of Pareto optimality, symmetry, linear invariance, and independence of irrelevant alternatives. For a typical profit maximisation problem, the fair solution maximises the Nash product which is given by (Harsanyi, 1977):

$$\Phi = \prod_{t \in T} (\pi_t - \pi_t^{SQ})^{\alpha_t} \quad (5)$$

where  $\alpha_t$  is the negotiation power of each player  $t$ . Hence, a player  $t$  enters the game only if  $\pi_t \geq \pi_t^{SQ}$ , i.e., if the profit entering the game is greater than its status quo point.

However, the objective function of Eq. (5) is nonlinear and nonconvex leading to computationally challenging problems. An approximate linearisation of the Nash product based on a separable programming approach has been proposed by Gjerdrum et al. (2001). Initially, a convexification is conducted by employing a logarithmic transformation and then the concave Nash product is linearised via a piecewise linear function of  $g$  prespecified grid points. So, the linear approximation of the Nash product is:

$$\tilde{\Phi} = \sum_{t \in T} \sum_{g \in G} \alpha_t \cdot \ln(\tilde{\pi}_{tg} - \pi_t^{SQ}) \cdot \lambda_{tg} \quad (6)$$

where by defining SOS2 variables  $\lambda_{tg}$  and variables for the profit of player  $t$  at grid point  $g$ ,  $\tilde{\pi}_{tg}$ , then the profit of each player  $t$  can be estimated as:

$$\pi_t = \sum_{g \in G} \tilde{\pi}_{tg} \cdot \lambda_{tg} \quad \forall t \in T \quad (7)$$

$$\sum_{g \in G} \lambda_{tg} = 1 \quad \forall t \in T \quad (8)$$

### 3. Methodology and mathematical developments

In the recent study by Bounitsis et al. (2022), performance and stability assessment of the deterministic models were conducted for obtained stochastic solutions from scenario trees of “slightly different sizes” (King and Wallace, 2012). However, results in Bounitsis et al. (2022) indicated that even user-defined parameters of the formulation may affect stability of the deterministic models. In this work, this remark is investigated by conducting stability tests for varying values of user-defined weights of the objective function. Moreover, aiming to improve stability, a modification of the DMP MILP SG model is proposed. The Nash bargaining approach is employed in the terms of the objective function and a separable programming based linearisation of the objective function (Eqs. (6)-(8)) assists to preserve the MILP formulation of the problem. Overall, the modified model can replace DMP MILP in the SG methodology by Bounitsis et al. (2022), which includes the preliminary steps of: (i) statistical analysis of uncertain data, (ii) simulation of distributions, (iii) copula sampling of original scenarios, (iv) clustering. Proposed Nash reformulation of the DMP MILP model is presented for the  $L^\infty$ -norm-based formulation. Let us denote as  $i \in I$  the uncertain parameters and  $m \in M$  the considered statistical moments. Moreover, we introduce an index  $t \in T = \{SM, ECDF, COV\}$  for the terms of the objective function of the original model, which correspond to the players of the game theoretic approach. The model can be written as:

$$\max \sum_{t \in T, g \in G} \ln(\pi_t^{MAX} - \tilde{\pi}_{tg}) \cdot \lambda_{tg} \quad (9)$$

s. t. Eqs. (7)-(8) and

$$\pi_t = r_t^{MAX} \quad \forall t \in T \quad (10)$$

$$r_{SM}^{MAX} \geq W_{im}^{SM} \cdot (d_{im}^+ + d_{im}^-) \quad \forall i \in I, m \in M \quad (11)$$

$$r_{COV}^{MAX} \geq W_{ii'}^{COV} \cdot (c_{ii'}^+ + c_{ii'}^-) \quad \forall i \in I, i' \in I: i < i' \quad (12)$$

$$r_{ECDF}^{MAX} \geq W_i^{ECDF} \cdot e_i \quad \forall i \in I \quad (13)$$

The rest necessary equations of the model and the definition of the deviations ( $d_{im}^+, d_{im}^-, e_i, c_{ii'}^+, c_{ii'}^-$ ) are identical to the original DMP MILP model which can be found in Bounitsis et al. (2022). Eq. (9) is the objective function and aims at the maximisation of the Nash product, which for this case is expressed using parameters  $\pi_t^{MAX}$  as the higher allowable statistical error for each term  $t$ . These can be obtained by solving the original DMP MILP for minimisation of each term separately. Then, Eqs. (11) – (13) capture the maximum errors for each term, i.e., statistical moments (SM), covariance matrix (COV) and ECDF. The weights on the deviations are defined as  $W_{im}^{SM} = \bar{w}_{im}^{SM} / |\bar{M}_{im}|$  and  $W_{ii'}^{COV} = \bar{w}_{ii'}^{COV} / |\bar{C}_{ii'}|$ , where  $\bar{M}_{im}, \bar{C}_{ii'}$  are the estimated statistical properties from the data and  $\bar{w}_{im}^{SM}, \bar{w}_{ii'}^{COV}$  are the user-defined weights. Finally,  $W_i^{ECDF}$  is also a user-defined weight for the error on ECDF.

#### 4. Case Study

A capacity planning problem under uncertainty is studied. Detailed description, model and data can also be found in Bounitsis et al. (2022). Briefly, it is an MILP profit maximisation problem, which selects between 11 candidate processes on the production of 5 final products from 5 raw materials. Uncertainty is revealed in the production yields of processes and/or demands of final products. The bias and stability of DMP NLP, MILP and its proposed Nash-based reformulation is investigated for sets of varying values of  $\bar{w}_{im}^{SM}, \bar{w}_{ii'}^{COV}, W_i^{ECDF}$ . Typical values 1, 10, 50 and, 100 are considered and 37 sets of values are constructed allowing 1 or 2 of the attributes to vary concurrently. Regarding Nash reformulation, 50 grid points are employed for the piece-wise linear approximations. Instances of increasing complexity are examined and are presented in Table 1. MILP models uses copula-based original scenarios as input.

**Table 1:** Instances of scenario reduction in the case study

Instances	Uncertain parameters	Original scenarios	Final scenarios
1	2	1,000	10
2	4	2,000	20
3	8	2,000	40

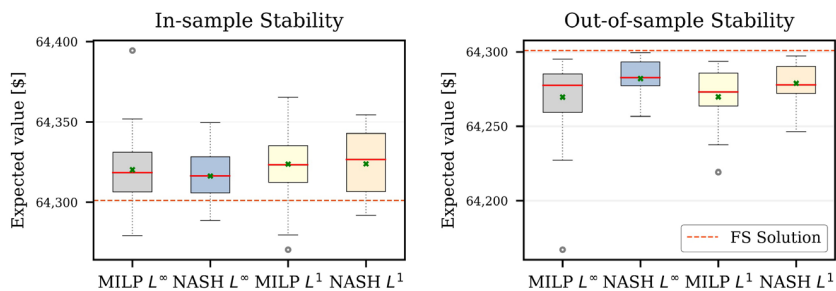
The executions are performed in a Dell workstation with Intel® Core™ i9-10900K CPU @ 3.70 GHz and 32.00 GB RAM. NLP model is solved using the BARON solver, while MILP models using the GUROBI 9.5 solver within GAMS 38.2 modelling system. Although calculation of lower bound constitutes a significant drawback of DMP MILP, the Nash-based reformulations can achieve near optimal solutions in short execution times. Towards a fair comparison, time limit of 1,200s and an optimality gap tolerance of 1% are set for the MILP models.

Firstly, the computational studies reflect the enhanced quality of stochastic solutions that the scenario sets by the considered MILP models provide compared to their NLP counterparts. In particular, NLP SG models suffer from under-specification issues and may generate scenario sets which lead to extreme bias values. Results among examined weights' sets are reported in Table 2 and are indicative of NLP models non-competitiveness. Moreover, Nash-based models lead to the lowest bias values.

**Table 2:** Average (%) bias values for the investigated instances.

Instance	NLP L <sup>1</sup>	NLP L <sup>∞</sup>	MILP L <sup>1</sup>	MILP L <sup>∞</sup>	NASH L <sup>1</sup>	NASH L <sup>∞</sup>
1	1.327	1.383	0.047	0.034	<b>0.027</b>	0.028
2	2.364	2.834	0.048	0.049	0.034	<b>0.029</b>
3	2.761	3.049	0.058	<b>0.042</b>	<b>0.042</b>	<b>0.042</b>

Focusing on the MILP SG methods, the distribution of their expected values for the in-sample and out-of-sample stability (as they are introduced in Eqs. (2) - (3), respectively) are visualised for instance 2 in Fig. 1. While all SG models display similar results for in-sample stability, the proposed Nash reformulations seem to improve the performance of the methods for out-of-sample stability testing, which pertains also to the quality of the provided stochastic solutions. Enhanced stability of the proposed reformulations can be interpreted by the considerably tighter interquartile ranges. Moreover, expected values lie closer to TSSP full-space (FS) solution and indicate improved quality of the stochastic solutions. Ultimately, the results attest that the DMP NLP model is the least stable of the investigated approaches. DMP MILP's stability has been studied in the work by Bounitsis et al. (2022) but is also validated in this work for the case of varying weights of the formulation. Finally, the results demonstrate that combination of DMP MILP with the Nash bargaining approach may further enhance the stability of the SG method.



**Figure 1:** In-sample & out-of-sample stability results for instance 2.

## 5. Conclusions and future directions

Integration of Nash game-theoretic concepts with DMP MILP model can lead to a more stable version of the scenario generation method and further alleviate the under-specification drawbacks of the NLP models for the user-defined parameters. Future work within our research group focuses on: (i) the implementation of optimisation and decomposition techniques to improve the computational performance of the proposed model while maintaining stability, and (ii) the extension to multi-stage setting.

**Acknowledgements:** Financial support from the UK EPSRC under project EP/V034723/1, EP/V051008/1 & EP/V050168/1 is gratefully acknowledged.

## References

- D. Bertsimas, N. Mundru, 2022. Optimization-Based Scenario Reduction for Data-Driven Two-Stage Stochastic Optimization. *Oper. Res.*
- G.L. Bounitsis, L.G. Papageorgiou, V.M. Charitopoulos., 2022. Data-driven scenario generation for two-stage stochastic programming. *Chem. Eng. Res. Des.* 187, 206–224
- B.A. Calfa, A., Agarwal, I.E. Grossmann, J.M. Wassick, 2014. Data-driven multi-stage scenario tree generation via statistical property and distribution matching. *Comput. Chem. Eng.*, 68, 7–23.
- V.M. Charitopoulos, A. Aguirre, L.G. Papageorgiou, V. Dua, 2018. Uncertainty aware integration of planning, scheduling and multi-parametric control. *Comput. Aid. Chem. Eng.* 44, 1171–1176.
- V.M. Charitopoulos, V. Dua, J.M. Pinto, L.G. Papageorgiou, 2020. A game-theoretic optimisation approach to fair customer allocation in oligopolies. *Optim. Eng.*, 21(4), 1459–1486.
- J. Gjerdrum, N. Shah, L.G. Papageorgiou, 2001. Transfer prices for multienterprise supply chain optimization. *Ind. Eng. Chem. Res.* 40, 1650–1660.
- J. Harsanyi, 1977. Morality and the Theory of Rational Behavior. *Soc Res.* 44, 623–656.
- K. Høyland, S. W. Wallace, 2001. Generating Scenario Trees for Multistage Decision Problems. *Manag. Sci.*, 47(2), 295–307.
- M. Kaut, 2021. Scenario generation by selection from historical data. *Comput. Manag. Sci.* 18, 411–429.
- M. Kaut, S. W. Wallace, 2007. Evaluation of scenario-generation methods for stochastic programming. *Pacific J. Optim.*, 3.
- A. J. King, S. W. Wallace, 2012. *Modeling with Stochastic Programming*. Springer Series in Operations Research and Financial Engineering.
- C. Li, I. E. Grossmann, 2021. A Review of Stochastic Programming Methods for Optimization of Process Systems Under Uncertainty. *Front. Chem. Eng.* 2, 622241.

# Global One-dimensional Root-finding Algorithm for Derivable and Continuous Functions

Andrea Galeazzi,<sup>a,b</sup> Kristiano Prifti,<sup>a,b</sup> Flavio Manenti<sup>a,b,\*</sup>

<sup>a</sup>*Dipartimento di Chimica, Materiali e Ingegneria Chimica “Giulio Natta”, Politecnico di Milano, Piazza Leonardo Da Vinci, 32, Milano, 20133, Italy*

<sup>b</sup>*Consorzio Interuniversitario Nazionale per la Scienza e Tecnologia dei Materiali, Via Giusti 9, Firenze, 50121, Italy*

\**flavio.manenti@polimi.it*

## Abstract

Efficient and robust root finding is a core problem of chemical and process engineering solutions. Thermodynamics, nonlinear systems, finite difference methods for partial differential equations, and optimization are just some examples in which root finding is fundamental. In this work, an algorithm that can be integrated with any local root-finding method is proposed to solve the problem of finding robustly and efficiently all the roots of an objective function in a domain. The method requires the function to be continuous up until the second-order derivative and it is here applied for the one-dimensional case. With the definition of a proper mesh, the first derivative is investigated to find the sub-domains that contain the zeros of the objective function. The study of the derivative guarantees to find those zeros that cannot be identified using only Bolzano’s existence criterion.

**Keywords:** One-dimension, Global Root-finding, Search Grid, Continuous and Derivable Function, Robust and Efficient Solution.

## 1. Introduction

This paper describes a method for solving the problem of computing all the simple roots of a given equation:

$$f(x) = 0 \tag{1}$$

in a predefined interval  $(a, b)$  where  $f: [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$  is continuous and derivable up to the second order derivative,  $f(x) \in C^2$ . The method proposed may be applied on top of any local root-finding algorithms and the information required is the definition of a search grid and the computation of the function and its continuous first-order derivative. Many different fields of science and technology make use of root-finding methods to solve various problems. One important example is the maximization or minimization of an objective function. In this case, a root-finding method can be applied to the first derivative of the objective function to find the optimum value. Another example of root-finding usage is in the numerical solution of non-analytical thermodynamic equations of state and the flashing problem (Bisotti et al., 2021; Quarteroni et al., 2007). Such problems are at the core of chemical process engineering solutions since the definition of the physical state of a mixture is the first step in fixing the degrees of freedom of a process simulation.

Being able to solve, efficiently and robustly, such thermodynamic problems brings an overall advantage in process modeling and a significant computational time reduction.

In this case, efficiency is defined in terms of the amount of functions evaluation needed to find a solution while higher robustness means a greater probability to find all the solutions in the investigated domain.

The method proposed in this work tries to tackle this challenge by first studying the function derivative to find with certainty all the domains in which the objective function has solutions and find them. This algorithm can be applied on top of any local root-finding solver, and it is based thoroughly on Bolzano's existence criterion (described in Section 2.1). The premise is that, given an appropriate search grid, or mesh, it guarantees to find all the solutions of the objective function inside a predefined domain.

## 2. Method

### 2.1. Basic Concepts

The method developed in this work is based on a very simple and well-known theorem which is Bolzano's existence criterion. It states that when  $f$  is a  $C^2$  function in  $[a, b] \subset \mathbb{R}$  and the following is true:

$$f(a) \cdot f(b) \leq 0 \quad (2)$$

at least a root is guaranteed to exist. However, the criterion does not give any information on the opposite case. If Equation (2) is not true it is impossible to state whether there is or not a solution between  $a$  and  $b$  since there could still be a zero. The application of this criterion to solve the root-finding problem is not new (Brent, 1971; Kearfott, 1987; Wood, 1991).

Thus, for this study, it has been decided to extend Bolzano's criterion using the function derivative since the key mathematical property is that between two adjacent roots  $(a, b)$  of the derivative of a  $f(x) \in C^2$  function, where  $f'(x) \geq 0 \vee f'(x) \leq 0 \forall x \in [a, b]$ ,  $f$  is strictly monotonic. Thus, if Equation (2) is true in the same interval  $[a, b]$ , a single root of  $f$  may only exist. Other authors have proposed different methods for solving this problem, for example by applying the concept of a Dominating Function of the derivative (Ying and Katz, 1989).

The definition of a search grid, or mesh, in this problem, is of utmost importance. Any possible failure of this whole method could be traced back mainly to this step. An appropriate mesh is defined such that it always contains no more than one single derivative root. To do so, a proper understanding of the function under study could be required. Otherwise, the most basic solution is to arbitrarily apply an extremely dense search grid. Some authors have addressed this problem by implementing an adaptive mesh refinement procedure (Gong et al., 2020; Razbani, 2015).

Thus, the case in which Equation (2) is not true that still has a solution can be ruled out if using an appropriate search grid in the first place since it can only happen when the roots of the derivative and the function are strictly coincident.

Should the search domain contain any equality or inequality constraints for  $x$  the problem must be split into multiple sub-problems and solved using the proposed method.

### 2.2. The Algorithm

The proposed algorithm can be briefly described in the following points:

- 1) Input:  $f(x)$ ,  $f'(x)$ , global search domain, local  $f'$  search sub-domains (grid).
- 2) Evaluate  $f$  and  $f'$  at each sub-domain node.
- 3) Select only sub-domains where  $f$  and  $f'$  show a sign change between the nodes.
- 4) Start a local root-finding in the selected sub-domains for  $f'$ ;

- 5) Check for coincident zeros between  $f$  and  $f'$  at local derivative roots.
- 6) Generate new sub-domains, with adjacent non-coincident  $f'$  roots and selected  $f$  sub-domains from Step 2 as nodes.
- 7) Select only sub-domains where  $f$  shows a change of sign between the nodes.
- 8) Start a local root finding in the new sub-domains for  $f$ .
- 9) Collect all the zeros of  $f$  (coincident and non-coincident roots).

A flowchart of this method is shown in Figure 1.

A more in-depth description of the algorithm is reported below:

*Step 1.* The necessary inputs to the problem are defined: the objective function  $f(x)$ , its derivative  $f'(x)$ , the extrema of the global search domain  $[a, b]$  where  $f(x) \in C^2$ , and the local search sub-domain nodes. The latter are crucial for the successful execution of this method. They may be defined arbitrarily but they must be generated in such a way as to always expect one and only one solution of  $f$  inside root-containing sub-domains. If this is accomplished then the robustness is guaranteed, otherwise, the search mesh should be refined accordingly. Having a good knowledge of the objective function to solve should result in a proper definition of the mesh.

*Step 2.* Each node of the local sub-domains is evaluated in  $f(x)$  and  $f'(x)$ .

*Step 3.* Bolzano's existence criterion is applied to evaluate which sub-domain contains a root of the function and its derivative. Thus, Equation (2) is applied to select derivative subdomains with  $f$  roots and  $f$  sub-domains with sign change not bracketed by  $f'$  roots. This redundancy is mandatory to overcome any possible saddle point where the derivative does not change sign but a root still exists.

*Step 4.* Each  $f'$  bounded sub-domain thus selected is investigated using a local root-finding method since it must contain a root. In this Step, it could be advised to apply a bracketing method since it mathematically guarantees convergence and the downside of the slower convergence rate with respect to the open methods should be mitigated given the smaller local sub-domains. However, even an open method can be applied knowing that, in case of non-convergence or convergence outside the selected subdomain, it could be possible to fall back to a bracketing method and restart the root-finding.

*Step 5.* By evaluating the roots found in Step 4 in  $f(x)$  it is possible to uncover any maxima or minima point that touch the axis at the extremum. They are coincident zeros of both  $f(x)$  and  $f'(x)$  and they are collected and saved for Step 9.

*Step 6.* The information gathered in the previous Step is now used to generate new sub-domains for the root-finding of the objective function with non-coincident roots as nodes. Each sub-domain is defined by nodes that are an adjacent pair of roots of  $f'$ , thus not containing a coincident zero in between.

*Step 7.* Again, as in Step 3, Bolzano's existence criterion, see Equation (2), is applied to rule out all the subdomains in which a change of sign is not appearing.

*Step 8.* Inside each newly defined search sub-domain, a local root-finding method can be launched. Also, in this case, it is guaranteed to find a zero which may not be an extremum point thanks to Steps 5 and 6. This guarantee is given by the fact that, if there is a significant change in Step 7 and the derivative is passing through zero at the nodes of the subdomain there is one and only one root available on the objective function. In this Step, the choice of using an open method or a bracketing method is mainly arbitrary. It could be advisable to apply an open method, such as the Newton method if it is possible to generate an appropriate guess value (a trivial guess value could be positioned in the middle of the local sub-domain) for the function under investigation.

*Step 9.* All the roots of the objective function  $f(x)$  found in Step 5 and Step 8 are then collected and the problem is solved.



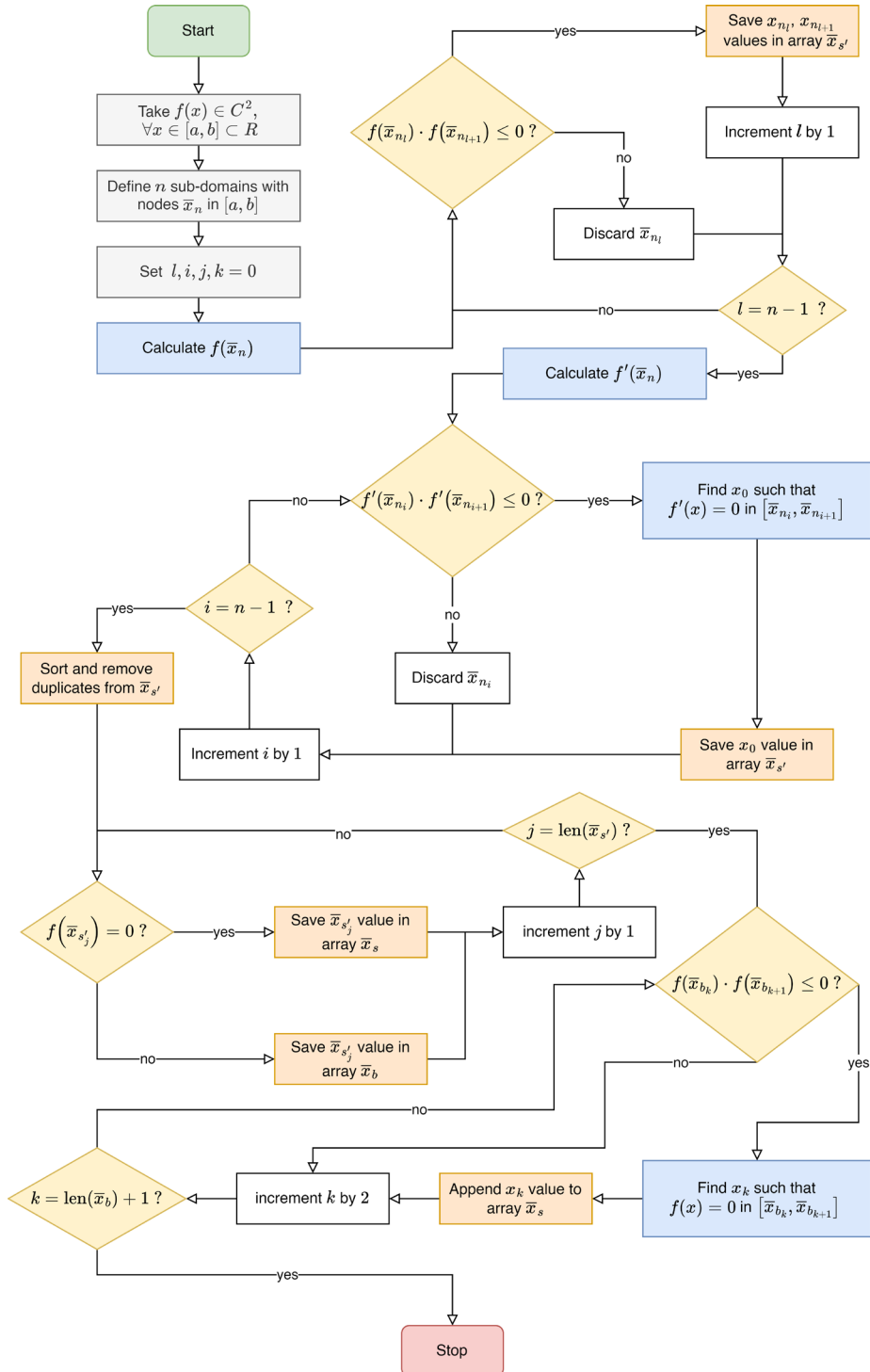


Figure 1: Flowchart of the global root finding algorithm.

### 3. Results & Discussion

#### 3.1. Prototype Function Example

The method described in Section 2 has been applied to the following prototype function:

$$f(x) = \sin^2(x) \cos(x) \tag{3}$$

$$f'(x) = 2 \cos(x)^2 \sin(x) - \sin(x)^3 \tag{4}$$

chosen because it has an analytical solution and shows many coincident solutions to better showcase the method described above.

In Figure 2, the algorithm Steps presented in Section 3.1 are reported for the example function of Equations (4) and (5). In Figure 2.a, the function and its derivative are shown along with the initial derivative search grid. The investigated domain is  $x \in [0,10]$  and the grid presented shows 30 equally distributed sub-domains. In Figure 2.b, Steps 2,3,4 are reported. The filled and dashed lines are the linear interpolation between each value of  $f'(x)$  at each node. The former, indicate sub-domains with a sign change. In Figure 2.c, Steps 5,6,7,8 are shown. The updated search grid is created using the roots of  $f'(x)$ . Grayed areas are the excluded sub-domains given that there is a coincident solution, and the two roots are not adjacent. In Figure 2.d, Step 9 is reported. All the solutions inside  $[0,10]$  are collected and the problem is solved.

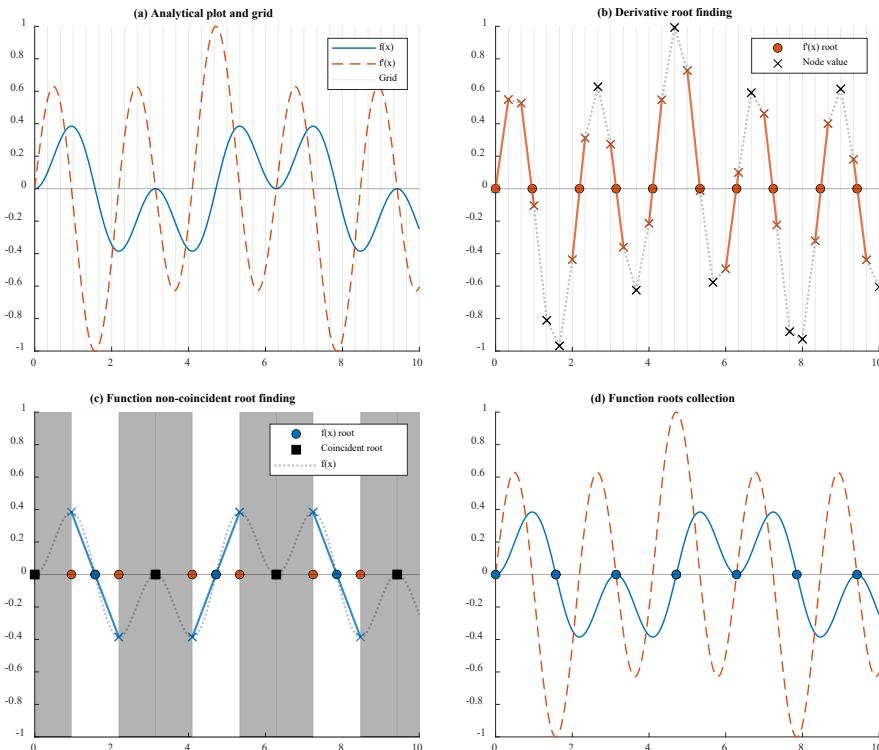


Figure 2: Condensed graphical representation of the algorithm Steps presented in Section 2.2 on the prototype function introduced in Section 3.1.

By using the proposed algorithm, it has been possible to solve the presented root finding problem by only using 12 local root findings (9 for the derivative and 3 for the objective function) and approximately 40 function evaluations (30 for the derivative and 10 for the objective function). If we consider a constant number of 10 iterations for every root-finding problem, we end up with a total of approximately 160 function evaluations. In comparison, by using a more naive solution, which consists of directly launching 30 root-finding solvers (that do not guarantee finding a solution) the results would be that of using 300 function evaluations. Thus, in this case, using this method the computational time required is almost halved.

#### 4. Conclusions

In conclusion, an algorithm has been proposed to solve reliably and efficiently the global root-finding problem of second-order continuous and derivable functions inside a domain. The method applied is based on an appropriate mesh and uses extensively Bolzano's existence criterion to define what subdomains to exclude from the root search. This method could benefit from a parallel computing development, possibly with a cloud approach (Galeazzi et al., 2021), to aid in the solution of chemical engineering problems, especially in thermodynamics.

#### References

- F. Bisotti, A. Galeazzi, L. Galatioto, F. Masserdotti, A. Bigi, P. Gritti, F. Manenti, 2021, Implementing robust thermodynamic model for reliable bubble/dew problem solution in cryogenic distillation of air separation units. *International Journal of Thermofluids*, 10, 100083.
- R.P. Brent, 1971, An algorithm with guaranteed convergence for finding a zero of a function, *The Computer Journal*, 14, 4, 422–425.
- A. Galeazzi, R. Nasti, G.L. Bozzano, L. Verotta, S. Marzorati, F. Manenti, 2021, A Cloud Computing Application for the Supercritical Carbon Dioxide Extraction Using Coffee Grounds Silverskin, *Computer Aided Chemical Engineering*, 50, 31 European Symposium on Computer Aided Process Engineering, 1035–1040.
- W. Gong, Y. Wang, Z. Cai, L. Wang, 2020, Finding Multiple Roots of Nonlinear Equation Systems via a Repulsion-Based Adaptive Differential Evolution. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 50, 4, 1499–1513.
- R.B. Kearfott, 1987, Some tests of generalized bisection, *ACM Trans. Math. Softw.*, 13, 3, 197–220.
- A. Quarteroni, R. Sacco, F. Saleri, 2007, Rootfinding for Nonlinear Equations, *Numerical Mathematics, Texts in Applied Mathematics*, 247–284.
- M.A. Razbani, 2015, Global root bracketing method with adaptive mesh refinement, *Applied Mathematics and Computation*, 268, 628–635.
- G.R. Wood, 1991, Multidimensional bisection applied to global optimisation. *Computers & Mathematics with Applications*, 21, 6, 161–172.
- X. Ying, I.N. Katz, 1989, A simple reliable solver for all the roots of a nonlinear function in a given domain, *Computing*, 41, 4, 317–333.

# A Demethanizer column Digital twin with non-conventional LSTM neural networks arrangement

Marta Mandis<sup>a</sup>, Roberto Baratti<sup>a</sup>, Jorge Chebeir<sup>b</sup>, Stefania Tronci<sup>a</sup>, José A. Romagnoli<sup>b</sup>

<sup>a</sup> *Dip. di Ingegneria Meccanica, Chimica e dei Materiali, Università degli Studi di Cagliari, Cagliari, 09123, Italy*

<sup>b</sup> *Department of Chemical Engineering, Louisiana State University, Baton Rouge, LA, 70809, USA*

## Abstract

This work aims to develop a digital twin for a demethanizer column and provide a useful tool for monitoring and quality control of the NGL recovery process. For this purpose, a digital data-driven model was proposed to mimic real dynamics of a cold residue reflux (CRR) unit through the incorporation of physical knowledge. A non-conventional LSTM network arrangement was developed considering training test and validation data sets generated by the process simulator Aspen HYSYS®. This simulation model was built by considering realistic measurement noises to mimic the actual measures in a real plant. The obtained surrogate model was evaluated considering its ability to recreate the operation of the actual distillation column, estimating the temperature and composition transient profiles of the bottom column product and of every stage of the column. Overall, the model developed with the proposed LSTM network arrangement proves capable of successfully reconstructing the actual profiles of all the considered variables.

**Keywords:** Natural gas liquids recovery, LSTM Neural Networks, Digital twin, Distillation Column, Dynamic process simulation

## 1. Introduction

Natural gas is a hydrocarbon mixture constituted mostly of methane and a variable fraction of heavier hydrocarbons known as natural gas liquids (NGL). The latter constitute a source of additional profit as can be sold separately with a higher selling price and used as industrial feedstock materials. Additionally, with NGL separation it is possible to reduce CO<sub>2</sub> emissions related to natural gas combustion. The NGL recovery process under steady-state conditions has been extensively studied in the literature (Kidnay et al., 2011; Chebbi et al., 2010; Getu et al., 2013; Kherbeck & Chebbi, 2015). The associated dynamics of the NGL separation process have been the subject of recent studies and generally focused on the dynamics and control of product quality indices (Luyben, 2013; Chebeir et al., 2019; Mandis et al., 2022). For this, it is essential to have information on the plant composition targets, which can be obtained with the use of expensive composition analyzers. These analyzers also introduce long-time delays in the control loops due to measurements delays. To overcome these drawbacks, the replacement of measured concentrations with accurate real-time estimations represents one of the most widespread approaches. In this type of task, temporal dependencies play a key role to obtain adequate estimations and with the goal of addressing time-dependent problems Recurrent neural networks (RNNs) (Elman, 1990) were developed. One of the most useful RNNs when addressing long time dependency is the Long-Short Term Memory network (LSTM) (Hochreiter & Schmidhuber, 1997), designed to overcome the

problem of the vanishing gradient. LSTM networks have been applied to different types of problems in chemical engineering including digital twin realization for plant monitoring and control (Qu et al., 2020; Zhu & Ji, 2022). A digital twin (Grieves, 2015) consists of a virtual model of whole plants or individual units which is interconnected with the actual equipment. The availability of reliable digital surrogate models is of great interest in the process industry. This tool exchanges information in real-time, and it can predict the dynamic evolution of a process over time. For this reason, a digital surrogate can provide a useful tool for performance monitoring, predictive asset maintenance, production optimization and advanced process control. This work aims to develop a data-driven digital model for a demethanizer column of a CRR unit, which is realized by using LSTM neural network with a non-conventional network architecture. The main goal is to obtain a surrogate model capable of approximating the real dynamics of the column with a much shorter computation time than the process simulator itself. In addition, it is intended to improve NGL recovery control strategies with real-time estimations of control variables.

## 2. Process Description

The dynamic operation of the CRR unit has been modelled in the process simulator Aspen HYSYS® considering realistic operating conditions (Chebeir et al., 2019). The raw gas is fed to the plant with a nominal flowrate of 4980 kmol/h, pressure and temperature of 5818 kPa and 35 °C, respectively, and a gas composition characterized by a low content of liquids (Chebbi et al., 2010). The main unit of the separation process, responsible for the removal of the light component of the raw gas, is the demethanizer column, given by a 30 stages distillation column with a reboiler. The raw gas entering the CRR unit is first cooled and then the liquid fraction is separated utilizing a flash tank. Three resulting feed streams enter the demethanizer column in the 2<sup>nd</sup>, the 8<sup>th</sup> and the 26<sup>th</sup> trays. The column reflux is provided by compressing part of the demethanizer top product with a cryogenic compressor, ensuing a reflux stream of nearly pure methane. Finally, the remaining part of the column product is recompressed for commercialization while the NGL bottom product of the demethanizer column is sent to further separation in the fractionation train.

## 3. Methods

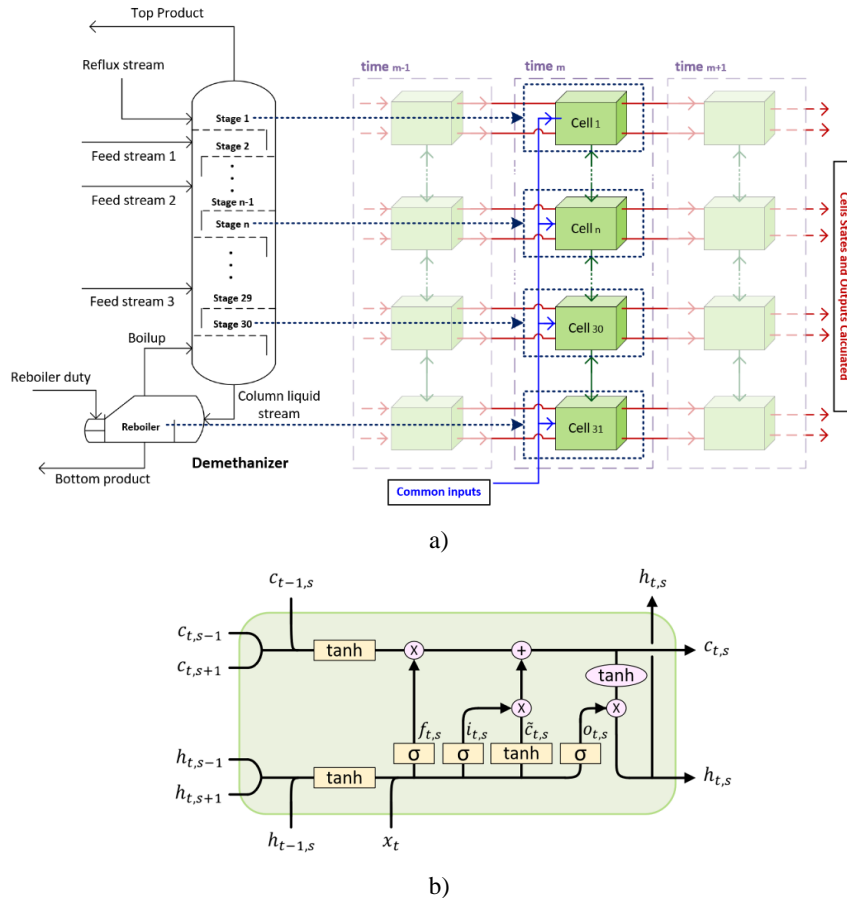
In this work, deep learning methods for the development of RNNs were used to build the surrogate data-driven model. The goals of the neural model were the achievement of accurate estimations for the prediction of the time evolution of temperature and compositions of all column trays and reboiler of the demethanizer column. The considered inputs were given by only easy-to-measure variables: pressure and temperature of all column feeds, reboiler duty, power of the cryogenic compressor and separator pressure and temperature. The selection of RNNs is based on the need to obtain an adaptive neural model capable to capture the existing temporal correlations between the involved variables. In particular, thanks to their ability to overcome the vanishing gradient problem, LSTM networks were used for digital twin realization.

### 3.1. Neural model architecture

In order to design a suitable model to represent the dynamics of the demethanizer column, the LSTM neural networks are arranged considering a particular architecture. The proposed network layout involves the utilization of a dedicated LSTM cell to predict the operations of each stage of the distillation column and the reboiler. Thus, each cell was demanded at modelling the dynamic separation of the corresponding tray, allowing the dynamics inherent in the separation process occurring in each column stage to be captured in the model. Furthermore, bidirectional connections have been applied by modifying the flow of information of a standard LSTM, interconnecting adjacent cells one to another to

account for the interactions between adjacent stages and emulate the action of internal flows within the distillation column. The described model layout, whose structure is schematically represented in Figure 1a, is responsible to approximate the demethanizer operation at a certain instant of time. A schematic representation of a modified LSTM cell is depicted in Figure 1b. As can be seen, in the considered cell (referred to by the subscript  $s$ ), the previous states for the cell are provided by combining the actual previous hidden states and internal cell states ( $h_{t-1,s}$  and  $c_{t-1,s}$ ) with the hidden states and internal cell states calculated by the adjacent LSTM cells (indicated with subscripts  $s - 1$  and  $s + 1$ ) in the current time.

For what concern the terminal cells of the network, mimicking the column top tray and the reboiler, the hidden states from the missing adjacent cell are replaced by the actual measurement of the corresponding variables of the reflux in a case and the liquid stream entering the reboiler in the other.



**Figure 1:** Schematic representation of: overall architecture a) and the modified LSTM cell b).

### 3.2. Datasets

All the data used in the work were generated using the simulated demethanizer column described in section 2 and collected in different datasets for the training, testing and validation campaigns. The datasets employed for training and testing of the neural models, considering a split ratio of 80:20, were two months of simulated plant operation.

Step changes of varying amplitude in the plant feed flow rate were applied with variation peaks of 10% of the plant feed nominal value. To validate the performance of the presented neural model, 2 days of plant operation were simulated by applying step variations of 5% in the plant feed flow rate nominal value.

All the datasets used in this work were registered by considering a sampling time of 20 seconds and without measurement delays, as the data were treated as historical plant data. Moreover, a measurement noise was applied to provide more realistic measured data. The values of the maximum allowed measurement noise applied are shown in Table 1.

**Table 1:** Values of maximum measurement noise considered for input and output data in train, test and validation datasets.

	Duty	Temperatures	Pressures	Concentrations
<b>maximum measurement noise</b>	1.3%*	0.1 [°C]	1.5%*	2%*

\* Referred to the maximum value of the considered variable for the given column stage.

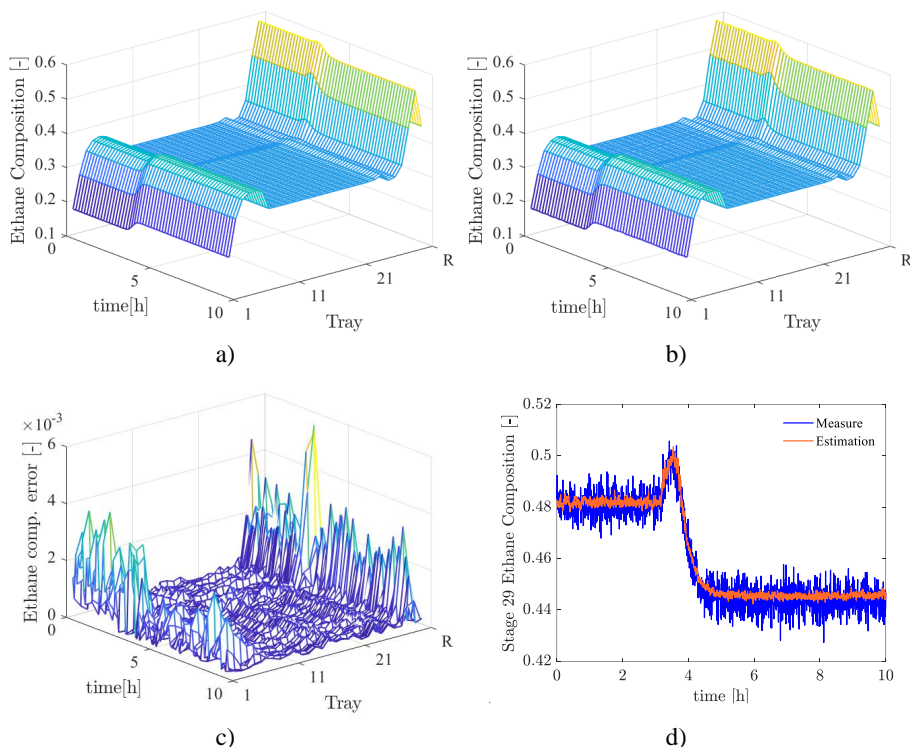
### 3.3. Neural model training

The LSTM neural networks described in section 3.1, have been implemented and trained in the Python programming environment by using the functions available in the open-source library PyTorch. The loss function considered in the optimization problem is given by the Mean Square Error (MSE), while the parameter updating was performed by the Adaptive Moment Estimation optimizer (ADAM) (Kingma & Ba, 2015). The network hyperparameters were obtained by multiobjective optimization using the NSGA-II algorithm available in the pymoo package (Blank & Deb, 2020).

## 4. Results

The performances of the developed neural model are evaluated by considering the column transient profiles predicted by the proposed neural model in the validation campaign for temperature, and key components composition. For sake of brevity, only the results obtained under the worst-case variation represented by the application of a 5% decrease in the feed plant nominal value for ethane composition are reported and shown in Figure 2. The ability of the model to estimate the remaining target variable is assessed by considering the Mean square error (MSE) between the actual and the predicted profiles in the same campaign. Figures 2a and 2b show respectively the actual and estimated ethane transient column profiles. The comparison of these two graphs gives a qualitative indication of the estimation performance of the model for the transient column profile of ethane composition. As it is possible to visualize the proposed model is able to predict the trend of ethane time evolution profile providing an estimation with no detectable deviations from the actual profile. To provide a quantitative measure of the accuracy of the obtained estimation, the estimation error of the column transient profiles for ethane composition is depicted in Figure 2c. Here, it is shown that the proposed neural model can accurately predict the ethane composition transient profiles for all the column trays. The estimation error increase where most of the ethane variations occur, indeed, it increases in the column terminal regions, where most of the separation takes place. The deviations of the predicted transient column profile are always within the maximum error reported in Table 1.

To visualize in detail the estimation obtained, Figure 2d shows the comparison between the estimated ethane composition profile and the measured profile in the tray where the higher estimation error is registered, the 29<sup>th</sup> column tray. As can be seen from the figure, the profile obtained with the neural model can correctly predict the evolution of the composition, with considerably lower variance than the considered measurement.



**Figure 2:** Ethane composition results obtained in the validation campaign under the decreasing variation of plant feed nominal value: actual column profile (a); predicted column profile (b); estimation error profile (c) and transient profiles of 29<sup>th</sup> tray (d).

**Table 2:** Mean square error obtained in the validation campaign under decreasing variation of 5% in the plant feed nominal value.

	Temp.	Methane	Ethane	Propane	Butanes	Pentanes	Hexane
MSE	0.0072	7.33e-7	4.31e-7	7.49e-8	1.22e-8	1.42e-9	2.27e-9

Analogue results are obtained for the temperature and the composition of the remaining components. Table 2 shows the MSE values obtained by considering the predictions of the transient composition profiles in all column trays and the actual profiles, for all the target variables in the validation campaign. As it can be noticed from the obtained MSE value, the use of the proposed model provides an accurate prediction of dynamic temperature profiles. Although it is not apparent from the MSE, it is reasonable to highlight that larger errors have been observed in the last three stages and in the reboiler, where most of the column temperature variation takes place. A possible action for reducing the mismatch is to increase the dataset to be used to train the network, simulating the system in different operating conditions (e.g., impose set-point changes). Concerning the prediction of the remaining components, also in these cases, the model is capable of accurately estimating the evolution over time with estimation errors resulting always lower than the assumed measurement errors, as can be inferred from the obtained MSE values, model's ability to filter the measurement noise is furthermore assessed.



## 5. Conclusions

In this work, a demethanizer column of a simulated CRR unit has been modelled by employing modified bidirectional LSTM neural networks exploring the possibility of its use as digital twins for the column. The results show that the developed neural model can approximate the column trend profiles of the target variables and has a good ability to filtrate measurement noise for the column composition profiles. It also proves to be able to predict the actual trend of the temperature transient profiles quite accurately for most of the column trays and to be a useful tool for temperature monitoring in the final stages of the column and the reboiler. As a result, the presented model may be a suitable candidate as a digital twin of the demethanizer column. On the other hand, the need for concentration measurements yields less attraction to be used as digital twins in a real plant.

## References

- Chebbi, R., Al-Amoody, N.S., Abdel Jabbar, N.M., Husseini, G.A., Al Mazroui, K.A. (2010). Optimum ethane recovery in conventional turboexpander process. *Chem. Eng. Res. Des.* 88, 779-787.
- Chebeir, J., Salas, S.D., Romagnoli, J.A. (2019). Operability assessment on alternative natural gas liquids recovery schemes. *J. Nat. Gas Sci. Eng.* 71, 102974.
- Elman, J. L. (1990). Finding structure in time. *Cognitive science*, 14(2), 179-211.
- Getu, M., Mahadzir, S., Long, N.V.D., Lee, M. (2013). Techno-economic analysis of potential natural gas liquid (NGL) recovery processes under variations of feed compositions. *Chem. Eng. Res. Des.* 91(7), 1272-1283.
- Grieves, M. (2015). Digital Twin: Manufacturing Excellence through Virtual Factory Replication.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- Blank, J., Deb, K. (2020). pymoo: Multi-Objective Optimization in Python, in *IEEE Access*, vol. 8, pp. 89497-89509.
- Kherbeck, L., Chebbi, R. (2015). Optimizing ethane recovery in turboexpander processes. *J. Ind. Eng. Chem.* 21, 292-297.
- Kidnay, A.J., Parrish, W.R., McCartney, D.G. (2011). *Fundamentals of Natural Gas Processing*, 2nd Edition. ed, Fundamentals of Natural Gas Processing. CRC Press, Boca Raton, FL.
- Kingma, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. *ICLR. 2015*. arXiv preprint arXiv:1412.6980, 9.
- Luyben, W.L. (2013). NGL demethanizer control. *Ind. Eng. Chem. Res.* 52, 11626-11638.
- Mandis, M., Baratti, R., Jorge, A., Chebeir, J.A., Tronci, S., Romagnoli, J. A. (2022). Performance assessment of control strategies with application to NGL separation units. *J Natural Gas Sci Eng* vol. 106, 104763.
- Qu, X., Song, Y., Liu, D., Cui, X., & Peng, Y. (2020). Lithium-ion battery performance degradation evaluation in dynamic operating conditions based on a digital twin model. *Microelectronics Reliability*, 114, 113857.
- Tao, F., Zhang, L., Nee, A. Y. C., & Pickl, S. W. (2016). Editorial for the special issue on big data and cloud technology for manufacturing. *The International Journal of Advanced Manufacturing Technology*, 84(1-4), 1-3.
- Zhang, S., Liu, C., Jiang, H., Wei, S., Dai, L., & Hu, Y. (2015). Feedforward sequential memory networks: A new structure to learn long-term dependency. arXiv preprint arXiv:1512.08301.
- Zhu, X., & Ji, Y. (2022). A digital twin-driven method for online quality control in process industry. *The International Journal of Advanced Manufacturing Technology*, 119(5), 3045-3064.

# Soft modelling of spruce conversion into bio-oil through pyrolysis – Note I: steam explosion and LPMO-activated enzymatic saccharification

Matteo Gilardi,<sup>a</sup> Filippo Bisotti,<sup>a,\*</sup> Olaf T. Berglihn,<sup>a</sup> Roman Tschentscher,<sup>b</sup>  
Vincent G.H. Eijssink,<sup>c</sup> Anikó Várnai,<sup>c</sup> Bernd Wittgens,<sup>a</sup>

<sup>a</sup>*SINTEF Industry-Process Technology, Richard Birkelands vei 2B, Trondheim, 7034, Norway*

<sup>b</sup>*SINTEF Industry-Process Technology, Forskningsveien 1, Oslo N-0373, Norway*

<sup>c</sup>*Faculty of Chemistry, Biotechnology and Food Science, NMBU, Chr. Magnus Falsens vei 18, Aas N-1433, Norway*

\* [filippo.bisotti@sintef.no](mailto:filippo.bisotti@sintef.no) (corresponding author)

## Abstract

Biorefinery converts biomass into energy and high-value bioproducts, such as bio-oil. The combined power and fuel generation is particularly attractive for energy transition and reduction of non-biogenic CO<sub>2</sub> emissions. Biomass conversion requires a series of complex treatments. The development of reliable models is essential for scaling up and designing. Soft-modelling approach provides a simplified description of the system without losing accuracy. This contribution presents a soft model for steam explosion (after impregnation with carbocation scavengers) and enzymatic saccharification (boosted with H<sub>2</sub>O<sub>2</sub>-driven oxidative depolymerization by LPMO enzymes). These are two key pre-treatments for the extraction of carbohydrate fraction prior to pyrolysis of lignin-rich solids to produce the bio-oil. The proposed saccharification model accurately predicts experimental data in the range of interest for steam explosion temperature and enzyme concentrations investigated. This approach can be considered general and is potentially applicable to every kind of biomass.

**Keywords:** biorefinery, steam explosion, enzymatic saccharification with LPMO, soft modelling in COCO, softwood cellulose and hemicellulose valorization.

## 1. Introduction and context

Softwood (such as Norwegian spruce) is a good feedstock for biorefinery. The main aim of the biorefinery is to convert the carbon captured by plants into valuable chemicals (fuels, for instance) or heat & power generation systems via several physical and (bio)chemical treatments (Cherubini, 2010). During downstream processing, biomass undergoes a series of pre-treatments to separate the cellulose/hemicellulose from the lignin fraction. Our model includes three main pre-treatment steps: 1) impregnation of spruce biomass with 2-naphthol, a carbocation scavenger that prevents lignin condensation during the subsequent steam explosion (Pielhop et al., 2016a); 2) steam explosion, which allows the breaking of biomass structure so that the hemicellulose and cellulose become accessible while partially solubilizing hemicellulose (Pielhop et al., 2016b); and 3) enzymatic saccharification, where the hemicellulose and cellulose

fractions are depolymerized into their elemental sugars (i.e., pentose and hexose) by selective catalysts (i.e., enzymes) under mild conditions. We selected this process because it selectively solubilizes the carbohydrate fractions and at the same time yields a less condensed and more reactive lignin for further downstream processes (Pielhop et al., 2017; Hansen et al., 2022). The efficiency of the steam explosion and saccharification steps influences the productivity and the overall yield of the complete biorefinery process. Therefore, the present article focuses on modelling these steps. In our model, we take a closer look at the impact of 2-naphthol impregnation on steam explosion as well as LPMO-aided saccharification of Norwegian spruce. Finally, a sensitivity analysis was performed to detect the most influential parameters in designing and modelling softwood biorefineries based on such upstream processing.

## 2. Purpose and material

The present work investigates the direct scale-up and the process configuration of the upstream processes in a biorefinery described by (Hansen et al. 2022), via direct translation of the reported laboratory-scale biomass treatments. The selected process layout (Figure 1) is based on the experimental setup by (Hansen et al. 2022) with the following modifications: (1) after the steam explosion, the pH of the biomass slurry was set to 5.0 (pH optimum of the cellulases in the subsequent saccharification step) in a single step using NaOH, instead of using an acetic acid/acetate buffer solution and then a caustic solution, for scalability. (2) A filtration stage was added after the saccharification step to separate the lignin-rich saccharification residue (solids) from the sugar solution (liquid/aqueous phase). This step is expected to reduce the residual moisture content of the lignin-rich residue and thus, lead to energy savings in the drying step before the pyrolysis unit.

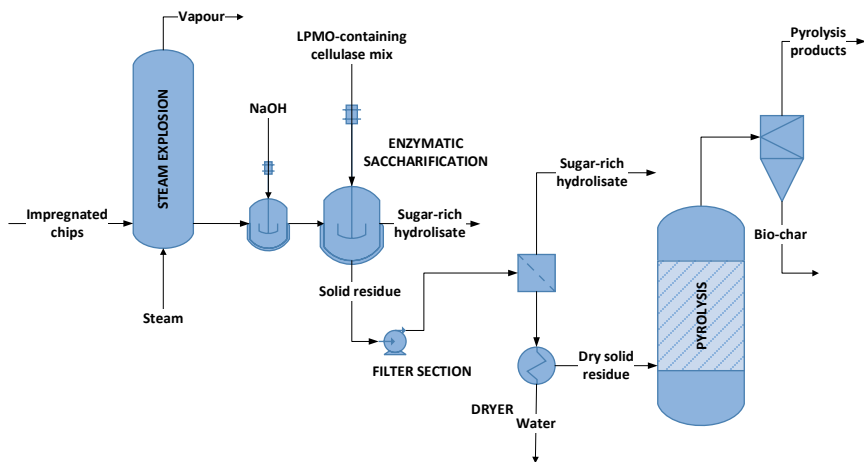


Figure 1 - Simplified layout of the investigated upstream processes of a softwood-based biorefinery. The picture does not include the impregnation step prior to the steam explosion.

## 3. Implementation

The steam explosion and enzymatic saccharification steps were modelled using COCO (a CAPE-open solution) simulation software (Figure 2). Due to the large complexities of bioprocesses, we adopted a soft-model approach instead of a first-principle one. The missing properties (i.e., enthalpy of formation, heat capacity, vapour pressure, and others)

for some components, namely the cellulose, hemicellulose, and lignin constituents, have been estimated using Gani’s methods (Constantinou and Gani, 1994). For the steam explosion step, we implemented a simplified kinetics scheme proposed by Lam (2009 and 2010), which has been suggested to be applicable for softwood in general. Due to the lack of measurements and data points collected during steam explosion, we do not have experimental data to validate the implemented model. For the enzymatic saccharification, on the other hand, the general expression of the soft model was taken from (Theuretzbacher et al., 2015), and it was validated and adjusted using the experimental saccharification data in (Hansen et al., 2022). The proposed model for saccharification conversion fits the experimental data (Figure 3), and the conversion model is a function of the steam explosion temperature ( $T_{SE}$  in  $^{\circ}C$ ) and overall Cellic CTec2 enzyme concentration  $[E]$  expressed in  $[mg/g_{biomass}]$ , as in (Hansen et al., 2022).

$$\gamma(t) = \gamma(0) + h \cdot [1 - \exp(-k_r \cdot t)] \quad (1)$$

Where  $\gamma$  is the conversion at the outlet of the saccharification reactor,  $\gamma(0)$  the conversion at the inlet (called “instantaneous”; in Equation 2),  $h$  an adaptative parameter accounting for the conversion “jump” from the inlet and outlet conditions for ideal infinite residence time (see Equation 3),  $k_r$  the reaction constant (see Equation 4), and  $t$  the residence time.

$$\gamma(0) = 0.8607 \cdot [E] + 4.6 \quad (2)$$

$$h(T_{SE}, [E]) = (0.435 \cdot T_{SE} - 88.227) \cdot [E] + (0.15 \cdot T_{SE} - 24.275) \quad (3)$$

$$k_r(T_{SE}, [E]) = (1.3 \cdot 10^{-4} \cdot T_{SE} - 0.0355) \cdot [E] + (-4.17 \cdot 10^{-3} \cdot T_{SE} + 1.0937) \quad (4)$$

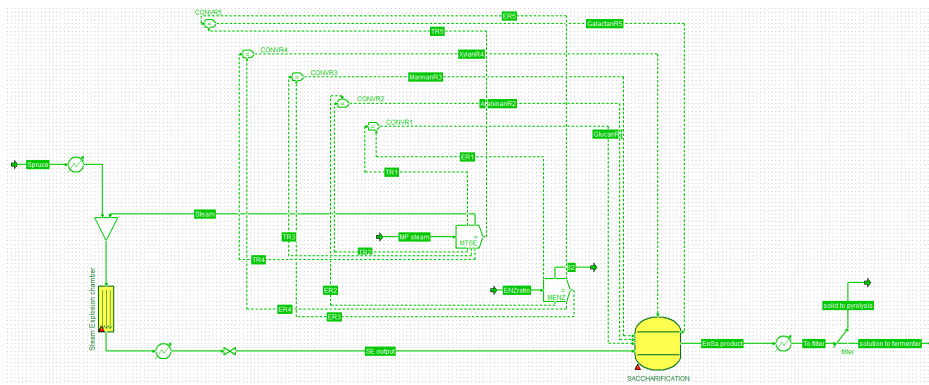


Figure 2 - Flowsheet for the steam explosion and enzymatic saccharification in COCO simulator. Yellow color indicates that the unit is handling a biphasic system of liquid and solid.

The saccharification curves in Figure 3 show that maximum glucose conversion is reached when the biomass undergoes a high-temperature steam explosion ( $220^{\circ}C$ ) and at the highest tested enzyme load ( $8 \text{ g/kg}_{biomass}$ ). The optimal residence time lies around 24 h, after which the saccharification yield reaches a plateau. Thus, longer saccharification times lead to a larger volume without substantial benefit to the process yield. As cellulose and hemicelluloses are depolymerized concomitantly by enzymes (Várnai et al., 2011), a similar yield trend is expected for the saccharification of the cellulose and hemicellulose fractions. Using the models here defined, the steam explosion chamber was simulated as a plug flow reactor and the saccharification reactor (which resembles a CSTR reactor with a large residence time) as a system at assigned conversion. The conversion was estimated

according to the steam explosion operating temperature (i.e., steam) and the concentration of the LPMO-rich enzyme cocktail fed into the saccharification reactor.

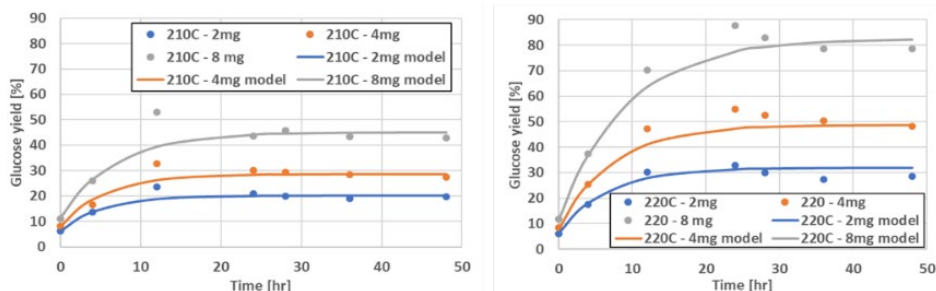


Figure 3 - Conversion model for enzymatic saccharification: experimental data (dots) and model (solid line) after the steam explosion (210°C–left and 220°C–right) using different enzyme doses expressed as  $\text{mg}_{\text{enzyme}}/\text{g}_{\text{biomass}}$  (2–blue, 4–orange, 8–grey).

#### 4. Sensitivity analysis and results

Additional parameters for the models were set as follows. The Norwegian spruce composition was adapted from (Wang et al., 2018). Hansen et al. report a residual humidity of 5 wt%. The steam explosion process is supposed to treat 100 kg/h of wet-impregnated spruce. The process requires steam at saturation pressure, corresponding to temperatures ranging from 210°C (18.5 bar) to 220°C (22.5 bar). The ratio between biomass and steam is 1:1 in mass. The steam explosion chamber volume has been set to 29 l (litres) and 58 l; with a fixed diameter of 0.3 m, hence the volume is directly proportional to the residence time. We performed a sensitivity analysis on the steam explosion temperature at two reactor volume levels to assess the impact of the operating conditions of this pre-treatment process on the overall upstream processing steps. Table 1 gathers the different case studies.

Table 1 - Case studies proposed in this work for the sensitivity analysis.

Case	Steam explosion temperature [°C]	Steam explosion reactor volume [l]
1	210	29
2	210	58
3	220	29
4	220	58

Table 2 gathers the increment or decrement in the mass yield of each product with respect to Case 1 (indicated in Table 1). Our main observations are that: (1) remarkably, the lignin fraction seems inert to both steam explosion and saccharification, thus the pyrolysis process is not affected; and (2) the composition of aqueous sugar solution depends on the steam explosion operating conditions. The latter one is partly because hemicellulose (arabinoxylan and galactoglucomannan) hydrolysis starts taking place already during steam explosion, depending on the severity. In Lam's model, hemicellulose hydrolysis is defined as a conversion to sugar (hexose and pentose) monomers and their corresponding degradation products HMF (hydroxymethyl-furfural) and furfural, respectively; the mass flow rate of the sugar monomer or secondary product being derived from the flowrate of its original corresponding polymer. The more severe the steam explosion conditions (i.e., high temperature and pressure) are, the larger the sugar yield from the cellulose and

hemicellulose fractions is after saccharification. Both HMF and furfural are valuable for bio-oil production. We adopted Case 1 as a reference because it represents the operating condition with the lowest yield, but also the one associated with the lowest costs. Figure 4 depicts the products' mass composition at the steam explosion outlet and enzymatic saccharification reactors in the proposed case studies.

Table 2 - Predicted conversion of hemicellulose, increment (positive) or decrement (negative values) in pentose and hexose sugars mass (on biomass dry matter basis), at the outlet of the enzymatic saccharification reactor using Case 1 as the reference case.

Case	Arabinose	Galactose	Xylose	Mannose	HMF	Furfural
2	+7.61%	-94.28%	-95.66%	+14.61%	+12.99%	+7.65%
3	+6.74%	-44.47%	-56.53%	+15.04%	+7.31%	+4.52%
4	+8.72%	-98.65%	-99.20%	+19.90%	+13.67%	+7.92%

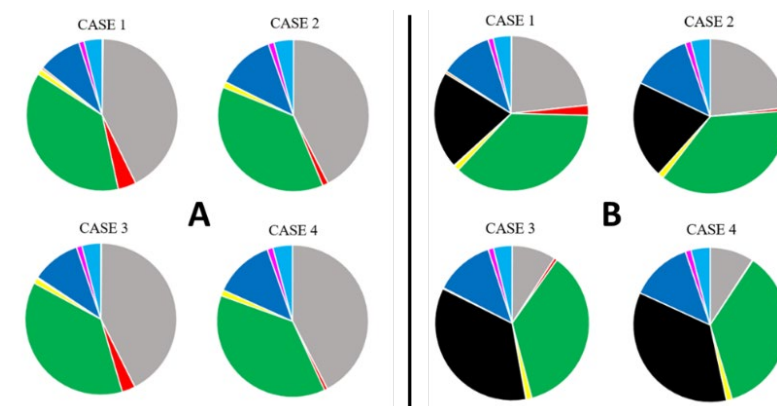


Figure 4 - Mass composition (dry weight basis) of the sugar solution obtained at the outlet of the steam explosion (A) and enzymatic saccharification (B) reactors in the four analyzed case studies (CASE). Legend: **lignin** (green), **glucan/cellulose** (grey), **glucose** (black), **mannan** (red), **mannose** (dark blue), **arabinose** (yellow), **HMF** (magenta), and **furfural** (light blue). Not mentioned species (arabinan, xylan, xylose, galactan, and galactose) have negligible content (< 0.1 wt%).

## 5. Conclusions and comments

The results show that the biomass pre-treatment model enables the scale-up of the impregnation–steam explosion–enzymatic saccharification process from laboratory to pilot scale while leading to the identification of the main parameters determining the selective and nearly complete removal of the polysaccharide fraction. At the projected scale, the steam explosion was the most determining step of the studied process setup since it had a direct influence on the subsequent enzymatic saccharification step. Correct tuning of the steam explosion operating conditions led to an optimal process yield. The hemicellulose and cellulose fractions were almost depleted after steam explosion and enzymatic saccharification. Hemicellulose (arabinoxylan and galactoglucomannan) was mainly degraded after steam explosion. The lignin fraction did not seem to undergo major chemical/physical transformation due to the impregnation pre-treatment. This emerges both from the experimental data and the model. Lignin remains in the insoluble fraction in the investigated process since its extraction and valorization require stronger operative conditions as in pyrolysis. The predicted outlet compositions from the steam explosion

and enzymatic saccharification reactors at different steam explosion temperatures (210°C and 220°C) and reactor volumes (29 l and 58 l) point to the following conclusions: (1) A proper design of steam explosion and enzymatic saccharification reactors can lead to a very efficient decomposition of the hemicellulose and cellulose fractions of biomass, respectively, thus generating a lignin-rich solid residue. (2) Increasing steam explosion temperature and reactor volume is a good strategy to substantially improve the solubilization of the polysaccharide content. (3) Glucomannan removal during the steam explosion is significantly slower than arabinoxylan (or galactoglucomannan) removal as the greatest yield increase is observable for the mannose. (3) Larger residence times in the steam explosion chamber are associated with the formation of substantial amounts of secondary products (HMF and furfural). The temperature increment is not enough to fully deplete the intermediate products (galactose and xylose). Even though these are by-products generated from successive pentose and hexose sugars oxidation, they are valuable chemicals which can be blended in biodiesel. (4) High-temperature steam explosion benefits the glucan conversion occurring within enzymatic saccharification. If the temperature is shifted from 210°C to 220°C, the global final glucose yield rises from 50.8% up to 87.7%. On the contrary, steam explosion reactor volume does not impact the performance of downstream unit operations.

### Acknowledgements

This work was supported by the Research Council of Norway through grants no. 257622 (Bio4Fuels) and 268002 (Enzymes4Fuels).

### References

- F. Cherubini, 2010, The biorefinery concept: Using biomass instead of oil for producing energy and chemicals, *Energy Conversion and Management*, 51(7), 1412-1421
- L. Constantinou et R. Gani, 1994, New group contribution method for estimating properties of pure compounds, *AIChE Journal*, 40(10), 1697-1710
- L.D. Hansen et al., 2022, 2-Naphthol impregnation prior to steam explosion promotes LPMO-assisted enzymatic saccharification of spruce and yields high-purity lignin. *ACS Sustainable Chemistry & Engineering*; 10(16), 5233-5242
- P. Kumar et al., Lignocellulose biomass pyrolysis for bio-oil production: A review of biomass pretreatment methods for production of drop-in fuels, *Renewable and Sustainable Energy Reviews*, 123, 109763
- P.S.W. Lam, 2009, Kinetic modeling of pseudolignin formation in steam exploded woody biomass, Conference proceeding of 8th world congress of chemical engineering, Montreal, Canada, 23-27 August 2009
- P.S.W. Lam, 2010, Chapter 5 – Modelling of polysaccharides depolymerization and pseudolignin formation of softwood during steam explosion, PhD Thesis
- T. Pielhop et al., 2016a, Autohydrolysis pretreatment of softwood - enhancement by phenolic additives and the effects of other compounds. *Green Chemistry*, 18, 5239-5247
- T. Pielhop et al., 2016b, Steam explosion pretreatment of softwood: the effect of the explosive decompression on enzymatic digestibility. *Biotechnology for Biofuels*, 9, 152
- T. Pielhop et al., 2017, Pilot-scale steam explosion pretreatment with 2-naphthol to overcome high softwood recalcitrance. *Biotechnology for Biofuels*, 10(1), 130
- F. Theuretzbacher et al., 2015, Steam explosion pretreatment of wheat straw to improve methane yields: Investigation of the degradation kinetics of structural compounds during anaerobic digestion, *Bioresource Technology*, 179, 299-305
- A. Várnai, et al., 2011, Synergistic action of xylanase and mannanase improves the total hydrolysis of softwood. *Bioresource Technology*, 102(19), 9096-9104
- Z. Wang et al., 2018, Chemical and structural factors influencing enzymatic saccharification of wood from aspen, birch and spruce, *Biomass and Bioenergy*, 109, 125-134

# Molecular Modelling and Optimization of Diesel Hydrotreating Processes

Xiaolin Bi,<sup>a</sup> Nan Zhang,<sup>a</sup> Robin Smith,<sup>a</sup>

*<sup>a</sup>Centre for Process Integration, Department of Chemical Engineering, University of Manchester, Oxford Road, Manchester, M13 9PL, UK*

## Abstract

Diesel is an important transportation fuel produced in oil refineries, for which the quality should be strictly controlled to meet the diesel engine requirement and minimize associated air pollution issues. A key refining process to improve diesel quality is the diesel hydrotreating (DHT) process. The molecular type homologous series (MTHS) matrix is developed for the molecular characterization of petroleum fractions. The kinetic modelling and reaction modelling are built and improved to discuss the influence of operating conditions on the DHT processes. This work innovates to combine the catalyst deactivation model with the kinetic model to continuously monitor the catalyst deactivation during the DHT processes, which is essential to optimize the reaction operating conditions and maintain constant product yields and quality.

**Keywords:** Diesel Hydrotreating Processes, Optimization, Molecular Characterization, Catalyst Deactivation.

## 1. Introduction

Diesel hydrotreating (DHT) has become one of the essential fundamental processes in the petroleum-refining industry from environmental, technical, and economic points of view for over 60 years (Hoehn et al., 2017). The DHT process is mainly to obtain products with high quality and low content of pollution. In order to fulfill the applicable regulations of product quality, and remove the undesirable compounds, including but not solely limited to sulfur, nitrogen, olefins, and aromatics, the operating conditions of the reactor need to be improved. Establishing reaction modelling and simulation is a useful way to have deep research on phenomena occurring in the DHT reactor. Attributed to effective modelling and simulations, operating conditions, reactor design, catalyst formulation and process configuration would be optimized. In addition, effects of operating variables could be investigated and integration with other technologies could be implemented. (Mederos et al., 2012).

The loss of catalytic activity is a problem of great and continuing concern in the practice of industrial catalytic processes. During the industrial production process, with the extension of the operation time of the unit, the catalyst of hydrodesulfurization will experience a decline in activity. The degree of catalyst deactivation depends mainly on the operating conditions and the properties of the feed (Kohli et al., 2016). The loss of catalytic activity must be compensated by periodic increases in the reaction temperature in order to maintain the constant product yields and quality (Moghadassi et al., 2011). When the catalyst is deactivated or the catalyst activity drops below the economical operation of the unit, it is necessary to shut down the operation for catalyst regeneration



or replacement with a new catalyst, which directly affects the economic benefits of the refinery. The annual cost to the industry for catalyst replacement and process shutdown is in the billions of dollars (Bartholomew et al., 2001). However, catalyst deactivation is inevitable. Therefore, it is significant to develop the catalyst deactivation model to investigate the catalyst deactivation regularity under ultra-deep desulfurization of diesel fuel, which shows practical guidance for refineries on predicting the catalyst usage periods and rationalising the production cycles of units.

In this research, the reaction network of a diesel hydrotreater is generated by incorporating most of the reaction types in this process, including hydrodesulfurization (HDS), hydrodenitrogenation (HDN), hydrodearomatization (HDA), and hydrocracking (HC). The current kinetic model of hydrodesulfurization has been proved by Vanrysselberghe et al. (1998) that expressed the reaction for dibenzothiophenes, which is the hardest sulfur component to remove. The three-phase hydrotreater model developed by Korsten and Hoffmann (1996) can be used in this research which contains a combination of algebraic and ordinary differential equations. The fourth-order Runge-Kutta method is used to numerically solve the first-order differential equations in GAMS.

## 2. Methodology

### 2.1. Molecular Type Homologous Series (MTHS) Method

The molecular type and homologous series method is used to characterize petroleum streams quickly and accurately at the molecular level. Pseudo-components could be generated based on their corresponding distillation profile and density using the pseudo-component approach, where each of them has specific boiling point, density and volume fraction. The molecular composition within each homologous series follows the gamma distribution shown in equations 1 and 2. The MTHS matrix is shown in Figure 1.

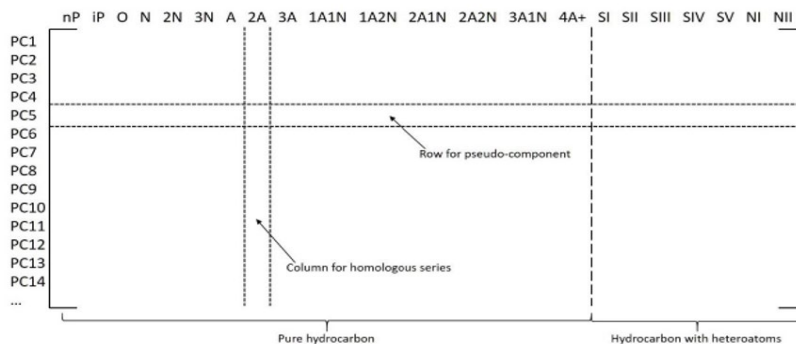


Figure 1 Pseudo-components based MTHS matrix

$$PDF(P) = \frac{(P-\eta)^{\alpha-1} \exp(-\frac{P-\eta}{\beta})}{\beta^{\alpha} \Gamma(\alpha)} \quad (1)$$

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt \quad (2)$$

where P is the property;  $\alpha$ ,  $\beta$ ,  $\eta$  are three parameters of gamma distribution; PDF is the probability density function.

The objective function is to minimize the difference between the measured data and the predicted data:

$$OBJ = \sum_V (W_1 \cdot \frac{T_V^{msd} - T_V^{pred}}{T_V^{msd}})^2 + \sum_{PC} (W_2 \cdot \frac{BP_{PC}^{msd} - BP_{PC}^{pred}}{BP_{PC}^{msd}})^2 + \sum_P (W_3 \cdot \frac{P_P^{msd} - P_P^{pred}}{P_P^{msd}})^2 + \sum_{f \in PIONA} (W_4 \cdot \frac{C_f^{msd} - C_f^{pred}}{C_f^{msd}})^2 + \sum_{e \in (S, N, O)} (W_5 \cdot \frac{X_e^{msd} - X_e^{pred}}{X_e^{msd}})^2 \quad (3)$$

where,  $T$  is the boiling point of the distillation curve;  $BP_{PC}$  is the boiling point of the pseudo-components;  $P$  is bulk properties;  $C$  is the composition of PIONA;  $X$  is heteroatoms content;  $W$  is the weighting factors for each type of the properties;  $msd$  is the measured values;  $pred$  is the predicted results based on the MTHS.

### 2.2. Kinetic Model of Hydrodesulfurization (HDS)

The kinetic model of HDS for the different sulfur compounds has been proposed by previous researchers. The Hougen-Watson rate equation is used to describe the hydrogenation and hydrogenolysis in the HDS process (Vanrysselberghe et al., 1998):

$$r_s = \frac{k_{s,\sigma} K_{H,\sigma} K_{S,\sigma} C_s C_{H_2}}{DEN_\sigma} + \frac{k_{s,\tau} K_{H,\tau} K_{S,\tau} C_s C_{H_2}}{DEN_\tau} \quad (4)$$

With

$$DEN_\sigma = (1 + \sum_i K_{i,\sigma} C_i + \sqrt{K_{H,\sigma} C_{H_2}})^3 \quad (5)$$

$$DEN_\tau = (1 + \sum_i K_{i,\tau} C_i + \sqrt{K_{H,\tau} C_{H_2}})^3 \quad (6)$$

where,  $r_s$  is the reaction rate of the sulfur compound;  $\sigma$ ,  $\tau$  are hydrogenolysis site and hydrogenation site;  $k_{s,\sigma}$  and  $k_{s,\tau}$  are the rate coefficients for the reactions of the sulfur compound on the two sites;  $K_{i,\sigma}$  and  $K_{i,\tau}$  are the adsorption coefficient of component  $i$  on the two sites;  $C_i$  is the liquid concentration of the component  $i$ .

### 2.3. Modelling of Three-Phase Diesel Hydrotreater

The three-phase catalytic reactor, the heart of the HDT unit, is the equipment that needs to be modelled. In HDT reactors, the gas phases ( $H_2$ ) and liquid phases (hydrocarbons) are contacted with the solid phase (catalyst). The reactions occur between the dissolved gaseous reactants and the liquid phase reactants on the surface of the catalyst (Mederos and Ancheyta, 2007). The reactor model created by Korsten and Hoffmann (1996) is used based on the following assumptions:

- The reactor operates under steady state conditions.
- No radial concentration gradients.
- Constant volumetric flowrates through the reactor.
- Constant pressure through the reactor.
- The catalyst particles are completely wetted.

#### 2.3.1. Mass Balance

For component  $i$  of the gas phase:

$$\frac{u_G}{RT_G} \frac{\partial p_i^G}{\partial z} + k_i^G a_L \left( \frac{p_i^G}{H_i} - C_i^L \right) = 0 \quad (7)$$

$$z = 0, p_i^G = p_i^{G,0}$$

For gaseous components in the liquid phase:

$$u_L \frac{dC_i^L}{dz} + k_i^L a_L \left( \frac{p_i^G}{H_i} - C_i^L \right) + k_i^S a_S (C_i^L - C_i^S) = 0 \quad (8)$$

$$z = 0, C_i^L = C_i^{L,0}$$

At catalyst surface:

$$k_i^S a_S (C_i^L - C_i^S) = -v_i \rho_B \xi \eta r_i \quad (9)$$

Where,  $u_G$  is the superficial gas velocity;  $p_i^G$  is the partial pressure of gaseous  $i$ ;  $Z$  is the length of the reactor;  $k_i^G a_L$  and  $k_i^S a_S$  are the gas-liquid liquid-solid mass-transfer coefficient;  $H_i$  is Henry's law coefficient;  $C_i^L$  and  $C_i^S$  are concentrations of component  $i$

in the liquid and on the surface of the catalyst;  $\eta$  is the effectiveness factor;  $r_i$  is the reaction rate.

### 2.3.2. Energy Balance

$$\frac{dT_R}{dz} = \frac{\xi n}{\rho_{BuL} C_p} \sum_{i=1}^n r_i (-\Delta H_i) \quad (10)$$

where,  $C_p$  is the specific heat capacity of the feed;  $T_R$  is the reactor temperature.

### 2.4. Catalyst Deactivation Model

Dong and Gao (2009) analyzed the sulfur content of the product oil samples at different times of catalyst operation to calculate the corresponding hydrodesulfurization reaction rate constants. The relative activity of the catalyst is defined as the ratio of the reaction rate constant at different times of operation to the initial reaction rate constant. The catalyst deactivation models are established as follows:

$$k = A \exp(-E_a/RT) \quad (11)$$

$$A = (1 - nt)A_0 \quad (12)$$

$$k = (1 - nt)k_0 \quad (13)$$

$$n = \rho^{61.305} T_{90\%}^{0.136} f_{Aro}^{0.111} \quad (14)$$

where,  $k$  is reaction rate constant;  $A$  is constant;  $E_a$  is activation energy;  $n$  is inactivation rate constant;  $f$  is total content of bicyclic and tricyclic aromatic hydrocarbons and colloid ( $\omega\%$ ).

## 3. Case study

The bulk properties and composition information of the SRGO stream are given in Table 1.

Table 1. Bulk properties of SRGO stream

Density at 15°C (kg/m <sup>3</sup> )	851.7	Distillation (D86)/ °C	
Sulfur content (m%)	1.51	IBP	220
Paraffin (m%)	44.4	10 vol%	279
Naphthene (m%)	21.9	30 vol%	295
Aromatics (m%)	17.3	50 vol%	310
Multi-ring aromatics (m%)	7.36	70 vol%	328
		90 vol%	359
		95 vol%	376
		FBP	379

The configuration of the reactor and its operating conditions are shown in the Table 2.

Table 2. Configuration of the reactor and operating conditions

Reactor configuration		Operating conditions	
Diameter (cm)	2.54	Temperature (K)	633
Length (cm)	45	Pressure (bar)	6.4
Density of bed ( $kg_{cat}/m^3$ )	780	H <sub>2</sub> /oil	500
Catalyst characteristics		LHSV (hr <sup>-1</sup> )	1.5
Equivalent diameter (m)	$1.30 \times 10^{-3}$	Hydrogen purity (%)	100
Density (kg/m <sup>3</sup> )	1420		
Specific surface area (m <sup>2</sup> /g)	200		

## 4. Results and Discussion

### 4.1. Feed Characterization

SRGO stream is divided into 25 pseudo-components, each having a 10K gap between 448K and 688K. Sulfur compounds are classified into 5 homologous series: SI (thiophene), SII (benzothiophene), SIII (dibenzothiophene), SIV (4-methylthiophene) and SV (4, 6-dimethylthiophene). The distribution of different types of sulfur compounds: thiophenes (TH) are commonly found at temperatures below 220°C, benzothiophenes (BT) make up the majority of the boiling range between 220°C and 257°C, and dibenzothiophenes (DBT) are substantially concentrated at temperatures above 330°C. Figure 2 and Table 3 show the comparison of distillation profiles and properties between the measured and predicted results. The biggest deviation of distillation profiles is at this point when the cumulative volume is 10%, and the biggest deviation of the properties is the content of multi-ring aromatics. These errors are still within acceptable limits.

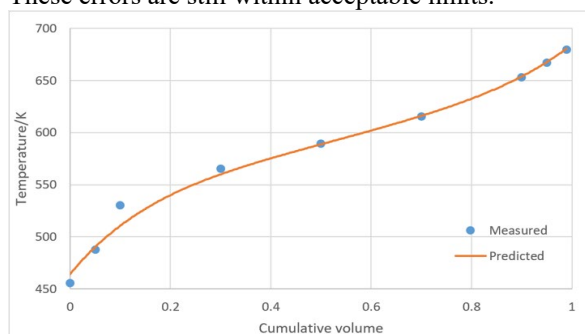


Figure 2. Comparison of distillation profiles between the measured and the predicted

Table 3. Comparison of the properties between the measured and predicted

Composition & Property	Measured	Predicted	Relative Error (%)
Density, kg/m <sup>3</sup>	851.7	850.8	-0.11
Paraffin(m%)	44.4	43.13	-2.86
Naphthene (m%)	21.9	22.4	2.15
Aromatics (m%)	17.3	18.14	4.93
Multi-ring aromatics (m%)	7.36	7.8	5.83
Sulfur content (m%)	1.51	1.514	0.26

### 4.2. The Pilot Plant Diesel Hydrotreater

The comparison of compositions and properties of the hydrotreated product between the measured and the predicted at the operating condition of 633K, 6.4bar and 1.5 hr-1 are shown in Table 4. The relative error between predicted results and measured results for the nitrogen content is slightly larger, but it is still within an acceptable range. The errors between the predicted and measured for other content are lower than 5%.

Table 4. Composition and properties of the measured and predicted product

	Measured	Predicted	Error (%)
Paraffin (m%)	55.2	53.4	-3.26
Naphthene (m%)	31.8	31.2	-1.89
Aromatics (m%)	12.9	13.4	3.88
N (wtppm)	8.77	9.29	5.93
S (wtppm)	6.02	6.16	2.32

The rising trend of temperature in the axial position of the reactor is shown in Figure 3. It can be seen from the figure that the temperature at the inlet of the reactor rises significantly, and the temperature is in a steady increase after 5% axial of the reactor. This is because the sulfides, especially thiophene and benzothiophene, react rapidly at the inlet of the reactor. Therefore, the temperature also improves rapidly.

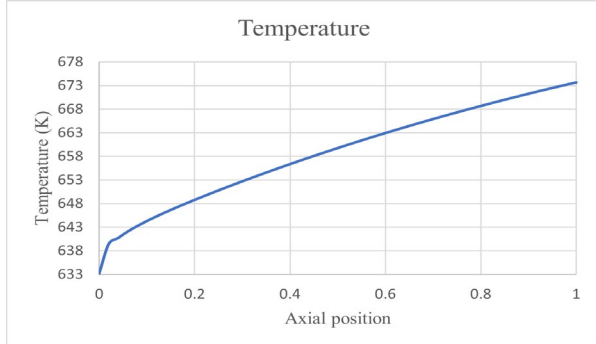


Figure 3. Temperature raise trend in the axial position of the reactor

## 5. Conclusion

In this research, the kinetic model and the mass balance of the three-phase diesel hydrotreater model are established based on the MTHS method. The feasibility of the model is confirmed in the case study by comparing the predicted results with the measured results, which is the basis for further work. Future work will include the optimization of the current modelling and the monitor continuously of the catalyst deactivation.

## References

- C. H. Bartholomew, 2001, Mechanisms of catalyst deactivation, *Applied Catalysis A: General*, 212(1-2), 17-60.
- K. Dong, X. Gao, 2009, Deactivation model of hydrotreating catalyst for ultra low sulfur diesel fuel production, *CHKI* 40(7).
- R. Hoehn, V. Thakkar, E. Yuh, 2017, *Hydroprocessing for clean energy: design, operation, and optimization*, John Wiley & Sons.
- K. Kohli, R. Prajapati, S. K. Maity, M. Sau, M. O. Garg, 2016, Deactivation of hydrotreating catalyst by metals in resin and asphaltene parts of heavy oil and residues, *Fuel*, 175, 264-273.
- H. Korsten, U. Hoffmann, 1996, Three-phase reactor model for hydrotreating in pilot trickle-bed reactors, *AIChE J.*, 42, p. 1350.
- F. S. Mederos, J. Ancheyta, 2007, Mathematical modeling and simulation of hydrotreating reactors: Cocurrent versus countercurrent operations, *Applied Catalysis A: General*, 332(1), 8-21.
- F. S. Mederos, J. Ancheyta, I. Elizalde, 2012, Dynamic modeling and simulation of hydrotreating of gas oil obtained from heavy crude oil, *Applied Catalysis A: General*, 425, 13-27.
- A. Moghadassi, N. Amini, O. Fadavi, M. Bahmani, 2011, Hydrocracking lumped kinetic model with catalyst deactivation in arak refinery hydrocracker unit.
- V. Vanrysselberghe, R. Le Gall, G. F. Froment, 1998, Hydrodesulfurization of 4-methylthiophene and 4, 6-dimethylthiophene on a CoMo/Al<sub>2</sub>O<sub>3</sub> catalyst: Reaction network and kinetics, *Industrial & engineering chemistry research*, 37(4), 1235-1242.

## Soft modelling of spruce conversion into bio-oil through pyrolysis – Note II: pyrolysis

Filippo Bisotti,<sup>a,\*</sup> Matteo Gilardi,<sup>a</sup> Olaf T. Berglihn,<sup>a</sup> Roman Tschentscher,<sup>b</sup>  
Vincent G.H. Eijssink,<sup>c</sup> Anikó Várnai,<sup>c</sup> Bernd Wittgens,<sup>a</sup>

<sup>a</sup>*SINTEF Industry-Process Technology, Richard Birkelands vei 2B, Trondheim, 7034, Norway*

<sup>b</sup>*SINTEF Industry-Process Technology, Forskningsveien 1, Oslo N-0373, Norway*

<sup>c</sup>*Faculty of Chemistry, Biotechnology and Food Science, NMBU, Chr. Magnus Falsens vei 18, Aas N-1433, Norway*

\* [filippo.bisotti@sintef.no](mailto:filippo.bisotti@sintef.no) (corresponding author)

### Abstract

Biorefinery converts biomass into valuable bioproducts. Cellulose and hemicellulose are degraded into simple sugars during enzymatic saccharification, while lignin remains unutilized. Lignin and residual cellulose can be subjected to pyrolysis, a key biorefinery step for improved valorization of biomass. During biomass pyrolysis thousands of compounds are formed, thus, a simplified approach must be adopted for the sake of modelling. Lumped kinetic models address a simplified description of the system by catching relevant features of the product mixture. This contribution presents a pyrolysis model where two previously validated kinetics are merged to get a more effective description of the product distribution. The model provides a quantitative characterization of the main product classes (aldehydes, carboxylic acids, phenols, furans, etc.) and can be used as a numerical tool for roughly scaling up pyrolysis reactors in a softwood-based biorefinery. This approach can be considered general and is applicable to all biomass types.

**Keywords:** biorefinery, pyrolysis, lumped kinetic, lignin valorization, bio-oil production.

### 1. Introduction

Biomass is composed of cellulose, hemicellulose, and lignin. Biochemical conversion of biomass, most commonly including steam explosion and enzymatic saccharification processes upstream, selectively depolymerizes and solubilizes the cellulose and hemicellulose fractions, leaving a lignin-rich solid residue. To increase the biorefinery yield, the lignin-rich residue can be subjected to thermochemical treatment to produce syngas and chemicals via gasification, heat and power via combustion, or bio-oil and biochar via pyrolysis (Kumar et al., 2020; Cherubini, 2010). The present work deals with the modelling of the pyrolysis of lignin-rich residue remaining after steam explosion and enzymatic saccharification of Norwegian spruce. An explicit, rigorous and descriptive model of the pyrolysis process, including thousands of species (considering also all intermediates) and reactions, would lead to a stiff and complex system, as exemplified by a few case studies where detailed kinetics have been adopted (e.g., Corbetta et al., 2015; Ranzi et al., 2014). Such models cannot be implemented in process simulators due to the

lack of thermodynamic properties for most species present in the kinetics and due to the huge dimensions of nonlinear equations which challenge the numerical stability of the software itself (Manca et al., 2001). On the other hand, the lumped kinetics approach brings a substantial simplification of the kinetic schemes without losing accuracy (Ranzi et al., 2008). Lumped kinetics approach entails the simplification of the schemes to a limited number of chemical species (generally less than one hundred) and reactions (less than fifty), which can be handled by simulation software. In this work, we used the lumped kinetics over experimental pyrolysis data from Hansen (Hansen et al., 2022) and verified its capacity to predict product mixture composition.

## 2. Materials and methods

The adopted lumped kinetics describing the pyrolysis of lignin-rich residue is composed of two different kinetics models. The first part describing the conversion of the residual cellulose and hemicellulose fractions was adapted from the model proposed by (Ranzi et al. 2008). The second part describing the conversion of the lignin fraction, on the other hand, was adapted from the kinetics model proposed by Dussan (Dussan et al. 2018). The latter, the lignin-specific part of the model, is composed of four pseudo-lignin building blocks (LIG-1 to LIG-4) as in the model by Dussan (Dussan et al. 2008) while the stoichiometry of lignin pyrolysis reactions was modified. Overall, our lumped kinetics model contained 16 reactions and 49 species describing lignin pyrolysis. The kinetics model has been implemented in MATLAB 2022a code. The pyrolysis reactions occur in a fluidized bed reactor, where solid particles are circulated by the generated gas, and the gas itself guarantees continuous mixing inside the reactor. Thus, we modelled the fluidized bed as an isothermal plug flow reactor (PFR), with the following assumptions to simplify the calculations: the generated gas drags and entrains the small solid particles in a plug flow, while the gas and solid phases are vigorously mixed in the system, and this leads to almost isothermal conditions in the reactor volume. These simplifications provided us with a general simplified model that could be incorporated into a COCO simulation flowsheet, instead of using a descriptive and compute-intensive computational fluid dynamics (CFD) model. Furthermore, this work aims at the preliminary validation of the lumped kinetic model using experimental data from (Hansen et al. 2022).

## 3. Model equations and implementation

The pyrolysis reactor (fluidized bed) has been modelled as an isothermal PFR as follows. The material balance for each compound in the reactor is given in Equation (1):

$$\frac{dn_k}{dV} = \sum_{j=1}^{NR} v_{kj} r_j \quad (1)$$

where  $n_k$  is the mole flow of a generic species and  $V$  is the volume of the reactor (domain of integration of the Ordinary Differential Equations system, ODE). The reaction rate ( $r_j$ ) is defined as:

$$r_j = k_{R_j}(T) \cdot c_{\text{reactant}}^\beta \quad (2)$$

and the reaction order depends on the kinetics model adopted: first order ( $\beta=1$ ) for the cellulose and hemicellulose reactants as in Ranzi (Ranzi et al. 2008) and second order ( $\beta=2$ ) for the lignin pyrolysis as reported in Dussan (Dussan et al. 2018). The concentration of the reactant refers to the reactor volume. Reactions occur both in the solid and gas phases, thus the concentration ( $c$ ) of the  $k$ -th specie should be defined as:

$$c_k = \frac{\dot{n}_k [\text{kmol}\cdot\text{s}^{-1}]}{Q [\text{m}^3\cdot\text{s}^{-1}]} \quad (3)$$

We accounted for the variation of the volumetric flow ( $Q$ ) which is the sum of both solid particles and gas species (in a plug flow the gas entrains the solid), as shown in (4):

$$Q = Q_{\text{sol}} + Q_{\text{gas}} \quad (4)$$

$$Q_{\text{sol}} = \frac{\sum \dot{m}_k^{\text{sol}}}{\sum \frac{\omega_k^{\text{sol}}(\dot{n}_{\text{sol}})}{\rho_k^{\text{sol}}}} \quad (5)$$

$$Q_{\text{gas}} = \frac{\sum \dot{m}_k^{\text{gas}}}{\rho_{\text{gas}}} \quad \text{with} \quad \rho_{\text{gas}} = \frac{P \cdot \sum x_i^{\text{gas}}(\dot{n}_{\text{gas}}) \cdot \text{MW}_i}{RT} \quad (6)$$

In the proposed expressions,  $\dot{m}_k^{\text{phase}}$  is the mass flow of a generic species ( $k$ ) in the corresponding phase (gas or solid),  $\omega$  is the mass fraction,  $x$  is the molar one, and  $\rho$  the density of the phase. Thus, the gas and solid volumetric flows, calculated using Equations (5) and (6), respectively, are continuously updated with the composition of the respective phases from the material balance as the pyrolysis of biomass progresses. For the gas phase, we assumed ideal gas mixture behavior due to the high temperatures reached in pyrolysis ( $T > 500^\circ\text{C}$ ). The model requires additional inputs such as the reaction dimensions and the minimum fluidization velocity. The minimum fluidization velocity was calculated using Fogler's method (Fogler, 2020). For the calculation, we assumed that the lignin-rich biomass particles have a spherical shape with a diameter of 1 mm. It is noteworthy that a spherical shape is responsible for the largest pressure drop requirement to spark fluidization, thus, this is a conservative assumption. Typical dimensions of a pyrolysis reactor for industrial applications were adapted from (Zhou et al., 2019). The kinetic model described in this section has been assembled using MATLAB R2022a. The feed inlet composition is the outcome of previous calculations related to other unit operations for biomass pre-treatment. As feed to the pyrolysis, we selected the biomass that had been steam exploded at  $220^\circ\text{C}$  and saccharified with LPMO-cellulase solution with a dose of  $8 \text{ g}_{\text{enzymes}}/\text{kg}_{\text{biomass}}$  (refer to Note I for details).

#### 4. Results

We compared our model's results with the experimental observation reported by Hansen (Hansen et al., 2022). Then, we performed a sensitivity analysis on the pyrolysis temperature (in the range between  $450$  and  $700^\circ\text{C}$ ). The pyrolysis process generates four main product streams: biochar (including unconverted biomass), syngas (including non-condensable components such as carbon monoxide, methane, and short aliphatic compounds), an aqueous phase, and bio-oil (comprising all other components). Table 1 gathers the results from the model under different scenarios. We considered our data obtained with  $450^\circ\text{C}$  and  $500^\circ\text{C}$  temperatures as relevant to contrast our results against the experimental observations (conducted at  $500^\circ\text{C}$ ) in Hansen's work. Furthermore, we tested different pseudo-lignin building block combinations (LIG-1, LIG-2, LIG-3, or LIG-4 alone or the equimolar combination of all four building blocks; see Table 1) to characterize the biomass treated in the experiments since no details are available to fully describe the lignin fraction with the pseudo-lignin components approach by Dussan. There was a considerable mismatch between the experimental data and the model, with a tendency to overestimate the mass ratio of the bio-oil fraction (Table 1). There are several plausible explanations that could justify these deviations to some extent. First, the simplified model does not account for secondary (decomposition) reactions that generate



lighter, non-condensable components in the gas phase, thus leading to an overestimation of the bio-oil and an underestimation of the gas fraction (Corbetta et al., 2015; Ranzi et al., 2014). Second, Dussan and Ranzi's models consider that a fraction of the gas and bio-oil constituents remain adsorbed to the solids even at high temperatures. Our model, on the other hand, assumes complete separation of the pyrolysis fractions, which, to some extent, could lead to underestimation of the biochar fraction and, at the same time, overestimation of the bio-oil and light gas fractions (compare the values to the ones in brackets in Table 1). Third, in the experimental setup by Hansen (Hansen et al. 2022), pyrolysis products are purged with cold nitrogen gas from a small fixed-bed system. The cold nitrogen may locally quench the solid bed, leading to a lower extent of pyrolysis and, consequently, to a higher amount of unconverted solid residue in the biochar fraction experimentally. Fourth, the mass ratio of the pyrolysis fractions substantially depends on the lignin model composition, i.e. the ratio of the four types of pseudo-lignin building blocks, as shown by the sensitivity analysis in Table 1. In the absence of available compositional data on the lignin-rich residue, the lignin model was approximated with the equimolar composition of the four pseudo-lignin building blocks based on the work by Dussan (Dussan et al. 2018), which may be suboptimal. The pseudo-lignin building blocks *LIG-4* and, to some extent, *LIG-2* are associated with a higher residual solid biochar fraction. Notably, (1) there are many other variables in the model that influence the pyrolysis fraction ratios, and (2) the differences between the outcomes with different building blocks are small compared to how far off we are from the experimental data. Moreover, these lignin building blocks (from *LIG-1* to *LIG-4* in Dussan) may not be optimal representatives of our system. Last but not least, the simplification of the process into a limited number of chemical species and reactions (i.e., the use of lumped kinetics) carries an additional source of error. The supporting material by Hansen (Hansen et al. 2022) reports a detailed product distribution and the classification of the thousands of reported species for the bio-oil fraction only. Estimating the distribution of especially low-boiling temperature species in the bio-oil and light gas fractions is not trivial. All in all, our results indicate that further refinement of the model is needed to reach the experimental outcome.

Table 1 - Comparison of the experimental mass distribution of pyrolysis product fractions between experimental data (Exp data) and model predictions with five pseudo-lignin building block combinations at 450°C and 500°C. Values before brackets assume that some species of the bio-oil and gas fractions remain entrapped in the solid bio-char fraction when recovering the pyrolysis fractions, while values in brackets assume complete separation of the pyrolysis fractions.

<b>T = 500°C</b>	<b>Exp data</b>	<b>Equal mix</b>	<b>Only LIG-1</b>	<b>Only LIG-2</b>	<b>Only LIG-3</b>	<b>Only LIG-4</b>
<b>Bio-oil</b>	0.35	0.708 (0.720)	0.800 (0.806)	0.710 (0.729)	0.761 (0.765)	0.675 (0.693)
<b>Gas</b>	0.20	0.081 (0.118)	0.133 (0.136)	0.078 (0.133)	0.082 (0.097)	0.044 (0.123)
<b>Char</b>	0.41	0.172 (0.124)	0.024 (0.015)	0.174 (0.100)	0.129 (0.110)	0.234 (0.137)
<b>Water</b>	0.04	0.039 (0.039)	0.044 (0.044)	0.039 (0.039)	0.030 (0.030)	0.048 (0.048)
<b>T = 450°C</b>						
<b>Bio-oil</b>	0.35	0.581 (0.598)	0.783 (0.793)	0.592 (0.629)	0.679 (0.685)	0.603 (0.620)
<b>Gas</b>	0.20	0.072 (0.108)	0.126 (0.132)	0.075 (0.145)	0.073 (0.089)	0.046 (0.114)
<b>Char</b>	0.41	0.310 (0.257)	0.051 (0.035)	0.293 (0.187)	0.221 (0.199)	0.299 (0.213)
<b>Water</b>	0.04	0.037 (0.037)	0.041 (0.041)	0.040 (0.040)	0.027 (0.027)	0.053 (0.053)

Next, we looked further into the impact of pyrolysis temperature on the conversion efficiency (described by the amount of unconverted biomass) and on the distribution of the types of chemicals being formed and consumed during pyrolysis (Figure 1). Pyrolysis was more complete at higher temperatures as indicated by a drop in the fraction of unconverted biomass remaining after pyrolysis (green bars in Figure 1), which can mainly

be attributed to pyrolysis being endothermic and thus thermodynamically favored at higher temperatures. Increasing the process temperature facilitated char formation (brown bars in Figure 1), although overall char formation remained limited (mass fraction <0.08 w/w). Notably, the decrease in the proportion of unconverted biomass at higher temperatures was not comparable to the increase in char formation, resulting in an apparent decrease in the proportion of the solid “biochar” fraction that comprises unconverted biomass and char. In our model, lignin-rich biomass was primarily converted to lighter gaseous compounds and bio-oil species (listed in Figure 2). Water content and thus the amount of aqueous fraction after pyrolysis remained low (mass fraction <0.05 w/w), irrespective of the pyrolysis temperature. The differences between the product distributions above 600°C were negligible. Thus, we can conclude that the optimal temperature window is found between 600–650°C, with the lignin conversion being higher than 90% at such operating conditions.

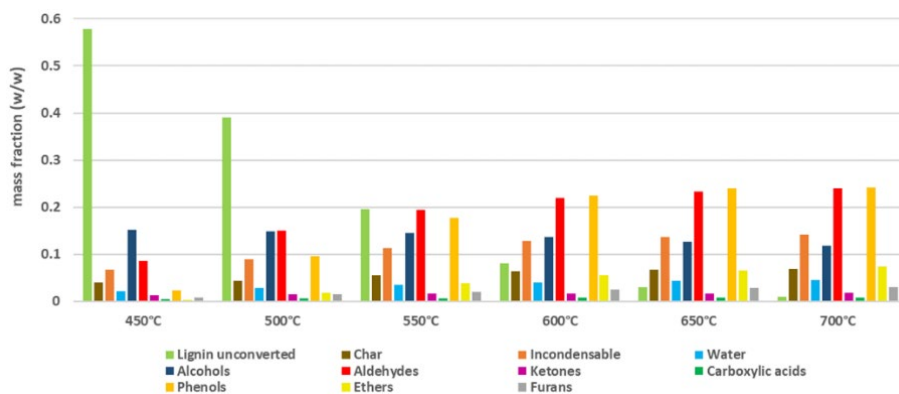


Figure 1 - Temperature effect on the distribution of pyrolysis products.

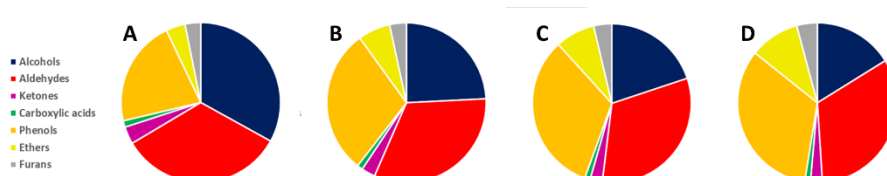


Figure 2 - The distribution of chemical compound types in the bio-oil generated at temperatures that are of interest for industrial pyrolysis: 500°C (A), 550°C (B), 600°C (C), and 700°C (D).

The relative proportion of bio-oil compounds formed at temperatures from 500 to 700°C is shown in Figure 2. We limited the upper temperature to 700°C because over that temperature lignin is almost depleted and secondary reactions result in excessive gas production, thus limiting the bio-oil production as described by Corbetta (Corbetta et al., 2015). In general, aldehydes, alcohols, and phenols, which are the main classes present in biofuels (Okolie et al., 2020) were the main constituents of the bio-oil (Figure 2). In particular, the alcohol fraction decreased when increasing process temperature, while the aldehyde and phenol contents followed opposite trends due to their relatively higher stability. Although phenols are valuable chemicals, their removal from bio-oil is preferred before fuel applications due to corrosiveness and low heating values (Fardhyanti et al., 2020). Furthermore, ethers and furans are appreciably formed above 450°C and make up

about one-tenth of the bio-oil from the comparison between Figure 1 and Figure 2. Ketones and carboxylic acids are also present in the bio-oil mixtures in smaller amounts, and the pyrolysis temperature does not seem to affect substantially their formation. To fully evaluate the potential of bio-oil production in a biorefinery concept, the storage stability of the generated bio-oil will have to be considered in addition to composition.

## 5. Conclusions

The present work merged two lumped kinetics models describing pyrolysis of lignocellulosic biomass published by Ranzi for the cellulose and hemicellulose part, and Dussan for the lignin fraction to characterize the thermal conversion of lignin-rich biomass to biofuels. The model shows a qualitatively good agreement over a limited number of experimental data applied to scale up a spruce-based biorefinery combining biochemical and thermochemical conversion routes as in Hansen (Hansen et al., 2022). Although lumped kinetics itself, which simplifies the thermo- and bio-chemical processes into a comprehensive model, presents some limitations to the accuracy of the model, it allows flowsheet implementation and thereby easier and more efficient workflow during scale-up. Sensitivity analysis of the pyrolysis process allowed us to identify the optimal temperature window for operating a large-scale pyrolysis reactor, and, remarkably, it provided a rough estimate of the residual solid (biochar), bio-oil, and light gas fractions, as well as of the product distribution in the bio-oil.

## Acknowledgements

This work was supported by the Research Council of Norway through grants no. 257622 (Bio4Fuels) and no. 268002 (Enzymes4Fuels).

## References

- F. Cherubini, 2010, The biorefinery concept: Using biomass instead of oil for producing energy and chemicals, *Energy Conversion and Management*, 51(7), 1412-1421
- M. Corbetta et al., 2015, Multi-scale kinetic modelling and experimental investigation of syngas production from coal gasification in updraft gasifiers, *Energy & Fuel*, 29(6), 3972-3984
- K. Dussan et al., 2019, A model of the chemical decomposition and pyrolysis kinetics of lignin, *Proceedings of the Combustion Institute*, 37, 2697-2704
- D.S. Fardhyanti et al., 2020, The separation of phenolic compounds from bio-oil produced from pyrolysis of corncobs, *AIP Conference Proceedings* 2243, 020005
- H.S. Fogler, 2020, Chapter 15 – Fluidized-bed reactors, in *Elements of Chemical Reactions Engineering*, fifth edition, Prentice Hall Editor
- P. Kumar et al., 2020, Lignocellulose biomass pyrolysis for bio-oil production: A review of biomass pre-treatment methods for production of drop-in fuels, *Renewable and Sustainable Energy Reviews*, 123, 109763
- L.D. Hansen et al., 2022, 2-Naphthol impregnation prior to steam explosion promotes LPMO-assisted enzymatic saccharification of spruce and yields high-purity lignin, *ACS Sustainable Chemistry & Engineering*, 10, 5233-5242
- D. Manca et al., 2001, Numerical problems in the solution of oxidation and combustion models, *Combustion Theory and Modelling*, 5(2), 185-199
- J.A. Okolie et al., 2021, Next-generation biofuels and platform biochemicals from lignocellulosic biomass, *International Journal of Energy Research*, 45, 14145-14169
- E. Ranzi et al., 2014, Kinetic modeling of the thermal degradation and combustion of biomass, *Chemical Engineering Science*, 110, 2-12
- E. Ranzi et al., 2008, Chemical kinetics of biomass pyrolysis, *Energy & Fuel*, 22, 4292-4300
- G. Zhou et al., 2019, Simulation of coal pressurized pyrolysis process in an industrial-scale spout-fluid bed reactor, *Advanced Powder Technology*, 30, 3135-3145

# Numerical Modeling of Gas-liquid Hydrodynamics in a U-loop Bioreactor

Johan Le Nepvou De Carfort,<sup>a</sup> Tiago Pinto,<sup>b</sup> Ulrich Krühne,<sup>a</sup>

<sup>a</sup> PROSYS, Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søtofts Plads Building 228A, Kongens Lyngby, Denmark

<sup>b</sup> R&D Department, UNIBIO A/S, Roskilde, Denmark

## Abstract

A numerical model of a U-loop bioreactor is developed, with the aim of replicating the mass transfer characteristics of the real system. The bioreactor model includes a multiphase fluid dynamic model, coupled with a population balance model, to simulate the dispersed bubbles in the loop and estimate the gas-liquid interfacial surface area. The results from the numerical model are compared to experimental measurements of the mass transfer coefficient in the downgoing leg of the U-loop. The numerical model of the U-loop bioreactor is able to replicate the qualitative behavior of  $K_L a$  as function of the operating parameters: liquid circulation, gas flow and bottom pressure, however the dependency on operational conditions is under-valued. Further validation of the individual models is required to develop more realistic and robust models, and will contribute to understanding the physical phenomena taking place in a U-loop. The numerical model also provides insights into the qualitative behavior of the multiphase flow in a U-loop bioreactor.

**Keywords:** Gas fermentation, Static Mixer, Computational Fluid Dynamics (CFD).

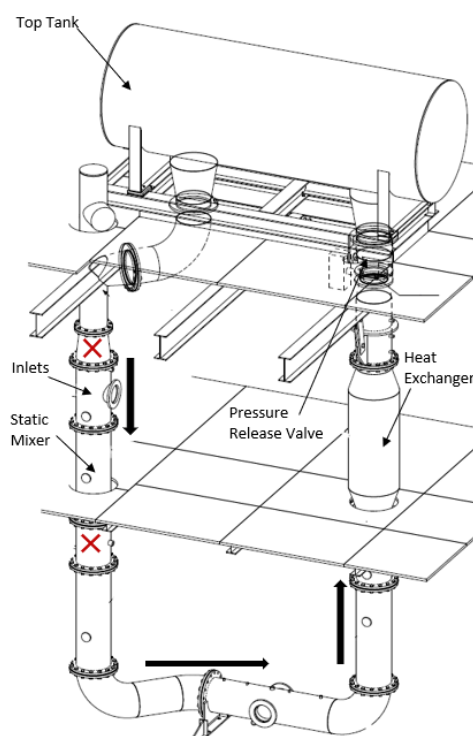


Figure 1 : Drawing of the 1725 L U-loop Bioreactor, UNIBIO A/S. Red cross: location of dissolved oxygen measurements. Black arrow: Direction of the flow.

## 1. Introduction

Fermentation has become an increasingly important chemical process in a wide range of applications, from pharmaceutical to bulk chemical and biofuel productions [1]. However, most fermentations are still run as batch processes and lack measurement-based control [2]. Other challenges with operating industrial scale fermentations relates to the inhomogeneous mixing conditions and mass transfer limitations, both related to the fluid dynamics of the system. This induces gradients in the fermenter, that lead to suboptimal conditions for the microorganisms and control issues, both having an impact on the

process economic metrics. Thus, understanding and accurately modeling the fluid dynamic behavior of bioreactors is a key element for predicting the overall behavior of the fermentation process. The main challenge with modeling fluid behavior in large-scale equipment originates from the complex nature of turbulent flow, which limits the accuracy of deterministic models in macroscopic simulations at high Reynolds numbers. Computational fluid dynamics has become a useful tool for investigating hydrodynamic conditions in multiphase systems through modeling of the turbulence and fine discretization of the flow equations. Interesting case studies are the U-loop bioreactors currently used by UNIBIO A/S, a Danish company specialized in the production of Single Cell Proteins through continuous aerobic fermentation of *Methylococcus capsulatus* using methane gas as primary carbon source. An aerobic fermentation using a gaseous substrate, makes mass transfer between the phases of crucial importance to the process. The alternative bioreactor design, U-loop, greatly improves volumetric mass transfer, however, it is not yet fully understood which underlying phenomena are contributing to the enhanced mass transfer. This study therefore looks at developing and validate a numerical model of a U-loop bioreactor to better understand the mass transfer mechanism in such a system.

## 2. Materials & Methods

### 2.1. Description of the system

The design of the U-loop bioreactor consists of two distinguishable parts: the loop (Plug flow) section and the top tank (or degassing unit). Gases introduced at the start of the down-going leg are carried through the loop by the liquid and are finally released in the top tank. A sketch of the system is shown in Figure 1. A pump, located before the gas inlets, is in charge of circulating the liquid through the loop. The U-loop is equipped with static mixers, and a heat exchanger. The pressure in the loop can be adjusted using an expansion valve located after the heat exchanger. This study uses a production scale U-loop with an active volume of 1725 [L], a liquid height of 6 [m] and a loop pipe diameter of 0.4 [m]. The selected operating conditions are similar to the conditions during production (e.g. flowrates and pressure), to mimic industrial process conditions. To investigate the multiphase hydrodynamics, water is used instead of fermentation broth, and the gas phase only includes air and pure oxygen. This simplification is based on the assumption that the mass transfer is mainly driven by convective transport (not diffusion driven) and thus the global liquid phase mass transfer coefficient ( $K_L a$ ) can be assumed equal for all compounds, and independent of the transfer direction. This simplification does consider the mass transfer of CO<sub>2</sub> produced during fermentation. The model can later be coupled with biokinetic equations to predict the transfer rates and resulting concentration profiles for each individual gas.

### 2.2. Experimental Setup

The  $K_L a$  in the U-loop is determined experimentally based on dissolved oxygen measurements. Due to the plug flow nature of the bioreactor, ideal mixing cannot be assumed for the entire system. Therefore, the component mass balance assuming ideal mixing (eq. 1), typically used to determine the  $K_L a$  in stirred tanks, cannot be applied for the U-loop. However,  $K_L a$  can be determined for a section of the loop based on two measurement points and the component mass balance over this loop section. Assuming steady state, and unidirectional mass transfer (i.e. from the gas phase to the liquid phase) the resulting mass balance (eq. 2) is an algebraic equation. Assuming perfect radial mixing, the dissolved oxygen measurements are used to determine the inlet and outlet oxygen concentrations,  $C_{in}$  and  $C_{out}$  respectively. The measurements are performed over

a 1[m] vertical pipe section on the down-going leg of the loop. The dissolved oxygen is measured at the top and bottom of the section, and the pressure is only measured at the top. The equilibrium saturation concentration  $C^*$  of oxygen in water is generally given by Henry's Law (eq.3) and is adjusted for changing pressures.

$$\frac{dC}{dt} = K_L a (C^* - C) \quad (eq.1)$$

$$0 = K_L a (C^* - C) V_L^{section} + F_{in} C_{in} - F_{out} C_{out} \quad (eq.2)$$

$$C^* = He P_{O_2} \quad (eq.3)$$

$He$  being the Henry Constant for oxygen in water (0.0013 [ $mol\ kg^{-1}\ bar^{-1}$ ] or 41.48 [ $g\ m^{-3}\ bar^{-1}$ ] at 25°C [3]), and  $P_{O_2}$  the partial pressure of oxygen (eq. 4).

$$P_{O_2} = y_{O_2} P \quad (eq.4)$$

For this study, the ratio between air and pure oxygen is adjusted to achieve an oxygen inlet concentration of 30%. Since the  $O_2$  liquid concentration  $C$  and saturation concentration  $C^*$  change over the pipe section, the driving force is estimated using the logarithmic mean (eq.5) of the driving force in (eq.6) and out (eq.7) of the pipe section.

$$(C^* - C)_m = \frac{(C^* - C)_{in} - (C^* - C)_{out}}{\ln\left(\frac{(C^* - C)_{in}}{(C^* - C)_{out}}\right)} \quad (eq.5)$$

$$(C^* - C)_{in} = (He P_{in} y_{O_2, in} - C_{in}) \quad (eq.6)$$

$$(C^* - C)_{out} = (He P_{out} y_{O_2, out} - C_{out}) \quad (eq.7)$$

The inlet pressure corresponds to the measurement of the top pressure sensor, and the outlet pressure is estimated from the height difference (eq.8) (neglecting the pressure drop in the pipe). The inlet oxygen fraction in the gas phase  $y_{O_2, in}$  is determined from the ratio of air and pure oxygen (eq.9) and the outlet oxygen fraction in the gas phase  $y_{O_2, out}$  is determined from the mass balance (eq.10):

$$P_{out} = P_{in} + \rho_L g \Delta h \quad (eq.8)$$

$$y_{O_2, in} = \frac{Q_{air} y_{air} + Q_{O_2}}{Q_{air} + Q_{O_2}} \quad (eq.9)$$

$$y_{O_2, out} = \frac{Q_{air} y_{air} + Q_{O_2} - (C_{in} - C_{out}) Q_{Uloop}}{Q_{air} + Q_{O_2} - (C_{in} - C_{out}) Q_{Uloop}} \quad (eq.10)$$

where  $Q_{Uloop}$  represents the circulation flow rate of the liquid phase in the loop,  $y_{air}$  the fraction of oxygen in air (21%),  $Q_{air}$  the inlet flowrate of air and  $Q_{O_2}$  the inlet flowrate of pure oxygen. From equations (eq. 2-10),  $K_L a$  is determined based on the dissolved oxygen measurements in and out, as well as the pressure measurement at the section inlet. The latter measurement is performed at different operating conditions, by changing the liquid circulation rate, the gas inlet flow rate, and the operating pressure in the loop.

### 2.3. Computational Model

The aim of the computational model is to replicate the multiphase physics taking place in the bioreactor to predict the mass transfer rates at different operating conditions. The model will focus on the loop section of the bioreactor, as the area of interest. Therefore, the fluid domain of the numerical model starts immediately after the pump, and ends immediately after the pressure reduction valve. For simplification, the heat exchanger,

and the pipe bends on the horizontal part of the loop are neglected. The static mixers in the loop are modelled as a porous domain for which the model parameters are fitted to match the measurements. In this way, the pressure drop, fluid mixing and impact on the bubbly flow resulting from the static mixers can be recreated, without having to resolve the geometry of the static mixers. Modeling the exact geometry of the static mixers and the complete U-loop would require a fine discretization to resolve the flow, adding to the computational cost of the simulation. The geometry of the computational domain is shown in Figure 2. The fluid domain, as well as the temporal dimension are then discretized, to solve the model equations in a numerical way. Ideally, the discretization should not influence the results. Therefore, a discretization study is performed, where the results from different meshes and timesteps are assessed. Once the geometry has been defined, and discretized, the governing equations to be solved are selected. Based on the liquid phase Reynold's Number (eq.11) of around  $4 \cdot 10^5$  in the pipe section, the flow regime is expected to be fully turbulent. The multiphase flow regime (e.g., Figure 3) depends on many factors (gas volume fraction, superficial gas velocity, pipe diameter and orientation, sparger type, etc. [5]) and is difficult to predict.



Figure 2: Fluid domain for the computational model

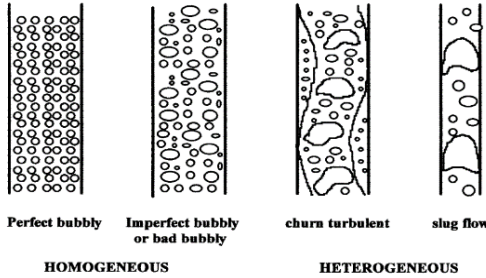


Figure 3: Gas-liquid flow regimes in a horizontal pipe (Adapted from [4])

parameters are shown in Table 1. The bubbles are assumed spherical, and the bubble size is modeled with a population balance model (PBM) describing the bubble coalescence and breakup. The bubble diameter is discretized into 5 bins: 0.86mm, 1.78mm, 2,7mm, 3.62mm and 4.54mm. The size fraction of each bin is determined from the population balance model. Finally, the mass transfer model is implemented. Based on the Film theory [11], the oxygen transfer rate can be expressed as shown in eq.13:

$$OTR = K_L a (C^* - C) \quad (eq. 13)$$

where  $OTR$  is the volumetric oxygen transfer rate [ $g m^{-3} s^{-1}$ ],  $C^*$  the saturation concentration of oxygen in water (or oxygen concentration at the surface of the bubble), and  $C$  the bulk oxygen concentration. The interfacial area  $a$  is determined from the PBM according to (eq.14) and (eq.15),  $d_p$  being the mean bubble Sauter diameter and  $r_\beta$  the volume fraction of air.

$$a = \frac{6\tilde{r}_\beta}{d_p} \quad (eq. 14)$$

$$\tilde{r}_\beta = \begin{cases} \max(r_\beta, r_{min}), & r_\beta < r_{max} \\ \max\left(\frac{1-r_\beta}{1-r_{max}} r_{max}\right), & r_\beta > r_{max} \end{cases} \quad (eq. 15)$$

The bubble Reynolds number (eq.12) can provide some insight into the turbulence level at the scale of a single bubble. However, it requires knowing the bubble size and relative velocity.

$$Re = \frac{\rho_L u L}{\mu_L} \quad (eq. 11)$$

$$Re_{bubble} = \frac{\rho_L d_b |\overline{U_G} - \overline{U_L}|}{\mu_L} \quad (eq. 12)$$

For this study, the gas phase is assumed fully dispersed resulting in a perfect bubbly flow. The simulation

Parameter	Value
Multiphase Model	Liquid phase: Continuous Gas phase: Fully dispersed + PBM (Ideal gas at 25°C)
Gravity	[0 -9.82 0] [ $m/s^2$ ]
Reference pressure	1 [atm]
Coalescence Model	Prince and Blanch [7]
Breakup Model	Luo and Svendsen [8]
Liquid phase turbulence	Shear Stress Transport model [9]
Gas-liquid surface tension	0.072 [ $N m^{-1}$ ]
Momentum Transfer Model	Schiller and Naumann [10]

Table 1: Simulation Parameters for the Numerical Study

$r_{min}$  and  $r_{max}$  are model parameters.  $\tilde{\beta}$  is used as a modified volume fraction to avoid the overprediction of the interfacial area at high gas volume fractions, where the assumption of fully dispersed gas phase is no longer valid. The liquid mass transfer coefficient  $K_L$  (eq.16) is determined from the derivation of the Higbie's penetration theory for turbulent Newtonian flow [6]:

$$K_L = 0.301 \left( \frac{D_\phi}{\nu} \right)^{-0.5} (\epsilon \nu)^{0.25} \quad (eq. 16)$$

where  $D_\phi$  is the diffusion coefficient of oxygen in water,  $\nu$  is the kinematic viscosity of water and  $\epsilon$  is the turbulence eddy dissipation generated by the turbulence model.

#### 2.4. Numerical Setup

The simulations are performed on ANSYS CFX 2019R2 on 24 hpc parallel cores, with three different conditions at the boundaries: the liquid phase inlet, the no-slip walls, and the pressure outlet. The relative pressure at the pressure outlet is adjusted to achieve the correct pressure profile in the loop (representing the pressure reduction valve). The gas phase is introduced into the system through a numerical source point with a bubble diameter of 2.7mm. All simulations are run in Transient mode for 200[s] simulation time, to ensure steady state is achieved (circulation time  $\approx 14$ [s]). The fluid is initialized stationary. Due to the unsteady nature of turbulent transient simulations, the results are time averaged over the last 50 [s] simulation time.

### 3. Results and Discussion

#### 3.1. Discretization Study

The simulation is performed for a range of generated meshes (from 15k to 140k mesh elements), each over a range of timesteps (0.001[s] to 0.1[s]). For each simulation, the gas holdup in the system is calculated and used to compare the results. In theory, the results should converge towards a single value as the discretization is made finer.

The optimal discretization is the coarsest discretization that provides similar results as the finer discretization. The results from the discretization study are presented in Figure 4. For the simulations performed with the coarse meshes (<60k elements) the results are inconsistent and depend on the mesh and timestep. For the two finest meshes (>80k elements) the results seem to be independent on the discretization. One outlier is the simulation performed with a timestep of 0.1 [s] yielded a gas holdup above 30 %, whereas the other simulations resulted in gas holdups around 10%. Based on this analysis, the optimal discretization is 83k mesh with timestep of 0.05 [s].

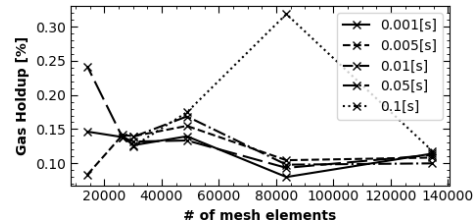


Figure 4: Gas Holdup as function of the number of mesh elements, for different



### 3.2. Model Validation

The measured and simulated  $K_L a$  values are shown in Figure 5. The first run is used for calibration, where the length of the porous domain and  $r_{max}$  were used as fitting parameters. The numerical model is able to replicate an increase or decrease of the  $K_L a$  at different operating conditions, however the

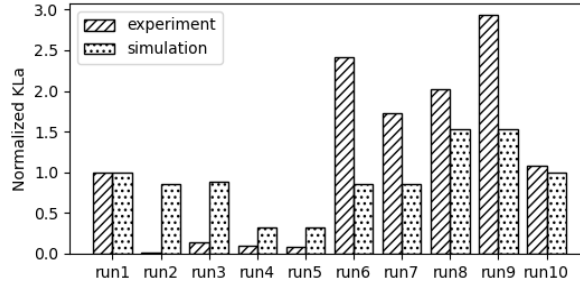


Figure 5: Measured and simulated  $K_L a$  values, at different operating conditions. (Results are normalized due to confidentiality of the experimental results)

dependency on operating conditions is under-valued. The discrepancy between the simulated and measured  $K_L a$  values can be explained by inaccuracies in the modeling, as well as inaccuracies in the experimental method. Regarding the numerical model, the choice of the models to be solved may be inadequate to provide realistic results. Here the choice of using a porous domain to simulate a static mixer may not be sufficient and requires further tuning, or resolving the exact geometry of the static mixer. The assumption of fully bubbly flow regime, which is the basis for the choice of multiphase model, may be inaccurate, which is discussed in the following section 3.3.. Lastly, the PBM may require additional bins and the bubble diameter range represented in the bins could be offset to the actual bubble size distribution in the loop. Regarding the experimental measurements, inaccuracy of the equipment (flowmeter, DO probes) due to calibration or malfunction is a potential source of error. Also, inaccurate assumptions used to derive  $K_L a$  from the experimental data may contribute to the discrepancy.

### 3.3. Discussion

It can be clearly noted that the gas is not distributed homogeneously in the loop. A gas fraction of 23% is reached in the down-going leg, whereas in the upgoing leg, the gas fraction is around 3%. In the horizontal section of the loop, the model predicts a large accumulation of gas in the top half of the pipe, indicating a phase split (gas volume fraction of 37%). This behavior can be expected in such a system, considering the buoyant forces, the bubbles will rise in water, with a velocity depending on the drag force [12]. This means the bubbles rise against the flow direction in the down-going leg, increasing the residence time of the bubbles, in the horizontal section the bubbles rise perpendicular to the flow direction, creating the phase split, and in the upgoing leg the bubbles rise along with the flow direction decreasing the residence time of the bubbles. This leads to an inhomogeneous distribution of  $K_L a$  in the system, which is expected to be larger in the down-going leg, due to the phase split in the horizontal section that leads to a small interfacial surface area, and the low gas holdup in the upgoing leg. This can also be observed in the simulation results, where the  $K_L a$  in the downgoing leg, the horizontal section and the upgoing leg is determined to 1692, 474 and 123 [ $h^{-1}$ ] respectively. The resulting gas volume fractions in some sections of the bioreactor indicated that the flow regime might be different from the assumed bubbly flow. The dispersed gas phase model requires that the bubbles are smaller than the mesh elements. If larger gas structures are expected, the dispersed model is not applicable, and the gas needs to be simulated as a continuous fluid where the gas-liquid interface is resolved. Here, it seems that in some regions of the bioreactor, the gas phase can be assumed fully dispersed, e.g. down- and up-going legs, whereas in other regions the interface needs to be resolved, e.g. in the horizontal loop section where a clear phase split occurs. The  $\tilde{r}_\beta$  parameter is used to

correct the interfacial area, but the momentum transfer model considers only bubbles. Moreover, the validation of the multiphase model and the PBM is crucial for further studies to develop more accurate numerical models. This requires measurements of the multiphase behavior at a deeper level to directly validate each sub-model involved simulation.

#### 4. Conclusion

The numerical model of the U-loop bioreactor is able to replicate the qualitative behavior of the measured  $K_L a$ , however the dependency on operational conditions is under-valued. As many sub-models are involved in the simulation of the bioreactor, it is difficult to evaluate the origin of the inaccuracy in predicting the measured  $K_L a$ . Therefore, further studies should include additional measurements of the multiphase flow (e.i. bubble size distribution, gas superficial velocity), to ensure a more rigorous validation of the sub-models. On the other hand, the numerical model provided insights into the qualitative behavior of the multiphase flow in a U-loop bioreactor, presenting some challenges related to inhomogeneous distribution of the mass transfer potential. The results can be used as a basis for upscaling and improving the design of the bioreactor. However, further studies including further validations are required for the model to be used for accurate prediction of the process, for process optimization and control purposes.

#### Reference

1. Capozzi, V., Fragasso, M., & Bimbo, F. (2021). Microbial resources, fermentation and reduction of negative externalities in food systems: Patterns toward sustainability and resilience. In *Fermentation* (Vol. 7, Issue 2).
2. Alarcon, C., & Shene, C. (2021). Fermentation 4.0, a case study on computer vision, soft sensor, connectivity, and control applied to the fermentation of a thraustochytrid. *Computers in Industry*, 128.
3. NIST Chemistry WebBook, Oxygen, Henry's Law data, Hensy's Law constant. URL: <https://webbook.nist.gov/cgi/cbook.cgi?ID=C7782447&Mask=10#Solubility>
4. Bouaifi, M., Hebrard, G., Bastoul, D., & Roustan, M. (2001). A comparative study of gas hold-up, bubble size, interfacial area and mass transfer coefficients in stirred gas-liquid reactors and bubble columns. *Chemical Engineering and Processing*, 40(2).
5. Hebrard, G., Bastoul, D., & Roustan, M. (1996). Influence of the gas sparger on the hydrodynamic behaviour of bubble columns. *Chemical Engineering Research and Design*, 74(3).
6. Kawase, Y., Halard, B., & Moo-Young, M. (1992). Liquid-Phase mass transfer coefficients in bioreactors. *Biotechnology and Bioengineering*, 39(11).
7. Prince, M.J. and Blanch, H.W. (1990), Bubble coalescence and break-up in air-sparged bubble columns. *AIChE J.*, 36: 1485-1499.
8. Luo, H. and Svendsen, H.F. (1996), Theoretical model for drop and bubble breakup in turbulent dispersions. *AIChE J.*, 42: 1225-1233.
9. Menter, F. R. (1994). Two-equation eddy-viscosity turbulence models for engineering applications. *AIAA Journal*, 32(8).
10. Schiller, L., & Naumann, Z. (1935). A drag coefficient correlation. In *Z. Ver. Deutsch. Ing* (Vol. 77, Issues 13–14).
11. Titchener-Hooker, N. (1995). Bioprocess Engineering Principles (by Pauline M. Doran). *TRENDS IN BIOTECHNOLOGY*, 13(1).
12. Kulkarni, A. A., & Joshi, J. B.(2005). Bubble formation and bubble rise velocity in gas-liquid systems: A review. In *Industrial and Engineering Chemistry Research* (Vol.44,Issue16).



# Effect of vaporisation models on the FCC riser modelling

Thabang W. Selalame, Raj Patel, Iqbal M. Mujtaba, Yakubu M. John

*Department of Chemical Engineering, University of Bradford, Bradford, BD7 1DP, West Yorkshire, UK*

## Abstract

Traditional approach to FCC riser modelling usually involves the modelling of the gas oil feed vaporisation since the cracking reactions occur in the gas phase. The majority of these studies assume instantaneous vaporisation of feed which greatly simplifies the modelling effort. A limitation of this approach is that it greatly overestimates the rates of vaporisation occurring in the riser which may lead to erroneous predictions from the model. In response to this, various models have been proposed in the literature to better model the finite vaporisation rates in the riser. This paper investigates the effect of these various vaporisation models on the predictions from riser models. Specifically, the infinite conductivity and the finite conductivity modelling frameworks are investigated. We show for the first time in this paper that vaporisation models are largely insignificant to the predicted gas oil conversion profiles and gasoline yield for usual operation conditions of FCC risers. Additionally, the effect of vaporisation models on conversion and yield predictions only becomes important for riser nozzles with feed droplets exceeding 2mm in diameter. This is a key finding as it proves the validity of instantaneous vaporisation models for droplets smaller than 2mm.

**Keywords:** Vaporisation, FCC unit, Riser, Modelling, Simulation

## 1. Introduction

Process modelling and simulation have always been important tools used by the petroleum industry to improve safety, efficiency and profitability. Owing to climate change, more emphasis has been put on the accuracy of these models in order to comply with increasingly stringent environmental regulations.

The Fluid catalytic cracking (FCC) unit is one of the most important units in petroleum refinery (Fernandes, Verstraete et al. 2007). For this reason, the modelling of the FCC riser has received substantial commercial and academic attention since the introduction of the process over five decades ago. Significant progress has been made to model the complex hydrodynamics and the reactions occurring in the riser. However, an area that has not been fully explored is that of vaporisation of the feed. Several models have been proposed to study vaporisation of fuel droplets, although most of these studies focus on single droplets evaporating into an infinite and quiescent gaseous atmosphere (Finneran 2021). This is far from the conditions in the riser, where droplets vaporise in complex gas-solid-liquid convective environment. This study investigates the effect of different vaporisation models, that have been previously developed in the literature, on the FCC riser model predictions for feed conversion and product yield. Two droplet frameworks are investigated, infinite thermal conductivity (ITC) and finite thermal conductivity

(FTC) (Miller, Harstad et al. 1998). In the former, spatial temperature variation in the droplet are ignored, while the latter assumes a finite value for the liquid phase and considers the spatial temperature distribution in the droplet as it rises. A majority of FCC riser models in the literature have adopted either the ITC framework (Nayak, Joshi et al. 2005) or assumed instantaneous vaporisation of the feed at the point of injection (John, Mustafa et al. 2019). Very few papers have used FTC in their riser models. As such, the impact of FTC on the riser predictions remain unknown until now.

## 2. Model Description

The catalytic cracking of heavy fraction gas oil in the FCC riser is a heterogeneous gas-solid reaction system. The complex hydrodynamics of the fast fluidised circulating bed have been described previously. The feed to the riser is introduced as atomised spray droplets that first undergoes vaporisation in the initial stages of the riser before reaction can occur. Various models have been proposed to describe the vaporisation dynamics in the riser environment. In this paper, we compare the influence of these models on the conversion and yield profiles in the riser.

The vaporisation problem being investigated in this situation is that of discrete spherical mono-component droplets evaporating in a gas-droplet-solid convective environment. The fluid and far field properties such as temperature, concentration and pressure are varying with the height of the riser due to the cracking reactions. Firstly, the assumption of infinite thermal conductivity (ITC) and diffusivity for the droplet phase is adopted, so that the temperature and concentration is the same everywhere on the droplet. The evolution of the droplet temperature, mass and size with height of the riser bed is given by:

$$m_d C_{pd} \frac{dT_d}{dz} = \frac{\dot{q} - \dot{m} L_v}{v_d} \quad (1)$$

$$\frac{dm_d}{dz} = -\frac{\dot{m}}{v_d} \quad (2)$$

$$\frac{dd_d}{dz} = -\frac{2\dot{m}}{\pi \rho_d d_d^2 v_d} \quad (3)$$

Where  $m_d$ ,  $d_d$  and  $T_d$  are the mass, diameter and temperature of the droplet,  $L_v$  is the latent heat of vaporisation,  $\rho_d$  is the liquid density and  $v_d$  is the velocity of the droplet.  $\dot{q}$  and  $\dot{m}$  are the rates of heat transfer and mass transfer at the droplet surface respectively. During the inert heating phase,  $\dot{m} = 0$ . Closures for  $\dot{q}$  and  $\dot{m}$  are provided via Nusselt and Sherwood number correlations. For convective environment (including the effect of Stefan flow) (Lupo 2017):

$$\dot{q} = \pi d_d^2 Nu k_g \frac{T_d - T_\infty}{d_d} \quad (4)$$

$$\dot{m} = \pi d_d^2 Sh \bar{\rho} D_{vap} \frac{Y_s - Y_\infty}{d_d} \quad (5)$$

$$Nu = \frac{Nu_0 \ln(1 + B_T)}{Le B_T} \quad (6)$$

$$Sh = \frac{Sh_0 \ln(1 + B_M)}{B_M(1 - Y_s)} \quad (7)$$

The subscript s and g represent the conditions at the droplet surface and the surrounding gas respectively. The heat and mass transfer is fully described by the dimensionless numbers:  $B_M = (Y_s - Y_\infty)/(1 - Y_s)$ ,  $B_T = (1 + B_M)^{\frac{1}{Le}} - 1$ ,  $Le = \frac{k_g}{\bar{\rho} D_{vap} C_{pg}}$ ,  $Nu = \frac{hd_d}{k_g}$ ,  $Sh = \frac{k_c d_d}{D_{vap}}$ . The boundary conditions are such that the far field properties (i.e  $Y_\infty, T_\infty$ ) are evaluated in the ambient hydrocarbon

gas mixture and  $Y_s$  and  $T_d$  are at the surface of the droplet. The vapour properties are evaluated at a reference temperature and composition using the 1/3 rule (Csemány and Józsa 2021). Subsequently, the gas and mixture properties are evaluated using the following mixing rules considering the vapour and ambient gas mixture:

$$C_{pg} = Y_{ref}C_{pv} + (1 - Y_{ref})C_{pm} \quad (8)$$

$$k_g = Y_{ref}k_v + (1 - Y_{ref})k_m \quad (9)$$

$$\mu_g = Y_{ref}\mu_v + (1 - Y_{ref})\mu_m \quad (10)$$

$$\rho_g = \left( \frac{Y_{ref}}{\rho_v} + \frac{1 - Y_{ref}}{\rho_m} \right)^{-1} \quad (11)$$

The subscripts “v” and “m” represent the vapour and ambient gas mixture respectively.  $C_p$ ,  $k$ ,  $\mu$  and  $\rho$  properties are evaluated at  $T_{ref}$  and  $Y_{ref}$  for the vapour and evaluated at  $T_\infty$  and  $Y_\infty$  for the ambient gas mixture. The mixing laws in Equations (8)-(11) are based on ideal gas conditions (Csemány and Józsa 2021).

The subscript “0” in Equations (6) and (7) indicates that the correlations for these transport numbers are evaluated in the limit of  $B_M = B_T = 0$  (i.e negligible evaporation). The classical models for  $Nu_0$  and  $Sh_0$  for moderate Reynolds numbers are those provided by Ranz (1952). In this work, we denote this the Classic model. The complexity with heat and mass transfer in the riser environment is that the flow field also consists of catalyst particles which enhance the heat transfer process. Several correlations have been suggested to account for this enhanced heat transfer. In this work the models by Buchanan (1994) and Nayak et al. (2005) are considered. A summary of these correlations is presented in Table 1.

Table 1: Nusselt number correlations for heat transfer to the droplets in the limit of negligible evaporation.

Model name	Reference	Model
Classic	Ranz (1952)	$Nu_0 = 2 + 0.552Re_d^{0.5}Pr^{0.33}$
Buchanan B*	Buchanan (1994)	$Nu_0 = \frac{\left(2 + 0.6Re_p^{\frac{1}{2}}Pr^{\frac{1}{3}}\right)}{\left(1 + \frac{c_{pg}(T_g - T_d)}{L_v}\right)^n}$
Nayak**	Nayak, Joshi et al. (2005)	$Nu_0 = Nu_{0,classic} + \frac{\emptyset \varepsilon_s L_v Pr}{4C_p(T - T_d)} \left( Re_d - Re_p \frac{d_d}{d_p} \right) \left( 1 - \frac{d_p}{d_d} \right)^2$
$Re_d = \frac{\rho_g  v_g - v_d  d_d}{\mu_g}$ , $Re_p = \frac{\rho_s \varepsilon_s  v_g - v_d  d_d}{\mu_g}$ , $Pr = \frac{c_{pg} \mu_g}{k_g}$ , Heat and mass transfer analogy is used to find $Sh_0$ from $Nu_0$ . *During heat up $n = 0$ and $n = 0.7$ during boiling. **The Nayak is the classic model with an extra term that represents the effect of catalyst particles. $\emptyset$ is an adjustable parameter of the model.		

### 3. Results

#### 3.1. Droplet size and temperature

Figure 1 shows the evolution of droplet temperature and size along the height of the riser from the point of feed injection. The Buchanan A model is the limit of infinitely fast heat transfer from catalyst particles to the droplets (Buchanan 1994).

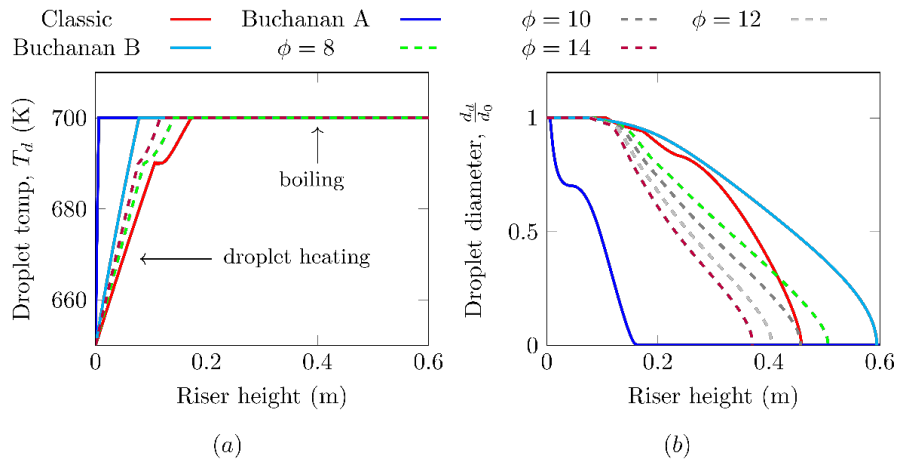


Figure 1: Simulation results for the steady state FCC riser ( inputs:  $\rho_d = 810$ ,  $\varepsilon_{d0} = 0.012$ ,  $d_d = 500\mu\text{m}$ ) comparing between different vaporisation models. (a) Evolution of droplet temperature from initial feed temperature to boiling, and (b) evolution of droplet size up the riser.  $\phi$  is the adjustable parameter in the Nayak heterogeneous model. The Buchanan models feature only inert heating and boiling, whereas Nayak and classic models have vaporisation beginning at  $T_{vap} = 690\text{K}$ .

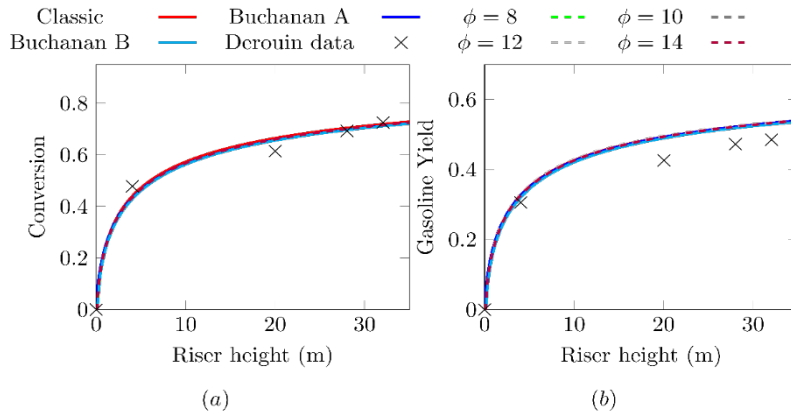


Figure 2: (a) Predictions of gas oil feed conversion, and (b) gasoline yielded along the riser for the various feed vaporisation models. ( inputs:  $\rho_d = 810$ ,  $\varepsilon_{d0} = 0.012$ ,  $d_d = 500\mu\text{m}$ ). Simulation results are compared against the plant data reported by Derouin, Nevicato et al. (1997).

Figure 1(a) shows that the Buchanan models have higher heat transfer rates in the heat up stages and therefore droplets reach boiling temperature faster. However, the factor  $\left(1 + \frac{c_{pg}(T_g - T_d)}{L_v}\right)^{-0.7}$  in the Nusselt number for the Buchanan B model, which accounts for Leidenfrost effects, greatly reduces the heat transfer during boiling and hence results in longer droplet penetration as shown in Figure 1(b). It is shown in Figure 1(b) that the different vaporisation models predict different droplet penetration at the bottom of the riser. Additionally, all the models predict complete vaporisation within 0.6m from feed injection point, which is  $< 2\%$  of the total height of

the riser. However, Figure 2 shows that these differences do not translate to differences in conversion and yield profiles along the riser. This is a significant finding because it proves that for the ITC vaporisation framework; (a) hydrocarbon cracking reactions in the gas phase are the rate limiting step in the cracking process, and (b) vaporisation models are largely insignificant to the predictions of conversion and yield in the FCC riser, which gives validity to the commonly used assumption of instantaneous vaporisation.

### 3.2. Effect of modelling framework

In Figure 3, simulation for the ITC and the FTC modelling frameworks are compared. Figure 3(b) shows that all the FTC models predict slower vaporisation rates compared to their corresponding ITC counterparts in Figure 1(b). This is due to the spatial temperature variation in the droplet and heat transfer resistance caused by the finite thermal conductivity of the liquid phase. This resistance causes  $T_d$  to take longer to reach  $T_{bp}$ .

Beyond the initial vaporisation of the droplet phase (i.e. in the range  $690 \leq T_d \leq 700$ ), a complex coupling of gas-liquid-solid phases results, which determines the evolution of the gas oil conversion in Figure 3(c). In this temperature range, evaporation is slower since the heat transfer to the droplet is also used for droplet heat up. The added heat transfer resistance in the liquid leads to more energy being used up the droplet phase which leads to lower ambient (surrounding gas and catalyst) temperature and hence lower conversion in the FTC models. We find that exit gas temperatures of the Classic model between ITC and FTC differ by  $\sim 15K$ . An interesting find, is that the FTC models underpredict the conversion of the feed. This is a result of the aforementioned lower ambient temperatures in the FTC models. A fix for this, which has not been shown in this current work, is to use the effective thermal conductivity (ETC) modelling framework, which was developed to correct the liquid thermal conductivity of droplets vaporising in convective environments. In such systems, shear at the surfaces of the droplets causes internal circulations which enhances the heat transfer in the droplet. A common problem in vaporisation modelling for the FCC riser is the difficult in validating the models since conditions are such that experimental determination of vaporisation rates is difficult.

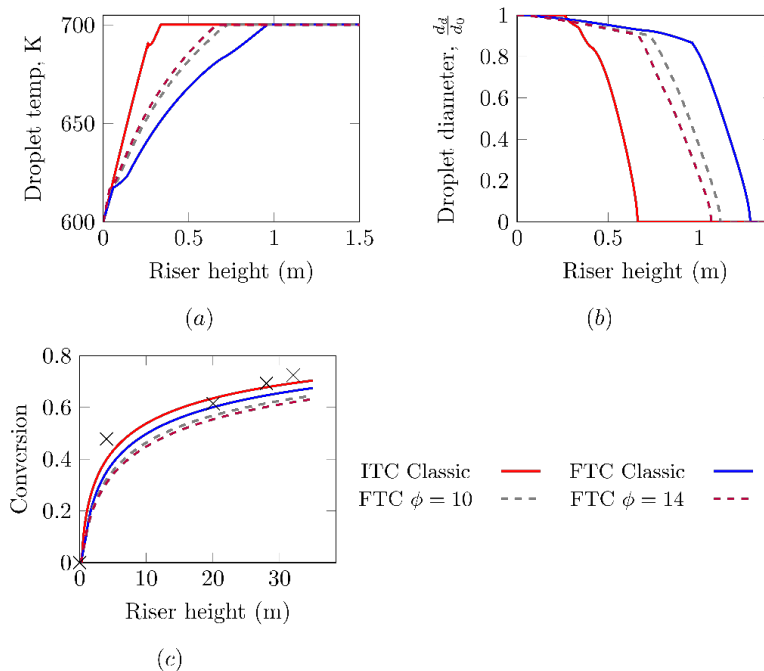




Figure 3:(a) Droplet temperature,  $T_d$ , (b) dimensionless droplet diameter,  $\frac{d_d}{d_0}$ , and (c) gas oil conversion for different models. All models use the same input values: ( $d_0 = 500\mu m$ ,  $T_{d0} = 600$ ,  $T_{vap} = 690$ ,  $T_{bp} = 700$ ,  $\rho_d = 810$ ,  $\varepsilon_{d0} = 0.012$ . For the FTC model framework, initial surface droplet temperature,  $T_{ds0} = 645$ ).

Using the conversion profile data, it is shown in Figure 3(c) that the ITC modelling framework best suits the data. It is also found in this work that this remains true for droplets  $< 2mm$ , beyond which the vaporisation starts slowing the cracking reactions.

#### 4. Conclusions

The ITC and FTC vaporisation modelling frameworks were investigated in the context of the FCC riser modelling and simulation. Both modelling frameworks are able to predict the penetration of the droplet feed up the riser from the point of feed injection. It is revealed that the ITC framework predicts faster vaporisation compared to the FTC, and also shown that the FTC over-predicts the energy transferred to the droplets over their lifetime and therefore leading to under-prediction of the conversion of feed. For most application, the available conversion profile data suggests that the ITC framework is recommended. Importantly within the ITC framework, the various homogeneous and heterogeneous Nusselt and Sherwood correlations do not affect the predicted yield and conversion profiles (which is only valid for  $d_{d0} < 2mm$ ). Beyond this droplet size range, the heterogeneous models should be used for better predictions of conversion.

#### 5. References

- Buchanan, J. S. (1994). "Analysis of Heating and Vaporization of Feed Droplets in Fluidized Catalytic Cracking Risers." *Industrial & engineering chemistry research* **33**(12): 3104-3111.
- Csemány, D. and V. Józsa (2021). "A Two-Parameter Corresponding States Method for Calculating the Steady-State Evaporation Rate of C2–C9 n-Alkane Droplets in Air for Elevated Pressures and Temperatures." *Flow, turbulence and combustion* **107**(2): 283-305.
- Derouin, C., D. Nevicato, M. Forissier, G. Wild and J.-R. Bernard (1997). "Hydrodynamics of Riser Units and Their Impact on FCC Operation." *Industrial & engineering chemistry research* **36**(11): 4504-4515.
- Fernandes, J. L., J. J. Verstraete, C. I. C. Pinheiro, N. M. C. Oliveira and F. R. Ribeiro (2007). "Dynamic modelling of an industrial R2R FCC unit." *Chemical Engineering Science* **62**(4): 1184-1198.
- Finneran, J. (2021). "On the evaluation of transport properties for droplet evaporation problems." *International journal of heat and mass transfer* **181**: 121858.
- John, Y. M., M. A. Mustafa, R. Patel and I. M. Mujtaba (2019). "Parameter estimation of a six-lump kinetic model of an industrial fluid catalytic cracking unit." *Fuel (Guildford)* **235**: 1436-1454.
- Lupo, G. (2017). "Detailed simulations of droplet evaporation."
- Miller, R. S., K. Harstad and J. Bellan (1998). "Evaluation of equilibrium and non-equilibrium evaporation models for many-droplet gas-liquid flow simulations." *International journal of multiphase flow* **24**(6): 1025-1055.
- Nayak, S. V., S. L. Joshi and V. V. Ranade (2005). "Modeling of vaporization and cracking of liquid oil injected in a gas–solid riser." *Chemical engineering science* **60**(22): 6049-6066.
- Ranz, W. E. (1952). "Evaporation from drops, Parts I & II." *Chem Eng Prog.* **48**: 141-146.

# A joint model-based design of experiments approach for the identification of Kriging models in geological exploration

Philipp Deussen<sup>a</sup>, Federico Galvanin<sup>a,\*</sup>

<sup>a</sup>*Department of Chemical Engineering, University College London, Torrington Place, London WC1E 7JE, United Kingdom*

*f.galvanin@ucl.ac.uk*

## Abstract

When exploring prospective mining locations, a central task is modelling rock attributes in the subsurface. The drilling needed to sample these locations is costly, so efficient sampling and reliable interpolation methods are needed. Kriging models (Gaussian Processes) are thus used, with the kernel and its parameters determined from data analysis of preliminary samples and expert judgement. New samples iteratively update the model, targeting exploitation (high ore grades) and exploration (minimising prediction variance).

The problem arises if the chosen kernel is incorrect or high uncertainty affects parameters. This paper thus suggests a joint model-based design of experiments (j-MBDoE) approach to target two objectives: maximising the distinguishability of candidate model predictions and reducing model uncertainty from parameter variance. Three different kernels in an Ordinary Kriging GP were used as candidate models. In-silico data was generated using one kernel and the optimal design strategy iteratively determined sampling locations to maximise model distinguishability with a constraint to ensure improved parameter estimates. Two models could be distinguished and the data approximated well with a limited number of drilling experiments while satisfactorily estimating kernel parameters.

**Keywords:** multiobjective optimization, joint model-based design of experiments, geostatistics, Gaussian processes, Kriging

## 1. Introduction

Geological exploration for metals is gaining relevance as industry seeks to secure supply chains for metals which are critical for industrial processes such as battery production. However, in the past decade, recovery rates and new discoveries have been low. The major bottleneck in mining value-dense metals (e.g. copper or nickel), is the identification and estimation of subsurface deposits. For resource and reserve models, Kriging models (Gaussian Processes) interpolate between sample data, which is obtained by drilling into the ground and taking sample cylinders of rock for visual and chemical analysis. Around targets of interest, closely spaced grids are drilled (infill drilling) for better estimates. Thus, optimal sampling can increase efficiency. Studies and industry practice show better performance of quantitative over qualitative design criteria (Jafrasteh and Suarez, 2020).

Model-based design of experiments (MBDoE) aims to design experiments for objectives such as obtaining the most informative data for parameter estimation (Franceschini and

Macchietto, 2008) or discrimination between candidate models (Schwaab et al., 2006). This paper proposes a joint experiment design approach (Galvanin et al., 2016) to locate samples to discriminate between three different candidate models for subsurface mineral concentrations while at the same time improving the precision of parameter estimates.

A metal concentration profile is generated in-silico by a function describing the spatial correlation of likely concentration grades and a preliminary model is fit based on ten initial samples. From there, the goal is to identify the valid model and its parameters more efficiently than methods such as closely spaced infill drilling. As a metric of success, model fit and parameter estimation statistics for candidate models may be compared.

## 2. Methodology

### 2.1. Fitting models using Gaussian Processes

A Gaussian Process (GP) interpolates between samples (Sahimi, 2011; Rossi & Deutsch, 2014). Its kernel is a correlation function  $R(h)$ , which relates the semivariance  $\gamma(h)$  of likely unsampled concentrations at distance  $h$  and the overall concentration variance  $\sigma_Z^2$ .

$$\gamma(h) = \sigma_Z^2 - R(h) \quad (1)$$

Ordinary Kriging (OK) is a form of GP (Sahimi, 2011) giving the best unbiased linear estimate of the mean expected concentration at unsampled locations  $Z$  using the estimator  $\hat{Z}$ , and the variance of estimates ( $\sigma_{OK}^2$ ) based on the expected square error to the mean.

$$\sigma_{OK}^2 = E[(Z - \hat{Z})^2] \quad (2)$$

The estimator is based on the samples  $Z_i$  and their relative importance weights  $w_i$ .

$$\hat{Z} = \sum_{i=1}^N w_i Z_i \quad (3)$$

The optimal weights  $w_i$  can be found by substituting (3) into (2):

$$\sigma_{OK}^2 = \sigma_Z^2 - 2 \sum_{i=1}^N w_i R(Z, Z_i) + \sum_{i=1}^N \sum_{j=1}^N w_i w_j R(Z_i, Z_j) \quad (4)$$

Introducing constraint (5) ensures the unbiasedness of the calculated predictions, as they no longer depend on a previous estimate of population variance. Substituting (5) into (4):

$$\sum_{i=1}^N w_i = 1 \quad (5)$$

$$\sigma_{OK}^2 = 2 \sum_{i=1}^N w_i \gamma(Z, Z_i) - \sum_{i=1}^N \sum_{j=1}^N w_i w_j \gamma(Z_i, Z_j) \quad (6)$$

Then, minimising the Kriging variance:  $\partial \sigma_{OK}^2 / \partial w_i = 0$ , results in:

$$\sum_{i=1}^N R(Z_i, Z_j) w_i = R(Z, Z_j) \quad (7)$$

This gives a solvable linear system of equations for all unsampled points. It comprises a matrix  $\mathbf{W}$ , for weights of samples at unsampled locations ( $n_s \times n_u$ ),  $\mathbf{P}$ , for correlations between these points ( $n_s \times n_u$ ) and  $\mathbf{A}$ , for correlations between sampled points ( $n_s \times n_s$ ).

$$\mathbf{W} = \mathbf{A}^{-1} \mathbf{P} \quad (8)$$

Finally, the predictions from (3) and (6) are multiplied with  $\mathbf{W}$  for all unsampled points.

### 2.2. Model-based design of experiments

The optimal experiment design comprises the objectives parameter estimation and model discrimination. Schwaab et al. (2006) propose a discrimination criterion,  $\Psi^{\text{MD}}$  (9), based on the difference in responses  $\hat{Z}$ , their uncertainty  $\sigma$  and the probabilities  $P$  of models  $m$

and  $n$ . New samples reveal information in a Bayesian sense, since, once sampled, the GP must reflect sample values and the variance at the sampled point collapses to zero. So, instead of estimating the probabilities,  $P_m$  (11-12), using the  $\chi^2$  value, the total prediction variance,  $V$  (10) is used. For each coordinate  $(x, y)$  and set of kernel parameters  $\vartheta$ :

$$\Psi^{MD}(x, y, \vartheta) = \sum_{m=1}^{M-1} \sum_{n=m+1}^M (P_m, P_n) [\hat{\mathbf{Z}}_m(x, y, \vartheta) - \hat{\mathbf{Z}}_n(x, y, \vartheta)]^T \mathbf{V}_{m,n}^{-1}(x, y, \vartheta) [\hat{\mathbf{Z}}_m(x, y, \vartheta) - \hat{\mathbf{Z}}_n(x, y, \vartheta)] \quad (9)$$

$$V_{m,n}(x, y, \vartheta) = 2V(x, y) + V_m(x, y, \vartheta) + V_n(x, y, \vartheta) \quad (10)$$

$$\phi_m(\vartheta) = 1 / \sum_{x=1}^X \sum_{y=1}^Y \sum_{k=1}^K V_m(x, y, k, \vartheta) \quad (11)$$

$$P_m(\vartheta) = \frac{\phi_m(\vartheta)}{\sum_{m=1}^M \phi_m(\vartheta)} \quad (12)$$

To quantify parametric uncertainty (Franceschini and Macchietto, 2008), the kernel parameters  $\boldsymbol{\vartheta}$  are perturbed from their initial estimates and the impact on the measured responses (Kriged concentration values),  $\mathbf{Z}$ , is stored in the sensitivity matrix  $\mathbf{Q}$  (13).

$$\mathbf{Q}_{x,y}(\mathbf{Z}, \boldsymbol{\vartheta}) = \begin{bmatrix} \frac{\partial Z_1}{\partial \vartheta_1} \\ \vdots \\ \frac{\partial Z_1}{\partial \vartheta_n} \end{bmatrix} \approx \begin{bmatrix} \frac{\hat{Z}_1 - \hat{Z}_1'}{\vartheta_1 - \vartheta_1'} \\ \vdots \\ \frac{\hat{Z}_1 - \hat{Z}_1'}{\vartheta_n - \vartheta_n'} \end{bmatrix} \quad (13)$$

However, in a GP, the information at a point depends on both the predicted mean and variance. A point with high variance is less informative than a point with low variance. Thus, analogous to the local uncertainty in  $\Psi^{MD}$ , an uncertainty (14) (standard deviation) for each model response and parameter perturbation is defined and stored in  $\mathbf{Q}^{Adj}$  (15).

$$\sigma_{\text{total}} \frac{\partial \hat{Z}_n}{\partial \vartheta_n}(x, y) = 2 \sigma_{\text{exp}}(x, y) + \sigma_{\hat{Z}_n, \vartheta_n}(x, y) + \sigma_{\hat{Z}_n', \vartheta_n'}(x, y) \quad (14)$$

$$\mathbf{Q}_{x,y}^{Adj}(\mathbf{y}, \boldsymbol{\vartheta}) = \begin{bmatrix} 1 & \frac{\partial y_1}{\partial \vartheta_1} \\ \sigma_{\text{total},1} & \vdots \\ 1 & \frac{\partial y_1}{\partial \vartheta_n} \\ \sigma_{\text{total},n} & \vdots \end{bmatrix} \approx \begin{bmatrix} \frac{1}{(2\sigma_{\text{exp}} + \sigma_{y_1, \vartheta_1} + \sigma_{y_1, \vartheta_1'})} \left( \frac{y_1 - \hat{y}_1}{\vartheta_1 - \vartheta_1'} \right) \\ \vdots \\ \frac{1}{(2\sigma_{\text{exp}} + \sigma_{y_n, \vartheta_n} + \sigma_{y_n, \vartheta_n'})} \left( \frac{y_1 - \hat{y}_1}{\vartheta_n - \vartheta_n'} \right) \end{bmatrix} \quad (15)$$

The Fisher information,  $\mathbf{H}_{x,y}(\boldsymbol{\vartheta})$  (16), is the local information about parameters  $\boldsymbol{\vartheta}$ . Its trace gives a scalar criterion known as A-optimality (Franceschini and Macchietto, 2008).

$$\mathbf{H}_{x,y}(\boldsymbol{\vartheta}) = \mathbf{Q}_{x,y}^{Adj} \mathbf{Q}_{x,y}^{Adj T} \quad (16)$$

The joint design criterion,  $\phi^{ID}$  (17), by Galvanin and co-workers (2016) combines  $\Psi^{MD}$  (9) and an A-optimal parameter precision criterion,  $\Psi^{PE}$  (18).

$$\phi^{ID} = \text{argmax}\{\Psi^{MD}\}_{\phi \in D} \quad (17)$$

s.t.

$$\Psi^{PE} = \sum_{j=1}^{N_M} \|\mathbf{H}_j\| / N_M \leq \varepsilon \quad (18)$$

$$\begin{aligned} \varepsilon_{\min} &= \max(\Psi^{PE}) & \text{s.t.} & \max(\Psi^{MD}) \\ \varepsilon_{\max} &= \max(\Psi^{PE}) & \text{s.t.} & \max(\Psi^{PE}) \end{aligned}$$

That is,  $\varepsilon_{\min}$  ensures that the design criterion addresses model discrimination, while  $\varepsilon_{\max}$  allows solutions that fully tilt towards parameter estimation. A slack variable,  $s_0$ , is added to avoid weakly efficient solutions (Mavrotas, 2009). The value of  $\varepsilon$  can be moved between the two extremes in order to explore the Pareto optimal set of suggested samples.

### 3. Case Study

#### 3.1. Definition of candidate models

The candidate models are defined through the correlation function used in the kernel. In these kernels, the parameter  $s$  is the partial sill of semivariogram, which is the maximum semivariance (or minimum correlation) the probabilities of mineral concentration at two points in space will reach, minus the nugget effect. The nugget effect,  $n$ , is the difference in measured value from one point to a point just adjacent to it (on a smaller scale than the model can capture),  $r$  is the range (characteristic length at which the maximum variance between the probabilities of mineral concentration at two points in space is reached) and  $a$  is a numerical parameter. With the bounds  $\delta_{(i,j)}(h)$  represented in a Boolean manner:

$$\delta_{(i,j)}(h) = \begin{cases} 1 & \text{if } h \in (i, j) \\ 0 & \text{otherwise} \end{cases}$$

Table 1 – Kernels used in formulation of candidate models

Type of Kernel (Source)	Kernel Expression
<b>M1:</b> Gaussian (Rossi & Deutsch, 2014)	$\gamma(h) = \left\{ (s) \left( 1 - e^{-\frac{1}{3} \frac{h^2}{r^2}} \right) \right\} \delta_{(0,r)}(h) + s \delta_{(r,\infty)}(h) + n \delta_{(0,\infty)}(h)$
<b>M2:</b> Exponential (Rossi & Deutsch, 2014)	$\gamma(h) = \left\{ (s) \left( 1 - e^{-\frac{1}{3} \frac{h}{r}} \right) \right\} \delta_{(0,r)}(h) + s \delta_{(r,\infty)}(h) + n \delta_{(0,\infty)}(h)$
<b>M3:</b> Spherical (Sahimi, 2011)	$\gamma(h) = \left\{ (s) \left( \frac{3h}{2r} - \frac{1}{2} \left( \frac{h}{r} \right)^3 \right) \right\} \delta_{(0,r)}(h) + s \delta_{(r,\infty)}(h) + n \delta_{(0,\infty)}(h)$

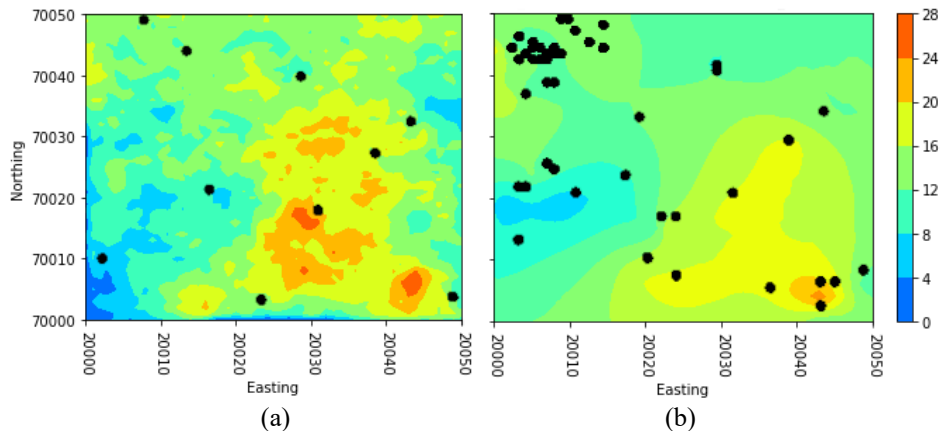


Figure 1 – Metal concentration (g/t) ground truth with initial samples marked in black (a) and (b) final model prediction of metal concentration with all samples marked in black.

### 3.2. Definition of Ground Truth and Design Space

The ground truth (Figure 1) was generated using the spherical kernel (M3 in Table 1) with a range  $r = 40$  hm, sill of  $s = 40$  (g/t)<sup>2</sup>, and nugget effect of  $n = 1$  (g/t)<sup>2</sup> over an area of 51E and 51N (5.1x5.1 km). The design variables are the coordinates of the new sample points (x,y). Ten preliminary samples were taken using Latin Hypercube Sampling.

### 3.3. Experimental Design

Fifteen optimal experiment designs were carried out based on the objectives in (17-18), adding 38 new sampling locations to the 10 preliminary samples. In each iteration, up to three samples were selected from the Pareto front (containing the optimal trade-offs between both objectives). One point was selected at either extreme ( $\epsilon_{\min}$  and  $\epsilon_{\max}$ ) and a third in the middle, unless the maxima for both objectives coincided or were so close that the grid resolution permitted selecting only one location.

### 3.4. Model Identification Results

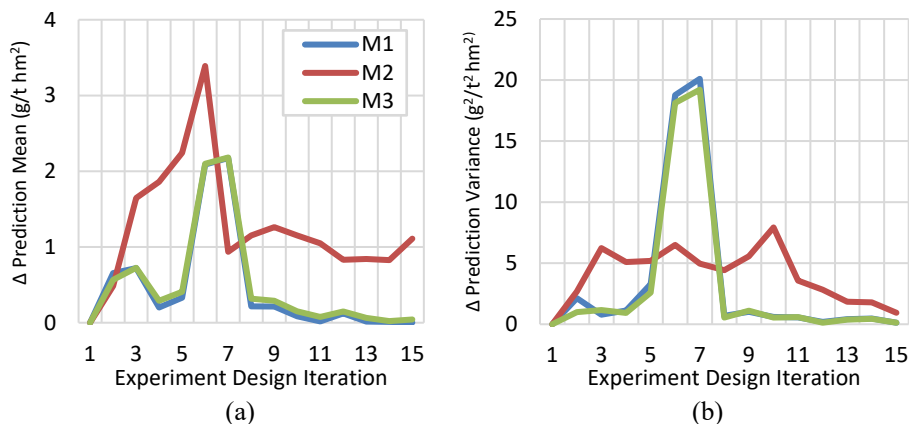


Figure 2 – Change in predicted mean (a) and variance (b) per hectare after 15 designs.

Each design iteration (1-15 in Figure 2) adds new samples that update  $\mathbf{W}$  and with it the predicted mean and variance. The change in their magnitude shows how similar model predictions and new samples are, i.e. a model’s predictive ability. As shown in Figure 2, M1 and M3 predict most new samples well and respond to them indistinguishably, because both kernels produce correlation structures that have similar shapes (and reflect the ground truth). The exception is a spike that results from suddenly improved parameter estimates (shown in Figure 3) that allow the model to capture a previously unknown characteristic of the data. M2 responds differently and shows constant large fluctuations, indicating a failure to represent the ground truth regardless of the design iteration.

The variance of parameter estimates is steadily reduced by new samples and starts converging after 10 design iterations as shown in Figure 3. Although the variances on the estimates for M1 are lower, they converge towards biased values. For M2, the variances are considerably higher, but they converge to the true value for the range and nugget effect. The estimates of the sill parameter, which is related to the maximum variability in the sample population, is systematically underestimated in all models. From Figure 1 (a) and (b) it becomes apparent that this is due to no samples hitting the small areas of tail concentration values. This is a known problem with Kriging interpolation, since tail values are likely to be less represented in the sample set (Rossi & Deutsch, 2014).

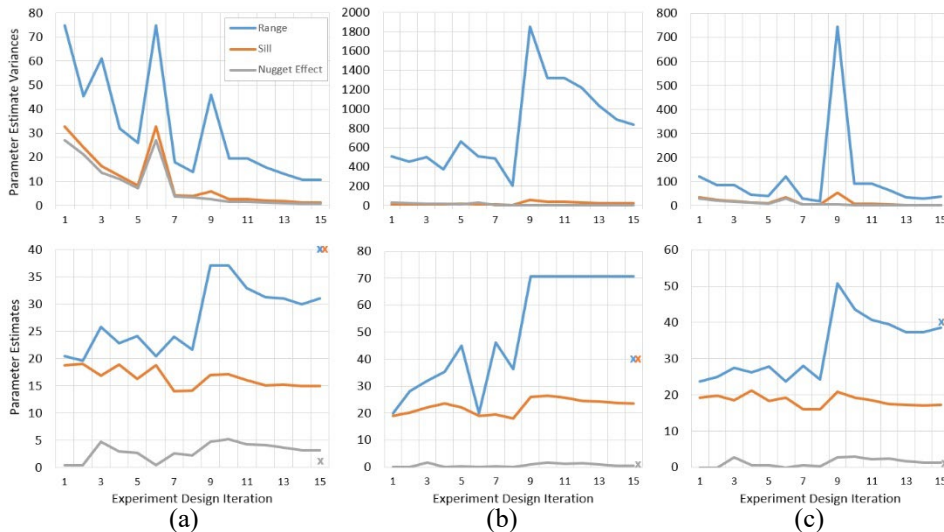


Figure 3 – Parameter estimates: variance (upper plots) and mean (lower plots) for M1 (a), M2 (b), M3 (c) following 15 sample designs. True parameter values are marked by the x.

#### 4. Conclusions

A multiobjective MBDoE approach was applied to 3 Kriging models of an in-silico metal deposit, generated using one of the 3 models. The behaviour of model predictions vs. new samples allowed the exclusion of model M2 that did not correspond to the ground truth. Two models (M1 & M3) had similar predictions and were thus indistinguishable. Parameter estimates steadily improved with design iterations. Only the sill (variance-related) was underestimated in all models, as the tail ends of the concentration spectrum were insufficiently sampled, which is a common problem with Kriging models in geological modelling. The parameter estimates converged to true values in M3 but to biased values in M1. The approach succeeded at improving both design objectives, identifying two models that could have generated the data. Future work will aim to improve the estimation of the sill, e.g. by aiming for high-grade samples, and improve the discrimination capability for similar model outputs, e.g. through explorative designs.

#### References

- Franceschini, G. and Macchietto, S., 2008, Model-based design of experiments for parameter precision: State of the art. *Chem Eng Sc*, 63, 4846-4872.
- Galvanin, F., Enhong, C., Al-Rifai, N., Gavriilidis, A. and Dua, V. (2016) A joint model-based design approach for the identification of kinetic models in continuous flow laboratory systems. *Comp Chem Eng*, 95, 202-215.
- Jafrasteh, B. and Suarez, A. (2021) Objective functions from Bayesian optimization to locate additional drillholes. *Computers and Geosciences*, 147.
- Rossi, M. and Deutsch, C. (2014) *Mineral Resource Estimation*. 1<sup>st</sup> Ed., Springer.
- Mavrotas, G. (2009) Effective implementation of the  $\epsilon$ -constraint method in multi-objective mathematical programming problems. *Appl. Mathematics and Computation*, 213, pp 455-465.
- Sahimi, M. (2011) *Flow and Transport in Porous Media*. 2<sup>nd</sup> Ed., Wiley-VCH Verlag.
- Schwaab, M., Silva, F.M., Queipo, C.A., Barreto, Jr., A.G., Pinto, J.C. (2006) A new approach for sequential experimental design for model discrimination. *Chem.Eng.Sci.*, 61, 5791-5806.

# Reinforcement Learning for inventory management in multi-echelon supply chains

Guoquan Wu,<sup>c</sup> Miguel Ángel de Carvalho Servia,<sup>b</sup> Max Mowbray<sup>a</sup>

<sup>a</sup> *Department of Chemical Engineering, University of Manchester, Oxford Road, Manchester, M1 3AL, UK.*

<sup>b</sup> *Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, London, SW7 2AZ, United Kingdom*

<sup>c</sup> *Department of Chemical and Biomolecular Engineering, National University of Singapore, 117585, Singapore*

## Abstract

Reinforcement Learning (RL) describes a set of model-free learning rules that enable identification of an approximately optimal feedback control policy function approximation for discrete, time uncertain process systems purely from data. This has led to interest in application of RL to supply chain management problems. However, many RL methods are dependent on estimating noisy first-order directions for policy improvement, and have been developed for video games and computer science applications. Here, we propose a deep RL method tailored for supply chain management. The algorithm deploys a derivative free approach to balance exploration and exploitation of a neural policy's parameter space, providing means to avoid low quality local optima. Additionally, the policy inherits ease in posing risk-sensitive formulations to learn a policy that respects constraints on the conditional value-at-risk. The performance of our algorithm is tested on a benchmark multi-echelon supply chain inventory management problem. The results demonstrate empirical improvements in performance over the first-order RL method, proximal policy optimization, and competitive performance with mathematical programming. Additionally, the risk-sensitive decisions generated can effectively handle low probability, high severity scenarios.

**Keywords:** Reinforcement Learning, multi-echelon supply chain management, inventory optimization,

## 1. Introduction

In recent years, many researchers have proposed methods to incorporate data-driven approaches into strategic and tactical supply chain, and scheduling, primarily for the associated benefits in accounting for uncertainty in decision-making. For example, in the context of long-term planning, Gao *et al.* (2019) proposed a two-stage distributionally robust optimization model, which takes advantage of industrial data to address the optimal design and operations of shale gas supply chains under uncertainty and distributional ambiguity. Distributionally robust optimization has also been exploited for tactical healthcare supply chain management as highlighted in Ash *et al.* (2022). The authors note how the hyperparameters of this approach (i.e. the size of the ambiguity set) could be used to bias the algorithm towards average cost performance or worst-case cost performance. This is a key observation, which underpins the flexibility of distributionally robust approaches. Others works have relaxed the assumption of distributional ambiguity and pursued stochastic optimization approaches. For example, Golari *et al.* (2017),



explored the development of a multi-stage stochastic programming model for inventory control under uncertainty in a renewable energy supply. Robust and multi-parametric optimization has also been investigated (Charitopoulos *et al.*, 2019).

In the context of tactical supply chain management, traditional operations research methods are still posed with practical challenges. In these problems, we are interested in identifying recourse decisions as the system's uncertainty is realized. The challenges primarily relate to accounting for exogenous and endogenous uncertainty in closed loop decisions and the tractability of solving the associated mathematical program recursively. These challenges are compounded by complex network relationships present within the supply chain, and the size of the decision problem tends to grow combinatorially with the supply chain size.

More recently, research focus has been directed towards the development of Reinforcement Learning frameworks to overcome some of the challenges mentioned. For example, a recent work by Hubbs *et al.* (2020) presented a suite of benchmark problems for multi-echelon supply chain inventory management problems and studied the performance of various nominal mathematical programming implementations. Perez *et al.* (2021) extended this work to study divergent multi-echelon supply chains and included benchmark to a number of stochastic mathematical programming implementations. Both works demonstrated application of the RL algorithm, proximal policy optimization (PPO), which relies on estimating first-order directions for policy improvement, and make use of overparameterised policy function approximations. In this work, we attempt to improve the performance of RL and investigate the identification of risk-sensitive policies by making explicit connection between RL and derivative free optimization.

## 2. Reinforcement Learning and Derivative Free Optimization

### 2.1. Markov Decision Processes

Reinforcement Learning describes a set of algorithms, which generally aim to parameterize an approximately optimal policy function approximation that provides control decisions in a closed loop feedback framework, known as the Markov Decision Process (MDP). Finite horizon MDPs are fully defined by a state space,  $\mathbb{X} \subseteq \mathbb{R}^{n_x}$ ; a control space,  $\mathbb{U} \subseteq \mathbb{R}^{n_u}$ ; a conditional probability density function,  $P_{xx'}: \mathbb{X} \times \mathbb{U} \times \mathbb{X} \rightarrow \mathbb{R}_+$ , descriptive of the uncertain, discrete time evolution of the system, from the current state,  $\mathbf{x}_t$ , and control,  $\mathbf{u}_t$ , to a future state,  $\mathbf{x}_{t+1}$ ; a stage cost function  $\varphi: \mathbb{X} \times \mathbb{U} \times \mathbb{X} \rightarrow \Phi_{t+1} \in \mathbb{R}$ , that ranks closed loop decisions at each control interaction; and a finite, discrete time horizon, such that  $t \in \{0, \dots, T\}$ . As decisions are identified within the context of state feedback, optimal decisions are provided by a policy,  $\pi: \mathbb{X} \rightarrow \mathbb{U}$ , which minimizes the expected sum of stage costs,  $Z = \sum_{t=0}^T \Phi_{t+1}$  such that:

$$\pi^* = \underset{\pi \in \Pi}{\operatorname{argmin}} \mathbb{E}_\pi[Z] \quad (1)$$

where  $\pi$  defines a policy that belongs to the set of policies,  $\Pi$ , that satisfy input constraints. It is also worth briefly noting that this implies  $Z$  is a random variable described via some distribution,  $Z \sim p_\pi(z)$ , with values realized together with a given process trajectory,  $\boldsymbol{\tau} = (\mathbf{x}_0, \mathbf{u}_0, \dots, \mathbf{x}_T)$ . In this work, we assume that the problem of identifying a centralized re-order policy for a multi-echelon supply chain can be posed as an MDP.

## 2.2. Reinforcement Learning and derivative-free optimization

RL describes a set of policy learning methods that aim to, in general, identify a parametric policy function,  $\pi(\mathbf{x}; \theta)$ , to approximate the optimal policy. This enables decision-making to generalize across continuous state spaces. The problem then reduces to identifying the policy parameters, such that:

$$\theta^* = \operatorname{argmin}_{\theta \in \Theta} \mathbb{E}_{\pi} [Z] \quad (2)$$

where  $\theta \in \mathbb{R}^{n_{\theta}}$  are the policy parameters and  $\Theta$  denotes the set of allowable policy parameters. The policy parameterization itself is identified through model-free parameter update rules. These are purely data-driven rules that estimate first-order directions to enable local gradient descent steps to improve the performance of policy parameters. However, importantly, although these directions may be estimated independently of gradient information from the dynamics themselves, they are often noisy and expensive to estimate. Additionally, the ultimate policy identified is known to be sensitive to policy initialization, indicating the optimization landscape is often characterized by multiple poor local minima.

Here, it is worth highlighting the strong connections between RL and derivative-free optimization (DFO). In DFO, one aims to identify an approximately optimal solution for a problem independently of gradient information, but rather purely via function evaluations. This is also property of DFO that enables it to optimize measures of objectives that are functions of random variables via, for example, a Monte Carlo method. Popular algorithms typically maintain a population of candidate solutions, evaluate the objective performance of each and then utilize information from the population to coordinate exploration and exploitation of the parameter space. This iterative process enables the DFO algorithm to identify better performing candidate solutions until a computational budget is exhausted or an exit tolerance satisfied. Examples include particle swarm optimization (PSO), evolutionary strategies (ES), and artificial bee colony (ABC), which can all be considered as heuristics inspired by phenomena observed within nature. It is well known that some algorithms are more explorative or exploitative than others. Hybridization of these algorithms has often been used to combine the respective strength of each algorithm to in one search process. Despite the potential application, these algorithms are known to perform poorly in highly dimensional spaces of  $n_{\theta} \geq 1000$ . This means we are required to construct a relatively low dimensional policy parameterization, given current RL practice is to massively overparameterize policy function approximations.

## 2.3. Risk-sensitive decision-making

As reflected in the discussion above, when deploying methods that consider uncertain variables in decision-making, typically one is posed with optimizing for some measure of the objective function. The choice of measure is typically guided by potential risk-tolerance of the operation. Optimizing in expectation is a popular choice, however, in some operations it may be desirable to consider what is happening in the tails of policy performance. For example, the conditional value-at-risk ( $CVaR_{\beta}$ ) defines the expected performance of the policy observed with at least a probability level  $\beta = (0,1)$  under the cumulative distribution function,  $F_{\pi}(z) = \mathbb{P}(Z \leq z | \pi)$  associated with the policy:

$$CVaR_{\beta} = \mathbb{E}_{\pi} [Z | Z \geq F_{\pi}^{-1}(\beta)] \quad (3)$$

$F_{\pi}^{-1}(\beta)$  defines the inverse cumulative distribution function evaluated for probability  $\beta$ . Risk sensitive formulations that leverage the conditional value-at-risk are widely discussed elsewhere. However, in this work, we are interested in a formulation that incentivizes performance in expectation, but provides some notion of probabilistic robustness as follows:

$$\theta^* = \underset{\theta \in \Theta}{\operatorname{argmin}} \mathbb{E}_{\pi}[Z] \quad \text{s. t. } \operatorname{CVaR}_{\beta} \leq b \quad (4)$$

where  $b \in \mathbb{R}$  is a bound that can be set by the implementation together with the probability level,  $\beta$ . This formulation is particularly appealing in supply chains because, for example, it enables one to place a bound on the expected costs observed with a given probability level under the policy.

#### 2.4. Derivative-free optimization for robust reinforcement learning in multi-echelon supply chain inventory management

In this work, we investigate the application of derivative-free optimization for policy learning. Specifically, we hybridize standard in-house python implementations of ABC, PSO, ES and simulated annealing (SA) to more efficiently explore and exploit the parameter space.

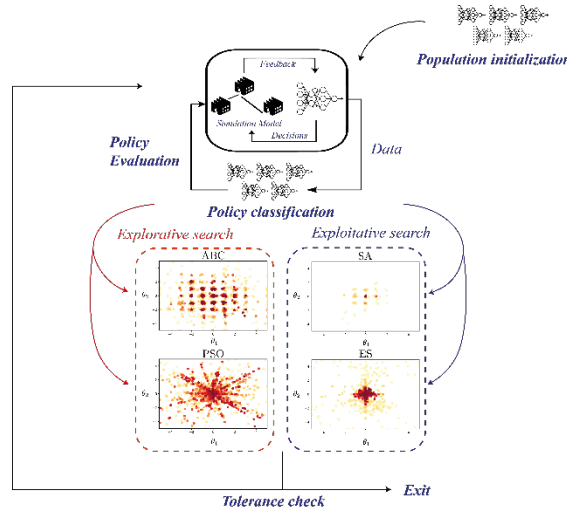


Figure 1: Figurative description of the proposed hybrid derivative free optimization algorithm. Poor performing candidate policies are passed to explorative search methods and high performing policies are passed to more exploitative methods.

Firstly a population of candidate policies is initialized. At each iteration of the algorithm, each of the candidate policies is evaluated within the supply chain inventory management problem and then the population is divided into disjoint sets according to policy performance. For example, the top 25% of policies comprise one subset, the policies ranked in the top 50-25% comprise an additional subset, and so on. The subsets of policies are then passed to one of the given DFO algorithms (ABC, ES, PSO and SA) depending on performance. The worst performing policies are passed to the most explorative algorithms, whereas the best are passed to the most exploitative. In this case, the bottom 50% are passed to ABC and PSO, with the top 50% passed to SA and ES. This is demonstrated by Fig.1.

In case study, we seek to benchmark the performance of our algorithm in a multi-echelon supply chain inventory management problem with integer reorder decisions against the state-of-the-art RL method, proximal policy optimization (PPO), and mathematical programming implementations including offline nominal MILP (i.e. with no recourse decisions) and online shrinking horizon linear programming (with assumption that the integer constraints on decisions can be relaxed in the model and then the solution rounded appropriately).

### 3. Case Study

#### 3.1. Operations Research Gym Benchmark

Here, we utilize the benchmarks and associated code provided by Hubbs *et al.* (2020). Specifically, we consider a supply chain network with  $n_s = 4$  production stages. Stage 0 deals with exogenous consumer demand uncertainty described via a Poisson distribution with a rate of 20 units of inventory. Satisfaction of this demand is handled via a central reorder policy that places inventory order requests from stage  $n$  to  $n+1 \forall n \in \{0, n_s - 1\}$ . Stage  $n_s$  is assumed to have an infinite inventory to supply demand from stage  $n_s - 1$  (and so in principle does not maintain a reorder policy itself). Production and transportation of the product is captured via a deterministic, constant delivery (lead) time.

The problem is defined as an MDP with discrete time dynamics described in Hubbs *et al.* (2020). The state is defined to preserve the Markov property and capture pipeline production and transportation. Specifically, the reorder history of the previous  $L_{max}$  time steps, where  $L_{max}$  is the maximum lead time associated with the delivery of product from all stages, are concatenated together with the current inventory level for stages with finite inventory. In this study  $L_{max} = 10$ . As a result, the state variables,  $\mathbf{x} \in \mathbb{X} \subseteq \mathbb{R}^{(n_s-1)(L_{max}-1)}$ , are defined by the current inventory and the history of  $L_{max}$  controls. The controls,  $\mathbf{u} \in \mathbb{U} \subseteq \mathbb{R}^{n_s-1}$ , are the reorder quantities made at each stage in the supply chain, which are integer values and bounded between zero and an upper bound, which is stage dependent and defined as  $\mathbf{u}_{ub} = [100, 90, 80]$ , respectively. The aim of supply chain management is to maximize the expected net present value (NPV) of the supply chain after 30 discrete time steps. At each time step the cost function,  $\varphi$ , allocates the negative of discounted profit made by the total chain in the most recent time increment, which is comprised by total sales and various costs of holding inventory, procurement and unfulfilled demand at the respective stages.

#### 3.2. Results

Here, we investigate two different formulations. The first formulation (F1) is provided by Eq. 2 and the second by Eq. 4 (F2). The optimization proceeds through unbiased Monte Carlo estimates of the objective measures. In the second formulation, we set the CVaR constraint to  $b = -380$  and probability level  $\beta = 0.9$ .

It is clear from Fig. 2a) that the derivative free approach is able to outperform the policy gradient, PPO method, over the same training budget. Over 1000 Monte Carlo evaluations, the final policy of the hybrid approach observed expected NPV of \$454.0, whereas the offline MILP strategy presented in Hubbs *et al.* (2020) achieved \$378.5, and the shrinking horizon LP achieved \$485.4. In Fig 2b), the results of the approach under the risk sensitive formulation (F2) is presented. It is clear that hybrid algorithm makes progress to maximize expected NPV (Monte Carlo evaluation of the final policy achieves

\$425) and satisfy the constraint on the CVaR (Monte Carlo evaluation of the final policy achieves \$392.4). This provides probabilistic robustness in the policy performance. However, this robustness is at the cost of a greater required computational budget in policy optimization (over formulation F1), and reduced performance in expectation.

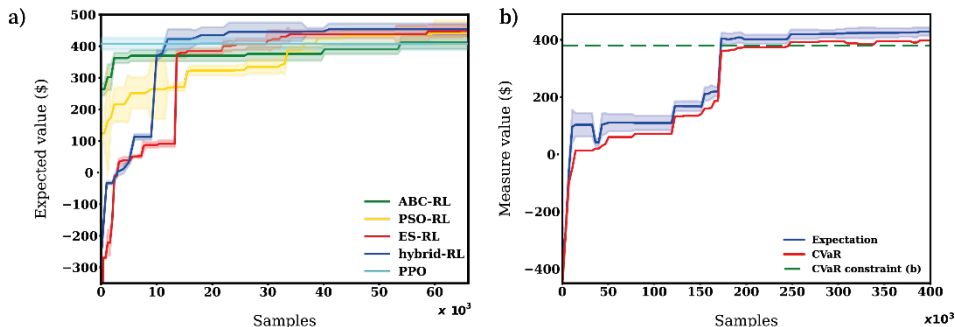


Figure 2: Policy optimization progress of a) the constituent stochastic search algorithms, together with the hybrid DFO algorithm (hybrid-RL) benchmarked against the final results of the state-of-the-art policy gradient method, PPO, evaluated over the same number of samples in F1; b) of the hybrid algorithm under F2. Presented from the perspective of NPV maximization.

#### 4. Conclusions

In this work, we have demonstrated that derivative-free optimization methods are capable approaches to policy optimization tasks in RL in problems observed commonly in PSE and that compact, well-constructed policy function approximations can perform competitively with mathematical programming. RL with function approximation typically provides a nonlinear optimization problem and so derivative free approaches have value in e.g. escaping local minima. Additionally, they provide flexibility in optimizing risk-sensitive formulations via unbiased Monte Carlo estimators. However, it is likely that the limitations of derivative-free approaches in high dimensions will become substantial in industrial-scale supply chain networks. In future work, we will explore representation learning methods and exploit the structure of the supply chain network to reduce the effective dimensionality of these instances. Additionally, thorough comparison to mathematical programming formulations which consider uncertainty will be made.

#### References

- [1] Gao, J., Ning, C. and You, F., 2019. Data-driven distributionally robust optimization of shale gas supply chains under uncertainty. *AIChE Journal*, 65(3), pp.947-963.
- [2] Ash, C., Diallo, C., Venkatadri, U. and VanBerkel, P., 2022. Distributionally robust optimization of a Canadian healthcare supply chain to enhance resilience during the COVID-19 pandemic. *Computers & Industrial Engineering*, 168, p.108051.
- [3] Golari, M., Fan, N. and Jin, T., 2017. Multistage stochastic optimization for production-inventory planning with intermittent renewable energy. *Production and Operations Management*, 26(3), pp.409-425.
- [4] Charitopoulos, V.M., Papageorgiou, L.G. and Dua, V., 2019. Closed-loop integration of planning, scheduling and multi-parametric nonlinear control. *Computers & Chemical Engineering*, 122, pp.172-192.
- [5] Hubbs, C.D., Perez, H.D., Sarwar, O., Sahinidis, N.V., Grossmann, I.E. and Wassick, J.M., 2020. Or-gym: A reinforcement learning library for operations research problems. *arXiv preprint arXiv:2008.06319*.
- [6] Perez, H.D., Hubbs, C.D., Li, C. and Grossmann, I.E., 2021. Algorithmic approaches to inventory management optimization. *Processes*, 9(1), p.102.

# Monte Carlo Simulation of Photo Induced Atom-Transfer Radical Polymerization for Microscopic Properties

Rui Liu<sup>a</sup>, Xiaowen Lin<sup>a</sup>, Antonios Armaou<sup>b, c, d</sup>, Xi Chen<sup>a, e</sup>

<sup>a</sup>*State Key Laboratory of Industrial Control Technology, College of Control Science and Engineering, Zhejiang University 310027, Hangzhou China.*

<sup>b</sup>*Chemical Engineering Department, University of Patras, Patras 26504, Greece.*

<sup>c</sup>*Chemical Engineering & Mechanical Engineering Departments, Pennsylvania State University, College Park, PA 16802 USA.*

<sup>d</sup>*College of Mechanical Engineering, Wenzhou University, 325035 Wenzhou China.*  
[armaou@psu.edu](mailto:armaou@psu.edu)

<sup>e</sup>*National Center for International Research on Quality-targeted Process Optimization and Control, Zhejiang University 310027, Hangzhou China.*  
[xi\\_chen@zju.edu.cn](mailto:xi_chen@zju.edu.cn)

## Abstract

Controlled radical polymerization (CRP) is a versatile way to control polymer microscopic structures through establishing a dynamic equilibrium between radical and dormant polymers. Atom transfer radical polymerization (ATRP), a typical case of CRP, has the ability to be conducted in mild conditions and produce uniformly distributed polymers. Moreover, the amount of metal addition can be greatly reduced in ATRP processes with the development of photochemical techniques. Photo induced ATRP (photoATRP) has a significant impact on sustainable production in industry. Modeling the microscopic quantities of polymers for photoATRP systems provides insights into the kinetic characteristics. In this work, a Monte Carlo (MC) approach, based on Gillespie's stochastic simulation algorithm, is developed for the photoATRP processes at microscopic resolution. The length dependency of the individual chains on the radical, dormant, and dead polymer information is captured. The polymerization rate of PhotoATRP is sensitive to light irradiation. The polymerization process model of photoATRP systems based on the MC method for temporal control over polymer growth is established. Kinetic characteristics are studied from the developed model under the condition of light source intermittency. The developed MC model is validated against the deterministic method of moments (MoM) and experimental results.

**Keywords:** Monte Carlo simulation; photochemical ATRP; microscopic properties; hybrid multiresolution model.

## 1. Introduction

In past decades, there has been a continuous evolution in controlled radical polymerization (CRP) for enabling to produce well-controlled polymers with narrow molecular weight distributions (MWD) (Pan et al., 2016). Atom transfer radical polymerization (ATRP) is a fundamental technique of CRP. The growth rate of radical polymers is controlled by manipulating activation and deactivation reactions in ATRP (Wang et al., 1995). Light irradiation as external stimuli maintains the stability of polymer

growth through regenerating activators in ATRP systems (Ribelli et al., 2014). These photo induced ATRP (photoATRP) systems provide well-defined and uniformly distributed polymers under mild conditions (Lorandi et al., 2022). The temporal and spatial control over polymer growth in photoATRP can be efficiently achieved upon regulating the applied light irradiation (Dadashi-Silab et al., 2020).

Identifying the microscopic properties of polymers is of great significance and demand for photoATRP processes. Monte Carlo (MC) simulation is a powerful tool to accurately predict the microscopic structure of polymers and provides insights into the kinetic characteristics (Liu et al., 2021). In this work, a dynamic MC model is developed for photoATRP systems at microscopic resolution. Kinetic characteristics of temporal control with light switching are provided based on the developed MC model.

## 2. PhotoATRP systems

In ATRP systems, transition metal complexes and alkyl halides are used to establish the equilibrium between activation and deactivation reactions. Low oxidation state metal complexes as activators activate dormant polymers, and high oxidation state metal complexes as deactivators transform radical polymers into dormant polymers. The mechanistic scheme of the ATRP processes is presented in Figure 1(a). The oxidation-reduction reaction pair presents a trend in the direction of activation in ATRP processes due to the fact that radical polymers are constantly consumed in termination reactions. The concentration of activators decreases as the reaction progress. To maintain the stability of the equilibrium between activation and deactivation, and the concentrations between radical polymers and dormant polymers, a continuous supplement of activators to the reaction system is required. However, in photoATRP processes, activators can be regenerated through light irradiation, as indicated in Figure 1(b). The amount of metal catalyst loadings is decreased in photoATRP (Chen et al., 2016).

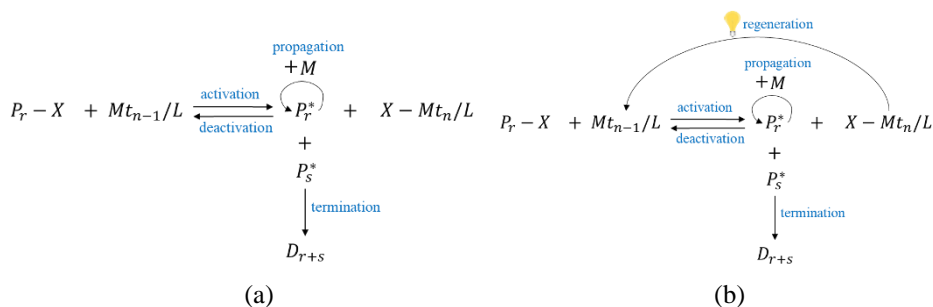


Figure 1 Mechanistic schemes for atom transfer radical polymerization (ATRP) processes. (a) Scheme of conventional ATRP. (b) Scheme of photoATRP.

Radical polymers are formed along with the regeneration of activators in photoATRP processes shown in Figure 1(b). Alkyl halides, ligands, and monomers can also lead to the formation of radicals in a photochemical way. A typical kinetic mechanism for photoATRP is listed in Table 1, where  $P_sX$ ,  $D_s$  and  $P_s^*$  denote dormant polymers, dead polymers, and radical polymers with chain lengths of  $s$ , respectively;  $Mt_{n-1}/L$  and  $Mt_nX/L$  denote activators and deactivators, respectively;  $M$  denotes monomers, and  $L$  denotes ligands. Photo induced reactions 14 – 17 in Table 1 only take place under light irradiation.

**Table 1.** Kinetic mechanism for photoATRP

	description	chemical equation	rate constant
1	activation	$P_0X + Mt_{n-1}/L \rightarrow P_0^* + Mt_nX/L$	$k_{a0} = 2.00 \times 10^3 \text{ mol}^{-1}\text{s}^{-1}$
2	deactivation	$P_0^* + Mt_nX/L \rightarrow P_0X + Mt_{n-1}/L$	$k_{d0} = 5.00 \times 10^7 \text{ mol}^{-1}\text{s}^{-1}$
3	activation	$P_sX + Mt_{n-1}/L \rightarrow P_s^* + Mt_nX/L$	$k_a = 2.00 \times 10^2 \text{ mol}^{-1}\text{s}^{-1}$
4	deactivation	$P_s^* + Mt_nX/L \rightarrow P_sX + Mt_{n-1}/L$	$k_d = 2.80 \times 10^8 \text{ mol}^{-1}\text{s}^{-1}$
5	propagation	$P_0^* + M \rightarrow P_1^*$	$k_{p0} = 7.30 \times 10^2 \text{ mol}^{-1}\text{s}^{-1}$
6		$P_s^* + M \rightarrow P_{s+1}^*$	$k_p = 1.56 \times 10^4 \text{ mol}^{-1}\text{s}^{-1}$
7	termination by combination	$2P_0^* \rightarrow D_0$	$k_{t0} = 2.00 \times 10^9 \text{ mol}^{-1}\text{s}^{-1}$
8		$P_0^* + P_s^* \rightarrow D_s$	
9		$P_s^* + P_r^* \rightarrow D_{s+r}$	
10	catalytic radical termination	$P_0^* + Mt_{n-1}/L \rightarrow D_0 + Mt_{n-1}/L$	$k_{tx0} = 0.00 \text{ mol}^{-1}\text{s}^{-1}$
11		$P_s^* + Mt_{n-1}/L \rightarrow D_s + Mt_{n-1}/L$	$k_{tx} = 4.00 \times 10^3 \text{ mol}^{-1}\text{s}^{-1}$
12	transfer	$P_s^* + L \rightarrow P_0^* + D_s$	$k_{trL} = 2.80 \times 10^3 \text{ mol}^{-1}\text{s}^{-1}$
13		$P_s^* + P_0X \rightarrow P_0^* + D_s$	$k_{trX} = 2.30 \times 10^2 \text{ mol}^{-1}\text{s}^{-1}$
14	photo induced radical regeneration	$L + M \xrightarrow{h\nu} 2P_0^*$	$k_{rLM} = 1.50 \times 10^{-9} \text{ mol}^{-1}\text{s}^{-1}$
15		$P_0X \xrightarrow{h\nu} 2P_0^*$	$k_{rX2} = 2.90 \times 10^{-9} \text{ s}^{-1}$
16		$P_0X + L \xrightarrow{h\nu} 2P_0^*$	$k_{rL2} = 6.20 \times 10^{-6} \text{ mol}^{-1}\text{s}^{-1}$
17		$Mt_nX/L + L \xrightarrow{h\nu} Mt_{n-1}/L + P_0^*$	$k_{rMt} = 1.00 \times 10^{-3} \text{ mol}^{-1}\text{s}^{-1}$

### 3. Dynamic Monte Carlo simulation

In the dynamic MC simulation, the propensity of the  $\mu^{\text{th}}$  reaction of the kinetic mechanism with  $M$  reactions is defined as:

$$R_\mu = k_\mu c_\mu, \forall \mu = 1, 2, \dots, M \quad (1)$$

where  $k_\mu$  is the kinetic rate constant, and  $c_\mu$  is the total concentration of the distinct molecular pairs that will react in the  $\mu^{\text{th}}$  reaction. The probability  $\mathbb{P}_\mu$  for reaction  $\mu$  to take place is given by:

$$R_{SUM} = \sum_{v=1}^M R_v \quad (2)$$

$$\mathbb{P}_\mu = \frac{R_\mu}{R_{SUM}} \quad (3)$$

A random number  $rnd_1$  with uniform distribution in the interval (0,1), i.e.,  $rnd_1 \in \mathcal{U}(0,1)$ , is generated to select one reaction  $\mu$  to occur in the current time interval based on the rule:

$$\sum_{\gamma=1}^{\mu-1} R_\gamma < rnd_1 \cdot R_{SUM} \leq \sum_{\gamma=1}^{\mu} R_\gamma \quad (4)$$

The dynamic time evolution of the kinetic system is described by determining the time interval between two consecutive reactions. To calculate the time interval  $\tau$ , another random number  $rnd_2 \in \mathcal{U}(0,1)$  is generated.



$$\tau = -\frac{1}{R_{SUM}} \ln(rnd_2) \quad (5)$$

The system states are updated after each reaction event, and accordingly, the simulation time is evaluated by the newly calculated  $\tau$ . This process repeats until a preset reporting time or conversion is reached. A detailed discussion can be found in ( Gillespie, 1977).

PhotoATRP processes are studied in this project through building dynamic MC models. In the implementation of the dynamic MC simulation for photoATRP systems, chain length information of radical polymers, dormant polymers, and dead polymers is stored in three individual arrays. The element value in the arrays denotes the chain length of one polymer. The arrays are updated following the selected reaction event. Specifically, if a propagation reaction is selected, an element in the array storing radical polymers is chosen at random, and the value of this element is increased by one. If an activation reaction is selected, one element in the dormant polymer array is chosen randomly and moved into the radical polymer array. If a termination by combination reaction is selected, two elements in the radical polymer array are chosen randomly and removed from the array, and the sum of the two element values is added to the dead polymer array. The instantaneous system states are recorded through tracking each radical, dormant and dead polymer chain.

#### 4. Results and discussions

Simulations for photoATRP systems are conducted through the dynamic MC method and the deterministic MoM. The dynamic MC simulation is capable to provide the microscopic structures of the produced polymers. The deterministic MoM cannot be directly applied to calculate microscopic information for photoATRP systems as the moment equations for all the polymers with diverse chain lengths are difficult to solve due to stiffness. Furthermore, investigations of temporal control over polymerization rate are studied using intermittent light-on and light-off periods. All simulations were performed using Jupyter Notebook on the Anaconda distribution of Python language version 3.8 on a personal computer with RAM 32 GB and Intel i7-10700 CPU core @2.90GHz running Windows 10 OS.

##### 4.1. Simulations of photoATRP for microscopic properties

A batch process where photoATRP takes place is simulated through developing a dynamic MC model and a deterministic method of moments (MoM) model. Tris 2-(dimethylamino)ethyl amine ( $\text{Me}_6\text{TREN}$ ), ethyl 2-bromoisobutyrate (EBiB), Methyl acrylate (MA),  $\text{CuBr}/L$  and  $\text{CuBr}_2/L$  are used as ligands  $L$ , alkyl halides  $P_0X$ , monomers  $M$ , activators  $Mt_{n-1}/L$ , and deactivators  $Mt_nX/L$ , respectively. The initial concentrations of monomers,  $P_0^*$ , deactivators and ligands are  $7.40 \text{ mol} \cdot L^{-1}$ ,  $2.47 \times 10^{-2} \text{ mol} \cdot L^{-1}$ ,  $7.40 \times 10^{-4} \text{ mol} \cdot L^{-1}$ , and  $3.70 \times 10^{-3} \text{ mol} \cdot L^{-1}$ , respectively. For brevity, the reaction conditions and kinetic rate constants can be found in (Ribelli et al., 2014).

As shown in Figure 2(a), the evolution of the average chain length calculated from the dynamic MC simulation is in good agreement with the results from the deterministic MoM and the experiment data. The evolution of concentrations for various reactants versus monomer conversion rate is shown in Figure 2(b). It can be observed that the concentrations of most of the reactants (monomer, activator, dormant polymer, dead polymer, deactivator, and ligand) from the developed dynamic MC model and the deterministic MoM are consistent with the literature (Ribelli et al., 2014). The

concentrations of dead polymers (solid and dashed lines in purple) obtained from this work and the literature diverge during the period of 0% to 20% conversion rate. The concentration lines of dead polymers from this work are smoother. This divergence most likely results due to inaccuracy in the adaptive time-step algorithm of the old version *Predici*, the software that was adopted to do the simulation in the literature. The MWD result of all types of polymers at 80% monomer conversion rate is shown in Figure 2(c). Individual MWD results of radical polymers, dormant polymers, and dead polymers can be calculated by the developed dynamic MC model. It is challenging to obtain such microscopic distribution using the deterministic MoM, which is achieved through deriving and solving a large system of equations in the order of thousands.

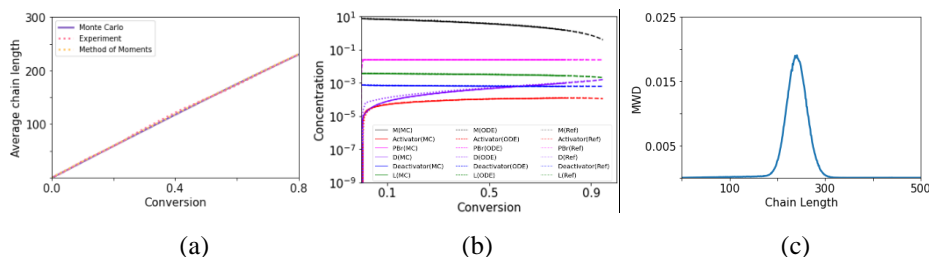


Figure 2 Simulation results for a photoATRP process. (a) Average chain length versus monomer conversion rate; the purple, yellow and red lines denote the results obtained from the MC simulation, MoM model and experiment, respectively. (b) Concentrations of reactants versus monomer conversion rate from different methods. (c) The molecular weight distribution (MWD) at 80% conversion obtained from the dynamic MC simulation, and the corresponding computational time is 9.79 hours. The MWD cannot be obtained from the deterministic MoM model.

#### 4.2. Simulations on the temporal control of photoATRP

Simulations on the temporal control of the photoATRP processes are further conducted with intermittent on/off periods with  $k_{ATRP} = 7.143 \times 10^{-5}$  ( $k_{ATRP} = \frac{k_a}{k_d}$ ). In the implementation of the simulations, the kinetic rate values of photo induced reactions (reactions 14 – 17 in Table 1) are set to zero to simulate the system in dark.

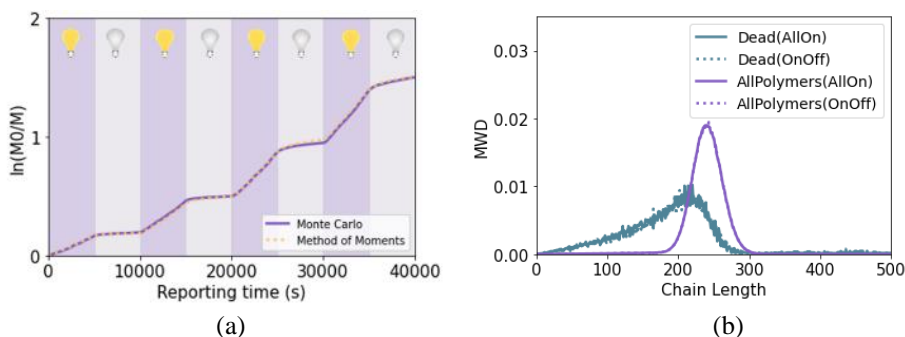


Figure 3 Simulation results of temporal control. (a) temporal control of polymer growth obtained from the dynamic MC simulation and the deterministic method of moments model. (b) MWDs of different polymer types obtained from dynamic MC simulations under conditions of light all-time on and light on/off.

The evolution of the concentration of monomers using the MC and the MoM models is presented in Figure 3(a). We observe that switching the light off decreases the

regeneration rate of radical polymers, leading to slower chain growth. The MWD results of dead polymers and all polymers are shown in Figure 3(b). The solid lines are the simulated results with all-time light irradiation, and the dotted lines are the results using intermittent 5000 second light-on and 5000 second light-off periods. As shown in Figure 3(b), reaction time can be affected by tuning the light on and off, and the MWDs are consistent under various light irradiation modes. It should be noted that the MWDs of each type of polymers at any reporting time and conversion rates can be addressed using the developed MC model. The kinetic characteristics of polymerizations such as the polymerization rate, dynamic average chain length, dynamic concentrations of reactants, under diverse irradiation conditions can be investigated through using the MC simulations.

## 5. Conclusions

In this work, dynamic Monte Carlo models and deterministic Method of Moments models are developed to simulate the photoATRP processes. The developed dynamic MC models are capable to predict the microscopic structures of polymers and provide the kinetic characteristics of the simulated mechanisms on demand. Temporal control over polymer growth was further investigated based on the developed models. As advanced photoATRP systems develop, MC simulation has great application potential in providing a clear understanding of the details of various mechanisms, optimizing the qualities of products, and designing polymerization processes at a microscopic level.

## Acknowledgments

The financial support of the "Pioneer" and "Leading Goose" R&D Program of Zhejiang (No. 2022C01176) is gratefully acknowledged by the first, second, and fourth authors. The financial support of the University of Patras, Medicus program (No. 81816) is gratefully acknowledged by the third author. The financial support from the China Scholarship Council is gratefully acknowledged by the first author. The authors would like to thank Prof. Krzysztof Matyjaszewski at Carnegie Mellon University and Prof. Dominik Konkolewicz at Miami University for offering experimental data, the simulated macroscopic information in their published paper, and analyzing the slight divergence in concentrations of dead polymers in Figure 2(b).

## References

- X. Pan, et al., 2016. Photomediated controlled radical polymerization. *Prog. Polym. Sci.* 62, 73-125.
- J. Wang, et al., 1995. Controlled/"living" radical polymerization. atom transfer radical polymerization in the presence of transition-metal complexes. *J. Am. Chem. Soc.* 117(20), 5614–5615.
- T. Ribelli, et al., 2014. How are radicals (re)generated in photochemical ATRP?. *J. Am. Chem. Soc.* 136(38), 13303–13312.
- F. Lorandi, et al., 2022. Atom transfer radical polymerization: A mechanistic perspective. *J. Am. Chem. Soc.* 144(34), 15413–15430.
- S. Dadashi-Silab, et al., 2020. Investigating temporal control in photoinduced atom transfer radical polymerization. *Macromolecules.* 53(13), 5280-5288.
- R. Liu, et al., 2021. Adaptable parallel acceleration strategy for dynamic monte carlo simulations of polymerization with microscopic resolution. *Ind. Eng. Chem. Res.* 60(17), 6173–6187.
- M. Chen, et al., 2016. Light-controlled radical polymerization: mechanisms, methods, and applications. *Chem. Rev.* 116(17), 10167-10211.
- D. Gillespie., 1977. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* 81(25), 2340–2361.

# Identification of a dynamic model of the N-vinylcaprolactam-co-glycidyl methacrylate microgel synthesis

Luise F. Kaven<sup>a</sup>, Thomas Nevolianis<sup>b</sup>, Nadja Wolter<sup>c,d</sup>, Adel Mhamdi<sup>a</sup>, Andrij Pich<sup>c,d</sup>, Kai Leonhard, Alexander Mitsos<sup>a,e</sup>

<sup>a</sup>Chair of Process Systems Engineering, RWTH Aachen University, Aachen, Germany

<sup>b</sup>Institut of Technical Thermodynamics, RWTH Aachen University, Aachen, Germany

<sup>c</sup>DWI – Leibniz Institute for Interactive Materials e.V., Aachen, Germany

<sup>d</sup>Functional and Interactive Polymers, Institute of Technical and Macromolecular Chemistry, RWTH Aachen University, Aachen Germany

<sup>e</sup>JARA-SOFT, Aachen, Germany

[amitsos@alumn.mit.edu](mailto:amitsos@alumn.mit.edu)

## Abstract

Incorporating methacrylate into microgels allows for their application in catalysis, drug release systems, and enzyme immobilization. Detailed insight into the occurring reactions during the microgel synthesis enables the precise control of the microgel properties. In this contribution, we identify a kinetic model of the copolymerization of N-vinylcaprolactam-based microgels with functional glycidyl methacrylate monomers. We include propagation reaction rate constants calculated via quantum chemical methods. First, we estimate the five most sensitive propagation and initiation and termination rate constants using the enthalpy transfer rate from calorimetry experiments. For the parameter estimation, we use the quantum chemically calculated values as start values and the determined error scopes as bounds in the optimization. We compare our approach of incorporating computed parameter values to similar procedures. The results show that our estimation approach enables a sufficient accuracy of parameter values. Finally, we predict the distribution of functional groups within the microgel with the fully identified model.

**Keywords:** parameter estimation, quantum chemistry, synthesis modeling

## 1. Introduction

Microgels are crosslinked polymer networks with a reversible stimulus-specific swelling response to temperature, pH, or ionic strength. Microgels based on N-vinylcaprolactam (VCL) are temperature responsive and interesting due to their biocompatibility (Saunders 2009). Incorporating comonomers into pVCL microgels allows tuning their functionalities. Glycidyl methacrylate (GMA) is a promising comonomer thanks to its reactive epoxy group that is versatile in subsequent reactions. Hence, pVCL microgels functionalized with GMA experience many applications (Schäfer 2020, Li 2013, Zou 2019).

GMA is more reactive than VCL in the copolymerization leading to the heterogeneous composition of the resulting microgels. Thus, the authors of (Qiu 2005) performed a gradual feeding technique to produce more homogeneous microgels. Under batch

conditions, (Häntzschel 2007) also showed via high-magnification SEM and TEM images that the microgels exhibit a core-shell structure with a GMA-rich core surrounded by a VCL-rich shell. Due to the complexity of the reaction network during microgel synthesis, understanding the relationship between the synthesis conditions and the final composition of the microgels is challenging. Thus, for designing tailored microgels, detailed reaction models provide insights and assist with developing microgels with defined monomer distributions.

In previous contributions, we have developed a mechanistic kinetic model for the aqueous microgel synthesis by precipitation polymerization (Janssen 2017, Jung 2019). To estimate the model parameters, we used a combination of quantum chemical calculations (Kröger 2017) and physical measurements (Jung 2019). We have applied the dynamic model to predict and guide the synthesis of microgels based on monomer N-isopropylacrylamide, comonomer vinyl ferrocene (VFc), and crosslinker N,N'-methylene bisacrylamide (BIS). In contrast to these monomer combinations, parameter values for the propagation rate constants and polymerization enthalpies of GMA/VCL and GMA/BIS systems in an aqueous solution remain unknown.

In our preceding work, we built a mechanistic model based on (Schneider 2020) for synthesizing VCL-based microgels functionalized with GMA. The model contained 14 unknown kinetic parameter values. For model identification, we followed a hybrid approach. First, using transition state theory and quantum chemical calculations, we computed 12 out of 14 parameters, namely the propagation rate constants and polymerization enthalpies of the GMA/VCL and GMA/BIS systems in an aqueous solution. We conducted experiments with four different initial compositions and measured the enthalpy transfer rate during the synthesis using a calorimetric device. Subsequently, we estimated the remaining parameter values of the five most sensitive propagation rates and the initiation and termination rates using the measured enthalpy transfer rate from the experiments. We used the quantum chemically calculated values as start values for the parameter estimation. We also added the computed error bounds from the quantum calculation as optimization constraints during the parameter estimation.

In this contribution, we further analyze the parameter estimation incorporating quantum chemically calculated parameter values. We estimate different combinations of unknown parameter values and conduct an identifiability analysis. Finally, we use the fully identified VCL/GMA microgel synthesis model to predict feeding strategies for a homogeneous microgel composition.

## 2. Materials and Methods

### 2.1. Microgel Synthesis and Reaction Calorimetry

We analyze the precipitation polymerization of VCL crosslinked with BIS and functionalized with GMA. As an initiator, we consider 2,2'-azobis(2-methylpropionamidine)dihydrochloride. Previously, we measured the enthalpy transfer

Table 1 Synthesis composition for experiments in the calorimeter.

Experiment	Molar ratio VCL:GMA	Molar amount of crosslinker BIS
1	90:10	2.0 mol%
2	90:10	2.6 mol%
3	93:7	2.0 mol%
4	95:5	2.6 mol%

rate during the reaction via reaction calorimetry (Nevolianis 2023). Details on the setup and compositions of the conducted experiments can be found in our previous work (Nevolianis 2023). In Table 1, we briefly outline the variations between the experiments used in the current work.

### *2.2. Quantum Chemical Calculations*

In our previous work (Nevolianis 2023), we incorporated propagation rate constants of the VCL/GMA/BIS system that were calculated via quantum chemistry methods. There, we applied conventional transition state theory to determine propagation rate constants. Further, we calculate the enthalpies of reaction using the Gibbs-Helmholtz equation. Computational details can be found in (Nevolianis 2023).

### *2.3. Kinetic Model of Microgel Synthesis*

We employ the reaction system of our kinetic synthesis model (Schneider 2020) and change the parameter values according to the VCL/GMA/BIS system. The main model assumptions are briefly summarized in the following. The dispersed system is modeled in a pseudo-homogeneous phase. The occurring reactions include initiation, propagation, termination, and crosslinking. The efficiency of the initiator decomposition is set to 50% (Jung 2020). The kinetics of the propagation reactions depend only on the radical at the terminal end of the chain, and propagation reactions make up for the enthalpy release. Termination only occurs due to disproportionation, and the kinetics of termination reactions do not depend on the terminal end of the active radical chains. The kinetics of the crosslinking reaction equals the kinetics of a propagation reaction with a BIS monomer.

### *2.4. Parameter Estimation and Identifiability Analysis*

We use the five reaction enthalpy values involving GMA from the quantum chemical computations without further treatment in the synthesis model. We then analyze the sensitivity to identify the most sensitive propagation rate ( $k_p$ ) parameters. Subsequently, we perform parameter estimations to determine different combinations of parameter values. The initiation and termination rates are estimated for all sets as they remain unknown after the quantum chemical calculations. We split the conducted estimations into the following groups:

- 1) Estimate all  $k_{ps}$  with arbitrary start values and bounds
- 2) Estimate all  $k_{ps}$  using the quantum calculations as start values and their error bounds as constraints
- 3) Estimate the five most sensitive  $k_{ps}$  using the quantum calculations as start values and their error bounds as constraints
- 4) Insert all  $k_{ps}$  into the model and only estimate  $k_t$  and  $k_d$

We perform the parameter estimation and the identifiability in Matlab, similar to (Jung 2019), using the presented copolymerization model and the calorimetric measurements. For part of the estimations, the calculated propagation rate constants act as start values, and the calculation uncertainties derived from Gaussian error propagation analysis of the quantum chemical method's uncertainty act as constraints in the optimization problem. For estimation (1), the start values for the propagation rates are 10, and the upper and lower bounds are  $10^{-3}$  and  $10^5$ , respectively.

### 3. Results and Discussion

#### 3.1. Quantum Chemically Calculated and Estimated Kinetic Parameters

The calculated propagation rate constants for the VCL/GMA/BIS system in an ideally diluted aqueous solution at a temperature of 343 K are shown in the second column of Table 3. The applied bounds are presented in the third/fourth column of Table 3. Using the optimized parameter values from estimation (3), our model agrees to 97% on average to the measured enthalpy transfer rate from the experiments.

The results in Table 3 show that the homopolymerizations of GMA and VCL are approximately in the same order of magnitude when calculated via quantum chemistry. The most significant reaction occurs between VCL-radical and GMA. The reaction between GMA-radical and VCL plays a negligible role in the polymerization system. All reactions consuming GMA occur relatively fast. Thus, the experimental observation of a GMA-rich microgel core fits our calculated findings.

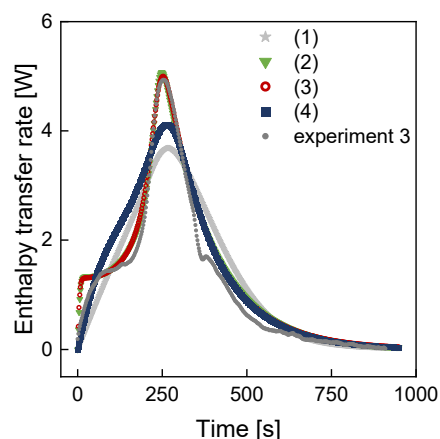


Figure 1 Enthalpy transfer rate for experiment 3 and the simulations with the estimated parameter values from estimation (1) to (4).

Table 1 Propagation rate constants [ $\text{m}^3 \cdot \text{mol}^{-1} \cdot \text{s}^{-1}$ ] at 343 K calculated via quantum chemistry and optimized in the parameter estimations with corresponding upper and lower error bound from the quantum chemical calculations. \* denotes estimated parameter values at the bounds.

Reactants	Calculated $k_p$	Lower error bound	Upper error bound	Optimized $k_p$ (1)	Optimized $k_p$ (2)	Optimized $k_p$ (3)	Optimized $k_p$ (4)
VCL-radical / VCL	86.8	10.4	727	28.7	23.8	116	
VCL-radical / GMA	516	61.6	4330	0.403	337	1300	
GMA-radical / GMA	65.5	7.8	549	$10^{-3}$ *	18.2	64.9	
GMA-radical / VCL	4.6	0.5	38.5	1.59	0.74	4.6	
BIS-radical / VCL	34.3	4.1	287	733	9.76	287	
BIS-radical / BIS	8.1	0.967	67.9	$10^{-3}$ *	67.9*		
GMA-radical / BIS	0.74	0.088	6.18	$10^{-3}$ *	2.44		
BIS-radical / GMA	129	15.4	1080	4210	15.4*		
$k_d$ [ $\text{s}^{-1}$ ]		$10^{-7}$	$10^{-3}$	$6.15 \cdot 10^{-7}$	$3.48 \cdot 10^{-5}$	$8.22 \cdot 10^{-6}$	$8.97 \cdot 10^{-7}$
$k_t$		$10^{-5}$	$10^5$	$10^{-5}$ *	244	1480	81.5

The enthalpy transfer rate simulated via the synthesis model containing the respective parameter values is shown in Figure 1, exemplary for Experiment 3. The parameter combinations from estimations (2) and (3) result in the best fit to the experimental data.

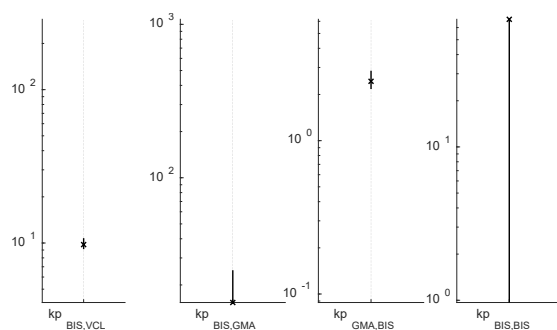


Figure 2 Identifiability analysis of selected parameters from estimation (2). The x indicates the optimal parameter value.

The estimation (1) allows a wide solution space with arbitrary start values and bounds gets stuck in a local minimum, indicating the need for deterministic global optimization. The simulation of estimation (4) only has two degrees of freedom and does not sufficiently follow the experimental trend. Altogether, estimation (3) is most suitable, as the fit to experimental data is eminent, and the computational effort is decreased due to reduced parameter values for estimation.

Figure 2 shows the identifiability result for the estimation set (2). For visualization, only the parameters with the most significant variability are shown. Here, all parameters

appear to be sufficiently identifiable except  $k_{p,BIS-BIS}$ .

### 3.2. Process Simulations for Feeding Strategy

The identified model with parameters from estimation (3) allows for the assessment of feeding strategies. We simulate different feeding strategies of GMA to the batch reaction. We focus on two aspects, namely, the ratio of GMA initially in the reactor to the dosing feed and the dosing flow rate. The simulation includes the composition of Experiment 1 (see Table 1) and assumes that 7.5 ml of the total water volume is used for dosing. The ratio of GMA fed to GMA in the reactor varies from 0:1 (all GMA is initially in the reactor) to 1:1 and 1:0 (all GMA is fed to the reactor). Simultaneously, the dosing flow rate is set to 1.5, 3, and 6 mmol/s.

In Figure 3, the results of the simulation scenarios are shown. Here, the combined reaction rate of GMA is depicted over time. The combined reaction rate comprises all reactions where GMA is consumed. When assuming monodisperse particle growth, the time of consumption of GMA correlates to the incorporated position in the microgel. Thus, the combined reaction rate of GMA represents its composition in the microgel.

The progressions of Figure 3 indicate that adding all GMA at a dosing flow of 1.5 mmol/s is likely to achieve a homogeneous composition of GMA in the microgel.

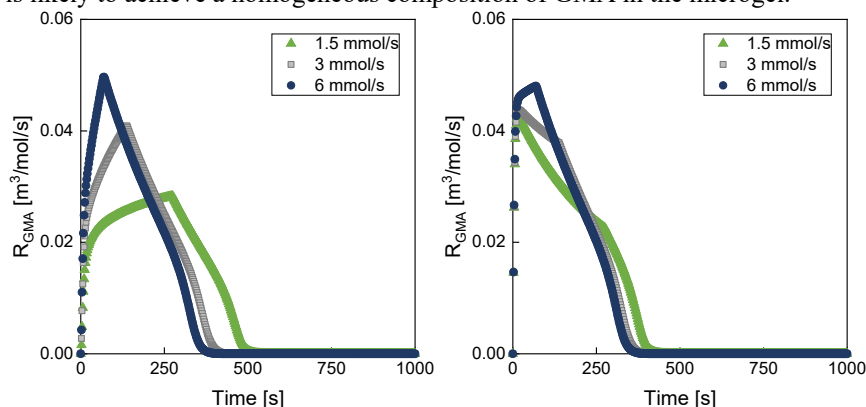


Figure 3 Left: Feeding strategy 1:0, all GMA is fed to the reactor, and the dosing flow rate varies. Right: Feeding strategy 1:1, half of the GMA is in the reactor at the start of the reaction, and half is fed at different dosing flow rates.



#### 4. Conclusions

We studied the precipitation polymerization of VCL/GMA/BIS in an aqueous solution focusing on parameter estimation approaches. We showed that incorporating quantum chemically computed parameter values and their calculation error into the estimation problem improves the optimization result. Based on the identified synthesis model, we were able to derive an improved feeding strategy of GMA to the batch reaction to achieve a homogeneous GMA composition in the microgels by testing different scenarios. Our hybrid approach of quantum chemistry and parameter estimation from experimental data shows the potential of integrating theoretical derivations and semi-empirical knowledge. Further, we laid the ground to design and optimize the functionalization of VCL-based microgels with GMA in the future.

#### References

- N. Häntzschel, F. Zhang, F. Eckert, A. Pich, M. A. Winnik, 2007, Poly(N-vinylcaprolactam-co-glycidyl methacrylate) aqueous microgels labeled with fluorescent LaF<sub>3</sub>:Eu nanoparticles, *Langmuir*, 23, 21, 10793-10800
- F. A. L. Janssen, M. Kather, L. C. Kröger, A. Mhamdi, K. Leonhard, A. Pich, A. Mitsos, 2017, Synthesis of poly(N-vinylcaprolactam)-based microgels by precipitation polymerization: process modeling and experimental validation, *Ind. Eng. Chem. Res.*, 56, 14545-14556
- F. Jung, F. A. L. Janssen, A. Ksiazkiewicz, A. Caspari, A. Mhamdi, A. Pich, A. Mitsos, 2019, Identifiability analysis and parameter estimation of microgel synthesis: a set-membership approach, *Ind. Eng. Chem. Res.* 58, 13675-13685
- F. Jung, A. Ksiazkiewicz, A. Mhamdi, A. Pich, A. Mitsos, 2020, *Ind. Eng. Chem. Res.*, 59, 20437-20446
- L. C. Kröger, W. A. Kopp, K. Leonhard, 2017, Prediction of chain propagation rate constants of polymerization reactions in aqueous NIPAM/BIS and VCL/BIS systems, *Physical Chemistry B*, 121, 2887-2895
- P. Li, R. Xu, W. Wang, X. Li, Z. Xu, K. W. K. Yung, P. K. Chu, 2013, Thermosensitive poly(N-isopropylacrylamide-co-glycidyl methacrylate) microgels for controlled drug release, *Colloids and Surfaces B: Biointerfaces*, 101, 251-255
- E. Mavroudakos, D. Cuccato, D. Moscatelli, 2015, On the use of quantum chemistry for the determination of propagation, copolymerization, and secondary reaction kinetics in free radical polymerization, *Polymers*, 7, 9, 1789-1819
- T. Nevolianis, N. Wolter, L. Kaven, L. Krep, C. Huang, A. Mhamdi, A. Mitsos, A. Pich, K. Leonhard, 2023, Kinetic Modeling of a Poly(N-vinylcaprolactam-co-glycidyl methacrylate) Microgel Synthesis - A Hybrid In-Silico and Experimental Approach, Submitted to *Ind. Eng. Chem. Res.*
- B. R. Saunders, N. Laajam, E. Daly, S. Teow, X. Hu and R. Stepto, Microgels: From responsive polymer colloids to biomaterials, *Adv. Colloid Interface Sci.*, 2009, 147-148, 251-262
- D. Schäfer, F. Fink, D. Kleinschmidt, K. Keisers, F. Thomas, A. Hoffmann, A. Pich, S. Herres-Pawlis, 2020, Enhanced catalytic activity of copper complexes in microgels for aerobic oxidation of benzyl alcohols, *Chemical Communications*, 56, 5601-5604
- S. Schneider, F. Jung, O. Mergel, J. Kammertz, A. C. Nickel, T. Caumanns, A. Mhamdi, J. Mayer, A. Mitsos, F. A. Plamper, 2020, Model-based design and synthesis of ferrocene containing microgels, *Polymer Chemistry*, 11, 315-325
- X. Qiu, S. A. Sukhishvili, 2005, Copolymerization of N-vinylcaprolactam and glycidyl methacrylate: Reactivity ratio and composition control, *Polymer Chemistry*, 44, 183-191
- Z. Zou, E. Gau, I. El-Awaad, F. Jakob, A. Pich, U. Schwaneberg, 2019, Selective Functionalization of Microgels with Enzymes by Sortagging, *Bioconjugate Chemistry*, 30, 11, 2859-2869

# A Machine Learning-assisted Hybrid Model to Predict Ribbon Solid Fraction, Granule Size Distribution and Throughput in a Dry Granulation Process

Yan-Shu Huang,<sup>a</sup> David Sixon,<sup>a</sup> Phoebe Bailey,<sup>a</sup> Rexonni B. Lagare,<sup>a</sup> Marcial Gonzalez,<sup>b,c</sup> Zoltan K. Nagy,<sup>a</sup> Gintaras V. Reklaitis<sup>a</sup>

<sup>a</sup>Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN 47907, USA

<sup>b</sup>School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA

<sup>c</sup>Ray W. Herrick Laboratories, Purdue University, West Lafayette, IN 47907, USA

[huan1289@purdue.edu](mailto:huan1289@purdue.edu)

## Abstract

A quantitative model can play an essential role in controlling critical quality attributes of products and in designing the associated processes. One of the challenges in designing a dry granulation process is to find the optimal balance between improving powder flowability and sacrificing powder tabletability, both of which are highly affected by ribbon solid fraction and granule size distribution (GSD). This study is focused on developing a hybrid machine learning (ML)-assisted mechanistic model to predict ribbon solid fraction, GSD, and throughput for the purpose of implementing model predictive control of an integrated continuous dry granulation tableting process. It is found that the predictability of ribbon solid fraction and throughput are improved when modification is made to Johanson's model by incorporating relationships between roll compaction parameters and ribbon elastic recovery. Such relationships typically are either not considered or assumed to be a constant in the models reported in the literature. To describe the nature of the bimodal size distribution of roller compactor granules instead of only using traditional  $D_{10}$ ,  $D_{50}$  and  $D_{90}$  values, the GSD is represented by a bimodal Weibull distribution with five fitting parameters. Furthermore, these five GSD parameters are predicted by ML models. The results indicate the ribbon solid fraction and screen size are the two most significant factors affecting GSD.

**Keywords:** dry granulation, roller compactor, machine learning, hybrid model

## 1. Introduction

The dry granulation process is an important route for producing a solid dosage form in the pharmaceutical industry. The roller compactor is the key unit operation in a dry granulation process. It includes two steps: (1) roll compaction in which powder blends are compressed between two counter-rotating rolls into a ribbon, and (2) milling in which these ribbons are crushed into granules. The benefits of dry granulation include improved blend uniformity and flowability by enlargement of particle size. Powder flowability plays a key role in determining the performance of the tablet manufacturing process and final drug product quality (Lagare et al., 2023). In addition, good powder flowability can reduce powder fouling and improve the on-line sensor performance, such as capacitance-

based particulate flow rate sensor (Huang et al., 2022). However, particle size over-enlargement or over-compression of powders can compromise the tabletability (Herting & Kleinebudde, 2008). Finding the optimal balance between improving powder flowability and sacrificing powder tabletability, which are highly affected by ribbon solid fraction and GSD, becomes one of the challenges in designing a dry granulation process. Therefore, quantitative models to predict ribbon solid fraction and GSD are essential to optimally operate the roller compactor.

Mechanistic models such as Johanson's model (Johanson, 1965) are typically used to describe roll compaction and further predict ribbon solid fraction. However, one of the reasons for the unsatisfactory prediction accuracy of ribbon solid fraction is due to elastic recovery, which is either not considered or assumed to be a constant in the models reported in the literature (Keizer & Kleinebudde, 2020). Population balance models (PBM) can account for the milling step and can be used to predict GSD, but it is complicated to determine the breakage function in the PBM purely based on ribbon fracture physics. Given the unknown physical nature of ribbon elastic recovery and GSD, machine learning (ML) is a preferred alternative to developing a mechanistic model. Moreover, ML and mechanistic model components can be combined into a hybrid model to maintain high physical interpretability and feasibility. Therefore, the primary objective of this work is to develop and validate a hybrid model for the purpose of implementing model predictive control of an integrated continuous dry granulation tableting process.

## 2. Methodology

### 2.1. Parameter Estimation

Given  $n$  experiments and  $m$  output variables, the procedure to estimate model parameters can be formulated as a constrained optimization problem:

$$\min_{\hat{\theta}} J = \sum_{i=1}^n \sum_{j=1}^m w_j (y_{j,i} - \hat{y}_{j,i})^2 \quad (1)$$

subject to  $f(u, \hat{\theta}, \hat{y}) = 0$ ,  $\hat{\theta} \in \Omega_{\theta}$ ,  $\hat{y} \in \Omega_y$

where  $u$  and  $y$  are measurements of input and output variables.  $\hat{\theta}$  and  $\hat{y}$  are estimated model parameters and predicted output variables, bounded in compact sets  $\Omega_{\theta}$  and  $\Omega_y$ , respectively.  $w_j$  is the weighting for residual of output variable  $y_j$ .

### 2.2. Machine Learning

Machine learning models are common alternatives when process outputs are difficult to predict by pure mechanistic models. The multiple linear regression (MLR) model and the neural network (NN) model are employed in this study. The NN models studied consist of one hidden layer, where the hyperbolic tangent function is applied.

### 2.3. Experimental Methods

The material used in this study was a blend of 90 % w/w microcrystalline cellulose Avicel PH102 and 10% w/w acetaminophen. In each experiment, the Alexanderwerks WP120 roller compactor was operated for three minutes to reach a steady state, at which point samples were collected. The ribbons of interest were collected after the powders were compressed into ribbons and then broken into smaller ribbons by the flake crusher. The granulator consists of a two-stage hammer mill with two different screen sizes (screen 1 and screen 2), which produces two granule samples (granule 1 and granule 2). An in-house flowrate sensor based on a Mettler Toledo ME 4001E weighing scale was located at the roller compactor exit to capture the throughput of granules.

In order to measure the ribbon thickness, a Fisherbrand Traceable digital caliper was used to measure thirty ribbons to provide a statistically significant sample size. Geopyc 1360

pycnometer was used to measure the ribbon envelop density ( $\rho_e$ ). The ribbon solid fraction ( $\gamma_R$ ) can be computed as:

$$\gamma_R = \frac{\rho_e}{\rho_t} \quad (2)$$

where  $\rho_t$  is the powder true density, which is measured by an Accupyc II 1340 pycnometer. The GSD of granule 1 and granule 2 samples were measured by the Cauty SolidSizer, which measures the size and area of each particle. In this study, the circular equivalent diameter was used and the cumulative frequency of GSD was volume-based.

### 3. Results and Discussion

The input and output variables of the roller compactor are shown in Figure 1. Ribbon splitting is a phenomenon causing additional uncertainty in the model accuracy and can be avoided when the roll gap and roll pressure are low.

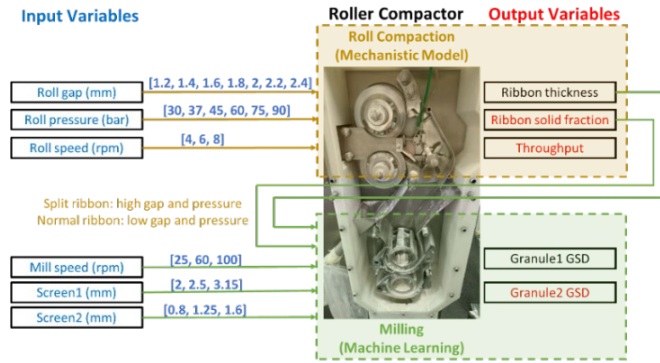


Figure 1. Roller compactor schematic.

#### 3.1 Roll Compaction

When the materials transform from the slip condition to the non-slip condition, the stress gradients in slip region and nip region are equal. The critical angular roller position at which this occurs is known as the nip angle  $\alpha$  and can be calculated by Johanson's model:

$$\frac{4\left(\frac{\pi}{2}-\alpha-\nu\right)\tan\delta_E}{\cot(A-\mu)-\cot(A+\mu)} - \frac{K\left(2\cos\alpha-1-\frac{S}{D_R}\right)\tan\alpha}{\cos\alpha} = 0 \quad (3)$$

$$\text{where } A = \frac{\alpha+\nu+\frac{\pi}{2}}{2}, \quad \nu = \frac{1}{2}\left[\pi - \sin^{-1}\left(\frac{\sin\phi_W}{\sin\delta_E}\right) - \phi_W\right], \quad \mu = \frac{\pi}{4} - \frac{\delta_E}{2} \quad (4)$$

Here,  $\delta_E$  is the effective angle of internal friction and  $\phi_W$  is wall friction angle,  $K$  is compressibility factor,  $S$  is roll gap, and  $D_R$  is roll diameter.

Given roll diameter  $D_R$  and roll width  $W$ , the peak pressure ( $P_{max}$ ) applied on the powders at the minimum roll gap  $S$  is computed as follows:

$$P_{max} = \frac{2P_H A_{cs}}{W D_R F} \quad (5)$$

with the force factor,  $F$ , given by

$$F = \int_0^\alpha \left[ \frac{\frac{S}{D_R}}{\left(1 - \frac{S}{D_R} - \cos\theta\right)\cos\theta} \right]^K \cos\theta \, d\theta \quad (6)$$

where  $P_H$  is hydraulic pressure (or roll pressure) and  $A_{cs}$  is area of the compact surface. Based on a compression power law, the ribbon solid fraction at the gap  $\gamma_G$  can be computed as follows:

$$\gamma_G = \gamma_0 (P_{max})^{\frac{1}{K}} \quad (7)$$

where  $\gamma_0$  is the pre-consolidation solid fraction. However,  $\gamma_G$  is not easily measured because ribbon elastic relaxation makes ribbon density decrease when ribbons are

released from the roll. Given the elastic recovery  $\beta$ , the ribbon solid fraction  $\gamma_R$  is represented as follows:

$$\gamma_R = \frac{\gamma_G}{\beta} \quad (8)$$

Considering mass balance around the roll gap and roll speed  $N_R$ , the mass throughput can be calculated as follows:

$$\dot{M} = \pi D_R W S N_R \rho_t \gamma_G = \pi D_R W S N_R \rho_t \beta \gamma_R \quad (9)$$

In this work, 15 sets of training data and 4 sets of test data are used to validate and evaluate three roll compaction models. First, a data-driven MLR model is built as a benchmark. Secondly, roll compaction mechanistic models are highly sensitive to powder compressibility  $K$  and powder pre-consolidation solid fraction  $\gamma_0$  (Toson et al., 2019). Instead of using common regression approaches to estimate these two model parameters, this study estimates parameters in a constrained optimization framework as follows:

$$\min_{\phi_W, \delta_E, K, \gamma_0, \beta} J = \sum_{i=1}^n [W_Y (\gamma_{R,i} - \hat{\gamma}_{R,i})^2 + W_M (\dot{M}_i - \hat{\dot{M}}_i)^2] \quad (10)$$

Thirdly, elastic recovery is known to be a function of roll compaction parameters instead of a constant. Under the assumption that measured elastic recovery is the ratio of measured ribbon thickness to roll gap, the predicted elastic recovery  $\beta$  is formulated as:

$$\beta_{Model3} = 0.96 + 0.12 \frac{S}{S_0} + 0.03 \frac{PH}{PH_0} + 0.01 \frac{NR}{NR_0} \quad (11)$$

where  $\beta_{Model3}$  has training error MAPE = 4.17 % and test error MAPE = 10.67%. To further improve the elastic recovery predictability, it is worth investigating ribbon splitting phenomenon and other nonlinear ML models in future studies. The prediction performances of the three roll compaction models are summarized in Table 1, which is calculated using the test sets. By incorporating the relationships between roll compaction parameters and ribbon elastic recovery, Model3 has the smallest mean absolute percentage error (MAPE) for both ribbon solid fraction and mass throughput, indicating the best model performance. The parity plot for Model3 is shown in Figure 2.

Table 1. Roll compaction model performance.

Model	MAPE( $\gamma_R$ ) [%]	MAPE( $\dot{M}$ ) [%]
Model1 (MLR)	5.68	9.98
Model2 (Johanson's with constant $\beta$ )	4.40	5.01
Model3 (Johanson's with $\beta = \beta_{model3}$ )	2.86	4.61

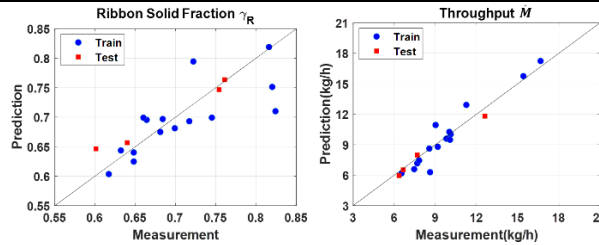


Figure 2. Performance of hybrid model of considering elastic recovery model (Model3).

### 3.2 Milling

The hammer milling step commonly produces granules with a bimodal size distribution, which is not adequately described by only using  $D_{10}$ ,  $D_{50}$  and  $D_{90}$  values. Therefore, the entire cumulative size distribution  $D_5, D_{10}, \dots, D_{95}$  measured with the Canty SolidSizer is represented by a bimodal GSD, which can be characterized by a cumulative bimodal Weibull distribution  $Q_3(x)$ :

$$Q_3(x) = a \left( 1 - e^{-\left(\frac{x}{p_1}\right)^{m_1}} \right) + (1 - a) \left( 1 - e^{-\left(\frac{x}{p_2}\right)^{m_2}} \right) \quad (12)$$

where  $a$  is the weighting of the two modes,  $p_1$  and  $p_2$  are the size parameters of the small mode and large mode, respectively, whereas  $m_1$  and  $m_2$  represent the shape parameters of the associated modes. Utilizing these statistical model parameters provides a more interpretable approach to describing GSD and reduces the GSD parameter set from 19 to 5. The bimodal Weibull distribution parameters can be computed by solving an optimization problem:

$$\min_{a, p_1, p_2, m_1, m_2} J = \sum_{i=5,10,\dots}^{95} \left( \frac{i}{100} - Q_3(D_i) \right)^2 \quad (13)$$

subject to  $0 \leq a \leq 1$ ,  $0 < p_1 < p_2$ ,  $1 < m_1, m_2$

To link the roll compaction and milling process, the ribbon solid fraction  $\gamma_R$  and ribbon thickness (*Rtck*) serve as inputs of the milling model. In addition, milling speed (*Mill*) and screen size(s) are also used to predict the GSD described by five bi-modal Weibull fitting parameters. The sensitivity analysis of both granule 1 and granule 2 are shown in Figure 3. The sensitivity analysis was determined by the MLR model coefficient, where inputs and outputs were both rescaled by dividing the minimum values. In terms of weighting  $a$ , ribbon solid fraction of granule 1 is dominated by ribbon solid fraction while that of granule 2 is more sensitive to screen size. For both granule 1 and granule 2, the size parameter  $p_1$  is highly related to ribbon solid fraction. On the other hand,  $p_2$  is dominated by screen size given that screen size determines the upper boundary of the particle size. Shape parameter  $m_1$  is less sensitive to all process inputs compared to the shape parameter  $m_2$ , which is highly impacted by the solid ribbon fraction.

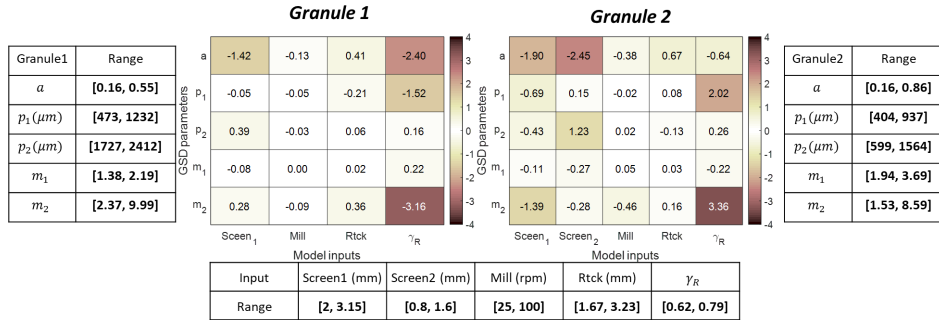


Figure 3. Sensitivity analysis of the milling process.

To evaluate the performance of the milling models, the MAPE of GSD is utilized and can be computed as:

$$MAPE = \sum_{i=1}^n \sum_{p=5,10,\dots}^{95} \left| \frac{D_{p,i}^{pred} - D_{p,i}}{D_{p,i}} \right| \times 100\% \quad (14)$$

The evaluation of the milling models is summarized in Table 2, which is based on 20 training sets and 6 test sets. MLR and NN models are used to predict GSD. The NN models have two neurons in the hidden layer. While increasing the number of neurons can readily make the training error of the NN model smaller than that of the MLR model, the test error can become much worse due to overfitting. The NN model generally can handle nonlinearity better than the MLR model, but some constraints might be hard to enforce, e.g., predicted shape parameters  $m_1$  and  $m_2$  might be less than 1 in the test sets. Figure 4 demonstrates the predictability of the granule 2 GSD by using the MLR or NN models based on six test sets. It is worth noting that there exists a significant mismatch between measurement and NN predictions in Exp 6, but the NN model prediction seems more reasonable considering that screen 2 is 1.25 mm and Exp 6 has a smaller ribbon solid fraction compared to Exp 3, which should result in a smaller GSD. In summary, the

hybrid model successfully predicts GSD, and further investigation of different ML models might be useful to enhance the model performance.

Table 2 Milling model performance.

Model	Granule1		Granule2	
	$MAPE_{train}$ (%)	$MAPE_{test}$ (%)	$MAPE_{train}$ (%)	$MAPE_{test}$ (%)
MLR	9.79	11.21	10.89	8.95
NN	9.66	12.20	13.03	8.37

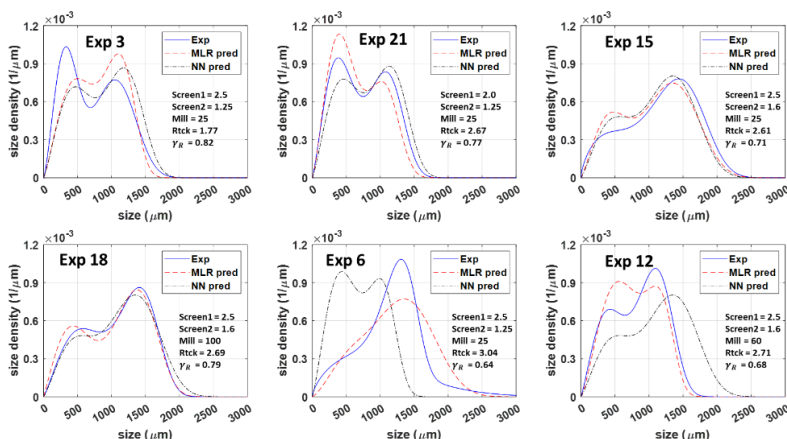


Figure 4. Model performance of granule2 based on test sets.

#### 4. Conclusion

A hybrid model for the roller compactor is proposed that demonstrates satisfactory predictability of ribbon solid fraction, throughput, and GSD. To further improve the model performance, investigation on ribbon elastic recovery and splitting phenomenon and incorporation constraints into ML model could be important. Future work will include relating ribbon solid fraction and GSD to the tablet properties and implementing model predictive control of the integrated dry granulation tableting process.

#### Acknowledgement

This work was supported by the National Science Foundation under Grant No. 2140452 - CMMI-EPSC: Right First Time Manufacture of Pharmaceuticals (RiFTMaP)

#### References

- Herting, M. G., & Kleinebudde, P. (2008). Studies on the reduction of tensile strength of tablets after roll compaction/dry granulation. *European Journal of Pharmaceutics and Biopharmaceutics*, 70(1), 372–379.
- Huang, Y. -S., Medina-González, S., Straiton, B., Keller, J., Marashdeh, Q., Gonzalez, M., Nagy, Z., & Reklaitis, G. v. (2022). Real-Time Monitoring of Powder Mass Flowrates for Plant-Wide Control of a Continuous Direct Compaction Tablet Manufacturing Process. *Journal of Pharmaceutical Sciences*, 111(1), 69–81.
- Johanson, J. R. (1965). *A Rolling Theory for Granular Solids Theory and Method of Calculation*.
- Keizer, H. L., & Kleinebudde, P. (2020). Elastic recovery in roll compaction simulation. *International Journal of Pharmaceutics*, 573.
- Lagare, R. B., Huang, Y.-S., Bush, C. O.-J., Young, K. L., Rosario, A. C. A., Gonzalez, M., Mort, P., Nagy, Z. K., & Reklaitis, G. v. (2023). Developing a Virtual Flowability Sensor for Monitoring a Pharmaceutical Dry Granulation Line. *Journal of Pharmaceutical Sciences*.
- Toson, P., Lopes, D. G., Paus, R., Kumar, A., Geens, J., Stibale, S., Quodbach, J., Kleinebudde, P., Hsiao, W. K., & Khinast, J. (2019). Model-based approach to the design of pharmaceutical roller-compaction processes. *International Journal of Pharmaceutics: X*, 1, 100005.

# A Data-driven Modeling Approach for Water Flow Dynamics in Soil

Zeyuan Song<sup>a</sup>, Zheyu Jiang<sup>a\*</sup>

<sup>a</sup>*Oklahoma State University, 420 Engineering North, Stillwater, Oklahoma 74074 USA*  
Corresponding author: [zheyu.jiang@okstate.edu](mailto:zheyu.jiang@okstate.edu)

## Abstract

Modeling and predicting soil moisture is essential for precision agriculture, smart irrigation, drought prevention, etc. Estimating root zone soil moisture from surface or near-surface soil moisture data is typically achieved by solving a hydrological model that describes water movement through soils. Advanced agro-hydrological models today use the Richards equation, a highly nonlinear, degenerate elliptic-parabolic partial differential equation that captures irrigation, precipitation, evapotranspiration, runoff, and drainage. State-of-the-art Richards equation solvers employ either a finite difference, finite element, or finite volume discretization framework in space. In this paper, we introduce a novel computational framework to solve generic  $n$ -dimensional Richards equation by introducing global random walk and deep neural network to a modified finite volume method (FVM). Furthermore, for  $n$ -dimensional Richards equation, we introduce multi-point flux approximation to the FVM framework. Through these innovations, our novel computational framework effectively utilizes the underlying physics behind the Richards equation, which enhances the speed and accuracy of the solution process. Through an illustrative case study, we demonstrate the efficiency and effectiveness of our computational framework and show that it correctly characterizes the physical relationships among soil moisture content, pressure head, and flux.

**Keywords:** Soil moisture, Richards equation, finite volume method, neural network, random walk

## 1. Introduction

Soil moisture is a key hydrological parameter that has significant importance to human society and environment. Accurate modeling and monitoring of soil moisture in crop fields, especially in the root zone (top 100 cm of soil), is essential for improving agricultural production and crop yield with the help of precision irrigation and farming tools. Recent studies also show that monitoring root zone soil moisture at suitable locations and adjusting irrigation schedules accordingly can reduce water use by 40-60% (Sadler et al., 2005). Improved irrigation infrastructures based on soil moisture knowledge could also prevent more than \$30 billion/year in drought related agricultural losses in the US (Khand et al., 2018). This is especially important to US states like Oklahoma, where 41% of total water use goes directly to agricultural irrigation and almost 90% of the land area suffers from drought throughout the year (Droughts.gov, 2022). Estimating root zone soil moisture is typically achieved by solving a hydrological model that describes water movement through soils. Most of the advanced agro-hydrological models today incorporate the Richards equation (Richards, 1931), which captures irrigation, precipitation, evapotranspiration, runoff, and drainage dynamics of water in saturated and unsaturated porous medium such as soil:



$$\frac{\partial \theta(\varphi)}{\partial t} = \nabla \cdot [K(\theta) \nabla(\varphi + z)] - S \quad (1)$$

where  $\varphi$  stands for pressure head,  $\theta$  denotes the soil moisture content,  $K$  is unsaturated hydraulic water conductivity,  $t \in [0, T]$  denotes the time,  $z$  denotes the vertical depth, and  $S$  is the sink term associated with root water extraction, which we ignore without loss of generality. The flux is given by the Darcy's law:  $q = -K(\theta) \nabla(\varphi + z)$ . For unsaturated flow, both  $K$  and  $\theta$  are highly nonlinear functions of  $\varphi$ . For instance, the widely used van Genuchten-Mualem correlations (van Genuchten, 1980) for  $\theta(\varphi)$  and  $K(\theta)$  are:

$$\theta(\varphi) = \begin{cases} \frac{\theta_s - \theta_r}{[1 + (\alpha|\varphi|)^n]^{\frac{n-1}{n}}} + \theta_r, & \varphi < 0 \\ \theta_s, & \varphi \geq 0 \end{cases} \quad (2)$$

$$K(\theta) = K_s \sqrt{\frac{\theta - \theta_r}{\theta_s - \theta_r}} \left\{ 1 - \left[ 1 - \left( \frac{\theta - \theta_r}{\theta_s - \theta_r} \right)^{\frac{n}{n-1}} \right]^{\frac{n-1}{n}} \right\}^2$$

where  $\alpha$  and  $n$  are van Genuchten parameters characterizing different soils, and  $\theta_s$ ,  $K_s$  and  $\theta_r$  denote saturation soil moisture, residual water content, and saturated hydraulic conductivity, respectively. As a result, exact analytical solutions to Equation (1) often do not exist, and numerical solutions rely on discretization of spatial and temporal domains. In particular, the finite volume method (FVM) adopts an integral form of the Richards equation, which captures some valuable physical insights about water flow dynamics (Rathfelder and Abriola, 1994). In addition, FVM is a flexible framework that can be coupled with various other techniques such as the predictor-corrector method (Lai and Ogden, 2015), linearization scheme (Radu et al., 2015), and global random walk approach (Vamos, 2013). Recently, we developed a computationally efficient FVM-based framework for solving the Richards equation by integrating it with adaptive linearization scheme, global random walk method, and multi-layer neural networks (Song and Jiang, 2023). In an illustrative example, we showed that our novel data-driven framework not only generated fast and accurate solutions to the 1-D Richards equation, but also implicitly captured the underlying physical relationships among soil moisture content, pressure head, and flux. In this work, we generalize this computational framework to successfully solving 3-D Richards equation for the first time in the literature.

## 2. FVM Discretization

To apply FVM to Richards equation, we first integrate both sides of Equation (1) to obtain an integral form of the Richards equation over a higher-dimensional control  $V$ :

$$\int_V \frac{\partial \theta(\varphi)}{\partial t} dV = \int_V \nabla \cdot [K(\theta) \nabla(\varphi + z)] dV \quad (3)$$

Next, we can apply the divergence theorem to convert the volume integral on the RHS of Equation (3) into a surface integral over  $S_V$  by introducing the outward pointing unit normal vector denoted as  $\mathbf{n}$ :

$$\int_V \frac{\partial \theta(\varphi)}{\partial t} dV = \int_{S_V} [K(\theta) \nabla(\varphi + z)] \cdot \mathbf{n} dS_V \quad (4)$$

By doing so, FVM is able to incorporate conservation laws, whereas other discretization methods cannot. Specifically, to obtain the discretized version of Equation (4) using FVM,

we discretize the volume integral on the LHS of Equation (4) into a total of  $N_V$  small cells  $V_i$  with  $i = 1, \dots, N_{\omega}$  whose volume is denoted as  $\vartheta_i$ . Each  $V_i$  is associated with surfaces  $\omega_{i,j}$  for  $j = 1, \dots, N_{\omega_i}$  on the RHS of Equation (4) whose area is denoted as  $\mathcal{A}(\omega_{i,j})$ , and we use  $[\cdot]_{\omega_i}$  and  $\mathbf{n}_{\omega_i}$  respectively to denote the operator and outward pointing unit normal vector associated with  $\omega_{i,j}$  upon discretization. Next, to discretize the time domain, we approximate the time derivative  $\partial\theta(\varphi)$  as  $\frac{\theta_i^{m+1,s+1} - \theta_i^m}{\Delta t}$ , where  $\theta_i^{m+1,s+1}$  represents the discretized  $\theta$  in the  $i$ th small cell at the next time step  $m + 1$  and iteration step  $s + 1$ , whereas  $\theta_i^m$  is the converged  $\theta$  value in the  $i$ th small cell at the current time step  $m$ . With this, the discretized version of Equation (4) using FVM becomes:

$$\frac{\theta_i^{m+1,s+1} - \theta_i^m}{\Delta t} \vartheta_i = \sum_{j=1}^{N_{\omega_i}} [K(\theta)\nabla(\varphi + z)]_{\omega_{i,j}} \cdot \mathbf{n}_{\omega_{i,j}} \mathcal{A}(\omega_{i,j}) \quad (5)$$

### 3. Data-Driven Global Random Walk Algorithm

In this section, we introduce a data-driven global random walk (DGRW) approach that seeks to implicitly encapsulate the physical knowledge describing particle movement and conservation within a control volume. Global random walk allows particles to move to neighboring cells simultaneously on any direction or stay in their current cells with a given probability. Thus, in 3-D case, there are a total of seven options for every particle. Let  $n_{i,j,k}^{m,s}$  be the number of particles in the cell  $(i, j, k)$  at fixed-point iteration step  $s$  and time step  $m$ . Correspondingly,  $\delta n_{i',j',k'}^{m,s}$ , where  $i' = i - 1, i, i + 1, j' = j - 1, j, j + 1$ , and  $k' = k - 1, k, k + 1$ , denotes the number of particles moving from cell  $(i', j', k')$  to cell  $(i, j, k)$ . Explicitly, we can write:

$$n_{i,j,k}^{m,s+1} = \delta n_{i+1,j,k}^{m,s} + \delta n_{i-1,j,k}^{m,s} + \delta n_{i,j+1,k}^{m,s} + \delta n_{i,j-1,k}^{m,s} + \delta n_{i,j,k+1}^{m,s} + \delta n_{i,j,k-1}^{m,s} + n_{i,j,k}^{m,s} \quad (6)$$

As we may expect, having such physical knowledge is important for solving Richards equation, which governs the movement of water molecules in unsaturated and saturated soil systems. In fact, Suciú et al. (2021) adopted global random walk concepts and proposed a numerical framework for solving 1- and 2-D Richards equation. In their numerical framework, the authors assumed that the pressure head  $\varphi$  is proportional to the number of particles  $n$  in a cell or control volume. With this assumption, soil moisture content in the cell of interest is simply proportional to the arithmetic mean of the number of particles. While this assumption is valid for diffusion equations (Vamoş et al., 2001), Richards equation is a highly nonlinear convection-diffusion equation, and the exact relationship between  $\varphi$  and  $n$  remains unclear. In fact, we have shown that the numerical framework proposed by Suciú et al. (2021) failed to obtain an accurate solution for 1-D Richards equation (Song and Jiang, 2023).

Since the relationship between  $\varphi$  and  $n$  may not be describable by any basic function, we decide to model the relationship using two multi-layer neural networks (MNNs). Although deep neural network with more layers and neurons could also be used, it is not required as MNN with less number of layers (e.g., 3) is sufficient for learning the relationship between  $\varphi$  and  $n$  given enough neurons in the neural network (Hornik, 1991). In one of the MNNs, we approximate  $\varphi$  as a function of  $n$ ,  $\varphi_{i,j,k} = f(n_{i,j,k})$ , at each fixed time step and iteration step. In the other MNN, we learn the inverse mapping

$f^{-1}$  from pressure head information to the number of particles,  $n_{i,j,k} = f^{-1}(\varphi_{i,j,k})$ . During offline training, we first obtain reference solutions from the global random walk solvers developed by Sucui et al. (2021) (code available at GitHub repository: <https://github.com/PMFlow/FlowBenchmark>). We then add Gaussian noise to these reference solutions to 1) account for the possibly nonlinear relationship between  $\varphi$  and  $n$ , and 2) enhance its generalization performance and the quality of solution obtained (Song and Jiang, 2023). Once offline training is complete, we substitute  $f$  to Equation (5) and derive the following data-driven random walk formulation for the discretized Richards equation:

$$\begin{aligned} n_{i,j,k}^{m,s+1} = & \left[ 1 - \left( r_{i+\frac{1}{2},j,k}^{m,s} + r_{i-\frac{1}{2},j,k}^{m,s} + r_{i,j+\frac{1}{2},k}^{m,s} + r_{i,j-\frac{1}{2},k}^{m,s} + r_{i,j,k+\frac{1}{2}}^{m,s} + r_{i,j,k-\frac{1}{2}}^{m,s} \right) \right] n_{i,j,k}^{m,s} \\ & + r_{i+\frac{1}{2},j,k}^{m,s} n_{i+1,j,k}^{m,s} + r_{i-\frac{1}{2},j,k}^{m,s} n_{i-1,j,k}^{m,s} + r_{i,j+\frac{1}{2},k}^{m,s} n_{i,j+1,k}^{m,s} + r_{i,j-\frac{1}{2},k}^{m,s} n_{i,j-1,k}^{m,s} \\ & + r_{i,j,k+\frac{1}{2}}^{m,s} n_{i,j,k+1}^{m,s} + r_{i,j,k-\frac{1}{2}}^{m,s} n_{i,j,k-1}^{m,s} \\ & + f^{-1} \left( \left( r_{i,j,k+\frac{1}{2}}^{m,s} - r_{i,j,k-\frac{1}{2}}^{m,s} \right) \Delta z - \frac{\theta^{m,s}(n_{i,j,k}^{m,s}) - \theta^{m-1}(n_{i,j,k}^{m-1,s})}{L_{i,j,k}^{m,s}} \right) \end{aligned} \quad (6)$$

where  $r_{i\pm\frac{1}{2},j,k}^{m,s} = \frac{K\left(\theta\left(\varphi_{i\pm\frac{1}{2},j,k}^{m,s}\right)\right)\Delta t}{(\Delta x)^2 L_{i\pm\frac{1}{2},j,k}^{m,s}}$ ,  $r_{i,j\pm\frac{1}{2},k}^{m,s} = \frac{K\left(\theta\left(\varphi_{i,j\pm\frac{1}{2},k}^{m,s}\right)\right)\Delta t}{(\Delta y)^2 L_{i,j\pm\frac{1}{2},k}^{m,s}}$ , and  $r_{i,j,k\pm\frac{1}{2}}^{m,s} = \frac{K\left(\theta\left(\varphi_{i,j,k\pm\frac{1}{2}}^{m,s}\right)\right)\Delta t}{(\Delta z)^2 L_{i,j,k\pm\frac{1}{2}}^{m,s}}$ . Here, we adopt an adaptive linearization scheme inspired by Mitra and

Pop (2019) by adding the term  $L_{i,j,k}^{m,s}(\varphi_{i,j,k}^{m,s+1} - \varphi_{i,j,k}^{m,s})$  on the LHS of Equation (5).

In DGRW algorithm, Equation (6) passes through the inverse mapping  $f^{-1}$  learned from MNN to generate Richards equation solutions in terms of the particle distribution in each cell,  $n_{i,j,k}$ . To convert these solutions to physically meaningful solutions such as the pressure head, flux, and soil moisture content, we apply the trained mapping  $f^{-1}$  and Equation (2). We have proven that the DGRW algorithm is convergent, and the total error

can be estimated using  $\varepsilon = \max_i \left\{ \frac{\|f(n_{i,j,k}^{m,s+1}) - f(n_{i,j,k}^{m,s})\|_F}{\|f(n_{i,j,k}^{m,s+1})\|_F} \right\} = \max_i \left\{ \frac{\|\psi_{i,j,k}^{m,s+1} - \psi_{i,j,k}^{m,s}\|_F}{\|\psi_{i,j,k}^{m,s+1}\|_F} \right\}$ , where  $\|\cdot\|_F$  is the Frobenius norm.

#### 4. An Illustrative Case Study

We modify the 2-D benchmark problem of Haverkamp (1977) that describes groundwater reservoir recharge from a drainage trench by extending it to 3-D. The boundary conditions of the problem are set up as  $\Omega = [0\text{m}, 2\text{m}]^3$ , and  $\Gamma_D = \Gamma_{D_1} \cup \Gamma_{D_2} \cup \Gamma_{D_3} \cup \Gamma_{D_4}$ , in which:

$$\begin{aligned} \Gamma_{D_1} &= \{(x, y, z) \in \partial\Omega | x \in [0\text{m}, 1\text{m}] \wedge y \in \Omega \wedge z = 2\text{m}\}, \\ \Gamma_{D_2} &= \{(x, y, z) \in \partial\Omega | x \in \Omega \wedge y \in [0\text{m}, 1\text{m}] \wedge z = 2\text{m}\}, \\ \Gamma_{D_3} &= \{(x, y, z) \in \partial\Omega | x = 2\text{m} \wedge y \in \Omega \wedge z = [0\text{m}, 1\text{m}]\}, \\ \Gamma_{D_4} &= \{(x, y, z) \in \partial\Omega | x \in \Omega \wedge y = 2\text{m} \wedge z = [0\text{m}, 1\text{m}]\}. \end{aligned}$$

The Dirichlet boundary condition on  $\Gamma_D$  results in the drainage process, whereas the Neumann boundary condition is applied on  $\Gamma_N = \partial\Omega \setminus \Gamma_D$ . The z-direction is point

downward into the ground. The initial conditions describing hydrostatic equilibrium are modified as follows:

$$\varphi(x, y, z, t) = \begin{cases} -2 + 2.2 \frac{t}{\Delta t_D} & \text{on } \Gamma_{D_1} \text{ and } \Gamma_{D_2}, T \leq \Delta t_D, \\ 0.2 & \text{on } \Gamma_{D_1} \text{ and } \Gamma_{D_2}, T > \Delta t_D, \\ 1 - z & \text{on } \Gamma_{D_3} \text{ and } \Gamma_{D_4}, \\ q(x, y, z, t) = 0 & \text{on } \Gamma_N, \\ \varphi(x, y, z, 0) = \begin{cases} 1 - z & \text{on } \Omega \setminus \Gamma_D \\ -2 & \text{on } \Gamma_D \end{cases} \end{cases}$$

We use the parameters listed in List (2016) to represent water flow dynamics in silt loam soil, in which  $\theta_s = 0.396$ ,  $\theta_r = 0.131$ ,  $\alpha = 0.423$ ,  $n = 2.06$ , and  $K_s = 0.0496$  in Equation (2). We choose  $\Delta t_D = \frac{1}{16} d$ ,  $\Delta t = \frac{1}{48} d$ , and  $T = \frac{3}{16} d$ . A rectangular 3-D mesh with 9261 nodes is used, where  $\Delta x = \Delta y = \Delta z = 0.1$ . The convergence tolerance  $\epsilon_0$  is set to be  $10^{-5}$ . Each MNN contains 3 layers and 10 neurons. Overall, 9261 noise-added reference solutions from global random walk solvers developed by Suciú et al. (2021) were obtained, with 70%, 15%, and 15% of them being used for training, validation, and testing, respectively. Bayesian regularization is used to train both MNNs.

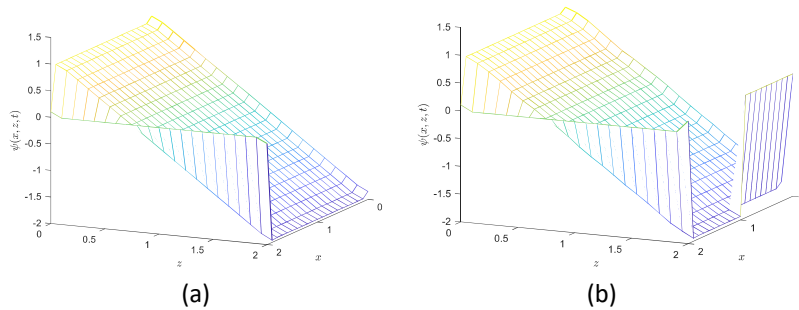


Figure 1. Pressure head obtained from (a) GRW algorithm, and (b) DGRW approach.

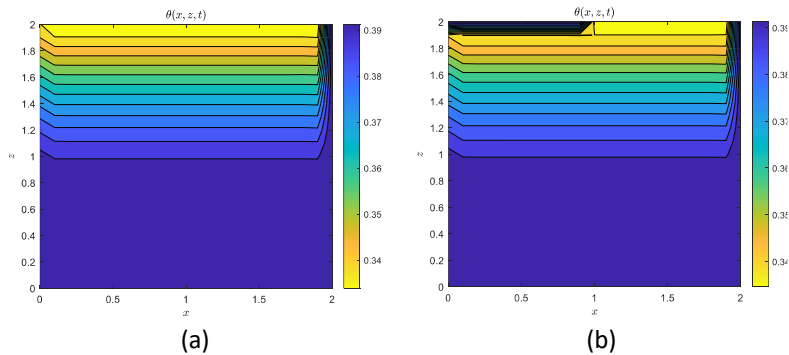


Figure 2. Pressure head obtained from (a) DGRW algorithm, and (b) GRW approach.

Figures 1 and 2 illustrate the pressure head and soil moisture results obtained using the GRW algorithm of Suciú et al. (2021) and our DGRW framework at  $y = 1.7m$  and  $t = T = \frac{3}{16} d$ . In general, our DGRW results show excellent agreement with the GRW results. Nevertheless, we observe that the GRW solutions did not produce jumps in pressure head and soil moisture content in the  $z$ -direction as  $x$  approaches  $1m$  ( $\Gamma_{D_1}$ ), whereas our

DGRW did predict these jumps. As the total simulated time is greater than  $\Delta t_D$ , the initial condition for pressure head should be 0.2m for  $x = [0m, 1m]$  and -2m for  $x = [1m, 2m]$  (Figure 1b). Similarly, jumps in soil moisture and flux in the z-direction are expected at  $x = 1m$  as well. Therefore, by effectively characterizing the complex relationship between pressure head and the number of particles, our DGRW approach is capable of capturing the water flow dynamics embedded in the Richards equation.

## 5. Conclusion

In this work, we propose a novel data-driven DGRW framework to accurately solve 3-D Richards equation for the first time. DGRW synergistically integrates data-driven global random walk theory and FVM to encapsulate physical knowledge and insights describing water flow dynamics in soil systems. We compare the solutions of our DGRW framework and state-of-the-art GRW approach on a 3-D benchmark problem and successfully validate the accuracy and usefulness of our approach in capturing the underlying physics of water flow dynamics.

## References

- Drought.gov, 2022, Current U.S. Drought Monitor Conditions for Oklahoma, <https://www.drought.gov/states/oklahoma>, accessed March 2022.
- R. Haverkamp, M. Vauclin, J. Touma, P.J. Wierenga, G. Vachaud, 1977, A comparison of numerical simulation models for one-dimensional infiltration, *Soil Science Society of America*, 41, 2, 285-294.
- K. Hornik, 1991, Approximation capabilities of multilayer feedforward networks, *Neural networks*, 4, 2, 251-257.
- K. Khand, S. Taghvaeian, A. Ajaz, 2018, Drought and Its Impact on Agricultural Water Resources in Oklahoma, <https://extension.okstate.edu/fact-sheets/drought-and-its-impact-on-agricultural-water-resources-in-oklahoma.html>, accessed March 2022.
- W. Lai, F.L. Ogden, 2015, A mass-conservative finite volume predictor–corrector solution of the 1D Richards’ equation, *Journal of Hydrology*, 523, 119-127.
- F.A. Radu, J.M. Nordbotten, I.S. Pop, K. Kumar, 2015, A robust linearization scheme for finite volume based discretizations for simulation of two-phase flow in porous media, *Journal of Computational and Applied Mathematics*, 289, 134-141.
- K. Rathfelder, L. M. Abriola, 1994, Mass conservative numerical solutions of the head-based richards equation, *Water Resources Research*, 30, 9, 2579-2586.
- L.A. Richards, 1931, Capillary conduction of liquids through porous mediums, *Physics*, 1, 5, 318-333.
- E. Sadler, R. Evans, K. Stone, and C. Camp, 2005, Opportunities for conservation with precision irrigation, *Journal of Soil and Water Conservation*, 60, 6, 371-378.
- Z. Song, Z. Jiang, 2023, A data-driven random walk approach for solving water flow dynamics in soil systems, In: *Proceedings Foundations of Computer-Aided Process Operations and Chemical Process Control Conference*. San Antonio, TX.
- N. Suci, D. Illiano, A. Prechtel, F.A. Radu, 2021, Global random walk solvers for fully coupled flow and transport in saturated/unsaturated porous media, *Advances in Water Resources*, 152, 103935.
- C. Vamos, N. Suci, H. Vereecken, O. Nitzsche, H. Hardelauf, 2001, Global random walk simulations of diffusion, In *Scientific Computing, Validated Numerics, Interval Methods*, pp. 343-354.
- M.T. van Genuchten, 1980, A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Science Society of America Journal*, 44, 5, 892-898.

# Numerical Simulation of No-Newtonian Fluids in Stirred-Tank with proximity impellers

Alvarado-Rodriguez C. E.<sup>a\*,b</sup>, Martinez-Herrera G.<sup>a</sup>, Cortez-Gonzalez J.<sup>c</sup>,  
Murrieta-Dueñas R.<sup>c</sup>

<sup>a</sup>*Departamento de Ingeniería Química, División de Ciencias Naturales y Exactas, Universidad de Guanajuato, Noria Alta S/N, Guanajuato, 36050, México*

<sup>b</sup>*Consejo Nacional de Ciencia y Tecnología, Avenida Insurgentes Sur 1582, Crédito Constructor, Ciudad de México, 03940, México.*

<sup>c</sup>*Departamento de Ingeniería Química y Bioquímica, Tecnológico Nacional de Mexico/ Instituto Tecnológico Superior de Irapuato, 36821 Irapuato, Mexico,*

*ce.alvarado@ugto.mx*

## Abstract

Many industrial applications involve stirred tanks with no-Newtonian rheology; hence the understanding of those mixing systems is required to improve final products. The present work aims to simulate the hydrodynamic no-Newtonian mixing system comprised of three impellers named: U-Type Anchor, Simple Hybrid, and Anchor with radial counterflow. As an industrial application, the water properties were considered for the Newtonian fluid, the ketchup and melted chocolate are used for shear-thinning and shear-thickening no-Newtonian fluids respectively. The performance is evaluated using power consumption, the homogeneity of concentration, velocity vector fields, and numerical mixing time. A modified version of the DualSPHysics code has been used to obtain the concentration fields and numerical mixing time. For the power consumption, the Herschel–Bulkley–Papanastasiou model was used to calculate the power for the different rheology. The numerical results suggest that the Simple Hybrid has the lowest power consumption and better performance.

**Keywords:** Stirred-Tank, Mixing times, CFD, Numerical simulation, SPH.

## 1. Introduction

Computational Fluid Dynamics (CFD) is a very useful tool to analyze in detail the hydrodynamics and mixing performance of stirred tanks in a variety of impellers. The CFD solves the fluid transport equations by computational methods. This can generate a huge amount of information that could not be obtained experimentally. It may also be used as an effective strategy to evaluate the performance of some processes. On mixing systems with a stirred tank, CFD allows obtaining axial and radial flow patterns, velocity profiles, mixing times, stagnation zones, or vortex formation in order to evaluate the performance of the impeller. In all processes, the agitation conditions vary with respect to the behavior of the different fluids, since these can present Newtonian or non-Newtonian behaviors. Recently, there is a great deal of interest in the performance of stirred tanks with no-Newtonian fluids using CFD. The existing literature on the no-Newtonian mixing system has been carried out by the eulerian-grid approach. Nevertheless, eulerian-grid numerical methods have certain drawbacks in dealing with deformable boundaries, moving interfaces, extremely large deformations, and especially

grid regenerating (Abdolahzadeh et al., 2019). The present work aims to simulate the mixing in different rheology using four not reported close-clearance impellers, namely U-type anchor, anchor with radial counterflow, and simple hybrid. Prior to undertaking the simulation cases, a convergence analysis is carried out to establish adequate particle numbers for developing the simulation cases. Once the particle resolution is obtained by convergence analysis, the performance of impellers is evaluated using the macromixing parameters numerical mixing time and power consumption analysis. Subsequently, the velocity fields within the stirred tank are visualized for identifying pattern flows and possible stagnation zones.

## 2. The SPH numerical model

The SPH numerical model is based on the DualSPHysics model (Dominguez et al., 2022) shown in Eq. (1) to (6)

$$\frac{d\rho}{dt} = -\rho\nabla \cdot \mathbf{v} , \quad (1)$$

$$\frac{d\mathbf{v}}{dt} = -\frac{1}{\rho}\nabla P + \Gamma + g , \quad (2)$$

$$P = B \left[ \left( \frac{\rho}{\rho_0} \right)^\gamma - 1 \right] , \quad (3)$$

where  $\rho$  is the density,  $t$  is time,  $\mathbf{v}$  is the velocity vector,  $P$  is the pressure,  $\Gamma$  is the viscosity dissipative term,  $B$  is a reference pressure,  $\gamma = 7$ ,  $\rho_0$  is the reference density and  $g$  is the gravity acceleration.

The no-Newtonian fluids were simulated using the model reported in the DualSPHysics code based on the generalized Herschel–Bulkley–Papanastasiou (HBP) model (Papanastasiou, 1987; Mitsoulis, 2007). The HBP model can be used to simulate shear thinning or thickening materials in the absence of yield strength.

$$\tau^{ab} = \eta_{app}\gamma^{ab} , \quad (4)$$

$$\eta_{app} = \mu|\gamma|^{n_{HB}-1} , \quad (5)$$

$$\gamma = [\nabla\mathbf{v} + \nabla\mathbf{v}^T] , \quad (6)$$

where  $\tau^{ab}$  is the viscosity stress tensor,  $\eta_{app}$  is the apparent viscosity,  $\mu$  is the viscosity or consistency index,  $n$  is the power law index,  $|\gamma|$  is the magnitude of the symmetric strain rate.

For this work the Eq. (7) was implemented in DualSPHysics follow the work reported by (Tartakovsky & Meakin, 2006) to include the mass transfer in the numerical simulations.

$$\frac{dC}{dt} = D\nabla^2 C - \nabla \cdot (C\mathbf{v}) + \mathbf{v} \cdot \nabla C , \quad (7)$$

where  $C$  is the concentration, and  $D$  is the diffusion coefficient.

The value of concentration is uploaded using the Verlet algorithm and the time step according to the Courant–Friedrichs–Lewy condition. The boundary conditions for Eq. (7) are implemented as Neumann conditions avoiding the mass transfer between the fluid particles to the boundary particles. With Eq. (7) is possible to calculate the mixing time considering advection-diffusion within a stirred tank to compare and evaluate the performance of the mixing between the impellers used in this work.

### 3. Geometry and impellers design

The diameter of tank is  $DT = 0.14$  m, and the height of tank is  $HT = 0.245$  m. Two elliptical capsules with 0.03 m of height are added to avoid stagnant zones on the bottom of the tank. In the Figure 1, it is shown the proposed unusual close-clearance impellers analyzed namely U-type anchor (1a), anchor with radial counterflow (1b) and simple hybrid (1c) defined by INOXPA as impellers for standard tanks. All impellers are used for a wide range of viscosities. The U-type anchor impeller has 4 scrapers that allow the removal of the material that adheres to the interior of the stirred tank. The recommended stirring rate works in a range of 3 and 120 rpm. The anchor with radial counterflow impeller is designed to facilitate radial flow. Its optimum working stirring rates range are from 3 to 120rpm. Finally, the self-designed modified simple hybrid impeller is intended to improve mixing times and fluid homogenization.

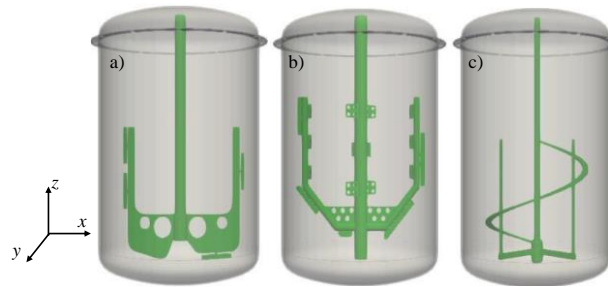


Figure 1. Design of the impellers described in section 3.1. a) U-type anchor, b) anchor with radial counterflow, c) simple hybrid.

### 4. Study Cases

The present work aims to simulate the hydrodynamic of no-Newtonian mixing systems using three impellers named: U-Type Anchor, Anchor with radial counterflow, and Simple Hybrid. As an industrial application, the water properties were considered for the Newtonian fluid, ketchup and melted chocolate are used for shear-thinning and shear-thickening no-Newtonian fluids, respectively. The performance is evaluated using power consumption, velocity vector fields, and numerical mixing time.

#### 4.1. Convergence test.

Before proceeding to examine power consumption analysis, numerical mixing time, and velocity fields, it is important to execute a convergence test. The aim of the convergence test is to determine the adequate number of total particles required in the simulation cases. This test consists of increasing the number of particles in order to find the appropriate resolution of the simulation. In this work, the used particle numbers or particle resolution is defined as follows 75, 000, 125, 000, 250, 000, and 500, 000. In order to calculate the velocity profiles at different resolutions, it was selected the next coordinates  $x = 0.036$  m,  $y = 0.036$  m, and  $z = 0.12$  m to evaluate the velocity at different times.

#### 4.2. The mixing time

The mixing time test is one of the most common procedures for determining the performance of stirred tank system. A longer mixing time is mainly associated with an increase in power consumption due to the operation time. Consequently, the cost of energy is also related to the mixing time. Besides, prolonged mixing time also has an adverse effect on the quality of the product, for instance in crystal grains or damaged microorganisms (Abdolahzadeh et al., 2019). The initial condition of the numerical mixing time test is comprised of one solution concentration of  $C_1 = 1 \times 10^{-6}$  kg/m<sup>3</sup> that



filled the superior fluid volume in the tank and another solution concentration of  $C_2 = 3 \times 10^{-6} \text{ kg/m}^3$  for the inferior fluid volume. The volume of each solution is the same and the initial condition of the mixing time system is shown in Fig. 2a. The criteria for fulfilling the mixing time test was that the final solution must have a concentration range between  $(0.49\text{--}0.51) \text{ kg/m}^3$  in 95 % of the fluid, according to reported literature (Delafosse et al., 2014). The diffusion coefficient of the solute was defined as  $D = 7.5 \times 10^{-7} \text{ m}^2/\text{s}$ .

#### 4.3. Power consumption

Power consumption displays the performance of the agitating process depending on the geometry of the impeller and the physical properties of the fluid. In the present work, power consumption ( $P_o$ ) is used as a parameter in order to evaluate the efficiency of the proposed new close-clearance impellers. The Power consumption was calculated using the measure tool post-processing in version 5 of the DualSPHysics code. For a shear-thinning fluid, the parameters of ketchup at 298.8 K were used. The pseudoplastic rheology parameters are defined as follows: a consistency index of  $K = 6.1 \text{ Pa}\cdot\text{s}$ , a power-law index of  $n = 0.4$ , and a density of  $\rho = 1,325 \text{ kg/m}^3$ . For a shear-thickening fluid, the parameters of the melted chocolate at 319.15 K were chosen. The dilatant rheology is defined as follows; a consistency index of  $K = 0.57 \text{ Pa}\cdot\text{s}$ , a power-law index of  $n = 1.16$ , and a density of  $\rho = 1,600 \text{ kg/m}^3$ , respectively. These parameters were taken from the work of (Joyner & Daubert, 2017).

## 5. Results

### 5.1. Convergence test

Figure 2 displays the velocity profiles as a function of time for different resolutions using the simple hybrid impeller. The maximum and minimum velocity values are developed in a range between 0.06 m/s and 0.22 m/s in a periodic way in time. In order to quantify adequate particle resolution, the root mean square error (RMSE) has been used to calculate the closeness between the velocity values for the highest particle resolution and the others lower particle resolutions.

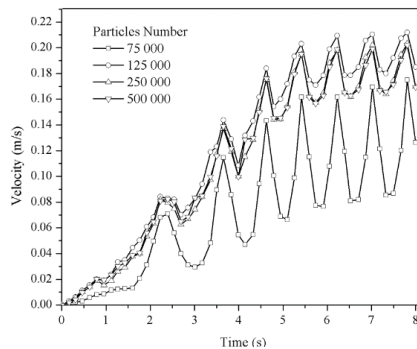


Figure 2. Design of the impellers described in section 3.1. a) U-type anchor, b) anchor with radial counterflow, c) simple Hybrid.

The obtained values of RMSE are  $4.99 \times 10^{-2}$ ,  $1.07 \times 10^{-2}$ , and  $3.14 \times 10^{-3}$  for total particles of 75, 000, 125, 000 and 250, 000, respectively. In summary, the RMSE values from this convergence velocity test suggest that there is no significant difference between the resolutions with 250, 000 or 500, 000 particles. Because of that, it was decided to use 250, 000 for reducing the runtime simulation.

5.2. The mixing time

The mixing time, according to the restriction reported in section 4.2, for each impeller are 306, 804, and 95 seconds for the U-type anchor, anchor with radial counterflow, and simple hybrid, respectively. The simple hybrid impeller has less mixing time compared to the other two impellers based on an anchor geometry. These results can be explained by the flow dynamics within the stirred tank shown in Figure 3. The impellers with anchor geometry produce a mean flow parallel to the wall tank, reducing the movement in the  $z$ -direction, however, the helical geometry produces vortices that induce the movement in the  $z$ -direction improving the blend within the tank.

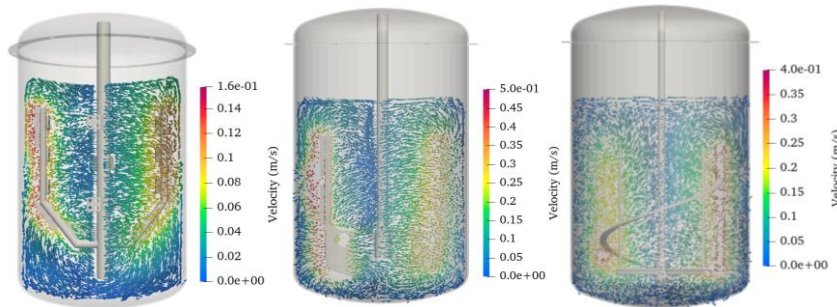


Figure 3. Velocity vectors within the tank. a) U-type anchor, b) anchor with radial counterflow, c) simple Hybrid.

5.3. Power consumption

In Figure 4 is shown the results obtained from the SPH simulations of the power consumption as a function of time for the Newtonian, shear-thickening, and pseudoplastic fluids at 75 rpm. The most striking result to emerge from Figure 4a is that the simple hybrid demonstrates the lowest power consumption among all impellers. In addition, the simple hybrid has also performed the lowest mixing time, as mentioned in section 5.2. In the opposite case, the anchor with radial counterflow impeller demonstrates the worst performance because it uses the highest amount of power energy consumption. It is suggested not to use Counter-current for this mixing system configuration in a Newtonian fluid. In Figure 4b are shown the results obtained from the power consumption of the ketchup as a pseudo-plastic fluid. What stands out in Fig 4b is the rapid increase of power consumption from 5 W to 45 W for the Simple Hybrid impeller in approximately 3 seconds.

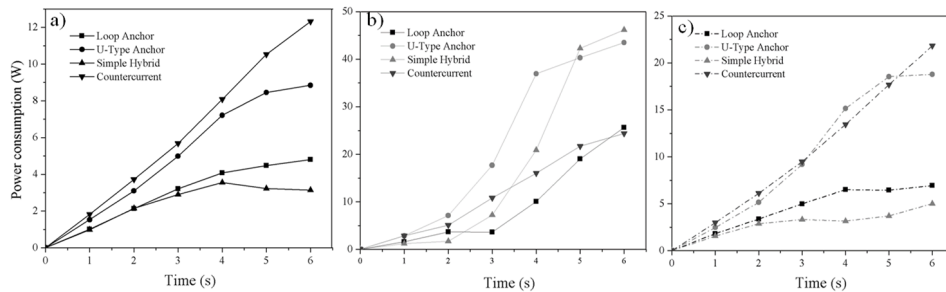


Fig. 4 Power consumption for each impeller and rheology studied.

Conversely, no difference greater than 20 W was observed using Counter-current. These results suggest that the Counter-current promotes the lowest power consumption in pseudo-plastic fluids. However, a mixing time test in a pseudo-plastic fluid may be

recommended in order to gain additional insights into those impellers for future works. In summary, it has been clearly observed from Figure 4b that power consumption in pseudo-plastic rheology is substantially greater than the Newtonian case in all impellers. Finally, in Figure 4c are shown the results obtained from power consumption as a function of time for mixing melted chocolate solution, a shear-thickening fluid, for  $N = 75$  rpm. First, what can be clearly seen in this figure is the similar trend of power input between shear-thickening and Newtonian fluids.

## 6. Conclusions

The aim of the present research was to assess the power consumption in a stirred tank using three studied impellers U-type anchor, anchor with radial counterflow, and simple hybrid for different rheology. The second aim of this study was to investigate the effects in the homogeneity evaluated by the numerical mixing time. The convergence test has determined that the adequate particle number for the proposed mixing configuration is 250, 000 approximately, this means a relationship between the initial distance particle and a diameter of the tank of 0.0162 approximately. On the other hand, the mixing time analysis reveals the importance of a helical blade in the impeller. Thus, the lack of a helical blade in the U-Type Anchor and anchor with radial counterflow impellers presents a poor performance. Thus, this study has contributed to understanding the performance of these unusual impellers. In relation to the study cases simulated, the numerical results suggest that a simple hybrid impeller has a better performance to homogenize and also demonstrates the lowest power consumption in Newtonian and shear-thickening fluids for a low stirring rate of 75 rpm. The investigation highlights the potential usefulness of the SPH method to evaluate the mixing system performance in a variety of rheology from Newtonian to no-Newtonian fluids.

## References

- T. C. Papanastasiou, (1987) Flows of materials with yield. *J Rheol (N Y N Y)* 31:385–404. <https://doi.org/10.1122/1.549926>
- E. Mitsoulis, (2007) Flows of viscoplastic materials: models and computations. *Rheol Rev*: 135–178
- A. M. Tartakovsky, P. Meakin, (2006), Pore-scale modeling of immiscible and miscible fluid flows using smoothed particle hydrodynamics, *Advances in Water Resources*, 29 (10):1464-1478.
- J. M. Domínguez, G. Fourtakas, C. Altomare, R. B. Canelas, A. Tafuni, O. García-Feal, I. Martínez-Estévez, A. Mokos, R. Vacondio, A. J. C. Crespo, B. D. Rogers, P. K. Stansby, M. Gómez-Gesteira, 2022, DualSPHysics: from fluid dynamics to multiphysics problems. *Computational Particle Mechanics*, 9(5), 867-895.
- M. Abdolazadeh, A. Tayebi, P. Omidvar, (2019), Mixing process of two phase non-newtonian fluids in 2d using smoothed particle hydrodynamics. *Computers & Mathematics with Applications* 78, 110–122.
- INOXPA, . Electronic references. URL: <https://docplayer.es/9359211-Agitadores-para-los-depositos-estandares.html>.
- H.S. Joyner, C. R. Daubert, (2017), Rheological Principles for Food Analysis. In: Nielsen, S.S. (eds) *Food Analysis*. Food Science Text Series. Springer, Cham. [https://doi.org/10.1007/978-3-319-45776-5\\_29](https://doi.org/10.1007/978-3-319-45776-5_29).
- A. Delafosse, M.L. Collignon, S. Calvo, F. Delvigne, M. Crine, P. Thonart, D. Toye, (2014) Cfd-based compartment model for description of mixing in bioreactors. *Chemical Engineering Science* 106, 76–85.

# Hybrid Dynamic Surrogate Modelling for a Once-Through Steam Generator

Sindre Stenen Blakseth,<sup>a,b</sup> Leif Erik Andersson,<sup>a</sup> Rubén Mocholí Montañés,<sup>a</sup>  
Marit Jagtøyen Mazzetti,<sup>a</sup>

<sup>a</sup>*Department of Gas Technology, SINTEF Energy Research, Trondheim, Norway*

<sup>b</sup>*Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway*

## Abstract

Four surrogate modelling techniques are compared in the context of modelling once-through steam generators (OTSGs) for offshore combined cycle gas turbines (GTCCs): Linear and polynomial regression, Gaussian process regression and neural networks for regression. Both fully data-driven models and hybrid models based on residual modelling are explored. We find that speed-ups on the order of 10k are achievable while keeping root mean squared error at less than 1%. Our work demonstrates the feasibility of developing OTSG surrogate models suitable for real-time operational optimization in a digital twin context. This may accelerate the adoption of GTCCs in offshore industry and potentially contribute towards a 25% reduction in emissions from oil & gas platforms.

**Keywords:** Digital Twin, Surrogate Modelling, Residual Modelling, Gaussian Process Regression, Neural Networks

## 1. Introduction

Installation of combined cycle gas turbines (GTCCs) offshore can reduce emissions from oil & gas platforms by up to 25% [Mazzetti et al. (2014)]. A digital twin (DT) framework for GTCCs may accelerate the adoption of GTCCs by increasing their reliability and performance. To this end, accurate and trustworthy yet computationally efficient dynamic models for crucial GTCC components are needed. The once-through steam generator (OTSG) has been identified as a particularly important component because its response to transients in the gas turbine load largely governs the GTCC's overall behavior. SINTEF Energy Research has previously developed highly accurate Modelica models for OTSGs [Montañés et al. (2021)]. In this work, we explore four data-driven techniques for creating high-speed, DT-suitable surrogate models based on the Modelica model: Linear regression (LReg), polynomial regression (PReg), Gaussian process regression (GPR), and neural networks (NNs). We use these techniques to develop both purely data-driven surrogate models and hybrid models utilizing the residual modelling technique. The goal of this work is to evaluate the feasibility of developing DT-suitable OTSG surrogate models, e.g., for real-time optimization and process control based on model predictive control schemes. The work is intended to contribute towards accelerating the adoption of GTCCs, thereby reducing emissions from offshore energy production. Secondly, it is of general interest to analyze the relative performance of the different techniques studied.

## 2. Once-Through Steam Generators

### 2.1 Background and Motivation

A waste heat recovery steam generator (HRSG) is a heat exchanger boiler system that recovers waste heat from a given heat source by producing steam from feedwater that is circulated through the heat exchanger. In the context of GTCCs, the heat source is the gas turbine exhaust, and the steam produced by the HRSG is passed through a steam turbine to generate electricity. Consequently, in comparison to standalone gas turbines, GTCCs offer significantly increased efficiency, and thus a corresponding reduction in CO<sub>2</sub> emissions for fixed power production.

In this work, we consider once-through steam generators (OTSGs), which are a particular type of HRSGs common in industrial applications. OTSG configurations are the preferred option in volume- and weight-constrained energy system environments, such as offshore oil and gas installations or floating production, storage and offloading systems. High-fidelity, physics-based, dynamic OTSG models facilitate simulation-based studies to better understand the inherent dynamics of the OTSG system, and to conduct control loop and control structure design studies. Traditionally, operation of OTSG systems has predominantly taken place under steady-state conditions. Therefore, it has been sufficient to consider transient conditions in the offline design phase. Faster models suitable for real-time operational optimization have consequently not received much attention. However, as intermittent renewable energy sources enter the offshore energy mix, computationally efficient models applicable to transient conditions are becoming increasingly important. This motivates the present study of surrogate models for OTSGs, i.e., low-fidelity OTSG models that are designed to capture the main characteristics of OTSG transient behavior while offering significant computational speed-up in comparison to traditional high-fidelity models.

## 2.2 High-Fidelity OTSG Modelling

Development of surrogate models generally requires data from which the main characteristics of the considered system can be extracted. In the present work, this data will stem from a previously published high-fidelity (hi-fi), dynamic OTSG model for offshore combined cycle applications [Montañés et al. (2021)]. The model is developed in the Modelica language and utilizes dynamic energy and mass balances to produce a set of differential algebraic equations describing the OTSG's transient behavior.

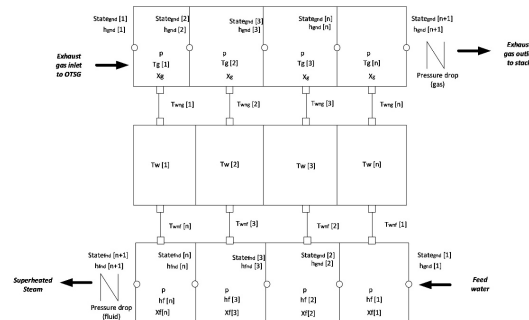


Figure 1: Discretization and main states of high-fidelity OTSG model [Montañés et al. (2021)].

Figure 1 (adapted from [Montañés et al. (2021)]) shows the OTSG model with discretization and main states. The model is based on 1-D lumped pressure flow pipes (for gas and water/steam sides) separated from each other by a wall model (representing the metal wall for heat transfer in a bundle of tubes). A total of  $n$  volumes and  $n+1$  nodes are used to discretize the system in the direction of the flow. Overall, the model includes

physical phenomena related to heat transfer, hydraulics, wall model thermal inertia, dynamic energy and mass balances on both the gas side and the water side [Montañés et al. (2021)]. We consider the OTSG operated under valve-throttling pressure control mode, i.e., with fixed pressure set point of the produced superheated steam [Nord and Montañés (2018), Zotică et al. (2022)]. Important boundary conditions for the model are gas turbine load (exhaust gas inlet mass flow and inlet specific enthalpy) and feed water mass flow rate (which is commonly a manipulated variable for OTSG operation).

### 3. Surrogate Modelling Techniques

In this section, we briefly describe the four surrogate modelling techniques considered in this work: linear and polynomial regression (Section 3.1), neural networks (Section 3.2) and Gaussian Process Regression (Section 3.3). We also describe how we generate the data needed to develop OTSG surrogate models using these techniques (Section 3.4) and how we evaluate the resulting models (Section 3.5). Common for all the surrogate models considered here is that their task is to predict the outlet temperature of the flue gas, as well as the outlet temperature and mass flow of the steam.

#### 3.1 Linear and Polynomial Regression

Polynomial regression (PReg) amounts to approximating some target function  $f(\cdot)$  using a polynomial constructed using  $N_i$  input data variables  $I_i$ , where  $i = 1, \dots, N_i$ . Mathematically, this can be expressed as finding coefficients  $\alpha_{i,n}$  such that

$$\hat{f}_{\text{reg}} := a_0 + \sum_i \alpha_{i,1} I_i + \alpha_{i,2} I_i^2 + \dots + \alpha_{i,N} I_i^d$$

is as close to the true target  $f$  as possible. We say that  $d$  is the degree of the interpolating polynomial. Linear regression (LReg) is simply the special case of  $d = 1$ . We used the Python package `lmfit` [Newville et al. (2014)] to implement our regression models.

#### 3.2 Neural Networks for Regression

Owing to their universal function approximation properties and acclaimed empirical successes, neural networks (NNs) are well suited for surrogate modelling. Here, we consider fully connected feed-forward NNs (cf. [Nielsen (2015)] for an introduction). Such a network with  $n_l$  so-called *layers* can be expressed as

$$\hat{f}_{\text{NN}} = \varphi_{n_l}(\dots(\varphi_2(\varphi_1(Iw_2 + b_1)w_2 + b_2) \dots)w_{n_l} + b_{n_l}),$$

where  $I$  is the NN's input vector, the matrices  $w_1, \dots, w_{n_l}$  (weights) and vectors  $b_1, \dots, b_{n_l}$  (biases) are tunable parameters, and  $\varphi_1, \dots, \varphi_{n_l}$  are non-linear functions known as activation functions. Typically, stochastic gradient descent is used to tune the weights and biases such as to minimize the difference between  $\hat{f}_{\text{NN}}$  and the target function  $f$ . We have implemented our NN models using the Python package `pytorch` [Paszke et al. (2019)]. We use the LeakyReLU activation function with slope parameter 0.01, `pytorch`'s `MSELoss` cost function, the Adam optimizer, and a learning rate of  $1e-5$ .

#### 3.3 Gaussian Process Regression

Gaussian Processes (GPs) are non-parametric, probabilistic kernel methods [Rasmussen and Williams (2006)] that aim to identify an unknown function  $f: \mathcal{R}^{n_u} \rightarrow \mathcal{R}$  from data. It is assumed that the noisy observation of  $f(\cdot)$  are given by

$$y = f(\mathbf{u}) + v,$$

where the noise  $v$  is Gaussian with zeros mean and variance  $\sigma_v^2$ , and  $\mathbf{u}$  is the input, which is assumed to follow a multivariate Gaussian distribution. Smoothness properties of the

underlying function  $f$  are enforced by the choice of mean and covariance function without relying on parametric assumptions [Snelson and Ghahramani (2006)]. A zero mean function and the automatic relevance (ARD) squared-exponential (SE) covariance function are chosen. GP depends on hyperparameters, which are usually unknown and need to be inferred from data. The marginal likelihood is used to estimate hyperparameters. The predictive distribution is the marginal of the normalized joint prior times the likelihood. The integral can be evaluated in closed form. The GP regression model was implemented in Python and the maximisation of the log marginal likelihood was solved with help of the SciPy package [Virtanen et al. (2020)].

### 3.4 Data Generation

Due to the unavailability of suitable operational data from a real GTCC, the high-fidelity OTSG model described in Section 2.2 will be used to create the reference data needed for developing and evaluating our surrogate models. As explained in Section 2.2, the gas turbine load is an essential boundary condition for the model and must therefore be prescribed. To cover a wide range of operating conditions, a randomized time series of one million data points at 1-second intervals, was generated. Every 100 seconds, a load change is occurring with 50% probability. The type of change (step or ramp), the new set point value and the duration of the transition (if it is a ramp) are randomized with the latter two drawn from uniform distributions on [40%, 100%] and [1s, 300s], respectively.

Using the generated gas turbine load sequence, the hi-fi model is used to generate two time series for each of the three outlet variables we aim to predict with our surrogate models. One time series is created using a coarse discretization in the hi-fi model (we call this the low resolution (LR) data), and the other is created using a fine discretization (denoted high resolution (HR) data). The HR data is used as ground truth, both during the model tuning process and the final evaluation. The LR data serves two purposes. Firstly, it can be used as additional input for the surrogate models. Secondly, it will allow us to explore the potential benefit of residual modelling. The concept of residual modelling is to model the residual  $\varepsilon = HR - LR$  instead of the HR data directly [Willard et al. (2022)]. Predictions are then constructed as  $\hat{HR} = LR + \hat{\varepsilon}$ , where  $\hat{\varepsilon}$  is the model's approximation of  $\varepsilon$ . Both the HR and LR data are split into three subsets. The first 980k data points are included in the *training set* (used for tuning model parameters), the next 10k data points go in the *validation set* (used for tuning certain hyperparameters), and the final 10k data points constitute the *test set* (used for evaluating the models).

### 3.5 Model Evaluation

All our surrogate models depend on one or more so-called hyperparameters, i.e., parameters that define a model's structure but are not used directly to compute predictions. Examples include choice of input variables, polynomial degree (PReg), number of layers and neurons per layer (NN) and the number of data points to use for tuning (GPR). To facilitate a fair comparison between the different kinds of models, we conducted a grid search to identify good hyperparameter choices for all models. For each model instance (corresponding to a particular choice of hyperparameters), the normalized root mean squared error (NRMSE) of the models' predictions was computed with respect to the HR test data for each of the three target variables. The sum of the three NRMSEs was taken to represent the overall predictive accuracy of any particular model. For models using old HR data as input, it is important to consider that, when predicting more than one time step into the future, HR data from the previous time step will not be available. Then, the surrogate model's prediction from the previous time step must be used in its place. This may lead to divergent behavior, as is observed for some of our models (cf. Table 1).

## 4. Results

### 4.1 Predictive Accuracy

For each modelling technique and model input choice, Table 1 lists the lowest total NRMSE error observed in our grid search for both normal and residual models using the specified technique and input choice. The available input choices are 1) only inlet data (in), 2) inlet data and LR outlet data (in+LR), 3) same as 2) but also including old HR data (in+LR+old), and 4) same as 2), but also including the residual HR-LR (in+LR+res).

We observe that the most accurate model is a GPR residual model using in+LR+old input. However, several other models achieve comparable performance, and the performance difference between the best PReg, NN and GPR models is generally insignificant from a practical perspective. Even the best LReg models are found to perform quite well, with a NRSME significantly smaller than 1% achievable for all three output variables. Moreover, while the use of LR data appears to be generally beneficial, we observe that only inlet data is sufficient to obtain NRSME values well below 1% for any given output variable. (Note that, in the table, we sum the NRMSE of each predicted variable.) Figure 2 shows the predictions of steam outlet mass flow made by the best model for each modelling technique (bold in Table 1). Based on Figure 2, it is difficult to identify any predominant failure mode for any of the different models.

	in		in+LR		in+LR+old		in+LR+res	
	Normal	Residual	Normal	Residual	Normal	Residual	Normal	Residual
LReg	0.0555	0.0154	0.0089	0.0089	<b>0.0088</b>	0.0514	19.4011	19.3999
PReg	0.0555	0.01525	0.0056	0.0056	<i><b>0.0031</b></i>	NAN	NAN	NAN
NN	0.0562	0.0107	0.0066	0.0061	0.0039	<b>0.0032</b>	0.0406	0.0107
GPR	<i>0.0369</i>	<i>0.0099</i>	<i>0.0045</i>	<i>0.0043</i>	0.0032	<i><b>0.0031</b></i>	<i>0.0036</i>	<i>0.0031</i>

Table 1: Normalized RMSE on the test set, summed over the three output variables, for given model type and input selection. Bold and italics are used for lowest error values in columns and rows, respectively. NAN indicates that the model diverged.

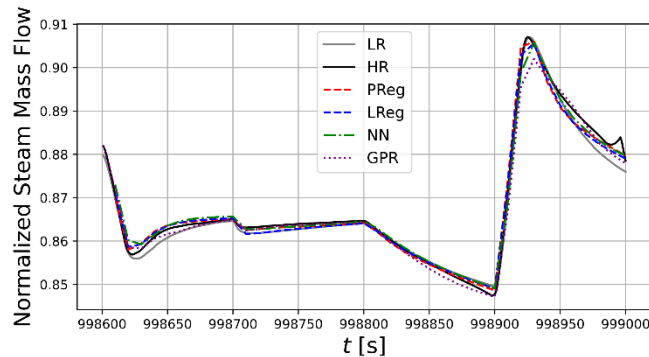


Figure 2: Predictions of steam outlet mass flow by best model within each category, shown along with the corresponding LR and HR data for a short interval within the test set.

### 4.2 Computational Efficiency

We use the number of CPU seconds (wall time) per simulated second of operations (CPUs/s), as measured on a standard, mid-range laptop, to quantify the computational efficiency of the various models. Using respectively the coarse and the fine spatial discretization, we measure 0.00289 CPUs/s and 0.1272 CPUs/s for the hi-fi-model. Thus, the reduced resolution yields a speed-up of roughly 44. In comparison, our linear regression model with only inlet data as input uses around 8e-6 CPUs/s, which



corresponds to a speed-up of roughly 15k. The PReg model is roughly as fast as the LReg model, while the NN and GPR models are roughly 1–1.5 orders of magnitude slower than the LReg model. Consequently, for residual models and models using LR data as input, the LR hi-fi component always dominates the computational expense.

## 5. Discussion

An interesting takeaway from our numerical experiment is that blindly relying on more advanced techniques like neural networks and GPR to be better than simple techniques like regression is not advised. The relation between a model's theoretical representation capacity and its empirical predictive accuracy is not necessarily linear. Indeed, our results show that the relation is not even strictly increasing in general, as 1) our best PReg model outperforms our best NN model, and 2) increasing the number of neurons in the NN (and thereby its representation power) did not improve its accuracy on the *test* set.

Our results also show that the value of hybrid modelling depends on the quality of the physics-based model component, and how that component is integrated into the fully hybrid model. In our case, using LR predictions as input was generally observed to be useful, while residual modelling yielded mixed results. This illustrates the obvious, but easily overlooked fact that residual modelling is only beneficial if the relation between the input data and the residual is simpler than that between the input data and the data to be predicted. This criterion is not always met in practice, as evidenced by our results.

Finally, we conclude that it is feasible to construct OTSG surrogate models suitable for use in real-time optimization procedures within a digital twin framework. This motivates further work, which could include exploration of more advanced neural network-based techniques, such as Long Short-Term Memory networks. Additionally, we believe that looking into the robustness of the methods with respect to noisy data would be valuable from a practical perspective.

## Acknowledgements

This work was supported by DIGITAL TWIN, RCN project no. 318899.

## References

- Mazzetti, M. J., Neksa, P., Walnum, H. T., & Hemmingsen, A. K. T. (2014). Energy-efficiency technologies for reduction of offshore CO<sub>2</sub> emissions. *Oil and gas facilities*, 3(01), 89–96.
- Montañés, R. M., Skaugen, G., Hagen, B., & Rohde, D. (2021). Compact Steam Bottoming Cycles: Minimum Weight Design Optimization and Transient Response of Once-Through Steam Generators. *Frontiers in Energy Research*, 9, 687248.
- Newville, M., Stensitzki, T., Allen, D. B., & Ingargiola, A. (2014). LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python (0.8.0). doi: 10.5281/zenodo.598352
- Nielsen, M.A. (2015). *Neural Networks and Deep Learning*. Determination Press.
- Nord, L. O., & Montañés, R. M. (2018). Compact steam bottoming cycles: Model validation with plant data and evaluation of control strategies for fast load changes. *Applied Thermal Engineering*, 142, 334–345.
- Paszke, A. et al. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Advances in Neural Information Processing Systems*, 32, 8024–8035.
- Snelson, E. & Ghahramani, Z. (2005). Sparse Gaussian processes using pseudo-inputs. *Advances in neural information processing systems*, 18.
- Virtanen, P. et al. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17(3), 261–272
- Willard, J., Jia, X., Xu, S., Steinbach, M., & Kumar, V. (2022). Integrating scientific knowledge with machine learning for engineering and environmental systems. *ACM Computing Surveys*, 55(4), 1–37.
- Williams, C. K., & Rasmussen, C. E. (2006). *Gaussian processes for machine learning* (Vol. 2, No. 3, p. 4). Cambridge, MA: MIT press.
- Zotică, C., Montañés, R. M., Reyes-Lúa, A., & Skogestad, S. (2022). Control of steam bottoming cycles using nonlinear input and output transformations for feedforward disturbance rejection. *IFAC-PapersOnLine*, 55(7), 969–974.

# PINN-based Design of Experiment Concept for Process Model Parameter Identification

Aike Aline Tappe <sup>a</sup>, Subiksha Selvarajan <sup>a</sup>, Caroline Heiduk <sup>b</sup>, Stephan Scholl <sup>b</sup>  
and René Schenkendorf <sup>a,\*</sup>

<sup>a</sup>*Automation & Computer Sciences Department, Harz University of Applied Sciences, Friedrichstr. 57-59, 38855 Wernigerode, Germany; {[atappe.sselvarajan@hs-harz.de](mailto:atappe.sselvarajan@hs-harz.de);*

<sup>b</sup>*TU Braunschweig, Institute for Chemical and Thermal Process Engineering, Langer Kamp 7, 38106 Braunschweig, Germany; {[c.heiduk,s.scholl@tu-braunschweig.de](mailto:c.heiduk,s.scholl@tu-braunschweig.de);*

<sup>\*</sup>*Correspondence: [rschenkendorf@hs-harz.de](mailto:rschenkendorf@hs-harz.de)*

## Abstract

In Process systems engineering, first-principles process models are standard. However, all model predictions, including any model-based design or control policy, suffer from imprecision. To increase the reliability of the process model, precise model parameters and model-based design of experiments (MBD<sub>oE</sub>) are essential. That is, experiments have to be designed for the specific purpose of model calibration. The same is true for data-driven engineering concepts. For example, in the case of physics-informed neural networks (PINNs), the first-principles model is incorporated into the structure of an artificial neural network, thus creating a hybrid framework in terms of model calibration. The PINN setting, in turn, results in different parameter sensitivities compared to the standard parameter identification approach, with profound consequences for the precision of parameter estimates. Also, due to the change in parameter sensitivities, traditional MBD<sub>oE</sub> concepts might fail. Alternatively, MBD<sub>oE</sub> strategies have to be adapted for PINN-based parameter identification and more precise parameter estimates. In particular, Monte Carlo simulations are combined with the PINN concept to choose appropriate experimental designs. As a case study, the proposed concept is demonstrated for a distributed-parameter system of an adsorption process model. The changes in parameter sensitivities and parameter uncertainties are critically discussed and, in terms of novelty, presented with the need for a mandatory MBD<sub>oE</sub> adaptation.

**Keywords:** design of experiments, data-driven engineering, physics-informed neural networks, sensitivities, uncertainty, partial differential equations, adsorption process

## 1. Introduction

Profit margins in the highly competitive and regulated pharmaceutical industry have continuously decreased in the past decades due to increased research and development costs (Behr et al., 2004). In the context of pharmaceutical manufacturing, process systems engineering can be used to design efficient and cost-effective production processes for drugs and other pharmaceutical products. This can involve optimizing the use of raw materials and energy, reducing waste and emissions, and ensuring that the final product meets the necessary quality and safety standards. In addition, process systems engineering can be used to model and simulate different production scenarios, helping to identify potential bottlenecks or inefficiencies in the manufacturing process. To gain an in-depth system understanding of the underlying physical processes, the temporal and spatial resolution of the underlying dynamic processes might be of relevance and interest. Thus,

research activities on distributed parameter systems (DPS) have gained importance over the years. DPSs can be modeled using partial differential equations (PDEs), where, in particular, the model parameters determine the dynamic behavior of the system. However, the literature on parameter identification of PDEs is sparse (Zhang et al., 2017). Alternatively, recent advances in machine learning and artificial intelligence provide new approaches to modeling strategies and parameter identification problems alike. For example, if the process under study cannot be described with first-principles models, ML algorithms are very promising to compensate for that knowledge gap resulting in hybrid process models. However, there is a proven workflow for concepts such as DoE in relation to traditional modeling strategies, but little is known about how these concepts can be applied to hybrid models involving machine learning. Traditionally, the numerical solutions of PDE developed over the past decades typically include the discretization of system domains with mesh-based approaches, including finite differences, finite elements, and finite-volume methods. This strategy has several disadvantages: The individual mesh section is solved separately, and high-resolution systems are computationally expensive. ML technologies to solve these problems are welcome, and PINN methods have recently received a lot of attention (Raissi et al., 2019). In addition to solving PDE systems, the PINN framework enables a new perspective in discovering DPS, i.e., deriving the governing equations and identifying the corresponding model parameters directly. Still, to increase the reliability of the process model, precise model parameters and model-based design of experiments (MBCoE) are essential (Abt et al., 2018). In the case of PINNs, first-principles models are technically incorporated into the structure of artificial neural networks, thus creating a hybrid framework in terms of model calibration (Yazdani et al., 2019) as well. The PINN setting, in turn, results in different parameter sensitivities compared to the standard parameter identification approach (Selvarajan et al., 2022), with profound consequences for the precision of parameter estimates, rarely addressed in the current literature.

## 2. Methodology

### 2.1. Distributed parameter system

Following the notation in (Shukla et al., 2022), partial differential equation systems (PDEs) with the solution  $u(\mathbf{x}, t)$ , including model parameters  $p$ , defined in a domain  $\Omega$  read as follows:

$$f\left(\mathbf{x}; \frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_d}; \frac{\partial^2 u}{\partial x_1 \partial x_1}, \dots, \frac{\partial^2 u}{\partial x_1 \partial x_d}; \dots; p\right) = 0, \quad \mathbf{x} = (x_1, \dots, x_d) \in \Omega, \quad (1)$$

with the boundary conditions:

$$B(u, \mathbf{x}) = 0 \text{ on } \partial\Omega. \quad (2)$$

In process system engineering, PDEs are used to describe the behavior of distributed parameter systems due to their ability to capture the spatial and temporal variation of the system's state variables. The heat equation and the wave equation, for example, are PDEs that describe how heat flows through a material or how waves propagate through a medium, respectively. However, distributed parameter systems can be challenging to analyze and control because the underlying PDEs often have complex solutions that cannot be obtained analytically. Instead, the solution needs to be approximated with numerical methods, which can be computationally intensive. However, several numerical methods, including finite difference methods, finite element methods, and spectral methods are standard in solving PDEs, but the choice of method depends on the specific form of the PDE and the desired accuracy of the solution.

## 2.2. Physics-informed neural networks

Physics-informed neural networks are beneficial for solving partial differential equations because they can handle complex, nonlinear equations that may be difficult to solve using classical numerical methods. These neural networks are trained using both the physical constraints of the problem and the data from the equation, allowing them to learn the underlying patterns and make accurate predictions. Additionally, physics-informed neural networks can be more efficient and scalable than classical numerical methods, particularly for high-dimensional or time-dependent problems. Technically, to solve a parametrized PDE system (Eq. (1)) through PINNs, a deep neural network  $\hat{u}(\mathbf{x}; \theta)$  with trainable meta-parameters  $\theta$  is used. Boundary and initial constraints of the PDE system (Eq. (2)) and the PDE system itself (Eq. (1)) define the system identification problem and are represented in a separate term of the loss function, that governs the neural network training. Emphasis will be taken on the loss function composition in the following section.

## 2.3. System Identification

In PINNs, the loss function is typically composed of multiple terms that reflect different aspects of the problem. The specific terms used in the loss function will depend on the specific problem being tackled, but some common terms include: *Data loss*; this term measures the difference between the predicted output of the neural network and the known data for the problem. For example, if the goal is to solve a partial differential equation with known boundary conditions, the data loss term would measure the difference between the predicted solution and the known boundary conditions. *Physics loss*; this term measures the difference between the predictions of the neural network and the known physical laws and constraints that apply to the problem. For example, if the goal is to simulate the motion of a particle in a gravitational field, the physics loss would measure the difference between the predicted motion of the particle and the known laws of gravity. By combining these different loss terms into a single function, PINNs can learn to predict physical phenomena while also incorporating prior knowledge about the problem. This can help the model to generalize better and make more accurate predictions on unseen data. According to (Shukla et al., 2022), the resulting loss function reads as:

$$\mathcal{L}(\boldsymbol{\theta}, p) = \mathcal{L}^{data}(\boldsymbol{\theta}) + \mathcal{L}^{pde}(\boldsymbol{\theta}, p), \quad (3)$$

with the data loss term  $\mathcal{L}^{data}$  defined for the measurements  $y^{data}$ :

$$\mathcal{L}^{data}(\boldsymbol{\theta}) = \sum_{i=1}^m \sum_{k=1}^K \|y^{data}(\rho_i, t_k) - h(\hat{u}(\rho_i, t_k, \boldsymbol{\theta}))\|_2^2, \quad (4)$$

with the measurement location  $\rho_i \in (0, L)$  at  $K$  discrete time points and the model output  $h(\cdot)$ . Moreover, the physics loss term  $\mathcal{L}^{pde}$  relates to the governing equation as follows:

$$\begin{aligned} \mathcal{L}^{pde}(\boldsymbol{\theta}, p) = & \frac{1}{|\mathcal{J}_f|} \sum_{\mathbf{x} \in \mathcal{J}_f} \left\| f(\mathbf{x}; \frac{\partial \hat{u}}{\partial x_1}, \dots, \frac{\partial \hat{u}}{\partial x_d}; \frac{\partial^2 \hat{u}}{\partial x_1 \partial x_1}, \dots, \frac{\partial^2 \hat{u}}{\partial x_1 \partial x_d}; \dots; p) \right\|_2^2 \\ & + \frac{1}{|\mathcal{J}_b|} \sum_{\mathbf{x} \in \mathcal{J}_b} \|B(\hat{u}, \mathbf{x})\|_2^2, \end{aligned} \quad (5)$$

Thus, the prediction of deep NN  $\hat{u}$  has to fulfill the governing equation system and fit the experimental data. Using the PINN framework for parameter identification relates to total least squares approach for a classical parameter identification problem, i.e., the combination of ordinary least squares (e.g., fitting the data) and input least squares (e.g., fulfilling the equation system), see (Selvarajan et al., 2022) and references therein. However, this also implies that standard MBDoe concepts based on local parameter sensitivities. Fisher information matrix, and the Cramér-Rao inequality may fail (Selvarajan et al., 2022). By deliberately controlling the experimental setup (e.g., sensor

positions, sampling rates, inlet profiles, etc.), a model-based design of experiments can provide more precise and accurate estimates of the model parameters. Practically, the basic prerequisite is that parameter sensitivity and uncertainty measures must be adapted to the underlying problem. In the case of PINN, the physical loss term (Eq. (5)) plays an important role.

### 3. Case study

The adsorption process separates and purifies chemicals and substances. It is used in the chemical, pharmaceutical, and food industries, as well as in environmental remediation and waste treatment, allowing the recovery of valuable resources. For better process understanding and process design, a reliable characterization of the kinetics is essential. In this study, we analyze the use of PINN-based kinetic parameter identification and its implications for the design of experimental concepts to ensure more precise parameter estimates with less experimental effort.

#### 3.1. Mathematical model of the adsorption process

The following assumptions are used in the derivation of a so-called lumped kinetic model (LKM) of the adsorption process: (i) Radial gradients in column concentration were neglected. (ii) The column is isothermal and homogeneously packed. (iii) Constant volumetric flow rate  $v$ . (iv) Lumped contribution of internal and external mass transport resistances represented with a mass transfer coefficient  $k_m$ . (v) The axial dispersion coefficient  $D_{ax}$  is also considered constant. Thus, the governing partial differential equation system of the LKM reads as:

$$\frac{\partial C}{\partial t} = -v_z \frac{\partial C}{\partial z} + D_{ax} \frac{\partial^2 C}{\partial z^2} - \frac{1 - \varepsilon}{\varepsilon} \frac{\partial q}{\partial t}, \quad (6)$$

$$\frac{\partial q}{\partial t} = k_m (q^* - q), \quad (7)$$

where  $t$  is the time and  $z$  is the space coordinate along the column axis;  $C$  and  $q$  are the concentrations in the liquid and solid phases, respectively,  $\varepsilon$  is the bed porosity,  $v_z$  is the average velocity of the fluid,  $k_m$  is the effective linear driving force mass transfer coefficient, and  $D_{ax}$  is the axial dispersion. Note that for the adsorption rate, a linear driving force model (Naidu and Mathews, 2021) is assumed. A detailed description of LKM is reported in the literature (Javeed et al., 2013). The boundary and initial conditions are given by:

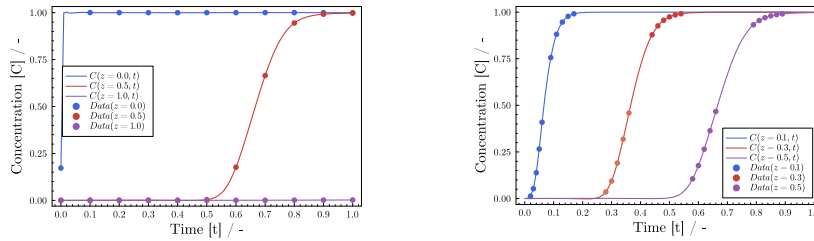
$$\begin{aligned} t = 0 \text{ and } 0 \leq z \leq L &\Rightarrow q = 0 \text{ and } C = 0 \\ t > 0 \text{ and } z = 0 &\Rightarrow C = C_0 \\ t > 0 \text{ and } z = L &\Rightarrow \frac{\partial C}{\partial z} = 0 \end{aligned} \quad (8)$$

Simulated data are generated based on the following parameter values:  $t \in (0,1)$ ,  $z \in (0,1)$ ,  $C_0 = 1$ ,  $D_{ax} = 0.4 \times 10^{-4}$ ,  $k_m = 10$ ,  $\varepsilon = 0.8$  and  $v_z = 0.1$  in a dimensionless form. In Figure 1a, the default experimental design is illustrated. Here, it is assumed that three sensor elements are available to measure the concentration  $C$ , where, for the default configuration, the concentration is selected at the inlet ( $z = 0.0$ ), outlet ( $z = 1.0$ ), and in the middle of the column ( $z = 0.5$ ). Furthermore, for the default experimental design, an equal time sampling with  $\Delta t = 0.1$  is postulated, as well as additive white measurement noise with a standard deviation of  $\sigma = 0.01$ .

#### 3.2. PINN-based parameter identification problem and experimental design

Specifically, the following multi-layer perceptron (MLP) setting is applied: two hidden layers, an input and output layer with 16 nodes each, and a classic sigmoid activation function. The PINN-based parameter identification problem was implemented in Julia,

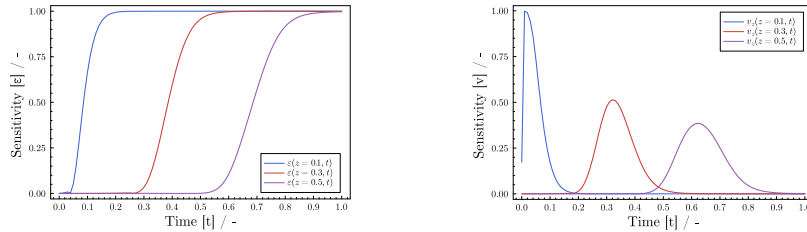
using the packages *Lux*, *NeuralPDE*, and *OptimizationOptimJL*. Furthermore, since the sensor elements at the inlet ( $z = 0$ ) and outlet ( $z = 1$ ) can only measure constant concentration profiles under the given boundary constraints and the considered time horizon, the positions of the sensor elements are optimized to cover the entire dynamic, i.e.,  $C(z_i) = 0$  to  $C(z_i) = 1$ . The optimized sensor layout is shown in Figure 1b with  $z_{opt} = [0.1, 0.3, 0.5]$ . In addition to sensor location, the sampling rate of time recordings is essential as well. Here, the PINN-based MBDoe makes direct use of Eq. (1), i.e., the Pde system does not have to be solved numerically, but the local parameter sensitivities are derived based on the right-hand side of Eq. (1). The resulting parameter sensitivities of  $\varepsilon$  and  $v_z$  are given in Figure 2. As a high parameter sensitivity is crucial for more precise parameter estimates, the sampling interval is adapted accordingly. To this end (Figure 1b), 6 sample points are selected for higher parameter sensitivity regions for  $\varepsilon$  and 5 sample points for  $v_z$ , i.e., the overall number of data points is similar to the default experimental design.



(a) Default experimental design

(b) Optimized experimental design

Figure 1: Dynamic profiles at three different sensor locations of the column/adsorption process. Time points of the actual measurement samples (simulated data) are highlighted.



(a) Parameter sensitivity of  $\varepsilon$

(b) Parameter sensitivity of  $v$

Figure 2: Model parameter sensitivity plot; higher parameter sensitivity are considered for adapting the time sampling of the measurements.

### 3.3. Results

In Figure 3, the resulting parameter estimates are shown. A Monte Carlo simulation concept was used to repeat the PINN-based parameter identification step 200 times for the default and optimized experimental design. These parameter estimates suggest that the changed configuration in the sensor location and the sampling rate leads to a significant reduction in parameter uncertainties when the optimized experimental design configuration is applied. The results also reveal that the PINN-based parameter identification framework has a non-neglectable offset/bias in the parameter estimates, i.e., the true parameter estimates are rarely covered by the scatter plot results. The related mean values and standard deviations are listed in Table 1, indicating a positive effect of the optimized experimental design i.e., lower parameter uncertainties while using the same amount of experimental data.

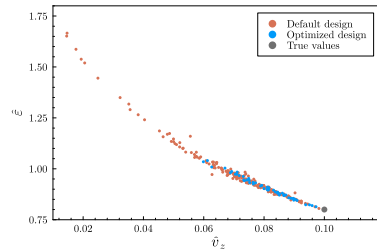


Figure 3: Scatter plot of the two estimated model parameters, including the result of the PINN-based parameter identification for the default and optimized experimental design.

Table 1: Uncertainties for the default and the optimized experimental design.

Parameter	True value	Mean & Std. (default)	Mean & Std. (optimized)
$v_z$	0.1	$0.0718 \pm 0.0164$	$0.0803 \pm 0.0091$
$\varepsilon$	0.8	$0.9779 \pm 0.1543$	$0.9101 \pm 0.0570$

#### 4. Conclusions

As a result, the data collected using an optimal design is more informative and can be used to estimate the model parameters more precisely for the PINN-based parameter identification framework. This can lead to a better understanding of the process under study and can enable more accurate predictions of the behavior of the system. Additionally, using an optimal design can also help to reduce the number of experiments needed to obtain reliable parameter estimates, which can save time and resources. As demonstrated, while PINNs can be a powerful tool for solving PDEs, they might be less optimal for precise parameter identification. Thus, further research on efficient uncertainty quantification and PINN-tailored design of experiment concepts is needed.

#### Acknowledgements

This research was funded by the German Research Foundation, grant number 444703025.

#### References

- V. Abt, T. Barz, M. N. Cruz-Bourmazou, C. Herewig, P. Kroll, J. Möller, R. Pörtner, R. Schenkendorf, 2018. Model-based tools for optimal experiments in bioprocess engineering. *Current Opinion in Chemical Engineering* 22, 244-252.
- A. Behr, V. Brehme, C. Ewers, H. Grön, T. Kimmel, S. Küppers, I. Symietz, 2004. New developments in chemical engineering for the production of drug substances, *Engineering in Life Sciences* 4(1), 15-24.
- S. Javeed, S. Qamae, W. Ashraf, G. Warnecke, A. Seidel-Morgenstern, 2013. Analysis and numerical investigation of two dynamic models for liquid chromatography. *Chemical Engineering Science* 90, 17-31.
- H. Naidu, A. P. Mathews, 2021. Linear driving force analysis of adsorption dynamics in stratified fixed-bed adsorbers. *Separation and Purification Technology* 257, 117955.
- M. Raissi, P. Perdikaris, G. Karniadakis, 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics* 378, 686-707.
- S. Selvarajan, A. A. Tappe, C. Heiduk, S. Scholl, R. Schenkendorf, 2022. Process model inversion in the data-driven engineering context for improved parameter sensitivities. *Processes* 10(9), 1764.
- K. Shukla, M. Xu, N. Trask, G. E. Karniadakis, 2022. Scalable algorithms for physics-informed neural and graph networks. *Data-Centric Engineering* 3.
- A. Yazdani, L. Lu, M. Raissi, G. E. Karniadakis, 2019. Systems biology informed deep learning for inferring parameters and hidden dynamics.
- X. Zhang, J. Cao, R. J. Carroll, 2017. Estimating varying coefficients for partial differential equation models. *Biometrics* 73(3), 949-959.

# Performance analysis of virtual platforms focused on multi-objective process optimization

Jaime David Ponce-Rocha,<sup>a</sup> César Ramírez-Márquez,<sup>b</sup> Ricardo Morales-Rodriguez,<sup>c</sup>

<sup>a</sup>*Dirección de Ingeniería y construcción de plantas, CIATEQ, A.C. Centro de Tecnología Avanzada, Av. Del Retablo 150, Constituyentes Fovissste, Querétaro, Querétaro, 76150, México*

<sup>b</sup>*Departamento de Ingeniería Química, Universidad Michoacana de San Nicolás de Hidalgo, Ciudad Universitaria, Morelia, Michoacán, 58060, México*

<sup>c</sup>*Departamento de Ingeniería Química, División de Ciencias Naturales y Exactas, Universidad de Guanajuato, Noria Alta S/N, Noria Alta, Guanajuato, Guanajuato, 36050, México*

[jaime.ponce@ciateq.mx](mailto:jaime.ponce@ciateq.mx)

## Abstract

Several systematic methodologies for designing and optimizing conceptual processes have been reported by combining different computer-aided tools, mainly based on process simulators and programming environments that can help to improve any processes significantly. Thus, the objective of this work involves the performance analysis for virtual platforms focused on multi-objective process optimization, using a framework based on identical objective functions (economic, exergetic, and environmental targets), similar algorithms and hyperparameters (NSGA-II), for the separation and purification of the benzene, toluene, m-xylene mixture (BTX). Two computational interfaces were developed: Aspen Plus-Matlab (A) and Aspen Plus-Python (B), both using direct communication based on COM. Results were compared in terms of computational time, objective functions, and state variables values. Finally, both computational interfaces have a similar performance in all terms: interface B showed better computational time, but the use of the utopia-tracking approach let to conclude that interface A has better objective function values.

**Keywords:** Simulation-optimization, MOGA, process simulator, programming environments.

## 1. Introduction

The challenges associated with climate change, population growth, and resource limitations, require innovative solutions due to their inherent complexity and interdisciplinary nature (socio-techno-economic nexus), in this connection, chemical engineers can contribute through Process Systems Engineering (PSE) discipline, that offers techniques and tools for process synthesis, design, optimization and/or control problems at multi-scale levels, based on physical and chemical phenomena (Cameron et al., 2019). PSE techniques have a specific role in problem resolution, where two solution pathways for the synthesis and design of a product/process could be employed: 1) reality: experiment-based solution approach, and 2) virtual reality: typical computer-aided



systems solution approach, which includes a) data and models, b) numerical solvers for simulation and/or optimization, and c) workflows: method or algorithm (Pistikopoulos et al., 2021). Several authors have reported methodologies based on virtual reality problem solutions focused on designing and optimizing chemical processes, combining different computer interfaces, linking process simulators and programming environments with several algorithms, mainly using direct communication through ActiveX-COM technology (Ponce-Rocha et al., 2022; Ruiz-Femenia et al., 2020; Sánchez-Ramírez et al., 2016). However, even though these methodologies have been applied in academic, research, and industrial fields, there is not a rigorous study to compare their performance in terms of objective function values and computational aspects. Thus, the objective of this work is to analyze the performance of two virtual platforms for multi-objective process optimization using identical objective functions (economic, exergetic, and environmental targets) with similar algorithms and hyperparameters (NSGA-II), for a separation and purification scheme, where the computational time, objective functions, and state variables values are used as performance criteria.

## 2. Methodology

### 2.1. Virtual platforms focused on multi-objective process optimization

Both virtual platforms proposed in this study were based on a systematic framework that includes four steps: 1) Optimization formulation: the multi-objective approach, state variables, and constraints definition, 2) Objective functions description: analysis and selection of objective functions associated with three targets: economy, exergy and environment (3E), 3) Optimal design: use of different computer-aided tools such as, Matlab and Phyton as programming environments, which allow the algorithm selection; both were combined with a process simulator (Aspen Plus), 4) Case study: separation and purification scheme.

### 2.2. Optimization formulation

The multi-objective optimization is represented as follows (1):

$$\underset{x}{\text{Min}} \quad \underline{Z} = [Z_k(\underline{x})] \quad k = 1, 2, \dots, k. \quad \text{s.t.} \quad \mathbf{h}(\underline{x}) = 0; \quad \mathbf{g}(\underline{x}) \leq 0 \quad (1)$$

Where  $\underline{Z}$  represents a vector with  $k$  objective functions (3E),  $\underline{x}$  is the vector of state variables: number of stages ( $N_S$ ), feed stages ( $F_S$ ), and molar reflux ratios ( $R_R$ ).  $\mathbf{h}(\underline{x})$  and  $\mathbf{g}(\underline{x})$  are the vectors to model equality and inequality constraints associated with the state variables ( $\underline{x}$ ), mass purities (P), and mass recoveries (R).

### 2.3. Objective functions description

Generally, chemical processes have intrinsic characteristics of high energy consumption associated with the use of utilities such as cooling, heating, and electricity. In this sense, the combination of metrics based on 4E analysis (energy-exergy-economy-environment) has given an improvement in terms of process sustainability (Li et al., 2021; Ponce-Rocha et al., 2022). In this study, three of the four components of 4E analysis were selected to perform an optimization task: economic, exergetic, and environmental aspects.

#### 2.3.1. Economy

The economic target was estimated using the Total Annual Cost (TAC), illustrated in Equation 2, which includes the equipment investment cost and payback period, and the sum of annualized utilities cost, referred to cooling, heating, and electrical services. To estimate the equipment cost the module factor approach introduced by Guthrie and modified by Ulrich was used, furthermore, the value of the chemical engineering plant cost index (CEPCI) was set at 801.30 (February-2022).

$$TAC = \frac{\text{Investment cost}}{\text{Payback period}} + \text{Utilities} \quad (2)$$

### 2.3.2. Exergy

The exergy target considers the exergy losses ( $\sigma T_0$ ), this metric is an adequate indicator to identify energy inefficiencies in the system, which considers the heating and cooling processes ( $Q$ ), separation ( $Sep$ ), and concentration changes ( $Conc$ ) in all the equipment ( $j$ ) and process streams ( $i$ ) (see Equation 3).

$$\sigma T_0 = \Delta \dot{E}x_{In} - \Delta \dot{E}x_{Out}$$

$$\sigma T_0 = \sum_{j=1}^n \sum_{i=1}^m [\Delta \dot{E}x_{Q,j} + \Delta \dot{E}x_{Sep,i} + \Delta \dot{E}x_{Conc,i}] \quad (3)$$

### 2.3.3. Environment

The Life-Cycle Assessment (LCA) methodology enables the computation of environmental impact associated with a process, product, or activity. This methodology is normally applied in four steps (ISO 14040): 1) Goal and scope definition, our target focuses on minimizing the process environmental impact, therefore, a fixed production rate of the feed and products was selected, (“cradle-to-gate” analysis). Ecoindicator-99 (EI-99) is integrated as a life cycle assessment tool, based on 11 categories and 3 specific damage categories: human health, ecosystem quality, and resources. 2) Inventory analysis, this step provides the inputs and outputs of mass and energy, construction materials, and utilities associated with the process, and all the burdens must be expressed per unit of reference (Life Cycle Inventory-LCI). 3) Impact assessment, in this step the process data are translated into environmental information, Equation 4 represents the EI99 compute. Where  $\beta_b$  represents the total amount of chemical  $b$  released per unit of reference flow due to direct emissions,  $\alpha_{b,k}$  is the damage caused in each category  $k$  per unit of chemical  $b$  released to the environment,  $\omega_d$  is the weighting factor for damage in category  $d$ , and  $\delta_d$  is the normalization factor for damage of category  $d$ . 4) Interpretation, the results are analyzed and a set of conclusions or recommendations for the system are formulated (Guillen-Gosalbez et al., 2007; Sánchez-Ramírez et al., 2016).

$$EI-99 = \sum_b \sum_d \sum_k \alpha_{b,k} \beta_b \delta_d \omega_d \quad (4)$$

## 2.4. Optimization tools

### 2.4.1. Multi-objective genetic algorithms

Both virtual optimization platforms were made considering the integration of a Multi-Objective Genetic Algorithm (MOGA) based on NSGA-II: Aspen Plus-Matlab (A) using the gamultiobj suite, and Aspen Plus-Python (B) using the PYMOO package. The hyperparameters for both algorithms were as follows: 50 generations, 200 individuals, 0.8 for mutation crossover, and the mutation function: feasible adaptable for gamultiobj, and polynomial for PYMOO.

### 2.4.2. Utopia-tracking approach

Once the optimal solutions set has been obtained by MOGA, the next challenge consists in selecting the optimal point, mainly when many objective functions exist, and the final solutions are non-dominated. The utopia-tracking approach was selected as an adequate tool to choose the best solution point through minimizing the normalized distances

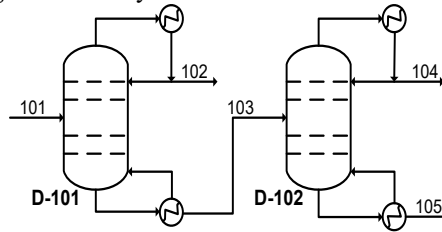
between the objective function values and the utopic point. Equation 5 shows the formulation for the utopia-tracking approach (Zavala and Flores-Tlacuahuac, 2012).

$$\min \|\Phi - \Phi^{L,s}\|_p = \left( \sum_{i=1}^n |\Phi_i - \Phi_i^{L,s}|^p \right)^{\frac{1}{p}} \quad (5)$$

Where  $\Phi$  is any point in the set of final solutions,  $\Phi^{L,s}$  is the utopia point associated with the minimum value for each individual objective function,  $n$  represents the number of the objective functions (three for this study), and  $p$  represents the  $p$ -norm value (one for this study).

### 2.5. Case study

The case study considered the separation and purification of the benzene, toluene, and m-xylene (BTX) mixture. The NRTL-SRK thermodynamic model was used to calculate the phase equilibria (Cui et al., 2019). Figure 1 shows the separation and purification scheme, based on a direct distillation sequence (D<sub>101</sub>, D<sub>102</sub>). The feed stream (101) used in this study has an equimolar flow of 300 kmol/h at 50 °C and 5 bar. Streams 102, 104, and 105 correspond to the recovered and purified compounds; benzene, toluene, and m-xylene, respectively. Furthermore, the vectors to model equality and inequality constraints are defined below. Two assumptions were made for the objective functions computing: for the TAC estimation a payback period was fixed to five years, in addition, for the EI-99 a process lifetime was adjusted to ten years.



$$15 \leq [N_{S_{D_{101}}}, N_{S_{D_{102}}}] \leq 100$$

$$0.01 \leq [R_{RD_{101}}, R_{RD_{102}}] \leq 15$$

$$1 \leq [F_{S_{D_{101}}}, F_{S_{D_{102}}}] \leq 100$$

$$[P_B, P_T, P_X] \geq [0.99, 0.99, 0.99]$$

$$[F_{S_{D_{101}}}, F_{S_{D_{102}}}] \leq [N_{S_{D_{101}}}, N_{S_{D_{102}}}]$$

$$[R_B, R_T, R_X] \geq [0.99, 0.99, 0.99]$$

Figure 1. BTX separation and purification scheme and optimization constraints.

## 3. Results

Table 1 shows the employed optimal function values for both virtual platforms (VP) based on 3D Pareto front and employing the utopia-tracking approach as a tool to make optimal decisions. Besides, the computational times for each evaluation during the optimization step are shown, as well as the optimal state values. All calculations were made on laptop with an AMD Ryzen 7 5800 H processor @ 3.2 GHz, and 16 GB RAM.

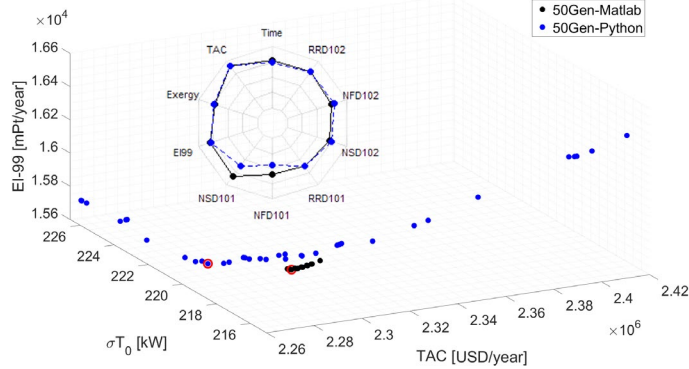
To determine the best optimal results, normalized values of the objective functions were used, thus, the values of  $\|\Phi - \Phi^{L,s}\|_p$  for the results of both virtual platforms were computed, getting a value of 0.45 (interface A), and 0.46 (interface B). For the selection of the optimal point, it was assumed that the objective functions have the same weight.

Table 1. Relevant indicators for the virtual platforms evaluated.

VP	Time (s)	TAC (USD/y)	$\sigma T_0$ (kW)	EI-99 (mPt/y)	N <sub>S</sub> D <sub>101</sub>	F <sub>S</sub> D <sub>101</sub>	R <sub>R</sub> D <sub>101</sub>	N <sub>S</sub> D <sub>102</sub>	F <sub>S</sub> D <sub>102</sub>	R <sub>R</sub> D <sub>102</sub>
A	0.97	2.28E6	216.43	1.59E4	62	30	1.56	53	37	1.97
B	0.94	2.26E6	219.10	1.58E4	47	22	1.56	56	40	1.99
$\Phi^{L,S}$	-	<b>2.26E6</b>	<b>215.03</b>	<b>1.57E4</b>	-	-	-	-	-	-

Figure 2 shows the results, where the black dots correspond to interface A, and the blue dots correspond to interface B. Figure 2.a shows the 3D Pareto front for the final optimal solutions obtained in both virtual platforms, as well as a spider chart to show the relationship between state variables values from both platforms; the dots inside the red circles represent the optimal design for both interfaces and the values are shown in Table 1. Figure 2.b is illustrated to improve the visualization of the three-dimensional plot, thus the objective functions were plotted in 2D arrays, where the values were normalized to have a similar reference system.

a)



b)

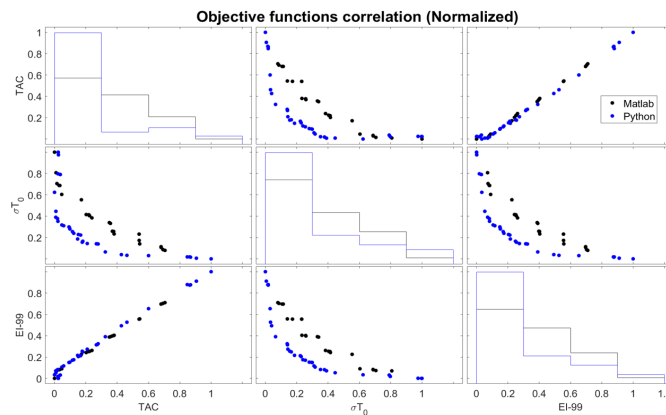


Figure 2. a) Comparative 3D Pareto plot. b) Objective functions normalized matrix.

The results showed that both computational interfaces have a similar performance in all terms, as well as a linear relation between economy-environment objective functions, and non-dominated solutions for economy-exergy, and economy-environment. In terms of computational time, (B) showed a better performance and, it might be possible to argue that Python is the best option because it has better performance in the TAC and the EI-99

objectives, but when the utopia-tracking approach is implemented, it is possible to observe that (A) had a lower value in terms of the three objectives normalized sum in the optimization problem, meaning that it is closer to the utopic value of the problem, which was generated through the minimum values obtained from both Pareto fronts. Based on the graphical results shown in Figure 2, it is likely to observe a close optimal solution for both algorithms, nevertheless, the limits and spread of both Pareto fronts are different for each solution and could be associated with the function coding of each step of the MOGA (fitness function selection, crossover, mutation, etcetera).

#### 4. Conclusions

The proposed tools for selecting the best VP allowed concluded that, based on the decision-making criteria results, it was observed a better computational time for the virtual platform based on Aspen Plus-Python (B) (0.94 vs 0.97 seconds). Regarding the objective function, both VP have similar values, but the interface based on Aspen Plus-Matlab (A) obtained a better-normalized distance to the utopic point ( $\Phi^{L-S}$ ) (0.45 vs 0.46). Although utopia-tracking offers a pathway to select the best solution in the optimal data sets from both Pareto fronts, a more detailed study should be carried out, which could tackle other case studies and focus on topics associated with the data set in the Pareto front and their performance indicators (e.g., distribution, convergence, etcetera).

#### Acknowledgments

Jaime David Ponce-Rocha acknowledges the financial support from “CIATEQ A.C. Advanced Technology Center”, for the development of this project.

#### References

- I.T. Cameron, S. Engell, C. Georgakis, N. Asprion, D. Bonvin, F. Gao, D.I. Gerogiorgis, I.E. Grossmann, S. Macchietto, H.A. Preisig, B.R. Young, 2019, Education in Process Systems Engineering: Why it matters more than ever and how it can be structured, *Comput. and Chem. Eng.*, 126, 102-112.
- C. Cui, X. Zhang, J. Sun, 2019, Design and optimization of energy-efficient liquid-only side-stream distillation configurations using a stochastic algorithm, *Chemical Engineering Research and Design*, 145, 48-52.
- G. Guillen-Gosalbez, J.A. Caballero, L.J. Esteller, M. Gadalla, 2007, Application of life cycle assessment to the structural optimization of process flowsheets, *Comput. Aided Chem. Eng.*, 24, 1163-1168.
- Y. Li, T. Sun, Q. Ye, J. Li, Y. Xu, X. Jian, 2021, Economic and environmental assessment for purification of acetonitrile and isopropanol by reactive coupling extractive distillation, *Separation and Purification Technology*, 275, 119133.
- E.N. Pistikopoulos, A. Barbosa-Povoa, J.H. Lee, R. Misener, A. Mitsos, G.V. Reklaitis, V. Venkatasubramanian, F. You, R. Gani, 2021, Process systems engineering – The generation next?, *Comput. and Chem. Eng.*, 147, 107252.
- J.D. Ponce-Rocha, M. Picón-Núñez, R. Morales-Rodríguez, 2022, A framework for optimal and flexible schemes design under uncertainty & sustainable aspects, *Comput. Aided Chem. Eng.*, 51, 757–762.
- R. Ruiz-Femenia, J. Javaloyes-Antón, R. Salcedo-Díaz, M.A.S.S. Ravagnani, J.A. Caballero, 2020, Integration of Chemical Process Simulators with Algebraic Modeling Languages, *Comput. Aided Chem. Eng.*, 48, 1891-1896.
- E. Sánchez-Ramírez, J.J. Quiroz-Ramírez, J.G. Segovia-Hernández, S. Hernández, J.M. Ponce-Ortega, 2016, Economic and environmental optimization of the biobutanol purification process, *Clean Technology Environmental Policy*, 18, 395-411.

# Mathematical Modeling of Anion Exchange Membrane Water Electrolyzer

Donggyun Lee<sup>a</sup>, Minsu Kim<sup>a</sup>, Il Moon<sup>a\*</sup>

*<sup>a</sup>Department of Chemical and Biomolecular Engineering, Yonsei University, Seoul 03722, Republic of Korea*

*email of the corresponding author: [ilmoon@yonsei.ac.kr](mailto:ilmoon@yonsei.ac.kr)*

## Abstract

Rising energy demand, depletion of fossil fuels and global warming are driving research into eco-friendly alternative energy technologies. Currently, interest in hydrogen as alternative energy is increasing. Among various hydrogen production technologies, there are three major methods for producing hydrogen by water electrolysis: proton exchange membrane water electrolysis (PEMWE), alkaline water electrolysis (AWE), and anion exchange membrane water electrolysis (AEMWE). AEMWE has been attracting attention in recent years because it has the advantage of overcoming the limitations of both electrolysis techniques; Low current density, low pressure, and use of expensive noble metal catalysts.

High-efficiency hydrogen production and durability are determined by the design of the serpentine and the performance of the catalyst and exchange membrane. That is, the distribution of the temperature and pressure in the electrolyzer and the hydrogen production are affected by the design of the serpentine. Manufacturing and experimenting with various serpentine designs are costly and time-consuming. This can be solved by constructing the water electrolyzer system through computational fluid dynamics (CFD). In addition, it is possible to check the flow phenomena, distribution of pressure, temperature, and current density inside the system.

In this study, a mathematical model of the AEMWE system was established. The model for the electrochemical reaction was implemented through user defined functions (UDFs), and reliability was increased through validation with actual experimental data. This model will be used to analyze the efficiency and durability of the water electrolyzer in various serpentine designs.

**Keywords:** Hydrogen production, AEM water electrolysis, Mathematical modeling, Computational fluid dynamics

## 1. Introduction

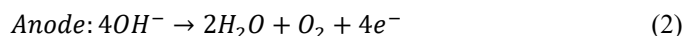
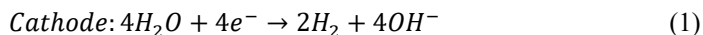
Since traditional energy sources such as fossil fuels generate greenhouse gas (GHG) (e.g., methane, carbon dioxide), there is a need for alternative fuels. Hydrogen is considered an alternative energy source because it has a higher energy content than fossil fuels and can significantly reduce GHG. The hydrogen production process is divided into three categories: grey hydrogen produced through reforming of methane and water vapor, blue hydrogen produced by combining grey hydrogen with carbon capture and storage (CCS), and green hydrogen produced through water electrolysis. In particular, green hydrogen significantly reduces carbon dioxide emissions compared to other methods and is considered a more promising method because water is an infinite source.

AEM electrolysis is a method that combines the advantages of alkaline electrolysis and PEM electrolysis, and has various characteristics: First, hydrogen can be produced at low cost by using a relatively inexpensive non-platinum group metal catalyst. Second, the system is easy to handle due to the low operating temperature.

Here, we analyzed various indices for improving performance and durability through mathematical modeling of AEM electrolysis. Through our developed model, it is possible to identify the exact contribution of mass, momentum, and energy transport.

## 2. Mathematical model

Fig. 1 shows the schematic view of AEM electrolysis cell (AEMEC). AEMEC consists of plate, gas diffusion layer (GDL), catalyst layer (CL), and membrane. As a reactant, water passes through the flow channel and GDL in the plate, and the reaction occurs at the electrode. At the cathode,  $OH^-$  ions are generated through the reaction of Eq. (1) and move to the anode. At the anode, water and oxygen are produced through Eq. (2).



The mixture model of ANSYS FLUENT was used to solve the two-phase flow, and the electrochemical model and source terms were applied as UDFs.

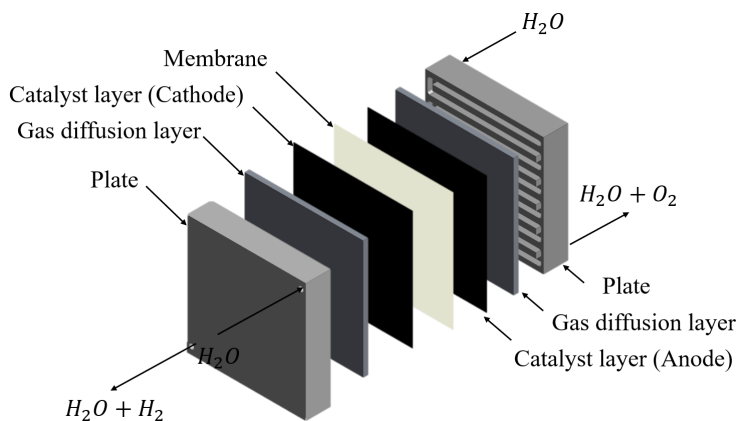


Fig. 1. Schematic view of AEM electrolysis cell.

### 2.1. Governing equations

The continuity equation for liquid-gas mixture is expressed as Eq. (3).  $\vec{v}_m$  is the mass-averaged velocity, and  $\rho_m$  is the mixture density.  $\alpha_k$  represents the volume fraction for the k phase.

$$\nabla \cdot (\rho_m \vec{v}_m) = S_m \quad (3)$$

$$\vec{v}_m = \frac{\sum_{k=1}^n \alpha_k \rho_k \vec{v}_k}{\rho_m} \quad (4)$$

$$\rho_m = \sum_{k=1}^n \alpha_k \rho_k \quad (5)$$

The momentum equation for a mixture can be obtained by summing the individual momentum equations for all phases. It is expressed as Eq. (6), where  $\mu_m$  denotes the viscosity of the mixture

$$\nabla \cdot (\rho_m \vec{v}_m \vec{v}_m) = -\nabla p + \nabla \cdot (\mu_m (\nabla \vec{v}_m + \nabla \vec{v}_m^T)) + \nabla \cdot \left( \sum_{k=1}^n \alpha_k \rho_k \vec{v}_{dr,k} \vec{v}_{dr,k} \right) + S_{mom} \quad (6)$$

$$\mu_m = \sum_{k=1}^n \alpha_k \mu_k \quad (7)$$

Drift velocity is related to relative velocity, which can be expressed as:

$$\vec{v}_{dr,k} = \vec{v}_k - \vec{v}_m \quad (8)$$

The energy equation for liquid-gas mixture is given by Eq. (9), where  $k_{eff}$  represents the effective conductivity.  $S_e$  includes volumetric heat sources.

$$\nabla \cdot \sum_{k=1}^n (\alpha_k \vec{v}_k (\rho_k E_k + p)) = \nabla \cdot (k_{eff} \nabla T) + S_e \quad (9)$$

The internal energy  $E_k$  for a compressible phase is given by Eq. (10), where  $h_k$  is the sensible enthalpy for k phase.

$$E_k = h_k - \frac{p}{\rho_k} + \frac{v_k^2}{2} \quad (10)$$

To solve the species conservation equation, the convection-diffusion equation for i th species predicts the local mass fraction of each species,  $Y_i$ . The species transport equation can be represented as follows:

$$\nabla \cdot (\rho \vec{v} Y_i) = -\nabla \cdot \vec{J}_i + S_i \quad (11)$$



$$\vec{J}_i = -\rho D_{i,m} \nabla Y_i \quad (12)$$

The source terms used in each governing equation are listed in Table 1.

**Table 1**  
Source terms of governing equations.

Source Terms	Anode	Cathode
Mass: liquid phase	$S_m = \dot{N}_{H_2O,m} - \dot{N}_{H_2O,evap} + \frac{iM_{H_2O}}{2F}$	$S_m = -\dot{N}_{H_2O,m} - \dot{N}_{H_2O,evap} - \frac{iM_{H_2O}}{F}$
Mass: vapor phase	$S_m = \frac{iM_{O_2}}{4F} + \dot{N}_{H_2O,evap}$	$S_m = \frac{iM_{H_2}}{2F} + \dot{N}_{H_2O,evap}$
Momentum	$S_{mom} = -\frac{\mu}{k} \vec{v}$	
Energy	$S_e = (V_{cell} - 1.48V)i - \Delta H_{vap} \frac{\dot{N}_{H_2O,evap}}{M_{H_2O}}$	
Species: $H_2$	$S_{H_2} = 0$	$S_{H_2} = \frac{iM_{H_2}}{2F}$
Species: $O_2$	$S_{O_2} = \frac{iM_{O_2}}{4F}$	$S_{O_2} = 0$

## 2.2. Electrochemical model

In the electrochemical model, the total voltage is expressed as the sum of open circuit voltage (OCV), activation overpotential, and ohmic overpotential. OCV is the voltage when no current flows and is derived from the Nernst equation.

$$V = E_{OCV} + \eta_{act} + \eta_{ohm} \quad (13)$$

$$E_{OCV} = 1.229 - 0.846 \times 10^{-3}(T - 298.15) + \frac{RT}{2F} (\ln p_{H_2} + 0.5 \ln p_{O_2}) \quad (14)$$

The activation overpotential is related to the electrochemical reaction at the electrode and is expressed by the Butler-Volmer equation (Eq. (15)).

$$\eta_{act} = \frac{RT_a}{2\alpha_a F} \operatorname{arcsinh}\left(\frac{i}{2i_{0,a}}\right) + \frac{RT_c}{2\alpha_c F} \operatorname{arcsinh}\left(\frac{i}{2i_{0,c}}\right) \quad (15)$$

$$i_0 = i_{0,ref} \exp\left[-\frac{E_{act}}{R} \left(\frac{1}{T} - \frac{1}{T_{ref}}\right)\right] \quad (16)$$

Ohmic overpotential is a term related to resistance when current flows and is expressed by ohmic's law.

$$\eta_{ohm} = R \times i \quad (17)$$

The relationship between current and voltage is expressed as Eq. (13), and the current according to the applied voltage is calculated through Newton's method through UDFs.

### 3. Results and discussion

Fig. 2 illustrates the validation with the experimental data. When the reference parameters were used in the electrochemical model, there was a difference from the experimental data. To solve this problem, the parameter values were estimated using a genetic algorithm and calibrated to the experimental data

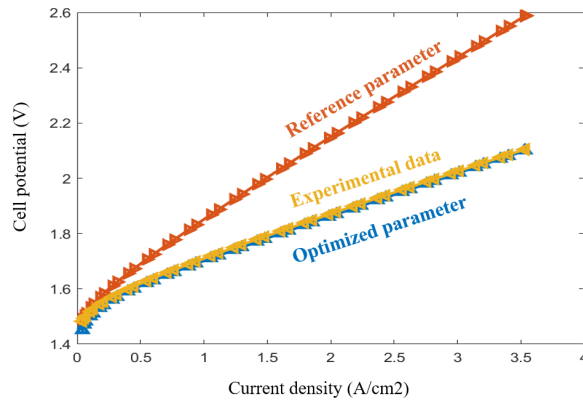


Fig. 2. Comparison between simulation results and experimental data.

Fig. 3 shows the gas volume fraction at the cathode using CFD simulation. In general, the gas volume fraction increases from inlet to outlet. Gas tends to accumulate at the tip of the cell.

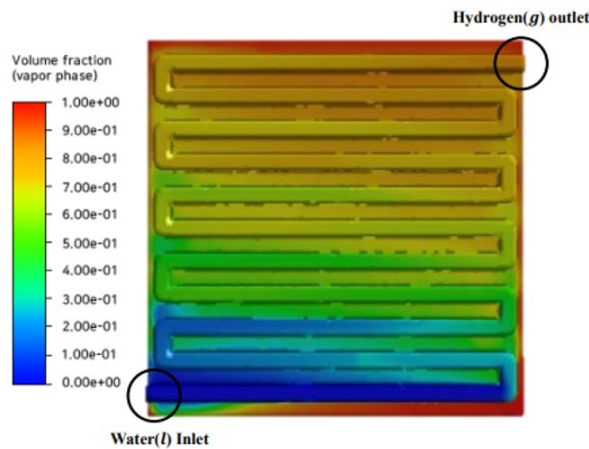


Fig. 3. Distribution of the gas volume fraction at the cathode.

In this study, 3-dimensional AEMEC model was constructed by combining a mixture model and an electrochemical model using UDFs. Durability and performance were analyzed by analyzing temperature and pressure distribution and hydrogen fraction in the cell. As future work, we plan to compare the effects of various serpentine designs, and

find the optimal design achieving the highest performance and durability of cell system by optimizing serpentine configurations.

## References

- Hamish A. Miller, 2022, Green hydrogen from anion exchange membrane water electrolysis, *Current Opinion in Electrochemistry*, 36, 101122.
- S. T. Revankar, 2010, Development of efficient flowsheet and transient modeling for nuclear heat coupled sulfur iodine cycle for hydrogen production, Purdue Univ.
- D. Xu, 2018, Earth-abundant oxygen electrocatalysts for alkaline anion exchange membrane water electrolysis: Effects of catalyst conductivity and comparison with performance in three electrode cell, *ACS Catalysis*, 9, 7.
- I. Pushkareva, 2020, Comparative study of anion exchange membranes for low-cost water electrolysis, *International Journal of Hydrogen Energy*, 45, 26070.
- H. A. Miller, 2020, Green hydrogen from anion exchange membrane water electrolysis: a review of recent developments in critical materials and operating conditions, *Sustainable Energy & Fuels*, 4, 2114.
- Chaochao Cheng, 2021, Numerical investigation on the feasibility of metal foam as flow field in alkaline anion exchange membrane fuel cell, *Applied Energy*, 302, 117555.
- Kui Jiao, 2014, Three-dimensional multiphase modeling of alkaline anion exchange membrane fuel cell, *International Journal of Hydrogen Energy*, 39, 5981.
- Zhiwen Ma, 2021, A comprehensive modeling method for proton exchange membrane electrolyzer development, *International Journal of Hydrogen Energy*, 46, 17627.
- Hao Deng, 2016, Modeling of hydrogen alkaline membrane fuel cell with interfacial effect and water management optimization, *Renewable Energy*, 91, 116.
- Dang Saebea, 2019, Model based evaluation of alkaline anion exchange membrane fuel cells with water management, *Chemical Engineering Journal*, 374, 721.
- D.S. Falcao, 2020, A review on PEM electrolyzer modelling: Guidelines for beginners, *Journal of Cleaner Production*, 261, 121184.
- Haluk Görgün, 2006, Dynamic modelling of a proton exchange membrane (PEM) electrolyzer, *International Journal of Hydrogen Energy*, 31, 29.

# An Aided Optimization Method for Molecular Reconstruction of Crude Oil Based on Pre-optimized Dataset

Chen Zhang,<sup>a,b</sup> Tong Qiu,<sup>a,b</sup>

<sup>a</sup> *Institute of Process Systems Engineering, Department of Chemical Engineering, Tsinghua University, Beijing 100084, China*

<sup>b</sup> *Beijing Key Laboratory of Industrial Big Data System and Application, Tsinghua University, Beijing 100084, China*  
[qiutong@tsinghua.edu.cn](mailto:qiutong@tsinghua.edu.cn)

## Abstract

As a method of converting easy-to-measure physical properties into detailed composition of feedstocks through distribution law and optimization calculation, molecular reconstruction method plays an important role in providing refineries with detailed composition data for molecular management. Introducing the knowledge of petroleum components, including the distribution pattern of homologues and the range of distribution parameters, can effectively enhance the effect of molecular reconstruction. We propose an aided optimization method to assist in the molecular reconstruction of crude oil. This method helps the optimization algorithm to obtain faster and more accurate results by using a pre-optimized dataset added to the initial population of the particle swarm optimization algorithm. The proposed aided optimization method was validated on 52 sets of crude oil data and achieved an average objective function reduction to 82.8% of the original and an average optimization time reduction to 87.5% of the original when using approximately 20% of the pre-optimized data points as aided data.

**Keywords:** Molecular reconstruction; Crude oil; Structure-oriented lumping; Pre-optimized dataset; Aided optimization method.

## 1. Introduction

The refining process is the separation, transformation and purification of oil molecules, in which the average molecular weight is gradually reduced, and various types of molecules are gathered together in an orderly manner. As the crude oil resources gradually become heavier, the government's environmental requirements for the production process become more and more stringent, and the description of the refining process becomes more and more detailed until it reaches the molecular level of description. Molecular management is a combined technical solution for petrochemical process system optimization based on the molecular level. Due to the limitations of analytical techniques and production costs, it is often difficult to perform an exhaustive experimental analysis of all oil products (Stratiev et al., 2019). Molecular reconstruction technology has gained wide research and application by using the physical properties that can be quickly measured to obtain detailed molecular compositions through optimization, which provide the basic data for molecular management (Wu and Zhang, 2010).

In recent years, besides the research on the methods and processes of molecular reconstruction itself (Bi and Qiu, 2019), a series of aided methods have been proposed to

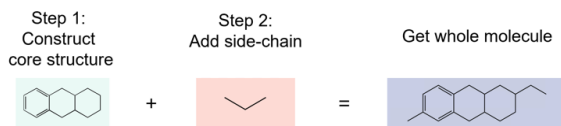
introduce information and knowledge in molecular reconstruction to improve the accuracy and computational speed. The aided methods mainly include (a) predicting the range of optimization parameters through historical data, to improve the speed and accuracy (Bi and Qiu, 2018); (b) based on the known distribution law of oil components, the virtual oil components are constructed to find similar oil components to assist in the optimization (Guan and Zhang, 2022); (c) achieving better prediction results by introducing neural network models in the optimization (Ma et al., 2022).

Considering that the distribution law of crude oil is not completely clear, an effective virtual oil fraction cannot be constructed. In this article, we propose an aided optimization method based on the results of molecular reconstruction of actual crude oil components. The similarity components of crude oil are found according to the similarity calculation method. By adding similar components to the initial particle population of the PSO algorithm, an enhancement to the optimization results is achieved. This aided optimization method not only demonstrates the enhancement effect of pre-optimized datasets on molecular reconstruction, but also shows the importance of introducing information and knowledge in molecular reconstruction.

## 2. Methods

### 2.1. Molecular construction based on SOL

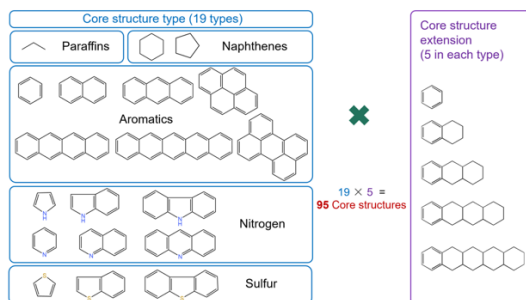
The first step in the molecular reconstruction process is to represent molecules appropriately. We chose the structure-oriented lumping (SOL) method as the way to represent molecules and build a molecular library. Based on previous work, we use 21 structural increments to describe the composition of the molecule (Chen et al., 2018). The structure of each molecule can be described by the number of these 21 structural increments. As shown in Figure 1, a two-step approach is used to construct the molecular library by first constructing representative core structures, followed by adding different lengths of side chains to each core structure to complete all the molecules. For the physical properties of the constructed molecules, the group contribution (GC) method is used to calculate the relevant physical properties.



**Figure 1.** Molecular construction steps.

### 2.2. Calculation of mole fraction of molecules

After constructing the molecular library, it is necessary to describe how to calculate the mole fraction of each molecule. Considering the complexity of crude oil, a two-layer homologous series distribution is used to describe the distribution of the molecular library. As shown in Figure 2, the first layer of homologous series distribution is the homologous series distribution of the core structure. There are 19 types of core structures, and each type has 5 core structures. The exponential distribution is used to describe the probability density function in each core structure type, and the histogram distribution is used to describe the mole fraction of each core structure type. The second layer of homologous series distribution is the length distribution of the side chains. Side chains with a length of 1-50 carbon numbers are added to each core structure, and the gamma distribution is used to describe the carbon number distribution of the side chains.

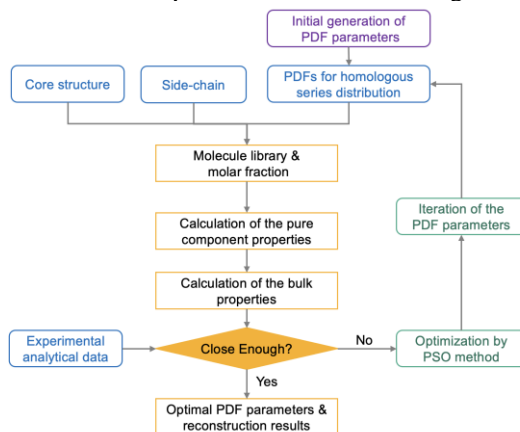


**Figure 2.** The core structure types and the homologous series distribution of the core structure.

According to the two-layer homologous series distribution and probability density function, the mole fractions of all molecules can be calculated. The physical properties of the reconstructed oil can be calculated based on the mole fractions of all molecules, the physical properties of all molecules and the mixing rule.

### 2.3. Optimization of molecular reconstruction of crude oil

The goal of the optimization of the molecular reconstruction is to make the difference between the calculated properties and the experimentally analyzed properties as small as possible, and the flow chart of the optimization is shown in Figure 3.



**Figure 3.** Flow chart of the optimization of molecular reconstruction.

The objective function is the weighted average of the absolute relative deviation values of physical properties, and the optimization variables are the parameters of the probability distribution, with a total of 227 optimization variables. The particle swarm optimization (PSO) is used as the optimization algorithm. The PSO can efficiently search for problems with large solution spaces and find candidate solutions without sufficient information. The particle swarm size is set to 200 based on the number of optimization variables, computational speed, and accuracy. The optimization process is calculated in MATLAB 2022a, and its own PSO algorithm is used for calculation.

### 2.4. Construction of pre-optimized dataset

Based on the molecular reconstruction proposed above, the molecular reconstruction optimization calculations are performed on the 52 sets of crude oil data we have, and the results (values of each optimization variable) are aggregated to obtain the pre-optimized dataset. The pre-optimized dataset will serve as the basic data used by the aided

optimization method to assist other crude oils in molecular reconstruction, and it will also be used as a benchmark for comparing the results of the aided optimization method.

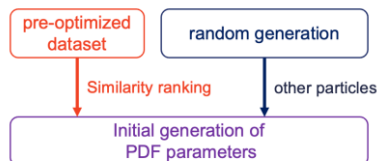
### 2.5. Using pre-optimized dataset to aid optimization

When new crude oil is available for molecular reconstruction, the process of using the pre-optimized dataset to aid in optimization can be divided into 3 steps.

Step 1, find similar data. Find data points in the pre-optimized dataset that are similar to the new crude oil. The Normalized Euclidean distance was used to measure the distance between the new crude oil data and data point in the pre-optimized dataset, which can effectively avoid the influence caused by different units of different properties and no pre-processing of data is required either. All 14 physical properties used in molecular reconstruction were used as the components of the distance calculation. The calculation formula of the Normalized Euclidean distance is shown in Equation 1, where  $u$  is the vector of a data point,  $v$  is the vector of another data point, and  $V$  is the variance of vectors  $u$  and  $v$ . According to the distance formula, the distance between the new crude oil data and all the data points in the pre-optimized dataset can be calculated, and the smaller the value, the more similar the two data set are.

$$distance = \sqrt{\sum_{i=1}^n \frac{(u_i - v_i)^2}{v_i}} \quad (1)$$

Step 2, introduce the pre-optimized data points. Based on the similarity data calculated in step 1, the pre-optimized data are sorted by similarity and the most similar pre-optimized data points are selected in preference. As shown in Figure 4, a certain number of pre-optimized data points are used as part of the initial particle population of the PSO algorithm in the solution of the optimization problem.



**Figure 4.** Initial particle and swarm construction process.

Step 3, perform the optimization calculation. After adding some of the pre-optimized data points to the initial particles, the remaining initial particles for the PSO algorithm are generated using a randomized approach as shown in Figure 4. Using these initial particles, the optimization computation is started following the process shown in Figure 3.

## 3. Results and discussion

The proposed aided optimization method is applied to the molecular reconstruction process of crude oil. Due to the limitation of the size of the available crude oil dataset (only 52 sets of crude oil with physical properties and calculation results data), we adopted the method of selecting one set of crude oil data as the test data one by one and the remaining 51 sets of data as the pre-optimized dataset for the aided optimization. The most similar 3/5/10/20 pre-optimized data points were selected as aided data points in the aided optimization calculation for each crude oil data, respectively. The optimization problem was solved in MATLAB 2022a, and the entire computation was performed on the Intel Xeon 4216 CPU @ 2.1GHz.

First, molecular reconstruction is performed on a set of crude oil data. The same values are used for all optimization parameters except the initial particle population. Calculations were performed for different number of aided data, and the results are shown in Table 1.

**Table 1.** The results obtained with the aided optimization method.

Number of Aided data	0	3	5	10	20
Objective Fun Value	100%	88.4%	82.5%	77.7%	97.4%
Optimization Time	100%	22.8%	19.8%	39.1%	34.4%
Property	Absolute relative error of the calculation result				
Carbon	0.27%	0.01%	0.01%	0.02%	0.11%
Hydrogen	1.70%	0.00%	0.00%	0.05%	0.60%
Sulfur	0.06%	0.00%	0.00%	0.00%	0.00%
Nitrogen	0.06%	0.00%	0.00%	0.00%	0.00%
Paraffins	0.05%	0.00%	0.00%	0.00%	0.00%
Naphthenes	0.06%	0.00%	0.00%	0.00%	0.00%
Aromatics	0.03%	0.00%	0.00%	0.00%	0.00%
5 vol% BP	5.65%	1.35%	0.21%	0.21%	5.65%
10 vol% BP	0.56%	0.13%	1.77%	3.86%	0.10%
30 vol% BP	4.54%	4.18%	0.40%	0.82%	0.31%
50 vol% BP	8.41%	0.17%	0.00%	7.40%	3.07%
70 vol% BP	0.01%	2.65%	3.49%	0.00%	2.31%
90 vol% BP	4.89%	7.17%	7.28%	4.32%	7.11%
95 vol% BP	7.55%	10.97%	10.11%	7.52%	10.11%

The objective function value and optimization time of the molecular reconstruction results without pre-optimized data are recorded as 100%, and the percentage share of the other results with pre-optimized data is calculated. It can be found that the inclusion of pre-optimized data points all reduces the objective function value and optimization time to some extent, respectively.

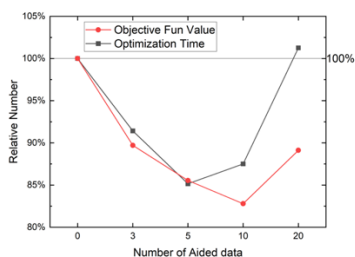
For the objective function, the value decreases gradually at first with the addition of pre-optimized data points and reaches the lowest level of 77.7% of the baseline value when 10 pre-optimized data are added (about 20% pre-optimized data). This indicates that the inclusion of pre-optimized data points effectively helps the whole population to find better results during the optimization. When 20 pre-optimized data were used (about 40% pre-optimized data), the objective function value increased to 97.4% of the baseline value. This situation indicates that only an appropriate proportion (around 20%) of all pre-optimized data points will work best, and the inclusion of more pre-optimized data points will not effectively help the population to find better results and generate positive effects due to the low similarity between them and the calculated data. By checking the similarity values, it was found that the pre-optimized data with distance values less than 4 had a great aided optimization effect, and the pre-optimized data with distance values greater than 4 had an insignificant or even counterproductive aided optimization effect.

For the optimization time, the addition of the pre-optimized data was also effective, which reached the lowest level of 19.8% of the baseline value when 5 pre-optimized data (about 10% pre-optimized data) were added. The addition of the pre-optimized data provides more information during the optimization and reduces the overall population disorder, allowing the algorithm to converge faster and more accurately. When 10 pre-optimized data are added, the optimization time rises to 39.1% of the baseline value, which is mainly due to the addition of more pre-optimized data, driving the algorithm to go for better results, which increases the number of iterations and makes the optimization time grow. Then, all the data were calculated separately with the aided optimization method, and the results were recorded. The results are compared with those obtained without using the pre-optimized data, and the percentage share of the results is calculated. The average calculation results of all data are shown in Table 2 and Figure 5.

**Table 2.** The average calculation result of all data.

Number of Aided data	0	3	5	10	20
Objective Fun Value	100%	89.7%	85.5%	82.8%	89.1%
Optimization Time	100%	91.4%	85.1%	87.5%	101.3%





**Figure 5.** The trend of calculation results with the number of aided data.

According to Table 2 and Figure 5, when 10 pre-optimized data points are added, the average objective function value reaches the lowest level, reaching 82.8% of the baseline value. When 5 pre-optimized data points are added, the average optimization time reaches the lowest level, reaching 85.1% of the baseline value. Considering the objective function value and optimization time together, the best results can be achieved when 10 pre-optimized data (or about 20% pre-optimized data) are added in molecular reconstruction.

#### 4. Conclusions

In conclusion, this novel work proposes an aided optimization method for molecular reconstruction of crude oil based on a pre-optimized dataset. With the pre-optimized dataset and similarity calculation, the results of historical molecular reconstruction can be introduced into the optimization of the molecular reconstruction as crude oil distribution knowledge, which plays an important role in enhancing the optimization results. By the characteristics of the PSO algorithm, the added pre-optimized data can effectively lead the whole population to a better solution, while the optimization time is simultaneously reduced due to the addition of more information and knowledge. The effectiveness of the proposed aided optimization method is verified on crude oil data and shows the potential for future application to real industrial scenarios to further improve the accuracy and speed of molecular reconstruction.

#### References

- Bi, K., Qiu, T., 2019. Novel Naphtha Molecular Reconstruction Process Using a Self-Adaptive Cloud Model and Hybrid Genetic Algorithm–Particle Swarm Optimization Algorithm. *Ind. Eng. Chem. Res.* 58, 16753–16760.
- Bi, K., Qiu, T., 2018. A high-performance molecular reconstruction method with parameter initialization based on PCA, in: *Computer Aided Chemical Engineering*. Elsevier, pp. 2005–2010.
- Chen, J., Fang, Z., Qiu, T., 2018. Molecular reconstruction model based on structure oriented lumping and group contribution methods. *Chinese Journal of Chemical Engineering* 26, 1677–1683.
- Guan, D., Zhang, L., 2022. Initial guess estimation and fast solving of petroleum complex molecular reconstruction model. *AIChE Journal* 68.
- Ma, F., Zheng, X., Han, C., Wang, J., Sun, W., 2022. Molecular Reconstruction of Naphtha based on Physical Information Neural Network. *IFAC-PapersOnLine* 55, 186–191.
- Stratiev, D., Shishkova, I., Tankov, I., Pavlova, A., 2019. Challenges in characterization of residual oils. A review. *Journal of Petroleum Science and Engineering* 178, 227–250.
- Wu, Y., Zhang, N., 2010. Molecular Characterization of Gasoline and Diesel Streams. *Ind. Eng. Chem. Res.* 49, 12773–12782.

# A Global Optimization Algorithm for the Solution of Mixed-Integer Quadratic Adjustable Robust Optimization Problems under Endogenous Uncertainty

Byungjun Lee<sup>a</sup>, Styliani Avraamidou<sup>a\*</sup>

*<sup>a</sup>Department of Chemical and Biological Engineering, University of Wisconsin-Madison, 1415 Engineering Dr, Madison, WI 53706, USA*

## Abstract

This work proposes an algorithm for the solution of the two-stage mixed-integer quadratic adjustable robust optimization (ARO) problem under endogenous uncertainty. This class of ARO problems is often used in Chemical Engineering application such as scheduling and control problems, but there are only a handful of approaches attempting to solve them. For our developed approach, the two-stage ARO problem is firstly reformulated as a tri-level mixed-integer quadratic programming problem, and the exact and global solution is obtained using multi-parametric programming. The proposed algorithm was verified and compared with other approaches such as the affine decision rules, the column-and-constraint generation algorithm, and full enumeration of extreme points. We showed that our algorithm has the ability to solve ARO problems that, to our knowledge, cannot be solved by any other algorithm. To assess the efficiency and the performance of the proposed algorithm, a computational study was performed.

**Keywords:** Adjustable robust optimization (ARO), Endogenous uncertainty, Multi-parametric programming, Tri-level programming

## 1. Introduction

Adjustable robust optimization (ARO) is an extension of robust optimization (RO), which includes recourse decisions in the second stage (Ben-Tal et al., 2009). In Chemical Engineering, the ARO problem formulation can be applied to solve a vast array of multi-stage decision making problems under uncertainty, such as batch scheduling problems (Shi and You, 2016) and explicit model predictive control (Tejeda-Iglesias et al., 2019). Although several methods have been proposed for the solution of this class of problems, including affine decision rules (Ben-Tal et al., 2004) and column-and-constraint generation algorithm (Zeng and Zhao, 2013), most solution methods are limited to the solution of continuous linear ARO problems, and solving the mixed-integer quadratic ARO problems still remains a challenge.

In this paper, we apply a multi-parametric programming approach to solve the two-stage mixed-integer quadratic ARO problems under endogenous uncertainty, as an extension of the previous work on the linear ARO problems (Avraamidou and Pistikopoulos, 2020). Section 2 presents the general formulation of the ARO problems and presents our solution algorithm. Then the proposed algorithm is verified with several numerical examples, and its advantages over the previously known algorithms are discussed in section 3. Lastly, section 4 presents the relation between the computational time and the total number of variables.

## 2. ARO through multi-parametric programming

The two-stage mixed-integer ARO problem with a convex quadratic objective function can be expressed as (1). The first-stage variable ( $x$ ) is called “here-and-now” decision, which is determined prior to the realization of uncertainty ( $u$ ). The second-stage variable ( $y$ ) is called “wait-and-see” decision, which is determined after the uncertainty is revealed.

$$\begin{aligned} & \min_x (Q_x x + c_x)^T x + \max_{u \in \Omega_u} \min_y (Q_y y + c_y)^T y \\ & \text{s.t. } A_x x + A_y y + A_u u \leq b \\ & \quad x \in X := \mathbb{R}^{p_c} \times \{0, 1\}^{p_b}, \quad y \in Y := \mathbb{R}^{q_c} \times \{0, 1\}^{q_b}, \quad u \in U := \mathbb{R}^{r_0}. \end{aligned} \quad (1)$$

In this work, uncertainty is assumed as a function of “here-and-now” decisions. This type of uncertainty is referred to as endogenous uncertainty, where “here-and-now” decisions constrain the feasible region of uncertainty (Lappas and Gounaris, 2018). At the initial step, problem (1) is reformulated as a tri-level mixed-integer quadratic programming (T-MIQP) problem (2).

$$\begin{aligned} & \min_x z^{(1)} = (Q_x x + c_x)^T x + (Q_y y + c_y)^T y \quad (1^{\text{st}} \text{ level}) \\ & \text{s.t. } \min_{u \in \Omega_u} z^{(2)} = -(Q_y y + c_y)^T y \quad (2^{\text{nd}} \text{ level}) \\ & \quad \text{s.t. } \min_y z^{(3)} = (Q_y y + c_y)^T y \quad (3^{\text{rd}} \text{ level}) \\ & \quad \text{s.t. } A_x x + A_y y + A_u u \leq b \\ & \quad \quad x \in X := \mathbb{R}^{p_c} \times \{0, 1\}^{p_b}, \quad y \in Y := \mathbb{R}^{q_c} \times \{0, 1\}^{q_b}, \quad u \in U := \mathbb{R}^{r_0}. \end{aligned} \quad (2)$$

The tri-level problem (2) is then solved using multi-parametric approach based on the generalized piecewise affine decision rules (Avraamidou and Pistikopoulos, 2019). First, the 3<sup>rd</sup> level MIQP problem is solved to obtain the parametric solution  $y(x, u)$  for each critical region. Similarly, resulting set of 2<sup>nd</sup> level MIQP problems are solved to find the parametric solutions  $u(x)$ . Then resulting set of single-level MIQP problems are solved for  $x$ , and the exact and global solution is selected after applying comparison procedure. The comparison procedure removes unfeasible critical regions and finds the optimum solution that minimizes the objective function  $z^{(1)}$  among the remaining critical regions. The proposed algorithm is also illustrated in Figure 1.

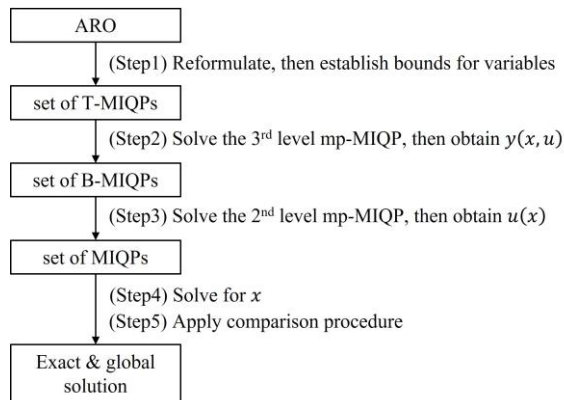


Figure 1. Algorithm to solve two-stage quadratic MI-ARO problem.

### 3. Numerical Examples

The first numerical example (3) has a convex, quadratic objective function in the first stage, which is a function of “here-and-now” decisions.

$$\begin{aligned}
 & \min_x 2x_1^2 + x_1x_2 + 3x_2^2 + \max_{u \in \Omega_u} \min_y 12y_1 + 18y_2 \\
 \text{s. t. } & x_1 + x_2 \leq 10 \\
 & y_1 \geq u_1 - x_1 \\
 & y_2 \geq u_2 - x_2 \\
 & x_1, x_2 \geq 0, y_1, y_2 \geq 0 \\
 & \Omega_u = \{u: 12.1 \leq u_1 \leq 26.3, 13.2 \leq u_2 \leq 30.4\}
 \end{aligned} \tag{3}$$

Problem (4) represents the tri-level formulation that is recast from the original two-stage ARO formulation (3).

$$\begin{aligned}
 & \min_x z^{(1)} = 2x_1^2 + x_1x_2 + 3x_2^2 + 12y_1 + 18y_2 \\
 \text{s. t. } & \min_{u \in \Omega_u} z^{(2)} = -(12y_1 + 18y_2) \\
 & \text{s. t. } \min_y z^{(3)} = 12y_1 + 18y_2 \\
 & \text{s. t. } x_1 + x_2 \leq 10 \\
 & y_1 \geq u_1 - x_1 \\
 & y_2 \geq u_2 - x_2 \\
 & x_1, x_2 \geq 0, y_1, y_2 \geq 0 \\
 & \Omega_u = \{u: 12.1 \leq u_1 \leq 26.3, 13.2 \leq u_2 \leq 30.4\}
 \end{aligned} \tag{4}$$

The proposed approach was verified by comparing it with both the affine decision rules and column-and-constraint generation algorithm. The objective function value and the corresponding optimum solution for each approach are summarized in Table 1.

Table 1. Objective function value and the corresponding optimum solution through multiparametric programming (MP), affine decision rules (AD), and column-and-constraint generation (CC).

Algorithm	$z^*$	$x_1^*$	$x_2^*$	$y_1^*$	$y_2^*$	$u_1^*$	$u_2^*$
MP	825.23	2.35	2.61	23.95	27.79	26.30	30.40
AD	825.23	2.35	2.61	23.95	27.79	26.30	30.40
CC	825.23	2.35	2.61	23.95	27.79	26.30	30.40

The second numerical example (5) has a convex, quadratic objective function in the second stage, which is a function of “wait-and-see” decisions.

$$\begin{aligned}
 & \min_x 3x_1 + 2x_2 + 1500x_3 + 6000x_4 \\
 & \quad + \max_{u \in \Omega_u} \max_y (6y_1^2 + 5y_2^2 + 9y_3^2 + 4y_1y_2 + 2y_1y_3 + 6y_2y_3) \\
 \text{s. t. } & x_1 + 2x_2 + 3x_3 + 4x_4 \leq 30 \\
 & y_1 \geq u_1 - x_1 - x_2 \\
 & y_2 \geq u_2 - 8x_3
 \end{aligned}$$

$$\begin{aligned}
 &y_3 \geq u_3 - 15x_4 \\
 &x_1, x_2 \geq 0, x_3, x_4 \in \{0, 1\}, y_1, y_2, y_3 \geq 0 \\
 &\Omega_u = \{u: 10 \leq u_1, u_2, u_3 \leq 20\}
 \end{aligned} \tag{5}$$

In this case, the affine decision rules are not applicable due to the nonlinear terms of uncertainties in the second stage, which are known to be computationally intractable in general. For the column-and-constraint generation algorithm, the sub-problem is non-convex, after using KKT conditions to transform the bi-level sub-problem into a single-level problem. Therefore, the multi-parametric approach has an advantage over the other two algorithms when the ARO problem includes the quadratic terms of “wait-and-see” decisions.

After the tri-level reformulation of the problem (5), POP toolbox (Avraamidou and Pistikopoulos, 2019) was used to obtain the third-level parametric solutions  $y(x, u)$  and second-level parametric solutions  $u(x)$ . The decision rules for each critical region in the 3<sup>rd</sup> level and 2<sup>nd</sup> level are summarized in the Table 2.

Table 2. Decision rules and critical region definition for the 3<sup>rd</sup> level and 2<sup>nd</sup> level.

Level	Critical Region	Decision Rules	Additional Critical Region Definition
3	CR <sub>1</sub>	$y_1 = -x_1 - x_2 + u_1,$ $y_2 = -8x_3 + u_3$ $y_3 = -15x_4 + u_4$	$-50 \leq x_1 + x_2 - u_1 \leq 0,$ $-50 \leq 15x_4 - u_3 \leq 0,$ $6x_1 + 6x_2 + 16x_3 + 15x_4 - 6u_1 - 2u_2 - u_3 \leq 0,$ $2x_1 + 2x_2 + 40x_3 + 45x_4 - 2u_1 - 5u_2 - 3u_3 \leq 0,$ $x_1 + x_2 + 24x_3 + 135x_4 - u_1 - 3u_2 - 9u_3 \leq 0$
	CR <sub>2</sub>	$y_1 = -x_1 - x_2 + u_1,$ $y_2 = -8x_3 + u_3$ $y_3 = 0$	$-50 \leq x_1 + x_2 - u_1 \leq 0,$ $15x_4 - u_3 \leq 0,$ $3x_1 + 3x_2 + 8x_3 - 3u_1 - u_2 \leq 0,$ $2x_1 + 2x_2 + 40x_3 - 2u_1 - 5u_2 \leq 0,$ $x_1 + x_2 + 24x_3 - u_1 - 3u_2 \leq 0$
	CR <sub>3</sub>	$y_1 = 0,$ $y_2 = -8x_3 + u_3$ $y_3 = -15x_4 + u_4$	$x_1 + x_2 - u_1 \leq 0,$ $-50 \leq 15x_4 - u_3 \leq 0,$ $16x_3 + 15x_4 - 2u_2 - u_3 \leq 0,$ $40x_3 + 45x_4 - 5u_2 - 3u_3 \leq 0,$ $8x_3 + 45x_4 - u_2 - 3u_3 \leq 0$
	CR <sub>4</sub>	$y_1 = 0,$ $y_2 = -8x_3 + u_3$ $y_3 = 0$	$x_1 + x_2 - u_1 \leq 0,$ $15x_4 - u_3 \leq 0,$ $8x_3 - u_2 \leq 0$
	Common	—	$x_1 + 2x_2 + 3x_3 + 4x_4 \leq 30,$ $-8x_3 + u_2 \leq 52,$ $x_1, x_2 \geq 0, x_3, x_4 \in \{0, 1\},$ $10 \leq u_1, u_2, u_3 \leq 20$
2	CR <sub>1,1</sub>	$u_1 = 20,$ $u_2 = 20,$ $u_3 = 20$	$x_1 + x_2 \leq 20,$ $3x_4 \leq 4$ $6x_1 + 6x_2 + 16x_3 + 15x_4 \leq 180,$ $2x_1 + 2x_2 + 40x_3 + 45x_4 \leq 200,$ $x_1 + x_2 + 24x_3 + 135x_4 \leq 260$
	CR <sub>2,1</sub>	$u_1 = 20,$ $u_2 = 20,$ $u_3 = 0$	$x_1 + x_2 \leq 20,$ $3x_1 + 3x_2 + 8x_3 \leq 80,$ $x_1 + x_2 + 20x_3 \leq 70$
	CR <sub>3,1</sub>	$u_1 = 0,$ $u_2 = 20,$ $u_3 = 20$	$x_1 + x_2 \geq 0,$ $3x_4 \leq 4$ $4x_3 + 5x_4 \leq 16,$ $8x_3 + 45x_4 \leq 80$
	CR <sub>4,1</sub>	$u_1 = 0,$ $u_2 = 20,$ $u_3 = 0$	$x_1 + x_2 \geq 0,$ $2x_3 \leq 5$
	Common	—	$x_1 + 2x_2 + 3x_3 + 4x_4 \leq 30,$ $x_1, x_2 \geq 0, x_3, x_4 \in \{0, 1\}$

Lastly, CPLEX was used to solve the resulting single-level mixed-integer quadratic programming problem. Table 3 summarizes the 1<sup>st</sup> level objective function values for all the critical regions, and only the first critical regions  $CR_1$  and  $CR_{1,1}$  were selected from the comparison procedure. Therefore, the optimum solution is  $x = [13, 7, 1, 0]^T$  and  $y = [0, 12, 20]^T$  with objective function value 7313. This value exactly matches with the optimum objective function value obtained from the full enumeration of extreme points.

Table 3. Objective function values and optimum solutions in the 1<sup>st</sup> level.

$CR_i$	$CR_j$	$z^*$	$x_1$	$x_2$	$x_3$	$x_4$	$y_1$	$y_2$	$y_3$	$u_1$	$u_2$	$u_3$
1	1	7313	13	7	1	0	0	12	20	20	20	20
2	1	2050	10	10	0	0	0	20	0	20	20	0
3	1	7260	0	0	1	0	0	12	20	0	20	20
4	1	2000	0	0	0	0	0	20	0	0	20	0

#### 4. Computational Implementation

To evaluate the performance of the proposed algorithm, computational time of randomly generated mixed-integer ARO problems was measured. The ratio of the number of variables for continuous  $x$ , binary  $x$ , continuous  $y$ , binary  $y$  was set to be 1 : 1 : 1 : 1, and the number of constraints is set to be equal to the total number of variables. Note that the solution time was measured using Intel Core i7 at 80GHz and 16GB RAM, MATLAB R2022b, and IBM ILOG CPLEX Studio 12.10.0.

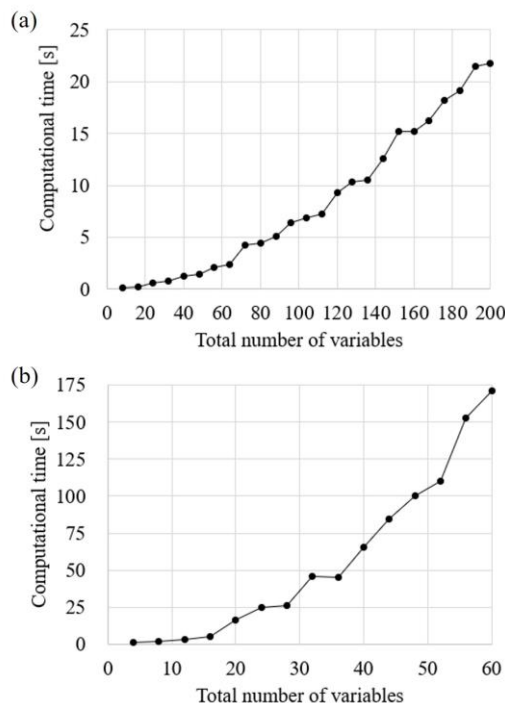


Figure 2. Measured computational time versus total number of variables for the (a) MILP-MILP structure and (b) MIQP-MIQP structure of the ARO problem.

Figure 2 represents the relation between the average computational time over five experiments and the total number of variables for the MILP-MILP structure and MIQP-MIQP structure of the ARO problem. Figure 2 also shows the exponentially increasing trend in the computational time as the total number of variables increases. In addition, it takes about 175s to solve the MIQP-MIQP structure of the ARO problem with 60 variables, about 70 times longer computational time compared to the MILP-MILP structure.

## 5. Conclusion

In this work, we developed an algorithm for the solution of the two-stage mixed-integer quadratic ARO problem under endogenous uncertainty. Initially, the ARO problem is reformulated into the tri-level mixed-integer quadratic programming problem, and it is solved multi-parametrically using the general piecewise affine decision rules. By solving the problem from the lowest level (3<sup>rd</sup> level) to the highest level (1<sup>st</sup> level), the exact and global solution is obtained after applying the comparison procedure. For the ARO problem with a quadratic objective function in the first stage, the proposed algorithm was verified by comparing its global solution with the global solutions from affine decision rules and column-and-constraint generation algorithm. In addition, we showed that these two well-known algorithms cannot solve the ARO problem with a quadratic objective function in the second stage, while our algorithm is able to solve this type of problem. Lastly, we analyzed the performance of the proposed algorithm by measuring the computational time for the randomly generated ARO problems with MILP-MILP structure and MIQP-MIQP structure.

## References

- S. Avraamidou, E. N. Pistikopoulos, 2019, Multi-parametric global optimization approach for tri-level mixed-integer linear optimization problems, *Journal of Global Optimization*, 74(3), 443-465.
- S. Avraamidou, E. N. Pistikopoulos, 2020, Adjustable robust optimization through multi-parametric programming, *Optimization Letters*, 14(4), 873-887.
- A. Ben-Tal, L. El Ghaoui, A. Nemirovski, 2009, *Robust optimization*, Princeton university press.
- A. Ben-Tal, A. Goryashko, E. Guslitzer, A. Nemirovski, 2004, Adjustable robust solutions of uncertain linear programs, *Mathematical programming*, 99(2), 351-376.
- N. H. Lappas, C. E. Gounaris, 2018, Robust optimization for decision-making under endogenous uncertainty, *Computers & Chemical Engineering*, 111, 252-266.
- H. Shi, F. You, 2016, Adjustable robust optimization for scheduling of batch processes under uncertainty, *Computer Aided Chemical Engineering*, Vol. 38, 547-552.
- M. Tejada-Iglesias, N. H. Lappas, C. E. Gounaris, L. Ricardez-Sandoval, 2019, Explicit model predictive controller under uncertainty: An adjustable robust optimization approach, *Journal of Process Control*, 84, 115-132.
- B. Zeng, L. Zhao, 2013, Solving two-stage robust optimization problems using a column-and-constraint generation method, *Operations Research Letters*, 41(5), 457-461.

# Cross-domain Fault Diagnosis for Chemical Processes through Dynamic Adversarial Adaptation Network

Ruoshi Qin,<sup>a</sup> Jinsong Zhao,<sup>a,b</sup>

<sup>a</sup>*State Key Laboratory of Chemical Engineering, Department of Chemical Engineering, Tsinghua University, Beijing 100084, China*

<sup>b</sup>*Beijing Key Laboratory of Industrial Big Data System and Application, Tsinghua University, Beijing 100084, China*

## Abstract

During the past decade, deep learning models have shown great potential in chemical process fault diagnosis. However, a major assumption accepted by default in these previous studies is that the training and testing data possess the same feature distribution, which is generally invalid in industrial applications. The ability of data-driven models to generalize on various operation conditions are limited due to their tendency to overfit the training dataset. Motivated by the development of transfer learning, domain adaptation is utilized in this paper to settle the problem of knowledge transfer under the distribution discrepancy between the training and test datasets. The proposed model Dynamic Adversarial Adaptation Network (DAAN) can dynamically and quantitatively evaluate the importance of these two distributions and achieve accurate classification performance. Empirical evaluations based on Tennessee Eastman process validate the applicability and practicability of DAAN, which achieves the state-of-the-art transfer results in terms of diverse operating conditions and fault types.

**Keywords:** Domain adaptation; Fault diagnosis; Transfer learning; Tennessee Eastman process

## 1. Introduction

Fault diagnosis plays a significant role in chemical processes to avoid severe accidents. In recent years, significant advances have been made towards fault diagnosis with the assistance of deep learning. However, the excellent performance of current intelligent fault diagnosis methods based on deep learning is mainly attributed to the availability of large amounts of labeled failure data which is laborious and heavily dependent on experience. The entire equipment usually runs in a normal state and accessible fault samples are scarce as well (Qin et al., 2022). Another limitation on the practical application of these fault diagnosis methods is the continual changes in operating conditions. The distribution discrepancy between source training data and target test data makes a high requirement for the transferability of the diagnosis model (Wu et al., 2020).

In industrial production, the variation of ambient temperature and raw material components contribute to diverse working modes. When the fault samples in the test dataset are lack of labels and the training and testing data are subject to different distributions, the performance of most intelligent diagnostic models deteriorates considerably (Qin et al., 2022). Transfer learning puts forward an emerging frame to address the above problems. The concept of domain in transfer learning represents the



corresponding data space and the data distribution. In this study, the source domain refers to the data space and distribution of the labeled training data, while the attribute of the unlabeled test data is known as the target domain. The challenge of transferring the learned knowledge from the source domain to the target domain is defined as the domain shift phenomenon. This problem makes the source domain classifier lose its generalization ability in the target domain. Domain adaptation is a preferred solution to the domain shift issue, aiming to develop a cross-domain classifier which can extract domain-invariant features to bridge the distribution discrepancy (Li et al., 2019). Fig.1 illustrates the principle of domain shift and domain adaptation.

The latest studies reveal that deep neural networks can learn transferable features sufficiently with unsupervised domain adaptation algorithms, such as maximum mean discrepancy (MMD), multiple kernel MMD and joint MMD (Long et al., 2017). Some innovative works raise deep transfer learning methods in the field of computer vision, including Domain Adaptive Neural Network (DANN) (Ganin et al., 2017) and Domain Adaptive Ensemble Learning (DAEL) (Zhou et al., 2021), but it is still rare to discover a powerful cross-domain fault diagnosis method for chemical processes.

In this paper, a novel adversarial domain adaptation transfer learning model called Dynamic Adversarial Adaptation Network (DAAN) is proposed for fault diagnosis in industrial chemical processes. The diagnostic framework is composed of four modules, a shared feature extractor for feature learning, a label classifier for category prediction, a global domain discriminator for marginal distribution alignment and local subdomain discriminators for conditional distribution alignment. Deep residual shrinkage network is designed as the backbone of the feature extractor to achieve multiscale characteristics from time-series data. The adversarial learning mode with an adaptive weighting strategy can dynamically and quantitatively evaluate the relative importance of marginal and conditional distribution. The alignment of the source and target feature space is finally realized through the overall learning proceed.

The rest of the paper is organized as follows. In Section 2, the module structures and the diagnosis framework of DAAN are introduced in detail. Section 3 elaborates the experiments on Tennessee Eastman process and compares the results among different popular transfer learning models. The conclusion and outlook of the proposed model are summarized in Section 4.

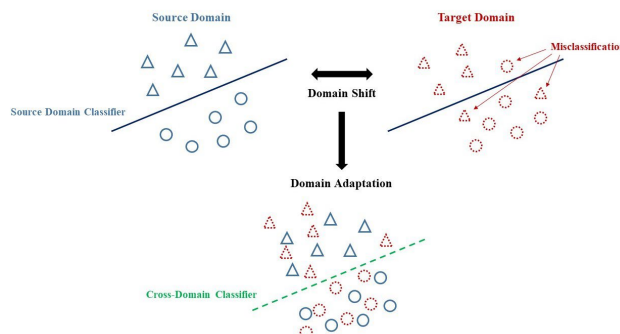


Fig.1. Illustration of domain shift and domain adaptation.

## 2. Dynamic Adversarial Adaptation Network

The proposed model Dynamic Adversarial Adaptation Network is displayed in Fig.2. Its skeleton consists of a feature extractor  $G_f$ , a label classifier  $G_y$ , a global domain discriminator  $G_d$  and a local subdomain discriminator  $G_d^c$ . In DAAN,  $G_f$  extracts the key features of data followed by  $G_d$  and  $G_d^c$  achieving the adaptation of marginal and conditional distributions respectively. The dynamic adversarial factor  $\omega$  performs a quantitative evaluation of these two distributions. Under the action of  $G_y$ , the DAAN model is trained efficiently with the Gradient Reversal Layer (GRL) (Yu et al., 2019).

### 2.1. Feature Extractor

The raw data from different domains are directly fed into a shared feature extractor for certain generic characteristics. Deep residual shrinkage network is employed as the extractor which can obtain powerful feature learning capability and eliminate the noise-related information. The feature extractor includes multiple residual learning blocks and each of them is composed of convolution operation, batch normalization, activation function, and short connection. The structure of residual shrinkage building unit is a variant of classical ResNet, and a range of units are stacked to form the deep network. Sigmoid and rectified linear unit (ReLU) as non-linear transform tools are adopted to promote model convergence and alleviate the problem of gradient vanishing.

### 2.2. Domain Adaptation Network

#### 2.2.1. Label Classifier

The label classifier  $G_y$  focuses on the discrimination of the input sample label from the source domain. With the effort of the supervised information and the extracted deep features, the classifier is trained based on the cross-entropy loss formulated as:

$$L_y = -\frac{1}{n_s} \sum_{\mathbf{x}_i \in \mathcal{D}_s} \sum_{c=1}^C P_{\mathbf{x}_i \rightarrow c} \log G_y(G_f(\mathbf{x}_i)) \quad (1)$$

where  $C$  is the number of classes,  $P_{\mathbf{x}_i \rightarrow c}$  is the probability of  $\mathbf{x}_i$  belonging to class  $c$ .

#### 2.2.2. Global Domain Discriminator

The global domain discriminator  $G_d$  takes the responsibility of align the marginal distributions between the source and target domains. Its loss objective function can be calculated in the formula:

$$L_g = \frac{1}{n_s + n_t} \sum_{\mathbf{x}_i \in (\mathcal{D}_s \cup \mathcal{D}_t)} L_d(G_d(G_f(\mathbf{x}_i)), d_i) \quad (2)$$

where  $L_d$  is the domain discriminator loss and  $d_i$  is the original label of the input sample.

#### 2.2.3. Local Subdomain Discriminator

The local subdomain discriminator  $G_d^c$  is designed to align the conditional distributions between the source and target domains. Its representation ability of merging multimode structure in two distributions makes the discriminator carry out more fine-grained domain adaptation. The local subdomain discriminator can be regarded class-wise to match the multi-domain data associated with the same class, and the global domain discriminator is able to be split into multi-class local subdomain discriminators. The output of the label predictor  $G_y(\mathbf{x}_i)$  indicates the matching degree of each data point  $\mathbf{x}_i$  to be categorized further. Therefore, the loss function of  $G_d^c$  can be described as:

$$L_l = \frac{1}{n_s + n_t} \sum_{c=1}^C \sum_{\mathbf{x}_i \in (\mathcal{D}_s \cup \mathcal{D}_t)} L_d^c(G_d^c(\hat{y}_i^c G_f(\mathbf{x}_i)), d_i) \quad (3)$$

where  $L_d^c$  represents the cross-entropy loss of the subdomain discriminator associated with class  $c$ .  $\hat{y}_i^c$  signifies the predicted probability distribution over the class  $c$  of the input sample  $\mathbf{x}_i$ .

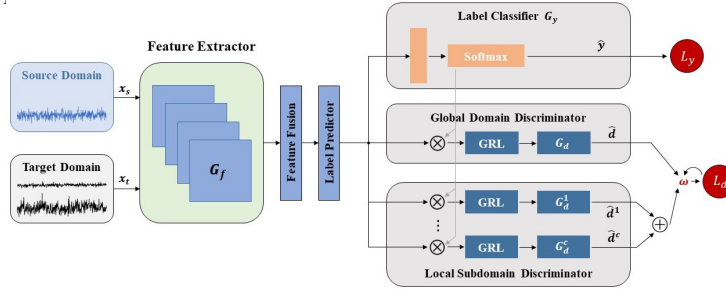


Fig.2. Architecture of Dynamic Adversarial Adaptation Network.

### 2.3. Dynamic Adversarial Training

Domain adversarial adaptation borrows the principle of Generative Adversarial Networks (GAN) to assist the model to learn transferable features. The adversarial learning procedure is a two-player game accompanied by the domain discriminator  $G_d$  and the feature extractor  $G_f$ . The former aims to distinguish the source domain from the target domain, while the latter keeps making puzzles for the former by extracting domain-invariant features. During this course, the parameters of  $G_f$  are optimized by maximizing the loss of  $G_d$ , and the ones of  $G_d$  are trained by minimizing the loss. With the loss of the label classifier  $G_y$  minimized, the loss function can be formalized as:

$$L(\theta_f, \theta_y, \theta_d) = \frac{1}{n_s} \sum_{x_i \in \mathcal{D}_s} L_y(G_y(G_f(x_i)), y_i) - \frac{\lambda}{n_s + n_t} \sum_{x_i \in (\mathcal{D}_s \cup \mathcal{D}_t)} L_d(G_d(G_f(x_i)), d_i) \quad (4)$$

where  $\lambda$  is a trade-off parameter. As the training converges, the parameters  $\hat{\theta}_f$ ,  $\hat{\theta}_y$ ,  $\hat{\theta}_d$  come to the optimized state.

$$(\hat{\theta}_f, \hat{\theta}_y) = \arg \min_{\theta_f, \theta_y} L(\theta_f, \theta_y, \theta_d), (\hat{\theta}_d) = \arg \max_{\theta_d} L(\theta_f, \theta_y, \theta_d) \quad (5)$$

The dynamic adversarial factor  $\omega$  is creatively built to evaluate the relative importance of the marginal and conditional distributions. In DAAN, the global  $\mathcal{A}$ -distance of the global domain discriminator and the local  $\mathcal{A}$ -distance of the local subdomain discriminator can be denoted as:

$$d_{\mathcal{A},g}(\mathcal{D}_s, \mathcal{D}_t) = 2(1 - 2(L_g)), d_{\mathcal{A},l}(\mathcal{D}_s^c, \mathcal{D}_t^c) = 2(1 - 2(L_l^c)) \quad (6)$$

where  $\mathcal{D}_s^c$  and  $\mathcal{D}_t^c$  represent samples from class  $c$ . Eventually, the dynamic adversarial factor  $\omega$  is estimated as:

$$\hat{\omega} = \frac{d_{\mathcal{A},g}(\mathcal{D}_s, \mathcal{D}_t)}{d_{\mathcal{A},g}(\mathcal{D}_s, \mathcal{D}_t) + \frac{1}{C} \sum_{c=1}^C d_{\mathcal{A},l}(\mathcal{D}_s^c, \mathcal{D}_t^c)} \quad (7)$$

$\omega$  starts to be initialized as 1 in the first epoch. The pseudo labels of the target domain can be obtained after each iteration, so the local distance for class  $c$  can be computed with cross entropy. Therefore, the learning objective of DAAN is concluded as follows:

$$L(\theta_f, \theta_y, \theta_d, \theta_d^c |_{c=1}^C) = L_y - \lambda((1 - \omega)L_g + \omega L_l) \quad (8)$$

### 2.4. DAAN-based fault diagnosis framework

The framework of the proposed DAAN-based fault diagnosis method is shown in Fig.3. The dual-stage procedures include offline training and online application. Historical data

collected from the chemical process are utilized for model design. Benefitting from the transferred knowledge, the diagnosis network monitors the online process and outputs the fault classification result.

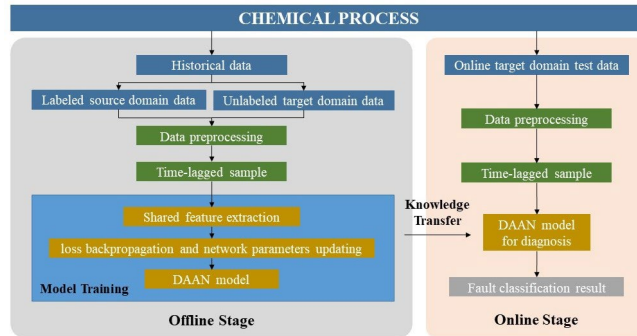


Fig.3. DAAN-based fault diagnosis framework.

### 3. Case Studies

In this section, the benchmark Tennessee Eastman process (TEP) is applied to evaluate the performance of the proposed model. TEP is widely used in chemical process monitoring tasks. The unit operations of TEP contain a reactor, a condenser, a recycle compressor, a vapor-liquid separator and a stripper. In this paper, the experiments are based on the revised version (Bathelt et al., 2015) which can be downloaded at <http://depts.washington.edu/control/LARRY/TE/download.html>. 19 fault types and 5 steady-state operating modes are introduced in this research to explore the multimode fault diagnosis performance. The fault types and working conditions are selected in the same way as the previously published work, and the datasets are prepared followed by its experience (Wu et al., 2020).

When the labeled fault data in the source mode and only one-batch unlabeled fault data in the target mode are involved in the training dataset, the unsupervised anomaly diagnosis tasks are presented to verify the model effectiveness. Different common transfer learning models are adopted in comparison with the proposed DAAN model, such as traditional deep transfer networks composed of stacked auto-encoder (DTN), domain adaptation networks along with CNN (DAN-CNN) and joint adaptation networks along with CNN (JAN-CNN). The fault identification performances are exhibited in Table 1. The average accuracy rates are calculated and summarized in Table 2. The classification results show the efficiency and robustness of DAAN.

Table.1. Unsupervised fault diagnosis performance comparison.

(a) DTN					
	$\mathcal{D}_t(0)$	$\mathcal{D}_t(1)$	$\mathcal{D}_t(2)$	$\mathcal{D}_t(4)$	$\mathcal{D}_t(5)$
$\mathcal{D}_s(0)$	-	91.4%	92.0%	82.3%	91.2%
$\mathcal{D}_s(1)$	92.9%	-	93.6%	85.3%	89.4%
$\mathcal{D}_s(2)$	92.8%	92.5%	-	84.1%	93.1%
$\mathcal{D}_s(4)$	80.6%	83.5%	82.9%	-	79.2%
$\mathcal{D}_s(5)$	89.4%	90.7%	92.9%	77.8%	-

(b) DAN-CNN					
	$\mathcal{D}_t(0)$	$\mathcal{D}_t(1)$	$\mathcal{D}_t(2)$	$\mathcal{D}_t(4)$	$\mathcal{D}_t(5)$
$\mathcal{D}_s(0)$	-	93.6%	93.1%	91.6%	91.8%

$\mathcal{D}_s(1)$	94.9%	-	95.4%	91.5%	95.6%
$\mathcal{D}_s(2)$	93.2%	93.3%	-	89.0%	<b>97.4%</b>
$\mathcal{D}_s(4)$	87.1%	84.4%	84.7%	-	82.2%
$\mathcal{D}_s(5)$	90.2%	92.6%	<b>97.4%</b>	86.3%	-

(c) JAN-CNN

	$\mathcal{D}_t(0)$	$\mathcal{D}_t(1)$	$\mathcal{D}_t(2)$	$\mathcal{D}_t(4)$	$\mathcal{D}_t(5)$
$\mathcal{D}_s(0)$	-	96.0%	96.4%	91.1%	<b>97.6%</b>
$\mathcal{D}_s(1)$	95.6%	-	<b>95.9%</b>	90.8%	96.4%
$\mathcal{D}_s(2)$	93.5%	95.6%	-	89.6%	97.3%
$\mathcal{D}_s(4)$	<b>88.8%</b>	89.7%	88.2%	-	90.6%
$\mathcal{D}_s(5)$	93.9%	95.8%	97.3%	86.9%	-

(d) DAAN

	$\mathcal{D}_t(0)$	$\mathcal{D}_t(1)$	$\mathcal{D}_t(2)$	$\mathcal{D}_t(4)$	$\mathcal{D}_t(5)$
$\mathcal{D}_s(0)$	-	<b>97.2%</b>	<b>96.8%</b>	<b>92.5%</b>	97.3%
$\mathcal{D}_s(1)$	<b>96.6%</b>	-	<b>95.9%</b>	<b>93.9%</b>	<b>96.7%</b>
$\mathcal{D}_s(2)$	<b>95.5%</b>	<b>96.1%</b>	-	<b>92.0%</b>	97.2%
$\mathcal{D}_s(4)$	87.2%	<b>91.7%</b>	<b>90.4%</b>	-	<b>91.1%</b>
$\mathcal{D}_s(5)$	<b>94.7%</b>	<b>96.0%</b>	97.3%	<b>88.4%</b>	-

Table.2. Average accuracy rate comparison of diagnostic models.

	DTN	DAN-CNN	JAN-CNN	DAAN
Average accuracy rate	87.9%	91.3%	93.4%	<b>94.2%</b>

#### 4. Conclusions

In the present work, Dynamic Adversarial Adaptation Network is designed for fault diagnosis application in the chemical process. This adversarial domain adaptation model mainly consists of three components, label classifier, global domain discriminator and local subdomain discriminator with a powerful feature extractor attached to the front. Comprehensive results demonstrate that the model can obtain superior diagnostic performance compared with other transfer learning models. Further work is being conducted to improve this model, such as further alleviating the need for labeled data in the source domain and tapping its potential for industrial process tasks.

#### References

- R. Qin, J. Zhao, 2022, High-Efficiency Generative Adversarial Network Model for Chemical Process Fault Diagnosis, *IFAC-PapersOnLine*, 55, 7, 732-737.
- H. Wu, J. Zhao, 2020, Fault detection and diagnosis based on transfer learning for multimode chemical processes, *Computers and Chemical Engineering*, 135, 106731.
- R. Qin, J. Zhao, 2022, Adaptive multiscale convolutional neural network model for chemical process fault diagnosis, *Chinese Journal of Chemical Engineering*, 50, 398-411.
- X. Li, W. Zhang, Q. Ding, 2019, Cross-Domain Fault Diagnosis of Rolling Element Bearings Using Deep Generative Neural Networks, *IEEE Transactions on Industrial Electronics*, 66, 7, 5525-5534.
- M. Long, H. Zhu, J. Wang, M.I. Jordan, 2017, Deep transfer learning with joint adaptation networks, *ICML'17: Proceedings of the 34th International Conference on Machine Learning*, 70, 2208-2217.
- Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, V. Lempitsky, 2017, Domain-Adversarial Training of Neural Networks, *Domain Adaptation in Computer Vision Applications*, 189-209.
- K. Zhou, Y. Yang, Y. Qiao, T. Xiang, 2021, Domain Adaptive Ensemble Learning, *IEEE Transactions on Image Processing*, 30, 8008-8018.

- C. Yu, J. Wang, Y. Chen, M. Huang, 2019, Transfer Learning with Dynamic Adversarial Adaptation Network, *2019 IEEE International Conference on Data Mining (ICDM)*, 19303062.
- A. Bathelt, N.L. Ricker, M. Jelali, 2015, Revision of the Tennessee Eastman Process Model, *IFAC-PapersOnLine*, 48, 8, 309-314.



# Design and performance analysis of energy efficient hydrogen liquefaction process

Ahmad Naquash<sup>a</sup>, Moonyong Lee<sup>a,\*</sup>

<sup>a</sup> *School of Chemical Engineering, Yeungnam University, Gyeongsan 38541, Republic of Korea*

*\*mynlee@yu.ac.kr*

## Abstract

Hydrogen (H<sub>2</sub>) has gained vital importance in reducing greenhouse gas emissions. Due to its low energy density, the long-time storage and long-distance transportation of H<sub>2</sub> is a challenging issue. Liquefaction is a practical approach to enhance energy density; however, it is an energy and cost-intensive technique. In addition, variation in H<sub>2</sub> composition (ortho H<sub>2</sub>/para H<sub>2</sub>) with respect to temperature is exothermic, adding heat load to the refrigeration cycle. Most conceptual studies do not incorporate this heat load in specific energy consumption (SEC). This study presents two case scenarios, i.e., with heat load (adiabatic) and without heat load (isothermal), to produce liquid H<sub>2</sub> (LH<sub>2</sub>) through stepwise cooling with an ortho to para conversion. The proposed adiabatic and isothermal processes show SEC of 12.2 and 6.9 kWh/kg, respectively. The adiabatic process requires a large quantity of refrigerant flowrates, i.e., 46 kg/s, compared to the isothermal process (26 kg/s). In both cases, 100% saturated LH<sub>2</sub> is produced with 99.7 % para conversion. This study can help achieve a sustainable green energy economy by enhancing the overall H<sub>2</sub> value chain.

**Keywords:** Hydrogen liquefaction, Clean energy, Ortho-para conversion, Adiabatic process, Isothermal process.

## 1. Introduction

The greenhouse gas emissions have now become a prime focus of industrialists and environmentalists owing to their negative impact on the atmosphere. These emissions can be reduced by introducing and promoting clean energy. Hydrogen (H<sub>2</sub>) is extensively considered a clean and multi-purpose energy carrier that can be used in every sector of life, such as transportation, electricity generation, and industry [1]. H<sub>2</sub> importance cannot be emphasized more than the global H<sub>2</sub> demand is increased by approx. threefold since 1975 [2]. Due to this increasing demand and clean characteristics, H<sub>2</sub> may take the place of conventional fuels in the near future. For that purpose, intercontinental or intercountry transportation of H<sub>2</sub> is vital for the global energy mix. Alongside transportation, long-time H<sub>2</sub> storage is also of primary importance, especially in the case of excess H<sub>2</sub> or excess electricity as both H<sub>2</sub> and electricity are interconvertible.

Commercially, H<sub>2</sub> is liquefied in a three-stage process: pre-cooling (up to -193 °C), cooling (up to -243 °C), and liquefaction (up to -253 °C) [3]. Liquid nitrogen (N<sub>2</sub>) is used as a refrigerant in the pre-cooling phase, while a combination of JT valves or expanders and H<sub>2</sub> itself is used in the subsequent stages. In recent times, mixed refrigerants (MR) have been suggested as an alternative to pure liquid N<sub>2</sub>. Numerous studies have suggested different compositions of low-boiling hydrocarbons and N<sub>2</sub> as constituents of MR [4].



Similarly, helium (He) and neon (Ne) are also considered alongside  $H_2$  in the cooling and liquefaction stages. The main objective behind trying different refrigerants and process configurations is to reduce the specific energy consumption (SEC) and improve the liquefaction process's thermodynamic (exergy) efficiency [5]. Industrial processes have SEC in the range of 12–15 kWh/kg $_{LH_2}$  and efficiency as low as 20–30% [6]. In an ideal case, the SEC for hydrogen liquefaction is  $\sim 2.7$  kWh/kg $_{LH_2}$  [7]. This large gap between the ideal and the real process motivates the researchers to design conceptual processes wherein different MR compositions, and process configurations have been proposed. The sole objective of such processes is to bring down the SEC value as close to the ideal case as possible.

The additional challenge associated with the  $LH_2$  process is ortho to the para conversion of  $H_2$  isomers. At 25 °C,  $H_2$  constitutes 75% ortho and 25% para having opposite rotational movement. The energy level of the ortho state is higher than the para-state, which means that by decreasing temperature, ortho  $H_2$  tends to convert to para  $H_2$  to maintain equilibrium by releasing energy. However, this ortho-to-para conversion (OPC) is slow due to the low rate of reaction, i.e.,  $0.0114\text{ h}^{-1}$ , and exothermic in nature, which releases a large amount of energy high enough to evaporate  $LH_2$  [8]. Recently, Riaz et al. [9] have discussed the significance of OPC reaction and its heat of reaction on the energy consumption for the production of  $LH_2$ . In another study, Son et al. evaluated different OPC reactor configurations to mimic the actual process configurations [10]. In commercial plants, catalysts are embedded inside heat exchangers. In this way, the heat of the reaction is dissipated inside the heat exchanger. Similarly, liquid  $N_2$  and liquid  $H_2$  are also used as fluid media for isothermal operation.

The OPC reaction heat is critical to the SEC of the process. It is pertinent to discuss the issues and solutions adopted by the researchers from a process systems viewpoint. The literature survey shows that many integral studies have employed isothermal OPC reactors ignoring the reaction heat. Only a few studies, such as Riaz et al. [9] and Son et al. [10], have discussed the importance of this OPC reaction heat. In the present study, two process configurations are proposed considering the heat of the reaction and process simplicity. The proposed process employs two refrigeration cycles and two reactor configurations: isothermal and adiabatic. Both these configurations are evaluated based on energy and exergy perspectives to evaluate their maximum thermodynamic potential.

## 2. $H_2$ liquefaction process

$H_2$  liquefaction is an energy-intensive process mainly because of low boiling point of  $H_2$  and exothermic OPC reaction. The process description and simulation details are provided in following subsections.

### 2.1. Process description

The proposed process is divided into two cases based on the reactor configurations.

Case I: Isothermal process

Case II: Adiabatic process

The  $H_2$  feed stream ( $n\text{-GH}_2$ ) is fed to a cryogenic heat exchanger (CHX-1) for pre-cooling. In CHX-1, the temperature is decreased from 25°C to  $-193.2^\circ\text{C}$  (see Figure 1). After that,  $H_2$  is fed to the equilibrium reactor (R-1) to convert 25% p- $H_2$  to 43% p- $H_2$ . In case I, stream H4 is fed to CHX-2, where the temperature is further reduced to  $-252^\circ\text{C}$ . In case II, stream H4 temperature increases to  $-183.2^\circ\text{C}$  due to exothermic OPC reaction (see Figure 2). For further cooling stream, H4 is sent back to the first heat exchanger, CHX-1, where the temperature is reduced to  $-193.2^\circ\text{C}$ . Then stream H5 is fed to CHX-2 to cool it till  $-252^\circ\text{C}$ . The outlet stream is fed to another equilibrium reactor (R-2) to convert

43% p-H<sub>2</sub> into 99.7% p-H<sub>2</sub>. The outlet stream H8 is at the temperature of -252°C due to heat being recovered through energy stream (Q) with a heat flow of -394.6 kW. Finally, H<sub>2</sub> is liquefied at 21 bar. The pressure of stream H8 is reduced to the LH<sub>2</sub> storage pressure of 1.3 bar by passing it through the expander K-5. At the K-5 outlet, 100% saturated liquid H<sub>2</sub> is obtained.

Case I is shown in Figure 1. In case I, the precooling MR (PMR) consists of C<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub>, n-C<sub>4</sub>, i-C<sub>4</sub>, n-C<sub>5</sub>, and N<sub>2</sub>. The PMR is first compressed and cooled in a series of compressors and coolers. Then, the PMR is cooled further in CHX-1 and expanded in expander. The outlet stream of expander is used to reduce temperature of hot streams.

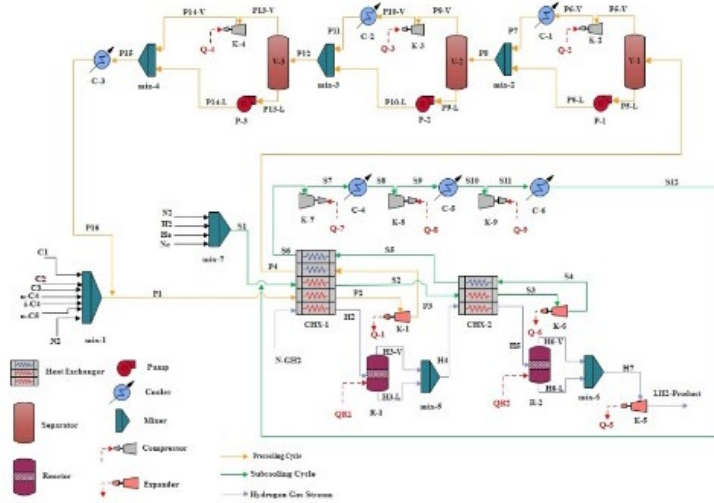


Figure 1 process flow diagram of case I

Case II is shown in Figure 2. In case II, the liquefaction MR (LMR) consists of H<sub>2</sub>, He, and Ne. The LMR is first compressed and cooled in a series of compressors and coolers. Then, the LMR is cooled further in CHX-1 and -2 and expanded in expander. The outlet stream of expander is used to reduce temperature of hot streams in CHX-1 and -2.

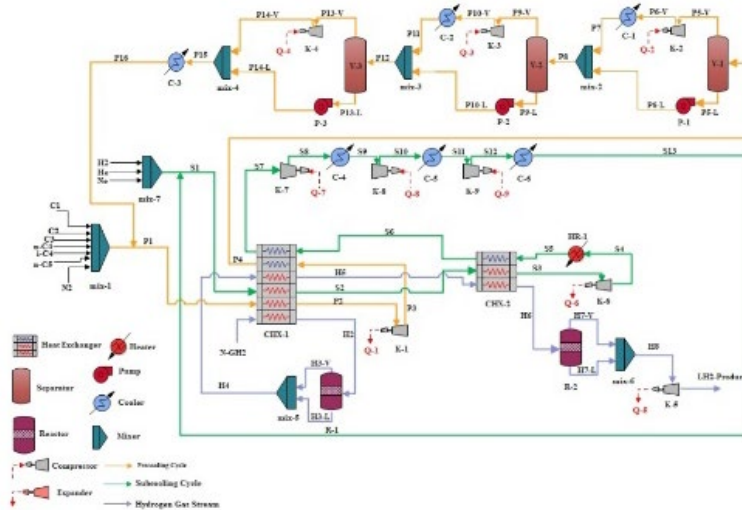


Figure 2 process flow diagram of case II

### 2.2. Process simulation

The proposed process was simulated in Aspen HYSYS v11 using the Peng–Robinson thermodynamic fluid package [11] for refrigerants and the Modified Benedict Webb Rubin equation of state for H<sub>2</sub>. The feed conditions and design parameters were selected based on the base case, as shown in Table 1. The efficiency of the compressor and expander was adjusted to 75% and 80%, respectively.

**Table 1** Design parameters for the H<sub>2</sub> liquefaction process [4]

Design parameters	Values
Feed temperature	25 °C
Feed pressure	21 bar
Feed flowrate	1 kg/s
Feed composition (mole fraction)	[oH <sub>2</sub> /pH <sub>2</sub> ] = [0.75/0.25]
Pressure-drop across aftercoolers	0 bar
Pressure-drop across CHXs	0 bar
Temperature aftercooler	22 °C
H <sub>2</sub> liquefaction rate	100 %
Liquid H <sub>2</sub> pressure	1.3 bar

### 2.3. Energy analysis

The energy analysis of this study is conducted in terms of SEC. The design variables (Table 2), which include refrigerant mass flow rates, suction pressure, and discharge pressure of refrigeration cycles, are varied to observe their effect on the SEC.

### 2.4. Exergy analysis

Exergy analysis is an excellent tool for analyzing process inefficiencies based on the second law of thermodynamics. Exergy is of mainly two types: physical and chemical, which is calculated as:

$$ex_{ph,i} = (h_i - h_o) - T_o(s_i - s_o) \quad (1)$$

$$ex_{ch,i} = \sum x_i e_i^{CH} + RT_o \sum x_i \ln x_i \quad (2)$$

Where  $ex_{ph,i}$  and  $ex_{ch,i}$  is denoted as physical and chemical exergy,  $x_i$  is denoted as the mole fraction of  $i^{\text{th}}$  component in a stream and  $e_i^{CH}$  is referred to as standard chemical exergy of  $i^{\text{th}}$  component taken from [12]. The physical exergy values are taken from Aspen Hysys<sup>®</sup> stream properties, whereas chemical exergy values are calculated using equation 2. The exergy analysis is performed by calculating the exergy destruction of each piece of equipment. Exergy destruction of equipment explains the deviation of equipment performance from an ideal scenario. The higher the exergy destruction, the lower the process performance. The formulas used to calculate the exergy destruction of equipment can be found elsewhere [13].

## 3. Results and discussion: Process analysis

### 3.1. Energy analysis

In this study, the design variables, including refrigerant flow rates and suction/discharge pressures, were analyzed with respect to SEC. Table 2 lists the design variables, constraints, and the SECs of the base case and the proposed cases.

**Table 2** Design variables and SEC of the base case, case I, and case II

PMR cycle	Base case	Case I	Case II
Suction / Discharge pressure (bar)	1.5 / 13.0	1.1 / 18.0	1.1 / 11.0
Refrigerants flowrate (kg/s)	84.7	19.3	31.4
CMR cycle			
Suction / Discharge pressure (bar)	2.9 / 35.6	-	-

Refrigerants flowrate (kg/s)	33.0	-	-
<b>LMR cycle</b>			
Suction / Discharge pressure (bar)	3.7 / 25.0	1.4 / 26.0	1.3 / 15.5
Refrigerants flowrate (kg/s)	1.9	7.1	16.4
<b>SEC (kWh/kg<sub>LH2</sub>)</b>	<b>6.45</b>	<b>6.96</b>	<b>12.19</b>

According to Table 2, case II has the highest energy consumption of 12.19 kWh/kg<sub>LH2</sub> among all cases. Compared to the base case, SEC has increased by 8% and 53% in proposed cases I and II, respectively. Since the proposed cases have only two cycles (PMR and LMR), CMR will not be discussed. In addition, the suction pressure is reduced in PMR and LMR to withstand the MITA value within the exact range of 1-2 °C. Also, it is observed that the refrigeration load on the PMR cycle increases when suction pressure decreases. As a result, SEC increases. However, high suction pressure cannot achieve the desired cooling temperature. Therefore, the pre-cooling temperature was adjusted from -153.2 °C to -193.2 °C, reducing the LMR cycle's energy load. Moreover, the PMR cycle of the base case shows the highest MR flow rate (84.7 kg/s), with C<sub>1</sub>, C<sub>2</sub>, and C<sub>4</sub> contributing the most. On the other hand, the overall refrigerants flow rate is 78% and 60% less in the proposed case I and II, respectively.

Further, in proposed case I, PMR cycle, 19.26 kg/s of MR is utilized, whereas 31.382 kg/s is consumed in proposed case II. Moreover, all PMR refrigerants' flow rates are reduced in proposed cases, except C<sub>3</sub>, N<sub>2</sub>, and n-C<sub>5</sub>, whose mass flow rate is increased. This is because only low boiling refrigerants are not energy efficient. So, an optimal blend of high and low boiling refrigerants is required to reduce energy consumption. In the LMR cycle, a new refrigerant was added, Ne, besides H<sub>2</sub> and He, to reach the required H<sub>2</sub> liquefaction temperature (-252.3°C). Here, a total mass flow rate of 7.06 kg/s and 16.4 kg/s is used in the proposed cases. Also, compared to the base case, a large amount of H<sub>2</sub> and He utilized 64.5% and 84% more in both proposed cases I and II, respectively.

### 3.2. Exergy analysis

Table 3 presents the exergy destruction values of each piece of equipment of the base case and the proposed cases.

**Table 3** Equipment-wise exergy destruction of the base case, case I, and case II

<b>Equipment</b>	<b>Case I (kW)</b>	<b>Case II (kW)</b>
Compressors	4158.32	7542.97
Pumps	0.99	0.1
Coolers	4174.19	12083.33
Separators	195.36	-294.5
Reactors	-693.61	479.42
Expanders	20087.02	5582.97
CHXs	4824.24	16170.68
<b>Total</b>	<b>32746.5</b>	<b>41564.97</b>

The results of the proposed cases cannot be compared with the base case because of different process configurations. Therefore, the only case I and II are discussed in this section. In case-I, the expanders share 61% (20087.02 kW) of exergy destruction out of the total equipment destruction because their low-temperature operation promotes high exergy destruction. The CHXs show 4824.24 kW of exergy destruction. On the other hand, a high exergy destruction source in case-II caused by CHXs with 38.9 % (16170.68 kW). The main reason for exergy destruction in heat exchangers is the huge temperature difference between inlet and outlet streams. For example, in both cases, feed enters CHX-1 at 25°C, and its outlet temperature is -193.2°C (pre-cooling). Similarly, other streams also have a temperature gap of more than 50°C between in and out streams. Also,

compressors share a significant amount of exergy destruction. Since OPC is an exothermic reaction, it contributes to exergy destruction as well. However, the energy gain from the OPC reaction is recycled, so it is calculated as inlet energy for reactors. Furthermore, the exergy destruction of reactors and phase separators in case-II shows a negative value which is thermodynamically not possible. Here the negative sign may appear due to the impact of other equipment. Since it is a cascaded process, all equipment is related to each other. Overall, case II has higher exergy destruction than case I.

#### 4. Conclusions

In this study, the proposed case I is energy efficient as compared case II. The case II is energy intensive because of additional heat load of OPC reaction on refrigeration cycles. However, the case I is energy intensive as compared to the base case. Similarly, the exergy destruction analysis results conclude that the case I is 21.2% efficient than the case II. In case II, the CHXs are the major source of exergy destruction.

#### References

- [1] IEA. GLOBAL TRENDS AND OUTLOOK FOR HYDROGEN. 2017.
- [2] International Energy Agency (IEA). The Future of Hydrogen - Seizing today's opportunities. 2019. <https://doi.org/10.1787/1e0514c4-en>.
- [3] Naquash A, Riaz A, Lee H, Qyyum MA, Lee S, Lam SS, et al. Hydrofluoroolefin-based mixed refrigerant for enhanced performance of hydrogen liquefaction process. *Int J Hydrogen Energy* 2022.
- [4] Qyyum MA, Riaz A, Naquash A, Haider J, Qadeer K, Nawaz A, et al. 100% saturated liquid hydrogen production: Mixed-refrigerant cascaded process with two-stage ortho-to-para hydrogen conversion. *Energy Convers Manag* 2021;246:114659. <https://doi.org/10.1016/j.enconman.2021.114659>.
- [5] Bi Y, Yin L, He T, Ju Y. Optimization and analysis of a novel hydrogen liquefaction process for circulating hydrogen refrigeration. *Int J Hydrogen Energy* 2021. <https://doi.org/10.1016/J.IJHYDENE.2021.10.012>.
- [6] Krasae-in S, Stang JH, Neksa P. Development of large-scale hydrogen liquefaction processes from 1898 to 2009. *Int J Hydrogen Energy* 2010;35:4524–33. <https://doi.org/10.1016/j.ijhydene.2010.02.109>.
- [7] Aasadnia M, Mehrpooya M. Large-scale liquid hydrogen production methods and approaches : A review. *Appl Energy* 2018;212:57–83. <https://doi.org/10.1016/j.apenergy.2017.12.033>.
- [8] Sherif SA, Goswami DY, Stefanakos EK, Steinfeld A. *Handbook of Hydrogen Energy*. Taylor & Francis; 2014.
- [9] Riaz A, Qyyum MA, Hussain A, Lee M. Significance of ortho-para hydrogen conversion in the performance of hydrogen liquefaction process. *Int J Hydrogen Energy* 2022. <https://doi.org/10.1016/j.ijhydene.2022.09.022>.
- [10] Son H, Yu T, Hwang J, Lim Y. Simulation methodology for hydrogen liquefaction process design considering hydrogen characteristics. *Int J Hydrogen Energy* 2022;47:25662–78. <https://doi.org/10.1016/J.IJHYDENE.2022.05.293>.
- [11] Peng D, Robinson DB. A New Two-Constant Equation of State. *Ind Eng Chem Fundam* 1976;15:59–64. <https://doi.org/10.1021/i160057a011>.
- [12] Szargut J. Chemical exergies of the elements. *Appl Energy* 1989;32:269–86. [https://doi.org/10.1016/0306-2619\(89\)90016-0](https://doi.org/10.1016/0306-2619(89)90016-0).
- [13] Naquash A, Qyyum MA, Islam M, Sial NR, Min S, Lee S, et al. Performance enhancement of hydrogen liquefaction process via absorption refrigeration and organic Rankine cycle-assisted liquid air energy system. *Energy Convers Manag* 2022;254:115200. <https://doi.org/10.1016/j.enconman.2021.115200>.

# Optimal design of a microbubble cap to enhance mass transfer efficiency using image processing

Yup Y.,<sup>a,b</sup> Jonghun L.,<sup>a,b</sup> Chonghyo J.,<sup>a,b</sup> Seongbin G.,<sup>a</sup> Sunghyun C.,<sup>a</sup> Jaewon L.,<sup>a</sup> Junghwan K.,<sup>a,b</sup> Hyungtae C.,<sup>a,\*</sup>

<sup>a</sup>*Green Materials and Processes R&D Group, Korea Institute of Industrial Technology, 55, Joga-ro, Ulsan 44413, Korea*

<sup>b</sup>*Department of Chemical and Biomolecular Engineering, Yonsei University, 50, Yonsei-ro, Seoul 03722, Korea*

[htcho@kitech.re.kr](mailto:htcho@kitech.re.kr) 82

## Abstract

Microbubble technology has been used in various industries to improve the mass transfer efficiency of applications. Although conventional bubble generators using venturi and swilling nozzles are useful, their low gas throughput is insufficient for industrial pollutant reduction. Hence, this study developed an optimal microbubble cap structure with a high gas throughput and microbubble generation rate using an image processing based overall mass transfer coefficient (OMTC) calculation model. Micro bubble cap with brim structure was developed, and the bubble generation pattern by varying the brim size was analyzed in a lab-scale microbubble system. Bubble diameters were measured using image processing and OMTC was calculated using bubble size. The observed bubble size range was approximately 0–250  $\mu\text{m}$ , and the bubble generation ratio was the highest at a 5-mm-brim size. When calculating the OMTC of  $\text{H}_2\text{S}$  in water of 50  $\mu\text{m}$  or less, the value increased by more than 50% compared to other brim-size cases. Therefore, at brim sizes of 5 mm, the dispersion and production of microbubbles were active, and they were expected to endow high mass transfer efficiency.

**Keywords:** microbubble, vortex, overall mass transfer coefficient, bubble cap, image processing

## 1. Introduction

Microbubbles are used in various fields, such as water treatment, reactors, nuclear power, and air pollutant reduction (Yup Y. et al., 2021). In the various industries, they are used either for the formation of a large interfacial area between the gas and liquid phases, the generation of radicals, or long interaction time in flow field. The principle of utilization of microbubble is that it creates an environment that enables gas–liquid interaction. Particularly, the gas–liquid interaction is a significant factor in the field of air pollutant reduction. Microbubbles have recently been used to remove air pollutants and increase reduction efficiency using bubble scrubber system (Hyundo P. et al., 2021).

Microbubbles have a role in absorbing air pollutants in a liquid state. Therefore, it is important to generate many small-sized bubbles and create an environment for enhancing mass transfer efficiency by expanding the interfacial area (Al-Mashhadani. et al., 2012).

Microbubbles are generated mainly using a small venturi structure or a sparger-type pressurized generation method (Ali, M. et al., 2013). Although this method is suitable for generating microbubbles, it is difficult to use in the air pollutant reduction field to quickly purify considerable gas owing to low gas throughput. Therefore, a novel microbubble generator is needed.

Confirmation of the bubble size must be an essential step to design the bubble generator. Because it is possible to calculate overall mass transfer coefficient (OMTC) from the bubble size data, which is a significant value in air pollutants removal process (Yup Y. et al., 2023). Nevertheless, previous studies cannot calculate OMTC because they cannot determine the volume of the flow field, which is three-dimensional data, due to limitations in image data. Therefore, we presented a OMTC calculation model that determines bubble size and volume of the flow field by using concept of camera depth of the field and image processing method. This study focuses on an optimal structure for a bubble cap is determined using an OMTC calculation model.

## 2. Methodology

### 2.1. Setup of lab-scale microbubble system

This study proposed a novel structure that induces bubble breakage by adding a brim structure that prevents large bubbles from rising. Bubble caps with various brim widths (0, 5, 10, 15, and 20 mm) were fabricated using 3D-printing to investigate the tendency of bubble generation pattern and OMTC to vary depending on the brim width.

A simple experimental device was created to confirm the bubble generation pattern by brim structure. The device has a rectangular structure with an acrylic body, and its width and height are 300 and 500 mm, respectively. Additionally, a vacuum suction blower draws air from the top of the rectangular structure. The induced airflow passes through the air velocity indicator and generates bubbles in the bubble cap installed in the center of the bottom of the experimental device. Water was filled to a height of 10 cm from the bottom of the experimental device. A detailed airflow diagram and component descriptions are shown in Fig. 1.

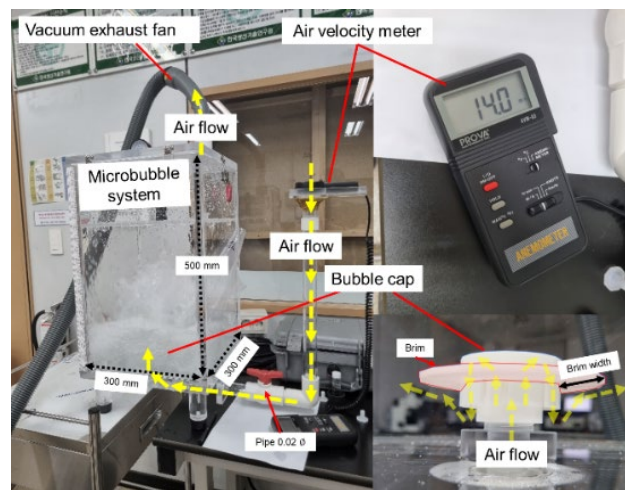


Fig. 1. Bubble generating system experimental setup

2.2. Image processing based overall mass transfer coefficient calculation model

A DSLR camera (SONY A7M3) and a 24 mm-Laowa macro-lens, were used to extract images of the bubbles generated. Pictures were taken using the camera at a height of 1 cm from the bottom of the experimental device. The photographed images were then processed to measure the bubble size. Image processing was performed using OpenCV-Python, and the shape of the bubble was detected using Gaussian blur, Canny edge detection, and morphology transformation functions. Figure 2 shows a sample image generated by the image processing process for bubble size measurement.

Gaussian blur step is a role in suppressing noise in the image, and the Canny edge detection step determines the bubble edge based on the gradient value of brightness. Then, in the morphology transformation step, the closed curve converts to a contour by filling the curve and determining the contour area to calculate the bubble's size. Finally, the extracted bubble size data calculates the number of bubbles and the average bubble size and displays it as the resulting image in Fig. 2.

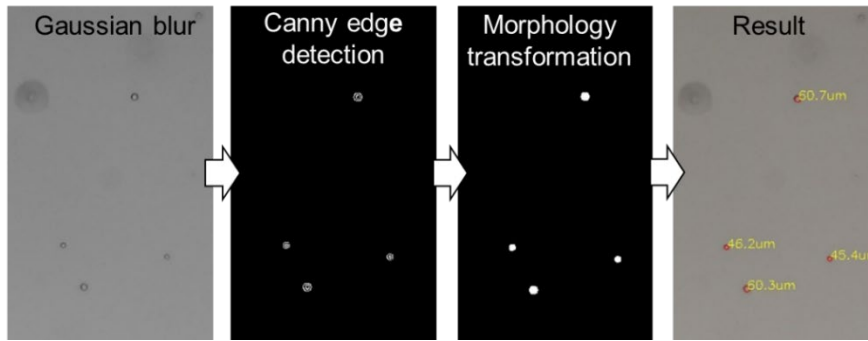


Fig. 2. Sample image of the image processing process for bubble size measurement

The mathematical model for determining OMTC to confirm the mass transfer efficiency due to the formation of microbubbles begins with the following equation.

$$\frac{\partial C_A}{\partial t} + \frac{\partial(uC_A)}{\partial z} = c \cdot A_{\text{surf}}(C_A^* - C_A) = k_{\text{ov}}(C_A^* - C_A) \quad (1)$$

$$A_{\text{surf}} = \frac{6\phi_{\text{gas}}}{d_{\text{sm}}} \quad (2)$$

where  $C_A$  represents the concentration of gas component A in a gaseous state,  $C_A^*$  represents the concentration of A in equilibrium with the current liquid,  $u$  represents the gas velocity, and  $A_{\text{surf}}$  represents the interfacial area concentration. Additionally,  $\phi_{\text{gas}}$  is the volume ratio of the bubble in the flow field, and  $d_{\text{sm}}$  is the Sauter mean meter of bubbles. Equations (1) and (2) can be used to extract bubble size data from several bubble images in a short time and determine OMTC.

$$\phi_{\text{gas}} = \frac{V_{\text{tb}}}{V_{\text{ff}}} \quad (3)$$

where  $V_{\text{tb}}$  is the sum of bubble volume in the flow field, and  $V_{\text{ff}}$  is the volume of flow field.  $V_{\text{ff}}$  is determined from the camera the depth of field.

$$V_{\text{ff}} = A_{\text{ip}} C_{\text{dof}} \quad (4)$$



$A_{ip}$  is the area where bubbles are detected in the image through image processing, and  $C_{dof}$  is the depth of field.

### 3. Results and Discussions

#### 3.1. Bubble size distribution analysis

The best way to analyze the bubble formation pattern is to determine the bubble size distribution (BSD). The BSD analysis method is a common method used for multi-phase processing. This study also checked the bubble size distribution to confirm the bubble formation pattern according to the bubble cap's brim width. The confirmed result is shown in Fig. 3.

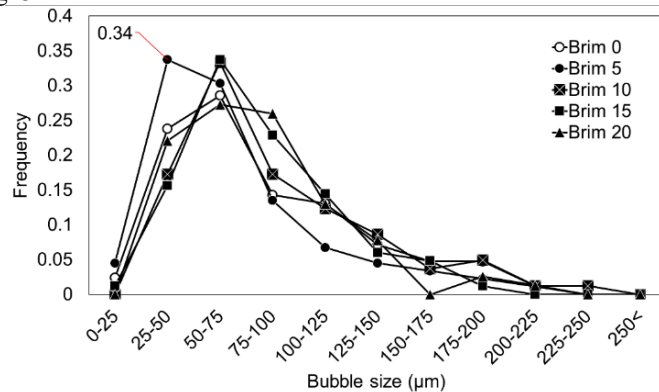


Fig. 3. Bubble size distribution graph according to the brim width of the bubble cap

As can be observed from the figure, when the brim width is 5 mm, bubbles between 25 and 50  $\mu\text{m}$  are generated the most. This shows a frequency (that is, 0.34), which is approximately 40% higher than that of bubble caps with other brim width structure, and the bubble generation ratio of 50  $\mu\text{m}$  or less was approximately 38%. Additionally, in the brim widths of 0, 10, 15, and 20 mm, bubbles between 50 and 75  $\mu\text{m}$  were generated with the highest frequency, and the frequency of bubbles between 75 and 100  $\mu\text{m}$  tended to increase as the brim width increased. The benefits of microbubbles with a size less than 50  $\mu\text{m}$  are numerous. Particularly, the interaction time between phases can be maintained for a long time because the buoyancy of small bubbles enables them to stay in the flow field for a long time without rising and disappearing.

### 3.2. Statistical analysis of bubble generating pattern

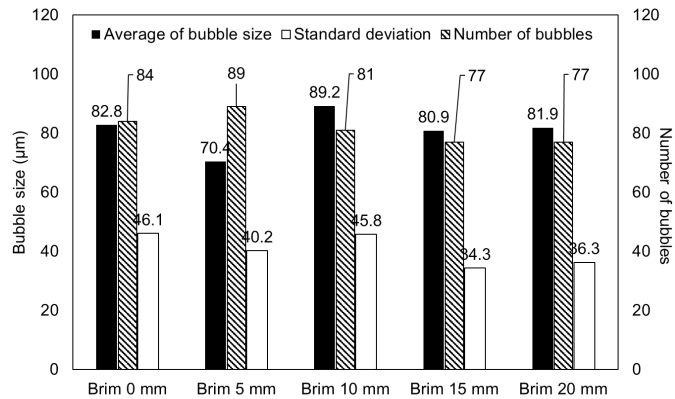


Fig. 4. Results of average bubble size, standard deviation, number of bubbles analysis for generated bubbles

Since the generation of bubbles is accompanied by complex mechanisms and microbubble is generated by various sizes, it is challenging to identify bubble generation patterns visually; hence, identifying the pattern of bubble formation using a statistical analysis method is essential. Therefore, the standard deviation for the average bubble size and number of bubbles was determined and is shown in Fig. 4. The average bubble size was 70.4 µm, which was the smallest in the 5-mm-brim structure. Additionally, 89 bubbles were generated in the observed flow field, which was confirmed to be higher than for other conditions. Where a brim structure creates bubbles of small average size, the standard deviation is lower than when the brim width is 0 and 10 mm. Hence, when the brim width is 5 mm, microbubbles can be well dispersed. Therefore, it is considered to have a characteristic of actively forming turbulence so that many bubbles can be generated.

### 3.3. Result of overall mass transfer coefficient

Identifying and determining the OMTC in the air pollutant removal process can be an index for selecting process conditions in which actual mass transfer is active. Thus, this index can be used to determine the mass transfer efficiency and improve the process's efficiency under various process conditions. Particularly, the mass transfer effect for the bubbles generated using the novel brim structure must be evaluated because the generation of bubbles of a size of 50 µm or less has a positive effect on the mass transfer efficiency. Figure 5 shows the OMTC for H<sub>2</sub>S, a representative air pollutant, which was determined for bubbles less than 50 µm.

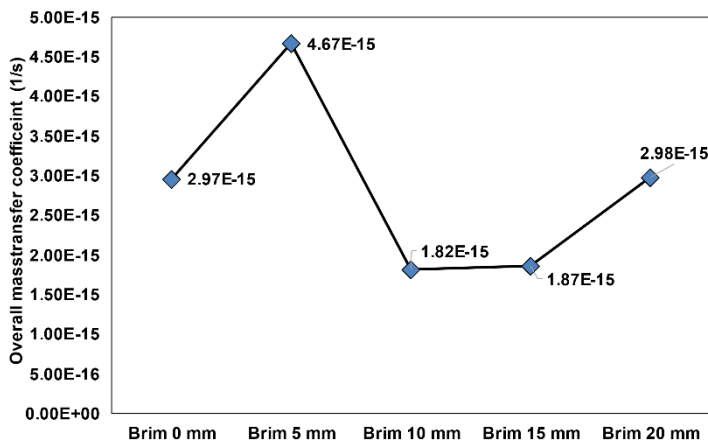


Fig. 5. OMTC calculation results for bubbles less than 50  $\mu\text{m}$

This outcome demonstrated that the OMTC was at its highest when the brim width was 5 mm. It showed an approximately 50% higher value than the other conditions.

#### 4. Conclusion

In this study, a bubble cap structure that generates a large number of microbubbles with high gas throughput is developed, and the optimal structure of the bubble cap is determined using image processing based OMTC calculation model. Bubble cap structure was then used to confirm the bubble generation pattern according to the width of the additional brim structure that prevents air from rising. A case study was conducted according to the bubble cap brim width (0, 5, 10, 15, and 20 mm) inside a bubble generator system made of acrylic.

Under the condition of brim 5 mm, the rate of formation of bubbles less than 50  $\mu\text{m}$ , which shows the microbubble characteristics, was as high as 38%. Additionally, the highest amount of bubble formation was observed under the 5-mm-brim condition, and the average bubble size was also determined to be the lowest at 70.4  $\mu\text{m}$ . Furthermore, the OMTC value was established to validate the effect of bubbles less than 50  $\mu\text{m}$  on  $\text{H}_2\text{S}$ , a representative air pollutant. OMTC, 4.67E-15 ( $\text{s}^{-1}$ ), was the highest in the 5-mm-brim condition, where bubbles of less than 50  $\mu\text{m}$  were frequently generated. According to these results, the 5-mm-rim bubble cap, in which bubbles of less than 50  $\mu\text{m}$  were induced by the dispersion and destruction of bubbles and appropriate location of turbulent flow, is considered the most suitable structure for removing air pollutants. The structure can only be used to create bubbles of specific size. Additionally, in future studies, we expect to enhance accuracy of OMTC calculation model through gas removal experiments.

#### Acknowledgment

This work was supported by the Korean Institute of Industrial Technology within the framework of the project "Development and application of AI based microbubble-scrubber system for simultaneous removal of air pollutants [grant number KM-22-0015]."

## **References**

- Ali, M. et al., 2013, Iodine removal efficiency in non-submerged and submerged self-priming venturi scrubber, *Nuclear engineering and technology*, 45, 2, 203-210.
- Al-Mashhadani., et al., 2012, CO<sub>2</sub> mass transfer induced through an airlift loop by a microbubble cloud generated by fluidic oscillation, *Industrial & Engineering Chemistry Research*, 51, 4, 1864-1877.
- Hyundo P., et al., 2021, Computational fluid dynamic modelling of optimal water level in low-pressure microbubbles scrubbers, *Chemical Engineering Transactions*, 86, 613-618.
- Yup Y., et al., 2021, Method for determining optimum operational conditions of microbubble scrubber using image processing, *Journal of Environmental Informatics*, 38, 2, 83-92.
- Yup Y., et al., 2023, Method for measuring bubble size under low-light conditions for mass transfer enhancement in industrial-scale systems, *International Communications in Heat and Mass Transfer*, 140, 106525.



# Design of a heterogeneous azeotropic distillation processes using stochastic optimization

Josué J. Herrera Velázquez,<sup>a</sup> Julián Cabrera Ruiz,<sup>a</sup> J. Rafael Alcántara Avila,<sup>b</sup>

<sup>a</sup>*Departamento de Ingeniería Química, Universidad de Guanajuato, Noria Alta S/N, Guanajuato, 36050, Mexico.*

<sup>b</sup>*Department of Chemical Engineering, Kyoto University, Kyoto 615-8510, Japan.*

## Abstract

Heterogeneous azeotrope distillation is often preferred industrially over homogeneous azeotrope distillation due to the ease of recovery of the entrainer and the transition across a distillation boundary in the decanter. In a previous study, the synthesis problem for separating an Ethanol and Water mixture by using benzene as an entrainer was solved using a low aggregation superstructure comprised of functional distillation and decanter modules where the utility cost was minimized (Okunishi, 2019). The solution was interpreted to obtain a realistic process. This work optimizes the interpreted solution using a modified Simulated Annealing Algorithm. The new solution makes it possible to reduce the overall energy consumption and the total annual cost of the process. Also, an interface between Python and Aspen Plus® is proposed, and it allows the reduction of convergence problems caused by the complexity of the optimized scheme.

**Keywords:** Stochastic Optimization, Azeotropic Distillation, Process Simulation, Aspen Plus – Python interface, modified Simulated Annealing Algorithm.

## 1. Introduction

A heterogeneous azeotropic mixture creates a new azeotropic point by adding a specific third component called entrainer to the binary system. Thus, in heterogeneous azeotropic distillation, a mixture having a composition that exists in the two liquid phase region is separated into an organic phase and an aqueous phase by a decanter. This work proposes the post-optimization of a synthesis problem that departs from the concept of unit operations and proposes functional modules. Hasebe et al. (2015) proposed distillation modules to separate a ternary mixture in previous work. The objective function was the utility cost minimization, and the optimal process structure was interpreted near the minimum reflux condition. Later, the methodology for the design of the separation scheme of the ethanol/water mixture using benzene as an entrainer that was implemented by Okunishi (2019) was taken through the solution of a superstructure. Since the interpreted solution was done assuming a large number of stages in each column section, it can be improved if the equipment cost is included to find a better compromise between equipment cost and operation cost. Therefore, the previous study is complemented with a rigorous simulation of the process that makes the process even more realistic, that is, to a real reflux relationship with finite dimensions.

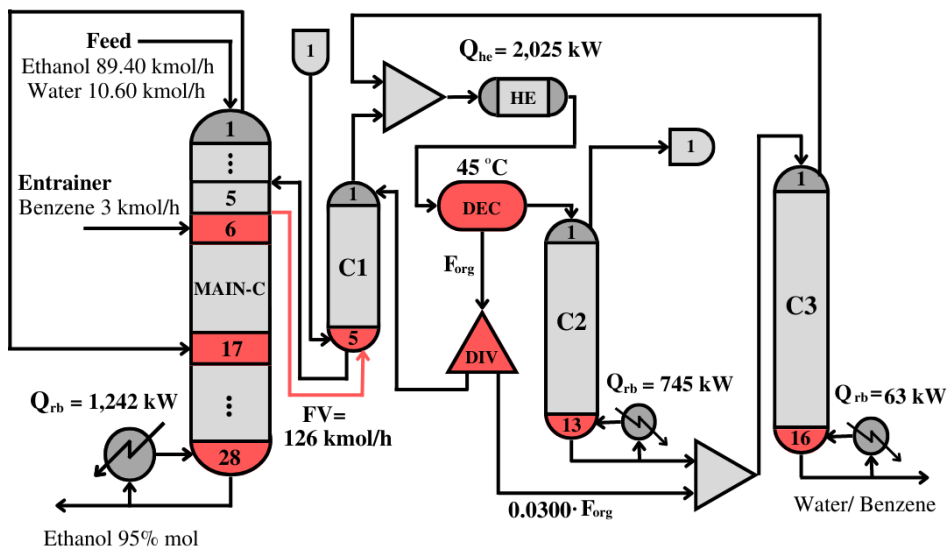
Three distillation columns formed the distillation process. The first column was a thermally coupled distillation column with a side rectifier, and the ethanol product was obtained at the bottom. Then, the top vapor entered the decanter, and part of the organic

stream was recycled to the first column. The aqueous stream was fed to the second distillation column, where water was concentrated at the bottom. In the third column, the water-rich stream was mixed with some of the organic stream to take all possible ethanol and benzene to concentrate water towards the column's bottom. For rigorous model simulation, this work uses an interface between the free access programming language Python and the chemical process simulation software Aspen Plus V12®. A modified Simulated Annealing Algorithm (m-SAA) stochastic algorithm was proposed by Cabrera et al. (2021). The problem was solved by minimizing the Total Annual Cost (TAC).

## 2. Methodology

### 2.1. Case Study

The scheme of the base case is shown in Figure 1. It was designed using Aspen Plus® software with the NRTL thermodynamic model. An ethanol/water mixture of 100 kmol/h (89.4% mol of ethanol, 10.6% mol of water) is fed to the system in the main column (MAIN-C). 3 kmol/h of fresh benzene is fed into the system. Optimization is done using the RADFRAC module for the distillation columns. The pressure of all columns is 1 atm, and the pressure drop was neglected. In MAIN-C, ethanol is obtained at a purity of 95 mol%, and in C3, water is obtained with the same purity.



**Figure 1.** Heterogeneous azeotropic distillation processes (Base case)

MAIN-C is the main column where the ethanol/water and benzene mixtures are feed. Here, the ethanol is also recovered in the bottom. C1 is the rectifier of MAIN-C. HE is the condenser of C1. DEC is the decanter after the C1 condenser that recovers the benzene and sends it to C2, and DIV is a splitter of the organic phase between C1 and C3. The reported value of the splitter is the fraction of flow that goes to the mixer that connects to C3. C3 is the column where water is removed from the process, and a part of the benzene is recovered. Figure 1 shows the solution with process optimal structure after interpreting the synthesis problem for separating mixtures forming heterogeneous azeotropes. Although the process structure is optimal, the number of stages, feed stages, or flows between columns have not been optimized but simulated at near minimum reflux

conditions. Therefore, this work solves a post-optimization problem that finds the best structural and operating conditions that minimize the Total Annual Cost (*TAC*).

In Fig. 1, the variables proposed to optimize the design problem are shown in red. In general, they represent a direct impact on the thermal load of the reboilers of each distillation column since the number of stages of each one and the conditions and amount of organic solution entering these will have a direct impact on these values. The design was proposed based on the number of modules present in the solution by Okunishi (2019).

## 2.2. Mathematical treatment

The case study was optimized according to the objective function shown in Eq. 1.

$$\begin{aligned} \text{Min}(TAC) &= f(N_{MAINC}, N_{C1}, N_{C2}, N_{C3}, Div_{DIV}, T_{DEC}, FV_{MAINC}, N_R, N_F) & 1 \\ \text{s. t.} \quad R_i(x) &\geq 99.5 \% \quad \& \quad CC = 1 \end{aligned}$$

$R_i(x)$  is the vector associated with the recovery percentage (for ethanol and water) that was 99.5 %. Since it is quite a complex simulation, placing a design specification on it will increase the number of iterations to reach convergence. For that reason, a penalty is proposed to the *TAC* in case the fixed purity of 95% mol is not met for the ethanol in MAIN-C. For the calculation of the cost (*FOB*) Eq. 2 is used.

$$FOB \left( \frac{USD}{year} \right) = OC + \frac{\left( \frac{CEPCI \text{ in } 2019}{CEPCI \text{ in } 2003} \right) (C_{Shell} + C_{Plates} + C_{Comp} + \sum_{j \in CS}^{i \in HS} C_{Exchanger \ i,j})}{\text{Recovery time (Year)}} \quad 2$$

The *FOB* modules were calculated using the Guthrie method (Turton et al., 2018) with a payback time of 5 years using the CEPCI value corresponding to the one for 2019. The adjustment is used under the  $F_C$  criterion that relates to the purity (In mole fraction) of the system, as shown in Eq. 3.

$$F_C = 1 + [0.95 - Purity_{EtOH}] \quad 3$$

This correction factor is multiplied by the *FOB*, so the *TAC* calculation is adjusted, as shown in Eq. 4.

$$TAC \left( \frac{USD}{year} \right) = FOB \cdot F_C \quad 4$$

This correction factor  $F_C$  allows us to adjust the *TAC* to the desired purity so that if the purity is higher, the technology will be less expensive. If the purity is low, the technology will be more expensive than the one calculated. The *CC* refers to a convergence criterion by the simulation in Aspen Plus V12®, which can take two values as described in Eq. 5.

$$CC = \begin{cases} 0 & \text{then } TAC = 2,000,000 \text{ \$/year} \\ 1 & \text{then } TAC \text{ calculated by Ec. 4} \end{cases} \quad 5$$



The value of  $CC$  equal to zero indicates that the simulation does not converge for the optimization variables proposed by the algorithm in that iteration. Therefore, it is penalized with a large  $TAC$  (i.e., 2,000,000 \$/yr.). The penalty value in Eq. 5 was determined by sensitivity analysis sufficiently large. The bounds on the variables for optimization are presented in Table 1.

**Table 1.** Optimization Variables

Variable	Description	Optimization bounds [lower, upper]	Variable type
$N_{MAINC}$	MAIN-C Stages	[20, 37]	Discrete
$N_{C1}$	C1 Stages	[3, 10]	Discrete
$N_{C2}$	C2 Stages	[6, 20]	Discrete
$N_{C3}$	C3 Stages	[8, 25]	Discrete
$Div_{DIV}$	Divider	[0.0150, 0.2000]	Continuous
$T_{DEC}$	Decanter Temperature (°C)	[40, 65]	Continuous
$FV_{MAINC}$	FV Flow (kmol/h)	[90, 135]	Continuous
$F_R$	Fresh Benzene Feed Stage	[7, $N_{MAINC}-3$ ]	Discrete
$F_F$	MAINC Dome Vapor Recirculation Feed Stage	[4, $N_R-2$ ]	Discrete

The optimization variables followed analogous to the work reported by Alcántara-Maciel et al. (2022) as shown in Eq. 6.

$$L_{Ni} = (L_{Upper} - L_i) / (L_{Upper} - L_{Lower}) \quad 6$$

$L_{Ni}$  is the value of the normalized variable  $L_i$ ,  $L_{Upper}$  is the upper limit of variable  $i$  and  $L_{Lower}$  is the lower limit of variable  $i$ . The algorithm works with random values from 0 to 1. For discrete variables, a rounding function is used. This technique ensures that the search range of the variables in the optimization does not go outside the limits defined in Table 1. Concerning dependent variables, this technique allows us to define some limits based on previous variables so that we have dynamic limits that do not limit the search space. The algorithm determines the variables in the descending order shown in Table 1.

The optimization algorithm used is the modified Simulated Annealing Algorithm (m-SAA) proposed by Cabrera et al. (2021), which is capable of manipulating integer variables more effectively than the typical continuous relaxation of the method. Four annealing of 225 iterations each were used, having a total of 900 iterations that ensured at least 100 iterations per variable to be optimized.

The Spyder IDE 5.2.2 (Spyder, 2022) was used to develop and execute the Python code to run the interface with Aspen Plus. The overall implemented interface is shown in the schematic presented in Figure 2. The computer used for this optimization has an Intel(R) Xeon(R) W-3235 CPU at 3.30 GHz with 48 GB of RAM.



**Figure 2.** Optimization-simulation software interface

### 3. Results and Discussion

The post-optimization results are shown in Table 2 and compared with the base case.

**Table 2.** Design variables and optimal operating specifications for the different cases.

Variable	Base Case	Proposed solution by the post-optimization
$N_{MAINC}$	28	25
$N_{C1}$	5	5
$N_{C2}$	13	6
$N_{C3}$	16	19
$Div_{DIV}$	0.0300	0.1082
$T_{DEC} (^{\circ}C)$	45.00	53.52
$FV_{MAINC} (kmol/h)$	126.00	91.84
$F_R$	17	20
$F_F$	6	4
$TAC (USD/year)$	572,551	462,636
Total heat duty (Reboilers only) (kW)	2,050	1,466
Condenser duty (kW)	2,025	1,488

Figure 3 shows the optimal scheme in post-optimization.

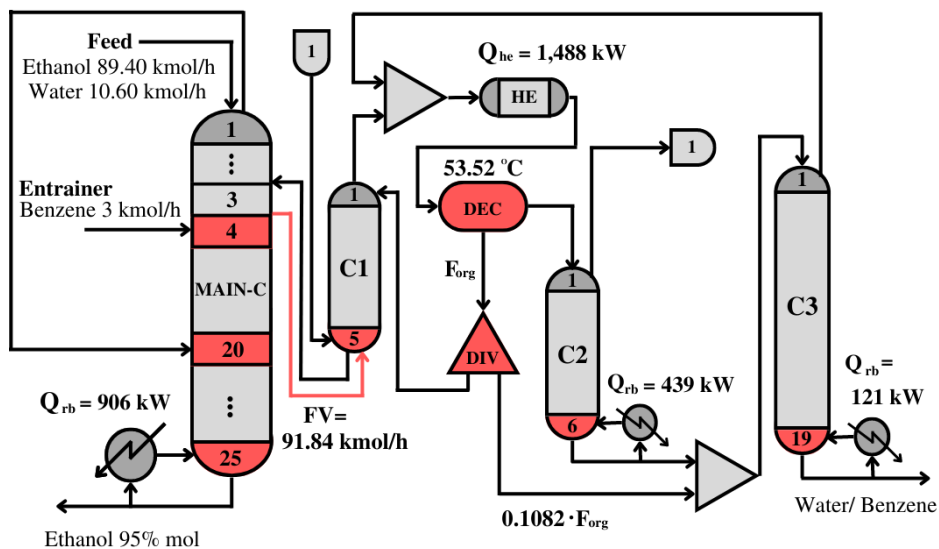


Figure 3. Optimal solution given by m-SAA

In the last solution, the number of stages in the C1 and C3 columns were similar, but the C2 column needed only a few stages. In the base case, the decanter temperature was assumed to be 45°C, but in the best case, a higher temperature (53.5°C) was better because decanters, in general, decrease their separation efficiency when they are operated at low temperatures, this is confirmed by the studies carried out by Skiborowski (2015). In our system, the recirculation flow (Organic) to C1 is less due to the division reflected by the splitter, and this generates lower heat duty in the reboiler of the main column (MAIN-C). The vapor interconnection flow of the first column was 27% lower than that obtained in the base case, which resulted in the reboiler duty reduction in columns MAIN-C and C2. Thus, the overall energy consumption of the process was reduced.

The recycles flows of column C3 (which is rich in organic phase), the one between column C2 and C1 and the recirculation of vapor in MAIN-C reduce the heat of the reboilers in MAIN-C in the same way and C2, the same occurs with the FV interconnection flow. This energy reduction directly reduce the area of the reboilers, and consequently minimize the cost of the TAC in post-optimization. In the case of the C3 column, the reboiler load increases because it receives a greater amount of organic phase flow, but this increase is small compared to all the energy saved in MAIN-C and C2.

#### 4. Conclusions

The proposed problem was solved using an interface with the Python programming language and the Aspen Plus V12® process simulator. The TAC base case was USD 572,551/year. The TAC of the optimized process using the m-SAA stochastic optimization algorithm was 462,636 USD/year, which implies a saving of 109,914 USD/year, which is a saving of approximately 20% for the process obtained at near minimum reflux condition. Using the m-SAA algorithm and rigorous models allows us to have a robust design, which in the case of this work, also implies a saving in its TAC. Using the proposed interface between Python and Aspen Plus V12® reduces the convergence problems caused by the complexity of the optimized scheme. The use of an

optimization algorithm as post-optimization was effectively used to improve the operating and structural conditions from the original design proposed by synthesis optimization.

## **References**

- F. D. Alcántara-Maciel, V. E. Casillas-Céspedes, J. A. López-García, J. Cabrera-Ruiz, C. Ramíez Marquez, J. R. Alcántara-Ávila, 2022, Economic optimization of a reactive distillation column with multiple reactive sections for silane production, *Computer Aided Chemical Engineering*, 51, 475–480.
- J. Cabrera Ruiz, S. Hasebe, J. R. Alcantara Avila, 2021, Improvement of the Optimal Design Procedure Using Randomized Algorithm and Process Simulators, *MATEC Web Conf* 2021, 333, 06004.
- R. Turton, J. Shaeiwitz, D. Bhattacharyya, W. Whiting, 2018, *Analysis, Synthesis, and Design of Chemical Processes*, 5th Edition. Boston, USA: Pearson.
- M. Okunishi, 2019, Development of systematic design method for azeotropic distillation process using decanter, Kyoto University Master Thesis (in Japanese).
- Spyder, 2021, <https://www.spyder-ide.org/> (accessed on 27 November 2022).
- H. Takase, S. Hasebe, 2015, Optimal Structure Synthesis of Ternary Distillation System, *Computer Aided Chemical Engineering*, 37, 1097–1102.
- M. Skiborowski, A. Harwardt, W. Marquardt, 2015, Efficient optimization-based design for the separation of heterogeneous azeotropic mixtures, *Computers and Chemical Engineering*, 72, 34-51.



# Single and multi-objective optimization of heat integrated distillation columns using PYMOO

Josué J. Herrera Velázquez,<sup>a</sup> Julián Cabrera Ruiz,<sup>a</sup> J. Rafael Alcántara Avila,<sup>b</sup>  
Salvador Hernández Castro,<sup>a</sup>

<sup>a</sup>*Departamento de Ingeniería Química, Universidad de Guanajuato, Noria Alta S/N, Guanajuato, 36050, Mexico.*

<sup>b</sup>*Department of Chemical Engineering, Kyoto University, Kyoto 615-8510, Japan.*

## Abstract

Python has become an open-source interpreted programming language that has recently gained popularity. The implementation of interface codes for process optimization between process simulators and closed-source programming languages has been widely reported. This paper presents an interface between the commercial chemical process simulation software Aspen Plus and the open-source programming language Python using the PYMOO (Multi-Objective Optimization in Python) package developed by Blank et al. (2020). This analysis was performed on a Heat-Integrated Distillation Column (HIDiC) previously optimized by Herrera-Velázquez et al. (2022) was taken as a case study. Multi-objective (NSGA-II) and single-objective (GA and PSO) optimization algorithms were used to minimize economic and energy criteria. Mathematical handling techniques such as normalization of variables and pseudo-normalizations of the objective functions were implemented to improve the solutions and to reduce the computation time during optimization.

**Keywords:** Global Optimization, Stochastic, PYMOO, Aspen Plus-Python Interface, HIDiC.

## 1. Introduction

Interfaces between process simulation and optimization software are increasingly becoming more utilized for developing new processes. Stochastic algorithms are particularly interesting because they can link process simulators, which contain material and energy balances represented by highly non-linear equations. The Python programming language has gained popularity among scientists and engineers to solve various mathematical problems. Nevertheless, its capabilities have not been exploited for solving chemical processes.

This work proposes an interface framework between the commercial chemical process simulation software Aspen Plus V12® and the open-source programming language Python. The PYMOO (Multi-Objective Optimization in Python) package developed by Blank et al. (2022) was used to optimize the studied process. This work takes the optimization of a Heat Integrated Distillation Column (HIDiC) to separate a binary mixture of benzene/toluene from a previous work (Herrera-Velázquez et al., 2022). For single optimization, the minimization of an economic criterion, the Total Annual Cost (TAC), was considered. The minimization of energy consumption was also considered for multi-objective optimization because there is a trade-off between energy reduction

resulting from heat integration and cost increase due to expensive electricity and compressor costs. Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) were used for single optimization. In contrast, the Non-Dominated Sorting Genetic Algorithm (NSGA-II) was used for multi-objective optimization. The NSGA-II algorithm, which is integrated with the PYMOO package, follows the general scheme of a genetic algorithm and contains a modified survival selection and binary mating, where each individual is compared by rank and crowding distance (Deb et al., 2002). The PSO (Particle Swarm Optimization) algorithm, which is also included in the PYMOO package, uses the concept of particle swarming to guide its search, each particle possessing a velocity and strongly influenced by the best solutions found, both locally and globally (Kennedy et al., 1995).

## 2. Methodology

### 2.1. Case Study

The process described in Figure 1 is simulated using Aspen Plus V12® software and the Peng-Robinson thermodynamic model was used. The feed is 50/50 kmol/h of benzene/toluene. By design specifications, each product must be obtained with a purity of 99.9% mol for both components. Optimization is done using the RADFRAC module for the distillation columns. The pressure in the STR column is 1 atm, while for the REC column, the pressure is PR times the STR pressure.

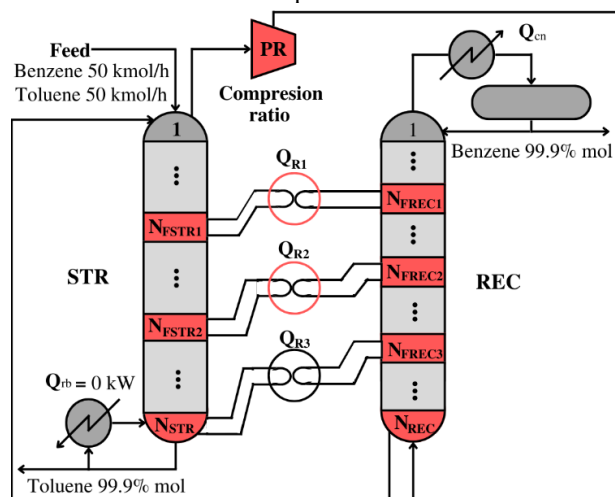


Figure 1. Heat Integrated Distillation Column

In the figure,  $N_{STR}$  and  $N_{REC}$  are the total stages of the stripping and rectification columns, respectively.  $PR$  is the compression ratio (i.e., the pressure in the rectification column divided by the pressure in the stripping column),  $Q_{R1}$  and  $Q_{R2}$  are the loads of the exchangers that perform the energy integration between the column zones,  $N_{FSTRj}$  and  $N_{FRECj}$  are the stages in which the integration is carried out for each exchanger. In the case of  $N_{FREC3}$ , it is the stage that connects with the exchanger that acts as a reboiler in STR.

### 2.2. Mathematical treatment

The case study was optimized according to the objective function shown in Eq. 1.

$$\text{Min}(TAC, Q_{COMP}) = f(N_{STR}, N_{REC}, PR, Q_{R1}, Q_{R2}, N_{FSTR1}, N_{FSTR2}, N_{FREC1}, N_{FREC2}, N_{FREC3}) \quad 1$$

$$s.t. \quad P_i(x) \geq 99.9 \% \text{ mol}, \quad R_i(x) \geq 99.5 \% \quad \& \quad CC = 1$$

$P_i(x)$  and  $R_i(x)$  are the constraint vectors associated with percent mol purity (99.9% mol for benzene and toluene) and percent recovery (99.5% for benzene and toluene), respectively. These constraints are integrated into the simulation as design specifications. The  $TAC$  objective function is given in Eq. 2:

$$TAC \left( \frac{USD}{year} \right) = OC + \frac{\left( \frac{CEPCI \text{ in } 2019}{CEPCI \text{ in } 2003} \right) (C_{Shell} + C_{Plates} + C_{Comp} + \sum_{j \in HS} C_{Exchanger \ i,j})}{Recovery \ time \ (Year)} \quad 2$$

OC is the operating cost. The  $TAC$  modules were calculated using the Guthrie method (Turton et al., 2018) with a payback time of 3 years using the 2019 CEPCI.  $Q_{COMP}$  is the value calculated by Aspen Plus of the module COMP (Compressor). Guthrie's method, and  $Q_{COMP}$  are a function of the variables presented in Eq. 1. The  $CC$  refers to a convergence status by the simulator Aspen Plus®, which can take two values as described in Eq. 3 and modify the Total Annual Cost ( $TAC$ ) for non-converged simulations.

$$CC = \begin{cases} 0 & \text{then } TAC = 2,000,000 \text{ \$/year} \ \& \ Q_{COMP} = 300 \text{ kW} \\ 1 & \text{then } TAC \ \& \ Q_{COMP} \text{ calculated by algorithm} \end{cases} \quad 3$$

The value of  $CC$  equal to zero indicates that the simulation does not converge for the optimization variables proposed in the iteration by the algorithm. The values of  $TAC$  and  $Q_{COMP}$  given in Eq. 3 are penalty values determined by sensitivity analysis, such that they are sufficiently larger than the values expected by the simulations that converge. The bounds on the variables for optimization are presented in Table 1.

**Table 1.** Optimization Variables

Variable	Optimization bounds [lower, upper]	Variable type	Units
$N_{STR}$	[15, 30]	Discrete	--
$N_{REC}$	[15, 30]	Discrete	--
$PR$	[2, 3]	Continuous	--
$Q_{R1}$	[100, 600]	Continuous	kW
$Q_{R2}$	[100, 600]	Continuous	kW
$N_{FSTR1}$	[2, $N_{STR}$ ]	Discrete	--
$N_{FSTR2}$	[ $N_{FSTR1}+1$ , $N_{STR}-1$ ]	Discrete	--
$N_{FREC1}$	[2, $N_{STR}-3$ ]	Discrete	--
$N_{FREC2}$	[ $N_{FREC2}$ , $N_{STR}-2$ ]	Discrete	--
$N_{FREC3}$	[ $N_{FREC3}$ , $N_{STR}-1$ ]	Discrete	--

The variables have been normalized similarly to Herrera-Velázquez et al. (2022), as shown in Eq. 4, to compare the multi-objective and single-objective results with respect to the previous study and to determine the advantages and disadvantages of the efficiency and effectiveness of the proposed platform

$$L_{Ni} = (L_{Upper} - L_i) / (L_{Upper} - L_{Lower}) \quad 4$$



where  $L_{Ni}$  is the value of the normalized variable  $L_i$ ,  $L_{Upper}$  is the upper limit of variable  $i$ , and  $L_{Lower}$  is the lower limit of variable  $i$ . The algorithm works with random values between 0 and 1, and a rounding function is used for discrete variables. The use of dynamic bounds ensures that the search range of the variables in the optimization does not go outside the limits defined in Table 1, which was done similarly to Alcántara-Maciél et al. (2022). This technique considers the bound of dependent variables. The algorithm defines the variables in the order shown in Table 1.

A case study was proposed in which a pseudo-normalization is performed, such that the  $TAC$  is intended to be left in the same order of magnitude as  $Q_{COMP}$ . For this, the first run of 100 iterations is made, and the rounded value of the mean  $TAC$  over  $Q_{COMP}$  is taken, as shown in Eq. 5.

$$FACTOR = (1/100) \sum_{i=1}^{100} (TAC_i / Q_{COMP,i}) \quad 5$$

For this case study, the  $FACTOR$  value is 5,000. The new objective function results from dividing Eq. 4 by the obtained  $FACTOR$  value.

The NSGA-II method was used for multi-objective optimization for 30 generations with 100 individuals with the Augmented Scaling Feature (ASF) decomposition method, a well-known metric in the multi-objective optimization with weights of 0.2, 0.4, 0.5, 0.6 and 0.8 (Wierzbicki., 1980). This study for this method has 3,000 iterations. The ASF method works since the Pareto front is already generated and does not interfere with the NSGA-II method. Therefore, the results from the weights are determined once the algorithm obtains the Pareto front.

The single-objective optimization was performed also using the PSO and GA method included in the PYMOO package for the same optimization variables, but only using the  $TAC$  objective function and, subsequently, the  $Q_{COMP}$  objective function. For the GA case, 30 generations of 100 individuals are used to obtain a sweep of 3,000 iterations. In the case of the PSO algorithm, freedom was given to run the optimization and terminate it if there was no improvement in the objective function after 500 iterations.

The Spyder IDE 5.2.2 (Spyder, 2022) was used to develop and execute the Python code to run the interface with Aspen Plus V12. The overall implemented interface is shown in the schematic presented in Figure 2. The computer used for this optimization has an Intel(R) Xeon(R) W-3235 CPU at 3.30 GHz with 48 GB of RAM.



Figure 2. Optimization-simulation software interface

### 3. Results and Discussion

Table 2 summarizes the time and convergence percentage of executed simulations during the  $TAC$  minimization using Eq. 2 and the results calculating the  $TAC$  using the proposed

pseudo-normalization. For the sake of comparison, the results for the base case by Herrera-Velázquez et al. (2022) are also reported.

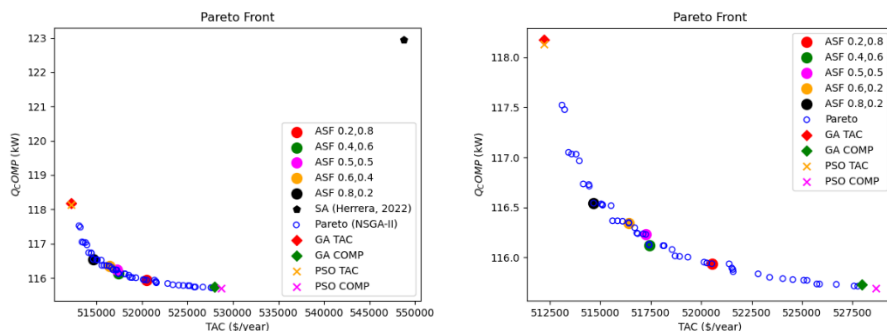
**Table 2.** Summary of efficiency results for the TAC pseudo-normalization proposal.

Optimization Algorithm	Use of pseudo-normalization of TAC	Execution Time (Hours)	Convergence Percentage
Herrera-Velázquez et al. (2022) - SA	NO	3.87	99.99%
PYMOO GA	NO	1.42	99.40%
	YES	1.18	99.30%
PYMOO PSO	NO	0.68	99.22%
	YES	0.69	99.22%
PYMOO NSGA-II	NO	1.19	98.70%
	YES	1.16	99.16%

For both GA and NSGA-II, the pseudo normalization reduces the optimization execution time of 16.90% and 2.52%, respectively. However, for PSO, the time increases 1.47%. The results reported in Table 3 for the TAC are those obtained by the pseudo-normalized TAC, the same for those reported in NSGA-II. In the case of the NSGA-II, only the best result between the values of the ASF weights is reported. Figure 3 shows the obtained Pareto Front, as well as each of the solutions obtained by each method.

**Table 3.** Design variables and optimal operating specifications for the different cases.

Variable	Herrera-Velázquez et al. (2022) SA	PYMOO PSO		PYMOO GA		PYMOO NSGA-II
	TAC	TAC	COMP	TAC	COMP	ASF (0.5,0.5)
$N_{STR}$	25	23	30	23	30	27
$N_{REC}$	22	21	30	21	30	24
$PR$	2.0014	2.0000	2.0000	2.0001	2.0000	2.0000
$Q_{R1}$	121.37	210.94	227.50	232.34	377.16	148.09
$Q_{R2}$	102.89	231.39	101.34	298.05	280.51	199.77
$N_{FSTR1}$	14	20	27	20	27	23
$N_{FSTR2}$	17	22	29	22	29	26
$N_{FRECI}$	4	2	2	2	2	2
$N_{FRECI2}$	14	3	3	3	3	3
$N_{FRECI3}$	15	4	4	4	9	4
TAC (USD/year)	548,786	512,196	528,674	512,188	527,952	517,249
$Q_{COMP}$ (kW)	122.94	118.10	115.69	118.17	115.73	116.23
Iteration for optimum	231 of 3,000	901 of 1,675	2,478 of 3,000	2,729 of 3,000	2,976 of 3,000	2,646 of 3,000
Time (Hours)	3.87	0.69	1.34	1.18	1.28	1.17



**Figure 3.** Pareto front solution for SA included (right) and SA not included (left)

#### 4. Conclusions

It was possible to successfully implement the Python-Aspen Plus interface, applying the PYMOO package for a single-objective (PSO and GA) and multi-objective (NSGA-II) optimization. The pseudo-normalization of the TAC improves, on average, from 2% to 17% of the computation time for the GA and NSGA-II algorithms. The best algorithm for the single-objective optimization is the PSO, while the best solution for the multi-objective problem was the ASF(0.5,0.5). The advantage of PYMOO was the reduction of computation times. They are reduced by more than 50% to those used with SA, which makes it an attractive option for its application to complex processes. The SA algorithm can be used for complex processes where convergence is the priority. This work can be extended without loss of generality to solve intensified chemical processes. It can be of interest for improving current interfaces used in Process Systems Engineering.

#### References

- F. D. Alcántara-Maciél, V. E. Casillas-Céspedes, J. A. López-García, J. Cabrera-Ruiz, C. Ramírez-Marquez, J. R. Alcántara-Ávila, 2022, Economic optimization of a reactive distillation column with multiple reactive sections for silane production, *Computer Aided Chemical Engineering*, 51, 475–480.
- J. Blank, K. Deb, 2020, Pymoo: Multi-Objective Optimization in Python, *IEEE Access*, 8, 89497–89509.
- K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, 2002, A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation*, 2, 182–197.
- J. J. Herrera-Velázquez, F. M. Zavala-Durán, L. A. Chávez-Díaz, J. Cabrera-Ruiz, J. R. Alcántara-Ávila, 2022, Hybrid two-step optimization of internally heat-integrated distillation columns, *Journal of the Taiwan Institute of Chemical Engineers*, 130, 1–9.
- J. Kennedy, R. Eberhart, 1995, Particle swarm optimization, *Proceedings of ICNN'95 - International Conference on Neural Networks*, 4, 1942–1948.
- R. Turton, J. Shaeiwitz, D. Bhattacharyya, W. Whiting, 2018, *Analysis, Synthesis, and Design of Chemical Processes*, 5th Edition. Boston, USA: Pearson.
- A. Wierzbicki, 1980, The Use of Reference Objectives in Multi-objective Optimization. *Multiple Criteria Decision Making Theory and Application*, 177, 468–486.
- Spyder, 2021, <https://www.spyder-ide.org/> (accessed on 26 November 2022).

## Design of multi-component gradient SMBs

Jesper Frandsen<sup>a</sup>, Jakob Kjøbsted Huusom, Krist V. Gernaey<sup>a</sup>, Jens Abildskov<sup>a</sup>

<sup>a</sup>*Dept. of Chemical and Biochemical Engineering, Technical University of Denmark, Søtofts Plads, Building 288 A, 2800 Kgs. Lyngby, Denmark.  
 ja@kt.dtu.dk*

### Abstract

In this work, a generic design method has been developed for the design of isocratic and gradient SMBs for multi-component systems. Based on the general method by (Migliorini et al., 2000b), the method is simplified and extended to enable multi-component systems, gradient SMB operation and any isotherm. The method has been verified through two case studies. For the first case study, an analytical solution is possible and the error using this method is insignificant. For the second case study, the system is solved through numerical simulation and the method is verified. With this method, the design of gradient SMBs can be performed with any isotherm and multi-component systems in a simple way.

**Keywords:** Design methods, Mechanistic Modelling, Gradient SMB

### 1. Introduction

In the biopharmaceutical industry, chromatography is a workhorse unit operation. Chromatography is typically operated in batch mode, although there are many benefits when transitioning to continuous operation such as reduced capital expenditure, higher capacity utilization, a better-controlled process, reduced utility consumption, higher purity and higher yield (Schmidt-Traub et al., 2020). For batch chromatography, applying a gradient can increase productivity and enable complete separation. This concept can also be transferred to a simulated moving bed (SMB), called gradient SMB. The corresponding True Moving Bed (TMB) is shown in Figure 1. The concept is that in section 1 and section 2, where desorption takes place, a low-affinity zone is created to lower the affinity of the components to the resin, whereas in section 3 and 4, a high-affinity zone is created. For liquid chromatography, this can be done by altering the solvent composition, salt concentration, pH or temperature. The most common gradient operation in SMB is based on altering the modifier concentration in the feed and desorbent. Comparing

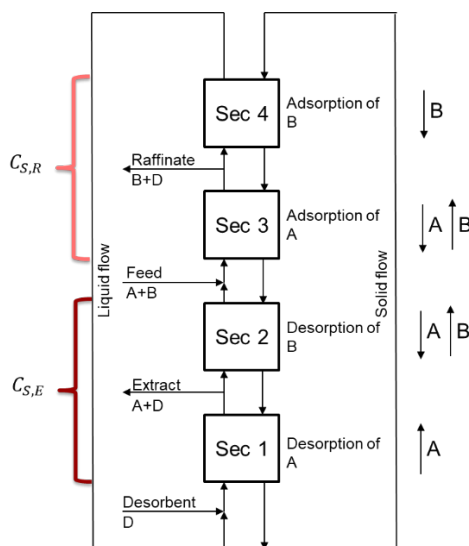


Figure 1: Schematic of a gradient TMB for a binary separation with component A and B and desorbent D. Sec refers to the section,  $C_{S,E}$  and  $C_{S,R}$  are the modifier concentration in extract zone (sections 1-2) and raffinate zone (sections 3-4), respectively.

gradient SMB to isocratic SMB, there are many advantages such as increased productivity, reduced eluent consumption, reduced production cost and increased product concentration both in closed-loop configuration and open-loop configuration (Kim et al., 2017). A particular challenge with implementing continuous chromatography such as gradient SMB is the design of the operation. Modelling chromatography and SMB processes is very complex as there are many phenomena taking place simultaneously which results in a set of complex coupled partial differential equations (Schmidt-Traub et al., 2020). Furthermore, there are multiple design variables which can make the optimal design computationally expensive. For the gradient SMB, there are two additional design variables i.e., the modifier concentration in the feed and desorbent. Therefore, the design of a SMB is typically performed by designing an ideal TMB, shown in Figure 1. A TMB has steady-state solutions whereas a SMB has cyclic steady-state solutions. The ideal TMB steady-state solutions to the quasi-linear first-order partial differential mass balance equations are solved, assuming isotherm equilibrium. Constraints are set up for each section of the TMB. Solving these equations results in operational conditions for each section of the TMB. For the mid-zone (sections 2 and 3 in Figure 1) where feed is added, the operational range for the internal flowrate ratios results in a ‘triangular diagram’ i.e., a region of complete separation, composed of three lines: w-b line, w-r line and r-a line. The feasible flow rate ratios can then be converted to SMB flow rates (Migliorini et al., 2000). Migliorini et al. developed a design method that works for any isotherm, but it is not stated explicitly how the method should be applied for neither gradient SMB nor multi-component systems (Migliorini et al., 2000b). Abel et al. developed a design method for gradient SMB for binary systems with Langmuir isotherms (Abel et al., 2004). Lee et al. developed a SMB design method for multi-component systems with any isotherm, but it is not specified explicitly how to incorporate gradients (Lee et al., 2021). Frederiksen developed a design method for any isotherm, and the method can be applied to gradient SMB but only for binary systems (Frederiksen, 2004). With this, there is a need for a design method that can be applied to multi-component systems, gradient SMB with any isotherm. This work proposes a design method for gradient SMB for multi-component systems and any isotherm. The method is verified with SMB simulations in CADET (Leweke & von Lieres, 2018).

## 2. Methodology

### 2.1 Design method

For designing the SMB, the dimensionless flowrate ratio,  $m_j$ , is often used. It describes the flow rate of the liquid phase per solid flow rate based on the TMB. The dimensionless flow rate ratio is given in eq. (1) (Migliorini et al., 2000b).

$$m_j = \frac{Q_j^* t^* - \epsilon_t V_c}{(1 - \epsilon_t) V_c} \quad (1)$$

Where subscript  $j$  refers to the section,  $Q^*$  is the SMB flowrate,  $t^*$  is the switching time for the SMB,  $\epsilon_t$  is the total porosity and  $V_c$  is the volume of the column.

For multi-component systems, the total number of solutes,  $U$ , are separated into raffinate components, 1 to  $n$ , and extract components,  $p=n+1$  to  $U$ .

For each section, there are constraints on the flow rate ratios that ensure complete separation. For section 1, the constraint on  $m_1$  should ensure that the solid phase is completely regenerated i.e., the most adsorbed component is completely desorbed, mathematically given in eq. (2) (Migliorini et al., 2000b).

$$m_1 \geq m_{1,min} = \frac{\partial q_U}{\partial C_U} \Big|_{C_U=0} \quad (2)$$

Where  $q$  is the solid phase concentration and  $C$  is the mobile phase concentration. The constraint for section 4 ensures that the least retained component does not overflow into section 1. The constraint is given in eq. (3) (Migliorini et al., 2000b).

$$\frac{\epsilon_p}{\epsilon_p - 1} \leq m_4 \leq m_{4,max} = \frac{q_C(C_{1,R})}{C_{1,R}} \quad (3)$$

Where subscript R refers to the raffinate concentration which can be obtained from an overall mass balance (Migliorini et al., 2000b). If having multiple components in the raffinate stream, determining  $m_{4,min}$  in eq. (3) requires setting up overall mass balances and isotherm equations for each component in the raffinate stream. The constraints for sections 2 and 3 ensure that the net flow of extract components, components p to U, is towards the extract outlet and the net flow of raffinate components, components 1 to n, is towards the raffinate outlet. Along the w-b line in the separation region, the minimum  $m_2$  required to desorb the most retained raffinate component, component n, is investigated. Hence, the w-b line is drawn by solving the mass balance over section 2 for component p to U and using the minimum  $m_2$  (Migliorini et al., 2000b), given in equation 4.

$$m_2 = m_{2,min} = \frac{\partial q_n}{\partial C_n} \Big|_{C_n=0} \quad (4a)$$

$$(m_3 - m_2)C_{i,F} = q_{i,2} - m_2 C_{i,2} \text{ for } i \in p \text{ to } U \quad (4b)$$

Where subscripts F and 2 refer to feed and section 2, respectively. In the determination of  $q_n$  in eq. (4), only the extract components and modifier concentration in section 2 ( $C_{S,E}$ ), are considered. Along the r-a line, the maximum  $m_2$  to desorb the least retained extract component, component p, is determined (Migliorini et al., 2000b). Combined with mass balances over section 2 for components p to U, the r-a line is drawn by solving equation 5.

$$m_2 = m_{2,max} = \frac{\partial q_p}{\partial C_p} \quad (5a)$$

$$(m_3 - m_2)C_{i,F} = q_{i,2} - m_2 C_{i,2} \text{ for } i \in p \text{ to } U \quad (5b)$$

where only the extract components and modifier concentration in section 2 are considered in the determination of  $q_p$  in eq. (5). The determination of the w-r line is more complicated as it describes the interaction between sections 2 and 3. Along the w-r line, the maximum  $m_3$  to let the least retained extract component, component p, overflow in section 3 is determined (Migliorini et al., 2000b). This is combined with a mixing balance at the feed node as well as with mass balances for extract components in section 2 and raffinate components in section 3 (Frederiksen, 2004). Whereas Frederiksen used this consideration for binary systems, it is here extended to multi-component systems. The equations for the w-r line are given in equation 6.

$$C_{p,B} = \frac{C_{p,2}m_2 + (m_3 - m_2)C_{p,F}}{m_3} \quad (6a)$$

$$m_3 = \frac{q_{p,3}}{C_{p,\beta}} \quad (6b)$$

$$(m_3 - m_2)C_{i,F} = q_{i,2} - m_2 C_{i,2} \text{ for } i \in p \text{ to } U \quad (6c)$$

$$(m_3 - m_2)C_{j,F} = m_3 C_{j,3} - q_{j,3} \text{ for } j \in 1 \text{ to } n \quad (6d)$$

Where subscript  $\beta$  refers to the point just after the feed node. For the w-r line, this method differs from the method of (Migliorini et al., 2000b) as this method does not involve a numeric integration. The solute concentrations and modifier concentrations used for the determination of the isotherms must be considered carefully and are given in Table 1.

## 2.2 Numeric SMB simulations

To verify the method, SMB simulations were set up in CADET (Leweke & von Lieres, 2018). A SMB setup with four columns in a 1:1:1:1 configuration was considered. The ideal mass balance for the SMB process for each component was used (Schmidt-Traub et al.,

Table 1: Concentrations to use in the isotherms for the determination of the w-r line.

Isotherm	Solute Concentrations	Modifier concentration
$q_{p,3}$	$C_{p+1..U,3} = 0$ $C_{p,3} = C_{p,\beta}$ $C_{1..n} = C_{1..n}$	$C_s = C_{s,3}$
$q_{i,2}$	$C_{p..U,3} = C_{p..U,3}$ $C_{1..n} = 0$	$C_s = C_{s,2}$
$q_{i,3}$	$C_{p..U,3} = 0$ $C_{1..n} = C_{1..n}$	$C_s = C_{s,3}$

2020). To convert the flow rate ratios to SMB flow rates, equation 1 was used together with mass balances over each section. The column parameters  $\epsilon_t$  and  $V_C$  were set to 0.83 and  $7.854e-5 \text{ m}^3$ , and a switching time of 180 seconds was used. The values for  $m_1$  and  $m_4$  were set conservatively such that  $m_1 = 1.5 m_{1,min}$  and  $m_4 = 0.5 m_{4,max}$ . Simulations were run until cyclic steady state was achieved. Complete separation was defined as when purity  $\geq 0.995$ .

## 2.3 Isotherms

To demonstrate that the method is not restricted to a single isotherm, the two most used isotherms, the Langmuir and Steric Mass Action (SMA), were considered (Kumar & Lenhoff, 2020). Assuming equilibrium, the Langmuir and the SMA isotherms are shown in eq. (7) and eq. (8), respectively (Brooks & Cramer, 1992; Schmidt-Traub et al., 2020).

$$q_i = \frac{H_i C_i}{1 + \sum_{h=1}^U K_{L,h} C_h} \quad (7)$$

$$q_i = C_i A_i \left( 1 - \sum_{h=1}^U \frac{q_h}{q_{max,h}} \right)^{v_i} \quad (8)$$

Where  $H$  is the Henry's constant,  $K_L$  is the Langmuir equilibrium constant,  $A = K_{eq} \left( \frac{\Lambda}{c_s} \right)^{v_i}$  is the initial isotherm slope,  $q_{max} = \frac{\Lambda}{\sigma + v}$  is the maximum binding capacity,  $K_{eq}$  is the SMA equilibrium constant,  $\Lambda$  is the ionic capacity,  $C_s$  is the modifier concentration,  $v$  is the characteristic charge and  $\sigma$  is the shielding factor.

### 3. Results and discussion

The design method was applied to two case studies. First, the method was verified by determining the separation regions for a multi-component isocratic SMB with the multi-component Langmuir isotherm. For this case, there is a design method based on the analytical solution of the TMB which the proposed design method was compared to (Migliorini et al., 2000a). The comparison of the two methods for determining the separation regions is shown in Figure 1 with parameters shown in the figure. The comparison shows that the proposed design method determines almost the same separation region as the method of (Migliorini et al., 2000a) with insignificant differences. This verifies that the design method can be used for multi-component systems with the competitive Langmuir isotherm. For the second case study, the separation region is determined for separating lysozyme from ribonuclease A and cytochrome c using gradient SMB with *NaCl* as modifier and parameters from (Meyer et al., 2020). The section 1-2 *NaCl* concentration was set to 200 mol/m<sup>3</sup> and the section 3-4 *NaCl* concentration was set to 300 mol/m<sup>3</sup>. The separation region and the numeric simulation results are shown in Figure 2 where the green dots indicate a complete separation and red pentagons indicate incomplete separation. Figure 2 shows that the separation region determined with the design method agrees well with the numerical SMB simulations. Some of the numerical simulations show complete separation outside the separation region. The deviation between the simulation results and the separation region from the design method is probably because the modifier concentrations fluctuate due to cyclic nature of the SMB, which has previously observed experimentally (Li et al., 2007) and not accounted for in the design method. Additionally, the assumption made by Frederiksen which avoids the numeric integration compared to Migliorini also leads to a small deviation, but this should not be significant (Frederiksen, 2004; Migliorini et al., 2000b). Despite the minor deviation, this verifies that the design method can be used to make an initial design of multi-component gradient SMBs using the SMA isotherm, though the design method is not limited to the SMA or Langmuir isotherms. Design methods based on the ideal mass balance form an initial step towards an in-depth design and optimization. In preparative chromatography, non-ideal phenomena such as mass

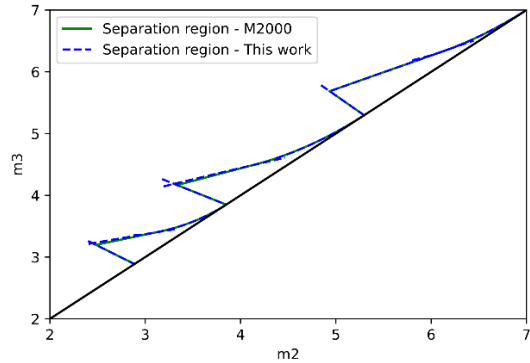


Figure 2: Separation region for  $m_2$  and  $m_3$  with parameters from (Lee et al., 2021) where  $H = [7.062, 5.297, 3.850, 2.888]$ ,  $K_L = [0.0321, 0.0321, 0.0175, 0.0175]$  L/g. M2000 refers to the method in (Migliorini et al., 2000a).

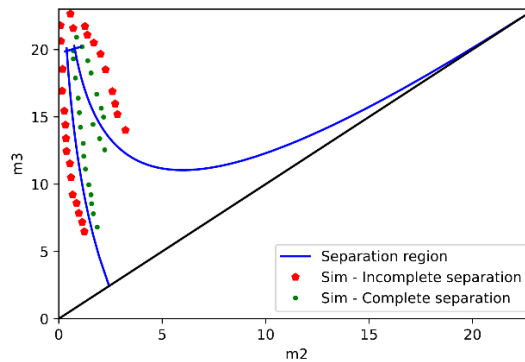


Figure 3: Separation region for  $m_2$  and  $m_3$  where the parameters  $\Lambda = 1200$  mol/m<sup>3</sup>,  $K_{eq} = [35.5e-3, 1.59e-3, 7.70e-3]$ ,  $\nu = [4.7, 5.29, 3.7]$ ,  $\sigma = [11.83, 10.6, 10.00]$  and  $C_{F,i} = 1$  mol/m<sup>3</sup> were used (Meyer et al., 2020).



transfer resistance and axial dispersion can be significant and can cause a spreading of the components, thus complicating the separation (Schmidt-Traub et al., 2020). For an in-depth design and optimization, a more complex model that takes these non-ideal phenomena into account is required. Here, this design method can provide good initial guesses for optimizing the operational flowrates.

#### 4. Conclusion

In this paper, a straightforward and versatile design method for isocratic SMBs and gradient SMBs was proposed. The design method can be applied to any isotherm and multi-component systems. Through two case studies, the design method was verified for the SMA and the Langmuir isotherms. This method can serve as an initial step in designing operational flowrates. For a more in-depth design and optimization, a complex model that takes phenomena such as axial dispersion and mass transfer resistance into account is required.

#### Acknowledgements

This study was financially funded by the Technical University of Denmark.

#### References

- S. Abel, M. Mazzotti, M. Morbidelli, 2004, Solvent gradient operation of simulated moving beds - 2. Langmuir isotherms. *J. Chromatogr. A*, 1026(1–2), 47–55
- C. A. Brooks, S. M. Cramer, 1992, Steric mass-action ion exchange: Displacement profiles and induced salt gradients. *AIChE Journal*, 38(12), 1969–1978
- S. S. Frederiksen, 2004, Computer Aided Development and Optimization of Chromatographic Separations, PhD thesis
- K. M. Kim, M. Lee, S. Kim, F. V. Santos da Silva, A. Siedel-Morgenstern, C. H. Lee, 2017, Advanced Operating Strategies to Extend the Applications of Simulated Moving Bed Chromatography, *Chem. Eng. Technol.* 40(12), 2163–2178
- V. Kumar, A. M. Lenhoff, 2020, Mechanistic Modeling of Preparative Column Chromatography for Biotherapeutics. *Annu. Rev. Chem. Biomol. Eng.*, 11, 235–255
- J. W. Lee, A. Kienle, A. Seidel-Morgenstern, 2021, Numerical Short-Cut Design of Simulated Moving Bed Chromatography for Multicomponent Nonlinear Adsorption Isotherms: Nonstoichiometric Langmuir Model. *Ind. Eng. Chem. Res.*, 60(29), 10753–10763
- S. Leweke, E. von Lieres, 2018, Chromatography Analysis and Design Toolkit (CADET). *Comput. Chem. Eng.*, 113, 274–294
- P. Li, G. Xiu, A. E. Rodrigues, 2007, Proteins Separation and Purification by Salt Gradient Ion-Exchange SMB. *AIChE Journal*, 59(4), 215–228
- K. Meyer, S. Leweke, E. von Lieres, J. K. Huusom, J. Abildskov, 2020, ChromaTech: A discontinuous Galerkin spectral element simulator for preparative liquid chromatography. *Comput. Chem. Eng.*, 141, 107012
- C. Migliorini, M. Mazzotti, M. Morbidelli, 2000a, Design of simulated moving bed multicomponent separations : Langmuir systems. 20, 79–96
- C. Migliorini, M. Mazzotti, M. Morbidelli, 2000b, Robust design of countercurrent adsorption separation processes: 5. Nonconstant selectivity. *AIChE Journal*, 46(7), 1384–1399
- H. Schmidt-Traub, M. Schulte, A. Seidel-Morgenstern, 2020, Preparative Chromatography for Separation of Proteins. In *John Wiley & Sons (3rd edition)*, John Wiley & Sons

# A Method for Finding a Design Space as Linear Combinations of Parameter Ranges for Biopharmaceutical Development

Thomas Oberleitner,<sup>a</sup> Thomas Zahel,<sup>b</sup> Christoph Herwig,<sup>b,c</sup>

<sup>a</sup>*Competence Center CHASE GmbH, Ghegastraße 3, Top 3.2, 1030 Vienna, Austria*

<sup>b</sup>*Körber Pharma Austria GmbH, PAS-X Savvy, Mariahilferstr. 88A/1/9, 1070 Vienna, Austria*

<sup>c</sup>*TU WIEN Research Area Biochemical Engineering, Getreidemarkt 9, 1060 Vienna, Austria*

[christoph.herwig@tuwien.ac.at](mailto:christoph.herwig@tuwien.ac.at)

## Abstract

A biopharmaceutical manufacturer might submit a DS definition as part of the regulatory approval application, in which case process parameter (PP) deviations within this space are not considered a change and do not trigger a regulatory post approval procedure. While several methods exist for finding a DS described by non-linear PP ranges, here we introduce a novel method for finding linear combinations – a DS definition often preferred due to its operational simplicity. We use a numeric optimizer to find the design space with the largest parameter space volume that results in critical quality attribute (CQA) boundaries within acceptance criteria, predicted by a regression model. A precomputed approximation of tolerance intervals is used in inequality constraints to facilitate fast evaluations of this boundary using a single matrix multiplication. Correctness of the method was validated against different ground truths with known design spaces.

**Keywords:** design space; biopharmaceutical development; ICH Q8; numeric optimization; parameter space

## 1. Introduction

The ICH Q8 guideline for pharmaceutical development defines the design space (DS) as “the multidimensional combination and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide assurance of quality” (ICH, 2017). The process parameters (PP) described here are generally identified in the risk assessment or process development phases and are considered critical when sufficient evidence was found that they affect the output of a unit operation, i.e., a critical quality attribute (CQA). A design space is comprised of the ranges of these critical process parameters that result in CQA values within acceptable limits. For the biopharmaceutical manufacturer a DS definition can be submitted as part of the regulatory approval application, in which case PP deviations within this space are not considered a change and therefore do not trigger a regulatory post approval procedure. For operators, the DS constitutes a valuable guideline document for controlling a process. While ICH Q8 does not recommend a specific form or method for describing a DS, it provides examples for non-linear and linear combinations of parameter ranges in the form of contour plots (appendix 2c in guide). Non-linear combinations describe the DS as a set of

rules, or parameter ranges conditioned on other parameters, e.g., “PP1 is allowed to move between -1 and 1 if PP2 is lower than 0.5”. Linear combinations of parameter ranges on the other hand are independent of each other. While the former description generally represents a larger space to operate in and methods for computing it can be found several publications (Kim & Kwak, 2002) (Kusumo, et al., 2020), the latter might be preferred due to its operational simplicity and is the subject of this contribution. Figure 1 shows the different types of design space graphically.

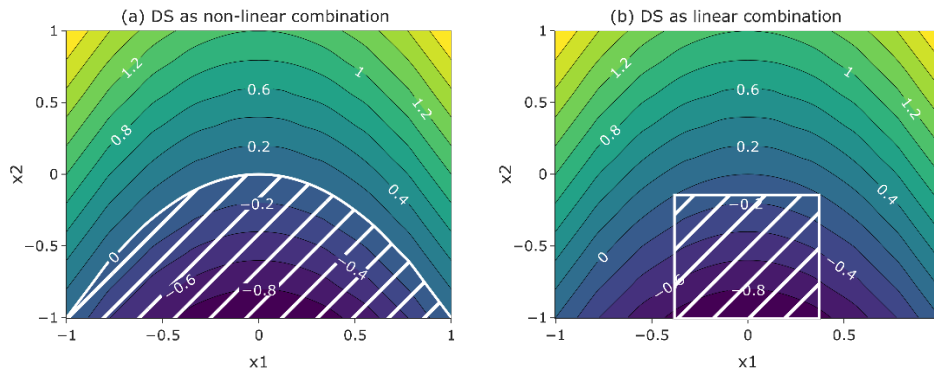


Figure 1: The design space for the function  $f(x_1, x_2) = x_1^2 + x_2$ , where  $f(x_1, x_2) \leq 0$ , shown in the contour plot as a non-linear (a) and linear (b) combination of input parameters  $x_1$  and  $x_2$ , as defined in (ICH, 2017).

So far, definition of a design space does not incorporate any measure of statistical uncertainty, i.e., the contour shown in Figure 1 directly represents the predicted mean CQA values from the model. In the context of biopharmaceutical process validation, we suggest a more conservative approach. To accurately quantify uncertainty inherent in the regression model due to analytical and process variability, we replace predicted CQA values with the upper and lower boundary of a tolerance interval that incorporates nominal levels of both confidence and coverage. These boundaries are then used to validate acceptance limits in a conservative manner.

In this contribution we present a method for finding a design space comprised of linearly independent parameter ranges, or hyper-rectangles when considered geometrically, while treating CQA predictions conservatively.

## 2. Method

### 2.1. Tolerance Interval Approximation

The method proposed here uses a series of numeric optimizers, inequality constraints and approximations specific to polynomial regression models commonly found in process characterization and validation in the biopharmaceutical domain. Such models regress CQAs on PPs and material attributes while considering parameter interactions and quadratic effects. These higher-order effects are relevant to the shape of the response surface and therefore to the optimization problem. While regression model predictions can be computed simply by multiplying the design matrix  $X$  by the vector of learned coefficients  $\hat{\beta}$ , the computation of tolerance intervals can be complex and generally involves multiple evaluations of probability density functions (PDF). For example, one

of the simplest equations for calculating TI boundaries around predictions is the one used in ordinary least-squares (OLS) models:

$$\hat{y} \pm \sigma \sqrt{\frac{(n-p)\chi_{1;\psi}^2 \left(\frac{1}{n_i^*}\right)}{\chi_{n-p;\alpha}^2}} \quad (1)$$

Like all tolerance intervals, this formulation contains nominal parameters for the proportion of the population to be included in the interval,  $\psi$ , as well as the confidence level  $\alpha$ . In addition, critical values for the  $\chi^2$  distribution are calculated as well as the “effective number of observations” vector  $n_i^* = \frac{\hat{\sigma}^2}{se(\hat{y}_i)^2}$ . As the evaluation of TI boundaries is used in several of the inequality constraints described in the following sections, this is considered a bottleneck and is replaced by an approximation. To that end, a central composite design (CCD) containing all possible 3-level combinations PPs (min, max, center point) is created and its columns extended to include all interaction and quadratic effects. A variable selection step eliminates all unnecessary effects, and we are left with a parsimonious model describing the relationship between PPs and TI. Consequently, interval boundaries can be computed by multiplying observations with the vector of TI model coefficients, here denoted as  $\tau_l$  and  $\tau_u$  to distinguish them from the coefficients of the original model,  $\hat{\beta}$ . Note that all steps involved in forming this approximation can be vectorized efficiently using linear algebra libraries, which is why the proposed method is much faster than computing the true TI in each iteration of the optimizer. To illustrate this, results of a performance evaluation are presented in section 3.

### 2.2. Optimization Objective and Constraints

With the TI approximation in place, we can now define the optimization objective:

$$\text{maximize} \quad f(x) = \prod_{i=1}^p x_{p+i} - x_i, \quad x \in \mathbb{R}^{2p} \quad (2)$$

$$\text{subject to} \quad c_1(x), \dots, c_{2^{p+1}+5p+2}(x) \geq 0 \quad (3)$$

Where  $x$  is the vector of parameters to be optimized, comprised of the lower parameter boundaries in the first  $p$  elements and the upper boundaries in the next  $p$  elements, with  $p$  being the number of factors in the model. One can see that (2) simply maximizes the volume of the hypercube spanned by lower and upper parameter ranges.

To meet all requirements for a valid design space, a total of  $2^{p+1} + 5p + 2$  inequality constraints is defined. First, the parameter space to be searched is constrained by ensuring that a parameter’s lower range boundary is smaller than its upper boundary.

$$c_i(x) = x_{p+i} - x_i, \quad i = 1, 2, \dots, p \quad (4)$$

Oftentimes the design space is required to contain each parameter’s setpoint, so that  $x_i \leq s_i \leq x_{p+1}$ . This is expressed as the inequality constrains

$$c_i(x) = s_i - x_i, \quad i = 1, 2, \dots, p \quad (5)$$

$$c_i(x) = x_i - s_i, \quad i = p + 1, \dots, 2p \quad (6)$$

Similarly, optimization of the parameter space should only be performed within the screening range boundaries  $b_l$  and  $b_u$ . As our main optimization algorithm does not support natural boundaries, this is implemented as inequality constraints:

$$c_i(x) = x_i - b_{l,i}, \quad i = 1, 2, \dots, p \quad (7)$$

$$c_i(x) = b_{u,i} - x_i, \quad i = p + 1, \dots, 2p \quad (8)$$

The remaining constraints address the evaluation of the TI. To that end, the approximation generated in the pre-optimization step is used to calculate boundaries around CQA predictions that capture model uncertainty. These boundaries are then compared against the upper and lower acceptance limits  $a_l$  and  $a_u$  in each of the  $2^p$  corner points of the hyperrectangle:

$$c_i(x) = x_* \tau_l - a_l, \quad i = 1, 2, \dots, 2^p \quad (9)$$

$$c_i(x) = a_u - x_* \tau_u, \quad i = 1, 2, \dots, 2^p \quad (10)$$

$x_*$  denotes a corner point, taken from all possible combinations of upper and lower parameter ranges in  $x$ . As the TI approximation is generally not a linear one but at least a second-order polynomial,  $x$  is expanded to  $x_*$  to contain the same quadratic and interaction effects as  $\tau_l$  and  $\tau_u$ .

Evaluating the TI at corners alone does not guarantee a valid design space, as curvature in the response surface might lead to parameter ranges between corner points that exceed acceptance limits. To resolve this problem, a final inequality constraint uses a nested optimization step to find minima and maxima of the TI boundaries inside the hypercube:

$$c(x) = \operatorname{argmin}_x(x_* \tau_l) \tau_l - a_l \quad (11)$$

$$c(x) = a_u - \operatorname{argmax}_x(x_* \tau_u) \tau_u \quad (12)$$

While these last constraints might seem to make (9) and (10) redundant, having both TI checks in place can improve convergence of the optimizer in certain scenarios. Furthermore, one or the other can be deactivated in practice, depending on the type of optimization problem.

### 2.3. Optimization Algorithms

From the problem described in section 2.2 one can infer the prerequisites for the optimization algorithm. To find a design space, we want to minimize a single, scalar-valued function that is subject to inequality constraints as well as boundaries, though this last requirement can also be implemented as inequality constraints, see (7) and (8). COBYLA (Constrained Optimization BY Linear Approximation) was chosen as the main optimization algorithm, as it meets those requirements and, as a gradient-free method, shows reasonable robustness against converging in local minima (Powell, 1994).

For the nested optimization problem in (11) and (12), where boundaries and no inequality constraints are needed, a fast L-BFGS-B algorithm was used to find minima and maxima within the hyperrectangle (Byrd, Lu, Nocedal, & Zhu, 1995).

Due to minor inaccuracies introduced by the TI approximation as well as COBYLA convergence conditions in certain cases, we recommend a final ‘‘refinement step’’ where

COBYLA results are passed to a gradient-based SLSQP optimizer (Kraft, 1988). Here, the same constraints are used as in the main optimization step, except that the true TI calculation is used, as opposed to an approximation. Although iterations are much slower, only a few are needed for convergence as we expect the input to be already near the ideal DS.

### 3. Results

To evaluate efficiency and effectiveness of the proposed method we compared results with a known design space, calculated by a grid-based method that discretizes the problem space and evaluates the TI at each point. This is illustrated in Figure 2.

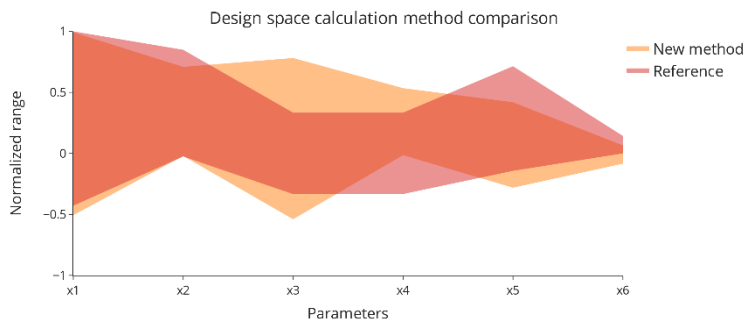


Figure 2: Parameter ranges calculated by both the proposed method and the reference method.

As is to be expected, the volume of the DS is larger in the method based on numeric optimization because it is not constrained by discrete grid points with finite limited resolution. Furthermore, due to parameter interactions, a change in one range can affect other ranges, which overall results in the substantial deviations of parameter ranges shown in in Figure 2. However, the main benefit of the proposed method is improved computing time. Performance evaluation results of an implementation in Python are reported in Figure 3. Due to the exponential increase of the parameter space over the number of parameters added, computing time of a discretization method also grows exponentially, which is not the case with the method based on numerical optimization.

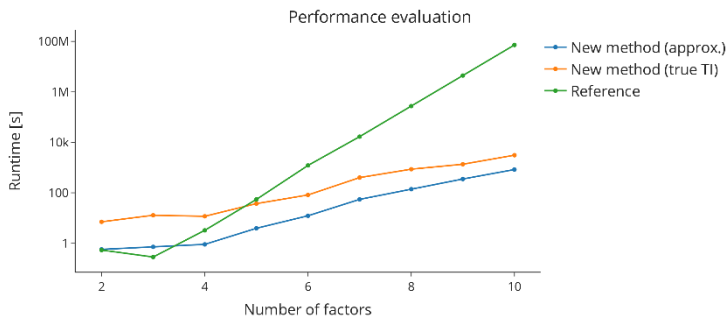


Figure 3: Performance evaluation of the proposed method using both the TI approximation and the true TI, as well as the reference method.

#### 4. Conclusion

In this contribution we outlined a novel method for finding a design space comprised of linear combinations of PP ranges. The relationship between CQA and PPs is represented as a polynomial regression model and its prediction is used to evaluate whether CQAs meet acceptance criteria. Conservative estimation methods are vital in the biopharmaceutical domain, which is why the boundaries of tolerance intervals are used for evaluation of ACs instead of the predicted CQA. As the TI calculation can be complex, an approximation is generated in a pre-optimization step that can be used to calculate the TI by performing a single matrix multiplication. COBYLA is used for the minimization of the main objective, and we suggest refining results with a SLSQP optimizer that uses the true TI instead of an approximation. Performance evaluations show that the proposed method results in design spaces with a larger volume when compared to PP space discretization methods, and that they can be calculated in a fraction of the time. We believe that this approach will facilitate a robust definition of the design space for biopharmaceutical development that reduces patient risk by employing conservative estimators while allowing manufacturers to maximize control ranges. Future work will include a case study where performance is evaluated in a real-world scenario.

#### 5. Acknowledgements

This work was conducted within the COMET Centre CHASE, funded within the COMET – Competence Centers for Excellent Technologies program by the BMK, the BMDW and the Federal Provinces of Upper Austria and Vienna. The COMET program is managed by the Austrian Research Promotion Agency (FFG). The authors acknowledge TU Wien Bibliothek for financial support through its Open Access Funding Program.

#### References

- Byrd, R., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on scientific computing*, 1190-1208.
- ICH. (2017). ICH guideline Q8 (R2) on pharmaceutical development. EMA.
- Kim, I., & Kwak, B. (2002). Design space optimization using a numerical design continuation method. *International Journal for Numerical Methods in Engineering*, 1979-2002.
- Kraft, D. (1988). A software package for sequential quadratic programming. *Forschungsbericht-Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt*.
- Kusumo, K., Gomoescu, L., Paulen, R., García-Muñoz, S., Pantelides, C., Shah, N., & Chachuat, B. (2020). Nested Sampling Strategy for Bayesian Design Space Characterization. *Computer Aided Chemical Engineering*, 1957-1962.
- Powell, M. (1994). A direct search optimization method that models the objective and constraint functions by linear interpolation. In *Advances in optimization and numerical analysis* (pp. 51-67). Springer.
- UCLA Statistical Consulting Group. (2021). *R Library Contrast Coding Systems for categorical variables*. Retrieved from UCLA Statistical Methods and Data Analysis: <https://stats.oarc.ucla.edu/r/library/r-library-contrast-coding-systems-for-categorical-variables/>

# Optimal operation of large gas networks: MILP model and decomposition algorithm

Lavinia Marina Paola Ghilardi<sup>a</sup>, Francesco Casella<sup>b</sup>, Daniele Barbati<sup>c</sup>, Roberto Palazzo<sup>c</sup>, Emanuele Martelli<sup>a,\*</sup>

<sup>a</sup>*Politecnico di Milano, Department of Energy, Via Lambruschini 4, Milano, 20154, Italy*

<sup>b</sup>*Politecnico di Milano, Dipartimento di Elettronica, Informazione e Bioingegneria, Via Ponzio 34/5, Milano, 20133, Italy*

<sup>c</sup>*Snam, San Donato Milanese (MI), 20097, Italy*

\**emanuele.martelli@polimi.it*

## Abstract

As of today, natural gas is one of the most widely deployed energy sources and its transport relies on large-scale infrastructures managed by the expertise of transmission system operators. In this framework, this paper proposes a Mixed-Integer-Linear-Programming model and a decomposition algorithm to optimize the operation of gas networks. The model aims to minimize the CO<sub>2</sub> emissions from compressor stations by optimizing the unit commitment of the gas-turbine-driven compressors, their loads and the dynamic operation of the network. The formulation includes the detailed linearization of the performance maps of the machines, their technical limitations, the dynamic conservation equations of the pipes, and the operating constraints of control valves. On top of this, the algorithm is capable of handling flow reversals in pipes, being suitable also for networks with cyclic topology. The resulting large scale MILP is extremely challenging to solve, and therefore we developed a decomposition algorithm to speed up the computational time. The decomposition algorithm is composed by two MILP levels, where the master level contains a simplified model of the compression stations and of pipe friction, while the lower level is the actual detailed model. The aim of the master level is to remove unused compression stations and to fix the flow direction in pipes in the lower level, thus reducing the size of this second problem. At the end of each bilevel iteration, an integer cut is added to the master problem to explore different combinations of committed stations. The algorithm is effectively tested on the Italian gas network case study, featuring 112 pipes and 9 compressor stations, each composed by 3-5 units.

**Keywords:** Gas network, MILP, decomposition, compression stations, Optimal operation.

## 1. Introduction

Natural gas is currently one of the most widely deployed energy sources and its transport is ensured by large-scale infrastructures, composed by pipelines, valves, and compressor stations. Nowadays, the management of these networks relies on the expertise achieved by the operators. Nevertheless, as the energy transition moves forward, transmission system operators (TSOs) are required to cope with new trends, making the role of optimization tools crucial, as they could help TSOs in their upcoming decisions.



The current literature on gas network optimization covers different optimization purposes, such as minimization of gas compression cost tackled by Liu et al. (2020) and Zlotnik et al. (2015), the stability of operation addressed by Hoppmann-Baum et al. (2021), or gas quality satisfaction. The optimization models can be divided in steady-state and dynamic, where only dynamic models are capable of capturing transient behaviors and considering the pipeline inventory. Nevertheless, very few dynamic approaches allow optimizing the on/off commitment of turbo-compressors, and for large networks they use heuristic rules to solve the optimization problem.

This work addresses the optimal dynamic operation of large gas network with the optimization of on/off commitment of turbo-compressors. The problem is formulated as a Mixed-Integer Non-Linear Program (MINLP), and then it is linearized into a Mixed-Integer Linear Program (MILP). Since also the MILP turns out to be extremely large to be solved in reasonable computational time, we developed a decomposition algorithm capable of finding solution in much shorter time.

## 2. Problem statement

The dynamic model addresses the optimal operation of the gas network so as to minimize the CO<sub>2</sub> emissions related to compression work. The formulation is as follows.

### “Given:

- the topology of the network;
- the performance maps of compressors and of the gas turbines driving them;
- import and consumption profiles across the network;
- air temperature forecasts;

### Determine:

- the operating mode of the compressor stations;
- the unit commitment and the operating point of each gas-turbine-driven compressor (GTDC);
- pressures and mass flow rates across the network;

### so as to minimize the CO<sub>2</sub> emissions and comply with the following constraints:

- conservation laws in pipes;
- technical limits of GT and natural gas compressors;
- compressor stations and control valves discrete operating modes;
- pressure bounds across the network.”

## 3. Optimization model

The MINLP formulation and the linearized MILP are described in Ghilardi et al. (2022). We report hereafter the main of features of the model.

### 3.1. Pipes and control valves

The gas dynamics within the pipes are governed by a system of differential equations depending on time  $t$  and space  $x$  (monodimensional flow). Mass (1), momentum (2) and energy conservations, and gas equation of state (3) depends on gas properties (density  $\rho$ , speed  $u$ , composition  $c$ , temperature  $T$ ), and pipe features (diameter  $D$  and elevation  $z$ ).

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} = 0 \quad (1)$$

$$\frac{\partial(\rho u)}{\partial t} + \frac{\partial(\rho u^2)}{\partial x} + \frac{\partial p}{\partial x} + \rho g \frac{\partial z}{\partial x} + \frac{c_f}{2D} \rho |u| u = 0 \quad (2)$$

$$p = Z(p, T, c) \cdot \rho \frac{R}{MM} T \quad (3)$$

The energy conservation equation can be omitted with the hypothesis of isothermal flow, while the kinetic and the inertial term of gas momentum equation can be neglected due to low speed of gas and large time discretization of the optimization (hourly basis). The differential system is discretized in space with the staggered grid method of Patankar (1980), and in time with backward Euler. As part of the linearization for the MILP model, we adopted an average  $\bar{Z}$  in (3), an average gas density for the friction term in (2), and a piecewise linear (PWL) approximation of the function  $|u|u$ . This second step is responsible for the introduction of the binary variable  $y_D$  related to flow direction.

Control valves (CV) are directed elements that continuously adjust the outlet pressure according to setpoints. They are commonly arranged in parallel with a bypass valve which, on the other hand, is undirected (backward flow can occur), and is discretely regulated, e.g. open/closed. The operation of this subsystem (control-bypass valves) can be mathematically modelled with three discrete modes: “regulating”, “closed” or “reverse-flow”. In “regulating” mode, the gas flows from the inlet to the outlet of the CV reducing its pressure to meet the setpoint; in “closed” mode both the bypass and the CV are closed, and the pressures at their boundaries are thus decoupled; in “reverse-flow” mode the gas flows backward in the bypass valve with a negligible pressure drop.

### 3.2. Compressor stations

Compressor stations consist of a set of GTDCs arranged in parallel, and, by assumption, they can compress NG along one direction only. The station can be “active” if there is at least one running machine, or, vice versa, it can run in “closed” or “bypass” mode.

Each GTDC is characterized by an operating point with a rotational speed  $N$ , inlet pressure  $p_{in}$ , mass flow rate  $w$ , outlet pressure  $p_{out}$ , and fuel consumption  $w_{fuel}$ . The variables  $p_{out}$  (4) and  $w_{fuel}$  (5) are functions of the other variables according to the polynomials  $f_1$  and  $f_2$ , resulting from the best-fit of technical maps. It must be noted that  $w_{fuel}$  depends also on ambient temperature  $T_0$ , due to its effect on GT efficiency.

$$p_{out} = f_1(w, N, p_{in}) \quad (4)$$

$$w_{fuel} = f_2(w, N, p_{in}, T_0) \quad (5)$$

These non-convex polynomials are linearized following the convex-hull approach of Makkonen and Lahdelma (2006), discretizing the operating region into binary sub-regions. Moreover, other linearized constraints are added: the choking and the surge curve for the centrifugal compressor, and the load range for the gas turbine, as a function of  $N$ .

### 3.3. Objective function

The objective function to be minimized is the CO<sub>2</sub> emitted by gas-turbines driving the NG compressors, defined by NG emission factor  $f_{CO_2}$  and fuel consumption  $w_{fuel}$  of each machine  $m$  at time  $t$ . Moreover, a penalty term is added to prevent frequent start-ups and account for the direct and indirect CO<sub>2</sub> emissions emitted during start-up.

$$\min \sum_{t \in \mathcal{T}} \left[ \sum_{m \in \mathcal{M}} f_{CO_2} \cdot w_{fuel}(t, m) + \sum_{m \in \mathcal{M}} f_{start} \cdot y_{start}(t, m) \right] \quad (6)$$

## 4. Decomposition algorithm

The presented MILP problem for the operation of half-day of the whole Italian gas network, with a time resolution of 1 hour and a spatial resolution of pipe finite volumes of 200 km features 4108 binary variables, 24453 continuous variables, 74992 constraints. Gurobi, one of the best available MILP solver, fails at finding a feasible binary solution in 10 hours. To address this issue, we developed a decomposition technique composed by two levels (Figure 1), where the master problem (MP) and the detailed lower-level (DP)

are solved over the iterations  $k$  to progressively refine the solution. The MP presents a simplified model of compressor stations and of pipe friction, and it is launched to determine the set of inactive stations  $\mathcal{S}_0^k$  and the direction of the flow  $y_D^k$  within the pipes  $p$ . Once these specifications are established, the DP can be solved in the resulting reduced space. Meanwhile, an integer cut is added to the MP to discard the combinations of active stations already explored from the next solutions. After a sufficient number of iterations  $\bar{k}$ , the algorithm is terminated and the best incumbent solution is maintained.

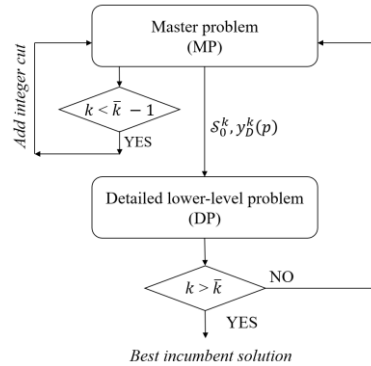


Figure 1. Block diagram of the bilevel decomposition algorithm

#### 4.1. Master problem

The master level MP is a simplification of the original MILP (1)-(6), with regard to the compressor station and the pipe friction model. These two simplifications are intended to reduce the complexity of the problem and especially to decrease the binary variables. Examining first the pipe model, the PWL representation of  $|u|u$  in the friction term is replaced by a straight line  $\bar{u}u$  with a constant slope  $\bar{u}$  for each pipe. This procedure eliminates the binary variable related to flow direction.

As far as the compressor station model is concerned, we defined for the problem MP a single convex hull (CH) for each station operating region (Figure 2), thus removing all individual GTDCs variables (unit commitment and operating point). In this way, the only remaining stations binary variables are those related to their configuration (see 3.3).

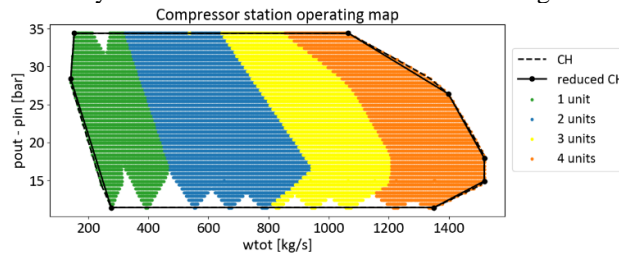


Figure 2. Operating map of a compressor station featuring 4 units projected into space ( $w_{tot}$ ,  $p_{out}$ - $p_{in}$ ) with its original convex hull (dashed), and reduced vertices convex hull (continuous).

The above cited convex hull is computed off-line by a series of (CHP) subproblems, one for each value of ambient temperature  $T_0$  and station operating point ( $p_{in}$ ,  $p_{out}$ ,  $w_{tot}$ ). Each subproblem (CHP) is obtained by applying the original MILP (4)-(6) to a single station, and by imposing, as boundary conditions, the coordinates of the operating point at stake. The (CHP) subproblem establishes whether the analyzed operating point is feasible and, if the response is positive, it computes the minimal fuel consumption  $w_{fuel,tot}$ . Therefore, the convex hull of the feasible points only is established and the

number of vertices of the resulting polyhedron is reduced to strike a balance between the number of variables and the accuracy the approximation. The graphical representation of this procedure into the bidimensional space ( $w_{\text{tot}}, p_{\text{out}} - p_{\text{in}}$ ) is provided by Figure 2 for a fixed  $p_{\text{in}}$  of 50 bar. Once the final CH of the operating region of each station is assessed, the compressor stations variables in MP – when the station is deemed as “active” – are computed as a convex combination of the vertices  $v$  of the polyhedron.

#### 4.2. Detailed lower-level problem and integer cuts

The lower level detailed problem (DP) exhibits all the constraints and variables described in (1)-(6), but it has a reduced size compared to the original problem. The DP defines the optimal unit commitment and load of the GTDCs within the resulting reduced space by the MP. Nevertheless, due to the approximations of this MP model, the induced choice of committable stations on DP might be sub-optimal. To mitigate this effect, at the end of each iteration  $k$ , an integer cut is progressively added to the MP to exclude the previous explored solutions. This cut (7) guarantees that, in the next iterations, at least one of the previously uncommitted stations ( $s \in \mathcal{S}_0^k$ ) is activated within the time horizon  $\mathcal{T}$ .

For the practical reasons, given the long computational time, the iterative algorithm stops when the best found solution is not improved for three consecutive iterations.

## 5. Results

The presented algorithm is applied to a simplified version of the Italian network (Figure 3), still featuring a very complex topology with many loops, 112 pipes, 20 GTDCs arranged in 9 stations, and 16 control valves.

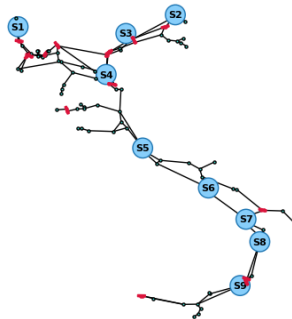


Figure 3. Italian network case study: compressor stations (blue), control valves (red) and pipes (black)

The MILP has been modelled with Pyomo and solved with Gurobi till a gap of 5% on a laptop computer with Intel i7 CPU and 32 GB RAM. The model is applied for the operation of 12 hours in a test scenario featuring a large number of uncommitted stations and backward flow in many pipes. For the practical implementation of the algorithm, the computational time must be limited to about 6 hours. Thus, the iterative MP-DP algorithm stops when the solution is not improved for 3 consecutive iterations.

With a time resolution of 1 hour, the first MP features 14781 continuous variables, 442 binary variables, 18927 constraints, while the DP features 8436, 696 and 74992, respectively. Table 1 depicts for each iteration  $k$  the computational time, the stations  $\mathcal{S}_1^k$  activated by the MP, and the objective function of the DP.

In all the iterations, the compressor station S2 (at the North-East boundary), the station S9 (near the South border), and S4 are always activated, while different choices of stations along the South-Central pipeline are explored using the integer cut. Among these

solutions, the optimal one is found in the first iteration, as shown in Table 1. Further investigations are needed to define the optimal number of iterations across different scenarios, finding a compromise between optimality and computational load. Nevertheless, it must be underlined that the overall computational time can be reduced in any case by adopting a coarser time grid in MP (without impacting the time resolution of the final solution), and by solving each  $MP^{k+1}$  in parallel with  $DP^k$  (Figure 1).

<b>k</b>	<b>MP time</b>	<b>DP time</b>	<b><math>S_1^k</math></b>	<b>Objective function</b>
1	1135 s	2125 s	S2, S4, S7, S9	586
2	1094 s	2580 s	S2, S4, S8, S9	646
3	1830 s	5811 s	S2, S4, S6, S9	659
4	1910 s	4489 s	S2, S4, S6, S8, S9	614

Table 1. Results of the Italian network case study. The table shows the computational time required to the master problem (MP), the detailed-lower level problem (DP), the activated stations  $S_1^k$  by the MP, the optimal CO<sub>2</sub> emissions [tonCO<sub>2</sub>].

## 6. Conclusions

This work presents a Mixed Integer Linear Programming algorithmic framework for optimizing the operation of large gas networks including commitment decisions of the turbo-compressors. The model is characterized by the dynamic equations of the network and the performance maps of gas-turbine-driven-compressors. Since the detailed MILP model is too large also for the best available MILP solvers, a decomposition algorithm has been developed. The simplified master problem is firstly launched to define the set of inactive stations and the flow direction within the pipes. Then, the detailed lower-level problem is solved in the resulting reduced space, providing a more accurate solution. Integer cuts are progressively added to the master problem to explore different combinations of active stations, and mitigate the possible sub-optimality of the decomposition. The proposed algorithm is applied to half-day operation of the Italian gas network, which consists of 112 pipes, 9 compressor stations, each one containing 3-5 gas-turbine-driven compressors. For the case study, the first iteration of the bilevel model is solved in less than 1 hour on a laptop computer, and proved to be the best solution across all explored iterations. These results suggest that the algorithm could be applied in decision support systems of transmission system operators. Further development of the research can apply the algorithm to different case studies to assess the optimal number of iterations to be adopted over different scenarios, and to develop some heuristics to speed up the solution.

## References

- L.M.P. Ghilardi et al., 2022, A MILP approach for the operational optimization of gas networks, IFAC-PapersOnLine, 55, 321–326.
- K. Hoppmann-Baum et al. 2021., Optimal Operation of Transient Gas Transport Networks, Optim. Eng., 22, 735–781.
- K. Liu et al., 2020, Dynamic optimization of natural gas pipeline networks with demand and composition uncertainty. Chem. Eng. Sci., 215, 115449.
- S. Makkonen, R. Lahdelma, 2006, Non-convex power plant modelling in energy optimisation, Eur. J. Oper. Res. 171, 1113–1126.
- S. V. Patankar, 1980. Numerical Heat Transfer and Fluid Flow, 1st ed. CRC Press .
- A. Zlotnik et al., 2015, Optimal control of transient flow in natural gas networks, Proc. IEEE Conf. Decis. Control 54rd IEEE, 4563–4570.
- GUROBI. <http://www.gurobi.com>.

# An Integrated Framework for the Geometric Optimization of Bio-catalytic Microreactor

Yueheng Han<sup>a</sup>, Runze Liang<sup>a</sup>, Zhihong Yuan<sup>a,\*</sup>

<sup>a</sup>*Department of Chemical Engineering, Tsinghua University, Beijing100084, China*  
*Corresponding authors: zhihongyuan@mail.tsinghua.edu.cn*

## Abstract

Biodiesel is an ideal substitute for traditional diesel under the requirements of sustainable and low-carbon issues because of its low-temperature chamber gas emission ratio and good adaptability to the characteristics of existing vehicle engines. However, biodiesel has yet to be fully industrialized due to the traditional batch production mode. Based on the rapid development of microreactor technology that can be used for continuous production, this paper studies the influence of shape parameters of a T-shaped microreactor on biodiesel yield through the enzymatic synthesis of biodiesel from biodiesel castor oil and methanol to prepare fatty acid methyl ester. An automatic simulation optimization platform is thus established based on computational fluid dynamics simulation and derivative-free optimization algorithms. The optimization work makes biodiesel yield increase by 86% compared with the initial result.

**Keywords:** Derivative free trust region algorithm, Microreactor, Geometric optimization, Automatic CFD simulation

## 1. Introduction

With the rapid global increase of market competition along with the requirements on safety and environmental issues, the search for sustainable chemical production systems is becoming increasingly important. Biodiesel has been widely regarded as a green alternative to fossil counterparts because of its low-temperature chamber gas emission ratio and good adaptability to the characteristics of existing vehicle engines. However, traditional biodiesel production is mainly carried out in the batch reactor, which takes a long time and consumes much energy and thus hinders its further development and potential large-scale industrial applications.

Benefiting from additive manufacturing and flow chemistry, the new continuous biodiesel production equipment represented by microreactors has started getting attention from academia and industry because of its excellent mass and heat transfer characteristics. However, the selection of microreactor geometric parameters is still determined by experience, lacking a systematic optimization approach (Liang, 2020). Based on the Derivative-Free Trust Region (DFTR) algorithm (Charles A et al., 2017) and the Computational Fluid Dynamics (CFD) simulation, this paper proposes an integrated framework for optimally discovering the geometric parameters of the T-shaped microchannel reactor. In general, the proposed framework can automatically run the CFD simulator based on the offered parameters by DFTR. Accordingly, the information from the CFD simulator feeds DFTR for seeking next-round parameters.

The proposed framework is applied to a biodiesel production example where triglyceride is chosen as the feedstock (Zarejousheghani et al., 2016). Facing the mixing problem caused by the two-phase laminar flow boundary in the reactor channel and the low diffusion coefficient of reactants, this example assumes that the laminar flow boundary can be broken by adding a groove to improve the mixing effect. The main objective is to maximize the yield of biodiesel. The results indicate the feasibility and effectiveness of the proposed framework. The yield of biodiesel associated with the designed microreactor is increased by 86% compared to the existing ones. The designed microreactor is now being fabricated to get further experimentally validated.

## 2. Reaction

As shown in Fig.1, the production method of enzyme-catalyzed reaction of triglyceride (castor oil) and methanol in a microreactor was selected, and the amount of fatty acid methyl ester obtained from the reaction was regarded as the final biodiesel production.

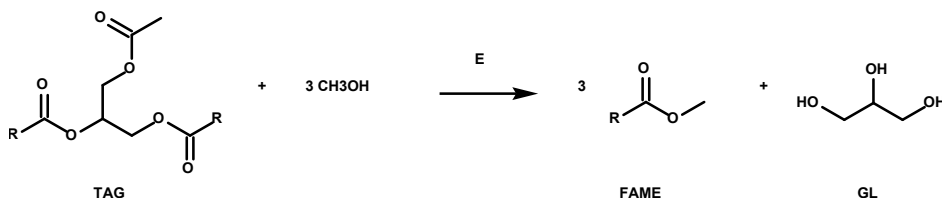


Figure 1. Reaction-producing biodiesel from triglyceride

Fig.2 illustrates that the reaction process can be explained by a ping-pong mechanism (Zarejousheghani et al., 2016): (1) The enzyme attacks triglycerides, producing and activating enzyme triglyceride conjugates. (2) The conjugates rearrange to form the enzyme and free fatty acid chain complex and release glycerol. (3) Alcohols link with enzymes in the activated state to release the final product, fatty acid methyl ester, and free enzymes.

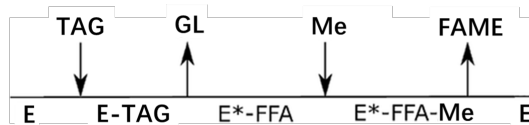


Figure 2. Reaction mechanism

Experimental data (Zarejousheghani et al., 2016) shows that the inhibition effect of methanol in the production process can be ignored so that the reaction mechanism can be characterized approximately by the ping-pong mechanism without inhibition.

$$r = \frac{k[E_0][TAG][Me]}{K_{mTG}[Me] + K_{mA}[TAG] + [TG][Me]} \quad (1)$$

$[E_0]$  represents the enzyme concentration in the reaction system,  $[TAG]$  and  $[Me]$  are the molar concentrations of triglyceride and methanol respectively. The meanings of other parameters and their estimated values are shown in Table 1.

Table 1. Reaction kinetics parameters

Parameter	Meaning	Value
-----------	---------	-------

$k(\text{mmol} \cdot \text{ml}^{-1} \cdot \text{s}^{-1})$	Maximum reaction rate	0.0208
$K_{mTAG}(\text{mmol} \cdot \text{ml}^{-1})$	Michaelis constant	10.54
$K_{mMe}(\text{mmol} \cdot \text{ml}^{-1})$	Michaelis constant	10900.42
$K_{iMe}(\text{mmol} \cdot \text{ml}^{-1})$	Inhibition constant	3976.48

In addition to reaction kinetics, the diffusion coefficient between reactants in the reaction system is also an indispensable parameter in the actual modeling process. In this work, the diffusion coefficient of alcohol in the oil phase is estimated by Wilke Chang's empirical formula. The diffusion coefficient of methanol to triglyceride is approximately  $1.8 \times 10^{-10} \text{ m}^2/\text{s}$ , the diffusion coefficient between lipase and other phases is approximately  $5 \times 10^{-11} \text{ m}^2/\text{s}$ , the diffusion coefficient between other substances is approximately  $10^{-9} \text{ m}^2/\text{s}$ .

### 3. Reactor

#### 3.1. Initial reactor

The traditional T-type microreactor is selected as the reactor for biodiesel synthesis. The raw materials enter from both ends of the reactor respectively, and the reaction and mixing are carried out in the reaction channel.

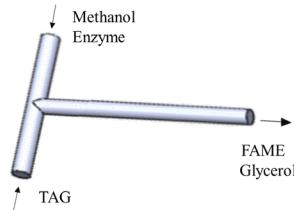


Figure 3. Initial T-shaped Microreactor

Table 2. Geometric parameters and feeding conditions of initial T-type reactor

Variable	Value
Pipe diameter	0.5 mm
Pipe length	10 mm
Branch length before the two-phase intersection	3 mm
Triglyceride feed flow rate	0.1 mm/s
Methanol feed flow rate	0.6 mm/s
The mass fraction of enzyme in the methanol feed solution	1%

#### 3.2. Reactor optimization

Due to the low interphase diffusion coefficient, the contact of reactants forms a laminar boundary, which hinders the diffusion and reaction process of substances. It is feasible to break the original laminar flow by adding grooves to improve the mixing condition. The optimization problem in this work is established by taking the geometric parameters related to the grooves as optimization variables.



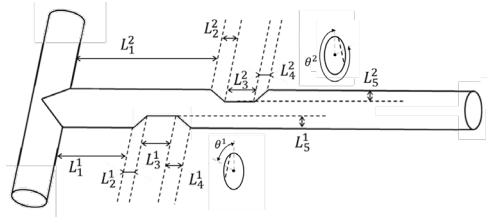


Figure 4. Geometric parameters of the groove of T-type reactor

Constrain the range of each variable: (1)The starting position of the groove is required to be 0.5 mm away from the intersection of two phases; (2)The length of the inclined front section, horizontal section, and rear inclined section of the groove shall be between 0.01 mm and 2 mm; (3)The depth of the two grooves is required to be between 0.05 mm and 0.45 mm to avoid the appearance of too small flow passage cross-section; (4)The deviation angle of the groove centerline is set between 0 and  $2\pi$ ; (5) To simplify the model, the second groove is required to be generated after the first groove.

$$\begin{cases} 0.5 \leq L_1^1 \leq 10, & 0.5 \leq L_1^2 \leq 10, & 0.01 \leq L_2^1 \leq 2, & 0.01 \leq L_2^2 \leq 2 \\ 0.01 \leq L_3^1 \leq 2, & 0.01 \leq L_3^2 \leq 2, & 0.01 \leq L_4^1 \leq 2, & 0.01 \leq L_4^2 \leq 2 \\ 0.05 \leq L_5^1 \leq 0.45, & 0.05 \leq L_5^2 \leq 0.45, & 0 < \theta^1 \leq 2\pi, & 0 < \theta^2 \leq 2\pi \\ L_1^2 + L_2^2 + L_3^2 + L_4^2 \leq 10, & L_1^1 + L_2^1 + L_3^1 + L_4^1 < L_1^2 \end{cases} \quad (2)$$

Table 3. Geometric parameters of the groove of T-type reactor

Variable	Meaning
$L_1^1/mm$	Distance from the beginning of groove 1 to the entrance
$L_2^1/mm$	Length of the front section of the inclined section of groove 1
$L_3^1/mm$	Length of the horizontal section of groove 1
$L_4^1/mm$	Length of rear section inclined section of groove 1
$L_5^1/mm$	Depth of groove 1
$\theta^1$	Deviation angle of groove 1 centerline
$L_1^2/mm$	Distance from the beginning of groove 2 to the entrance
$L_2^2/mm$	Length of the front section of the inclined section of groove 2
$L_3^2/mm$	Length of the horizontal section of groove 2
$L_4^2/mm$	Length of rear section inclined section of groove 1
$L_5^2/mm$	Depth of groove 2
$\theta^2$	Deviation angle of groove 2 centerline

### 4. Integrated framework

#### 4.1. Computational fluid dynamics modeling

Empirical formulas in the macro process cannot be directly applied to a micro-scale reactor. The numerical simulation method of solving the fundamental equations in fluid mechanics, computational fluid dynamics (CFD), is used to model the reaction process in the microreactor. Ansys series software was used to complete the modeling process of geometric modeling, mesh generation, and computational fluid dynamics.

#### 4.2. Derivative-free trust-region algorithm

CFD can only obtain the output results under given input conditions, so the optimization process needs an optimization method that does not rely on derivative information. We use a derivative-free trust-region algorithm to solve the problem.

Under the variables and constraints given in section 3.2 to be optimized, the constraint problem  $\max f(x)$  is constructed to maximize biodiesel yield. Table 4 shows the algorithm framework:

Table 4. The framework of the derivative-free trust-region algorithm

1. Initial point  $x_0$ , trust region radius  $\Delta_0$
2. Establish approximate function model  $q(x_k + s) = f(x_k) + g^T s + \frac{1}{2} s^T B_k s$
3. Solve the problem  $\max q(x_k + s)$  to obtain  $s$
4. Calculate the ratio of the actual objective function descent to the approximate function model descent
5. Update the trust region radius according to the ratio
6. Output the final result if the termination conditions are met. If not, return to step 2.

#### 4.3. Framework

A fully automatic simulation optimization platform is established based on the modeling and optimization method. The framework is shown in Figure 5.

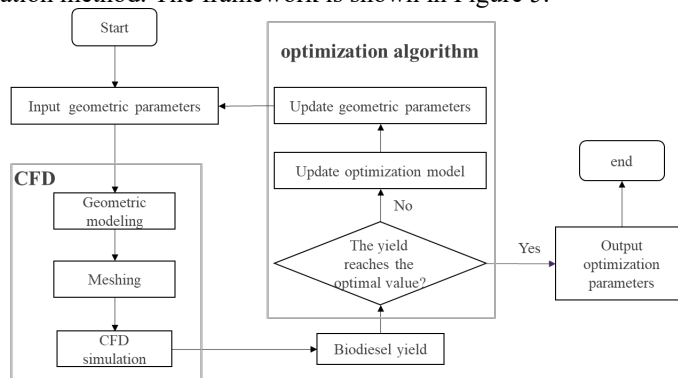


Figure 5. Operation framework of simulation optimization platform

## 5. Analysis and discussion

With the goal of maximizing the yield of biodiesel, this work calculates the conversion rate of triglyceride at the outlet interface obtained by CFD simulation as the objective function:

$$\max f = \frac{\sum c_{FAME}}{\sum c_{FAME} + 3 \sum c_{TAG}} \quad (3)$$

Forty feasible groove shapes were selected by Monte Carlo random sampling, with the corresponding diesel yield obtained through CFD simulations. The quality of grids generated in the automatic CFD simulation process is higher than 0.4, the quantity of grids is about 500000, and the calculation accuracy error is about 2%. The groove shape with the highest biodiesel yield was selected as the initial point, and the biodiesel yield was 4.86%.

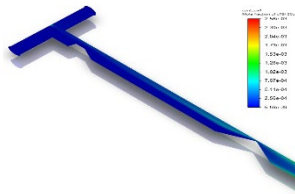


Figure 6. The product distribution of the initial reactor

After the normalization of variables, the initial trust region radius is selected as 0.05, and the trust region radius less than 0.001 is taken as the termination condition of the trust region algorithm.

As shown in Figure 5 above, according to the geometric parameters and trust region, select the points in the trust region to carry out CFD simulation to establish the proxy model, calculate the proxy model to obtain the best geometric parameters, and carry out CFD simulation to get the calculation results. If the results do not meet the convergence conditions, update the radius of the trust region and the proxy model, and repeat the process. If the result meets the convergence condition, output the optimized geometric parameters and terminate the operation. Each iteration requires about 91 function value calculations. After about 1000 times function value calculation, iterative convergence is achieved. The final yield of biodiesel is 9.03%.

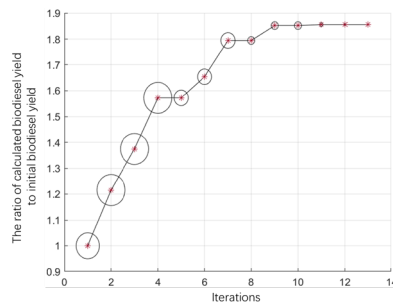


Figure 7. The ratio of optimization result to initial result varies with the number of iterations

The circle size represents the trust region radius size.

Figure 8 shows the geometry of the microreactor under the final optimization results and the product distribution in the reactor at this time. The simulation results show that the

optimization process can effectively improve the two-phase mixing condition and promote the reaction.

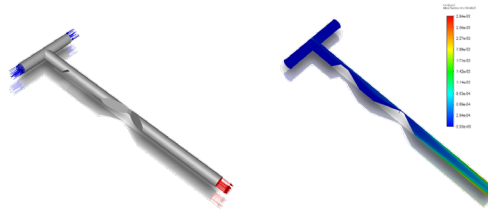


Figure 8. The geometric shape and product distribution of the optimized microreactor

Table 5. Geometric parameters of the optimized microreactor

Parameter	Value	Parameter	Value
$L_1^1/mm$	1.266	$L_1^2/mm$	3.754
$L_2^1/mm$	0.328	$L_2^2/mm$	1.168
$L_3^1/mm$	0.437	$L_3^2/mm$	0.187
$L_4^1/mm$	1.371	$L_4^2/mm$	0.998
$L_5^1/mm$	0.297	$L_5^2/mm$	0.450
$\theta^1$	141.5°	$\theta^2$	257.7°

## 6. Summary

An automatic simulation optimization platform has been established, and the biodiesel yield has been successfully improved using a derivative-free trust-region algorithm and CFD. The optimization platform can be extended to more extensive geometric shape optimization research in the future.

## References

- Liang RZ and Yuan ZH, (2020). Computational Shape Optimization of Microreactors based on CFD Simulation and Surrogate Model-driven Optimization[C]. The 30<sup>th</sup> European Symposium on Computer Aided Process Engineering, PTS A-C, 48:925-930
- Zarejousheghani F, Kariminia HR, Khorasheh F (2016). Kinetic modeling of enzymatic biodiesel production from castor oil: Temperature dependence of the Ping Pong parameters [J]. Canadian Journal of Chemical Engineering, 94(3):512-7
- Charles Audet, Warren Hare, et al., (2017). Derivative-Free and Blackbox Optimization[M]. Springer International Publishing.



# A genetic algorithm-active learning approach in energy systems optimization

Rafael Castro Amoedo,<sup>a</sup> Julia Granacher,<sup>a</sup> Yi Zhao,<sup>a</sup> and François Maréchal,<sup>a</sup>

<sup>a</sup> *Industrial Process and Energy Systems Engineering group, École Polytechnique Fédérale de Lausanne, Rue de l'Industrie 17, 1950 Sion, Switzerland*  
*rafael.amoedo@epfl.ch*

## Abstract

Recently, active learning methods have been used to solve large and complex problems. The underlying principle is the construction of computationally cheap surrogate models, which are able to mimic the original problem with a high-level of accuracy. In complex energy systems, (non-linear) simulation models (e.g. generated using Aspen Plus) have grown larger. In the field of process system engineering, these models need to be repeatedly used to optimize operating conditions. The use of surrogate models reduces the required optimization time. However, the quality and the way to obtain these surrogates is crucial. In this work an active learning approach is coupled with a genetic algorithm, that ensures the selection of points that are relevant from the optimization perspective. Nine different surrogate methods are tested. Our results show that artificial neural networks and random forest are the best performing algorithms, with mean square error up to two orders of magnitude better than other more simplistic approaches. A biogas separation case-study is used to validate the approach.

**Keywords:** Active Learning, Surrogate models, Genetic algorithm, optimization, superstructure

## 1. Introduction and state-of-the-art

In the field of Process system engineering (PSE), superstructure-based optimization is a common approach for both synthesis and design stages<sup>1</sup>. The backbone of a superstructure is a set of interconnected units (or blocks), that combined are able to generate solutions that meet desired specifications and constraints. Such units are often modeled as black-boxes (BB); thus their accuracy is critical for generating insightful results. These BB models try to capture, via a mathematical formulation, the relation between inputs and outputs. This can be a challenging task, especially if complex mathematical models (e.g., CFD) or large flowsheeting diagrams (e.g. Aspen or Belsim Vali) are required to describe conversion processes. Lengthy computational times and converge issues are among the main concerns with such modelling strategies.

Surrogate models (or metamodels) are a way to by-pass the limitations of complex simulations, while keeping a high-level of accuracy. This type of supervised machine learning is able to explore features between input and output data in a computationally cheap way. They can then be used in superstructure-based optimization, allowing one to obtain fast and reliable results. In the field of PSE, the most used surrogate strategies are Kriging models (or Gaussian processes) and Artificial Neural Networks (ANN)<sup>2</sup>, justified by the built-in error measure of the former, and the adaptability to practically any type of data structure of the latter. Examples of such applications can be found for large scale superstructure-based optimization<sup>3</sup> or even to replace state-of-the-art solvers in deterministic global optimization<sup>4</sup>.

Building surrogates can be a simple task if large amounts of useful and reliable data are available. Unfortunately, this is often not the case. One technique is to run the simulation model for an extremely large number of points. This is the domain of Deep Learning (DL). Provided enough time and results, DL can build a very accurate (and cheap to run) surrogate model. However, obtaining such data can be (computationally) expensive. Active learning (AL) is another domain of machine learning that leverages the capacity of an algorithm to learn and ‘smartly’ decide which sample points need to be simulated. Points are selected in regions for which the surrogate level of accuracy is the lowest, and therefore running the expensive simulation program will bring the largest amount of information. The algorithm decides based on the evaluation of a specific function (called learning function)<sup>5</sup>.

An AL approach is by itself a good improvement compared with greedy DL strategies. For optimization purposes, one would like to have an extremely accurate surrogate model in the points that are assessed during the optimization stage - rather than a good overall method, that underperforms in the most critical points. In this regard, a way to cover the most ‘interesting’ space domains for optimization is by using a genetic algorithm (GA) - a nature-inspired meta-heuristic. Such evolutionary strategy has been extensively used for optimization. A review on its properties, advantages and limitations is given in<sup>6</sup>.

This work departs from the literature by coupling two ‘smart’ approaches with the goal of building surrogate models. A GA complements the use of an AL approach, particularly by guiding the surrogate building strategy in the direction that is most interesting from the optimization perspective. The method was tested and validated in a superstructure-based optimization approach for the chemical-absorption separation of biogas.

## 2. Methodology

### 2.1. Mathematical formulation and overview

The goal of the GA-AL approach is to create the function  $f$  that is able to accurately predict the output(s) of interest. An overview of the methodology is shown in Figure 1. There are four macro blocks, explored below.

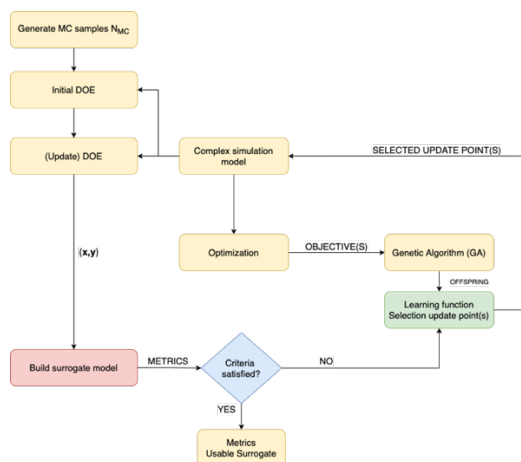


Figure 1 – Simplified block diagram of the GA-AL approach.

#### 2.1.1. Design of experiments (DOE)

The DOE is the section responsible for the initial number of (expensive) simulation runs and the connection between the surrogate building framework and the simulation code. The initial number of samples is a topic of debate. We have adopted the  $\max\{n_s + 2, 12\}$ , in which  $n_s$  is the number of parameters (also known as

dimensions), following the suggestion of Hong et al<sup>7</sup>.

### 2.1.2. Surrogate methods

Table 1 summarizes the list of methods used, comprising simple ones (such as least squares and inverse distance weighting), Kriging-based approaches and more advanced ones such as ANN or Random forests. For this stage, ML python libraries were used, together with the python package ‘smt’ based on the work of Bouhlel et al.<sup>8</sup>, where the pros and cons of each strategy are also discussed and the complete mathematical formulation provided.

Table 1 - Surrogate methods used

Acronym	Surrogate method
LS	Least squares
KRG	Kriging
KPLS	Kriging + partial least squares (PLS)
KPLSK	2-step Kriging + PLS
GEKPLS	Gradient-enhanced Kriging + PLS
IDW	Inverse distance weightings
RBF	Radial basis function
ANN	Artificial Neural Network
RF	Random Forest

### 2.1.3. Learning function

The learning function decides which point ( $\mathbf{x}$ ) (which is in reality a  $n_s$ -dimension vector) needs to be integrated in the DOE set. An adaptation of the potential risk function (PRF) as described in Hong et al.<sup>9</sup> is used (eq. 1). It consists of an exploration ( $e_r$ ) and exploitation ( $e_i$  in blue) approach, that is weighted by the parameter  $\tau$ . The numerator is the distance between any future point and the DOE points, whereas the denominator is the maximum distance within the current DOE points (of size  $m$ ). For the exploitation,  $y(x_i)$  is the real simulation result on the DOE points,  $\xi$  is the limit value (equivalent to the limit-state value) and  $r(x)$  is given by eq. 2, where  $\hat{y}$  represents the surrogate model answer.

$$e_{PRF}(\mathbf{x}) = e_r(\mathbf{x}) \cdot e_i(\mathbf{x}) = \frac{\min_{x_i \in \mathbf{x}} |\mathbf{x}_i - \mathbf{x}|}{\max_{x_i, x_j \in \mathbf{x}} |\mathbf{x}_i - \mathbf{x}_j|} \cdot \exp\left(-\tau \frac{r(\mathbf{x})}{\max_{x_i \in \mathbf{x}} y(x_i) - \xi}\right), i, j = 1, 2, \dots, m, \tau > 0 \quad (1)$$

$$r(\mathbf{x}) = \begin{cases} \hat{y}(\mathbf{x}) - \xi, & \hat{y}(\mathbf{x}) \geq \xi \\ \xi - \hat{y}(\mathbf{x}), & \hat{y}(\mathbf{x}) < \xi \end{cases} \quad (2)$$

### 2.1.4. Stopping criteria

The stopping criteria (eq.3) defines the level at which the surrogate model is good enough to be used without further refinement. The criteria is a convergence based one, adapted from Hong et al.<sup>9</sup>;  $\hat{\mu}_i$  and  $\hat{\sigma}_i$  are, respectively, the estimated expected value and standard deviation, obtained using the surrogate model, in the GA offspring points. The threshold ( $\varepsilon_{th}$ ) is set to  $10^{-5}$ . In addition, a stopping criterion on the number of runs (100 runs) is included to avoid extremely lengthy surrogate training.

$$\varepsilon = \max\left(\frac{\hat{\sigma}_i}{\hat{\mu}_i}\right) \leq \varepsilon_{th} \quad \forall i \in \text{Offspring} \quad (3)$$



## 2.2. Case-study

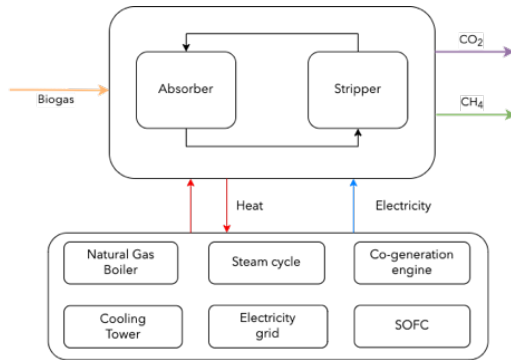


Figure 2 – Superstructure-based approach for biogas separation. The set of technologies was modeled according to Castro-Amoedo et al.<sup>10,11</sup> The average annual (2021) energy price was queried using the approach described in Santecchia et al.<sup>12</sup>

The GA-AL method is applied to the total cost minimization of a biogas (60% CO<sub>2</sub>, 40% CH<sub>4</sub>) chemical absorption separation (Figure 2). This chemical separation was modeled according to Sharma et al.<sup>13</sup>; the parameters used for surrogate building are described in Table 2, and the total heat duty is the surrogate output. The superstructure problem is formulated as a mixed-integer linear programming (MILP), following the approach described in Castro-Amoedo et al.<sup>14</sup>.

### 2.2.1. Performance evaluation

The evaluation of the GA-AL approach is based on **four** criteria. The first is the number of total (expensive) model evaluation per surrogate method ( $N_{mtd}^{eval}$ ); the second is the beta approximation ( $\beta_{rank}$ ) (eq. 4), which measures how close a given method (mth) is from the reference expected value ( $\beta_{ref}$ ), which was obtained using a Monte Carlo simulation with  $10^5$  points. The third (eq. 5) is a trade-off between the first and second criteria, with  $N_{med}$  being the median number of model evaluations. Lastly, MSE (eq. 6) represents the mean square error. To avoid the bias introduced by the set of initial points, twenty replications were performed for each surrogate method.

Table 2 – Parameters, range and output

Inputs	Range
Absorber temperature	40-60°C
Stripper temperature	90-120°C
Absorber diameter	15-20 m
Stripper diameter	5-12 m
Absorber tray space	0.6-1.2 m
Stripper tray space	0.6 –1.8m

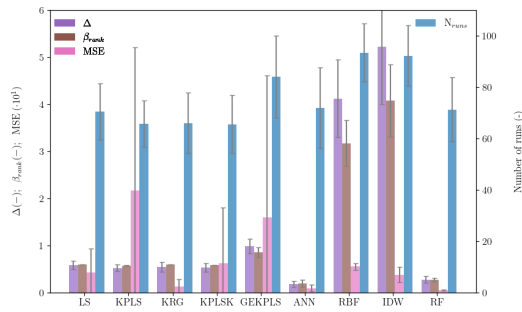
$$\beta_{rank} = \left| \frac{\hat{\beta}_{mtd} - \beta_{ref}}{\beta_{ref}} \right|, \text{ with } \hat{\beta} = \frac{\hat{\mu}_{mth}}{\hat{\sigma}_{mth}} \quad (4)$$

$$\Delta = \beta_{rank} \cdot \frac{N_{mtd}^{eval}}{N_{med}} \quad (5)$$

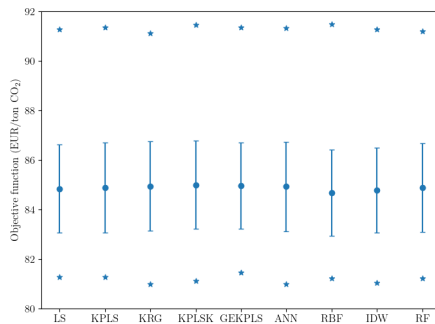
$$MSE = \sum_{i=1}^{N_m^{eval}} (y(x_{test,i}) - \hat{y}(x_{test,i}))^2 \quad (6)$$

### 3. Results and Discussion

The systematic evaluation of each approach is based on evaluating the performance criteria for the different strategies. Figure 3 (a) summarizes the results for the set of surrogate models tested. As suggested, random forest and artificial neural networks outperform the others, scoring the lowest MSE and being close to the reference expected value ( $\beta_{ref}$ ). This demonstrates the capacity of such algorithms to adapt to any kind of mathematical structure, even with limited amount of data.



(a)



(b)

Figure 3 – (a) surrogate models performance based on the four criteria performance criteria; (b) MILP objective function range

noting that additional features, such as enhanced-gradient (GEKPLS) or the use of partial least squares (KPLS) do not bring advantage. Indeed, for the same average number of runs, an increase (by one order of magnitude) in MSE is recorded.

The surrogate model is part of a superstructure that minimizes the total cost of separating CO<sub>2</sub> from a biogas mixture. At the level of the MILP, the quality of the surrogate influences the GA choice of points and consequently the range of costs. This is visible in Figure 3 (b). Regardless of the surrogate model, the fine tuning of the operating parameters bound the separation process between 81 and 92 EUR/ton of CO<sub>2</sub>.

### 4. Conclusions

In this work, a genetic algorithm is coupled with an active learning methodology to generate surrogate models. Such combined approach harvests the best of both methods:

As suggested, random forest and artificial neural networks outperform the others, scoring the lowest MSE and being close to the reference expected value ( $\beta_{ref}$ ). This demonstrates the capacity of such algorithms to adapt to any kind of mathematical structure, even with limited amount of data.

On the other hand, RBF and IDW underperform, due to the combination of a large number of runs required but also a poor approximation to the reference. Their MSE is one order of magnitude greater than the best performing methods, despite the larger number of samples. This might be due to the rather simplistic approach of both methods.

The standard Kriging (KRG) is, after ANN and RF, the best performing algorithm. It brings advantage compared to the least squares (i.e. linear regression) approach. However, it is worth

the genetic algorithm guides the active learning search to the points that are more relevant to the optimization problem at hand, whereas the active learning ensures finding the most interesting point to be simulated and then executed at the MILP optimization interface. Nine surrogate methods were tested in a biogas separation case study. Artificial neural networks and random forest are the best performing algorithms, followed by a standard Kriging formulation. On the other hand, the radial basis function and the inverse distance weighting algorithms have shown the worst performance. Moreover, the coupling of Kriging metamodel with additional features such as enhanced-gradient or partial least squares did not prove advantageous. The cost of separating CO<sub>2</sub> and CH<sub>4</sub> from a biogas mixture was simultaneously assessed, with costs ranging between 81 and 92 EUR/ton of CO<sub>2</sub>.

The approach here presented is modular and can accommodate any type of surrogate method, learning function or stopping criteria. Future work will address a more comprehensive set of stopping criteria and their effect on surrogate accuracy, as well as a systematic ranking system. In addition, high-dimensional problems (with more than twenty parameters) will be tested, show-casing the use of this framework for any type of PSE problem.

## References

1. Granacher, J., Nguyen, T.-V., Castro-Amoedo, R. & Maréchal, F. Overcoming decision paralysis—A digital twin for decision making in energy system design. *Applied Energy* **306**, 117954 (2022).
2. Cozad, A., Sahinidis, N. V. & Miller, D. C. Learning surrogate models for simulation-based optimization. *AIChE Journal* **60**, 2211–2227 (2014).
3. Granacher, J., Kantor, I. D. & Maréchal, F. Increasing Superstructure Optimization Capacity Through Self-Learning Surrogate Models. *Frontiers in Chemical Engineering* **3**, (2021).
4. Schweidtmann, A. M. & Mitsos, A. Deterministic Global Optimization with Artificial Neural Networks Embedded. *J Optim Theory Appl* **180**, 925–948 (2019).
5. Moustapha, M., Marelli, S. & Sudret, B. Active learning for structural reliability: Survey, general framework and benchmark. *Structural Safety* **96**, 102174 (2022).
6. Lee, S., Kim, J., Kang, H., Kang, D.-Y. & Park, J. Genetic Algorithm Based Deep Learning Neural Network Structure and Hyperparameter Optimization. *Applied Sciences* **11**, 744 (2021).
7. Hong, L., Li, H. & Peng, K. A combined radial basis function and adaptive sequential sampling method for structural reliability analysis. *Applied Mathematical Modelling* **90**, 375–393 (2021).
8. Bouhlel, M. A. et al. A Python surrogate modeling framework with derivatives. *Advances in Engineering Software* **135**, 102662 (2019).
9. Hong, L., Li, H. & Fu, J. A novel surrogate-model based active learning method for structural reliability analysis. *Computer Methods in Applied Mechanics and Engineering* **394**, 114835 (2022).
10. Castro-Amoedo, R., Damartzis, T., Granacher, J. & Marechal, F. M. A. System Design and Performance Evaluation of Wastewater Treatment Plants Coupled With Hydrothermal Liquefaction and Gasification. *Front. Energy Res.* **8**, (2020).
11. Castro-Amoedo, R., Morisod, N., Granacher, J. & Maréchal, F. The Role of Biowaste: A Multi-Objective Optimization Platform for Combined Heat, Power and Fuel. *Frontiers in Energy Research* **9**, 417 (2021).
12. Santicchia, A., Kantor, I., Castro-Amoedo, R. & Maréchal, F. Industrial Flexibility as Demand Side Response for Electrical Grid Stability. *Frontiers in Energy Research* **10**, (2022).
13. Sharma, S., Castro-Amoedo, R., Hemrle, J. & Maréchal, F. Assessment of Carbon Capture Technologies for Waste-to-Energy System. in *Computer Aided Chemical Engineering* (eds. Montastruc, L. & Negny, S.) vol. 51 871–876 (Elsevier, 2022).
14. Castro-Amoedo, R., Dahmen, A., Barbosa-Povoa, A. & Maréchal, F. Network design optimization of waste management systems: the case of plastics. in *Computer Aided Chemical Engineering* (eds. Türkay, M. & Gani, R.) vol. 50 185–190 (Elsevier, 2021).

# Data-driven Distributionally Robust Joint Chance-Constrained Optimization for Industrial Utility systems under Uncertainty

Hanxiu Li, Liang Zhao\*

*Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China*

## Abstract

The motivation to improve the energy efficiency of utility systems is resonating more intensely since the issue of carbon neutrality was put forward. However, the energy demand uncertainty brings challenges for utility systems modeling and optimization. This paper proposed a multi-objective optimization framework for minimizing operating cost and environmental impact simultaneously. The environmental impact is defined using the life cycle assessment method, and the operating cost consists of the resource consumption of the utility system. A data-driven distributionally robust joint chance-constrained (DRJCC) approach is proposed to deal with the uncertainty of steam demands. The data-driven ambiguity set is constructed based on Wasserstein distance, which contains the empirical distribution and is incorporated into the DRJCC framework. The dual representation for the worst-case probability transforms the problem into a deterministic non-convex reformulation. The proposed method is verified by a case study from an industrial ethylene plant. The results demonstrate that the total cost and environmental impact reduce as the confidence coefficient increases. In addition, the Pareto frontier enables us to select the optimal solutions flexibly, which may minimize the operational cost while the environmental standard is satisfied.

**Keywords:** industrial utility system, Wasserstein ambiguity set, distributionally robust joint chance-constrained program, life cycle assessment, multi-objective optimization.

## 1. Background

Emission reduction and energy saving have become the main concerns for the petrochemical industry in the carbon-neutral era. New technology and method applied to improve efficiency and decrease environmental burden is necessary for sustainable development in the next few decades. Process modeling, optimization, and environmental assessment have been focused on for this issue. An improved simulation and optimization method was presented for the sustainable development of utility systems (Varbanov et al., 2020). Ioannou et al. reviewed the role of process systems engineering tools in assessing and optimizing production patterns for alternative chemicals based on renewable resources (Ioannou et al., 2021). A consequential life cycle optimization framework was developed to determine the optimal chemical recycling technology (Zhao and You, 2021). Čuček et al. reviewed the methods of footprints which are defined as indicators and used to measure sustainability (Čuček et al., 2012). With continuous research on system optimization under uncertainty in recent years, stochastic and robust optimization methods were proposed to improve the system's robustness. As a combination of robust and stochastic programming, the distributionally robust chance-

constrained program provides a strategy that can adjust the relationship between system reliability and the optimal cost (Ding et al., 2022; Xie and Ahmed, 2016).

As one of the indispensable roles in the process industry, the utility system is generally responsible for providing power and heat for the industrial manufacturing process. Therefore, the deterministic multi-objective model based on life cycle assessment is constructed first. Then, the uncertain parameters of steam demand are collected from the historical data. An improved data-driven DRJCC framework with a Wasserstein ambiguity set is presented, and the deterministic model is reformulated in the dual form for the utility system. Finally, a case study of the utility system is presented to verify the effectiveness of the proposed method.

## 2. Deterministic Multi-objective Model Based on Life Cycle Assessment

A deterministic multi-objective framework for getting the optimal solutions with sustainability based on economic and environmental factor is proposed in this section.

### 2.1 Objective function

The operating cost and environmental indicator based on LCA method are defined as sustainability objectives. According to the list of the energy and material consumed in the system, we can calculate operating cost as:

$$\begin{aligned}
 COST = & \sum_{bo} \sum_i P_{i,bo}^{Fuel} \cdot M_{i,bo}^{Fuel} + p^{Ele} \cdot \left( \sum_{mo} M_{mo}^{Ele} + \sum_m M_m^{Ele} \right) + \\
 & p^{Water} \cdot \left( \sum_{lv} M_{lv}^{Water} + \sum_m M_m^{Water} \right) + \sum_m p_m^{De} \cdot M_m^{De} \quad (1) \\
 & i = \{Coal, NG\}, m = \{FGD, SCR\}
 \end{aligned}$$

The life cycle assessment method of CML 2001 was applied for evaluating the environmental load and helped to design the environment-friendly system (Guinée et al., 2001). The total emissions produced in chemical production are composed of direct emissions and indirect emissions. Direct emissions are mainly from the burning of fossil fuels. Equation (2) states the emission of CO<sub>2</sub> and equation (3) indicates the emissions of SO<sub>2</sub> and NO<sub>x</sub> under controlled conditions of the flue gas treatment process. The emissions generated by the exploitation, processing, transportation of upstream resources and power production are the indirect emission, which is presented in (4). By weighting the four impact factors, a single environmental assessment index is finally obtained in (5).

$$DE_{CO_2} = \sum_i F_{CO_2,i} \cdot M_i^{Fuel} + F_{CO_2,FGD} \cdot M_{FGD}^{De} \quad (2)$$

$$DE_k = Produce_k \cdot (1 - \theta_k), \quad \theta_k \in [0,1], k = 1, 2, \dots, 12 \quad (3)$$

$$IE_k = \sum_j LCI_{k,j} \cdot M_j, \quad j = \{Coal, NG, Ele, Water, Lim, Amm\} \quad (4)$$

$$IMPCT = \sum_n \frac{\sum_k LCI_{k,n} \cdot (IE_k + DE_k)}{F_n^{Normalization}}, \quad n = \{ACI, GWP, EUT, HTP\} \quad (5)$$

### 2.2 Constraints

The steam demand constraints, power demand constraints and variables range constraints of the proposed utility system are given in equation (6)-(8)(Zhao and You, 2019).

$$W_t^E + W_t^{ST} \geq W_t^{Demand}, \forall t \quad (6)$$

$$\sum_i f_{le,i}^{in} - \sum_j f_{le,j}^{out} \geq d_{le}, le \in \{HS, MS, LS\} \quad (7)$$

$$x^{\min} \leq x \leq x^{\max} \quad (8)$$

### 3. Data-Driven DRJCC Model for the Utility System

#### 3.1 Data-driven uncertainty sets for DRJCC

The chance-constrained approach can be applied to control the system reliability to a predefined level directly, and help to decide the optimal solutions. The distributionally robust chance-constrained approach over Wasserstein ambiguity set, defined in equations (9) and (10), is presented (Ho-Nguyen et al., 2021) to handle this problem. The ambiguity sets  $F_N(\theta)$  are sets of probability measures and defined as the  $\theta$ -radius Wasserstein ball of distributions around the empirical distribution  $\hat{\mathbb{Q}}_N$ . It ensures that all probability distributions of uncertain constraints satisfy with a given probability threshold within the selected Wasserstein distance from the empirical distribution.

$$d_W(\mathbb{Q}, \mathbb{Q}') := \inf_{\Pi} \left\{ \begin{array}{l} \mathbb{E}_{(\tilde{\mathbf{u}}, \tilde{\mathbf{u}}') \sim \Pi} [\|\tilde{\mathbf{u}} - \tilde{\mathbf{u}}'\|_*], \\ \Pi \text{ has marginal distributions } \mathbb{Q}, \mathbb{Q}' \end{array} \right\} \quad (9)$$

$$F_N(\theta) := \left\{ \mathbb{Q} : d_W(\hat{\mathbb{Q}}_N, \mathbb{Q}) \leq \theta \right\} \quad (10)$$

However, the uncertainty involves three levels of steam demand constraints, which can be considered as a joint chance constraint. The joint chance constraint with right hand side uncertainty is reformulated as an individual chance constraint by enforcing the worst-case CVaR approximation method (Chen et al., 2022). The set (11) with demand uncertainty is reformulated as (12).

$$S(\mathbf{x}) := \left\{ \mathbf{x} : \mathbf{b}_p^T \tilde{\mathbf{u}}_i + d_p - \mathbf{a}_p^T \mathbf{x} > 0, p \in [P] \right\} \quad (11)$$

$$\chi_{DR}(S) := \left\{ \mathbf{x} \in \mathcal{X} : \max \left\{ \begin{array}{l} \exists t \geq 0, \mathbf{r} \geq \mathbf{0}, \\ 0, \min_{p \in [P]} \frac{\mathbf{b}_p^T \tilde{\mathbf{u}}_i + d_p - \mathbf{a}_p^T \mathbf{x}}{\|\mathbf{b}_p\|_*} \end{array} \right\} \geq t - r_i, i \in [N], \right. \\ \left. \varepsilon t \geq \theta + \frac{1}{N} \sum_{i \in [N]} r_i \right\} \quad (12)$$

where  $\mathbf{a}_p \in \mathbb{R}^K$ ,  $\mathbf{b}_p \in \mathbb{R}^L$ ,  $d_p \in \mathbb{R}$  for all  $p \in [P]$ , and  $\|\cdot\|_*$  is the dual norm (Boyd and Vandenberghe, 2006).  $\theta$  is used to adjust the size of Wasserstein ball.  $\varepsilon$  is the desired safety factor between 0 and 1, which requires a group of  $m$  inequalities affected by uncertainty jointly satisfied the probability of at least  $1 - \varepsilon$ .

#### 3.2 Multi-objective optimization with DRJCC framework

We use the weighted-sum method to find the Pareto frontier (Deb and Deb, 2014), the operating cost and environment impact objectives are formulated as follows,

$$\min \lambda \text{COST} + (1-\lambda) \text{IMPACT} \tag{13}$$

where  $\lambda$  is the weighting coefficient and satisfying  $\lambda \in (0,1)$ . The solutions of the multi-objective problem will strongly depend on the changes of weighting coefficient.

We present the DRJCC formulation considering uncertain steam demands distributions (Chen et al., 2022; Xie, 2021). All constraints involving demands  $\tilde{d}$  are reformulated in the dual form of DRJCC and shown as follows.

$$\varepsilon t \geq \theta + \frac{1}{N} \sum_{i \in [N]} r_i, \tag{14}$$

$$M(1 - z_i) \geq t - r_i, i \in [N], \tag{15}$$

$$\sum_{i \in [N]} z_i \leq \lfloor \varepsilon N \rfloor, \tag{16}$$

$$\left( \sum_t a_{p,t}^{in} f_t^{in} + a_{p,t}^{ext} f_t^{ext} \right) + \sum_j a_{p,j}^M (1 - y_j) F_j^S + \sum_{lv} (a_{p,lv}^{in} f_{lv}^{in} + a_{p,lv}^{out} f_{lv}^{out}) - d_{p,i} \Big) / \|m_p\|_* \tag{17}$$

$$+ (d_{p,i} - q_p) / \|m_p\|_* \geq t - r_i, i \in [N]_p, p \in [P],$$

$$\left( \sum_t a_{p,t}^{in} f_t^{in} + a_{p,t}^{ext} f_t^{ext} \right) + \sum_j a_{p,j}^M (1 - y_j) F_j^S + \sum_{lv} (a_{p,lv}^{in} f_{lv}^{in} + a_{p,lv}^{out} f_{lv}^{out}) \Big) / \|m_p\|_* \geq t, \tag{18}$$

$$p \in [P],$$

$$f_t^{in, \min} \leq f_t^{in} \leq m_t^{in, \max}, \forall t \in T, \tag{19}$$

$$f_t^{exh, \min} \leq f_t^{in} - f_t^{ext} \leq f_t^{exh, \max}, \forall t \in T \tag{20}$$

$$\mathbf{z} \in \{0,1\}^N, t \geq 0, \mathbf{r} \geq 0. \tag{21}$$

Then, the multi-objective problem with joint demand uncertainties is converted into a deterministic MINLP problem, which can be solved by the existing solvers easily.

### 4. Case study

The case of the utility system with four steam grades (SS, HS, MS, LS) for an ethylene plant is presented (Zhao et al., 2014). There are three boilers, four extractions turbines, sixteen backpressure turbines with backup motors, two let-down valves in the given utility system. The demand parameters are collected from the process historical database. The data is not uniformly distributed over a specific range of operating conditions, which can be regarded as a multi-modal, interrelated and asymmetric uncertain parameter set. The superstructure of the utility system and uncertain parameters are given in figure 1.

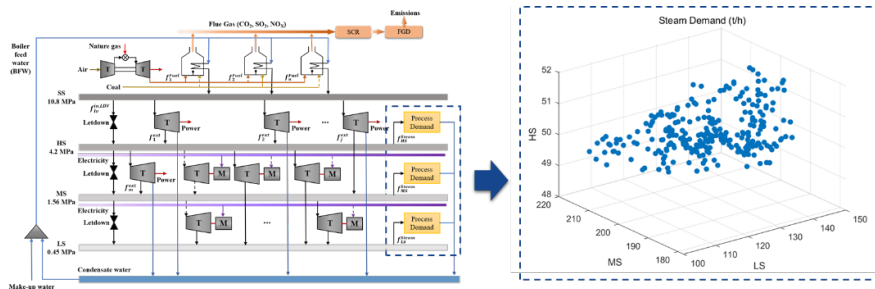


Figure 1. The uncertainty parameters from historical operation data.

Resources consumed by the system and their prices are shown in Table 1. The environmental impact categories and indicators are obtained from the CLCD and openLCA (Ciroth, 2007; Liu et al., 2010).

Table 1 The resources consumed in the system.

	Coal (ton)	Nature gas (ton)	Electricity (kWh)	Water (ton)	Limestone (ton)	Ammonia (ton)
Price (CNY)	800	2700	0.55	0.16	350	3500

We test four joint risk tolerance  $\varepsilon$  which are 0.05, 0.1, 0.15 and 0.2 and set the Wasserstein radius to  $\theta \in \{0.05, 0.1, 0.15, 0.2\}$ . All problems are solved with BARON solver. Figure 2 shows the trend of cost and environmental impact with changes of  $\varepsilon$  and  $\theta$ .

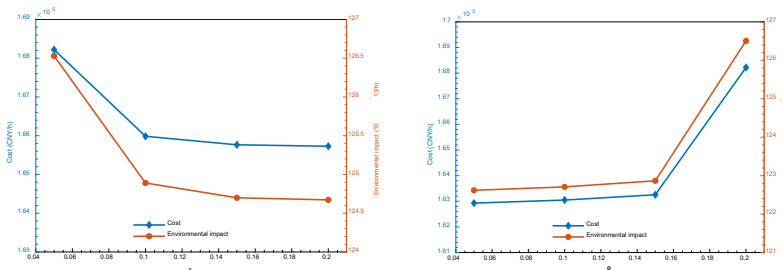


Figure 2. The trend of cost and environmental impact.

The operating cost and environmental impact increase along with the system reliability requirements when the value of risk tolerance decreased from 0.1 to 0.05. On the other hands, extending the Wasserstein radius also increases the economic and environmental burden, which means that the conservation of the solution increases.

Figure 3 shows the Pareto frontier of the multi-objective solutions when  $\varepsilon=0.05$  and  $\theta=0.05$ . Points A and C represent the minimal operating cost environmental impact solutions. Point B is the optimal solution which is supported by the TOPSIS method.

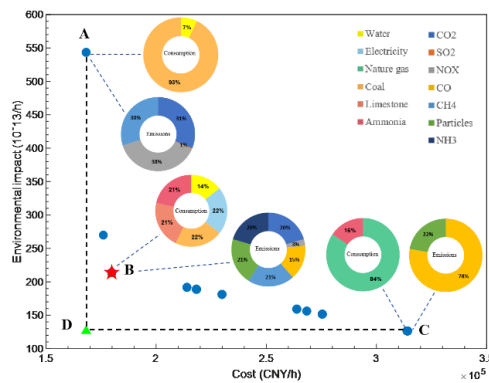


Figure 3. The Pareto frontier of operating cost and environmental impact.



Furthermore, the resources consumption and emissions of different solutions are shown in the pie charts. The total cost ranges from 168,220 to 314,176 CNY per hour, and environmental impact ranges from 126.53 to 543.42 units per hour.

## 5. Conclusion

An improved DRJCC method with Wasserstein ambiguity set framework was proposed for the utility system optimization based on LCA. Under steam demand uncertainties, it sets economic and environmental objectives and optimizes chance constraints to obtain a trade-off between operating cost, and environmental impact. A case study of utility system in ethylene plant is presented to verify the effectiveness of the proposed method.

## Acknowledgment

This work was supported by National Key Research & Development Program - Intergovernmental International Science and Technology Innovation Cooperation Project (2021YFE0112800), National Natural Science Foundation of China (Key Program: 62136003, 22178103, 62173144) and Shanghai AI Lab.

## References

- S.P. Boyd, L. Vandenberghe, 2006. Convex Optimization. *IEEE Transactions on Automatic Control* 51, 1859-1859.
- Z. Chen, D. Kuhn, W. Wiesemann, 2022. Data-Driven Chance Constrained Programs over Wasserstein Balls. *Operations Research*.
- A. Ciroth, 2007. ICT for environment in life cycle applications openLCA — A new open source software for life cycle assessment. *The International Journal of Life Cycle Assessment* 12(4), 209.
- L. Čuček, J.J. Klemeš, Z. Kravanja, 2012. A Review of Footprint analysis tools for monitoring impacts on sustainability. *Journal of Cleaner Production* 34, 9-20.
- K. Deb, K. Deb, 2014. Multi-objective Optimization, in: Burke, E.K., Kendall, G. (Eds.), *Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques*. Springer US, Boston, MA, pp. 403-449.
- Y. Ding, T. Morstyn, M.D. McCulloch, 2022. Distributionally Robust Joint Chance-Constrained Optimization for Networked Microgrids Considering Contingencies and Renewable Uncertainty. *IEEE Transactions on Smart Grid* 13(3), 2467-2478.
- J.B. Guinée, R. Heijungs, G. Huppes, A.d. Koning, L. Oers, A.W. Sleswijk, U.d. Haes, R.v. Duin, E. Lindeijer, 2001. *Life cycle assessment An operational guide to the ISO standards*.
- Ho-Nguyen, N., Kılınç-Karzan, F., Küçükyavuz, S., Lee, D., 2021. Distributionally robust chance-constrained programs with right-hand side uncertainty under Wasserstein ambiguity. *Mathematical Programming*.
- I. Ioannou, S.C. D'Angelo, Á. Galán-Martín, C. Pozo, J. Pérez-Ramírez, G. Guillén-Gosálbez, 2021. Process modelling and life cycle assessment coupled with experimental work to shape the future sustainable production of chemicals and fuels. *Reaction Chemistry & Engineering* 6(7), 1179-1194.
- X. Liu, H. Wang, J. Chen, Q. He, P. Hou, 2010. Method and basic model for development of Chinese reference life cycle database. *Acta Scientiae Circumstantiae* 30(10), 2136-2144.

P.S. Varbanov, J. Škorpík, J. Pospíšil, J.J. Klemeš, 2020. Sustainable Utility Systems: Modelling and Optimisation. De Gruyter.

W. Xie, 2021. On distributionally robust chance constrained programs with Wasserstein distance. *Mathematical Programming* 186(1), 115-155.

W. Xie, S. Ahmed, 2016. On Deterministic Reformulations of Distributionally Robust Joint Chance Constrained Optimization Problems. *SIAM Journal on Optimization* 28.

H. Zhao, Y. Feng, X. Dong, G. Rong, 2014. Integration Optimization of Production and Utility System for Refinery-wide Planning. *IFAC Proceedings Volumes* 47(3), 9599-9604.

L. Zhao, F. You, 2019. A data-driven approach for industrial utility systems optimization under uncertainty. *Energy* 182(SEP.1), 559-569.

X. Zhao, F. You, 2021. Consequential Life Cycle Assessment and Optimization of High-Density Polyethylene Plastic Waste Chemical Recycling. *ACS Sustainable Chemistry & Engineering* 9(36), 12167-12184.



## Online real-time multi-parameter optimization solution based on parallel EGO algorithm

Xuerong Gu<sup>a</sup>, Siyu Yang<sup>a, b\*</sup>

<sup>a</sup> School of Chemistry and Chemical Engineering, South China University of Technology, Guangzhou, 510640, PR China

<sup>b</sup> Guangdong Key Laboratory of Green Chemical Products Technology, South China University of Technology, Guangzhou, 510640, PR China

Email address: cesyyang@scut.edu.cn

### Abstract

Chemical process simulation and optimization have the characteristics of high-dimensional and non-linear, which make the simulation higher computational cost or even more difficult to converge. Using surrogate models to replace mechanistic models is an effective way to reduce computational complexity and ensure the accuracy of results. Though the Kriging surrogate model has a stronger nonlinear approximation, it is still difficult to deal with high-dimensional problems. Therefore, this paper investigates the parallel EGO algorithm integrated with the surrogate model, and applies the model to the typical chemical process. Taking eight multi-peak test functions as test cases, it is found that the accuracy and convergence speed of the parallel EGO algorithm is obviously improved. Finally, it is applied to the study case of the two-stage ammonia absorption refrigeration process. The results show that the simulation error is less than 0.01%, and the optimization time is reduced from 9846 s to 3705 s.

**Keywords:** surrogate model; parallel efficient global optimization algorithm; Multi-parameter optimization; process simulation.

### 1. Introduction

Process simulation and optimization are the basis for achieving energy saving and emission reduction in the chemical field. Due to the complex engineering design problems with high-dimensional and nonlinear characteristics, the optimization of operating parameters using just traditional process simulation software is often ineffective. To effectively solve this problem, some researchers [1-2] proposed building surrogate models relying on data-driven. The computationally fast surrogate model was used to replace the computationally time-consuming mechanism model, thus greatly reducing the number of computations of time-consuming objectives and constraints. The surrogate models [3-7] commonly used in engineering are polynomial response surface (PRS), radial basis function (RBF), artificial neural network (ANN), support vector regression (SVR), and Kriging surrogate model. Rosma et al [8-9] established PRS and RBF surrogate models for turbines and complex fitness functions and used particle swarm optimization algorithms to dynamically update the models. Yang et al [10] used the ANN surrogate model to optimize the process operation parameters of the crude oil distillation unit to maximize the oil production rate. Singh et al [11] used three surrogate models (SVR, RBF, Kriging) combined with an efficient multi-objective optimization algorithm to identify the Pareto frontier of cyclone separator design with the minimum number of simulations. Compared with the traditional optimization methods, the above-mentioned surrogate-based optimization algorithms can reduce the time-consuming simulation computation in the optimization process. However, this optimization

algorithm is too dependent on the sample sets and often requires a large number of samples to be fitted or trained to obtain more reliable results.

To solve the shortcomings of the above optimization algorithms, Jones et al [12] proposed the EGO (Efficient Global Optimization) algorithm. This approach uses fewer sample points to train the Kriging surrogate model and continuously selects the point with the largest value of the EI (Expected Improvement) function as the update point during each iteration. The Kriging surrogate model is continuously updated to find the global optimal solution. This approach can be used to solve the black-box problem but also to speed up the convergence and obtain higher accuracy and computing efficiency.

Based on the superiority of the EGO algorithm, the parallel EGO (Pseudo Efficient Global Optimization) algorithm is proposed based on it in this paper. The algorithm uses parallel computing techniques to select multiple update points for computation during each iteration, thus speeding up the optimization speed of the EGO algorithm. It was applied to 8 multi-peak test functions and the two-stage ammonia absorption refrigeration operation optimization case. The results show that the approach can optimize all variables online in real-time with high accuracy and efficiency, which has a greater value for engineering applications.

## 2. Parallel EGO algorithm

### 2.1 Kriging surrogate model

The Kriging model [13] consisting of  $m$  samples has a total of  $2m+2$  parameters, which can be obtained by the maximum likelihood estimation [14]. Finally, the predicted value and variance of any point  $x$  can be obtained:

$$\begin{cases} \hat{y} = \hat{\mu} + r^T R^{-1}(y - \hat{\mu}) \\ s^2(x) = \hat{\sigma}^2 \left[ 1 - r^T R^{-1} r + \frac{(1 - R^{-1} r)^2}{R^{-1}} \right] \end{cases} \quad (1)$$

### 2.2 EI criterion

For the unknown point, it can be regarded as a normally distributed random variable with mean  $\hat{y}(x)$  and standard deviation  $s(x)$ , it can be expressed as  $Y(x) \sim N(\hat{y}(x), s(x))$ . If the minimum objective function value in the current sample point is  $f_{\min}$ , then the improvement of the unknown point to the current optimal solution can also be regarded as a random variable:

$$I(x) = \max(f_{\min} - Y(x), 0) \quad (2)$$

Solving the mathematical expectation of  $I(x)$  can get the expression of the EI function as:

$$EI(x) = (f_{\min} - \hat{y}(x)) \Phi \left( \frac{f_{\min} - \hat{y}(x)}{s} \right) + s(x) \phi \left( \frac{f_{\min} - \hat{y}(x)}{s} \right) \quad (3)$$

where  $\phi$  and  $\Phi$  represent the probability density function and the cumulative distribution function of the standard normal distribution. It can be found from equation (3) that the EGO algorithm mainly searches for the minimum value  $\hat{y}(x)$  or maximum value  $s(x)$  as an update point by maximizing the EI function in each iteration, and finds the optimal solution after several iterations.

### 2.3 PEI criterion

The main role of the influence function is to approximate the changes brought to the EI function when iteratively updating the Kriging surrogate model, and its equation is as follows:

$$IF(x, x^{(h)}) = 1 - Corr[\varepsilon(x), \varepsilon(x^{(h)})] \quad (4)$$

where  $x$  is an arbitrary unknown point and  $x^{(u)}$  is an update point. The PEI (Pseudo Expected Improvement) function is constructed by multiplying the initial EI function by the impact function of the update point, and its equation is as follows:

$$\begin{aligned}
 PEI(x, q-1) &= EI(x) \cdot IF(x, x^{(m+1)}) \cdot IF(x, x^{(m+2)}) \cdots IF(x, x^{(m+q-1)}) \\
 &= EI(x) \cdot \prod_{i=1}^{q-1} [1 - \text{Corr}[\varepsilon(x), \varepsilon(x^{(m+i)})]] \\
 &= EI(x) \cdot \prod_{i=1}^{q-1} \left[ 1 - \exp\left(-\sum_{k=1}^d \theta_k |x_k - x_k^{(m+i)}|^{p_k}\right) \right]
 \end{aligned}
 \tag{5}$$

From the above derivation, it is clear that the parallel EGO algorithm can achieve parallel computation by maximizing the PEI function to get  $q$  update points in each iteration. This provides a new breakthrough for solving time-consuming chemical simulation optimization problems and improves optimization efficiency.

### 3. Numerical experiments

The proposed parallel EGO algorithm is compared with the EGO algorithm. The number of iterations required to find the global optimal solution for eight representative classical multi-peaked test functions is shown in Figure 1. Each data point in the figure was obtained by averaging each point by testing 10 times using 10 different initial designs.

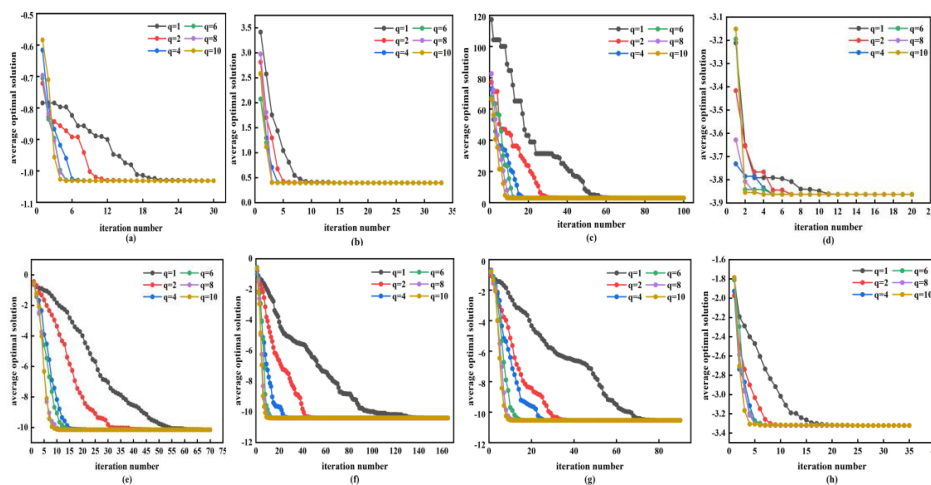


Fig.1 Iteration history curves of different optimization algorithms on test functions. (a) Six-hump. (b) Branin. (c) GoldPrice. (d) Hartman3. (e) Shekek15. (f) Shekek17. (g) Shekek10. (h) Hartman6.

From Figure 1, it can be found that the proposed parallel EGO algorithm performs significantly better than the EGO algorithm no matter how many update points are selected in each iteration. Meanwhile, the convergence speed is significantly faster with the number of update points increasing.

It can be found from the time variation curves in Figure 2 that with guaranteed accuracy, the proposed parallel EGO algorithm is higher efficiency and stronger adaptability when more points are added in each circulation.

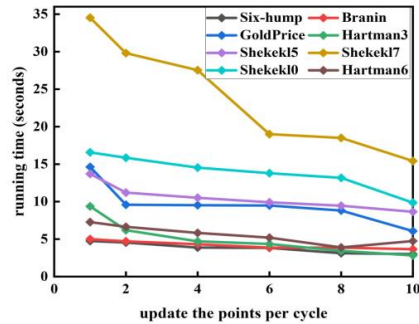


Fig.2 Running time of different multi-peak test functions

## 4. Multi-parameter optimization of two-stage ammonia absorption refrigeration

### 4.1 process description

The two-stage ammonia absorption refrigeration unit is an important part of the low-temperature methanol washing waste heat and cooling power co-generation system. Based on this, Aspen Plus is used to model and simulate it. The flowchart of its is shown in Figure 3.

Based on the analysis of key parameters by Liu et al [15], 15 decision variables to be optimized were determined. And the coefficient of performance COP (the ratio of evaporator load to the sum of flasher and distillation tower load) was chosen as the optimization objective for optimization.

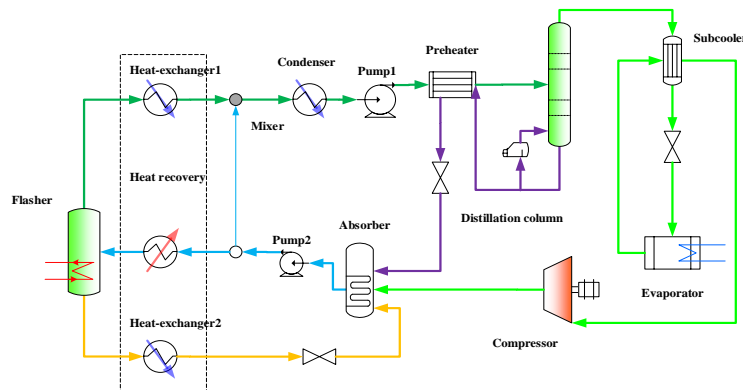


Fig.3 Schematic diagram of the two-stage ammonia absorption refrigeration process

### 4.2 Results and Discussions

The predicted values using the Kriging surrogate model are compared with the actual values as shown in Figure 4. The mean square error of the model is  $2.35 \times 10^{-3}$ , indicating that the constructed Kriging surrogate model predicts very accurately and can replace the original two-ammonia absorption refrigeration simulation model. Meanwhile, the time required for the numerical calculation of the original simulation model is 21013s, while Kriging only requires 0.383s, so the calculation time is significantly reduced.

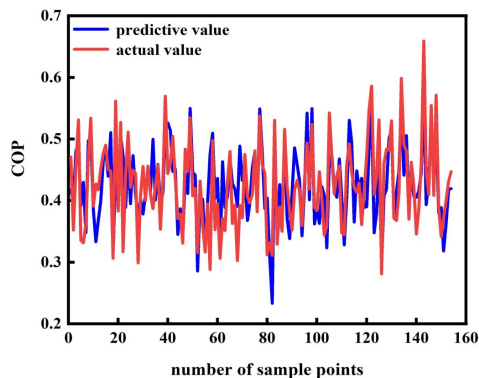


Fig.4 Comparison between predicted and actual values of Kriging model

The iteration curves of the standard EGO algorithm and the parallel EGO algorithm are shown in Figure 5. It can be found from the figure that the standard EGO algorithm requires 330 iterations to converge. In contrast, for the parallel EGO algorithm, the number of iterations gradually decreases to 29 as the number of points updated per circulation increases. The optimization convergence speed is improved by 91.2%. The results show that the parallel EGO algorithm converges faster than the standard EGO algorithm; the optimal solution obtained by the proposed parallel EGO algorithm is also better than the standard EGO algorithm when 600 iterations are ended.

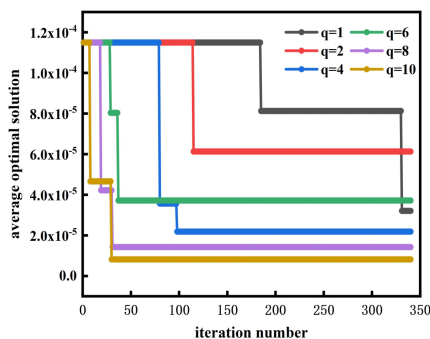


Fig.5 Optimization iteration curves

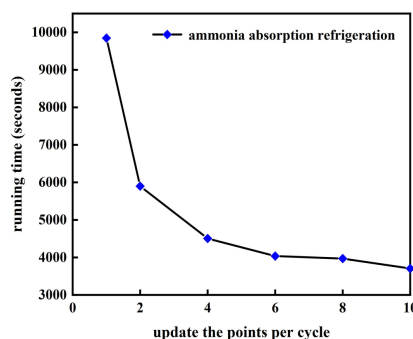


Fig.6 Running time

The relationship that the computation time required by each algorithm relative to the number of update points per iteration is shown in Figure 6. It can be found that under the condition that the maximum number of iterations is set to 600, the time required by the standard EGO algorithm is 9846s, and the time required by the parallel EGO algorithm is 5898-3705s. It is obvious that the parallel EGO algorithm requires less computation time than the standard EGO algorithm, and the operation efficiency is improved by 62.37%.

## 5. Conclusions

This paper proposed an online real-time parallel EGO optimization algorithm based on Kriging surrogate model for solving computationally intensive and time-consuming optimization problems of chemical process operations. The proposed parallel EGO algorithm uses the PEI criterion to select multiple update points in each iteration for parallel computation. The algorithm is tested on eight test



functions of dimensions 2 to 6. The experimental results show that the parallel EGO algorithm proposed in this paper can significantly improve the convergence speed of the standard EGO algorithm; it has higher efficiency and robustness than the standard EGO algorithm in all eight test problems. In addition, when the proposed parallel EGO algorithm is applied to the parameter optimization of two-stage ammonia absorption refrigeration, under the condition of ensuring higher accuracy, using relatively less training data to approach the multi-parameter optimal solution, the obtained solution is also significantly better than the standard EGO algorithm. The convergence speed is improved by 91.2% compared with the standard EGO algorithm.

## References

- [1] M. Tang, Y. Liu, L. J. Durlofsky, 2020, A deep-learning-based surrogate model for data assimilation in dynamic subsurface flow problems[J]. *Journal of Computational Physics*, 413: 109-456.
- [2] R. M. M. Slot, J. D. Sørensen, B. Sudret, et al, 2020, Surrogate model uncertainty in wind turbine reliability assessment[J]. *Renewable Energy*, 151: 1150-1162.
- [3] S. Wang, J. Xiao, J. Wang, et al, 2018, Application of response surface method and multi-objective genetic algorithm to configuration optimization of Shell-and-tube heat exchanger with fold helical baffles[J]. *Applied Thermal Engineering*, 129: 512-520.
- [4] Y. Deng, X. Zhou, J. Shen, et al, 2021, New methods based on back propagation (BP) and radial basis function (RBF) artificial neural networks (ANNs) for predicting the occurrence of haloketones in tap water[J]. *Science of The Total Environment*, 772: 145534.
- [5] J. Eason, S. Cremaschi, 2014, Adaptive sequential sampling for surrogate model generation with artificial neural networks[J]. *Computers & Chemical Engineering*, 68: 220-232.
- [6] M. W. Ahmad, M. Mourshed, Y. Rezgui, 2018, Tree-based ensemble methods for predicting PV power generation and their comparison with support vector regression[J]. *Energy*, 164: 465-474.
- [7] D. Zhan, J. Qian, Y. Cheng, 2017, Pseudo expected improvement criterion for parallel EGO algorithm[J]. *Journal of Global Optimization*, 68(3): 641-662.
- [8] G. E. González, R. Y. Qassim, P. C. C. Rosman, et al, 2016, Optimization of hydrokinetic turbine array layouts via surrogate modeling [J]. *Renewable Energy*, 93: 45-57.
- [9] H. Yu, Y. Tan, C. Sun, et al, 2016, An adaptive model selection strategy for surrogate-assisted particle swarm optimization algorithm[C]//2016 IEEE symposium series on computational intelligence (SSCI). Los Alamitos: IEEE Computer Society Press.
- [10] C. K. Liau, C. K. Yang, M. T. Tsai, 2004, Expert system of a crude oil distillation unit for process optimization using neural networks[J]. *Expert Systems with Applications*, 26(2): 247-255.
- [11] P. Singh, I. Couckuyt, K. Elsayed, et al, 2016, Shape optimization of a cyclone separator using multi-objective surrogate-based optimization[J]. *Applied Mathematical Modelling*, 40(5-6): 4248-4259.
- [12] D. R. Jones, M. Schonlau, W. J. Welch, 1998, Efficient global optimization of expensive black-box functions[J]. *Journal of Global Optimization*, 13(4): 455-492.
- [13] D. G. Krige, 1951, A statistical approach to some basic mine valuation problems on the Witwatersrand[J]. *Journal of the Southern African Institute of Mining and Metallurgy*, 52(6): 119-139.
- [14] S. N. Lophaven, H. B. Nielsen, J. Søndergaard, 2002, DACE: a Matlab kriging toolbox[M]. IMM, Informatics and Mathematical Modelling, The Technical University of Denmark.
- [15] S. Liu, Y. Qian, D. Li, et al, 2022, Multi-Scenario scheduling optimization for a novel Double-Stage ammonia absorption refrigeration system incorporating an organic Rankine cycle[J]. *Energy Conversion and Management*, 270: 116170.

# Optimal design of Water Treatment Networks: effluent and piping disaggregated modelling

Francisco J. G. Patrocínio,<sup>a,b</sup> Hugo M. D. Carabineiro,<sup>c</sup> Henrique A. Matos<sup>b</sup> and  
Nuno M. C. Oliveira,<sup>a\*</sup>

<sup>a</sup>*CIEPQPF, Dep. Chemical Engineering, University of Coimbra, 3030-790 Coimbra, Portugal*

<sup>b</sup>*CERENA, Dep. Chemical Engineering, IST, University of Lisbon, 1049-001 Lisbon, Portugal*

<sup>c</sup>*Petrogal, S.A., 7520-952 Sines, Portugal*

\**nuno@eq.uc.pt*

## Abstract

Water availability is of crucial importance to industry; therefore, optimal management of this resource has garnered increased importance. This work focuses on industrial water treatment networks (WTN) optimal design models, and can be divided in two sections. The first section addresses the modification of typical mathematical routines to implement efficient and rigorous effluent formulations, using an integer disaggregated formulation. The second section considers network piping expenditures, employing a two-step methodology to account for the discontinuous classes of commercialized piping.

**Keywords:** WTN optimization, MINLP, effluent modelling, piping, disaggregated model

## 1. Introduction

WTN models have been the subject of various publications in the past (e.g., Galan and Grossmann, 1998), although the effluent description has been somewhat simplified. Field observation of WTNs brings various additional features: a network can have multiple discharges; these can account for a significant fraction of the network OPEX; and they can be tariffed according to discontinuous concentration classes. Patrocínio *et al.* (2022) considers more detailed modelling of effluents employing indicator variables. This formulation has the downside of introducing new non-linearities in the mathematical model. In this work, a disaggregated formulation is suggested to rigorously model the effluent section of WTNs.

The second section of this work addresses the introduction of piping related expenses in the optimal grassroots design of WTNs, considering the nature of the diameter classes in which pipes are commercialized. Some authors (Alnouri *et al.*, 2014; Caballero and Ravagnani, 2019) already consider this characteristic. In this work, a two-step disaggregated methodology is proposed to model these equipment expenses, according to discontinuous pipe diameter classes.

## 2. WTN disaggregated model

The water treatment network superstructure (Figure 1) is composed by water flow producer nodes (set  $i$ ) and consumer nodes (set  $j$ ). Subsets of wastewater sources ( $ww \subset i$ ), treatment units ( $tu \subset i \wedge tu \subset j$ ) and multiple effluents ( $ef \subset j$ ) are defined. Superstructure connections between the same treatment unit are forbidden (with the set  $par_{i,j}$ ). Eq. (1) models that outlet flow from nodes  $i$  ( $F_i^{nd}$ ) is the sum of flows from  $i$  to  $j$

( $F_{i,j}$ ) and eq. (2) models the inlet flow in nodes  $j$ . Eq. (3) establishes that the flow from wastewater sources is defined by the parameter  $f_{ww_i}$ .

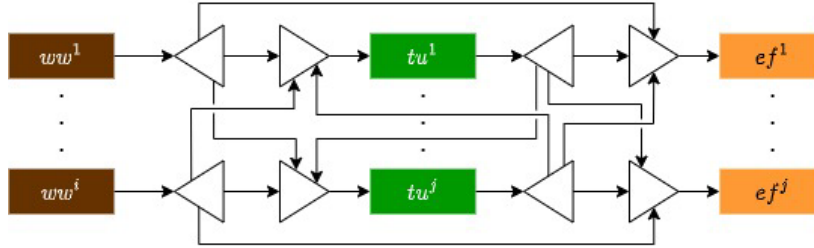


Figure 1 WTN superstructure.

The problem is further specified by a set of network contaminants (set  $co$ ), with eq. (4) as the mass balance of contaminants in the inlet of consumer nodes ( $C_{j,co}^{in}$ ,  $C_{i,co}^{out}$  are the inlet and outlet concentration of nodes  $j$  and  $i$ ). Eq. (4) contains the only non-linearity of the model, as two continuous variables are multiplied. From eq. (5), the outlet concentration of wastewater sources is fixed by the parameter  $c_{i,co}^{outWW}$  and eq. (6) defines the removal of contaminants in treatment units, through parameter  $rr_{j,co}$ . Eq. (7) establishes an upper limit for the inlet concentration of  $j$  ( $c_{j,co}^{lim}$ ). Note that each treatment and effluent node may contain inlet concentration limits. Effluents can be further classified into regular or special effluents ( $se \subseteq j$ ). The costs of the former are simply a tariff ( $cost_j^{RE}$ ) multiplied by the effluent flow, and special effluent costs are related to discontinuous contaminants concentration classes (set  $tcl$ ) of specific contaminants ( $tar \subseteq co$ ), with the highest activated class as the one tariffed. Here a disaggregated formulation for modelling the discharge of special effluents is used, avoiding the non-linearities resulting from the multiplication of integer and continuous variables. The inlet flow and concentration in the special effluent nodes are partitioned ( $F_{j,tcl}^{nd-tcl}$ ,  $C_{j,co,tcl}^{cl}$ ), according with the tariff classes (eqs. (8)-(9)) and only one partition is activated (eq. (10)), introducing the binary variable  $Y_{j,tcl}^{tcl}$ . Eqs. (11)-(12) guarantee that the concentration for the activated class is lower or equal to the upper bound concentration for this class ( $ub_{j,co,tcl}^{cl}$ ). Eqs. (13)-(14) assign the inlet flow of the node to the activated class, using an upper bound flow for node  $j$  ( $f_j^{nd-up}$ ). The lowest possible class is activated, as higher cost tariff classes are penalized by the objective function (eq. (15)). This function includes the operational costs of the treatment units ( $cost_j^{treat}$ ) multiplied by their inlet flow, and the costs of regular and special effluents.

$$F_i^{nd} = \sum_{j|par_{i,j}} F_{i,j}, \forall i \quad (1)$$

$$F_j^{nd} = \sum_{i|par_{i,j}} F_{i,j}, \forall j \quad (2)$$

$$F_i^{nd} = f_{ww_i}, i \in ww \quad (3)$$

$$\sum_{i|par_{i,j}} F_{i,j} C_{i,co}^{out} = F_j^{nd} C_{j,co}^{in}, \forall j, \forall co \quad (4)$$

$$C_{i,co}^{out} = c_{i,co}^{outWW}, i \in ww, \forall co \quad (5)$$

$$C_{j,co}^{in} rr_{j,co} = C_{j,co}^{out}, j \in tu, \forall co \quad (6)$$

$$C_{j,co}^{in} \leq c_{j,co}^{lim}, \forall j, \forall co \quad (7)$$

$$C_{j,co}^{in} = \sum_{tcl} C_{j,co,tcl}^{cl}, j \in se, \forall co \quad (8)$$

$$F_j^{nd} = \sum_{tcl} F_{j,tcl}^{nd-tcl}, j \in se \quad (9)$$

$$\sum_{tcl} Y_{j,tcl}^{cl} = 1, j \in se \quad (10)$$

$$C_{j,co,tcl}^{cl} \geq 0, j \in se, co \in tar, \forall tcl \quad (11)$$

$$C_{j,co,tcl}^{cl} \leq ub_{j,co,tcl}^{cl} Y_{j,tcl}^{tcl}, j \in se, co \in tar, \forall tcl \quad (12)$$

$$F_{j,tcl}^{nd-tcl} \geq 0, j \in se, \forall tcl \quad (13)$$

$$F_{j,tcl}^{nd-tcl} \leq f_j^{nd-up} Y_{j,tcl}^{cl}, j \in se, \forall tcl \quad (14)$$

$$Cost^{Network} = \sum_{j \in tu^j} cost_j^{treat} F_j^{nd} + \sum_{\substack{j \notin tu^j \\ j \in se^j}} cost_j^{RE} F_j^{nd} + \sum_{\substack{j \in se^j \\ tcl}} cost_{j,tcl}^{SE} F_{j,tcl}^{nd-tcl} \quad (15)$$

### 3. Piping investment and operational expenditures

The grassroots design of WTNs must account for piping expenses, with a cost that depends on the flows and distances between the network nodes. Furthermore, pipes are manufactured according to discontinuous diameter classes. This work proposes a two-stage procedure to incorporate these discontinuous classes in the WTN design model, performing a correspondence between the flows and the pipe classes costs and incorporating that correspondence in a mathematical optimization model, using a disaggregated strategy.

#### 3.1. Pre-processing stage

The pre-processing stage consists in an *a priori* correspondence between flows and the pipes cost. Firstly, the pipe optimal diameter ( $d_p^{opt}$  - Peters *et al.*, 2003) is computed (eq. (16)) using a set of  $p$  wastewater flows values ( $F_p$ ). Secondly, the optimal diameter is corresponded with the commercial pipe diameter, with a simple mixed integer linear model – eqs. (17)-(19), where  $d_p^{opt}$  is assigned to a pipe class (set  $pcl$ ), using a binary variable  $Z_{p,pcl}^{Pipe-Class}$ . Eq. (17) forces the assigned class diameter to be greater-or-equal than the optimal diameter, eq. (18) ensures that one pipe class is corresponded to each flow  $p$ , and the objective function (eq. (19)) imposes that the difference between the assigned class diameter and optimal diameter is as small as possible.

$$d_p^{opt} = \left( \frac{F_p}{3600} \right)^{0.45} 0.593, \forall p \quad (16)$$

$$Z_{p,pcl}^{Pipe-Class} (d_{pcl}^{class} - d_p^{opt}) \geq 0, \forall p, \forall pcl \quad (17)$$

$$\sum_{pcl} Z_{p,pcl}^{Pipe-Class} = 1, \forall p \quad (18)$$

$$\min \sum_{p,pcl} Z_{p,pcl}^{Pipe-Class} (d_{pcl}^{class} - d_p^{opt}) \quad (19)$$

The correspondence obtained will be similar to Figure 2, relating individual sch. 40 pipe costs, (per meter length) for several inner diameter classes.

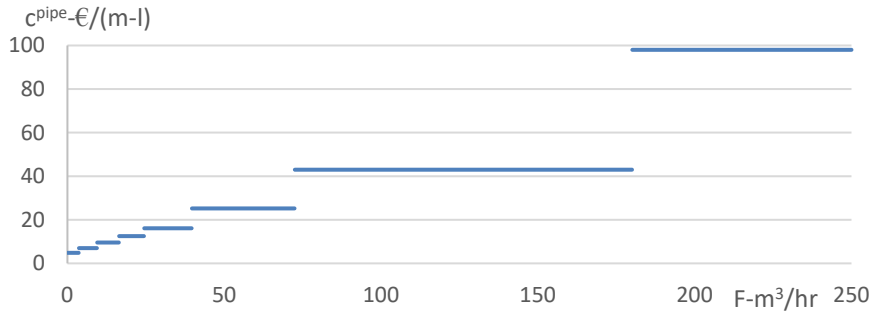


Figure 2 Individual pipe costs as a function of flow.

### 3.2. WTN disaggregated piping costs model

The proposed disaggregated technique is similar to the one presented for special effluents (section 2). The flow between nodes  $i,j$  is partitioned ( $F_{i,j,pcl}^{pcl}$ ) regarding the number of piping classes (eq. (20)). A new binary variable ( $Y_{i,j,pcl}^{pipe}$ ) is introduced to assign  $F_{i,j,pcl}^{pcl}$  to a class  $pcl$  (eqs. (21)-(22)), using the lower and upper bounds for each class ( $f_{pcl}^{lb}$ ,  $f_{pcl}^{ub}$ ). Eq. (23) assures that, at most, one class is activated. The inequality operator is important as it allows that a null flow does not activate a piping class. Yearly costs of installing a pipe in connection  $i,j$  ( $Cost_{i,j}^{Pipe-install}$ ) are modelled by eq. (24), considering an annualization factor ( $af$ ), labor/material cost factors ( $bm^{Pipe}$ ), distance between nodes ( $d_{i,j}^{nd}$ ) and cost of activated pipe class ( $cost_{pcl}$ ). Eq. (25) expresses the operational costs of pipe between nodes  $i,j$  ( $Cost_{i,j}^{Pipe-oper}$ ), where a maintenance factor ( $of^{Pipe}$ ) is multiplied by  $Cost_{i,j}^{Pipe-install}$ .

$$F_{i,j} = \sum_{pcl} F_{i,j,pcl}^{pcl}, \forall i, \forall j \quad (20)$$

$$F_{i,j,pcl}^{pcl} \geq f_{pcl}^{lb} Y_{i,j,pcl}^{pipe}, \forall i, \forall j, \forall pcl \quad (21)$$

$$F_{i,j,pcl}^{pcl} \leq f_{pcl}^{ub} Y_{i,j,pcl}^{pipe}, \forall i, \forall j, \forall pcl \quad (22)$$

$$\sum_{pcl} Y_{i,j,pcl}^{pipe} \leq 1, \forall i, \forall j \quad (23)$$

$$Cost_{i,j}^{Pipe-install} = af \cdot bm^{Pipe} \cdot d_{i,j}^{nd} \sum_{pcl} Y_{i,j,pcl}^{pipe} \cdot cost_{pcl}, \forall i, \forall j \quad (24)$$

$$Cost_{i,j}^{Pipe-oper} = of^{Pipe} Cost_{i,j}^{Pipe-install}, \forall i, \forall j \quad (25)$$

$$TCost^{Yearly} = H \cdot Cost^{Network} + \sum_{i,j} (Cost_{i,j}^{Pipe-install} + Cost_{i,j}^{Pipe-oper}) \quad (26)$$

The original objective function is replaced by eq. (26), and total network cost is computed ( $TCost^{Yearly}$ ) considering the yearly operation of the wastewater network (multiplying the original WTN model objective function by the number of operational hours, in a year –  $H$ ), and adding the piping yearly installation and operational expenditures.

### 4. Application Examples

The results presented in this section were obtained using an AMD Ryzen 7 4800H computer with 16 GB of RAM, and GAMS 40.3.0.

#### 4.1. Large industrial example of Galan and Grossmann (1998)

The last example of Galan and Grossmann (1998) was used to create the first application example of these techniques. The original problem specification was extended to account for different treatment unit operational expenditures (Table 1) and 2 effluent nodes with different limiting concentrations (Table 2). Effluent 1 is a special effluent with different concentration tariff classes (Table 3). Effluent 2 has a discharge cost of 0.1 €/t.

Table 1 Treatment units operating cost.

Unit	X1	X2	X3	X4	X5
Cost (€/t)	0.1	0.18	0.25	0.25	0.8

Using the WTN disaggregated model, the optimal network has an hourly operation expenditure of 106.97 €/hr, and the special effluent EF1 represents 53.8% of the total network cost, with 128.1 t/hr and concentrations within Class 1. Effluent EF2 has a flow of 12.9 t/hr, and corresponds to 1.2% of the total network cost. Here, the *ANTIGONE* solver was used, with a time limit of 1 hour, resulting in an optimality gap of 20.8%.

To consider the pipe investment together with the operational expenditures, using the procedure of section 3, distances between nodes were randomly generated within [10;300] meters, for correspondences  $(ww^{nd};tu^{nd})$  and  $(ww^{nd};ef^{nd})$ , and [5;50] meters for correspondences  $(tu^{nd};tu^{nd'})$  and  $(tu^{nd};ef^{nd'})$ . Piping classes were constructed based on Figure 2, but aggregating several classes and using median cost values as represented in Table 4. *af* was assumed as 0.207, *bm<sup>Pipe</sup>* as 17 (per recommendation of Woods, 2007), *of<sup>Pipe</sup>* as 0.2 (Peters *et al.*, 2003) and *H* as 8000 operating hours.

Table 2 Effluents concentration limits.

	C1 (ppm)	C2 (ppm)	C3 (ppm)	C4 (ppm)	C5 (ppm)	C6 (ppm)
ef1	100	100	100	100	100	100
ef2	0	0	200	500	200	0

Table 3 ef1 tariff classes.

	Cost (€/t)	C2 (ppm)	C3 (ppm)	C4 (ppm)
c11	0.44	[0;50[	[0;30[	[0;30[
c12	0.54	[50;85[	[30;55[	[30;55[
c13	0.72	[85;100]	[55;100]	[55;100]

When computed *a posteriori* for the solution of the WTN disaggregated model, piping investment and operational expenditures correspond to an annual network cost of 975.1 k€. Utilising the WTN disaggregated model with piping costs included, the optimal solution found improves 2.75%, with an annual network cost of 948.3 k€. The local solver *SBB* was used in this case, with an imposed optimality gap of 0.1%, requiring 11.15 hours of wall clock time for this solution.

Table 4 Pipe classes costs and flow intervals.

	pcl1	pcl2	pcl3	pcl4
Cost (€/m-l)	6.98	12.51	16.13	43.02
F (t/hr)	[0;9.4]	]9.4;24.4]	]24.4;39.5]	]39.5;141]

$d_{i,j}^{nd}$  is a crucial parameter regarding the savings reported by the WTN disaggregated piping costs model (in comparison to the computation of those expenditures *a posteriori* with the WTN disaggregated model), and the computational solving time. When the originally generated distance is decreased by 20% ( $0.8 d_{i,j}^{nd}$ ), the savings and wall clock computation time of the WTN disaggregated piping costs model drop to 1.96% and 3.03 hours. In a 30% decrease ( $0.7 d_{i,j}^{nd}$ ), the drop is to 1.5% and 1.43 hours.

#### 4.2. Sines Refinery wastewater treatment network

The proposed methodologies were applied to an industrial scenario, using data from the Sines Refinery wastewater network, in Portugal, composed by 11 wastewater sources, 5 treatment units, two effluent nodes and 6 contaminants. One of the effluent lines is a special effluent node, with 5 tariff classes, regarding 5 of the contaminants, and the costs associated with that node account for 42% of the WTN expenditures. Applying the optimization model of the first section achieves a 33.6% reduction in the total wastewater network costs. The solver utilized was *ANTIGONE*, with a wall clock time limit of 1 hour, resulting in an optimality gap of 14.2%.

To account for piping expenditures,  $d_{i,j}^{nd}$  was obtained with the refinery plot plan and, using a similar procedure to the previous example, 5 pipe classes were considered, as well as the parameters regarding  $af$ ,  $bm^{Pipe}$ ,  $of^{Pipe}$  and  $H$ . When these expenditures are considered, posteriorly and independently of the optimization routine, WTN savings drop to 7.9%. Optimizing the network using the WTN disaggregated piping costs model allows a further increase of WTN savings to 10.4% (with local solver SBB, and an imposed optimality gap of 0.1%, expending 3.04 hours of wall clock time).

## Conclusion

This work demonstrates that WTN effluents tariffs cannot be disregarded, in WTN optimization models, focusing in the relevance of discontinuous effluent discharges, as demonstrated by the example of Sines Refinery. Furthermore, piping expenditures are modelled in a more realistic manner, avoiding introducing further non-linearities in the routines. The proposed methodology has the downside of demanding a moderate degree of computational effort, which should be addressed in future work.

**Acknowledgments:** The authors gratefully acknowledge the financial support provided by Galp Energia and Fundação para a Ciência e Tecnologia, through grant COVID/BDE/152674/2022, and data from INTERFLUIDOS, Lda.

## References

- Alnouri S. *et al.* 2014. “Water Integration in Industrial Zones: A Spatial Representation with Direct Recycle Applications.” *Clean Technologies and Environmental Policy* 2014 16:8 16 (8): 1637–59.
- Caballero J.A. and Ravagnani M.A.S.S. 2019. “Water Distribution Networks Optimization Considering Unknown Flow Directions and Pipe Diameters.” *Computers & Chemical Engineering* 127 (August): 41–48.
- Galan B. and Grossmann I.E. 1998. “Optimal Design of Distributed Wastewater Treatment Networks.” *Industrial and Engineering Chemistry Research* 37 (10): 4036–48.
- Patrocínio F. *et al.* 2022. “Water Network Optimisation in Chemical Complexes: A Refinery Case Study” *Computer Aided Chemical Engineering* 51 (January): 817–22.
- Peters M. *et al.* 2003. “9. Optimum Design and Design Strategy.” In *Plant Design and Economics for Chemical Engineers*, 3rd ed., 401–6. McGraw-Hill Education.
- Woods, D.R. 2007. “Appendix D: Capital Cost Guidelines.” In *Rules of Thumb in Engineering Practice*, 376–436. Weinheim: Wiley-VCH.

# Evaluation of the potential of a deep eutectic solvent for liquid-liquid extraction of furfural using optimization-based process design

Kai Fabian Kruber<sup>a,b</sup>, Mariann Kroll<sup>b</sup>, Christoph Held<sup>b</sup>, Mirko Skiborowski<sup>a,b</sup>

<sup>a</sup>*Hamburg University of Technology, Institute of Process Systems Engineering, Am Schwarzenberg-Campus 4, 21073 Hamburg, Germany*

<sup>b</sup>*TU Dortmund University, Laboratory of Fluid Separations, Emil-Figge-Straße 70, 44227 Dortmund, Germany*  
*mirko.skiborowski@tuhh.de*

## Abstract

Reducing our environmental impact is the greatest challenge of our time and affects all areas of the chemical, biochemical, and pharmaceutical industries. The use of sustainable and biodegradable solvents in downstream processes, such as deep eutectic solvents (DESs), is one essential step in this direction. While many studies have focused on the investigation of- the thermo-physical properties of DESs, relatively little attention has been paid to solvent regeneration in the context of a closed-loop process, assuming a simple recovery due to the negligible vapor pressure of the DES. In order to evaluate the suitability of this assumption, the current work performs an optimization-based process design for the recovery of furfural from a diluted fermentation broth by liquid-liquid extraction, comparing a previously screened DES with toluene as an established organic solvent. The comparison reveals important limitations of the DES due to considerable water content in the raffinate and the high boiling point resulting in a more complex and energy-demanding process.

**Keywords:** process optimization, liquid-liquid extraction, distillation, solvent selection, green solvent.

## 1. Introduction

While deep eutectic solvents (DESs) were first proposed as an alternative to ionic liquids less than 20 years ago (Abbott et al., 2003), with 200 publications during the first ten years (Smith et al., 2014), publications have rapidly increased to more than 7000 in the last ten years, primarily addressing synthesis and solvent properties. DESs can be considered as molten salts with a eutectic point near or below room temperature due to hydrogen bonding and have been promoted with excellent application potential in the chemical, biochemical, and pharmaceutical industries (Hansen et al., 2021), showing low vapor pressure, production cost, and toxicity, as well as good biodegradability. These characteristics deem DESs environmentally friendly and cost-effective competitors to established organic solvents and ionic liquids. Most studies on fluid separation process applications focus on the investigation of thermodynamic properties, such as vapor pressures, densities, boiling and decomposition temperatures, or partition coefficients and selectivities (Salleh et al., 2017). On the other side, studies on the evaluation of closed-loop process performance, considering solvent recovery, of DES-based processes are still scarce. Larriba et al. (2018) studied the dearomatization of gasoline by liquid-liquid



extraction with six different DESs based on choline chloride. Based on experimental investigations paired with simulation studies, they selected a promising sustainable alternative to commonly applied organic solvents. However, the study does not consider the investigation of a complete process flowsheet with solvent recovery and a closed recycle. Han et al. (2022) examined such a closed-loop flowsheet in Aspen Plus for the dehydration of acetonitrile with glycolic acid:choline chloride(3:1) by extractive distillation. They found a substantial cost reduction compared to the benchmark process considering ethylene glycol as solvent. Shang et al. (2019) studied the dehydration of ethanol by extractive distillation using choline chloride:urea(1:2) as mass separating agent. The results based on a closed-loop process indicate a potential reduction in the annualized costs as well as the energy consumption of the process.

While the low vapor pressure implies a simple thermal regeneration of the DES, the high boiling temperature and limitations regarding the degradation temperature are potentially detrimental factors that need to be considered for the closed-loop process, including solvent regeneration. To evaluate the tradeoff of these factors, a rigorous optimization of the closed-loop flowsheet, including solvent recovery, is required. In the current study, hybrid extraction-distillation processes for the purification of furfural from an aqueous fermentation broth using decanoic acid:thymol(1:1) as DES and toluene as established organic solvent are evaluated as optimal design problems. Furfural is widely applied as a herbicide, fungicide, and insecticide in agricultural production and a flavoring agent in the food industry (Yong et al., 2020). Additionally, Furfural is one of the top bio-based chemicals due to its numerous reaction products, like fuel additives and value-added chemicals (Yan et al., 2014).

## 2. Methodology

As a first step for the targeted evaluation, a suitable DES for the considered separation has to be identified. Thermodynamic properties need to be modeled before a comparison with the established organic solvent can be made based on rigorous model-based process optimization.

### 2.1. Selection of a Deep Eutectic Solvent

For liquid-liquid extraction, the applicability of a solvent is restricted by the acceptable loss in the raffinate stream, while the economic performance is strongly linked to the amount of solvent required for a feasible separation. Therefore, the distribution coefficient is commonly considered a key performance indicator for extraction performance in an early evaluation step. Thermal recovery of the solvent by means of distillation has to adhere to the boiling and degradation temperature of the DES, which may limit the operating range of a distillation column. Consequently, to prevent degradation and the resulting product pollution and loss of solvent, the DES needs a comparably high degradation temperature to enable an efficient solvent recovery.

For the current example of furfural purification, Dietz et al. (2019c) provide the necessary data for selecting a suitable DES. In an experimental investigation of ten different DESs, decanoic acid:thymol(1:1) (Deca-Thy) was identified as the solvent with the highest distribution coefficient for liquid-liquid extraction of furfural (Dietz, 2019a). Based on investigations of the thermal decomposition of comparable DESs (Dietz, 2019a), a degradation temperature of 190 °C was assumed for Deca-Thy, which is only slightly lower than the highest decomposition temperature reported by Dietz et al. (2019a), for dodecanoic acid:atropine(2:1). However, Deca-Thy provides a distribution coefficient that is about 12 times higher compared to dodecanoic acid:atropine(2:1). Thus, it is

selected as most interesting DES for the current application and compared with toluene, as widely applied benchmark organic solvent (Dietz, 2019c).

### *2.2. Thermodynamic Properties*

While the previously described data allows for an initial selection of promising solvent candidates, a variety of thermodynamic properties have to be modeled reliably for a model-based assessment of the respective processes. Besides the pure component ideal gas heat capacity, vapor pressure, and enthalpy of evaporation, which are needed to evaluate the energy balance and thermal regeneration in the hybrid extraction-distillation process, especially the binary interactions for the liquid-liquid and vapor-liquid equilibrium (LLE, VLE) have to be known.

Parameters for pure component properties and activity coefficient models for water, furfural, and toluene are available and exported from Aspen Plus databases DB-PURE32 and APV88 LLE-ASPEN/APV88 VLE-IG, respectively. For modeling the DES and its interactions, similar to other studies (Han and Chen, 2018; Dietz et al., 2019b; Dai et al., 2022), it is assumed that the DES can be modeled as a single pseudo component, for which the ideal gas heat capacity is estimated by the group contribution method by Joback and Reid (1987) for decanoic acid and thymol individually and averaged for the DES. The DIPPR equation for the ideal gas heat capacity is fitted to the obtained data for the process optimization. The vapor pressure is predicted using the perturbed-chain statistical associating fluid theory (PC-SAFT) (Gross and Sadowski, 2001) with the pure component parameters published by Dietz et al. (2019b) and a ratio-specific individual constituent approach. The extended Antoine equation is fitted to the predicted data and used in the optimization-based design. Based on the vapor pressure, the enthalpy of evaporation is assumed to be constant and calculated using the Clausius-Clapeyron equation.

In order to avoid the application of the comparably complex PC-SAFT model in the process optimization, the non-random two-liquid (NRTL) model is considered as a surrogate model to describe the LLE and VLE within the process optimization. Therefore, the NRTL parameters are fitted to predicted equilibrium data determined by the PC-SAFT model. The binary interactions of furfural and Deca-Thy are determined by regressing VLE data at pressures of 0.05, 0.1, and 0.2 bar to ensure a correct phase behavior at vacuum conditions around 0.1 bar. In contrast, the binary interactions of water and Deca-Thy are obtained by regressing LLE data at a pressure of 1 bar. The  $c_{ij}$  parameter is fixed at 0.3 and 0.2 for VLE and LLE regression, respectively.

### *2.3. Rigorous Process Optimization*

In order to perform a techno-economical evaluation of the processes with the considered solvents, toluene, and Deca-Thy, a rigorous superstructure optimization based on equilibrium-stage models is performed. The general process configurations do, however, differ due to the different solvent properties. Since toluene is light-boiling compared to furfural, it is obtained as the top product in the solvent recovery column (cf. Figure 1). In contrast, Deca-Thy is heavy boiling and, as such, recovered as the bottoms product in the solvent recovery column (cf. Figure 2). More importantly, furfural is the intermediate boiling component in this mixture. On the basis of the individual process structures, a superstructure model is built, allowing for individual sizing of the extraction column and the rectifying and stripping sections of the distillation column. The respective mathematical models are extended with constraints on product purities as well as sizing and costing models (Biegler et al., 1997), resulting in mixed-integer nonlinear programming (MINLP) problems for the techno-economical designs, which are solved by a series of successively relaxed nonlinear programming (NLP) problems with the respective equilibrium calculations performed via a set of external functions

(Skiborowski et al., 2015). The models and the polyolithic solution strategy are implemented in GAMS 34.3.0, and SNOPT is used to solve for local optimality.

### 3. Process Optimization

For both processes, a feed stream of 5 mol% furfural and 95 mol% water with a flow rate of 100 mol/s (about 7900 kg/h) is assumed. For the wastewater stream comprising the raffinate and the water-rich phase of the decanter, a furfural content  $x_{FF} < 0.01$  mol% is defined, and the product stream is constrained to a purity of  $x_{FF} > 99.9$  mol%. A depreciation time of ten years with an interest rate of 6% and 8000 working hours per year are set for the calculation of the total annualized costs (TAC). For both solvents, a price of 1000 \$/t (846.5 €/t) is assumed to account for solvent make-up streams. As utilities, cooling water at 25 °C (0.05 €/t), low-pressure steam at 10 bar (14 €/t), and high-pressure steam at 25 bar (16 €/t) are considered. The superstructure consists of a liquid-liquid extraction column with a maximum number of 15 (initial) stages and a distillation column with a maximum number of 80 (initial) stages.

Figure 1 shows the optimization results for the benchmark solvent toluene. The TAC of the process is about 194.4 k€/y and is largely determined by the operating costs, which represent about 69%. The furfural purity in the optimized process even exceeds the desired specification, avoiding unnecessary solvent makeup. Apparently, there is no reduction in the number of stages for the extraction column due to the comparably low investment costs, considering only a column shell with simple sieve tray inlets. In contrast, the cost of potential solvent makeup and increased flow rates in the distillation column would increase the TAC significantly compared to the reduced cost of the extraction column.

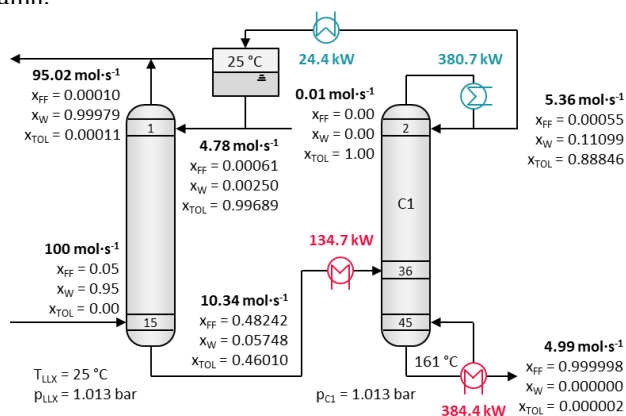


Figure 1: Optimized process flowsheet for the separation of furfural (FF) and water (W) using toluene (TOL) as the solvent in the liquid-liquid extraction (LLX) column

The results for the Deca-Thy-based process are presented in Figure 2. The high boiling temperature of Deca-Thy requires an alternative process design resulting in a three-column setup, including a liquid-liquid extraction, a distillation column for solvent recovery, and a heteroazeotropic distillation column for further purification of furfural. The solvent recovery is carried out in a vacuum column with a pressure of 0.1 bar to avoid the degradation temperature of Deca-Thy. The heteroazeotropic distillation is optimized separately based on the model and solution strategy presented by Kruber and Skiborowski (2022). Due to the furfural losses in the heteroazeotropic distillation, the overall furfural recovery drops slightly from 99.8% to 99.3%, whereas the furfural losses would also add

up to the wastewater stream. The TAC of the process is about 246.4 k€/y with almost the same capital expenditures as for the toluene process of about 60.5 k€/y. The increase in the TAC originates exclusively from an increase of 39% in operational expenditures, which is caused by the increased energy demand as well as the use of high-pressure steam (25 bar) in the vacuum column. Notably, the solvent-to-feed (S/F) ratio is significantly lower for the DES, which aligns with the increased distribution coefficient reported by Dietz et al. (2019c). Yet, the thermodynamically derived low S/F ratio should be considered with care, as phase ratios between 1:10 and 10:1 are usually considered for practical application (Goedecke, 2006).

As for toluene, the extraction column is utilized at its maximum number of stages due to the low investment costs. However, the distillation column uses less than half the stages compared to toluene, which is expected based on the extremely low vapor pressure of the DES. The main limitation of the DES is caused by the high cross-solubility of water in the solvent phase, due to which the additional purification in the heteroazeotropic distillation column is required, which counteracts the savings due to the smaller vacuum column. This not only causes a more complex and energy-demanding process but would also complicate the operation of the process.

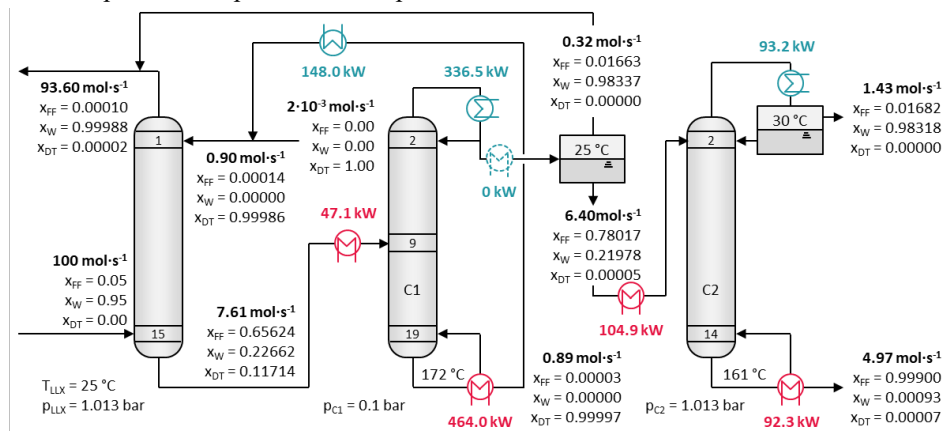


Figure 2: Optimized process flowsheet for the separation of furfural (FF) and water (W) using decanoic acid:thymol (DT) as the solvent in the liquid-liquid extraction (LLX) column

#### 4. Conclusion & Outlook

In order to evaluate the potential benefits of a DES, this work considers the purification of a diluted furfural stream by hybrid extraction-distillation processes by means of a rigorous optimization-based design. DESs are often examined based on simple thermodynamic properties and, therefore, recommended based on their low toxicity, flexibility, and particularly low vapor pressure, which implies an easy thermal regeneration. While all these properties are favorable for solvents, the high boiling temperature can result in a higher energy demand and the necessity for more costly heating utilities. The results illustrate the importance of a holistic consideration of solvent and process design to evaluate if the desired thermodynamic properties of the DES really translate into economic benefits for the overall process. It was shown that thermodynamic properties alone could not directly indicate economic savings. Although the recovery of Deca-Thy from the extract phase is simple, the high cross solubility of water requires further product purification and results in higher TAC of 246.4 k€/y compared to 194.4 k€/y for the toluene-based process. Especially for the purification of high boiling

organic compounds in diluted aqueous streams, a volatile solvent may be a favorable choice since the value product will not be an intermediate boiling component that can be separated without additional unit operations. In addition, the assumed degradation temperature for Deca-Thy is rather optimistic, and experimental studies indicate a much lower thermal stability resulting in even higher process cost.

In conclusion, DESs provide many favorable properties which make them interesting solvent candidates, but their suitability should be evaluated based on closed-loop process flowsheets that consider the effort for solvent recovery for each solvent candidate to gain more reliable information with respect to the process performance. In the future, the consideration of rigorous process models during the design and selection of DESs for separation processes should be further investigated, extending the scope to applications in extractive distillation and heteroazeotropic distillation. With the knowledge gained for the furfural purification, alternative DESs can be considered to improve the process performance and find a green alternative to the presented benchmark process and other interesting solvent candidates, such as butyl chloride (Nhien et al., 2017).

## References

- A.P. Abbott, G. Capper, D.L. Davies, R.K. Rasheed, V. Tambyrajah, 2003, *Chem. Comm.*, 1, 70–71
- L.T. Biegler, I.E. Grossmann, A.W. Westerberg, 1997, Prentice-Hall, Upper Saddle River, NJ
- Y. Dai, X. Chu, Y. Jiao, Y. Li, F. Shan, S. Zhao, G. Li, Z. Lei, P. Cui, Z. Zhu, Y. Wang, 2022, *Chem. Eng. Sci.*, 264, 118179
- C.H.J.T. Dietz, 2019a, Ph.D. Thesis, Chemical Engineering and Chemistry, Eindhoven University of Technology
- C.H.J.T. Dietz, J.T. Creemers, M.A. Meuleman, C. Held, G. Sadowski, M. van Sint Annaland, F. Gallucci, M.C. Kroon, 2019b, *ACS Sustainable Chem. Eng.*, 7, 4047–4057
- C.H.J.T. Dietz, F. Gallucci, M. van Sint Annaland, C. Held, M.C. Kroon, 2019c, *Ind. Eng. Chem. Res.*, 58 (10), 4240–4247
- Goedecke, R., 2006, *Fluidverfahrenstechnik*, Wiley-VCH, Weinheim
- J. Gross, G. Sadowski, 2001, *Ind. Eng. Chem. Res.*, 40 (4), 1244–1260
- D. Han, Y. Chen, 2018, *Chem. Eng. Process.*, 131, 203–214
- D. Han, Y. Chen, D. Mengru, 2022, *Comput. Chem. Eng.*, 164, 107865
- B.B. Hansen, S. Spittle, B. Chen, D. Poe, Y. Zhang, J.M. Klein, A. Horton, L. Adhikari, T. Zelovich, B.W. Doherty, B. Gurkan, E.J. Maginn, A. Ragauskas, M. Dadmun, T.A. Zawodzinski, G.A. Baker, M.E. Tuckerman, R.F. Savinell, J.R. Sangoro, 2021, *Chem. Rev.*, 121, 3, 1232–1285
- K.G. Joback, R.C. Reid, 1987, *Chem. Eng. Commun.*, 57 (1-6), 233–243
- K.F. Kruber, M. Skiborowski, 2022, 10 (8), 1482
- M. Larriba, M. Ayuso, P. Navarro, N. Delgado-Mellado, M. Gonzalez-Miquel, J. García, F. Rodríguez, 2018, *ACS Sustainable Chem. Eng.*, 6 (1), 1039–1047
- L.C. Nhien, N. Van Duc Long, S. Kim, M. Lee, 2017, *Biotechnol. Biofuels*, 10 (81)
- Z. Salleh, I. Wazeer, S. Mulyono, L. El-blidi, M.A. Hashim, M.K. Hadj-Kali, 2017, *J. Chem. Thermodyn.*, 104, 33–44
- X. Shang, S. Ma, Q. Pan, J. Li, Y. Sun, K. Ji, L. Sun, 2019, *Chem. Eng. Res. Des.*, 148, 298–311
- M. Skiborowski, A. Harwardt, W. Marquardt, 2015, *Comput. Chem. Eng.*, 72, 34–51
- E.L. Smith, A.P. Abbott, K.S. Ryder, 2014, *Chem. Rev.*, 114, 21, 11060–11082
- K. Yan, G. Wu, T. Lafleur, C. Jarvis, 2014, *Renewable Sustainable Energy Rev.*, 38, 663–676
- T.L.-K. Yong, K.F. Pa'ee, N. Abd-Talib, N. Mohamad, 2020, in: Inamuddin, A.M., Asiri (Eds.), *Advanced Nanotechnology and Application of Supercritical Fluids*, Springer International Publishing, Switzerland, 53–73.

# Profit Allocation in Industrial Symbiosis Networks: Utility Exchanges

Fabian Lechtenberg<sup>a</sup>, Antonio Espuña<sup>a</sup>, Moisès Graells<sup>a\*</sup>

<sup>a</sup> *Department of Chemical Engineering, Universitat Politècnica de Catalunya, Barcelona 08019, Spain*

\*[moises.graells@upc.edu](mailto:moises.graells@upc.edu)

## Abstract

Industrial symbiosis networks enable efficient resource sharing between companies and, thereby, generate otherwise unachievable profits. This work studies the multi-actor process integration problem that arises when addressing utility exchange involving different companies. The aim is to find stable profit allocations by combining process integration principles with game theory methods. The profit table provides the necessary information to solve the cooperative game that takes place in the industrial park. For the case study, a simplified utility exchange model is developed and the proposed decision-making framework is applied to it, in order to highlight the importance of knowledge about the potential pitfalls of the employed profit allocation methods.

**Keywords:** cooperative game theory, process integration, decision-making, steam

## Introduction

Collaboration in industrial symbiosis networks potentially leads to profits over standalone operation. These profits may be of any kind (economic, environmental, social ...). While non-cooperative interaction between the actors may be analyzed to find optimal decision-making strategies (Hjaila et al., 2017), it is not ensured that the combined profit of all actors is efficient, i.e. as high as if the actors were selflessly cooperating. Especially, in industrial parks it is reasonable to assume that companies may bargain in a cooperative manner: they can form binding agreements to fairly share the generated profits, regardless of whether an individual member would suffer losses if no allocation occurs. Fair cost and profit allocation was studied before for various case studies (Chin et al., 2021; Hiete et al., 2012) but no study considered utility exchanges via steam transfer so far. To that end, a utility exchange model and the corresponding case study are contributed. Moreover, this work makes a significant step towards a generalized framework for decision-making in general multi-actor process integration problems.

## Problem Statement

For the given multi-actor process integration problem (*utility exchanges*) in an industrial park find the highest possible total economic profit and apply allocation methods to fairly distribute the profit among the participants. Check if the allocations are stable, i.e. that an individually rational actor would be likely to accept the proposal.

## Methodology

The allocation of profits requires the prior knowledge of the potential profits that can be obtained in the grand coalition and all sub coalitions. To that end, a simplified utility network model is proposed based on works by Galvan-Cara et al. (2022) and Kim et al. (2010). The original model complexity is reduced by omitting the multi-period problem, reducing the number of steam levels, etc. Material, energy and economic balances are maintained. The resulting set of equations is summarized in (1)-(13) and it contains no binary decisions. Thus, the minimization problem is a linear program (LP).

- High-pressure steam (HS) and low-pressure steam (LS) headers:

$$\dot{m}_{HS,i}^{demand} + \dot{m}_{HS,i}^{vent} + \sum_j \dot{m}_{HS,i,j}^{exc} + \dot{m}_{HS,i}^{turbine} + \dot{m}_{HS,i}^{letdown} = \sum_{b_{HS}} \dot{m}_{i,b_{HS}}^{boiler} \quad \forall i \quad (1)$$

$$\dot{m}_{LS,i}^{demand} + \dot{m}_{LS,i}^{vent} + \sum_j \dot{m}_{LS,i,j}^{exc} = \dot{m}_{LS,i}^{turbine} + \dot{m}_{LS,i}^{letdown} + \sum_{b_{LS}} \dot{m}_{i,b_{LS}}^{boiler} \quad \forall i \quad (2)$$

- Boiler mass balance:

$$\dot{m}_{i,b}^{boiler} = Cal_b \cdot \dot{m}_{fuel,i,b}^{purchase} \cdot x_{i,b} \quad \forall i, b \in b_{HS} \cup b_{LS} \quad (3)$$

- Turbine and letdown units:

$$\dot{m}_{HS,i}^{turbine} = \dot{m}_{LS,i}^{turbine} + \dot{m}_i^{cond} \quad \forall i \quad (4) \quad \dot{m}_{HS,i}^{letdown} = \dot{m}_{LS,i}^{letdown} + \dot{m}_i^{letdown} \quad \forall i \quad (7)$$

$$\dot{m}_{LS,i}^{turbine} = Split \cdot \dot{m}_{HS,i}^{turbine} \quad \forall i \quad (5) \quad \dot{m}_i^{letdown} = Fill \cdot \dot{m}_{HS,i}^{letdown} \quad \forall i \quad (8)$$

$$P_i^{turbine} = Yield \cdot \dot{m}_{HS,i}^{turbine} \quad \forall i \quad (6)$$

- Water units:

$$\dot{m}_{water,i}^{demand} + \sum_b \dot{m}_{i,b}^{boiler} + \dot{m}_i^{letdown} = \dot{m}_{water,i}^{purchase} + \dot{m}_i^{cond} \quad \forall i \quad (9)$$

- Electricity units:

$$P_i^{demand} + \sum_b P_{i,b}^{boiler} = P_i^{turbine} + P_i^{purchase} \quad \forall i \quad (10)$$

- Economic balances:

$$C_i^{raw} = C^{elec} \cdot P_i^{purchase} + C^{water} \cdot \dot{m}_{water,i}^{purchase} + \sum_b C_b^{fuel} \cdot \dot{m}_{fuel,i,b}^{purchase} \quad \forall i \quad (11)$$

$$C_i^{transport} = \frac{1}{2} \cdot \sum_{s \in \{HS, LS\}} \sum_j \beta_s \cdot (\dot{m}_{s,i,j}^{exc} + \dot{m}_{s,j,i}^{exc}) \cdot L_{i,j} \quad \forall i \quad (12)$$

$$C = \sum_i (C_i^{raw} + C_i^{transport}) \quad (\text{subject to minimization}) \quad (13)$$

Where  $i, j$  is the set of plants and  $b$  is the set of boilers. Figure 1 shows the superstructure of the utility system in a plant and visualizes the balances from the model. Utility costs for fuel, water and electricity are considered. Steam exchanges can lead to savings in all utility purchases at the cost of pipeline construction and operation.

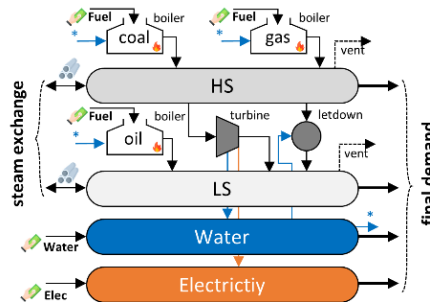


Figure 1. Simplified utility system superstructure for each plant.

### Cooperative Game Theory and Profit Allocation:

Multi-actor process integration problems such as the utility-exchange problem presented herein can be adjusted for every existing sub coalition by introducing binary parameters that restrict flows to only allow exchanges between the sub coalition members. The resulting profits are always positive with respect to the standalone case because any participant may decide to not collaborate if no individual profit is in sight. Figure 2 shows the steps followed in this work which is part of a general framework that is to be presented in future work. The profit table in the Figure is an arbitrary example with three players.

The profit table results from solving the process integration problem for the  $2^N$  coalitions, where  $N$  is the number of players. Using the information about the worth of each coalition, various solution concepts can be applied to fairly distribute the profits to the members of the coalition. We developed the Python package “*pyCoopGame*” which implements various well-known single-point solution concepts and stability criterion checks for any given profit table. In this work we employ four solution approaches of which the following descriptions briefly outline the underlying intuition.

The core concept (Gillies, 1959) is not a solution approach but a stability criterion that checks if a given profit allocation is efficient and acceptable for individually rational players. That means, an allocation is stable and lies in the core if the total profit equals the profit of the grand coalition and no sub coalition can ensure higher payoff for all members. The Shapley value (Shapley, 1988) makes use of the profit table to determine the marginal contribution of every member to every coalition and calculates a weighted average of the marginal contributions. The Nucleolus (Puerto & Perea, 2013) is the allocation that minimizes the largest dissatisfaction of all coalitions. Dissatisfaction is the difference between the worth of the coalition and the sum of the profits allocated to the coalition members. The  $\tau$ -value (Tijs & Driessen, 1986) is the efficient allocation that lies between the utopian allocation where each player receives their marginal contribution to the grand coalition and the minimum rights allocation. The least core solution (Maschler et al., 1979) is the profit allocation where the relaxed  $\varepsilon$ -core set is just composed of a single point.

It should be noted that these four approaches are selected based on their popularity but others may be chosen depending on the use case. Luo et al. (2022) give a good overview of recent developments and applications, as well as current challenges, of cost and profit allocation solution concepts applied to operations management.

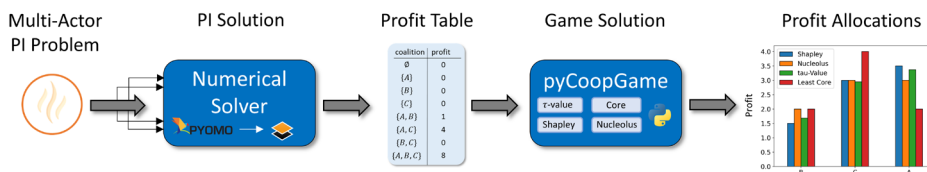


Figure 2. Solving a profit allocation problem in a multi-actor process integration case.

### Case Study

The multi-actor process integration problem studied in this work is a utility exchange problem adapted from the Yeosu industrial complex as shown in Figure 3. The game is redefined to include three players (10, 12 and 15) and the utility system is adapted to the previously presented simplified model.



The resulting parameters are summarized in Tables 1, 2 and 3. Demand values for VHS and HS as well as MS and LS are aggregated over the four time periods.

Table 1. Final annual demand values

	HS [kton]	LS [kton]	Water [kton]	Electricity [MWh]
10	5707	0	1908	952
12	0	111	476	396
15	3878	1042	1588	1272

The economic parameters shown in Table 2 are updated to match current market conditions. The fuels are coal (1), methane (2) and fuel oil (3).

Table 2. Economic parameters and calorific content of fuel

Parameter	Value	Unit	Parameter	Price [\$/ton]	$Cal_b$ $\left[ \frac{kg\ steam}{kg\ fuel} \right]$
Water Price	0.500	\$/ton	Fuel 1	205	8.9
Electricity Price	150.0	\$/MWh	Fuel 2	290	18.5
Pipe Operation ( $\beta$ )	0.436	\$/km/t	Fuel 3	1174	15.9

The distance between plants is kept unchanged.

Table 3. Distance between companies

Plants	Distance	Plants	Distance	Plants	Distance
10 ↔ 12	2.319 km	10 ↔ 15	2.193 km	12 ↔ 15	1.237 km

The utility system of each plant is changed to correspond to the simplified utility model. As can be seen from Figure 3 (right), in scenario 1 plant 10 has a HS coal boiler and a turbine, plant 12 has a LS fuel oil boiler and plant 15 has a HS methane boiler, a turbine and a LS fuel oil boiler. In scenario 2 plant 10 has a methane boiler but no turbine. Results for the whole complex using the original parameters from Galvan-Cara et al. (2022) are also provided.

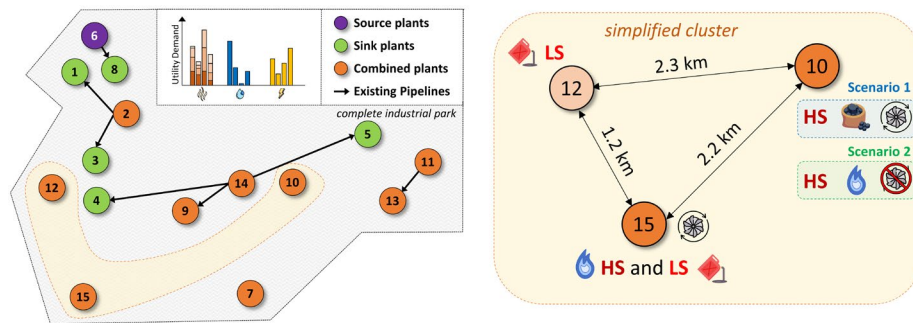


Figure 3. Yeosu industrial complex (Kim et al., 2010) and scenarios.

## Results and Discussion

Solving the process integration problem for all coalitions in scenario 1 and 2 leads to the profit tables as shown in Table 4 and 5. Most of the profit in scenario 1 is due to the fuel that plant 10 saves by not using its coal boiler and importing instead HS from plant 12, which has a more cost-efficient methane boiler. Further savings are generated by LS exchange of plant 12 and 15 with 10. In scenario 2 the big savings from plant 10 are not feasible because plant 15 already has a methane boiler. Thus, the only savings come from LS exchange with plant 12, which has a limited demand for this utility.

Applying the single-point solution concepts leads to the profit allocations in Figure 4 and 5. In scenario 1, Shapley, Nucleolus and  $\tau$ -value yield similar results that follow a simple intuition: plant 10 and 15 both need to be present to generate the large benefits, so both should receive similar payoffs. Plant 12 also has a contribution but its impact is way smaller so it can't claim as much payoff. It is not surprising that these three payoff vectors are indeed stable. However, it is non-intuitive that plant 15 and 10 should receive such asymmetric payoffs according to the least core solution. This outcome follows from the property of the least core method which always yields a solution on the edge of the feasible core allocation set, granted it is non-empty.

In scenario 2 the profit is largely determined by plant 12 accepting LS from either 10 or 15 until its demand is satisfied. Plant 10 is in the most powerful position and can claim the largest payoff. The results of the Nucleolus,  $\tau$ -value and least core methods are stable and they allocate the payoffs to plants 12 and 15. The latter is preferred over plant 10 due to the slightly closer geographical proximity. Plant 10 is not engaged in the final utility exchange so not assigning payoff to this player is intuitive. However, the Shapley value assigns a significant share of the payoffs to both players 10 and 15 due to their marginal contribution to the sub coalitions. This result is not stable and reveals one of the main weakness of the Shapley value, the most popular approach used in literature. These two scenarios demonstrate that allocation methods should be selected with care, being aware of their potential pitfalls.

Figure 6 shows the results for the whole industrial park. The Nucleolus is not included here as its calculation for games with many players becomes computationally unaffordable. The remaining profit allocations are very different and this is due to the fact that the core for this 15-player game is empty. In this kind of situations, the rules of the game may be adjusted by means of taxation and incentives (Chin et al., 2021).

Table 4. Profit Table (scenario 1)

Coalition	Profit
{10,12}	5,290,864 \$/year
{10,15}	36,673,525 \$/year
{12,15}	6,127,690 \$/year
{10,12,15}	42,804,345 \$/year

Table 5. Profit Table (scenario 2)

Coalition	Profit
{10,12}	6,074,326 \$/year
{10,15}	0 \$/year
{12,15}	6,127,690 \$/year
{10,12,15}	6,127,690 \$/year

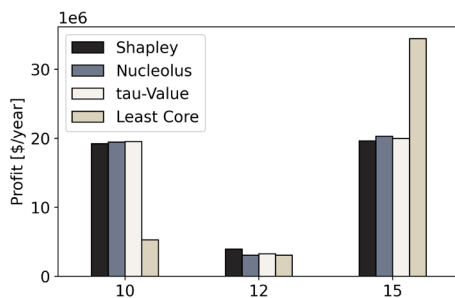


Figure 4. Allocations (scenario 1)

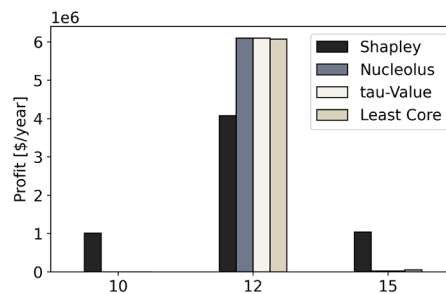


Figure 5. Allocations (scenario 2)

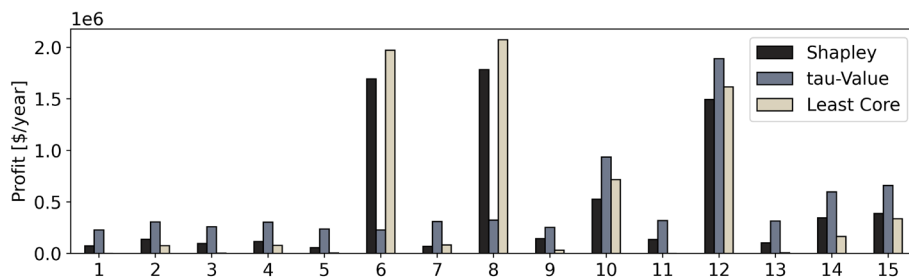


Figure 6. Allocations (complete industrial park). Total profit: 7,157,500 \$/year

## Conclusions

The application of cooperative game theory methods (and the associated profit allocation) to industrial symbiosis networks facilitates the decision-making and may create trust between the actors by proposing payoff vectors that are intuitive and stable. The presented study demonstrates this on a simplified utility exchange case study. Moreover, a significant step towards the generalization of a profit allocation framework for multi-actor process integration problems is done. Such a framework requires knowledge about strength and weaknesses of the available methods. Future work will further formalize this framework, characterize the allocation methods and demonstrate on a wide variety of process integration problems how decision-making can be systematically improved using cooperative game theory.

## Acknowledgements

Financial support received from the Spanish “Ministerio de Ciencia e Innovación” and the ERDF (project CEPI-PID2020-116051RB-I00) is fully acknowledged. Fabian Lechtenberg gratefully acknowledges the “Departament de Recerca i Universitats de la Generalitat de Catalunya” for the financial support of his predoctoral grant FI-2022.

## References

- A.L. Galvan-Cara, M. Graells, A. Espuña, 2022, Application of Industrial Symbiosis principles to the management of utility networks. *Applied Energy*, 305
- C. Luo, X. Zhou, B. Lev, 2022, Core, shapley value, nucleolus and nash bargaining solution: A Survey of recent developments and applications in operations management. *Omega*, 110
- D.B. Gilies, 1959, Solutions to general non-zero-sum games. In *Contributions to the Theory of Games IV*, 40
- H.H. Chin, P.S. Varbanov, J.J. Klemeš, S. Bandyopadhyay, 2021, Subsidised water symbiosis of eco-industrial parks: A multi-stage game theory approach, *Computers & Chemical Engineering*, 155
- J. Pueto, F. Perea, 2013, Finding the nucleolus of any n-person cooperative game by a single linear program. *Computers & Operations Research*, 40
- K. Hjäila, L. Puigjaner, J.M. Laíney, A. Espuña, 2017, Integrated game-theory modelling for multi enterprise-wide coordination and collaboration under uncertain competitive environment. *Computers & Chemical Engineering*, 98
- L.S. Shapley, 1988, A value for n-person games, *The Shapley Value: Essays in Honor of Lloyd S. Shapley*
- M. Hiete, J. Ludwig, F. Schultmann, 2012, Intercompany Energy Integration: Adaptation of Thermal Pinch Analysis and Allocation of Savings. *Journal of Industrial Ecology*, 16
- S.H. Kim, S.G. Yoon, S.H. Chae, S.Park, 2010, Economic and environmental optimization of a multi-site utility network for an industrial complex. *Journal of Environmental Management*, 91
- S.H. Tijs, T.S.H. Driessen, 1986, Game Theory and Cost Allocation Problems. *Management Science*, 32

# Operational optimisation of steam power system under uncertainty using time series prediction optimization model with **Markov** chain

Kenian Shi, Jingyuan Zheng, Siyu Yang\*

*School of Chemical Engineering, South China University of Technology, Guangzhou, 510640, China*

## Abstract

The stochastic programming and robust optimisation methods have been proposed to solve the operational optimisation of industrial energy system under uncertainty. However, these methods can not comprehensive the stability and economy at the same time. This paper proposes a two-stage time series prediction optimization model with **Markov** chain to solve this problem. The steam demand value is predicted by the method of scenario generation and reduction. Taking the steam power system of a coal-to-gas enterprise as an example, the corresponding optimization model is established, and the predicted steam value is put into the optimization model to solve. The results show that the optimization method in this paper combines the advantages of high economy of stochastic programming and high stability of robust optimization, which provides a new way to solve the uncertain optimization problem of steam power system.

Keywords: Steam demand uncertainty, time series prediction optimization, Markov chain, Scenario generation and reduction.

## Introduction

In the operation of steam power system, there are various uncertain factors (uncertain changes of operating temperature, pressure and flow rate), which will eventually lead to uncertain changes of steam demand <sup>[1]</sup>. The uncertainty of steam demand can lead to the blind production of steam power system, producing more or less steam than the demand. The former will lead to the venting of steam and waste resources, while the latter will affect the safety and stability of the production. Therefore, it is of great theoretical and practical significance to study the operational optimisation of steam power system under uncertainty.

At present, the methods to solve the uncertain optimization problem include stochastic programming <sup>[2]</sup> and robust optimization <sup>[3]</sup>. The main difference between these uncertain optimization methods lies in the description of uncertain parameters and solving algorithm <sup>[4]</sup>. Stochastic programming is only applicable to the uncertain optimization problems whose uncertain parameters can be expressed as probability distribution, but not to the uncertain optimization problems with non-probability distribution <sup>[5]</sup>, and it cannot avoid high risks in the process of system planning <sup>[6]</sup>. The robust optimization expresses the uncertain parameter as the uncertain parameter set and considers the optimization scheme in the worst case <sup>[7]</sup>.

Each method has its own advantages and disadvantages. Robust optimization has high stability but is too conservative. Stochastic programming has high economy but poor stability. At present, few optimization methods can comprehensive the stability and economy at the same time. This paper

proposes a two-stage time series prediction optimization model with Markov chain to solve the coexistence problem of economy and stability.

### 1. Methodology

The time series prediction of steam demand can be divided into two stages. In the first stage, multiple working conditions are divided based on the distance of three levels of steam. In the second stage, the steam demand is predicted based on the probability of state transition under each working condition, which mainly includes scenario generation and scenario reduction.

#### 1.1. Multiple working condition division

Due to the high volatility of the three levels of steam, it is necessary to classify them to improve the overall efficiency and economy of the steam power system. In this paper, a clustering algorithm is used to classify the annual steam demand [8] The annual demand for three levels of steam can be divided into five operating conditions.

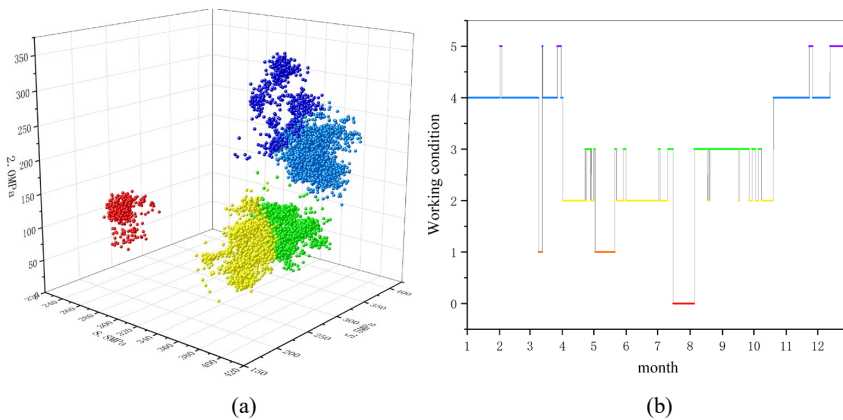


Fig.1 (a)、(b) -Results of steam demand division

The division results of annual steam demand conditions were plotted with time as the abscissa, as shown in Figure 1. There were six operating conditions including parking conditions (operating condition 0).

#### 1.1.2 Scenario generation

The principle of scenario generation is based on the random process of Markov chain, and the probability of the random occurrence is determined by the probability of transition between different states in the research object [9]. Suppose the state of steam demand sequence in the current scenario is  $\dots X_{t-2}, X_{t-1}, X_t, X_{t+1}, X_{t+2} \dots$ , Then the probability of the steam demand state at time  $X_{t+1}$  only depends on the state  $X_t$  at the previous moment:

$$P(X_{t+1} | \dots X_{t-2}, X_{t-1}, X_t) = P(X_{t+1} | X_t) \tag{1}$$

In this paper, the steam state is evenly divided into 10 states, the probability transition matrix of three levels of steam can be constructed by statistics of probabilities of transition under different working conditions. According to the state transition probability matrix, the state of the next hour can be predicted continuously until the set prediction time is generated. The prediction time in this paper is 24h.

#### 1.1.3 Scenario reduction

The number of scenarios generated is very large, if all of them are considered, the efficiency of subsequent optimization problems will be seriously affected, and feasible solutions cannot even be

obtained [10]. Therefore, the scenario reduction method is needed to eliminate redundant scenarios. The scenario reduction method used in this paper is the simultaneous backward reduction method [11]. The principle of this method is to minimize the probability distance between the initial scenario set and the scenario set saved after reduction, which makes the reduction process easier [12]. The Dr Distance model is shown below:

$$\min D_r(P, Q) = \left\{ \sum_{i \in I} \sum_{j \in J} p_i \|\varepsilon_i - \tilde{\varepsilon}_j\|^r \right\}^{1/r} \quad (3)$$

Where,  $\|\cdot\|$  is norm, the Euclidean distance is used here, so  $r=2$ .

The number of reserved scenario cuts is generally 10. Figure 2 shows the result of reducing 2000 scenarios generated by the scenario generation to 10. As you can see in Figure 2, the reduced results include all representative cases from minimum to maximum.

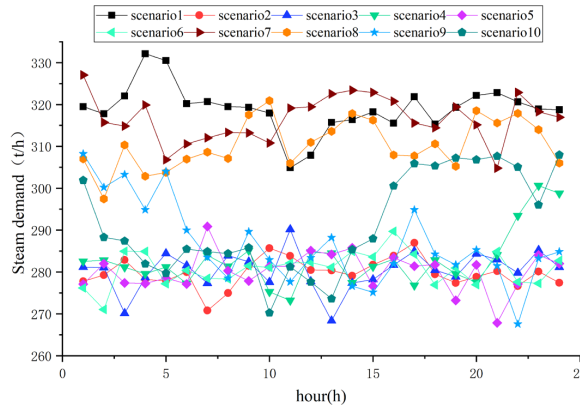


Fig.2 The results of scenario reduction

#### 1.4 Determine the number of initial scenarios $S_p$

Take the maximum and minimum values of reserved 10 scenarios at each moment to form the upper and lower bounds of the prediction interval. The prediction results of the three levels of steam are shown in Figure 3.

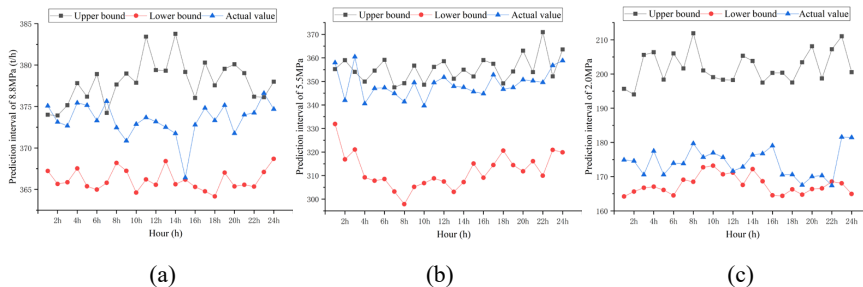


Fig.3 (a), (b), (c)-Prediction interval of 8.8MPa、5.5MPa、2.0MPa

From the prediction intervals of the three level of steam, it can be seen that the actual values of the predicted results of 8.8MPa, 5.5MPa and 2.0MPa are well within the prediction intervals. In order to meet the actual steam demand, the upper bound value is taken as the predicted value of each level of steam.

The upper bound of the prediction interval is related to the number generated in the scenario generation process. The range of the scenario generation number  $S_p$  is 500–5000. When the scenario

generation number is 5000, the upper bound value tends to be stable. Therefore, by comparing the number of other generated scenarios with the upper bound of  $Sp=5000$ , the number of scenarios generated for different levels of steam is determined, as you can see in Figure 4.

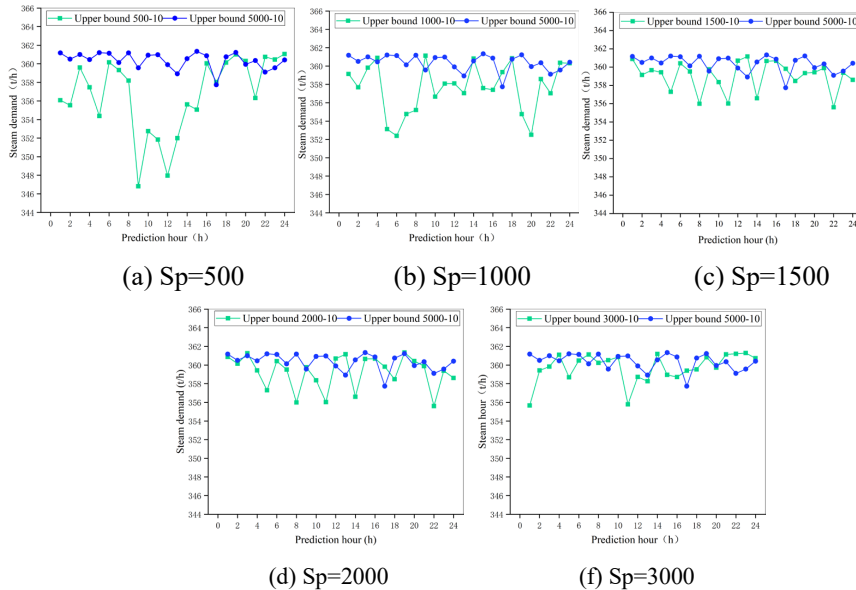


Fig.4 Comparison between the number of different generated scenarios and  $Sp=5000$

For 8.8MPa, the number of scenarios  $Sp$  is 500. In the same way, the number of scenarios  $Sp$  is 1500 for 5.5MPa, the number scenarios  $Sp$  is 1000 for 2.0MPa.

## 2. Case study

### 2.1 The prediction results of one month steam demand

The example is the steam power system of a coal-to-gas plant. According to the extracted historical data of steam demand, the prediction results for one month are as follows in Figure 5.

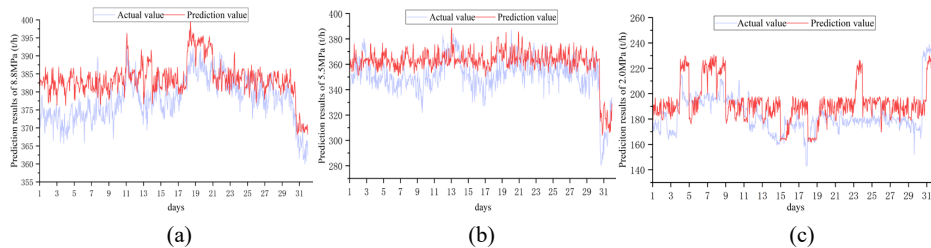


Fig.5 (a)、(b)、(c)-Prediction results of 8.8MPa、5.5MPa and 2.0MPa

As can be seen from Figure 5, the stability (the predicted value is greater than or equal to the actual value) of the three levels of steam is respectively 87.1%、79.4% and 76.2%.

### 2.2 Steam power system

The flow diagram of steam power system is shown in Figure 6. B1-B4 is the boiler. CN100T is 100MW high pressure extraction type direct air cooled turbine unit, CB30T is 30MW high pressure back pressure turbine unit. Insufficient power and high pressure steam are purchased from outside.

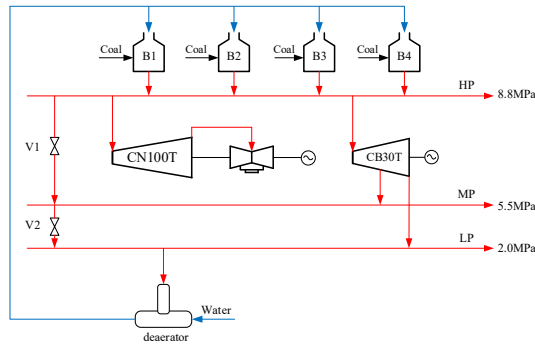


Fig.6 The flow diagram of steam power system

2.3 Analysis of optimization results

In this paper, sequential quadratic programming method was adopted to solve the steam power system model. The objective function is the operating cost, and the monthly steam demand value predicted based two-stage time series prediction model with Markov chain is brought into the optimization model. The comparison of various costs before and after optimization of each optimization method is shown in Table 1.

Table 1 Cost of each optimization method before and after optimization

Cost /10 <sup>4</sup> ¥	Before optimization	This paper optimization	Robust optimization	Stochastic programming
Coal	5871	5721	5828	5694
Water	867	834	859	827
Purchased electricity	0	0	0	0
Purchased steam	0	0	0	0
The total cost	6738	6555	6687	6521
Cost saving	—	183	51	217

As shown in Table 1, the cost savings of the three optimization methods were 1.83 million yuan, 510,000 yuan and 2.17 million yuan respectively.

The optimized stability and economy of the three optimization methods are shown in Figure 7.

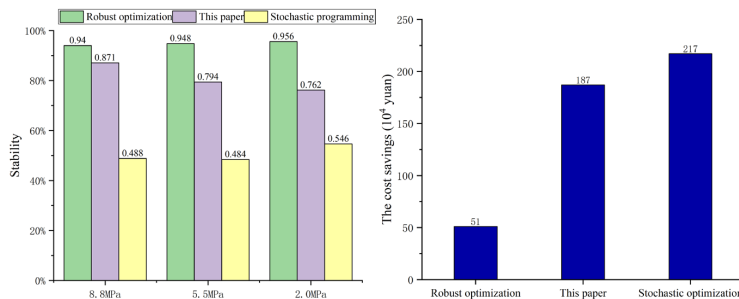


Fig.7 The optimized stability and economy of the three optimization methods

By comparing the stability and economy of the three optimization methods, the two-stage time series prediction optimization model with Markov chain in this paper takes into account the advantages of high economy of stochastic programming and high stability of robust optimization, and its stability and



economy are between stochastic programming and robust optimization.

### 3. Conclusion

In this paper, a two-stage time series prediction optimization model with Markov chain method is proposed. The first stage is based on spatial distance expression to divide multiple operating conditions, and the second stage is based on state transition probability expression to predict the demand value of steam 24h before the day through the method of scenario generation and reduction. This method is used to solve the operation optimization problem of a steam power system of coal-to-gas. The optimization results show that:

- (1) The stability of the three levels of steam demand prediction is between that of stochastic programming and robust optimization.
- (2) The cost savings of one-month prediction optimization can save 1.83 million yuan, and its economy is between the stochastic programming and robust optimization.
- (3) The proposed method combines the advantages of robust optimization and stochastic programming and has better performance.

### References

- [1] LI S, JIANG X B, HE G H, et al, 2017, Research progress for flexible design and multi-objective optimization of steam power network [J], *Chemical Industry and Engineering Progress*, **36**, *6*, 1989-1996.
- [2] ALIPOUR M, ZARE K, SEYEDI H , 2018, A multi follower Bi-level stochastic programming approach for energy management of combined heat and power micro-grids [J], *Energy*, **149**, *15*, 135-146.
- [3] Saeedi M, Moradi M, Hosseini M , Emamifar A , Ghadimi N, 2019, Robust optimization based optimal chiller loading under cooling demand uncertainty[J], *Apply Thermal Engineering*, **148**, 1081–1091.
- [4] Sun L, Gai L, Smith R, 2017, Site utility system optimization with operation adjustment under uncertainty[J], *Applied Energy*, **186**, 450-456.
- [5] Xie Y L, Huang G H, Li W, et al, 2013, An inexact two-stage stochastic programming model for water resources management in Nansihu Lake Basin, China[J], *Journal of environmental management*, **127**, 188-205.
- [6] Niu T, Yin H, Feng E, 2022, An interval two-stage robust stochastic programming approach for steam power systems design and operation optimization under complex uncertainties[J], *Chemical Engineering Science*, **253**, 117533.
- [7] Shen F, Zhao L, Du W, et al, 2020, Large-scale industrial energy systems optimization under uncertainty: A data-driven robust optimization approach[J], *Applied Energy*, **259**, 114199.
- [8] Sinaga K P, Yang M S, 2020, Unsupervised K-means clustering algorithm[J], *IEEE access*, *8*: 80716-80727.
- Balster H. Markov chain models for vegetation dynamics[J]. *Ecological modelling*, **126**, *2-3*, 139-154.
- [9] Lee D, Baldick R, 2016, Load and wind power scenario generation through the generalized dynamic factor model[J], *IEEE Transactions on power Systems*, **32**, *1*, 400-410.
- [10] Dupačová J, Gröwe-Kuska N, Römisich W, 2003, Scenario reduction in stochastic programming[J], *Mathematical programming*, **95**, *3*, 493-511.
- [11] Chen Z, Yan Z, 2018, Scenario tree reduction methods through clustering nodes[J], *Computers & Chemical Engineering*, **109**, 96-111.

# **An approach for modelling simultaneous fluid-phase and chemical-reaction equilibria in multicomponent systems via Lagrangian duality: The Reactive HELD algorithm.**

Felipe A. Perdomo<sup>a</sup>, George Jackson<sup>a</sup>, Amparo Galindo<sup>a</sup>, Claire S. Adjiman,<sup>a</sup>

*<sup>a</sup>Department of Chemical Engineering, Sargent Centre for Process Systems Engineering, Institute for Molecular Science and Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom*

## **Abstract**

An approach for the calculation of simultaneous phase and chemical-reaction equilibria and stability that does not require any assumptions on the number of stable phases and is applicable to any number of reactions is presented. It is based on an extension of the dual extremum principle concept for non-reactive systems (HELD) formulated in terms of the Helmholtz energy, where additional constraints are introduced that relate molar amounts and extents of reaction. The extended R-HELD algorithm is applied to reacting mixtures such as the esterification of acetic acid and the formation of methyl tert-butyl ether.

**Keywords:** Phase stability, reacting mixtures, SAFT- $\gamma$  Mie equation of state.

## **1. Introduction**

The calculation of equilibria in multiphase, multi-reaction systems is an important subject of study in chemical engineering thermodynamics. The inclusion of chemical reaction equilibria in multiphase equilibria gives rise to significant challenges in terms of thermodynamic consistency and numerical complexity. In reacting systems, stable phases may vanish or emerge as a consequence of the chemical reactions, making it difficult to determine the number and compositions of the phases coexisting at equilibrium. Two main classes of approaches have been presented to overcome these issues (Smith, 1980; Seider and Widagdo, 1996). In the first class, a non-linear programming approach is used to minimize the Gibbs energy function with a fixed number of phases (Castillo and Grossmann, 1981; Lantagne, et al., 1988). In the second class, phases are added (or removed) iteratively and the stability of each solution is tested (Gautam and Seider, 1979). This has led to approaches to solve simultaneously multiphase stability and equilibria in the presence of chemical reactions; (Michelsen, 1989; Gupta et al., 1991; Smith et al., 1993; Jiang et al, 1996). These approaches allow for a good compromise between effectiveness and efficiency, but the search for the correct number of stable phases can lead to numerous iterations or convergence to local minima. In our current work, an approach is developed to obtain the stable state comprising one or more phases without making initial assumptions on the number of phases nor restricting the number of components in each phase. To this end, we extend the seminal work of Mitsos and Barton (2007) on the determination of the phase stability of non-reactive multicomponent mixtures via the dual-extremum principle (DEP). We adopt the Helmholtz energy formalism of Pereira et al. (2010; 2012), who recast the DEP in terms of temperature,

pressure, and mole numbers as natural variables, a formulation which is better-suited to the use of state-of-the-art equations of state such as SAFT- $\gamma$  Mie. Building on the DEP, a new phase stability and equilibria algorithm, HELD (Helmholtz energy Lagrangian dual) algorithm is developed, which we now extend to the reactive systems (R-HELD). In Section 2, we outline the extension of the DEP to reacting systems. In Section 3, we embed this stability criterion within the R-HELD algorithm. The performance of the algorithm on reacting systems with one or two phases is illustrated in Section 3.

## 2. A phase stability criterion for systems at chemical equilibrium

The extension of the DEP of Mitsos and Barton (2007) and of its Helmholtz energy counterpart (Pereira et al. 2010) to reacting systems is presented without proof. We formulate an optimization (dual) problem such that the set of all global solutions admitted by the dual problem is exactly equal to the set of stable phases of the mixture of interest. We consider the (trivial) single-phase minimisation of the Gibbs energy for a reacting system. Given  $C$  components,  $R$  independent reactions, temperature  $T_o$ , pressure  $P_o$  and total mole fractions  $N_{i,0}$ ,  $i = 1, \dots, C$ , the minimisation of interest can be expressed using the molar volume,  $V$ , and the vector  $\Xi$  of (molar) extents of reaction,  $\xi_r$ ,  $r = 1, \dots, R$ , as independent variables and including a material-balance constraint:

$$G_V = \min_{\Xi, V} A(T_o, P_o, \mathbf{N}(\Xi)) + P_o V$$

$$\text{s. t. } N_i(\Xi) - (N_i^0 + \sum_{r=1}^R \nu_{i,r} \xi_r) = 0, i = 1, \dots, C, \quad (1)$$

$$V \in [\bar{V}, \bar{\bar{V}}], \Xi \in \mathbb{R}_+^R$$

where  $A$  is the Helmholtz energy,  $N_i$  is the number of moles of species  $i$ ,  $\nu_{i,r}$  is the stoichiometric coefficient of component  $i$  in reaction  $r$ ,  $\bar{V}$  and  $\bar{\bar{V}}$  are lower and upper bounds on the volume, respectively. The application of the first-order optimality conditions for Problem (1) results in

$$-\left(\frac{\partial A(T_o, V^*, \mathbf{N}(\Xi^*))}{\partial V}\right)_{T, \xi_{r=1, \dots, R}} = P(T_o, V^*, \mathbf{N}(\Xi^*)) = P_o, \quad (2)$$

indicating that the molar volume at the solution  $(\Xi^*, V^*)$  corresponds to pressure  $P_o$  (Nagarajan et al., 1991), and

$$\sum_{i=1}^C \mu_i(T_o, V^*, \mathbf{N}(\Xi^*)) \nu_{i,r} = 0, \quad r = 1, \dots, R, \quad (3)$$

enforcing chemical-reaction equilibria for all  $R$  reactions. Thus, these conditions do not need to be stated explicitly as constraints, and Problem (1) is equivalent to the minimisation of the Gibbs energy at  $(T_o, P_o, \mathbf{N}_0)$  for a phase at chemical equilibria.

We now recast Problem (1) in terms of the set of transformed compositional variables proposed by Ung and Doherty (1995). These  $C - R$  compositional variables exhibit useful properties such as independence of the extents of reaction, summation to unity, and thermodynamic consistency. Using the summation to unity property, variable  $C - R$  can be eliminated. The resulting vector of  $C' = C - R - 1$  variables is denoted by  $\hat{\mathbf{x}}$ . The primal problem of Pereira et al. (2010) can then be reframed as

$$\widehat{G}_V^p = \min_{\hat{\mathbf{x}}, V} A(T_o, V, \hat{\mathbf{x}}) + P_o V$$

$$\text{s. t. } \hat{x}_{i,0} - \hat{x}_i = 0, i = 1, \dots, C'$$

$$\hat{\mathbf{x}} \in \mathcal{X}, V \in [\bar{V}, \bar{\bar{V}}] \quad (4)$$

where  $\mathcal{X}$  defines the domain of the compositional variables. The Lagrange function, corresponding to Problem (5), is given by

$$\widehat{L}^V(\widehat{\mathbf{x}}, V, \boldsymbol{\lambda}) = A(T_o, V, \widehat{\mathbf{x}}) + P_o V + \sum_{i=1}^{C'} \lambda_i (\widehat{x}_{i,o} - \widehat{x}_i), \quad (5)$$

and the dual of Problem (4) is given by

$$\widehat{G}^D = \max_{\boldsymbol{\lambda} \in R^{C'}} \min_{\widehat{\mathbf{x}} \in \mathcal{X}, V \in [\overline{V}, \overline{V}]} \widehat{L}^V(\widehat{\mathbf{x}}, V, \boldsymbol{\lambda}) \quad (6)$$

Under the assumption that the solutions of the inner problem in Problem (6), i.e., the minimisation of the Lagrange function, lie in the interior of sets  $\mathcal{X}$  and  $[\overline{V}, \overline{V}]$ , the solution set of Problem (6) describes the stable state of the reacting mixture at  $(T_o, P_o, \mathbf{N}_o)$ .

### 3. The reactive HELD (R-HELD) algorithm for reacting systems

Using Problem (6), we extend the implementation of the HELD algorithm (Pereira et. al, 2012) to enable the identification of all stable phases of systems undergoing chemical reactions. The implementation makes use of a multistart approach combined with a SQP solver to solve any nonlinear problem. Thus global optimality cannot be guaranteed.

#### 3.1. Stage 1: Initialization.

Primal problem (4) admits a trivial solution at  $\widehat{\mathbf{x}} = \widehat{\mathbf{x}}_o$ , where  $\widehat{\mathbf{x}}_o$  is found by solving for chemical equilibria at the specified conditions. This yields an upper bound on the minimum Gibbs energy of the system. The tangent-plane distance criterion is applied at  $\widehat{\mathbf{x}}_o$  to determine whether the single phase is stable. If a negative value of the tangent-plane distance is found, the algorithm proceeds to Stage 2. Otherwise, it terminates.

#### 3.2. Stage 2: Identification of candidate stable phases

An iterative search for all global solutions of Dual problem (6) is carried out. At iteration  $k$ , (i) given a guess vector  $\boldsymbol{\lambda}^k$ , the (inner) minimisation of the Lagrange function yields a lower bound (LBD<sup>V</sup>) on the minimum Gibbs free energy and a solution  $(\widehat{\mathbf{x}}^k, V^k)$ , which is used to construct an approximation of the inner problem by substitution into Eq. (5) and, (ii) an approximate Dual problem, constructed from all solutions to the inner problem up to iteration  $k$  is solved with respect to the  $C'$  Lagrange multipliers. This outer problem yields an upper bound (UBD<sup>V</sup>) on the minimum Gibbs energy and a new guess  $\boldsymbol{\lambda}^{k+1}$ . The iterations continue until the gap between the lower and upper bounds converges to a specified tolerance,  $\varepsilon$ , i.e.,  $|\text{UBD}^V - \text{LBD}^V| \leq \varepsilon$ .

#### 3.3. Stage 3: Material balance

The total Gibbs energy, calculated using all the phases identified so far, is minimised, subject to the overall mass balance. If the problem is found to be feasible, further criteria are applied, including checking the gap between the total Gibbs energy and the upper bound found in Stage 2 and ensuring that the necessary conditions for phase and chemical equilibria are satisfied across all species. This latter check is important to ensure the accuracy of the calculations – the convergence tolerance can be tightened if this is not met and Stage 2 to achieve increased accuracy / identify missing phases.

## 4. Results

We assess the performance of the extended R-HELD algorithm for several multicomponent reactive systems at conditions where the system is stable as in a single phase or coexists in several phases in the presence of chemical reactions. Our approach

only requires knowledge of the stoichiometry of each reaction and the enthalpies and entropies of formation of each species at the selected standard state, here, the ideal gas at 298.15 K and 1.013 bar. All mixtures are modelled with SAFT- $\gamma$  Mie equation of the state (Papaioannou et al., 2014) with parameters reported in Haslam et al. (2021) and in the supporting information.

#### 4.1. Single stable phase with reaction equilibria

We consider the liquid phase synthesis of methyl tert-butyl ether (MTBE) from isobutene (IB) and methanol (MeOH) at 16 bar. The predicted chemical-reaction equilibria surfaces (CES) are presented in Figure 1 for temperatures from 323.15 K to 353 K and initial methanol/isobutene molar ratio from 0.1 to 1.6, with data from Izquierdo et al. (1992).

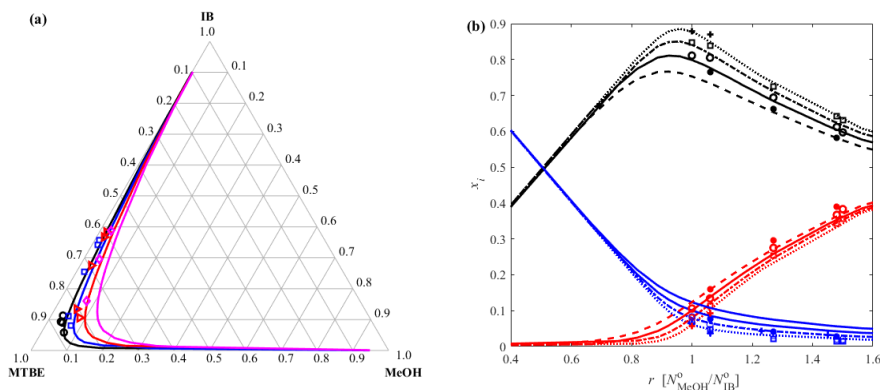


Figure 1: Reaction equilibria for the system isobutene, methanol, and methyl tert-butyl ether at 16 bar: (a) CES at 323.15 K (black circles), 333.15 K (blue squares), 343 K (red triangles), and 353 K (magenta diamonds). (b) Composition profiles for reactants and products as a function of initial MeOH/IB ratio: IB (blue), MeOH (red), and MTBE (black) at 323.15 K (crosses), 333.15 K (squares), 343 K (circles), and 353 K (filled circles). Continuous curves are predictions using SAFT- $\gamma$  Mie and R-HELD and symbols are experimental data (Izquierdo et al. 1992)

The quaternary reactive mixture arising from the esterification of acetic acid with ethanol acetic acid [(Ac.Ac) + ethanol (EtOH)  $\rightleftharpoons$  ethyl acetate (Et.AcO) + water (H<sub>2</sub>O)] is modelled with R-HELD at conditions where one phase is present (1.013 bar and 293.15 K). The predicted CES is presented in Figure 2 as a concentration tetrahedron and is seen to be in excellent agreement with the experimental data of Golikova et al. (2017). The CES is seen to limit feasible compositional space that can be reached.

#### 4.2 Simultaneous phase and chemical reaction equilibria

We present in Figure 3 the calculation at conditions at which phase and chemical equilibria interact for the mixture IB, MeOH, MTBE (5 bar, 342.25 to 378.29 K), together with the molecular simulation data of Lisal et al. (2010). In Figure 3a, the CESs in the vapour and liquid phase (blue) and the bubble and dew curves (black) are seen to intersect at the solutions of the simultaneous phase- and chemical-equilibrium problem.

## 5. Conclusions and outlook

We have extended the dual extremum principle as a criterion for fluid-phase stability, as proposed by Mitsos and Barton (2007), to include chemical reactions. Specifically, we have extended the Helmholtz energy formalism proposed by Pereira et al. (2010), combined with the compositional transforms of Ung and Doherty (1995). The approach,

presented here without proof, is consistent, general, and can be applied to multicomponent and multi-reaction systems coexisting at phase equilibria. We have implemented this within the framework of the HELD algorithm for non-reactive systems (Pereira et al., 2012). The extended R-HELD algorithm has been tested successfully with ternary and quaternary mixtures with one or two stable phases, using the SAFT- $\gamma$  Mie equation of state. Broadening the scope of the HELD algorithm allows us to consider challenging mixtures that exhibit simultaneous phase and reaction equilibria.

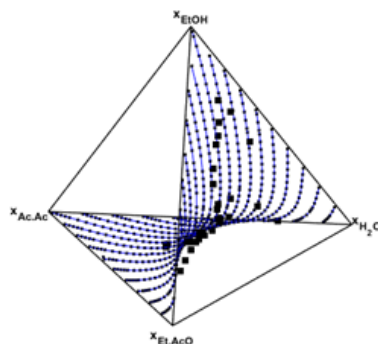


Figure 2 - CES for the esterification reaction of acetic acid 323.15 K and 1.013 bar computed with SAFT- $\gamma$  Mie and R-HELD (continuous-dotted envelope) and measured experimentally (symbols)

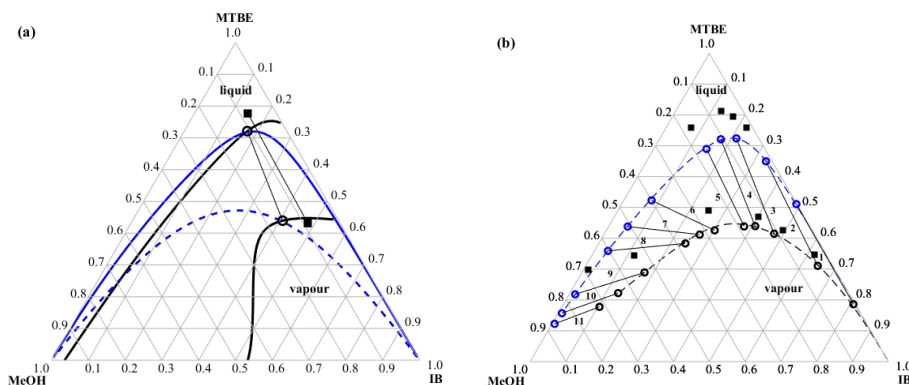


Figure 3: Mixture of IB, MeOH and MTBE at 5 bar (a) CES at 360 K. Black continuous curves indicate non-reactive VLE; continuous and dotted blue curves represent the prediction of the single-phase chemical reaction equilibria (CES) in the liquid and vapour phase respectively. The light solid lines represent reactive tie lines. (b) Complete simultaneous reaction + phase equilibria, with reactive tie-lines (circles connected by solid lines) 1) 342.25 K, 2) 351.8 K, 3) 358.4 K, 4) 360.15 K, 5) 360.65 K, 6) 363.25 K, 7) 365.19 K, 8) 367.15 K, 9) 372.81 K, 10) 376.13 K, 11) 378.29 K. Filled black squares are molecular simulation data (Lisal, et al., 2000).

### Acknowledgments

The authors are grateful for financial support from Eli Lilly & Company and the UK Engineering and Physical Sciences Research Council (EP/T005556/1).

## Data statement / Supplementary information

All data related to this article and not published elsewhere can be found at doi: 10.5281/zenodo.7418447 and used under the Creative Commons Attribution licence.

## References

- J. Castillo, and I.E., Grossmann, 1981. Computation of phase and chemical equilibria. *Comput. Chem. Eng.* 5(2), 99-108.
- A. Golikova, A. Samarov, M. Trofimova, S. Rabdano, M. Toikka, O. Pervukhin, and A. Toikka, 2017. Chemical equilibrium for the reacting system acetic acid–ethanol–ethyl acetate–water at 303.15 K, 313.15 K and 323.15 K. *J. Solution Chem.*, 46(2), 374-387.
- A.K. Gupta, P.R. Bishnoi, and N. Kalogerakis, 1991. A method for the simultaneous phase equilibria and stability calculations for multiphase reacting and non-reacting systems. *Fluid Phase Eq.*, 63(1-2), 65-89.
- A.J. Haslam., A. González-Pérez, S. Di Lecce, S.H. Khalit, F.A. Perdomo, S. Kournopoulos, M. Kohns, T. Lindeboom, M. Wehbe, S. Febra, and G. Jackson, 2020. Expanding the Applications of the SAFT- $\gamma$  Mie Group-Contribution Equation of State: Prediction of Thermodynamic Properties and Phase Behavior of Mixtures. *J. Chem. Eng. Data*, 65(12), 5862-5890.
- J.F. Izquierdo, F. Cunill, M. Vila, J. Tejero, and M. Iborra, 1992. Equilibrium constants for methyl tert-butyl ether liquid-phase synthesis. *J. Chem. Eng. Data*, 37(3), 339-343.
- Y. Jiang, G.R. Chapman, and W.R. Smith, 1996. On the geometry of chemical reaction and phase equilibria. *Fluid Phase Eq.*, 118(1), 77-102.
- G. Lantagne, B. Marcos, and B. Cayrol, 1988. Computation of complex equilibria by nonlinear optimization. *Comput. Chem. Eng.* 12(6), 589-599.
- M. Lísal, W.R. Smith, and I. Nezbeda, 2000. Molecular simulation of multicomponent reaction and phase equilibria in MTBE ternary system. *AIChE J.*, 46(4), 866-875.
- M.L. Michelsen, 1989. Calculation of multiphase ideal solution chemical equilibrium. *Fluid Phase Eq.*, 53, 73-80.
- A. Mitsos, and P.I. Barton, 2007. A dual extremum principle in thermodynamics. *AIChE J.*, 53(8), 2131-2147.
- N.R. Nagarajan, A.S. Cullick, and A. Griewank, 1991. New strategy for phase equilibrium and critical point calculations by thermodynamic energy analysis. Part I. Stability analysis and flash. *Fluid Phase Eq.*, 62(3), 191-210.
- V. Papaioannou, T. Lafitte, C. Avendaño, C.S. Adjiman, G. Jackson, E.A. Müller, and A. Galindo, 2014. Group contribution methodology based on the statistical associating fluid theory for heteronuclear molecules formed from Mie segments. *The J. Chem. Phys.*, 140(5), 054107.
- F.E. Pereira, G. Jackson, A. Galindo, and C.S. Adjiman., 2010. A duality-based optimisation approach for the reliable solution of (P, T) phase equilibrium in volume-composition space. *Fluid Phase Eq.*, 299(1), 1-23.
- F.E. Pereira, G. Jackson, A. Galindo, and C.S. Adjiman., 2012. The HELD algorithm for multicomponent, multiphase equilibrium calculations with generic equations of state. *Comput. Chem. Eng.* 36, 99-118.
- W.D. Seider, and S. Widagdo, 1996. Multiphase equilibria of reactive systems. *Fluid Phase Eq.*, 123(1-2), 283-303.
- W.R. Smith, 1980. The computation of chemical equilibria in complex systems. *Ind. Eng. Chem. Fund.*, 19(1), 1-10.
- J.V. Smith, R.W. Missen, and W.R. Smith, 1993. General optimality criteria for multiphase multireaction chemical equilibrium. *AIChE J.*, 39(4), 707-710.
- S. Ung, and M.F. Doherty, 1995. Theory of phase equilibria in multireaction systems. *Chem. Eng. Sci.*, 50(20), 3201-3216.
- S. Ung, and M.F. Doherty, 1995. Vapor-liquid phase equilibrium in systems with multiple chemical reactions. *Chemical Engineering Science*, 50(1), 23-48.

# Optimization of the Biomass Supply Chain for Power Generation with the Steam Rankine Cycle

Monika Dokl, Zdravko Kravanja, Lidija Čuček

*Faculty of Chemistry and Chemical Engineering, University of Maribor, Smetanova ulica 17, 2000 Maribor, Slovenia*  
*lidija.cucek@um.si*

## Abstract

The steam Rankine cycle (SRC) is a mature technology adapted to high temperatures and has a high electrical efficiency. However, the cost-effective and environmentally-sound utilisation of biomass in SRC requires efficient management of the supply chain and overcoming challenges such as complex conversion processes and uncertainties in biomass quality and supply. In this study, the optimisation of biomass utilisation using SRC to produce electricity is conducted considering the entire biomass supply chain. The model is developed as a combined biomass supply chain network and SRC process flowsheet and optimized in GAMS to achieve the most effective power generation while optimising the network economics. The concept of the SRC process-based approach to biomass supply network is evaluated through a demonstrative case study of a relatively small region.

**Keywords:** supply chain optimisation, mathematical modelling, biomass utilisation, steam Rankine cycle

## 1. Introduction

The depletion of fossil fuels, their negative impact on the environment and the security of energy supply have stimulated many countries to focus on more extensive use of renewable energy resources to reduce environmental burdens and achieve energy security and fuel diversification. Among renewable energy sources, biomass is becoming increasingly important, as it can be used for heat and power generation and biofuel production. Apart from biofuels, the most widely used technology for biomass utilisation is direct combustion, where the resulting high-temperature gases are typically used to generate high-pressure steam. Power generation using steam cycles such as the steam Rankine cycle (SRC) is suitable for large-scale power plants where thermal efficiencies of over 30 % can be achieved. SRC uses water as working fluid, which has the advantage of being highly available and non-toxic, but it precludes its wide range of applications, especially for low and medium grade heat sources (Abbas et al., 2020).

In many situations, biomass suitable for power generation is available as a feedstock in large quantities with low density and largely distributed production, characterised by its seasonal availability. Achieving competitive electricity generation through biomass utilization requires an efficient management of the biomass supply network, which primarily addresses biomass harvesting, transportation, storage, and utilisation. Biomass supply chain optimisation focuses on supply chain performance and configuration, component sizing and operations scheduling, usually with an objective of minimising cost (Nunes et al., 2020).



In this study a new supply network concept is proposed that considers integrated biomass supply network formulation and detailed optimization of SRC process. The objective of the study is maximising the revenue from selling heat and electricity and subtracting distance-dependent biomass transportation cost and cost of piping network to distribute heat.

## 2. Methodology

The proposed approach consists of optimizing the integrated biomass supply network and SRC process flowsheet system. First, the proposed integrated design is explained, and further mathematical formulation is briefly presented.

### 2.1. Biomass Supply Network and Steam Rankine Cycle

An integrated biomass supply network and SRC process, schematically represented in Fig. 1, consists of 3 main layers (L1-L3). The first layer (L1) represents the supply of feedstocks to the plant and accounts for the regional characteristics of each zone ( $i$ ) and biomass storage at locations  $s$ . The second layer (L2) represents the SRC plants which are located at sites  $j$ . SRC processes are formulated as a flowsheet structure consisting of two heat exchangers (boiler and condenser) and two pressure changers (turbine and pump). The formulation enables consideration of details, such as optimization of temperatures, pressures, flowrates, specific enthalpies and energy flows. Generated products by SRC plants (heat and electricity) are then distributed to the third layer (L3) at sites  $d$ . Road transport is considered as the transportation mode for biomass to SRC utilisation sites, while pipeline is used for delivery of hot water to demand sites.

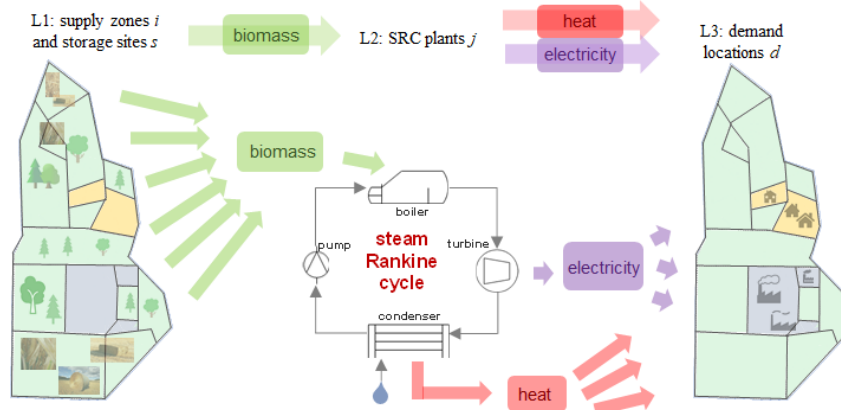


Figure 1: Schematic representation of integrated biomass supply network and SRC

### 2.2. Model Formulation

The developed model for optimization of biomass supply chain for maximal profit of sold electricity and heat produced by SRC reduced by variable transportation cost comprises material and energy balances and economic constraints (transportation cost, revenue) for the supply and demand layers, interconnections supply - process and process - demand and optimization of SRC which accounts for optimizing the operating conditions of the cycles. The approach for simultaneous consideration of supply network and process flowsheet optimization is based on the work by Kegl et al. (2021) and the approach for optimizing operating conditions in the SRC is adapted from Dokl et al. (2022).

For simplification, the main objective ( $Z$ ) is to maximize revenue obtained from SRC ( $R_{SRC,i}$ ) subtracted by transportation cost generated from the supply network ( $c_{tr}$ ) (Eq. (1)).

$$Z = \max(\sum_{j=1}^n R_{\text{SRC}_j} - c_{\text{tr}}) \quad (1)$$

Such objective is selected to identify a simple economic bound of the system (Biegler et al., 1997), while considering only such costs that the selected distances between possible connections will be minimal. Transportation cost is for simplification thus calculated only based on the distance variable costs, and accounts for factors related to return trips and tortuosity factor (Egieya et al., 2019). The obtained revenue of plant  $j$  is based on production of electricity and heat (Eq. (2)). Electricity generation in the turbine depends on the fluid flow rate ( $q_j^{\text{turb}}$ ), specific enthalpy difference ( $h_j^{\text{turb,out}} - h_j^{\text{turb,in}}$ ) and turbine efficiency ( $\eta_j^{\text{turb}}$ ), while the obtained heat in the form of hot water at the condenser outlet is described by the heat duty of the condenser ( $Q_j^{\text{cond}}$ ) reduced by the transportation heat loss ( $L_{\text{tr}}^{\text{heat}}$ ) between plant  $j$  and demand site  $d$  ( $D_{j,d}$ ).

$$R_j = (q_j^{\text{turb}} \cdot (h_j^{\text{turb,out}} - h_j^{\text{turb,in}}) \cdot \eta_j^{\text{turb}}) + (Q_j^{\text{cond}} \cdot (1 - L_{\text{tr}}^{\text{heat}})^{D_{j,d}}) \quad (2)$$

The optimized operating conditions are obtained by correlating thermophysical properties with temperature and/or pressure. Such properties are specific enthalpies of water and steam (Fig. 2), vapor liquid pressure, specific heat capacity at constant pressure and heat capacity ratio.

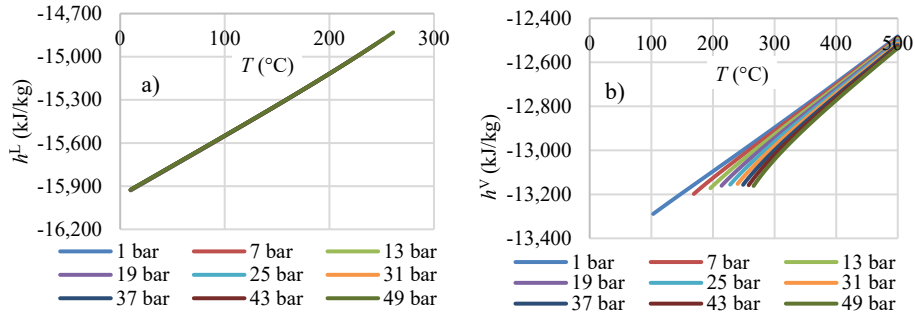


Figure 2: Specific enthalpy of a) water and b) steam, changing with temperature and pressure

### 3. Case Study and Results

The integrated biomass supply network – SRC flowsheet optimization model is implemented on a demonstration case study of a relatively small region. For simplification, the model assumes single time-period and the average values of parameters in a year. The map of the case study is shown in Fig. 3, with a total area of 906.6 km<sup>2</sup>. Up to 17 supply ( $i_1 - i_{17}$ ), 8 storage ( $s_1 - s_8$ ), 5 plant ( $j_1 - j_5$ ) and 5 demand locations ( $d_1 - d_5$ ) (built-up and industrial areas) could potentially be selected.

The actual areas for growing biomass in zones (forest and agricultural areas) are between 10 and 90 % of the zone areas. The yield of wood is based on MKGP (2011) and of agricultural residues from Jejičič and Verbič (2019). Only the main biomass feedstocks  $pb$  are considered, which are wood, corn stover, grain straw and oilseed straw with the yields of 3.21 t/(ha·y) for wood, 5.41 t/(ha·y) for grain straw, 5.59 t/(ha·y) for corn stover and 5.72 t/(ha·y) for oilseed straw. The distribution of area for biomass in the municipality is 87.3 % for wood, 8.6 % for grain straw, 3.9 % for corn stover and 0.2 % for oilseed straw.

For the SRC, all the characteristics of process units and streams were optimized to satisfy the given objective. The only assumptions are minimum temperature difference in condenser, which is 5 °C for all the cycles, 85 % mechanical efficiency of turbines, maximum flow rate of water in the condenser of 200 kg/s (Yıldırım et al., 2006), and temperature of water at condenser inlet (50 °C) and outlet (90 °C) (Čuček et al., 2015). The MINLP model consists of 3852 single equations, 5026 single variables, and 25 binary variables and was solved using DICOPT solver in GAMS (GAMS, 2022). The solutions are provided in around 10 s on a personal computer with an Intel® Core™ i7-10750 H CPU @ 2.60 GHz processor with 16 GB RAM. For the calculation of the revenue from sold electricity and heat, price of electricity of 222.48 €/MWh was assumed as the average day-ahead Slovenia auction trading price of November 2022 (BSP Energy Exchange, 2022) and price of heat of 101.07 €/MWh, which is a recent price for the supply of natural gas for business consumption (Energetika Ljubljana, 2022).

The results from the breakdown of optimal quantities of feedstocks selected and related transportation cost together with optimal conditions of selected SRC plants ( $j_1$ ,  $j_3$  and  $j_4$ ) are shown in Table 1. From Table 1 it can be seen that optimal biomass supply chain structure and operating conditions of SRC plant can generate the total annual revenue of 67.89 M€, of which the highest revenue is obtained from heat (29.12 M€/y) and electricity (8.53 M€/y) by SRC plant  $j_3$ .

Optimal turbine working conditions are the same for all selected SRC plants, where steam enters the turbine with temperature of 284.7 °C and pressure of 5.8 bar and passes to the condenser with temperature of 103.9 °C and pressure of 1.1 bar. Higher heat and electricity rates in plant  $j_3$  correspond to higher flows of working fluid and water in the condenser, reaching the upper limit of 200 kg/s. SRC plant  $j_4$  has for 14 M€/y lower revenue than plant  $j_3$ , while with plant  $j_1$  only 6.6 M€/y of revenue is obtained. Distance-dependent transportation cost presents 7.98 % of revenue for plant  $j_1$ , 2.2 % for plant  $j_3$ , and only 0.72 % for  $j_4$ . Fig. 4 further shows optimised supply chain superstructure, presenting relation between layers with optimal selection and connections between sites of biomass supply, storage sites, utilisation plants, and demand sites.

Three storage locations are selected for storing feedstocks, where  $s_2$  receives biomass from 8 zones,  $s_6$  from 9 zones, one of which ( $i_{16}$ ) transport biomass also to  $s_8$ . Feedstock is further transported to the selected utilisation plant  $j_1$ ,  $j_3$ , and  $j_4$ , delivering produced heat to demand sites  $d_1$ ,  $d_3$ , and  $d_4$ .

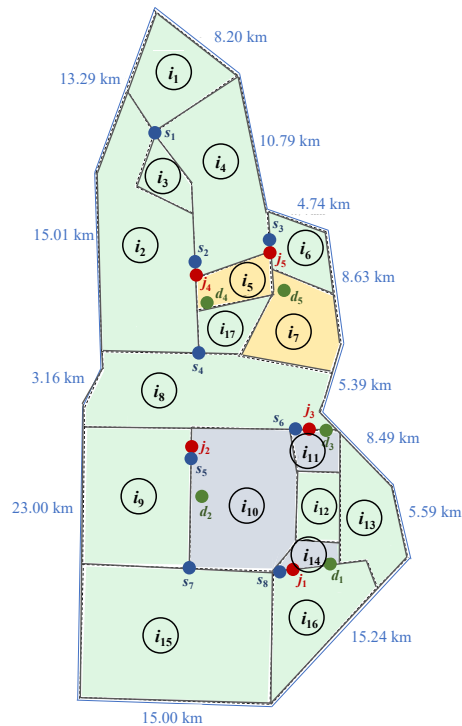


Figure 3: Map of demonstrative case study with areas of zones, storage, plant and demand sites

Table 1: Optimal values of variables for selected SRC plants

Optimal values	$j_1$	$j_3$	$j_4$
<b>Steam Rankine cycle</b>			
$Q_{\text{boiler}}$ (kW)	6879.1	38,670.9	24,420.2
$W_{\text{turbine}}$ (kW)	778.4	4375.7	2763.2
$F_{\text{steam}}$ (kg/s)	2.6	14.8	9.3
$F_{\text{water}}$ (kg/s)	35.6	200.0	126.3
$T_{\text{turbine}}^{\text{in}}   T_{\text{turbine}}^{\text{out}}$ (°C)		284.7   103.9	
$p_{\text{turbine}}^{\text{in}}   p_{\text{turbine}}^{\text{out}}$ (bar)		5.8   1.1	
<b>Economic parameters</b>			
Road transport cost (M€/y)	0.414	0.771	0.093
Pipeline cost (M€/y)	0.113	0.056	0.077
Revenue from heat (M€/y)	5.08	29.12	18.26
Revenue from electricity (M€/y)	1.52	8.53	5.39
Objective value (M€/y)	66.37		

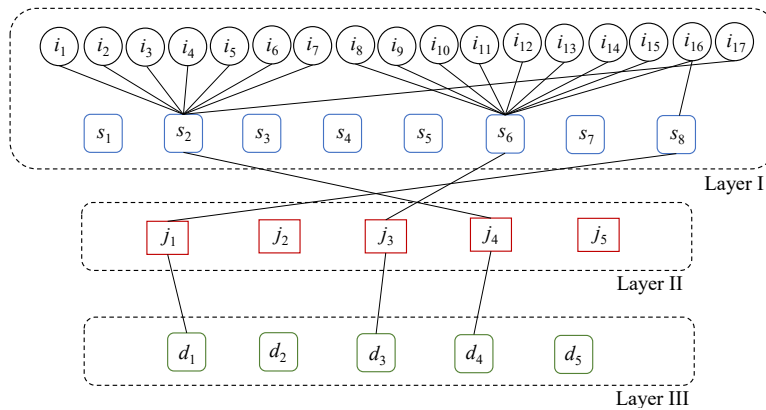


Figure 4: Biomass supply chain superstructure

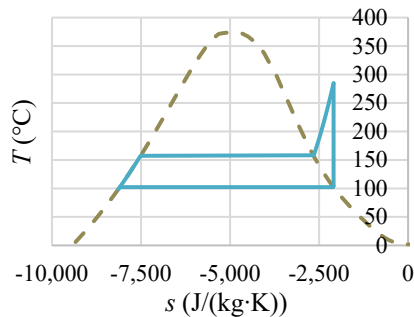


Figure 5:  $T-s$  diagram for SRC at selected plants

In Fig. 5 temperature vs. entropy ( $T-s$ ) diagram for all the selected SRC plants is shown, representing energy transfer within the cycles. The lower horizontal blue line illustrates steam condensing in the condenser, which is further passed to the boiler, where its temperature is increased to 284.7 °C, extending high above saturation curve (grey dashed curve) as it is being superheated for about 126 °C. The expansion in the turbine is shown by vertical blue line, also closing the cycle by entering the condenser again.

#### 4. Conclusions and Recommendations

In this paper the integration of biomass supply network and SRC flowsheet is performed by optimising the revenue from selling electricity and heat subtracted by transportation

cost, and demonstrated on an illustrative case study of a relatively small region. The optimal structure of biomass supply network at optimal operating conditions of selected SRC plants can annually produce almost 8 MW of electricity and about 1300 t/h of hot water for district heating at supply and return temperatures of 90 and 50 °C. Optimal supply chain consists of 17 biomass supply zones, 3 storage locations, 3 SRC plants for biomass utilisation and 3 demand sites of heat consumption. In the future, a multiperiod model could be developed to account for the seasonal characteristics of biomass supply and the complete economic and environmental performance of the studied superstructure could be evaluated.

### Acknowledgments

The authors are grateful for funding support from the Slovenian Research Agency (research core fundings No. P2-0412, P2-0421 and P2-0414 and project No. J7 3149).

### References

- T. Abbas, M. Issa, A. Ilinca, 2020, Biomass cogeneration technologies: A review, *Journal of Sustainable Bioenergy Systems*, 10(1), 1-15.
- L.T. Biegler, I.E. Grossmann, A.W. Westerberg, 1997, *Systematic Methods of Chemical Process Design*, Prentice Hall PTR, New Jersey, USA
- BSP Energy Exchange, 2022, Day-ahead Trading Results: Auction Trading, available at: <bsp-southpool.com/day-ahead-trading-results-si.html> (accessed 29.11.2022)
- M. Dokl, R. Gomilšek, L. Čuček, B. Abikoye, Z. Kravanja, 2022, Maximizing the power output and net present value of organic Rankine cycle: Application to aluminium industry, *Energy*, 239, Part E, 122620
- J.M. Egieya, L. Čuček, K. Zirngast, A.J. Iafiade, B. Pahor, Z. Kravanja, 2019, Synthesis of biogas supply networks using various biomass and manure types, *Computers and Chemical Engineering*, 122, 129-151
- L. Čuček, V. Mantelli, J. Y. Yong, P. S. Varbanov, J. J. Klemeš, Z. Kravanja 2015, A procedure for the retrofitting of large-scale heat exchanger networks for fixed and flexible designs applied to existing refinery total site, *Chemical Engineering Transactions*, 45, 109-114
- Energetika Ljubljana, 2022, Price list of public company Energetika Ljubljana d.o.o. for the supply of natural gas for business consumption, available at: <energetika.si/sites/www.jhl.si/files/cenik/datoteke/plin\_-\_cenik\_za\_dobavo\_-\_redni\_1.\_10.\_2022\_poslovnj\_odjem\_0.pdf> (accessed 29.11.2022) (in Slovenian)
- GAMS, 2022, available at: <gams.com/39/docs/> (accessed: 8.12.2022)
- V. Jejčič, J. Verbič, 2019, Utilization of biomass for energy purposes and potentials to 2050, Overview of biomass potentials from agriculture, available at: <podnebnapot2050.si/wp-content/uploads/2019/03/2b\_BIOMASA-IJS-LIFE-CLIMATE-PATH-2050-1-2019.pdf> (accessed: 5.11.2022) (in Slovenian)
- T. Kegl, L. Čuček, A. Kovač Kralj, Z. Kravanja, 2021, Conceptual MINLP approach to the development of a CO<sub>2</sub> supply chain network – Simultaneous consideration of capture and utilization process flowsheets, *Journal of Cleaner Production*, 314, 128008
- Ministry of Agriculture, Forestry and Food of Slovenia (MKGP), 2011, Strategy for utilization of biomass from agriculture and forestry for energy purposes, available at: <arhiv2014.skupnostobcin.si/fileadmin/sos/datoteke/pdf/Barbara/PREDLOGI\_PREDPISOV/Kmetijstvo/Strategija\_biomasa\_priloga.pdf> (accessed: 5.11.2022) (in Slovenian)
- L.J.R. Nunes, T.P. Causer, D. Ciolkosz, 2020, Biomass for energy: A review on supply chain management models, *Renewable and Sustainable Energy Reviews*, 120, 109658.
- N. Yıldırım, M. Toksoy, G. Gökçen, 2006, District heating system design for a university campus, *Energy and buildings*, 38(9), 1111-1119.

# Hybrid modelling and data-driven parameterization of monoclonal antibody cultivation processes: Shifts in cell metabolic behavior

Sara Badr,<sup>a</sup> Kota Oishi,<sup>a</sup> Kozue Okamura,<sup>a</sup> Sei Murakami,<sup>b</sup> Hirokazu Sugiyama<sup>a\*</sup>

<sup>a</sup>*Department of Chemical System Engineering, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, 113-8656, Tokyo, Japan*

<sup>b</sup>*Manufacturing Technology Association of Biologics, 2-6-16, Shinkawa, Chuo-ku, 104-0033, Tokyo, Japan*  
*sugiyama@chemsys.t.u-tokyo.ac.jp*

## Abstract

Process developments to achieve more efficient monoclonal antibody (mAb) production processes include cell line modifications to obtain high-performance alternatives. Cell lines with higher growth rates can exhibit higher sensitivities to variations in carbon sources in the cell culture leading to shifts in metabolic behavior. Current kinetic process models often rely on static model parameters like specific cell consumption and production coefficients. Such models can be inadequate for describing process performance in cases of shifts in cell metabolism. Further understanding of the complex biological phenomena is required to develop more robust process models. In this work, pilot-scale experimental results were obtained for a novel high-performance cell line exhibiting switches in metabolic behavior. State estimation-based methods were used to evaluate the model fit over the course of the run and to identify regions where parameter updates are required. Clustering of the underlying experimental conditions was performed to isolate the sources of variations and correlate them with the different regions of model fit. Alternative empirical formulations are accordingly proposed for the varying model parameters. This approach allows for the development of more interpretable hybrid models by using the data-driven insights to categorize and describe the underlying variations in operating parameters and their influence on cell behavior and process performance.

**Keywords:** Biopharmaceutical production, cell culture, state estimation, hierarchical density-based clustering

## 1. Introduction

The exponential rise in demand for monoclonal antibodies (mAb) over the past decade (Butler & Meneses-Acosta, 2012; Walsh, 2018) has been driving the search for innovative solutions to increase production efficiency. In addition to exploring different operation modes, e.g., continuous cultivation, and optimizing operating conditions, developments were achieved through modifying the cell lines used (Badr & Sugiyama, 2020). The newly established CHO-MK cell line is a high-performance alternative offering rapid cell growth and higher mAb productivity. This cell line is, however, sensitive to carbon source

concentrations, e.g., glucose and lactate, and could exhibit fluctuating behavior in cell metabolism. A shift from lactate production to net consumption has been observed with this cell line. Such shifts have been reported for other cell lines in literature, but the exact mechanism leading to such a switch is still unverified (Hartley et al., 2018). Understanding and manipulating these shifts is important for better process controllability and achieving higher efficiency.

Developing versatile process models to adequately describe the process operation and changes is essential to reach optimal solutions for process design, operation, and control. Kinetic models currently available for describing cell cultivation often rely on static parameters to describe cell-related factors, such as specific productivity or nutrient consumption rates. For stable cell lines, like the standard CHO-K1 cells, these models could show relatively good fit to the experimental and pilot-scale data obtained (Badr et al., 2021). In industrial or experimental applications, cells may temporarily take different metabolic pathways. The reliance on static parameters in the models often leads to discrepancies with industrial applications, especially in the case of sensitive cell lines as the CHO-MK cell line. Discrepancies can be sometimes only partial and not over the course of the entire experimental run. Okamura et al. (2022) have demonstrated the use of a hybrid modelling technique, which inserts a data-driven module to get updated parameter values to better describe shifts to lactate consumption. The input to the data-driven module included the online sensor measurements and the initial predictions of the kinetic model.

In this work, we investigate a data-driven approach to aid the development of more comprehensive and robust mechanistic models for process design and control. This work aimed to isolate the operating conditions correlated to the shifts in metabolic behavior and a change in state. The model fit was analyzed to identify regions where parameter updates are required. Data clustering of the online measurements was used to isolate separate phases of the underlying experimental conditions. The results were combined to suggest new forms for influential model parameters.

## 2. Methods

### 2.1 Experimental data

Pilot-scale experimental data were obtained from the Kobe GMP consolidated lab of the Manufacturing Technology Association of Biologics in Japan. CHO-MK 9E-1 cells were cultivated in a 50 L stirred tank in fed-batch mode for 7 days. Supplementary media feeding was started on day 2. Feeding rates were doubled on day 3 and then adjusted on the following days to maintain a glucose concentration of 3 g/L in the solution. Nutrient and metabolite concentrations were measured daily. Online measurements included dissolved oxygen, pH, temperature, pressure, solution weight, air and oxygen sparging, and agitation rate.

### 2.2 Kinetic model

The model used to describe fed-batch cultivation was based on the one presented by Badr et al. (2021) and adapted by Okamura et al. (2022). The model involves a set of mass balance equations for important system components and Monod expressions for cell growth and death as shown in Eq. (1-7).

$$\frac{d(VX_V)}{dt} = (\mu - \mu_d) VX_V \quad (1)$$

$$\frac{d(VP)}{dt} = Q_P VX_V \quad (2)$$

$$\frac{d(V[GLC])}{dt} = - \left( \frac{\mu}{Y_{X_V/glc}} + m_{glc} \right) VX_V + F_{in} C_{in} \quad (3)$$

$$\frac{d(V[LAC])}{dt} = Q_{lac} VX_V \quad (4)$$

$$\frac{d(V)}{dt} = F_{in} \quad (5)$$

$$\mu = \mu_{max} \left( \frac{[GLC]}{K_{glc} + [GLC]} \right) \left( \frac{KI_{lac}}{KI_{lac} + [LAC]} \right) \quad (6)$$

$$\mu_d = k_d \left( \frac{[LAC]}{KD_{lac} + [LAC]} \right) \left( \frac{KD_{glc}}{KD_{glc} + [GLC]} \right) \quad (7)$$

where,  $X_V$  is the viable cell density and  $V$  is the culture volume.  $F_{in}$  is the supplementary feed flows rate.  $P$ ,  $[GLC]$ , and  $[LAC]$  represent the antibody product, glucose and lactate concentrations, respectively.  $C_{in}$  is the glucose concentration in the supplementary feed streams.  $\mu$  and  $\mu_d$  are cell growth and death rates respectively, while  $\mu_{max}$  and  $k_d$  are their corresponding maximum values.  $Q_P$  is the specific productivity of mAb, while  $Q_{lac}$  is the specific productivity of lactate.  $Y_{X_V/glc}$  and  $m_{glc}$  represent specific glucose consumption for cell growth and maintenance, respectively.  $KI_{lac}$ ,  $K_{glc}$ ,  $KD_{lac}$ ,  $KD_{glc}$  are the Monod model parameters. In the previous works,  $Q_{lac}$  [ $\text{mmol cells}^{-1} \text{h}^{-1}$ ] has been treated as a constant. However, to account for shifts between lactate production and consumption,  $Q_{lac}$  should either be treated as a variable that can vary in magnitude and direction, or an additional term for lactate consumption should be added to Eq. 4.

### 2.3 Data Analysis

State estimation methods have been extensively used in the chemical industry for control applications where process models are well established (Mohd Ali et al., 2015). Observers can be used to estimate hard to measure state variables. A recent study demonstrated the application of the extended Kalman filter (EKF) with a hybrid model for monitoring and control of mammalian cell cultures with limited mechanistic understanding (Narayanan et al., 2020). In this work, we explored the use of a similar concept for the development of improved process models.

Parameter estimation was first conducted given the measured experimental values of system concentrations. Initial guesses for model parameters were provided and a one step forward solution was obtained. A solution was found to minimize the error between the model estimates and experimental values. The model fit was assessed during different regions of the experimental run. This approach allows us to isolate the regions where parameter updates are required. Further analysis of the experimental conditions enables matching the regions with required parameter updates to the underlying varying conditions thus giving a basis for the generation of modified model equations.

Hierarchical density-based clustering (HDBSCAN) was then applied to the online experimental measurements. This method was chosen due to its resilience to noise in the data and its ability to identify clusters of nonuniform densities. It also requires minimal parameter finetuning or assumptions about the number of clusters in the data. The minimum number of points per cluster was set to 50, which is low compared to the size

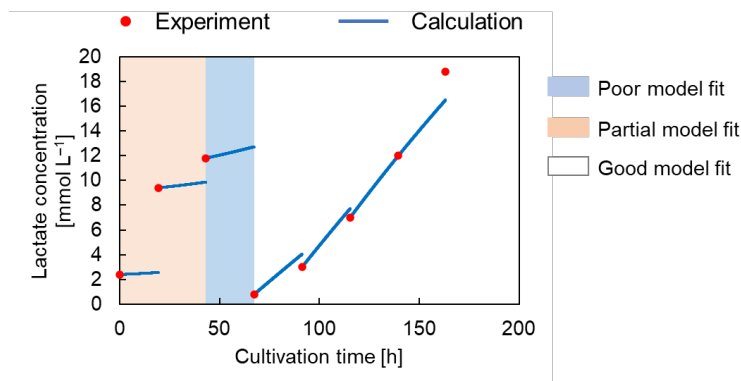


of the online measurements taken every minute, effectively unrestricted the number of clusters formed. The resulting clusters were compared to the different regions of varying model fit and the underlying varying experimental conditions were accordingly isolated. Experimental measurements of lactate concentrations were interpolated using the Akima spline method and were used to regress the values for  $Q_{lac}$ . Alternative formulations for  $Q_{lac}$  were then suggested as a function of the identified varying experimental conditions.

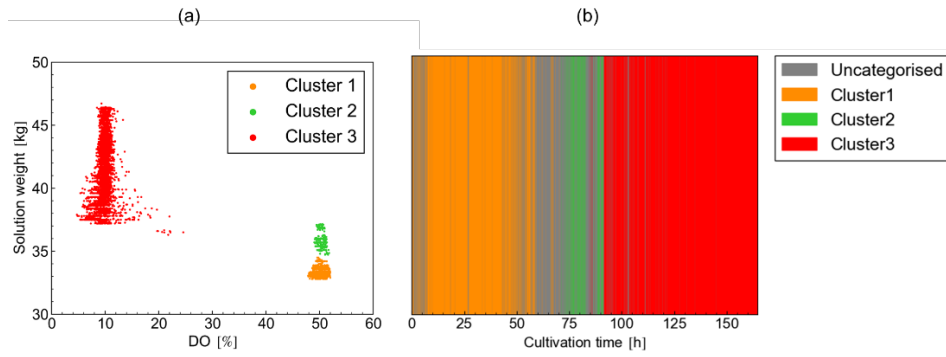
### 3. Results and discussion

**Figure 1** shows an example of the obtained results of the final one step forward simulation using the updated parameters estimated by the implemented algorithm. The figure shows three regions of different model fit to estimate lactate concentrations. Lactate was the component with the highest discrepancies between model estimations and the experimental results relative to all predicted concentrations. There were areas of good model fit observed towards the latter half of the experimental run, areas of partial fit, where model predictions were correct in direction but not in magnitude, and finally areas of poor fit, where the model failed to capture both magnitude and direction of lactate concentration development. Regions of poor fit were those where the cell shifted to net lactate consumption instead of production, this generally coincided with nutrient depletion in the solution.

The results of the clustering step for the operating conditions are shown in **Figure 2**. Three clusters were identified by the algorithm. The strongest separation of the different clusters was observed as a function of the solution weight and the dissolved oxygen (DO) levels as shown in **Figure 2(a)**. Nutrient addition during the fed-batch run leads to changes in the solution weight. The dependency on the solution weight indicates a strong role of nutrient levels in determining the different clusters. In addition, other conditions that showed an impact on cluster formation were the pH and the  $O_2$  sparging levels. **Figure 2(b)** shows a representation of the distribution of the cluster points along the course of the experimental run. The identified clusters were closely related to the different regions of model fit. As a result, it can be concluded that the identified conditions were correlated with the changes in the cell metabolic behavior, resulting in net lactate consumption.



**Figure 1.** Measured lactate concentrations in comparison to the results of the one step forward simulation using the estimated model parameters to assess the model fit.

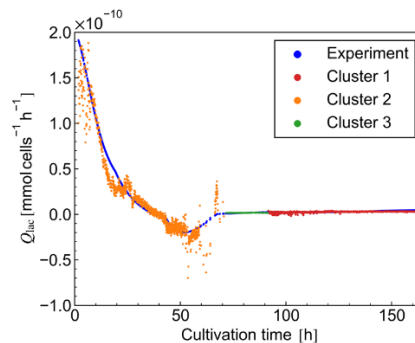


**Figure 2.** Results of the HDBSCAN clustering. (a) solution weight plotted against dissolved oxygen, (b) the distribution of the clusters with respect to experimental run time.

Regression of the  $Q_{lac}$  values from the experimental values and Eq. 4 show that  $Q_{lac}$  is only constant in the latter regions of the experimental run (coinciding with clusters 2 and 3). Whereas in the earlier regions (cluster 1), values of  $Q_{lac}$  decreased until negative values were reached, indicating net lactate consumption. **Figure 3** shows the regressed values of  $Q_{lac}$  along the experimental run. The variations in  $Q_{lac}$  values in the earlier regions (cluster 1) showed a strong correlation with variations in the  $O_2$  sparging levels. Therefore, an empirical formula for calculating  $Q_{lac}$  in the variable region as a function of the  $O_2$  sparging rate [ $L\ min^{-1}$ ] was proposed as follows in Eq. 8. **Figure 3** shows the correspondence between the calculated and the experimentally obtained values of  $Q_{lac}$ , where an  $R^2$  value of 0.93 was obtained.

$$Q_{lac} = \frac{k_0}{k_1 + O_2 \text{ sparging rate}} + k_2 \quad (8)$$

The influence of the  $O_2$  sparging rate on the  $Q_{lac}$  must still be studied in further detail to draw more concrete mechanistic conclusions. Oxygen sparging is assumed to impact dissolved oxygen levels in the system in addition to affecting the general mixing conditions in the reactor. While the sparging rates varied rapidly in the system, DO levels remained fairly constant in the bulk levels during cluster 1 conditions (**Figure 2 (a)**). This could suggest an impact of sparging levels on local variations in nutrient, metabolite, and



**Figure 3.** Regressed values of the specific lactate productivity factor ( $Q_{lac}$ ) from the experimental data and comparison to calculated values for each of the identified clusters.

oxygen concentrations around the cell. Another aspect to be investigated is the impact of changes in pH on  $Q_{lac}$  and on the cell metabolic behavior. Lactate concentrations strongly impact pH levels, while concurrently variations in pH have been reported to influence the switch from lactate production to net consumption (Hartley et al., 2018). Further work also includes determining the role of depletion of different nutrients, e.g., glucose and glutamine, on variation of  $Q_{lac}$ . Further experimental runs are required to confirm the findings and resolve the uncertainties.

#### 4. Conclusions

This work exploits state estimation concepts to develop an understanding of model fit in cases where shifts in cell metabolic behavior result in a need for model parameter updates. Clustering of experimental conditions is used to isolate the underlying variations in experimental conditions leading to a change of cell metabolic behavior and state. This approach allows the development of more interpretable hybrid models, where the data-driven insights are used to describe correlations between experimental conditions and variations in model parameters. The derived insights in this approach can be applied to develop more flexible and representative mechanistic models which can increase the interpretability of the results and help achieve further understanding of the complex biological phenomena.

#### Acknowledgements

This work was supported by the Japan Agency for Medical Research and Development (AMED) [grant No. JP21ae0121015, JP21ae0121016].

#### References

- Badr, S., Okamura, K., Takahashi, N., Ubbenjans, V., Shirahata, H., & Sugiyama, H. (2021). Integrated design of biopharmaceutical manufacturing processes: Operation modes and process configurations for monoclonal antibody production. *Computers & Chemical Engineering*, *153*, 107422.
- Badr, S., & Sugiyama, H. (2020). A PSE perspective for the efficient production of monoclonal antibodies: integration of process, cell, and product design aspects. *Current Opinion in Chemical Engineering*, *27*, 121–128.
- Butler, M., & Meneses-Acosta, A. (2012). Recent advances in technology supporting biopharmaceutical production from mammalian cells. *Applied Microbiology and Biotechnology*, *96*(4), 885–894.
- Hartley, F., Walker, T., Chung, V., & Morten, K. (2018). Mechanisms driving the lactate switch in Chinese hamster ovary cells. In *Biotechnology and Bioengineering* (Vol. 115, Issue 8, pp. 1890–1903). John Wiley and Sons Inc.
- Mohd Ali, J., Ha Hoang, N., Hussain, M. A., & Dochain, D. (2015). Review and classification of recent observers applied in chemical process systems. In *Computers and Chemical Engineering* (Vol. 76, pp. 27–41). Elsevier Ltd.
- Narayanan, H., Behle, L., Luna, M. F., Sokolov, M., Guillén-Gosálbez, G., Morbidelli, M., & Butté, A. (2020). Hybrid-EKF: Hybrid model coupled with extended Kalman filter for real-time monitoring and control of mammalian cell culture. *Biotechnology and Bioengineering*, *117*(9), 2703–2714.
- Okamura, K., Badr, S., Murakami, S., & Sugiyama, H. (2022). Hybrid Modeling of CHO Cell Cultivation in Monoclonal Antibody Production with an Impurity Generation Module. *Industrial and Engineering Chemistry Research*, *61*(40), 14898–14909.
- Walsh, G. (2018). Biopharmaceutical benchmarks 2018. *Nature Biotechnology*, *36*(12), 1136–1145.

# Optimal Design of Anion-pillared Metal-organic Frameworks for Gas Separation

Xiang Zhang,<sup>a,\*</sup> Kai Sundmacher<sup>a,b</sup>

<sup>a</sup>*Department for Process Systems Engineering, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, D-39106 Magdeburg, Germany*

<sup>b</sup>*Chair of Process Systems Engineering, Otto-von-Guericke University Magdeburg, Universitätsplatz 2, D-39106 Magdeburg, Germany*

*\*E-mail address of corresponding author: [zhangx@mpi-magdeburg.mpg.de](mailto:zhangx@mpi-magdeburg.mpg.de)*

## Abstract

Pressure swing adsorption (PSA) has been widely studied for various gas separation tasks. The efficiency of PSA processes is influenced by the properties of adsorbents and the process operating conditions. As innovative adsorbents, metal-organic frameworks (MOFs) have been designed and synthesized. As special family of MOFs, anion-pillared (AP) MOFs where anion groups (e.g.,  $\text{SiF}_6^{2-}$ ,  $\text{TiF}_6^{2-}$ ,  $\text{AlF}_5^{2-}$ , etc.) are used as linkers have been proposed for the separation of light hydrocarbon. The pore size and framework topology (e.g., interpenetrated and non-interpenetrated) of AP MOFs have a dramatic impact on the adsorptive behavior. Considering the diversity of MOF chemistry and only a limited number of AP MOFs reported, a large design space is worth to be explored. In this work, a computational approach is developed to design new hypothetical AP MOFs. The objective is to find high-performance AP MOFs that can be used in PSA processes for propane/propene separation.

**Keywords:** metal-organic frameworks, anion-pillared MOF, computational material design, organic linkers, C3 separation

## 1. Introduction

Pressure swing adsorption (PSA) has been widely applied in the chemical and energy industries for gas separation. It exploits the different affinity of gaseous components to solid adsorbents at different pressures to achieve gas separation. Hence, the adsorbent properties strongly influence the separation efficiency. As well known, multiple types of porous materials can be used as adsorbents such as metal-organic frameworks (MOFs), zeolites, or porous polymers. Among them, MOFs are formed via the self-assembly of different building blocks (i.e., metal nodes and organic linkers) in different topologies. Due to the large variety of available building blocks, MOFs have a nearly infinite design space (Yao et al. 2021). In addition, MOFs have many superior properties such as large porosity, tunable pore geometry, functional pore surface, and importantly modular-ordered. Thus, it is desirable to develop high-performance MOFs that can enhance process efficiency in adsorption-based gas separation (Zhang et al., 2020).

When MOFs are used as adsorbents, four alternative separation mechanisms exist: equilibrium separation, kinetic separation, molecular sieving, and gate-opening separation. Among these, equilibrium separation is the most common method, which enables separation based on the difference of gas equilibrium loadings. So far, an increasing number of MOFs with diverse isotherm characteristics have been synthesized

by varying MOF chemistry and structure (e.g., pore geometry and topology). Given certain separation tasks, it has been found that many MOFs fail to achieve the pre-specified separation requirements, not to mention good process performance (Burns et al. 2020). To find promising MOF adsorbents efficiently, various computational screening strategies have been developed. In general, a set of MOF candidates are collected. Their adsorption and diffusion properties are computed via molecular simulations. Then, simple performance metrics (e.g., working capacity and selectivity) are calculated to rank top candidates. Although this approach can be used to screen a large pool of MOFs, using simple performance metrics cannot tell whether specific separation requirements are satisfied, particularly when tough separation specifications (e.g., purity  $\geq 99.95\%$  in  $H_2$  purification) exist. Thus, a computational methodology is developed in this work to design and screen MOFs that can best serve the process for achieving pre-specified gas separations.

Anion-pillared (AP) MOFs where anion groups (e.g.,  $SiF_6^{2-}$ ,  $TiF_6^{2-}$ ,  $AlF_5^{2-}$ , etc.) are used as linkers have been highlighted for light hydrocarbon separations. The pore sizes of AP MOFs can be fine-tuned by rationally designing their chemistry. It has been found that small changes of pore sizes can have a dramatic impact on the adsorptive behavior of AP MOFs. Moreover, AP MOFs generally have two types of cuboid networks: non-interpenetrated and interpenetrated. The non-interpenetrated network is the standard primitive cubic (i.e., *pcu*) topology, while the interpenetrated one is composed of doubly interpenetrated nets that are isostructural to the non-interpenetrated network (Lin et al. 2011). Evidently, the non-interpenetrated network can form a small pore size with a large organic linker. So far, only a limited number of AP MOFs have been reported and a large design space has not been explored yet considering the diversity of MOF chemistry. This inspired us to design new AP MOFs for efficient gas separation.

As an essential building block for polypropylene production, propene (PE) is a key olefin raw material. It is widely known that polypropylene synthesis requires high-purity propylene (over 99.5%). Thus, the removal of other hydrocarbons, especially propane (PA), from propene is one of the critical gas separation tasks in the modern chemical industry. Currently, this is achieved by the energy-intensive cryogenic distillation. As energy-efficient alternatives, adsorption-based separation methods are of great interest. In this case, the development of advanced adsorbents that can achieve low energy consumption is crucial. Thus, the objective of this work is to use the proposed methodology to design new hypothetical AP MOFs for efficient PA/PE separation.

## 2. Workflow of Anion-pillared MOF Design

Figure 1 shows the explicit workflow for computational AP MOF design. First, the reported AP MOFs are collected and decomposed to identify the feasible alternatives of metal nodes and anion pillars. Meanwhile, all the MOFs in the 2019 CoRE MOF database are deconstructed. Referring to the common characteristics found in AP MOFs, feasible organic linkers are selectively extracted. Then, based on the topology templates and the collected building blocks, we constructed hypothetical AP MOFs using ToBaCCo that are modified to generate MOFs with two types of organic linkers (i.e., anion pillars and the general organic linker). Afterwards, the single-component adsorption isotherms of hypothetical AP MOFs are predicted using grand canonical Monte Carlo (GCMC) simulations performed in RASPA. To improve the reliability, new Lennard-Jones (LJ) parameters that are suitable for AP MOFs are calibrated based on experimental isotherm data. The obtained single-component isotherms are used to fit the multi-component

isotherm models that can be directly applied in a short-cut PSA model. This model can efficiently evaluate whether the new AP MOFs are feasible to satisfy the pre-specified separation requirements. After this feasibility check, rigorous PSA optimization are performed on the feasible AP MOFs to obtain the optimal adsorbent together with optimal PSA process operating conditions.

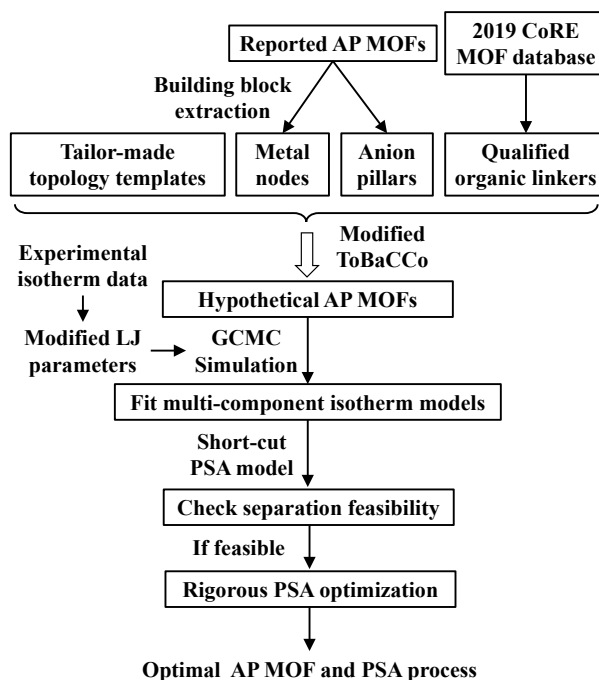


Figure 1. Workflow of computational AP MOF design.

### 3. Build Blocks Construction

AP MOFs consists of metal nodes and two types of organic linkers (i.e., anion pillars and general organic linkers serving as donor and acceptor, respectively) in certain topologies. For creating sound feasible AP MOFs, proper building blocks and topology templates have to be specified.

#### 3.1. Two topology templates

So far, AP MOFs with limited types of topologies (e.g., *pcu*, *pcu-i*, *mno*, *pto*, etc.) have been successfully synthesized. Since the non-interpenetrated *pcu* and interpenetrated *pcu-i* are the two most common topologies, only these two types are considered here. The *pcu* template can be found in the original ToBaCCo platform, while the *pcu-i* template is made using Material Studio.

#### 3.2. Sixty-four feasible organic linkers

The general organic linkers that can be used in AP MOFs are extracted from the CoRE MOFs. After removing the CoRE MOFs with structural defects, 9268 CoRE MOFs are left and 14107 organic linkers can be extracted. Then, referring to the most reported AP MOFs, it is found that the general organic linkers should be symmetric and only consist of carbon, nitrogen, and oxygen atoms (Li et al. 2022). Considering these common characteristics, 64 unique symmetric organic linkers are finally obtained.

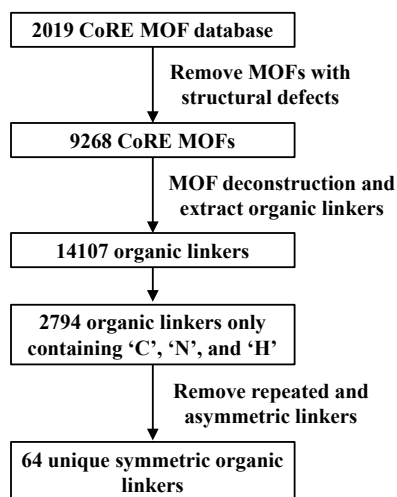


Figure 2. Procedures of extracting general organic linkers used in AP MOFs.

### 3.3. Five metal nodes and eight anion pillars

After decomposing the reported AP MOFs with *pcu* and *pcu-i* topologies, it is found that the metal nodes are simply single-metal atoms and five metal atoms are commonly involved (i.e., Zn, Fe, Cu, Co, Ni). In addition, 8 kinds of anion pillars are found including  $\text{SiF}_6^{2-}$ ,  $\text{TiF}_6^{2-}$ ,  $\text{SnF}_6^{2-}$ ,  $\text{GeF}_6^{2-}$ ,  $\text{ZrF}_6^{2-}$ ,  $\text{FeF}_5^{2-}$ ,  $\text{AlF}_5^{2-}$ , and  $\text{NbOF}_5^{2-}$ . With these, the collected AP MOFs in the form of cif files are imported into Material Studio to extract the corresponding metal nodes and anion pillars.

### 3.4. Bond lengths between building blocks

For MOF construction, building blocks are connected via single-bonds. Given the collected building blocks above, metal-fluorine bonds are formed when connecting metal nodes and anion pillars. Metal-nitrogen bonds are formed when connecting metal nodes and general organic linkers. Since each atom-atom pair has different bond length, the bond lengths between metal-nitrogen and metal-fluorine should be specified explicitly. Based on the reported AP MOFs, Table 1 lists the average bond lengths between metals and two connection atoms. These data can be directly used to adjust the sizes of the five metal nodes.

Table 1. Bond lengths between metals and connection atoms in AP MOFs.

Metal	Metal-fluorine distance (Å)	Metal-nitrogen distance (Å)
Co	2.04	2.15
Cu	2.25	2.04
Fe	1.99	2.19
Ni	1.97	2.08
Zn	2.09	2.13

## 4. Design of anion-pillared MOFs for PA/PE separation

### 4.1. ToBaCCo-based AP MOF generation and isotherm prediction

The prepared building blocks and topology templates are imported into the modified ToBaCCo to generate hypothetical AP MOFs. In this process, a series of MOF construction feasibility rules will be verified and the cell sizes will be optimized. After this, over 3000 hypothetical AP MOFs in the form of crystallographic information framework (.cif) files are obtained.

The cif files are directly used to perform GCMC simulations using the open-source software RASPA. The reliability of GCMC simulations relies on the force field parameters (i.e., LJ parameters). It is well-known that generic force fields such as DREIDING force field and Universal Force Field (UFF) are widely used in MOF screening studies. However, for AP MOFs, the PA and PE isotherms cannot be predicted accurately using those general force fields. Thus, a new set of LJ parameters should be specified based on experimental isotherm data. After careful literature review, the single-component PA and PE experimental isotherm data at 298 K of three AP MOFs (i.e., GEFIX-2-Cu-i, SIFSIX-2-Cu-i, and NbOFFIVE-i-Ni) are collected (Wang et al. 2020; Cadiau et al. 2016). On the basis of UFF, the LJ parameters of three atoms (i.e., Ni, Nb, and Ge) and two groups (i.e., CH<sub>3</sub>- and CH<sub>2</sub>- for PE) should be modified so that the predicted isotherms can match the corresponding experimental data well.

### 4.2. Process-based AP MOF screening

After specifying the new LJ parameters, the single-component PA and PE adsorption loadings at four pressure levels (i.e., 0.1, 0.5, 1, and 2 bars) and 313 K are calculated for all hypothetical AP MOFs. The obtained GCMC results are used to fit the multi-component dual-site Langmuir isotherm models for the purpose of process-based screening and optimization.

Clearly, it brings great computational burdens to perform rigorous PSA optimization for all the AP MOFs. Thus, for fast screening, a short-cut PSA model is proposed to identify whether an AP MOF can meet the pre-specified separation specifications (i.e., 99.5% PE purity and 65% PE recovery). The short-cut model is based on a one-bed, four-step PSA process that consists of pressurization, adsorption, rinsing, and desorption steps. It is assumed that the column is initially saturated with 99.95% PE at a low pressure ( $P_{low}$ ). The column goes through the pressurization step where PA/PE gases at a high pressure ( $P_{high}$ ) are fed into the bottom of the column. The PE mole fraction in the feed gas is set to 85%. The top is closed and the internal pressure gradually increases. When the column's top reaches an intermediate pressure level (i.e.,  $0.85P_{high}$ ), the top is opened and the adsorption step begins. The adsorption is stopped when the stream leaving the top is the same as the feed gas (i.e., breakthrough state). Afterwards, based on the mass balance and PE recovery, a certain amount of 99.5% PE gas is fed into the bottom for rinsing. Finally, in the desorption step, a constant low pressure ( $P_{low}$ ) is enforced at the bottom of the column where purified PE can be collected. If the collected PE stream has a purity above 99.5%, the column state at the end of desorption can presumably be equal to the initial state (i.e., 99.5% PE at the low pressure). This implies that the adsorbent can meet the pre-specified separation requirements. Otherwise, the adsorbent is deemed as infeasible one. Clearly, this short-cut model is a batch PSA process model with two variables (i.e.,  $P_{high}$  and  $P_{low}$ ). Grid search on the two variables can be easily performed for a complete feasibility check. Hence, the proposed short-cut model is applied to screen all the generated AP MOFs. It can be found that a set of potential AP MOFs can survive such as NbOFFIVE-15-Fe-i and SIFSIX-2-Co. By minimizing the energy consumption



and maximizing the productivity, finally a multi-objective rigorous PSA optimization can be performed on those feasible AP MOFs for final evaluation.

## 5. Conclusions

This work presented a computational approach to systematically design and screen hypothetical AP MOFs for efficient PA/PE separation in PSA processes. The workflow starts with extracting various building blocks from the reported AP MOFs and the CoRE MOFs. Together with two topology templates, hypothetical AP MOFs are constructed using the modified ToBaCCo. For each generated AP MOF, GCMC simulations are performed using newly calibrated LJ parameters to predict the single-component adsorption loadings of PA and PE. With the predicted loadings, multi-component DSL isotherms are fitted and used in a short-cut PSA model for fast screening. Finally, for the survived candidates, rigorous multi-objective PSA optimization are carried out in order to find the real optimal AP MOFs and the corresponding process operating conditions. Clearly, as demonstrated via the PA/PE separation example, hypothetical AP MOFs leading to better process performance can be computationally synthesized and screened. The major novelty of the present work is the use of modified computational tools and models to design optimal AP MOFs efficiently. Since AP MOFs with specific topologies have similar synthesis routes, this provides great conveniences for the subsequent experimental validation.

## References

- TD. Burns, KN. Pai, SG. Subraveti, SP. Collins, M. Krykunov, A. Rajendran, TK. Woo, 2020, Prediction of MOF performance in vacuum swing adsorption systems for postcombustion CO<sub>2</sub> capture based on integrated molecular simulations, process optimizations, and machine learning models, *Environmental Science & Technology*, 54, 4536-4544
- A. Cadiou, K. Adil, P.M. Bhatt, Y. Belmabkhout, M. Eddaoudi, 2016, A metal-organic framework-based splitter for separating propylene from propane, *Science*, 353, 6295
- YJ. Colón, DA. Gómez-Gualdrón, RQ. Snurr, 2017, Topologically guided, automated construction of metal-organic frameworks and their evaluation for energy-related applications. *Crystal Growth & Design*, 17, 11, 5801-5810
- X. Li, H. Bian, W. Huang, B. Yan, X. Wang, B. Zhu, 2022, A review on anion-pillared metal-organic frameworks (APMOFs) and their composites with the balance of adsorption capacity and separation selectivity for efficient gas separation, *Coordination Chemistry Reviews*, 470, 214714
- MJ. Lin, A. Jouaiti, N. Kyritsakas, MW. Hosseini, 2011, Molecular tectonics: control of interpenetration in cuboid 3-D coordination networks, *CrystEngComm*, 13, 776-778
- X. Wang, P. Zhang, Z. Zhang, Li. Yang, Q. Ding, X. Cui, J. Wang, H. Bing, 2020, Efficient separation of propene and propane using anion-pillared metal-organic frameworks, *Industrial & Engineering Chemistry Research*, 59, 3531-3537
- Z. Yao, B. Sánchez-Lengeling, NS. Bobbitt, BJ. Bucior, SGH. Kumar, SP. Collins, T. Burns, TK. Woo, OK. Farha, RQ. Snurr, A. Aspuru-Guzik, 2021, Inverse design of nanoporous crystalline reticular materials with deep generative models, *Nature Machine Intelligence*, 3, 1, 76-86
- X. Zhang, T. Zhou, K. Sundmacher, 2021, Integrated MOF and P/VSA process design: Descriptor optimization, *AIChE Journal*, 68, e17524

# Probabilistic neural networks for mechanical properties and tensile behavior prediction of polymer composites considering uncertainty in tensile testing

Jinkyung Son,<sup>a</sup> Tim B. Jaewook Lee,<sup>a</sup> Dongil Shin,<sup>a,b</sup>

<sup>a</sup>*Department of Chemical Engineering, Myongji University, Yongin, Gyeonggi-do, 17058, Korea*

<sup>b</sup>*Department of Disaster and Safety, Myongji University, Yongin, Gyeonggi-do, 17058, Korea*

## Abstract

Polymer Matrix Composite (PMC) is gaining attention for its excellent modeling processability, light weight, low unit cost, and various mechanical properties by reinforcing the heat resistance and strength of plastics with fillers. However, to this day, PMC development relies on designers' intuition, and the cost of repeated testing is quite high. Therefore, in order to enhance PMC's development competitiveness, an AI model that predicts mechanical properties according to the combination and composition of constituent materials is essential. Since these mechanical properties can be grasped using the Stress-Strain curve (S-S curve) obtained from tests such as tensile, creep, and bend, the S-S curve is a kind of mechanical property evaluation index that can accurately know the mechanical properties.

However, Finite Element Method (FEM)-based simulators or Machine Learning (ML)-based predictive models that calculate S-S curve are deterministic models that provide only one S-S curve and do not have a predictive model that considers the uncertainty (variance) of PMC tensile data. Therefore, in this study, after collecting S-S curve data of PMC, Using the Probabilistic Neural Network (PNN) model, we develop a model that predicts the confidence interval along with the mean value of the tensile S-S curve of the PMC. The performance of the developed model is evaluated on the  $R^2$  basis, and it is confirmed that not only the elastic section, which is relatively easy to predict, but also the S-S curve of the plastic section can be predicted. Based on the results of future research, it is expected that this model can be used to build an AI-based PMC reverse engineering system.

**Keywords:** Polymer Matrix Composite (PMC), Mechanical behavior, Gaussian Process Regression (GPR), Probabilistic Neural Network (PNN), Feed-Forward Neural Network (FNN)

## 1. Introduction

The mechanical behavior of the material can be expressed in the S-S curve and not only the mechanical properties such as tensile strength, modulus of elasticity and maximum load are evaluated, but also the deformation from the elastic section to the plastic section. However, polymeric materials have a high uncertainty in the S-S curve due to the inherent

nonlinearity properties of the material, which shows a wide distribution. Therefore, in general, when designing polymers, the average value of the data obtained after conducting at least 10 experiments after uniform formulation and manufacturing conditions is used as the S-S curve of the material to consider 'uncertainty'.

However, if the average value of the repeated experiment is used, there is a possibility of over-estimating or under-estimating the physical properties of the composite material because it ignores the variance of the data. Prior studies that directly predicted the S-S curve, which is the result of such a PMC tensile test, [G. Chen et al. \(2021\)](#). In previous studies, the microstructure of composite materials is simulated and analyzed through Convolutional Neural Networks (CNN) models. However, the results of this study produce only one predicted S-S curve and are difficult to apply to composite materials such as PMC with a wide distribution. Therefore, as the first step in the AI-based material reverse engineering system, this study proposes a model that predicts the confidence interval of the S-S curve by reflecting the uncertainty of the PMC tensile S-S curve. First, Domain Knowledge is used to obtain information such as density, porosity, and Poisson's ratio of constituent materials to build a feature set. Next, feature engineering is performed to input a sufficient amount of information to predict the S-S curve of the PMC by the ML-based model and at the same time prevent falling into the dimensional curse. As a feature engineering technique, Feature elimination is used to secure an optimal feature set by selecting only useful features that are most involved in the performance of the model from the feature sets obtained. Using the obtained optimal feature set and tensile test data of PMC, model training is carried out, and a model is designed to predict the S-S curve of PMC using Gaussian Process Regression (GPR), a representative ML model that provides variance of regression results.

## 2. PMC Tensile Test Data

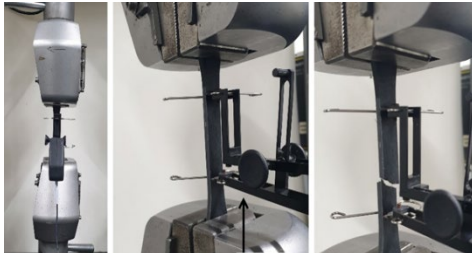


Figure. 1. Universal Testing Machine (UTM)

Matrix	Polypropylene (PP), Polycarbonate (PC), Polyamide 6 (PA6), Polyamide 6,6 (PA6,6)
Filler	Al <sub>2</sub> O <sub>3</sub> , Al <sub>2</sub> O <sub>5</sub> Si, Boron Nitride (BN), Si <sub>3</sub> N <sub>4</sub>
Matrix wt%	90 ~ 40

Figure. 2. Combination and composition of subject to tensile test

### 2.1. Data collection

A training database of an integrated model for predicting mechanical properties was established by integrating literature material information data, material information bank data, and data factory collection data. As for the tensile test data, the test was conducted with a universal testing machine as shown in Fig. 1. The material to be subjected to the tensile test was manufactured by selecting a PMC, which is widely used as a lightweight material. As shown in Fig. 2, the matrix was set to have Polypropylene (PP), Polycarbonate (PC), Polyamide 6 (PA6), Polyamide 6,6 (PA6,6) and filler were Al<sub>2</sub>O<sub>3</sub>, Al<sub>2</sub>O<sub>5</sub>Si, Boron Nitride (BN), and Si<sub>3</sub>N<sub>4</sub>, and the composition ratio of the polymer and filler was set to between 90% and 40%. According to the combination and composition set in this way, a total of 85 PMC specimens were produced and a tensile test was conducted.

As mentioned earlier, the uncertainty of the tensile test data of the PMC is high due to the high randomness of the polymer itself.

As can be seen from Fig. 3, the high dispersion of the PMC tensile test data shows different S-S curves even when manufactured and tested with the same recipe. Fig. 3 shows the tensile test data of the PMC manufactured with 40wt% PA6,6 and 60wt% Al<sub>2</sub>O<sub>3</sub>, 90wt% PC and 10wt% Al<sub>2</sub>O<sub>3</sub>. As can be seen from the figures, the tensile test data of each PMC has a high uncertainty of 7% to 24%.

Generally, when designing a PMC, the average value of physical properties obtained after performing five or more repeated experiments is determined and used as the physical properties of the polymer composite water. However, in this study, since the model design considering the uncertainty of the data is the goal, the collected S-S curve raw data was used as it is, not the S-S curve calculated as an average.

## *2.2. 2<sup>nd</sup> Or Heading*

### *2.2.1. 3<sup>rd</sup> Order Heading*

main text

main text

## **References**

Z. Allen, Year, Article or Chapter Title, Journal or Book Title, Volume, Issue, Pages

Y. Brown, Year, Article or Chapter Title, etc.

X. Cullem, Year, etc.



# Data-Driven Indication of Flooding in an Industrial Debutanizer Column

Martin Mojto,<sup>a</sup> Karol Lubušký,<sup>b</sup> Miroslav Fikar,<sup>a</sup> Radoslav Paulen,<sup>a</sup>

<sup>a</sup>*Slovak University of Technology in Bratislava, Bratislava, Slovakia*

<sup>b</sup>*Slovnaft, a.s., Bratislava, Slovakia*

*martin.mojto@stuba.sk*

## Abstract

The profitability and sustainability of process industries are affected by the performance of each unit involved. A key measure of a unit's performance is based on whether it operates in a desired production window or whether it trips into an abnormal condition. In this contribution, we study flooding of industrial distillation columns. We aim to improve the performance of an industrial debutanizer column by designing a data-driven flooding indicator. The design of the indicator consists of three steps; (a) the data treatment, (b) a priori labeling, and (c) indicator design. The prior knowledge about flooding within the column is used to design a reference indicator. This knowledge is either unused or fully exploited during the design of the indicator. We compare various design methods and show the potential of data-driven approaches for flooding indication.

**Keywords:** Debutanizer Column, Flooding Indicator, Soft Sensor, Subset Selection

## 1. Introduction

Flooding is an undesired phenomenon in industrial distillation columns. It occurs when the liquid rises above a tray because of foaming or excessive downcomer fill-up (King, 2016). This state causes a high loss in tray separation efficiency and, hence, plant profitability. Early flooding detection is thus crucial for a profitable and sustainable plant.

Several works dealt with the problem of flooding detection. They considered correlation of the flooding effect with internal process variables, especially the pressure difference (drop) across the column (Peiravan et al., 2020) and the time derivative of the pressure drop (Pihlaja, 2008). Industrial experts use these results and combine them often with an insight into the principal triggering cause of flooding, creating a tailored solution for each column. This requires some effort that could be saved using machine learning (ML).

Several ML approaches (Mojto et al., 2021; Oeing et al., 2021; Fuentes-Cortés et al., 2022) were employed to aid decision making in industrial columns. Unsupervised ML approaches, such as *k*-means clustering (Forgy, 1965) or principal component analysis (PCA) (Pearson, 1901), consider no knowledge about the model outcome. The supervised ML techniques, e.g., subset (feature) selection (SS) (Smith, 2018) or support vector machine (SVM) (Boser et al., 1992), use knowledge on the desired outcome for training.

This paper investigates the design of data-driven flooding indicators for an industrial debutanizer. Performance of indicators designed via data-driven approaches (unsuper-

vised and supervised ML) is assessed against the reference indicator (considered as ground truth), which is designed according to the industrial experience about flooding.

## 2. Problem Statement

Flooding indication is essentially a binary classification problem. We aim to design an indicator  $\mathbf{I}$  that assigns a categorical label  $\hat{y}$  via the classification model  $f(x)$  as:

$$\hat{y} = \begin{cases} +1 \text{ (flooding)}, & \text{if } f(x) \geq 0, \\ -1 \text{ (normal operation)}, & \text{if } f(x) < 0, \end{cases} \quad (1)$$

where  $x \in \mathbb{R}^{n_p}$  represents a subset (sparse representation,  $n > n_p$ ) of all online plant measurements  $\xi \in \mathbb{R}^n$  at one time instant. We consider a linear classifier in the form:

$$f(x) = w^\top x + w_0, \quad (2)$$

where  $w \in \mathbb{R}^{n_p}$  represents a classifier normal vector and  $w_0$  is a classifier off-set.

### 2.1. Industrial Debutanizer Column

We study a debutanizer (distillation) column that is a part of the FCC unit of the refinery Slovnaft, a.s. in Bratislava, Slovakia. The column (40 trays) separates the C4/C5 fraction into the C4-fraction-rich distillate product and the C5-fraction-rich bottom product.

The available dataset involves the measurements from January 2019 to April 2021. In this study, we consider 30-minute moving average values of the available minutely data. Overall, the dataset involves 34,297 measurements. The measurements from two plant shutdowns (May – July 2019 and December 2020) are excluded.

The following 41 input variables are directly measured by online sensors at the column:

$$\xi = (vo_B, vo_{D,1-3}, vo_R, vo_{reb,h}, T_{col,1-5}, T_{B,1-2}, T_{D,1-3}, T_F, T_{reb,h,1-2}, Q_{con}, p_{col}, p_{D,1-4}, p_{df,col}, p_{con}, F_R, F_B, F_{D,1-4}, F_F, F_{reb,h,1-2}, L_{reb,1-2}, L_{con,1-3}), \quad (3)$$

where  $vo$ ,  $T$ ,  $Q$ ,  $p$ , and  $F$  are a valve opening, temperature, heat input, pressure, and flow rate, respectively. Indices B, D, F, R, col, con, reb, and df represent a bottom section, distillate section, feed section, reflux section, column section, condenser section, reboiler section, and cross-column difference, respectively. Note that exact location of sensors cannot be disclosed due to the confidentiality reasons. The input vector is extended with ratios ( $F_R/F_F$ ,  $F_B/F_F$ ,  $Q_{con}/F_F$ ) and pressure compensated temperatures ( $PCT_B$ ,  $PCT_D$ ).

The studied debutanizer column usually operates within the desired operating regime. At times, however, the operating conditions within the unit induce flooding. The envisioned low-cost solution to the flooding problem is to design a reliable indicator. The key aspect of this approach is that the indicator is not only used for monitoring the plant, but it is connected directly to the advanced process controller that can provide a fast response.

The dataset does not contain any direct indication of flooding that could be used to label the data. However, it is possible to attribute flooding occurrence to the increased values of  $p_{df,col}$ ,  $F_R$ ,  $F_{reb,h,2}$ , and  $T_{col,4}$  and decreased values of  $T_{reb,h,1}$ . We use this knowledge to design the reference indicator to provide the ground truth of the flooding for our study.

## 3. Methodology

Data-driven indicators are designed using unsupervised ( $\mathbf{I}^{Uns}$ ) and supervised ( $\mathbf{I}^{Sup}$ ) ML approaches. For training ( $\mathbf{I}^{Sup}$ -type indicator) and testing, the ground truth is provided by the aforementioned reference indicator resulting from industrial expert knowledge.

### 3.1. Indicator Design

The design procedure of the data-driven indicator consists of three sequential steps:

1. Data processing (data filtering/treatment, distribution to training/testing dataset).
2. A priori labeling of the training dataset (only applied for  $\mathbf{I}^{\text{Uns}}$ -type indicators).
3. Training of a classifier (fit the  $f(x)$  parameters on the labeled training dataset).

After the standardization of the data set (zero mean and unit variance), the aim of the data treatment (the 1<sup>st</sup> step) is to reduce the number of outliers. Due to the non-ideal (yet close normal) noise distribution within the industrial dataset, minimum covariance determinant (MCD) (Hubert and Debruyne, 2010) is applied. The outlier detection is performed using the  $F$ -distribution, retaining data with 99.9999% probability. The high probability follows from the need to eliminate only the most deviated measurements while maintaining the data representing the flooding, which can be otherwise seen as outliers.

It is optional to smoothen the dataset by filtering out the high-frequency noise that does not represent slower effects of flooding. Subsequently, as flooding is characterized by the changes of the variables, we extend the dataset (here, 46 variables) by time differences:

$$\Delta\xi_i(k) = \xi_i(k) - \xi_i(k-1), \quad \forall_i = \{1, 2, \dots, n\}, \quad (4)$$

where  $k$  is a time instant. The resulting dataset considers both, the original dataset and time differences, i.e., 92 variables. Effectively, we assign  $\xi \leftarrow (\xi, \Delta\xi)$  in this step.

The 2<sup>nd</sup> step, applied to label the data for  $\mathbf{I}^{\text{Uns}}$ -type approaches, is performed by  $k$ -means clustering (Forgy, 1965) with the elbow method to determine the optimal number of clusters. The clusters with a low cardinality but large distance between the cluster center and the dataset mean are considered to represent the debutanizer flooding.

The training phase needs to choose an appropriate indicator input space ( $\mathbb{R}^{n_p}$ ) among all the process variables and their time differences. The methods used in this study are:

1. Industrial patent by (Pihlaja, 2008), which uses  $\Delta p_{\text{df,col}}$  only (referred to as  $\mathbf{I}_{\text{pat}}$ ).
2. Industrial experience (specific to the studied debutanizer) using  $\Delta p_{\text{df,col}}$ ,  $\Delta F_R$ ,  $\Delta F_{\text{reb,h,2}}$ ,  $\Delta T_{\text{col,4}}$ , and  $\Delta T_{\text{reb,h,1}}$  (referred to as  $\mathbf{I}_{\text{ref}}$ ).
3. PCA approach (Pearson, 1901) that chooses a number principal components that explain at least 95% of variance in the dataset (referred to as  $\mathbf{I}_{\text{PCA}}$ ).
4. SS approach (Smith, 2018), which determines the best subset of inputs via cross-validation and comparison of structures with  $n_p = \{1, 2, \dots, 5\}$  (referred to as  $\mathbf{I}_{\text{SS}}$ ).

In the 3<sup>rd</sup> step, the linear classifier (see Eq. (2)) is designed according to the chosen inputs ( $x \in \mathbb{R}^{n_p}$ ). To this end, we use support vector machines (SVM) (Boser et al., 1992).

### 3.2. Performance Assessment

The outcome of an indicator can fall into four categories: true positive (TP) and false positive (FP), when flooding is indicated correctly and incorrectly, respectively, and, vice versa, true negative (TN) and false negative (FN), for indicating of normal operation. We use some well-known normalized performance criteria for the designed indicators:

$$AC = \frac{TP + TN}{TP + FP + TN + FN}, PR = \frac{TP}{TP + FP}, RC = \frac{TP}{TP + FN}, F1 = \frac{2 \times PR \times RC}{PR + RC}, \quad (5)$$

here AC (accuracy) is a measure of how often the classifier makes the correct prediction, PR (precision or correctness) is a measure of how precisely the true prediction is achieved, RC (recall or sensitivity) is a measure of how actual observations are predicted correctly. F1-score (F1) is a harmonic mean between PR and RC. In the industry, it is important to warn about the potential of flooding and thus low value of FN (high RC) is preferred.



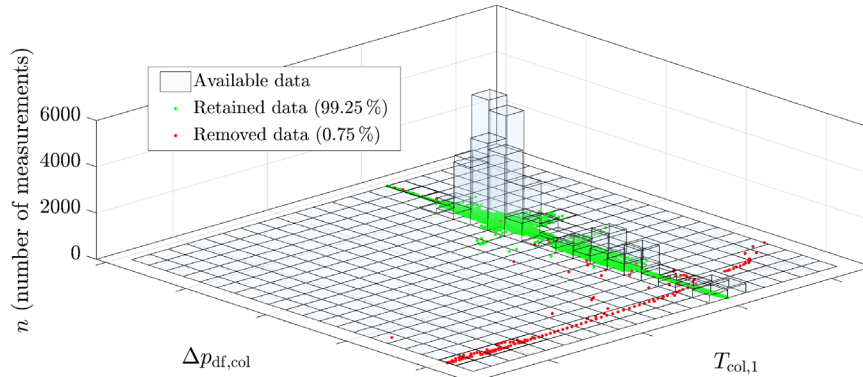


Figure 1: Histogram of two variables from the debutanizer dataset treated by the MCD method.

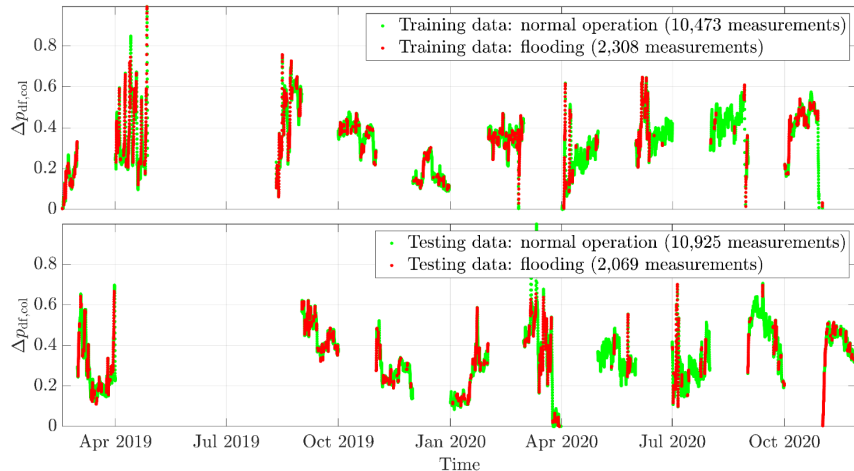


Figure 2: Visualization of training and testing datasets and ground truth labels.

## 4. Results

The results of data treatment using the MCD method are shown in Fig. 1. The data values are anonymized for confidentiality reasons. Only the most deviated measurements (0.75%) are considered as outliers, and the rest of the measurements (99.25%) is retained. The dataset is further smoothed by filtering using a 10th-order low-pass Butterworth filter with a cut-off frequency of 0.028 mHz (with zero-phase distortion).

For fairness of assessment, we distribute the retained data chronologically on an alternating monthly basis into the training and testing datasets (see Fig. 2). From the entire dataset (25,775 points), 12,781 and 12,994 points are assigned to the training and testing dataset, respectively. Fig. 2 illustrates the training-testing data division together with (ground truth) labels assigned based on industrial experience with the reference indicator.

### 4.1. Training of Data-Driven Indicators

Design of the data-driven flooding indicators for the debutanizer column is conducted via MATLAB based on the methods from Section 3.1. MATLAB built-in routines for  $k$ -means clustering, PCA, and SVM are exploited. We design indicators based on unsupervised ML ( $\mathbf{I}_{pat}^{Uns}$ ,  $\mathbf{I}_{ref}^{Uns}$ ,  $\mathbf{I}_{PCA}^{Uns}$ ,  $\mathbf{I}_{SS}^{Uns}$ ) and supervised ML ( $\mathbf{I}_{pat}^{Sup}$ ,  $\mathbf{I}_{ref}^{Sup}$ ,  $\mathbf{I}_{PCA}^{Sup}$ ,  $\mathbf{I}_{SS}^{Sup}$ ). A main difference between the approaches is the use of  $k$ -means (used for  $\mathbf{I}^{Uns}$  indicators).

Table 1: The comparison of the true positives (TP), false positives (FP), true negatives (TN), false negatives (FN), accuracy (AC), precision (PR), recall (RC), and complexity (no. of input variables  $n_p$ , no. of principal components  $n_{pc}$ ) of the designed data-driven indicators on the testing dataset.

ML method	Unsupervised learning			Supervised learning				
	Structure	$\mathbf{I}_{pat}^{Uns}$	$\mathbf{I}_{ref}^{Uns}$	$\mathbf{I}_{PCA}^{Uns}$	$\mathbf{I}_{pat}^{Sup}$	$\mathbf{I}_{ref}^{Sup}$	$\mathbf{I}_{PCA}^{Sup}$	$\mathbf{I}_{SS}^{Sup}$
TP		1,784	1,704	618	1,192	2,031	1,720	2,029
FP		3,828	2,823	3,147	1,358	4	168	0
TN		7,097	8,102	7,778	9,567	10,921	10,757	10,925
FN		285	365	1,451	877	38	349	40
AC		68.3	75.5	64.6	82.8	99.7	96	99.7
PR		31.8	37.6	16.4	46.7	99.8	91.1	100
RC		86.2	82.4	29.9	57.6	98.2	83.1	98.1
F1		46.5	51.7	21.2	51.6	99	86.9	99
$n_p/n_{pc}$		1	5	17	1	5	17	2

A key to success of unsupervised ML is an appropriate data labeling. The results indicate that, unsurprisingly, the best results are obtained when the  $k$ -means clustering is performed on a dataset with reduced dimensionality (e.g., one variable for  $\mathbf{I}_{pat}^{Uns}$  indicator or seventeen principal components determined for  $\mathbf{I}_{PCA}^{Uns}$ ), with appropriate input structure. The clustering method reveals 4–5 clusters out of which 1–2 clusters are selected to represent flooding. This result suggest that merging of steps 1 and 2 mentioned in Section 3.1 is a sensible approach to successful indicator design. We can, therefore, expect PCA-based approaches to give inferior performance. We exclude  $\mathbf{I}_{SS}^{Uns}$  from the assessment as its performance would suffer from the inaccurate data labelling. Only a more advanced design method (iterating over steps 1–3 from Sec. 3.1) would construct a useful indicator.

The performance assessment of the designed indicators on the testing dataset is shown in Tab. 1, taking into account the so-called confusion matrix elements (i.e., TP, FP, TN, and FN) and performance criteria (i.e., AC, PR, RC, and F1). The complexity of designed indicators is represented by the number of principal components  $n_{pc}$  for PCA-based approach and by the number of input variables  $n_p$  for the rest of approaches. We can directly see that the supervised ML approaches outperform the unsupervised ones when we compare similar structures. The only exception appear to be the RC criterion when evaluated for  $\mathbf{I}_{pat}$  indicator. There are two reasons for this performance drop: 1. RC is given up in training for the AC and precision as the dataset is more populated with data points of normal operation; 2. the industrial data labels indicate flooding based on other variables than pressure (the sole input to  $\mathbf{I}_{pat}$  indicator) and thus  $\mathbf{I}_{pat}$  indicator falls short in terms of model (input) adequacy (some extra input variables would explain flooding better). Note that, the first reason can be remedied by a modification to SVM objective and some proper tuning, which, however, is beyond the scope of this study.

Among the  $\mathbf{I}^{Uns}$ -type approaches, it is interesting that, although the structure of the reference indicator is optimal, the highest RC criterion (low FN) is achieved by  $\mathbf{I}_{pat}^{Uns}$ . Of course, this is paid off by worse accuracy as the classifier mis-indicates flooding (high FP) more often. The PCA-based indicator appears to be the least effective (in all criteria). This is attributed to the aforementioned inappropriate labelling in high dimensions.

Unlike for the unsupervised learning approaches, the performance of the  $I_{PCA}^{Sup}$  indicator is sufficient. It also appears that the  $I_{PCA}^{Sup}$  is more efficient compared to the  $I_{pat}^{Sup}$  indicator viewed by each performance criterion. The highest efficacy among supervised learning approaches is achieved for  $I_{ref}^{Sup}$  and  $I_{SS}^{Sup}$  indicators. These approaches already consider or can find the best possible input structure. It is noteworthy that  $I_{SS}^{Sup}$  achieves the best performance (almost 100% in all performance criteria) using a very simple structure. This effectively tells that the reference structure is overly complicated (some inputs are redundant) and that it is possible to indicate flooding with data from just two sensors. It is also a very interesting result as it allows the industrial practitioners to concentrate efforts regarding sensor maintenance towards smaller subset of online sensors. Surprisingly, pressure is not among the inputs selected for the best indicator. The input structure involves reflux flow  $\Delta F_R$  and the time difference of heating medium flow in the reboiler  $\Delta F_{reb,h,2}$ , which are both part of the reference indicator structure. It is possible that the two selected flow rates are measured with better precision and that they do not involve high-frequency fluctuations as pressure measurements do. The results need further validation in an industrial setup to validate the reference indicator (ground truth).

## 5. Conclusions

This contribution is focused on the design of a data-driven flooding indicator for an industrial debutanizer column. As the ground truth, a reference indicator is used that is designed according to the industrial experience. The effectiveness of unsupervised and supervised machine learning (ML) approaches was evaluated by considering various input structures. The results showed that the unsupervised learning approach can provide sufficient flooding indicators if appropriate input variables are used. The supervised ML approaches achieved higher effectiveness compared to unsupervised ML approaches, resulting from the direct usage of reference indicator labels. The results of supervised learning approaches revealed that the most accurate estimate of the debutanizer flooding is provided by the reflux flow rate and heating medium flow in the reboiler. Future work involves design of indicators for the chosen performance criterion in a multi-objective fashion. It would also be possible to design an unsupervised approach capable of choosing the optimal indicator structure much like the structure achieved with supervised ML.

## References

- B. E. Boser, I. M. Guyon, V. N. Vapnik, 1992. A training algorithm for optimal margin classifiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory. COLT '92. Association for Computing Machinery, New York, NY, USA, pp. 144–152.
- E. Forgy, 1965. Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics* 21, 768–769.
- L. F. Fuentes-Cortés, A. Flores-Tlacuahuac, K. D. P. Nigam, 2022. Machine learning algorithms used in pse environments: A didactic approach and critical perspective. *Ind. Eng. Chem. Res.* 61 (25), 8932–8962.
- M. Hubert, M. Debruyne, 2010. Minimum covariance determinant. *WIREs Computational Statistics* 2, 36–43.
- M. King, 2016. *Process Control: A Practical Approach*. Wiley.
- M. Mojto, K. Ebušský, M. Fikar, R. Paulen, 2021. Data-based design of inferential sensors for petrochemical industry. *Computers & Chemical Engineering* 153, 107437.
- J. Oeing, L. M. Neuendorf, L. Bittorf, W. Krieger, N. Kockmann, 2021. Flooding prevention in distillation and extraction columns with aid of machine learning approaches. *Chem. Ing. Tech* 93 (12), 1917–1929.
- K. Pearson, 1901. LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2 (11), 559–572.
- H. Peiravan, A. R. Ilkhani, M. J. Sarraf, 2020. Preventing of flooding phenomena on vacuum distillation trays column via controlling coking value factor. *SN Applied Sciences* 2 (10), 1–11.
- R. K. Pihlaja, 2008. Detection of distillation column flooding. U.S. Patent No. US 8.216.429 B2.
- G. Smith, 2018. Step away from stepwise. *Journal of Big Data* 5, 32.

# Can we deploy carbon capture, utilization and storage at a scale?

Ahmed Alhajaj<sup>a,b</sup>

<sup>a</sup>*Research and Innovation center on CO<sub>2</sub> and H<sub>2</sub> (RICH), Khalifa University, PO Box 127788, Abu Dhabi, United Arab Emirates*

<sup>b</sup>*Chemical Engineering Department, Khalifa University, PO Box 127788, Abu Dhabi, United Arab Emirates*

*ahmed.alhajaj@ku.ac.ae*

## Abstract

CO<sub>2</sub> emissions are set to rise in view of continuous growth in economic and supporting infrastructure. This means that the indispensable approach to decarbonize the economy in the medium term is through carbon capture, utilization and storage (CCUS). The main challenges in the large-scale implementation of CCUS network are costs of carbon capture, compression and transportation, energy requirements for CO<sub>2</sub> utilization and conversion, and limited market demand for end products. The development of strategic designs for such CCUS infrastructure at this scale is a non-trivial activity, which must take account of local industry specifics and development plans. A multiscale system modeling and optimization approach is developed to assess alternative CCUS processes and understand transient network behavior. The potential of CCUS in main potential markets in the United Arab Emirates such as simultaneous CO<sub>2</sub> storage with Enhanced Oil Recovery (CO<sub>2</sub>-EOR), CO<sub>2</sub> capture using microalgae and mineralization have been examined.

**Keywords:** CO<sub>2</sub> capture, Optimization, Multiscale Modeling, CCUS, techno-economic.

## 1. Introduction

The most recent Intergovernmental Panel for Climate Change (IPCC) report indicates that climate change is unequivocal and that without mitigation, average global temperatures may increase by 2-6 °C by 2100. As a result, most major economies have announced targets for reductions in GHG emissions. UAE is amongst the nations wishing to take a lead in this area and aspires for a flourishing low carbon economy. Sustainable development is a major aspect of Abu Dhabi's Vision 2030, and this is closely linked with achieving significant reductions in carbon emissions, maintaining economic growth and building a circular economy with minimum ecological footprint. Lately, the government of the UAE announced the country's net zero carbon strategic initiative in 2050. This aligns with Paris agreement signed in 2015 by 195 countries including the UAE, which aims to limit the increase in global average temperature to 1.5°C. In addition, the 2<sup>nd</sup> intended nationally determined contribution of the UAE is to reduce its greenhouse gas (GHG) emissions for the year 2030 by 23.5%, relative to the Business-As-Usual (BAU) scenario (i.e., a reduction of 70 million tons of CO<sub>2</sub> per year in 2030). These ambitious CO<sub>2</sub> reduction targets should go in line with developing technologies and markets for CO<sub>2</sub> capture, utilization and storage (CCUS) along with energy efficiency,

renewables, energy storage, and others. CCUS involves CO<sub>2</sub> capture from air or flue gas emitted from power plants and industrial sources using solvents or adsorbents, and then use it as a feedstock in an economically chemical, biological and industrial processes. There are many CO<sub>2</sub> utilization routes that has different market share, temporal storage potential, technological readiness level (TRL) and product value. The main utilization pathways at a scale (Fig. 1) are (1) chemicals; (2) building materials; (3) CO<sub>2</sub>-enhanced oil recovery (EOR); (4) soil carbon storage; (5) microalgae; and (6) fuel.

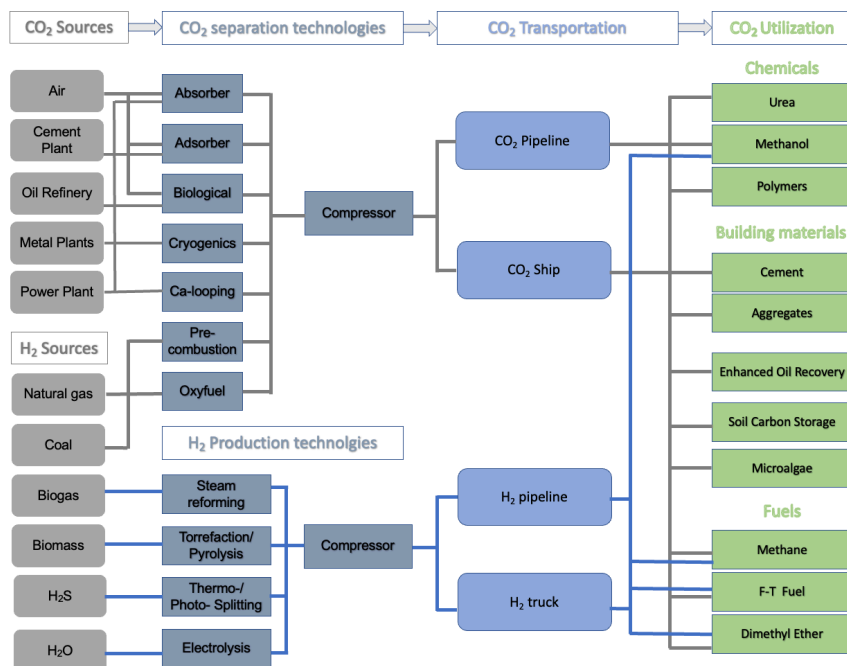


Figure 1: Superstructure of potential CCUS networks

CO<sub>2</sub> capture of around 100 kt yr<sup>-1</sup> and utilization for urea production has already been implemented commercially in the UAE. Globally, 140 Mt yr<sup>-1</sup> of CO<sub>2</sub> is used to produce 400 Mt yr<sup>-1</sup> of urea (Perez-Fortes et al., 2014). Other emerging chemical is polycarbonate polyols that will have a limited market of 10-50 Mt yr<sup>-1</sup> in 2050 (Hepburn et al., 2019). CO<sub>2</sub> fixation in waste streams from exhausted slag-based and carbon-based sorbents to produce carbonate and silicate aggregates for material construction is a promising long-term utilization route to reduce the overall ecological footprint of CCUS process. This has a large-scale potential in countries climbing the development ladder such as the UAE. One of the main anthropogenic CO<sub>2</sub> utilization route in oil exporting companies such as the UAE is CO<sub>2</sub>-Enhanced oil recovery (EOR). EOR is a tertiary approach applied to mature oil reservoirs to improve oil recovery. There are three main temporal phases in oil extraction: primary recovery, at the early stage of oil production, geologic pressure and pumps can push oil from the wellbore to the surface; secondary recovery maintains the pressure of the reservoir and improves oil productivity through injection of water or gas; tertiary recovery, which can recapture 30 to 60% of the original oil in place (OOIP). CO<sub>2</sub> injection decreases the viscosity of the remaining oil, swell it and detach it from the formation. This allows it to move freely in the reservoir reaching the production well,

while trapping CO<sub>2</sub> in the reservoir, leading to CO<sub>2</sub>-EOR and storage. Part of this CO<sub>2</sub> is produced and recycled back in the operation making it complex to quantify the potential of CO<sub>2</sub> utilization. In this work, a multiscale modeling approach based on system dynamics is demonstrated to examine the potential of combined CO<sub>2</sub>-EOR and storage. Another emerging utilization route is soil carbon storage (SCS). Understanding the carbon balance in the soil is essential to evaluate the potential of SCS. Soil stores organic carbon comprising of decaying plants, fungal, bacterial and animal matters. This also enhances the soil structure and reduces the erosion and carbon losses. CO<sub>2</sub> fixation in waste sorbents and carbon rich materials such as char can also play an important role in enhancing the carbon content and water within the soil enhancing the productivity of growing plants. An important utilization route in arid environments such as the UAE is to cultivate microalgae for simultaneous fixation of carbon dioxide (CO<sub>2</sub>) and production of high value-added products such as vitamins, antioxidants, and carotenoids for pharmaceutical use. Microalgae are organisms that are able to fix CO<sub>2</sub> with solar energy at higher rates than other plants with no need for arable land and freshwater resources — it can be cultivated using wastewater or salt water medium. Here, a multiscale model is demonstrated to examine the potential and scale of Microalgae cultivation in the UAE. Other utilization routes through recycling CO<sub>2</sub> back into different fuels (Fig. 1) involves production of green hydrogen at a reasonable price as most of the fuel cost comes from hydrogen production (i.e., >85% for methanol (Atsonios et al., 2016)). This is highly linked to the availability of renewables and cost of power. In the UAE, a low power production price was recorded at US\$0.015/kWhr. This should go in line with dispatching hydrogen economy at a scale, increasing efficiency and reducing CAPEX of electrolyzers. A price of US\$1.0/kg of hydrogen is required to deploy this pathway at a scale. It is worth noting that the fuel generated using this option will be burnt again and emits CO<sub>2</sub>, which emphasis an open carbon cycle. This might limit the impact of adoption of power to fuel technologies at a scale while looking at the impact of storage unless carbon is being captured and utilized again in a closed cycle. This work demonstrates the significance of multiscale model as a guide to examine the potential of three major utilization pathways namely CO<sub>2</sub>-EOR, microalgae and mineralization.

## **2. Multiscale modelling methodology**

This work builds upon multiscale modelling approach for CCUS system connecting models across different length and time scales (Fig. 2). At picosecond and angstrom scales, molecular dynamics simulations are performed to examine the potential of each sorbent to attract solutes within the available sites in the material. This helps in generating the missing data required to generate isotherm models for CO<sub>2</sub> capture in adsorbents. A continuum model based on partial differential and algebraic equations systems at a scale of seconds and length of meters are used to find the optimal scheduling of different adsorption steps. This helps in establishing the performance of different sorbents through capturing the relationships and trade-offs between structure-property and key performance indicators such as cost, purity, and recovery. Moving to the km scale of reservoir size and decades of EOR operations, a system dynamics model has been developed to quantify the demand of CO<sub>2</sub> for EOR, amount of CO<sub>2</sub> stored and recycled. This integrated multiscale models also applies to other utilization routes such as CO<sub>2</sub> capture using microalgae at the country level. This integrates genomic scale models based on flux balance analysis that describes the change in inter-cellular metabolites with kinetic based models that describes the interactions with extra-cellular environment. The optimal CCUS network can be obtained through Mixed Integer Linear Programming

(MILP) (Alhajaj and Shah, 2020) or Mixed Integer Non-Linear Programming (MINLP) approach.

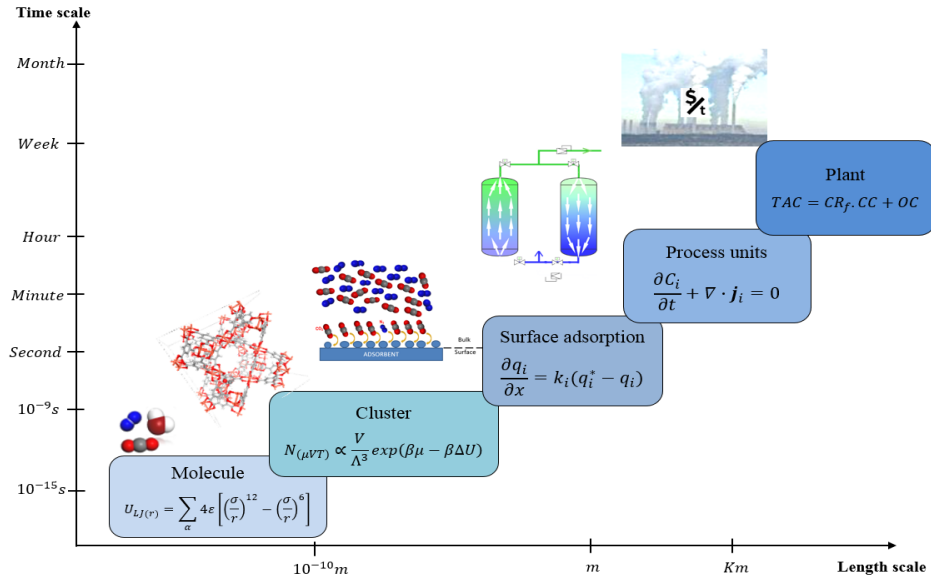


Figure 2: Illustration of multiscale model (Reproduced from (Balogun et al., 2021) with permission of the Royal Society of Chemistry)

### 2.1. Problem definition

A multiscale model integrating models across different length and time scales have been used to find the potential of CCUS considering main markets in the UAE, namely CO<sub>2</sub>-EOR, microalgae and mineralization. The main objective was to examine the potential in meeting the 2<sup>nd</sup> national reduction target standing at 70 Mt CO<sub>2</sub> yr<sup>-1</sup>. The multiscale CCUS supply-chain network model was formulated as an MILP problem, similar to our earlier work (Alhajaj and Shah, 2020). The selection of CO<sub>2</sub> sources, main capture technologies with varying configurations, absorbents, and adsorbents in addition to microalgae, transportation modes and CO<sub>2</sub> utilization routes have been considered in this work. The model has been used to find the cost-optimal network that meets a specific CO<sub>2</sub> decarbonization target with a positive long-term environmental impact. The latter is considered through performing life cycle analysis in utilization pathways. This captures the energy requirements for synthesis and upgrading of fuels and materials in addition to the duration of CO<sub>2</sub> storage.

In this work, a multiscale model is used to obtain realistic data for the maximum storage capacity and utilization potential per year, presented as inequality constrains in the MILP problem. This has been used to capture the system dynamics of CO<sub>2</sub>-EOR enterprise. The adoption of anthropogenic CO<sub>2</sub> as a solvent to enhance oil recovery in aging reservoirs combined with CO<sub>2</sub> storage to mitigate climate change is a complex process due to the interplay of environmental, technical and market targets. A system dynamics model has been adopted to account for the complex interactions between the dynamics of oil production, climate change mitigation targets, CO<sub>2</sub> sources market, oil production targets, carbon credits, policies, and oil market subsystems (Santos et al., 2021). This has been used to obtain the yearly demand of fresh CO<sub>2</sub> and the potential of CO<sub>2</sub>-EOR storage in mitigating greenhouse gas emissions in the UAE and meeting the 2<sup>nd</sup> national reduction target.

To examine the potential of embracing biological carbon capture and utilization using microalgae at a large scale, bio-process model imbedded with genomic scale model have been used to examine the potential of biological CO<sub>2</sub> capture and utilization from 400 MW combined cycle gas turbine (Alhajaj et al., 2016). This has identified the pond size (length scale) and time scale required for deep decarbonization of flue gas using microalgae. All data obtained from the detailed analysis of utilization pathways and capture technologies are fed into the supply-chain network to obtain realistic and cost-optimal decarbonization pathways.

### 3. Preliminary Results

The results of the multiscale model have identified CO<sub>2</sub>-EOR as a major utilization and storage route in the UAE. This can mitigate 10-20% of the 2<sup>nd</sup> national reduction target of the UAE (i.e., 13.6-25.6 Mt CO<sub>2</sub> trapped in the reservoir). On contrary, microalgae CO<sub>2</sub> capture and cultivation from combined cycle power plant require a massive land with a size of Dubai Expo (i.e., >10000 hectares) to capture 75% of the flue-gas. A higher capture rate (Fig. 3) is difficult due to the high evaporation rates and emissions to atmosphere resulting from the exponential growth in number of ponds. This demonstrates the complexities in applying microalgae at a scale to mitigate climate change due to the slow process of CO<sub>2</sub> fixation. The main driver to implement microalgae projects should be targeted to small scale high-value and green products.

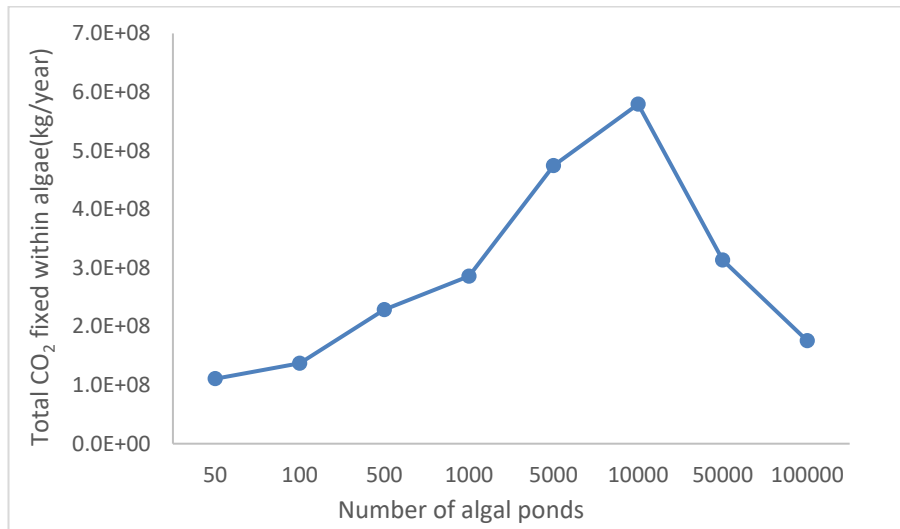


Figure 3: Illustration of microalgae CO<sub>2</sub> capture and utilization from large scale power plant

The results of this study have demonstrated the difficulties of implementing solo-CCUS route at a scale. Therefore, synergized adoption of CCUS with storage should be implemented to mitigate climate change. Emerging markets such as CCUS in building materials and soil carbon storage can complement large scale storage in reservoirs and reactive Basaltic rocks. Other enabling technologies such as hydrogen economy should be supported to create the mass for economies at a scale. This will help in creating new large-scale markets for CO<sub>2</sub>-Fuel, especially for aviation.



#### 4. Conclusion

Implementing large scale Carbon Capture, Utilization and Storage (CCUS) projects is required to mitigate climate change and meet national CO<sub>2</sub> reduction targets. In this work, we have developed multiscale models that has helped in examining the potential of large scale CCUS pathways such as CO<sub>2</sub>-EOR, mineralization and microalgae in the UAE. The results of this study have highlighted the potential of implementing CCUS at a large scale in the UAE with positive environmental impacts. CO<sub>2</sub>-EOR with storage can play an important role in developing CCUS infrastructure, which helps in establishing a network and economies of scale for wider adoption. Emerging markets such as mineralization, soil storage and saline aquifers can help in accelerating the rate and scale of CCUS implementation in the UAE. Microalgae and CO<sub>2</sub>-based fuels have an important market such as hard-to-abate aviation but further decarbonization strategies should be implemented to ensure larger net CO<sub>2</sub> reductions. This depends on the availability of enabling resources such as renewables, heat, green hydrogen, and storage sites. These CCUS pathways are region-specific and might not be replicated at a large scale elsewhere. There is also a need for policies and incentives to de-risk the CCUS business and liability, which will assure deep decarbonization of the society. Overall, this study has demonstrated the need for integrated tools and models across different length and time scales to capture the system dynamics using various key performance indicators such as cost and life cycle analysis. This integrated tool can guide policy makers and governments on selecting right incentives and policies that help in wider adoption of CCUS with desired social benefits.

#### 5. Acknowledgements

Funding for this work has been provided by Khalifa University of Science and Technology under projects CIRA-2021-067 and RC2-2019-007. Computational resources from RICH are also acknowledged.

#### References

- A. Alhajaj, N. Shah, , 2020, Multiscale design and analysis of CO<sub>2</sub> networks, *International Journal of Greenhouse Gas Control*, 94, 102925.
- A. Alhajaj, N. Mac Dowell, N. Shah, A techno-economic analysis of post-combustion CO<sub>2</sub> capture and compression applied to a combined cycle gas turbine: Part I. A parametric study of the key technical performance indicators, *International Journal of Greenhouse Gas Control*, 44, 26-41
- K. Atsonios, K. Panopoulos, E. Kakaras, 2016, Investigation of technical and economic aspects for methanol production through CO<sub>2</sub> hydrogenation, *International Journal of Hydrogen Energy*, 41, 4, 2202-2214.
- H. Balogun, D. Bahamon, S. AlMenhali, L. Vega, A. Alhajaj, 2021, Are we missing something when evaluating adsorbents at the system level?, *Energy & Environmental Science*, 14, 12, 6360-6380.
- C. Hepburn, E. Adlen, J. Beddington, E. Carter, S. Fuss, N. Mac Dowell, J. Minx, P. Smith, C. Williams, 2019, The technological and economic prospects for CO<sub>2</sub> utilization and removal, *Nature*, 575, 7781, 87-98.
- M. Perez-Fortes, A. Bocin-Dumitriu, E. Tzimas, 2014, CO<sub>2</sub> utilization pathways: Techno-economic assessment and market opportunities, *Energy Procedia*, 63, 7968–7975.
- R. Santos, S. Sgouridis, A. Alhajaj, 2021, Potential of CO<sub>2</sub>-enhanced oil recovery coupled with carbon capture and storage in mitigating greenhouse gas emissions in the UAE, *International Journal of Greenhouse Gas Control*, 111, 103485.

# Prediction of capacitance using artificial neural networks for carbon nanofiber-based supercapacitors

Kiran Donthula, Naresh Thota, Sarath Babu Anne, Manohar Kakunuri\*  
Department of Chemical Engineering, National Institute of Technology  
Warangal, Telangana, India

\*[manohar@nitw.ac.in](mailto:manohar@nitw.ac.in)

## Abstract

Carbon nanofibers (CNF) with different morphologies and microstructures were widely used as electrode material for supercapacitors since they exhibit high electrolyte-electrode interface, flexibility, and tunable physicochemical properties. This CNFs-based electrode electrochemical performance is greatly influenced by their surface area, pore structure, and fiber diameter. No credible physical model is available that can predict the performance of supercapacitors based on the CNFs physicochemical properties. In this work, we have used a data-driven Artificial Neural Network (ANN) model to predict the performance of CNF electrode-based supercapacitors based on electrode material microstructural properties and electrochemical operational parameters. The proposed model demonstrates its feasibility in predicting specific capacitance with a correlation coefficient (R-square) value close to 0.95. A sensitivity analysis was also conducted to understand the effect of independent input parameters on a single output parameter (specific capacitance).

**Keywords:** ANN, Supercapacitor, Data-driven, Energy storage, Machine learning.

## 1. Introduction

Recent developments in supercapacitors have attracted great attention due to their superior power density and cyclic stability when compared to conventional batteries (Panda et al., 2020). The use of carbon nanomaterials, such as activated carbons, nanoparticles, nanofibers, etc., as electrode materials for supercapacitors, has been widely studied because of their high surface area, high conductivity, and low production cost (He et al. 2020; Lokhande et al. 2020; Mathis et al. 2019). Compared to other electrode materials, CNFs are unique in that they are both flexible and possess tunable surface area, porosity, and fiber diameter. These characteristics make them ideal electrode materials for flexible supercapacitors. Hence, these CNFs have been extensively studied as electrode materials for supercapacitors (Chandu et al., 2020). However, the improvement in the specific capacitance is inadequate due to charge storage at the electrode-electrolyte interface. Activation and physical etching methods were used to engineer the microstructural properties of these CNFs to enhance their electrochemical performance. The microstructural properties, micro (<2 nm) and mesopore (2nm-50nm) volumes, specific surface area (SSA), and fiber diameter greatly rely on the polymer precursor and carbonization process. In addition to these microstructural properties, supercapacitor performance depends on voltage window, scan rate, cell configuration, and electrolyte characteristics. As the performance of CNF-based supercapacitors depends on these many

parameters, there is a significant need to develop a model that correlates these parameters to the performance of the supercapacitors. It is imperative to note that the development of a model does not only assist in predicting the performance of the supercapacitor, but also in planning the experiments to achieve the desired microstructural properties for enhancing its performance. The qualitative deviation of electrochemical performance with electrode SSA, micropore, and mesopore volumes has been examined through experimental studies. However, no physical model has been developed that predicts how these parameters interact to affect electrochemical performance in a synergetic manner. Alternatively, experimental data may be used to develop a data-driven machine-learning model that addresses this problem.

In recent years, various machine learning models have been applied to determine the synergetic effects of multiple variables on the electrochemical performance of energy storage devices. Among different machine-learning models, an ANN was widely used to develop a model using complex and nonlinear data (Himanenmanen et al. 2019; Farouji et al. 2022). In a recent study, Rahimi et al., 2022, developed an ANN-based data-driven model that correlated microstructural characteristics of activated carbon electrodes, electrode synthesizing procedures, and operational conditions that affect the electrode's performance. The developed model accurately predicted the performance of activated carbon electrodes with an R-square value of 0.9. To our knowledge, there are no studies on CNF-based electrodes using data-driven machine-learning models. In this work, we have developed an ANN model to predict the performance of CNF-based electrodes using various independent input variables. Our goal is to examine the effect of input variables independently and synergistically on the prediction of specific capacitance of the CNF-based electrode.

## 2. Material and methods

### 2.1 Experimental data collection and preprocessing

The ANN model in this work was trained using seven input parameters, which include carbonization temperature ( $C_t$ ), SSA, total pore volume ( $V_t$ ), micropore volume ( $V_{mic}$ ), mesopore volume ( $V_{meso}$ ), nanofiber diameter (NFD), and voltage window (VW) and a single output parameter (Specific capacitance). Data for both input and output parameters were collected from 70 research articles reporting experimental data, each reference paper we took one single data set. The data was limited to the pristine CNFs-based electrode studies with the aqueous electrolytes and incorrect or outdated data can also impact the accuracy of the model so we considered only recent research articles data and inconsistent formatting or scaling of data can affect the model's accuracy so as a preprocessing step, all the parameters were normalized between 0 and 1. This normalization enables the ANN to learn more quickly since the output is more sensitive to changes in the input (Rafiq et al. 2001). The following function was used to normalize the data in this study:

$$X_{nor} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

where  $X$  denotes the variable value and  $X_{nor}$  denotes the value of the normalized variable,  $X_{min}$  and  $X_{max}$  correspond to the minimum and maximum values of the variables, respectively.

#### 2.1.2 Artificial neural network

It is possible for an ANN to learn rather than be programmed to make predictions. The ANN consists of input parameters, output parameters, and several hidden layers. Depending on the weight ( $W_{ij}$ ) and bias ( $B_i$ ), the hidden layer will augment or cripple the received information. Herein, our goal is to use this ANN and the experimental data

collected from the literature to build a data-driven model as shown in Fig. 1 that can predict the specific capacitance performance of a CNF electrode-based supercapacitor. Our model was trained and tested using input and output data collected using the supervised machine learning method and the neural network toolbox in MATLAB 2021a. The artificial neural network consists of one hidden layer with a different number of neurons (11 to 23), a log-sigmoid activation function in a hidden layer, and a purelinear transfer function at the output layer.

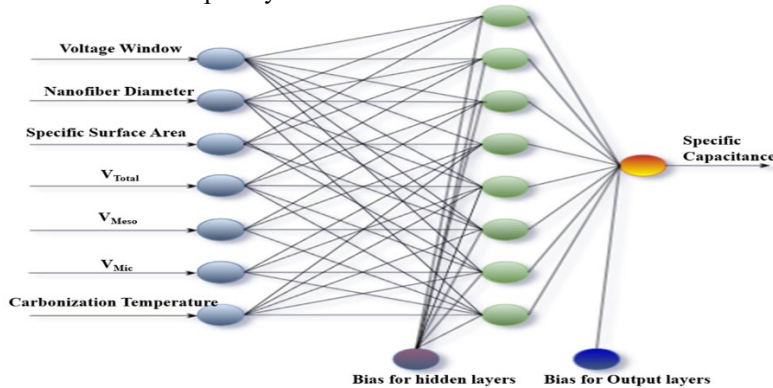


Figure 1. The neural network model of ANN used in this study.

An ANN model was trained using Levenberg–Marquardt with backpropagation of errors (LM-BP)(Willis et al.1991) Levenberg-Marquardt algorithm appears to be the fastest method for training moderate-sized feedforward neural networks (up to several hundred weights). It also has an efficient implementation in MATLAB software, because the solution of the matrix equation is a built-in function. The data was divided into three subsets of 70%, 15%, and 15% for training, testing, and validation, respectively. The schematic illustration of the dataset preparation procedures and modeling implementation is shown in Fig. 2.

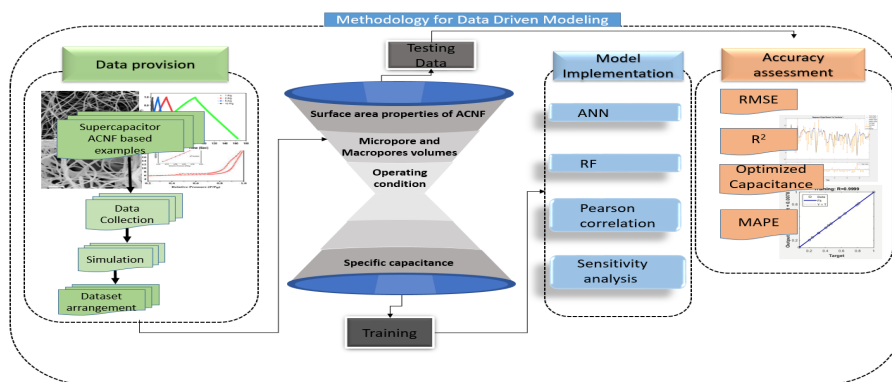


Figure 2. Schematic representation of data-driven modeling methodology

The number of neurons in an ANN model's hidden layer also significantly affects the prediction. To avoid overfitting (a high number of neurons) and offsetting (a low number of neurons), the number of neurons in the ANN model were optimized(Karsoliya, 2012). In this study, the number of neurons in hidden layers (hidden layer size) varied between 11 and 23. After the training, the performance of ANN was evaluated using the test data

set. The error of ANN is determined by the mean square error (MSE) and correlation coefficient (R-Square).

### 3. Results and Discussion

#### 3.1 Result of ANN model

The Pearson correlation coefficient (PCC) was used to investigate the linear relationship and co-occurrence of input and output parameters, as shown in Fig.3. The PCC analysis reveals the dependency of the specific capacitance (output parameter) on all the input parameters. The reliance is in the decreasing order of  $SSA > V_{meso} > V_t > V_{mic} > C_t > VW$ . NFD, an independent input variable, showed a negative correlation (-0.16) with specific capacitance in the group of seven input parameters. Due to this, linear/inverse relationships cannot be used accurately to predict capacitance.

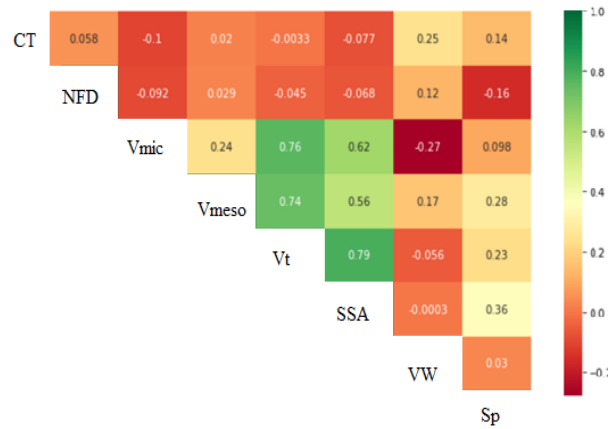


Figure 3. Pearson correlation matrix of the seven input variables and single output variable (specific capacitance).

Table 1 summarizes the effect of hidden layer size on the performance (as MSE & R-square) of the training algorithm LM-BE. Increasing the number of neurons in the hidden layer from 11 to 19 resulted in a moderate decrease in the MSE from 0.0554 to 0.0031, and a further increase in neurons from 19 to 23 resulted in a moderate increase in the MSE to 0.006. As a result, a network with 19 neurons with least MSE value (0.0031) was used for sensitivity analysis studies. Furthermore, the network with 19 neurons demonstrated excellent correlation coefficients, with R-square values exceeding 0.95 for both training and testing achieved performance metrics are consistent across multiple training repetitions.

Table 1: Performance of ANN vs. Number of neurons in hidden layer

Hidden layer size	Training		Testing	
	MSE	R <sup>2</sup>	MSE	R <sup>2</sup>
11	0.01550	0.786	0.05540	0.918
13	0.00117	0.888	0.00700	0.921
15	0.00392	0.873	0.00310	0.967
17	0.00238	0.878	0.00835	0.910
19	0.00035	0.955	0.00286	0.955
21	0.00703	0.916	0.00160	0.854
23	0.00648	0.867	0.00200	0.866

Based on the results presented in Fig. 4, it can be concluded that substantial agreement exists between the actual and predicted values of specific capacitance during both the training and testing phases. Mostly, ANN models can be compared to other machine learning models such as Decision Trees, Random Forests, Support Vector Machines, etc., so Random Forest (RF) model is used to compare the ANN accuracy results. ANN model showed better accuracy than RF model.

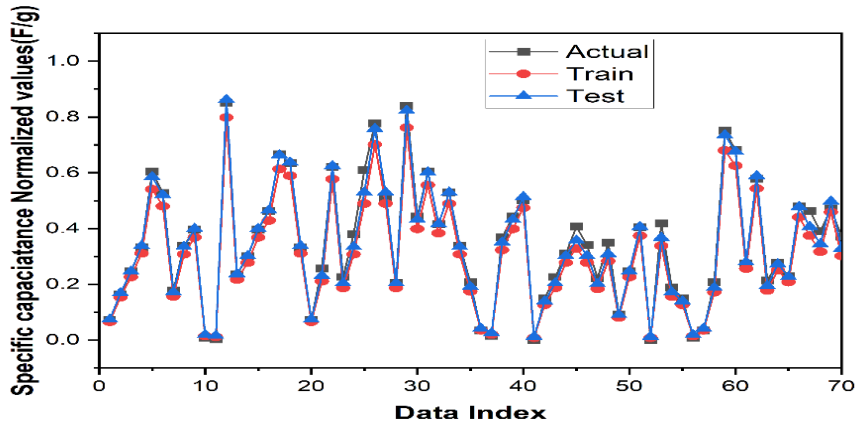


Fig.4 The predicted and actual values of specific capacitance in both test and training, using the ANN.

### 3.2 Sensitivity analysis

A sensitivity analysis was conducted by selecting input parameters precisely to determine how they affect the performance of CNF-based supercapacitors. The sensitivity analysis is presented in detail in table 2. First, all input variables were considered in this analysis, and then a single input parameter was excluded from other analyses. Microstructural variables appear to have a substantial impact on the electrochemical performance of CNF-based supercapacitors. According to this analysis, specific capacitance is greatly influenced by  $V_{meso}$  and SSA, in agreement with experimental observations (Xie et al. 2020). In addition, the mean absolute percentage error (MAPE) was also used to quantify the degree of dependency. The MAPE values for the input set excluding  $V_{meso}$  and the input set excluding SSA rose to 4.1% and 4.74%, respectively. This increase in error value indicates the importance of these microstructural properties on the performance of CNF-based supercapacitors. The input set without NFD, however, showed lower sensitivity, with MAPE of 1.57.

Table 2 Sensitivity analysis of the ANN model versus independent variables.

Model Input	Input set	MAPE		
		Train	Test	Total
	All	0.56	1.10	0.58
Structural	All exclude Vmic	3.20	3.56	2.93
	All exclude Vmeso	4.63	4.20	4.10
	All exclude SSA	5.14	5.96	4.74
	All exclude Vt	2.10	2.90	2.01
	All exclude NFD	1.90	1.00	1.57
Carbonization	All exclude Ct	2.88	1.69	2.41
Operational	All exclude VW	3.12	3.39	2.84

#### 4. Conclusion

The ANN model predicts the CNF-based electrode performance accurately using microstructural properties and operational parameters as input. In addition, the proposed model with 19 neurons showed a high R-square value of 0.95 and a lower MSE value of 0.003. The results of a sensitivity analysis on seven input parameters suggested that SSA and  $V_{\text{meso}}$  have a significant impact on electrode performance. In contrast, the input parameter NFD has a relatively small effect. The developed model can be employed to design nanofiber-based electrode materials experimentally, which reduces the need for labor-intensive experimentation to optimize the microstructural properties.

#### References

- P. K. Panda, A. Grigoriev, Y. K. Mishra, and R. Ahuja, 2020, Progress in supercapacitors: roles of two dimensional nanotubular materials *Nanoscale Adv.*, 2, 70–108.
- P. E. Lokhande, U. S. Chavan, and A. Pandey, 2020, Materials and Fabrication Methods for Electrochemical Supercapacitors: Overview, *Electrochemical Energy Reviews*, Springer Singapore, 155–186.
- J. He, D. Zhang, Y. Wang, J. Zhang, B. Yang, H. Shi, K. Wang, Y. Wang, 2020, Biomass-derived porous carbons with tailored graphitization degree and pore size distribution for supercapacitors with ultra-high rate capability, *Appl. Surf. Sci.*, 515, 146020.
- T. S. Mathis, N. Kurra, X. Wang, D. Pinto, P. Simon, Y. Gogotsi, 2019, Energy Storage Data Reporting in Perspective-Guidelines for Interpreting the Performance of Electrochemical Energy Storage Systems, *Adv. Energy Mater.*, 9.
- V. V. Chandu. Muralee Gopia, R. Vinodha, S. Sambasivamb, I. M. Obaidatb, 2020, recent progress of advanced energy storage materials for flexible and wearable supercapacitor: From design and development to applications, *J. Energy Storage*, 27, 101035.
- M. S. Farouji, H. V. Thanh, Z. Dai, A. Mehbodniya M. Rahimi, U. Ashraf, A. E. Radwan, 2022, Exploring the power of machine learning to predict carbon dioxide trapping efficiency in saline aquifers for carbon geological storage project *Journal. Clean. Prod.*, 372, 133778.
- L. Himanen, A. Geurts, A. Foster, P. Rinke, 2019, Data-Driven Materials Science: Status, Challenges, and Perspectives *Advanced. Sciences*, 6, 1900808.
- M. Rahimi, M. H. Abbaspour-Fard, A. Rohani, 2022, synergetic effect of N/O functional groups and microstructures of activated carbon on supercapacitor performance by machine learning, 2022, *J. Power Sources*, 521, 230968.
- M. Y. Rafiq, G. Bugmann, D. J. Easterbrook, Neural network design forengineering applications, 2001, *Comput. Struct.*, 79, 1541–1552.
- M. J. Willis, C. Di Massimo, G. A. Montague, M. T. Tham, and A. J. Morris, 1991, Artificial neural networks in process engineering *IEE Proc. D Control Theory Appl.*, 138, 256–266.
- S. Karsoliya, 2012, Approximating Number of Hidden layer neurons in Multiple Hidden Layer BPNN Architecture *Int. J. Eng. Trends Technol.*, 3, 714–717.

# Model size determination using data analysis tools in the context of reaction network identification

Carolina S. Vertis,<sup>a</sup> Belmiro P.M. Duarte,<sup>a,b</sup> Nuno M.C. Oliveira<sup>a,\*</sup>

<sup>a</sup>*Univ Coimbra, CIEQPFF, Department of Chemical Engineering, Rua Sílvia Lima — Pólo II, 3030–790 Coimbra, Portugal*

<sup>b</sup>*Polytechnic Institute of Coimbra, ISEC, Department of Chemical & Biological Engineering, Rua Pedro Nunes, 3030–199 Coimbra, Portugal*

\**nuno@eq.uc.pt*

## Abstract

Data dimensionality analysis or model size determination (identifying the true dimension of the underlying data variant space from experimental data) is a key step for the accurate identification of reaction networks and kinetic modeling of chemical systems. Here, several data analysis tools are compared regarding their relative performance in the identification of the model size from simulated noisy data concerning a challenging case study that presents low reaction extents. Depending on the characteristics of the data used, distinct choices of analytic methods can be suggested. The complementary application of different methods to reinforce the confidence of this analysis is also considered.

**Keywords:** Data dimension, Model size, Reaction networks, Systematic methods, Statistical approaches.

## 1. Introduction

Accurately describing chemical reaction systems, including their dynamics, generally requires using systematic frameworks and a detailed set of experimental data. Models under and over-parameterized might present unsatisfactory predictive abilities during the calibration (training) and validation (testing) phases, thus compromising their subsequent applications such as process optimization, design, and control (Bonvin et al., 2016). A common concern to modelers is to maximize the extent of the information explained by the model, while avoiding the addition of model features that can be essentially attributed to noise, thus preventing overfitting.

One crucial step during the modeling task is the identification of the underlying data variant space, whose dimensionality can be described as the model size for the system. The integration of various existing methods within a systematic methodology for chemical reaction model development has been proposed and tested (Vertis, 2022). In this context, the model structure (or the system stoichiometry) can be further identified by determining an abstract (latent) space without using any information about the system kinetics (Bonvin and Rippin, 1990). Here, we compare classical empirical methods such as the scree test (Cattell, 1966), the fractional variance analysis (Bro and Smilde, 2014), and the Kaiser test (Cliff, 1988), in addition to the parametric F-test (Malinowski, 1989) and the procedure for non-parametric cross-validation by eigenvectors (Wise and Ricker, 1991), concerning their ability to identify the “true” model size from noisy reaction data. These data represent cumulative extents corrupted with (white) noise generated from in silico experiments, simulated at different experimental conditions. This type of problem may become challenging if the concentration of some species is of the same order of



magnitude as the system noise, leading to various model identifiability issues. As an illustration example, a simulated system including 6 species and 4 reactions is considered. The results allow a comparison of the efficiencies of the methods in capturing the correct model size, and their requirements in terms of the data characteristics. These results also evidence that the choice of the best method to identify the model size depends on the data itself, especially on the characteristics of the underlying noise component.

## 2. Methods considered

The information from the Singular Value Decomposition (SVD) is very useful to determine the model size  $R$ , since it allows the data decomposition in directions ordered by decreasing variance. Thus, many tools for model size identification rely on the analysis of (squared) singular values / eigenvalues. A collection of them is briefly listed next.

### 2.1. Heuristic approaches

This section introduces three empirical (*ad hoc*) methods, available in the literature, for determining  $R$ : the scree test, the Kaiser rule, and the analysis of fractional variances.

#### 2.1.1. Scree test

The scree test assumes that the relevant information can be distinguished from random noise in terms of singular values and that the magnitude of the random noise variation seems to stabilize linearly with the number of components (Cattell, 1966). The common graphical analysis based on heuristic principles requires a visual analysis of the eigenvalues plot and the use of a given threshold.

#### 2.1.2. Kaiser rule

Kaiser's rule assumes that if a component has an eigenvalue greater than one, it explains the variation of more than one original variable (Cliff, 1988). This led to the rule of selecting all components with eigenvalues greater than one; however, recent works demonstrate this rule can be inadequate (Bro and Smilde, 2014).

#### 2.1.3. Fractional variance analysis

Another rule of thumb is to include the first  $R$  eigenvectors such that the sum of their variances (represented by eigenvalues) is up to an amount of the total, such as 95% or 99%, which can be seen as an empirical rule to match model accuracy and parsimony.

### 2.2. Statistical $F$ -test

Malinowski (1989) proposed a parametric test to determine  $R$  (within a specified significance level) based on the Fisher variances ratio obtained from pools of samples with normal distribution and i.i.d.

### 2.3. Cross-validation procedure

The goal of the method selected for cross-validation consists of finding the number of model components that leads to the best overall prediction in the testing phase, established through the minimum PRESS-CV (predicted residual error sum of squares evaluated in a cross-validating approach). The original data matrix  $\mathbf{D}[n_o \times n_s]$  is randomly sampled (without replication) into  $K$  data files of size  $[n_c \times n_s]$  and  $[n_v \times n_s]$ , respectively, with  $n_c + n_v = n_o$ . For clarification, let  $\mathbf{D}^{-k}[n_c \times n_s]$  and  $\mathbf{D}^k[n_v \times n_s]$  be the matrices of data used, respectively, for model training and testing in  $k^{\text{th}}$  iteration ( $k \in \{1, \dots, K\}$ ). Then, the SVD of  $\mathbf{D}^{-k}$  is carried out resulting in  $\mathbf{U}[n_c \times n_c]$ ,  $\mathbf{\Sigma}[n_c \times n_s]$  and  $\mathbf{V}[n_s \times n_s]$  orthogonal matrices where  $\mathbf{U}$  and  $\mathbf{V}$  contain, respectively, left and right eigenvectors of  $\mathbf{D}^{-k}$ , and  $\mathbf{\Sigma}$  is a rectangular diagonal matrix containing the singular values. This structure is then used (in an internal loop to that of  $k$ ; the counter is  $s$ ) to construct a projection matrix  $\mathbf{M}$  from  $\mathbf{V}_{R^*}[(n_s - 1) \times R]$  which, on turn, is obtained from the columns 1 to  $R$  of  $\mathbf{V}$  by stripping

the  $s^{\text{th}}$  row. For every  $s = 1, \dots, n_s$  a column vector  $\mathbf{d}_s^k [n_v \times 1]$  containing the variable  $s$  is removed from  $\mathbf{D}^k$ , resulting in a smaller matrix  $\mathbf{D}_*^k [n_v \times (n_s - 1)]$ . Afterwards, the matrix of scores for data validation, say  $\mathbf{T} [n_v \times R]$ , is then predicted by projecting  $\mathbf{D}_*^k$  onto  $\mathbf{M}$ . Finally, the variable  $s$  in  $\mathbf{D}^k$  can be estimated as a linear combination of  $\mathbf{T}$  using the excluded row vector  $\mathbf{v}_{R,s} [1 \times R]$ .

This strategy is commonly designated as *cross-validation by eigenvector* to distinguish it from classic cross-validation where the goal is estimating a model using (partial) least-squares. In contrast, in principal components models (or models from SVD of data) the left-out elements and their predicted values are not independent, resulting in a decreasing PRESS-CV with the number of components (never reaching a minimum) if the traditional (k-fold or leave-one-out) cross-validation approaches are considered (Bro et al., 2008). This result would lead to the incorrect conclusion that all components are significant, thus overestimating the optimal model size. Here, in practice, the problem is handled by predicting the missing variable from the model eigenvectors extracted from the training dataset by excluding one of the columns of the testing data set such that the expectation is maximized (Wise and Ricker, 1991).

### 3. Case study

The example considered for comparison consists of a catalytic reaction system where 6 chemical species are linked by 4 chemical reactions – the reaction network in Figure 1 (right panel). The reaction kinetics follow the Langmuir-Hinshelwood model (1) where species A, B, and C are adsorbed in the catalyst with a single active site. In this equation, for the reaction  $j = 1, \dots, 4$ ,  $k_j$  is the kinetic constant in  $\text{min}^{-1}$ ,  $C_{\text{reac}_j}$  and  $\alpha$  are, respectively, the normalized concentration and the stoichiometric number of moles of the reactant species, and  $K_s$  (inverse molar ratio units, dimensionless) and  $C_s$  (molar ratio) are, respectively, the adsorption equilibrium constant and the normalized concentration respecting the adsorbed species  $s = A, B, C$ . Both  $k$  and  $K$  parameters are temperature-dependent following the Arrhenius (2) and van't Hoff (3) equations, where  $A_0$  (in  $\text{min}^{-1}$ ) is the pre-exponential factor and  $E_a$  (in  $\text{J}\cdot\text{mol}^{-1}$ ) is activation energy of the reaction,  $R$  (in  $\text{J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$ ) is the universal gas constant,  $T$  (in K) is the experimental temperature, and,  $\Delta H_{\text{ad}}$  (in  $\text{J}\cdot\text{mol}^{-1}$ ) and  $\Delta S_{\text{ad}}$  (in  $\text{J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$ ) are the enthalpy and entropy changes of the species adsorption.

$$r_j = \frac{k_j C_{\text{reac}_j}^\alpha}{(1 + K_A C_A + K_B C_B + K_C C_C)}, \quad j = 1, \dots, 4 \quad (1)$$

$$k_j = A_{0j} \exp\left(\frac{-E_{aj}}{RT}\right), \quad j = 1, \dots, 4 \quad (2)$$

$$K_s = \exp\left(\frac{-\Delta H_{\text{ad}_s}}{RT} + \frac{\Delta S_{\text{ad}_s}}{R}\right), \quad s = A, B, C \quad (3)$$

The in-silico experiments were simulated at 50, 70, and 90 °C, respectively, lasting for 30 min with the reaction starting from pure A; thus, three data sets were produced by sampling the process every 2 min. White noise with 0.015 standard deviation was added to the entire set of concentration data. In Figure 1 the model dynamics for a single experiment is presented. As can be seen, the species E is present only in very small concentrations and is strongly affected by noise, leading to potential model identification issues.

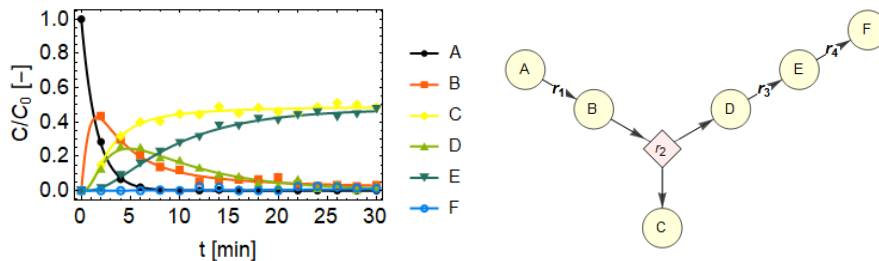


Figure 1 – Model simulation (continuous curves) and noisy experimental data (discrete points) at 70 °C, and model structure (reaction network) on the right-hand side.

The methods described in Section 2 were applied to cumulative molar changes (species extent data) to compare their performance in identifying the correct model size. Note that apart from the measurements, no additional information about the reaction system is used in this task. Also, we considered the cumulative molar changes as they allow minimizing the error in estimation of changes, compared to the estimation of derivatives. In Figure 2 the results obtained for the heuristic approaches are shown. The distribution of the logarithm of the variances in (a) does not present a clear sharp curve, leading to an inconclusive result. Plot (b) shows that the Kaiser test overestimates the model size, indicating a 5-dimensional model. Finally, the plots (c) and (d) for the fractional variance analysis indicate a model with 3 components. However, if 1% of the noise injected was used as a threshold criterion, only a single model component would be selected.

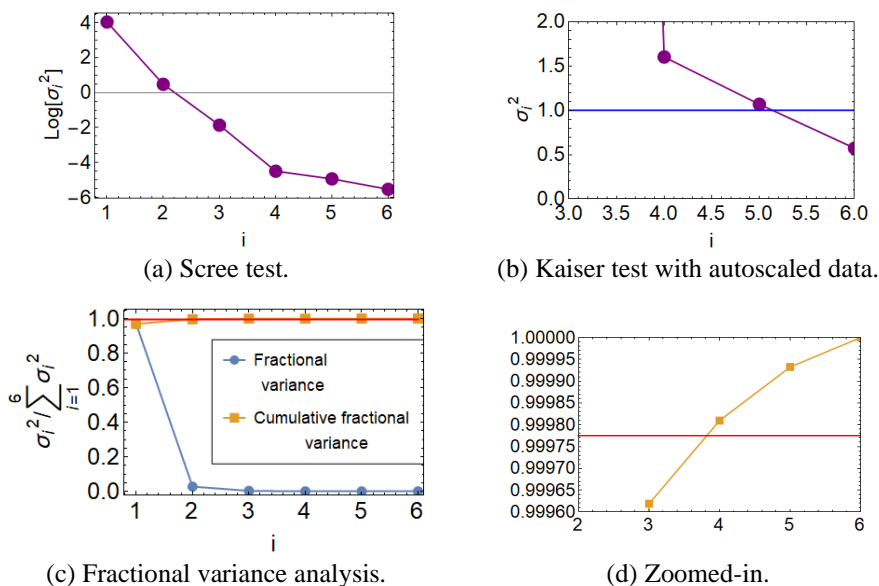


Figure 2 – Heuristic methods results based on data variance analysis.

The F-test indicates that 3 model components are statistically significant. The results are shown in Table 1 and Figure 3. The  $F$ -distribution plots indicate the respective PDF and CDF profiles for different pool sizes. The results demonstrate that  $F < F_c$  until the 2-dimensional pool size analysis (where the vertical red dashed line is located to the right of the vertical black dot-dashed line and the horizontal red dashed line is located above

the horizontal black dot-dashed line). When the 3<sup>rd</sup> lowest component is included, the opposite behavior is observed, i.e.,  $F > F_c$ . Therefore, according to these results, the 3<sup>rd</sup> model component contains structural information related to the data variability that is greater than the noise variability, and hence, it must not be discarded from the pool of noisy components.

Table 1 - F-test results.  $F_c$ : critical  $F$  value for  $\alpha = 5\%$ .  $\alpha^\ddagger$  indicates the percentage of the remain CDF function for each  $F$  observed (cutoff value if that  $F$  was considered critical).

$R$	$n_s - R$	$v(R) \times 10^5$	$v(0) \times 10^5$	$F(1, n_s - R)$	$F_c(1, n_s - R)$	$\alpha^\ddagger[\%]$	$H_0$
5	1	8.763	9.904	0.885	161.448	51.95	Failed to reject
4	2	8.856	4.568	1.939	18.513	29.84	Failed to reject
3	3	90.174	2.998	30.077	10.128	1.19	Rejected
2	4	731.095	10.560	69.234	7.709	0.11	Rejected
1	5	21021.1	55.807	376.677	6.608	0.00	Rejected

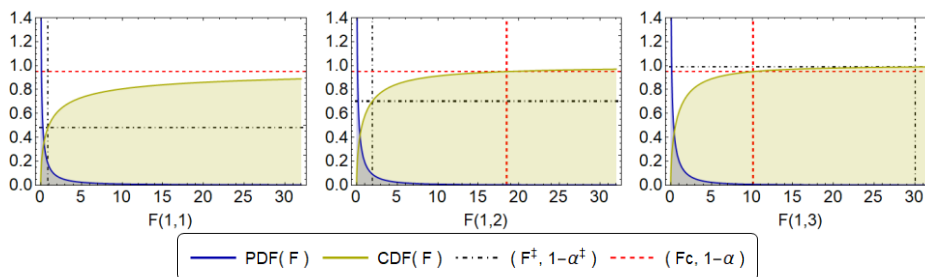


Figure 3 -  $F$ -distributions for several sizes of pools. The red dashed lines indicate  $F_c$  (x-axis) concerning  $1-\alpha = 95\%$  (y-axis). The black dot-dashed lines indicate the observed  $F$  (x-axis) and the respective  $1-\alpha^\ddagger$  (y-axis).

Figure 4 presents the PRESS-CV for the adopted cross-validation ( $k$ -fold) approach. In this plot, a flat behavior of the PRESS-CV curve can be observed between model sizes from 2 to 4, indicating that the predictive ability of the data is similar for models including 2 to 4 components. In a purely data-driven-based approach, 2 components would be selected as it allows maximizing the parsimony. However, only from the fourth dimension the PRESS-CV sharply increases, showing the presence of noise in the fifth component and consequent data overfitting. Although there is no significant improvement in the prediction with the addition of a 3<sup>rd</sup> and a 4<sup>th</sup> component to model structure, these are necessary for the complete description of the reaction system through a structure that represents the production/consumption of the entire set of observed species. Thus, from a fundamental perspective where the model needs to be able to describe the underlying phenomena, it is preferable to remain with 4 dimensions in the model, which can be seen as the most plausible model size.

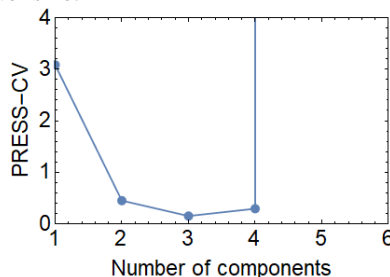


Figure 4 - Cross-validation results for 5 groups of validation datasets.

#### 4. Conclusions

The correct identification of model size from experimental data is important to accurately describe the dynamics of reaction systems. Several methods can be applied, without using any additional system information (like the assumed system kinetics or stoichiometry). The choice of the method must consider the noise, the number of measurements available, the subsequent model application, and the complexity of its structure. When large datasets are considered, the PRESS-CV method tends to perform well and shows a high likelihood of correctly capturing the “true” model size, since the chance of sampling bias is small in this case. The F-test seems preferable to heuristic-based methods, since the former has some theoretical support. When chemical reactions only occur in low extents (i.e., when the signals have variances of the same magnitude as the noise variance) the identification of the “real” size of the model becomes a hard problem. Within this setup, the use of non-heuristic-based methods is advised as they are grounded on theoretically consistent approaches. The results also highlight the advantage of combining different methods to determine the data characteristic space and the complementary space (i.e., the model size and the set of variables included in model representation, and the complementary set of variables). Given the possible synergies of the metrics, the combination of methods can increase the degree of confidence in the identification of the “true” dimensionality of the underlying system. In addition, the results obtained serve as preliminary guidelines for finding new model-based experiments for identifying the real size of the model. Table 2 presents the adequacy of the methods tested (here “+++” indicates a significant preference, “++” an average preference, and “+” a lower preference) and the basic characteristics of the data sets required for its application.

Table 2 - Preference of applicability of the selected methods and respective constraints.

Preference	Method advised	Requirements for successful application
+++	Cross-validation by eigenvectors	Large datasets
++	F-test	---
+	Heuristic approaches*	Low noise content

\*When the noise content is known the fractional variance analysis is preferred.

#### References

- D. Bonvin and D. Rippin, 1990, Target factor analysis for the identification of stoichiometric models. *Chemical Engineering Science* 45, 3417–3426.
- D. Bonvin, C. Georgakis, C. Pantelides, M. Barolo, M. Grover, D. Rodrigues, R. Schneider, D. Dochain, 2016, Linking models and experiments, *Industrial & Engineering Chemistry Research*, 55, 6891–6903.
- R. Bro, K. Kjeldahl, A. K. Smilde, H. Kiers, 2008, Cross-validation of component models: a critical look at current methods. *Analytical and bioanalytical chemistry* 390, 1241–1251
- R. Bro and A. K. Smilde, 2014, Principal component analysis, *Analytical Methods*, 6(9), 2812–2831.
- R. B. Cattell, 1966, The scree test for the number of factors, *Multivariate behavioral research*, 1(2), 245–276.
- N. Cliff, 1988, The eigenvalues-greater-than-one rule and the reliability of components, *Psychological Bulletin*, 103(2), 276.
- E. R. Malinowski, 1989, Statistical F-tests for abstract factor analysis and target testing, *Journal of Chemometrics*, 3(1), 49–60.
- C. S. Vertis, 2022, Identification and modeling of chemical reaction networks, Ph.D. thesis, Chemical Engineering Department of University of Coimbra.
- B. Wise, and N. Ricker, 1991, Recent advances in multivariate statistical process control: improving robustness and sensitivity, In *Proceedings of the IFAC, ADCHEM Symposium*, 125–130.

# Modeling of particle formation in pan granulators with sieve-mill recycle

Robert Dürr<sup>a</sup>, Mateusz Przywara<sup>b</sup>, Eric Otto<sup>c</sup>, Dorota Antos<sup>b</sup>, Achim Kienle<sup>c,d</sup>

<sup>a</sup>*Magdeburg-Stendal University of Applied Sciences, Breitscheidstraße 2, 39114 Magdeburg, Germany*

<sup>b</sup>*Rzeszow University of Technology, al. Powstańców Warszawy 12 35-959 Rzeszów, Poland*

<sup>c</sup>*Otto von Guericke University, Universitätsplatz 2, 39106 Magdeburg, Germany*

<sup>d</sup>*Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany*

## Abstract

Agglomeration of powders in pan granulators is a widespread production process, in many industries, e.g., agricultural and pharmaceutical. A continuous operation provides certain desired advantages over batch operation including constant throughput with constant properties of the product particles. An extended setup including a sieve mill recycle can improve the sustainability of overall process. As shown for related process, e.g., fluidized bed granulation, such a setup can exhibit interesting dynamic phenomena like sustained oscillation of the product particle properties, which must be researched in more detail to identify (un-)favorable process regimes. In this contribution, a model for fertilizer production in such a setup is presented. Simulation results indicate the potential benefits not only on increased product particle yield but also manipulation of particle properties.

**Keywords:** Population balance modeling, pan granulation, numerical solution, digital process twin, fertilizer manufacturing.

## 1. Introduction

Agglomeration is a particle formation process in which at least two particles are fused. This principle is used in many industries, e.g. fertilizer production. The properties of the formed agglomerates, e.g. size, shape and porosity, significantly affect certain end-use properties, e.g., dissolubility, processability and storeability (Bück & Tsotsas, 2016). In particular for fertilizer production and minerals processing, agglomerates are often formed in drums or pans (Litster & Ennis, 2004) in continuous operation which may provide constant throughput with constant product quality (in terms of specific agglomerate properties) during steady-state operation. Besides sophisticated experiments, mathematical modeling is also important to come to a deeper understanding and for potential development of model-based process control and intensification. In this report, focus is on population balance modelling of pan granulation which is also known under the term dish or disc granulation. The process setup is given as follows: a powder is fed constantly to an inclined rotating dish into which additional liquid binder is sprayed. Wet granules are sticking together and form larger ones. Flux from the pan does not only contain product particles but also fines and oversized particles. To improve sustainability of the process the non-product fractions can be recycled using a sieve-mill recycle. It is well-known that for size-enlargement processes like pan granulation individual

properties, like characteristic size and porosity, differ from particle to particle. As an alternative to Monte-Carlo modeling approaches, population balance modeling (PBM) (Ramkrishna & Singh, 2014) represents an established framework to model such processes. For standard monivariate PBEs, accounting mostly for characteristic particle size or volume, efficient and accurate numerical solution algorithms have been developed, see e.g. Kumar et al. (2008), Dürr & Bück (2020), and the references therein. In this contribution, a PBM for the continuous pan granulation process with sieve-mill recycle is presented based on a recently reported model for operation without recycle. The model is also extended for a mass controller to guarantee stable process operation. In different scenarios the advantage of recycling is emphasized and also the effects on the product particle characteristics in terms of size distribution is shown.

## 2. Process Model

### 2.1. Population Balance Model

Starting from the model for continuous pan agglomeration presented in Otto et al. (2021) the model equations must be adapted to account for the screen-mill-recycle. The volume density distribution function of the particles in the pan can be described with the following population balance model

$$\frac{\partial n(t,v)}{\partial t} = \dot{n}_{\text{feed}}(t,v) + \dot{n}_{\text{rec}}(t,v) - \dot{n}_{\text{out}}(t,v) + \dot{n}_{\text{agg}}(t,v)$$

Therein, the expressions on the left hand side represents the accumulation of particles in the pan while the right hand side describes contributions of the powder feed to the pan, the recycled particle stream, particles withdrawn from the pan and new particles generated by agglomeration. It is furthermore assumed that the process parameters are chosen such that particle formation via aggregation is dominant and thus additional effects from breakage of granules is negligible. In the following, all expressions on the right hand side are described in more detail.

### 2.2. Modeling of Powder Feed and Agglomeration

For the description of the powder feed to the pan the approach for the process without screen-mill-recycle presented in Otto et al. (2021) is used.

$$\dot{n}_{\text{feed}}(t,v) = N_{\text{in}}(t) \widetilde{n}_{\text{feed}}(v)$$

Here, it is assumed that the powder particle volume characteristics do not change over time and can be represented by the normalized feed powder number density function. Furthermore, denotes the time dependent feed rate.

The change in the particles NDF by formation of a new agglomerate of volume  $v$  by binary fusion of two primary particles of volumes  $u$  and  $v-u$ , respectively, is described with the following expression

$$\dot{n}_{\text{agg}}(t,v) = \frac{1}{2} \int_0^v \beta(t,u,v-u) n(t,u) n(t,v-u) \, du - \int_0^\infty \beta(t,u,v) n(t,v) n(t,u) \, du$$

Generally, the formation of new particles by agglomeration depends on multiple factors, e.g. temperature, binder composition, and particle volume. A common approach to describe such dependencies is a so-called separation approach for the agglomeration kernel function

$$\beta(t, v, u) = \beta_0(t)\beta_v(v, u)$$

with the first factor, the coalescence frequency, accounting for time-dependency due to process conditions or operating parameters and the second for volume dependency. The latter function can either be derived from micro-scale mechanistic modelling, heuristic assumptions or abstract data-driven approaches. In this manuscript, focus is on steady-state operation and thus constant process conditions can be assumed. In consequence,

$$\beta_0(t) = \beta_0 = \text{const.},$$

the so-called Kapur-Kernel is used

$$\beta_v(u, v) = \frac{(u+v)^a}{(u+v)^b}.$$

### 2.3. Modeling of Outflux Rate

The outflux rate comprises granules smaller than product range, granules in product range and also particles larger than the desired product range.

$$\dot{n}_{\text{out}} = \dot{n}_{\text{fines}} + \dot{n}_{\text{product}} + \dot{n}_{\text{over}}.$$

Assuming a sieve with two screens as presented in *Neugebauer et al. (2019)* the rates above can be modeled as

$$\begin{aligned} \dot{n}_{\text{over}} &= T_{\text{screen},1} \dot{n}_{\text{out}} \\ \dot{n}_{\text{product}} &= (1 - T_{\text{screen},1}) T_{\text{screen},2} \dot{n}_{\text{out}} \\ \dot{n}_{\text{fines}} &= (1 - T_{\text{screen},1})(1 - T_{\text{screen},2}) \dot{n}_{\text{out}}. \end{aligned}$$

With sieving functions given as

$$T_{\text{screen},i}(L) = \frac{\int_0^L \exp\left(-\frac{(\xi - L_{\text{screen},i})^2}{2\sigma_{\text{screen},i}^2}\right) d\xi}{\int_0^\infty \exp\left(-\frac{(\xi - L_{\text{screen},i})^2}{2\sigma_{\text{screen},i}^2}\right) d\xi}.$$

In contrast to the previous characterization of granules by volume, the sieving function is described in terms of characteristic particle size and thus reformulation under the assumption of spherical granule shape, i.e.,

$$v = \frac{\pi}{6} L^3$$

When then distribution function is reformulated it must be taken specific care for mass conservation in a subsequent computational implementation, i.e.,

$$n(t, v) dv = n_L(t, L) dL$$

### 2.4. Modeling of recycle rate

The recycle rate includes particles which are out of the desired product particle range with oversized particles being milled before being fed back into the pan

$$\dot{n}_{\text{rec}}(t, v) = \dot{n}_{\text{fines}}(t, v) + \dot{n}_{\text{mill}}(t, v)$$

Assuming an idealized mill, the milled particles follow a normal distribution (*Neugebauer et al., 2019*)



$$n_{L,mill}(t, L) = \frac{\int_0^L \exp\left(-\frac{(\xi - L_{mill})^2}{2\sigma_{mill}^2}\right) d\xi}{\int_0^\infty \exp\left(-\frac{(\xi - L_{mill})^2}{2\sigma_{mill}^2}\right) d\xi} n_{L,over}(t, L)$$

### 2.5. Powder feed manipulation for mass control

The extension of the process setup for the recycle requires control of the overall volume/mass of particles in the pan to guarantee a constant product stream and omit situations with too much or too few particles in the pan. Different control variables could be used to manipulate the overall mass of the pan content:

- Pan inclination,
- Angular speed of the pan,
- Recycle stream,
- Powder feed rate.

Manipulation of the first two variables are favored for the manipulation of the granule properties and for improved sustainability the goal is to use the full recycle stream in the process. Thus, manipulation of the powder feed rate is supposed to be the best option to obtain a stable mass/volume of the pan content. The change of particle mass in the pan is given as the sum of all the in- and outgoing particle fluxes

$$m_{pan} \dot{=} m_{feed} + m_{rec} - m_{out}$$

One pragmatic choice of control would be to use a two-point controller for manipulation of the feed stream. However, the mass of particles in the pan can not be assessed easily and thus an alternative strategy is proposed: The process is run without recycle at startup until steady-state conditions are obtained for a given mass stream of feed-particles  $m_{feed,nom}$ . Afterwards the recycle is added and the feed is manipulated as follows

$$m_{feed} \dot{=} m_{out} - m_{rec}$$

## 3. Results

In the following, simulation scenarios for the setup with recycle are summarized. Feed particle size distribution, separation function and agglomeration kernel parameters are the same as in the non-recycle setup presented in *Otto et al. (2021)*. All additional parameters are summarized in Table 1. Simulations were carried out in MatLab2021a. A detailed description of the simulation algorithm is found in *Otto et al. (2021)* and the references therein. In Fig. 1 – 4 results are shown for four different scenarios: operation without recycle (I), with only milled particles recycled (II), only fines recycled (III) and both recycled (IV).

It can be observed, that recycling of particles which are out of the product specification is favorable in terms increased product particle mass flow. However, the recycling rate also has a clear effect on the characteristic particle size distribution of the product fraction.

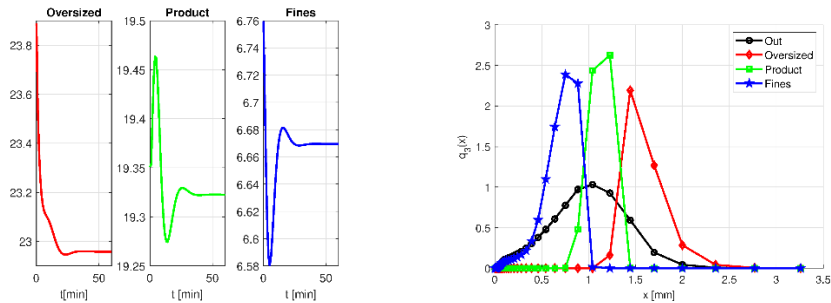


Figure 1: Scenario (I) without recycle stream: mass flows of different fractions (left) and normalized particle size distribution (right)

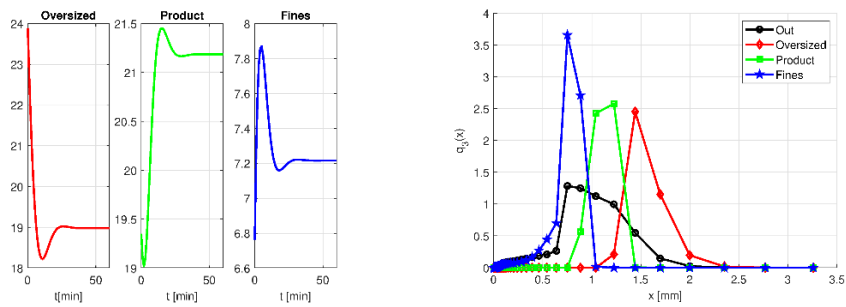


Figure 2: Scenario (II) recycle of milled-particles: mass flows of different fractions (left) and normalized particle size distribution (right)

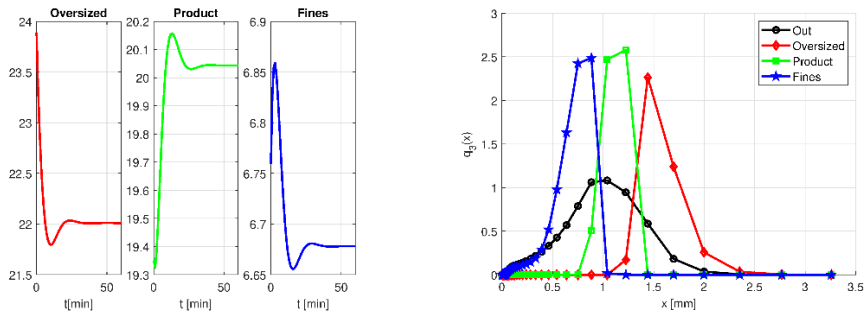


Figure 3: Scenario (III) recycle of undersized particles: mass flows of different fractions (left) and normalized particle size distribution (right)

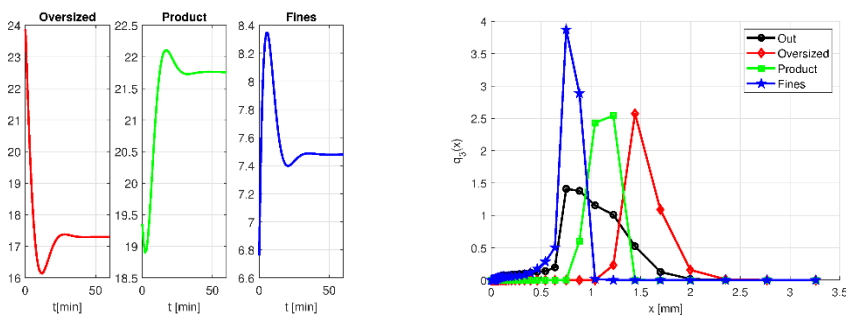


Figure 4: Scenario (IV) recycle of milled and undersized particles: mass flows of different fractions (left) and normalized particle size distribution (right)

#### 4. Summary and Outlook

In this contribution a population balance model for continuous pan granulation with sieve-mill-recycle was presented. Here, it must be mentioned that a mass controller is necessary to guarantee a stable process operation. Simulation results indicate that increased product mass flows are obtained yet also the product particle size distribution and thereby the product properties are affected. Therefore, further research is necessary on this issue as well as identification of (un-)favorable process regimes and interesting dynamic phenomena, such as sustained oscillations, e.g., by bifurcation analysis (Otto et al., 2022). Additionally, particle breakage could be taken into account requiring the adaption of the presented concepts. Moreover, the model may be used for the design of model-based sensors or controllers. Of course the previous mentioned research perspectives require model adaption to experimental data (Otto et al., 2021b; Przywara et al., 2021)

#### References

- J. Litster, B. Ennis. *The Science and Engineering of Granulation Processes*, 2004, Vol. 15 Springer Science & Business Media
- C. Neugebauer, E. Diez, A. Bück, S. Palis, S. Heinrich, A. Kienle. On the dynamics and control of continuous fluidized bed layering granulation with screen-mill cycle. 2019, *Powder Technology*, Vol. 354, p. 765 – 778.
- E. Otto, R. Dürr, M. Przywara, D. Antos, and A. Kienle. Population balance modelling of pan granulation processes. 2021, *Proceedings to the 31<sup>st</sup> European Symposium on Computer Aided Process Engineering*, p. 965-970
- D. Ramkrishna and M. R. Singh. Population balance modeling: Current status and future prospects. *Annual Review of Chemical and Biomolecular Engineering*, 5:123–146, 2014.
- J. Kumar, M. Peglow, G. Warnecke, and S. Heinrich. An efficient numerical technique for solving population balance equation involving aggregation, breakage, growth and nucleation. *Powder Technology*, 182(1):81 – 104, 2008.
- A. Bück and E. Tsotsas. Agglomeration. In Benjamin Caballero, Paul M. Finglas, and Fidel Toldrá, editors, *Encyclopedia of Food and Health*, pages 73 – 81. Academic Press, Oxford, 2016.
- R. Dürr and A. Bück. Approximate Moment Methods for Population Balance Equations in Particulate and Bioengineering Processes, *Processes*, 8(4), 414, 2020.
- E. Otto, R. Dürr, A. Kienle. Bifurcation analysis of combined agglomeration and layering granulation in fluidized bed spray processes, *Computer Aided Chemical Engineering* 51, 691-696, 2022.
- M. Przywara, R. Dürr, E. Otto, A. Kienle, D. Antos. Process behavior and product quality in fertilizer manufacturing using continuous hopper transfer pan granulation—Experimental investigations, *Processes* 9 (8), 1439, 2021
- E. Otto, R. Dürr, G. Strenzke, S. Palis, A. Bück, E. Tsotsas, A. Kienle. Kernel identification in continuous fluidized bed spray agglomeration from steady state data, *Advanced Powder Technology* 32 (7), 2517-2529, 2021b

# Synthesis and optimization of NGL separation as a complex energy-integrated distillation sequence

Qing Li,<sup>a,b</sup> Adrian J. Finn,<sup>c</sup> Stephen J. Doyle,<sup>b</sup> Robin Smith,<sup>b</sup> Anton A. Kiss<sup>a\*</sup>

<sup>a</sup> *Department of Chemical Engineering, Delft University of Technology, Van der Maasweg 9, 2629 HZ, Delft, The Netherlands*

<sup>b</sup> *Centre for Process Integration, Department of Chemical Engineering, The University of Manchester, Sackville Street, Manchester, M13 9PL, United Kingdom*

<sup>c</sup> *Costain, Costain House, 1500 Aviator Way, Manchester Business Park, Manchester, M22 5TG, United Kingdom*

*E-mail: [tonykiss@gmail.com](mailto:tonykiss@gmail.com)*

## Abstract

The synthesis of heat-integrated distillation sequences for energy-efficient separation of zeotropic multicomponent mixtures is complex due to the many interconnected design degrees of freedom. This paper explores the basis on which reliable screening can be carried out. To solve this problem, a screening algorithm has been developed using optimization of a superstructure for the sequence synthesis using shortcut models, in conjunction with a transportation algorithm for the synthesis of the heat integration arrangement. Different approaches for the inclusion of heat integration are explored and compared. Then the best few designs from this screening are evaluated using rigorous simulations. A case study for the separation of NGL is used to compare options. It has been found that separation problems of the type explored can be screened reliably using shortcut distillation models in conjunction with the synthesis of heat exchanger network designs. Unintegrated designs using thermally coupled complex columns show much better performance than the corresponding designs using simple columns. However, once heat integration is included the difference between designs using complex columns and simple columns narrows significantly.

**Keywords:** Distillation sequencing; energy efficiency; process synthesis and design; process optimization

## 1. Introduction

Distillation is by far the most widely applied separation technology but is recognized as the most energy intensive operation in the chemical industries due to its inevitable energy degradation. Therefore, to genuinely enhance the energy efficiency of the distillation system, it is necessary that all the degrees of freedom are manipulated simultaneously (Kiss and Smith, 2020), i.e. the basic separation configuration including thermally coupled complex columns, column pressures, reflux ratios, feed conditions, condenser types and heat integration arrangement. For all but the simplest separation problems an exhaustive search is infeasible. The method most often used to address larger problems is to use short-cut distillation models to screen the large number of structural options and determine the best few potential designs, followed by more detailed examination of the best few designs using rigorous simulation. However, this raises many questions regarding the screening procedure. It is not clear whether short cut distillation models have the necessary accuracy for reliable screening, and whether

heat integration must be included in the screening. Moreover, if heat integration is included, it is not clear whether the integration can be represented by the energy targets of pinch analysis, or it is necessary to use detailed heat exchanger network designs for screening. This work addresses the problem of multicomponent separation by distillation sequences, by developing an integral systematic optimization method which simultaneously synthesizes sequences, optimizes simple and complex distillation arrangements, heat integration, operating conditions, as well as closed cycle heat pump for refrigeration systems.

The fractionation of natural gas liquids (NGL) is presented here to demonstrate the developed methodology. NGL fractionation, which demands a sequence of several distillation columns, is one of the most energy-demanding processes in the oil and gas industry, producing marketable products such as ethane, propane, butanes, and heavier hydrocarbons. NGL feed is typically 2.5 million tons per year, and approximately 77 GJ per barrel of NGL feed is required, which is equivalent to 4.65 tons of CO<sub>2</sub> emissions per barrel (Manley, 1998). Several attempts based on process intensification principles have been made on the NGL fractionation process to generate more energy efficient distillation sequence design with better separation performance and lower energy use. However, these previous studies made simplifications to reduce model complexities to achieve the challenging optimization: only simple columns are taken into account (Yoo et al., 2016), column pressures are fixed, relative volatilities of the initial feed stream are used for all sequence columns (Ramapriya et al., 2018), or assume only saturated liquid or vapor-phase feed and product streams (Tamuzi et al., 2020). These simplified assumptions on key operating variables lead to loss of accuracy, and none of the previous works considered all degrees of freedom, which risks missing the promising solutions.

## 2. Problem statement

The optimization of the NGL fractionation system is motivated by the benefits of energy savings in reducing costs and CO<sub>2</sub> emissions. Its design requires many interconnected design degrees of freedom, as well as column arrangement possibilities (simple column or complex column arrangements), which lead to a challenging optimization problem. The present study solves this problem through a novel and systematic optimization approach to evaluate all the complex energy integrated NGL fractionation sequence alternatives simultaneously, eliminating simplifications on key design variables. The best sequences can be ranked based on techno-economic analysis.

## 3. Methodology and optimization approach

For a given multicomponent mixture, a superstructure is initially created that contains all possible sequences using simple columns, and then the superstructure is extended to merge any two simple columns in series to form different complex column arrangements, including side draw columns, side stripper, side rectifier, prefractionator, Petlyuk and dividing wall column, as illustrated in Figure 1 (a) (Smith, 2016). Therefore, based on the same starting defaults of each simple column sequence, the modified superstructure generates complex column sequences accordingly. Shortcut Fenske-Underwood-Gilliland models are used to design the columns in this task, which can screen the very many options and avoid dramatic computational difficulties created by nonlinear, large tray-based combinations of rigorous design formulation. The sequences screened by shortcut models are then evaluated by Aspen Plus rigorous simulations.

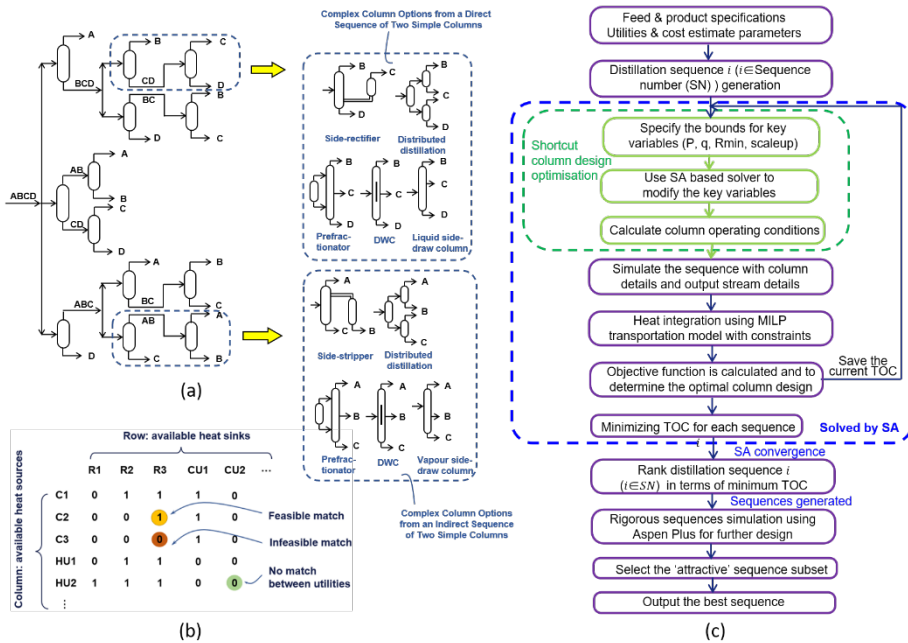


Figure 1 Extended superstructure for sequence synthesis (a), heat integration matrix (b) and the overall methodology flowchart (c)

Heat integration opportunities including feed preheating or precooling and column to column heat recovery are optimized with sequence synthesis and sequence design simultaneously. All heat sinks and heat sources are compiled into an incidence matrix (as shown in Figure 1 (b)) and solved by the MILP transportation method. Binary variables are introduced to make discrete decisions for each match between rows and columns of the matrix. The optimal heat exchanger network structure to minimize operating costs is determined. Liquids and gases are pressurized by pumps and multistage-compressors and depressurized by valves and turbo expanders. The separation is assumed to be sharp split.

A decomposed optimization methodology is created and illustrated by a flowchart shown in Figure 1 (c). The overall model is formed as an MINLP problem and solved by simulated annealing (SA). The shortcut distillation models equations are built into the optimizer, acting as a ‘simulation’ section to eliminate the non-linearity in the optimization. For a given mixture, the SA creates ‘moves’ in the optimization that allow for changes in the distillation structure, for example, changing two adjacent simple columns to be a complex column, and the operating variables in their permitted bounds. The values are then sent to the objective function which uses them as inputs for the simulation model. After each move, the sequences are simulated, heat exchanger network designed, and the corresponding total operating cost (TOC) is evaluated by the current column design details and then by the heat integration. The objective value is then returned to the optimizer. The operating variables can subsequently be changed by the optimizer and the process repeated until convergence. This optimization methodology can be carried out within the ColSeq software developed and available at the Centre for Process Integration, The University of Manchester (CPI Suite, 2022).

## 4. Results and discussion

The conventional configuration of the NGL fractionation uses a train of 4 simple distillation columns: deethanizer (30.98 bar), depropanizer (17.50 bar), debutanizer (5 bar) and deisobutanizer (6.50 bar) to generate 5 products (Manley, 1998; Long et al., 2018). In this work, the NGL fractionation is optimized by applying the developed optimization model to enhance its energy benefits. Multiple utilities considered for heating, cooling and refrigeration and their prices are taken from Turton et al., (2018). The corresponding details of feed and products are taken from Long et al., (2016). For heat exchangers, the minimum temperature difference between hot streams and cold streams (including utilities) is assumed to be 10 K for above ambient heat transfer, whereas for sub-ambient it is assumed to be 4 K (Kiss and Smith, 2020). The optimization is carried out using a desktop with CPU - *Intel<sup>(R)</sup> Core i7-10700 CPU 2.90 GHz* with 8 cores, and the average time resource is 16 hours.

By optimization, a total of 14 simple sequences using 4 simple columns, and the corresponding 14 complex column sequences are generated. These sequences are developed by two optimization routes for the NGL fractionation: sequences without heat integration and integrated sequences with detailed heat exchanger network design.

### 4.1. Optimization results evaluation and analysis

Firstly, the developed approach is validated demonstrating by non-heat integrated simple column sequences. Figure 2 (a) shows sequence rankings in terms of the total operating cost. The ranking top 5 sequences (Zone-I) shown in Figure 2, with relatively small differences in TOC (less than 5%), are considered as the most economically attractive sequence family identified by the screening method based on shortcut models. To evaluate the optimized sequences based on shortcut models, the short-cut design details (reflux ratio, trays, condenser temperatures and distillate rate) are extracted as the initialization parameters, and then all optimized sequences are rigorously simulated using Aspen Plus rigorous distillation models. The sequences S7, S6, S4 are the top ranked sequences identified by rigorous simulations in terms of total operating costs, which have all been preselected in the attractive families by the developed pre-screening approach. These results indicate that despite the deviation between shortcut and rigorous models, the developed fast screening method using shortcut models is validated and comparable. Compared to the conventional distillation sequence (Manley, 1998), the optimized operating conditions based on the same separation structures lead to lower column pressure for all four columns without changing the refrigerant level of the deethanizer, and lower condenser and reboiler duties with 2.37 M\$/yr TOC savings.

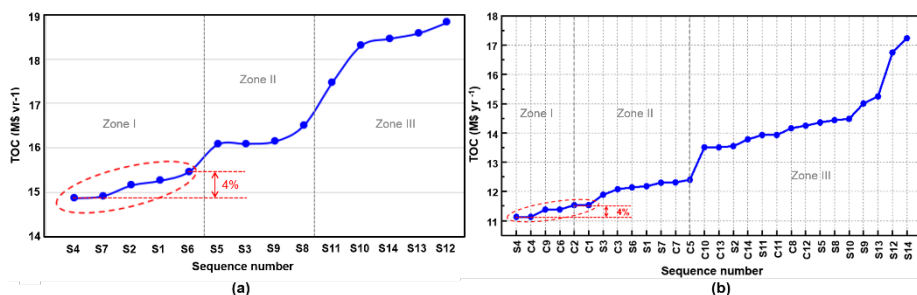


Figure 2 Total operating costs ranking of non-heat integrated simple column sequences (a) and heat integrated sequences (b)

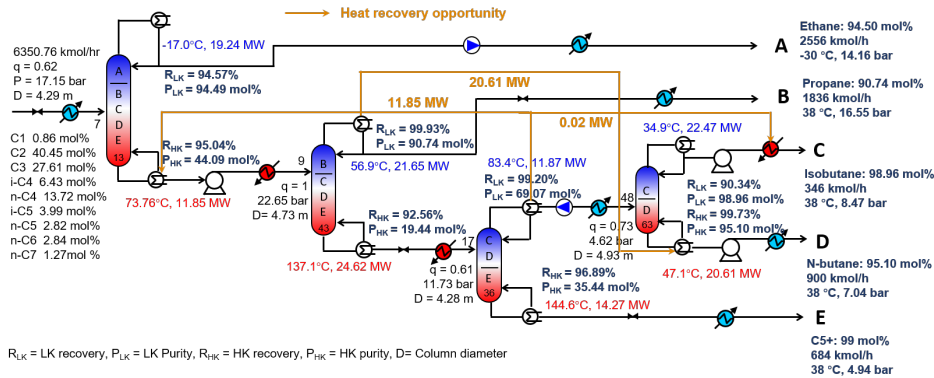


Figure 3. Design details of heat integrated simple column sequence S4

#### 4.2. Energy integrated distillation sequences

Energy intensified sequences using complex columns can bring energy savings. For the non-heat integrated sequences, introducing complex columns can achieve up to 21% operating cost savings compared with the simple column sequences. The conventional sequence (which is the same structure as the simple sequence S4 in this case) that is ranked as the most economical sequence in a simple column sequence, is ranked 11 out of 28 taking into account complex column sequences. The best sequence is now the complex sequence 6, which separates isobutane by a side rectifier, followed by two simple columns separating C2 / C3 and n-butane / C5+, with the TOC of 12.78 M\$/yr.

To achieve further energy savings, heat integration is adopted, achieved by optimizing heat exchanger network with column operating conditions simultaneously for both simple and complex sequences. Figure 2 (b) shows the overall rank of heat integrated simple and complex column sequences. The total operating cost saving brought by heat integration on the optimized non-heat integrated simple columns is up to 26%, and the operating cost saving from optimizing both heat integration and complex columns can achieve up to 38%. The top 6 sequences are considered as the most economically attractive sequences family with a relatively small operating cost difference. Further designs can be taken within this attractive families based on different requirements.

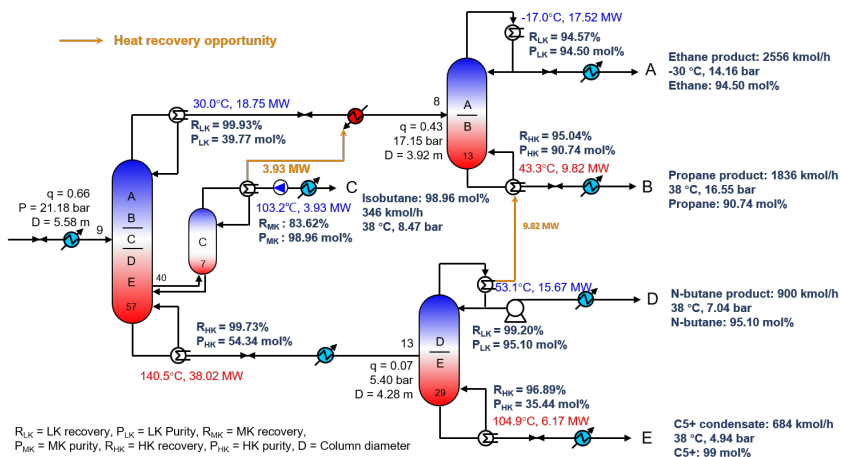


Figure 4. Design details for heat integrated complex column sequence C9



It is shown that when heat integration is adopted, the simple sequence 4 is ranked the best (with the detailed design in Figure 3), followed by the complex sequence C9 (as shown in Figure 4). Simple column sequences benefit to a larger extent from heat integration compared to complex column sequences; simple column sequences can achieve an average energy saving of 16% by introducing heat integration, but complex columns gain an average 11% energy saving by introducing heat integration. This is because complex columns could lose potential heat integration possibilities when constraining the combination of heat sinks and sources during the merging, so that the number of design degrees of freedom of complex columns decrease. The complex sequence C9 may get more benefits when taking into account capital costs due to the reduced equipment sizes. The capital cost is not considered at this stage, as capital costs may not be reflected accurately using approximate capital cost calculation methods.

## 5. Conclusions

The systematic screening approach developed in this study was successfully used for simultaneously optimizing sequence structures, accounting for all design degrees of freedom and heat integration in the distillation sequences. By optimizing the whole sequencing process, new NGL fractionation sequences have been explored and ranked techno-economically. The generated economically attractive sequences (ranked by the operating costs) are shown to be comparable with those of rigorous simulations using the Aspen Plus RadFrac model. For a given feed composition and product requirements, the ranking of the detail designed energy integrated complex sequences for NGL fractionation are listed, and the most energy efficient and economically attractive sequences family have been selected. These results demonstrate that screening of the search space of energy-efficient distillation sequences, considering synthesis and optimization of the many design degrees of freedom, is achievable with the newly proposed approach.

## References

- CPI Suite, 2022, The University of Manchester, Centre for Process Integration, Online: <https://www.ce.manchester.ac.uk/cpi/research/resources/software/>
- A.A. Kiss, R. Smith, 2020, Rethinking energy use in distillation processes for a more sustainable chemical industry, *Energy*, 203, 117788
- N.V.D. Long, L.Q. Minh, T.N. Pham, A. Bahadori, M. Lee, 2016, Novel retrofit designs using a modified coordinate descent methodology for improving energy efficiency of natural gas liquid fractionation process. *Journal of Natural Gas Science and Engineering*, 33, 458e68.
- D. Manley, 1998, Thermodynamically efficient distillation: NGL fractionation, *Latin American Applied Research*, 28, 211-216
- G.M. Ramapriya, A. Selvarajah, L.E. Jimenez Cucaita, J. Huff, M. Tawarmalani, R. Agrawal, 2018, Short-cut methods versus rigorous methods for performance-evaluation of distillation configurations, *Industrial & Engineering Chemistry Research*, 57, 7726-7731
- R. Smith, 2016, *Chemical process design and integration*, Chichester, West Sussex, United Kingdom, John Wiley & Sons, Inc
- A. Tamuzi, N. Kasiri, A. Khalili-Garakani, 2020, Design and optimization of distillation column sequencing for NGL fractionation processes, *Journal of Natural Gas Science and Engineering*, 76, 103180
- R. Turton, R.C. Bailie, W.B. Whiting, J.A. Shaeiwitz, 2008, *Analysis, synthesis and design of chemical processes*, Pearson Education
- H. Yoo, M. Binns, M.G. Jang, H. Cho, J.K. Kim, 2016, A design procedure for heat-integrated distillation column sequencing of natural gas liquid fractionation processes, *Korean Journal of Chemical Engineering*, 33, 2, 405-415

# Multi-scale model of solid oxide fuel cell: enabling microscopic solvers to predict physical features over a macroscopic domain

Hamid Reza Abbasi,<sup>a</sup> Masoud Babaei,<sup>a</sup> Arash Rabbani,<sup>b</sup>

Constantinos Theodoropoulos,<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, The University of Manchester, Manchester, M13 9PL, United Kingdom*

<sup>b</sup>*School of Computing, The University of Leeds, UK*  
*k.theodoropoulos@manchester.ac.uk*

## Abstract

In this study, a multi-scale model is developed to address the transport of gases, ions, and electrons inside a solid oxide fuel cell. Transport phenomena and electrochemical reactions inside the active layer of anode are simulated at the nanometer scale, while the rest of cell is simulated macroscopically at a coarser resolution. The microstructure of the anode side is reconstructed using pluri-gaussian random fields and the triple phase boundary (TPB), where all the electrochemical reactions are assumed to take place, is captured from the discretized (in voxels) image of the microstructure. Therefore, this multi-scale framework offers a much finer level of description for transport phenomena and for electrochemical reactions, significantly increasing the accuracy of numerical modelling of entire solid oxide fuel cells.

**Keywords:** Solid Oxide Fuel Cell, Multi-scale model, Triple phase boundary

## 1. Introduction

Fuel cells, as an alternative power generation device, have shown tremendous potential in clean, small- and large-scale electricity generation in recent decades. Fuel cells convert the chemical energy of a fuel (typically hydrogen) into electricity through electrochemical reactions. The energy conversion (chemical to electrical) is not limited by any thermodynamic law. Theoretically they can have 100% efficiency, whereas the efficiency of e.g., coal-powered stations is limited by the first and second law of thermodynamics and the maximum efficiency attainable is the Carnot efficiency. If hydrogen is to be utilized inside a cell as fuel, the only by-products of the electrochemical reactions would be oxygen gas and water.

Solid oxide fuel cells as the name suggests use a solid ceramic compound as the electrolyte to generate clean electricity (Singh, Zappa, and Comini 2021). SOFCs are well known for their high operating temperatures, which eliminate the need for metallic catalysts. Due to the convoluted interplay between different transport phenomena and chemical/ electrochemical reactions inside the cell, which span different length scales (Grew and Chiu 2012), complex models are needed to capture their behaviour. The

electrode microstructure has a substantial effect on the overall electrochemical performance of the cell (dos Santos-Gómez et al. 2021). However, the size of an entire fuel cell is far larger than the microscopic details of the electrodes to be considered in a macroscopic model. It is not, therefore, practically possible to numerically solve the mathematical equations governing the SOFC behaviour over an entire cell while maintaining a fine degree of description of microscopic details. This issue is thoroughly discussed by Grew and Chiu (Grew and Chiu 2012). Hence, most relevant works in the literature are focused either on macroscopic descriptions [4], component-level modelling [5], or on pore-scale modelling [6] of SOFC electrodes. These classes of models make multiple assumptions about the unresolved physics of the “other” scale. To bridge the gap between these two classes of numerical models, an integrated multi-scale numerical framework has been developed in this work by enabling a macroscopic model to obtain information directly from a pore-level model.

The macroscopic model uses the finite-element approach, and its domain covers the entire cell (fuel/air channels, electrolyte, and electrodes, and interconnects). In addition, a microscopic model is developed to simultaneously solve the coupled mass, ion, and electron transport equations, and chemical/electrochemical reactions in a three-dimensional microstructure of the SOFC electrode. Microscopic simulations are only performed on small segments (patches) of the microstructure. A “patch” [7] is a small spatial computational domain extracted from a larger volume/surface of interest, in order to facilitate the efficiency of complex computations. The contribution of this work is that the developed integrated, multi-scale model can resolve the physics linked to the fine microscopic topological features of the electrode microstructure over an entire length of an electrode in a particular fuel cell. This can improve the overall performance prediction of SOFCs, which is highly dependent on both their microscopic and macroscopic features.

## 2. Methodology

In this section, the two different models that are developed to address the multiscale problem are briefly discussed. The micro model solves the gas/electron/ion transport inside the active electrochemical layer. Using Fick’s and Ohm’s laws, mass/charge conservation equations can be written as follows. Knudsen diffusivity should also be considered due to small radius of pores. It can be approximated using kinetic theory of gases.

$$\nabla(-D_{H_2}\nabla c_{H_2}) = -\frac{\rho_{TPB}J}{2F}M_{H_2} \quad (1)$$

$$\nabla(-\sigma_{el}\nabla V_{el}) = -\rho_{TPB}J \quad (2)$$

$$\nabla(-\sigma_{io}\nabla V_{io}) = \rho_{TPB}J \quad (3)$$

Here,  $c_{H_2}$ ,  $V_{el}$ ,  $V_{io}$  are the hydrogen concentration, electronic potential, and ionic potential, respectively.  $\rho_{TPB}$  is the density of the triple phase boundary and  $J$  is charge transfer rate. This rate is only non-zero at the TPB. Therefore, all three governing equations (1-3) reduce to the Laplace equation anywhere else in the domain. For the microscopic model, these equations are solved over a representative volume element of a microscopic domain with regular hexahedral mesh. Each grid element represents a pixel in the tomographic

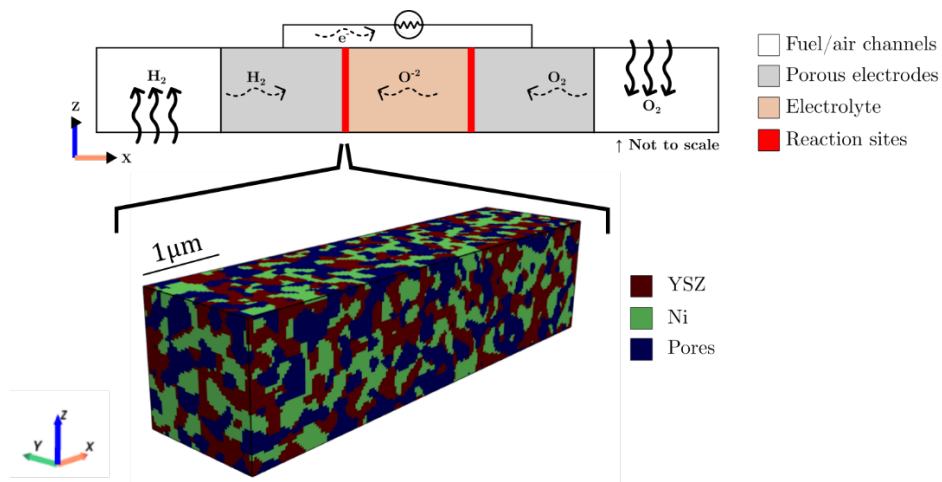
image obtained from 3D imaging. Finite-volume method is employed to discretize equations in 3D space. Boundary conditions for micromodel consist of constant ion and electron potential at microstructure surfaces facing electrolyte and anode, respectively. Hydrogen concentration is also assumed constant at the boundary facing anode side. All other boundary conditions are zero-flux. Three coupled systems of linear equations resulting from discretizing equations 1-3 over the entire microscale domain are then solved until convergence using GMRES algorithm implemented in python.

The macroscopic model solves the mass and charge conservation equations inside an entire cell. The governing equations are the same as mentioned above. Since electrochemical reactions take place on the TPB in microscale, they cannot be resolved with the macroscale model. Instead, these reactions are modeled in the microscale model and then volume-averaged reaction rates are imported into macroscopic model as boundary conditions on the interfaces between electrolyte and porous electrodes. The charge transfer rate is assumed to follow the Butler-Volmer model.

$$J = J_0 \rho_{TPB} \left( \exp\left(\frac{2\beta F}{RT} \eta_{act}\right) - \exp\left(-\frac{2(1-\beta)F}{RT} \eta_{act}\right) \right) \quad (4)$$

Where  $J_0$  [A/m] is the lineal exchange current density taken from data provided by Prokop et al. [8].

A schematic of the macroscopic and the microscopic model and their interactions is shown in Figure 1. The macroscopic domain contains fuel/air channels, porous electrodes, and electrolyte. The microscopic model, on the other hand, is limited to the active electrochemical site between the porous anode and the electrolyte. The microstructure is generated using pluri-gaussian random fields described in [9], and the location of TPB inside the microstructure is calculated with the algorithm discussed by Vivet et al. in [10].



**Figure 1** – Schematic of the multi-scale numerical framework. The macroscopic model spans an entire cell, whereas the microscopic model only covers the electrochemical active layer between electrolyte and electrodes

The transport equations are solved until convergence both in macro and in micro models. The information that is shared between two models are the physical features of the porous media (porosity, permeability), the effective current density exiting the electrolyte layer, operating voltage, and amount of overpotentials.

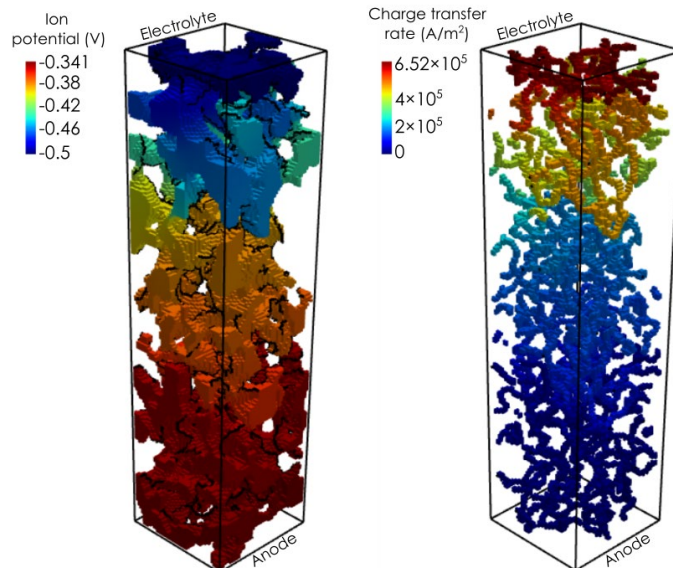
Geometrical parameters of the cell and physical properties of Nickel and YSZ are taken from papers by Tseronis et al. [4], and Shearing et al. (Shearing et al. 2010).

The numerical mesh that is used for the microstructure model is essentially the same as that of the voxelized image. Since the voxelized image is the finest description available of the real porous geometry, it has been decided that no smoothing should be performed on the reconstructed image. If the reconstructed geometry is to undergo smoothing, a finer numerical mesh can be constructed, which however would not be useful as its features would have been finer than the finest level of microstructure description available.

### 3. Results and discussion

In this section, the results of the multi-scale model combining the macro- and micro-scale models are presented and discussed. For the simulations, the operating temperature of the cell is assumed to be 900 °C. The density of the triple phase boundary is measured to be  $9.456 \mu\text{m}^2$ . The variation of the ion potential and current density in the microstructure electrochemical active layer of the anode side is shown in Figure 2. The YSZ phase inside the microstructure is on the left-hand side of the figure with the TPB lines depicted on the outer surface of the YSZ phase as black lines. On the right-hand side of the figure, the variation of the current density is shown on the TPB lines. To visualize the TPB lines, it was assumed that they have an artificial thickness of two pixels on each perpendicular direction. The top interface of the microstructure in this figure is adjacent to the electrolyte and the bottom side faces towards the porous anode layer. Due to the electrochemical reaction taking place on the TPB sites, the ion potential increases from the electrolyte interface (-0.341 V) to the porous anode interface (-0.5 V). The highest charge transfer rate is on the electrolyte interface, and it gradually reduces to zero as the distance from the electrolyte increases.

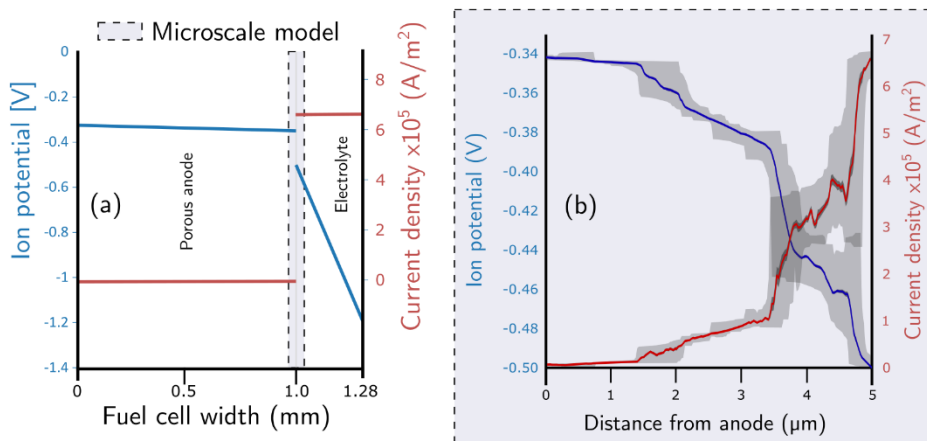
Changes in hydrogen concentration and electronic potential in the microstructure were insignificant, so they were not shown here. The micromodel is also capable of computing the local activation and concentration overpotential in the microstructure.



**Figure 2** – spatial variation of ionic potential and current density, respectively in the YSZ phase and on the TPB.

Results of the multi-scale model are illustrated in Figure 3 below. Figure 3a depicts results from the macroscopic model, which the mass and charge transport through the entire cell, whereas Figure 3b depicts the results of the microscopic model. As mentioned above, the microscopic model solves the same transport equations as of the macroscopic model, with the exception that the micro-scale model also computes the electrochemical reactions on the TPB sites. As shown below, this enables the micro-scale model to capture the rate of charge transfer locally with nanometer resolution. To demonstrate the capability of the multi-scale model, changes of ion potential and current density in the electrolyte and the porous anode are shown in figure 3a, whereas Figure 3b depicts the local variation of the same variables on the interface between electrolyte and porous anode, with much finer (nanoscale) resolution. For comparison, the size of each computational node in the macro-model is roughly 0.1 mm, whereas each computational node of the micro-model is only 25 nm. In Figure 3b, the minimum and maximum value, and the 95% confidence band of ion potential and current density in each slice of the microstructure in x direction (refer to Figure 1) is also shown.

The microscale model covers the electrochemical active layer only. Therefore, the rate of electrochemical reactions (or consequently current density, red curves in Figure 3) approaches zero when the distance from the electrolyte increases in the micromodel. The maximum current density on the electrolyte surface is then transferred to the macroscale model to simulate large-scale mass/charge transport. Potential difference across the micromodel (blue curves in Figure 3) implies the sum of activation and concentration overpotentials in the anode side. This potential difference is forced in the microscale model by macroscale simulation. As a result, until the mass and charge conservation criteria between the micro and macro domains are satisfied, there is an iterative loop between the micro and macro models. This link between micro and macro models allows macroscopic description of mass/charge transport inside a single cell while maintaining a microscopic resolution to explain electrochemical reactions at TPBs. It ultimately creates a framework to investigate the effects of microscopic features on macroscopic performance of the cell in our future studies.



**Figure 3** – Variation of ion potential and current density across the width of the fuel cell both in macro- and micro-models with different numerical resolutions. (a) macroscale model. (b) microscale model

#### 4. Conclusions

A multi-scale numerical framework has been developed to study the mass/ion/electron transport inside the entire solid oxide fuel cell, from fuel channel to air channel with the capability of resolving the nanometer scale features inside the active electrochemical layer. This can help to achieve a more accurate description of electrochemical reactions taking place on the TPB and also on the double phase boundaries in the case of mixed ionic-electronic conductors, which ultimately increases the accuracy of the numerical model for predicting the performance of the entire cell or even an entire stack of cells. The future research direction is to extend this multi-scale numerical framework to use multiple microstructures in parallel (patches) to capture the nanometric features and rate of electrochemical reactions across the length of the entire fuel cell, and even across multiple cells, significantly expanding the computational capabilities of the multi-scale model.

#### References

- Bertei, A., V. Yufit, F. Tariq, and N.P. Brandon. 2018. 'A Novel Approach for the Quantification of Inhomogeneous 3D Current Distribution in Fuel Cell Electrodes'. *Journal of Power Sources* 396 (August): 246–56. <https://doi.org/10.1016/j.jpowsour.2018.06.029>.
- Grew, Kyle N., and Wilson K.S. Chiu. 2012. 'A Review of Modeling and Simulation Techniques across the Length Scales for the Solid Oxide Fuel Cell'. *Journal of Power Sources* 199 (February): 1–13. <https://doi.org/10.1016/j.jpowsour.2011.10.010>.
- Kevrekidis, Ioannis G, C William Gear, James M Hyman, Panagiotis G Kevrekidis, Olof Runborg, Constantin Theodoropoulos, and others. 2003. 'Equation-Free, Coarse-Grained Multiscale Computation: Enabling Microscopic Simulators to Perform System-Level Analysis'. *Commun. Math. Sci* 1 (4): 715–62.
- Moussaoui, H., J. Laurencin, Y. Gavet, G. Delette, M. Hubert, P. Cloetens, T. Le Bihan, and J. Debayle. 2018. 'Stochastic Geometrical Modeling of Solid Oxide Cells Electrodes Validated on 3D Reconstructions'. *Computational Materials Science* 143 (February): 262–76. <https://doi.org/10.1016/j.commatsci.2017.11.015>.
- Prokop, Tomasz A, Janusz S Szmyd, and Grzegorz Brus. 2018. 'Developing Micro-Scale Heterogeneous Numerical Simulation of a Solid Oxide Fuel Cell Anode'. *Journal of Physics: Conference Series* 1101 (October): 012027. <https://doi.org/10.1088/1742-6596/1101/1/012027>.
- Santos-Gómez, Lucía dos, Javier Zamudio-García, José M Porras-Vázquez, Enrique R Losilla, and David Marrero-López. 2021. 'Recent Progress in Nanostructured Electrodes for Solid Oxide Fuel Cells Deposited by Spray Pyrolysis'. *Journal of Power Sources* 507: 230277.
- Shearing, P.R., Q. Cai, J.I. Golbert, V. Yufit, C.S. Adjiman, and N.P. Brandon. 2010. 'Microstructural Analysis of a Solid Oxide Fuel Cell Anode Using Focused Ion Beam Techniques Coupled with Electrochemical Simulation'. *Journal of Power Sources* 195 (15): 4804–10. <https://doi.org/10.1016/j.jpowsour.2010.02.047>.
- Singh, Mandeep, Dario Zappa, and Elisabetta Comini. 2021. 'Solid Oxide Fuel Cell: Decade of Progress, Future Perspectives and Challenges'. *International Journal of Hydrogen Energy* 46 (54): 27643–74.

- Tseronis, K., I. Bonis, I.K. Kookos, and C. Theodoropoulos. 2012. 'Parametric and Transient Analysis of Non-Isothermal, Planar Solid Oxide Fuel Cells'. *International Journal of Hydrogen Energy* 37 (1): 530–47. <https://doi.org/10.1016/j.ijhydene.2011.09.062>.
- Tseronis, K., I.S. Fragkopoulos, I. Bonis, and C. Theodoropoulos. 2016. 'Detailed Multi-Dimensional Modeling of Direct Internal Reforming Solid Oxide Fuel Cells'. *Fuel Cells* 16 (3): 294–312. <https://doi.org/10.1002/fuce.201500113>.
- Vivet, N., S. Chupin, E. Estrade, T. Piquero, P.L. Pommier, D. Rochais, and E. Bruneton. 2011. '3D Microstructural Characterization of a Solid Oxide Fuel Cell Anode Reconstructed by Focused Ion Beam Tomography'. *Journal of Power Sources* 196 (18): 7541–49. <https://doi.org/10.1016/j.jpowsour.2011.03.060>.





# Parameter estimation combined with model reduction techniques for identifiability analysis of biological models

Michael Binns,<sup>a</sup> Alessandro Usai,<sup>b</sup> Constantinos Theodoropoulos,<sup>b</sup>

<sup>a</sup>*Department of Chemical and Biochemical Engineering, Dongguk University-Seoul, 30 Pildong-ro 1-gil, Jung-gu, Seoul 04620, Republic of Korea*

<sup>b</sup>*Department of Chemical Engineering, Biochemical and Bioprocess Engineering Group, The University of Manchester, Manchester M13 9PL, United Kingdom  
[k.theodoropoulos@manchester.ac.uk](mailto:k.theodoropoulos@manchester.ac.uk)*

## Abstract

Parameter estimation is typically used as part of model development to determine the values of unknown parameters. However, depending on the model complexity the number of parameters can also vary. High complexity models have large numbers of parameters requiring more computational effort to determine them and are also prone to overfitting. Low complexity models have smaller numbers of parameters but may have reduced accuracy. Based on available experimental data cross-validation can be used to compare different complexity models and determine the most appropriate complexity (James et al., 2013). Alternatively, it is possible to look at the identifiability of parameters based on experimental data which considers the sensitivity and correlation between parameters. Both these types of methodologies can be used to reduce the complexity of models such that insensitive and/or dependent/correlated can be removed or re-estimated and an alternative set of parameters can be computed. In this work both types of methods are explored with examples. Cross-validation combined with a Least Absolute Shrinkage and Selection Operator (LASSO) regularisation method is used to reduce the complexity of linear empirical equations for predicting the performance of downdraft biomass gasification (Binns and Ayub, 2021). Sensitivity and identifiability methods utilizing the Fischer Information Matrix (FIM) are used to reduce the complexity of a nonlinear system of partial integral differential equations describing a population balance model for microalgae cultivation (Usai et al., 2022). Application of these methods allows the number of parameters to be reduced depending on the tolerance and/or accuracy required.

**Keywords:** Model reduction, Parameter estimation, Optimisation, Identifiability, LASSO.

## 1. Introduction

Model development typically starts from some knowledge derived from existing experimental data and the equations which have previously been used to model similar systems. Based on this starting point a model might be suggested which includes all known physical variables and parameters which affect the system outputs. This approach might lead to a complex model for which there are a large number of unknown parameters. For this reason model reduction and identifiability methods can be used to reduce the complexity (Baker et al., 2015). In this study we compare two different approaches for model reduction combined with parameter estimation: LASSO based regularization (James et al., 2013; Binns and Ayub, 2021) and a sensitivity and

identifiability analysis utilizing the Fischer information matrix and multi-objective optimisation to reduce correlations and improve identifiability.

## 2. Parameter estimation methods

Parameter estimation involves minimising or reducing the difference between model predictions and experimental values by changing the values of parameters in the models. This is typically achieved by minimising a least-squares type function such as equation 1 where  $y_z$  is the experimental value,  $\hat{y}_z$  is the corresponding model prediction and  $N$  is the number of data points.

$$RSS = \sum_{z=1}^N (y_z - \hat{y}_z)^2 \quad (1)$$

The number of parameters to be estimated depends on the complexity of the terms used in the models to predict  $\hat{y}_z$ .

### 2.1. Parameter elimination

Reducing the complexity of models may be possible through the elimination of certain parameters which are either insensitive or do not affect the model outputs. There are potentially three ways by which parameters can be eliminated

- Set parameter to 0
- Set parameter to 1
- Set parameter to a fixed value

In cases where a fixed value is used this is defined using existing knowledge of the system. Alternatively setting values to 0 or 1 will typically allow the model equations to be written in a simpler form.

### 2.2. LASSO regularisation

$$RSS = \sum_{z=1}^N (y_z - \hat{y}_z)^2 + \lambda \sum_{i=1}^n |\beta_i| \quad (2)$$

Regularisation is the methodology where the objective function from equation 1 is modified to include a second set of terms including the sum of the values of  $n$  parameters,  $\beta$ , multiplied by a factor,  $\lambda$ . When this function is minimised the values of parameters are also minimised and if the absolute values of  $\beta$  are used (as shown in equation 2) the effect is that a number of parameters are set equal to zero (James et al., 2013). Varying the value of  $\lambda$  will change the number of parameters set to zero. A very small  $\lambda$  value leads to the same solution obtained by minimising equation 1, a very large value  $\lambda$  will lead to a solution where all the parameters are set to zero. Additionally, this value will also affect the model accuracy, so cross-validation is required to find the most appropriate value of  $\lambda$ . In particular cross-validation involves dividing the training set into a number of sections or folds. One of the folds is selected for testing and the remainder are used for training. Sequentially repeating this for every fold will then lead to an overall cross validation mean square error. This cross validation is then repeated for every value of  $\lambda$ , being considered giving a range of different model solutions with varying accuracy and varying numbers of non-zero parameters.

### 2.3. Sensitivity and identifiability analysis

#### 2.3.1. Sensitivity based parameter reduction

As mentioned in section 2.1 a parameter might be eliminated by setting its value to either zero, one or some pre-defined fixed value. Starting from some existing full model containing possibly a very large set of parameters the change in model output ( $y$ ) resulting from and a small change in each parameter ( $P$ ) will give sensitivity values as shown in equation 3. Then sequentially considering the elimination of each parameter by setting it to some fixed value and checking if the resulting model error exceeds some specified tolerance leads to a reduction of the number of parameters.

$$S_p = \frac{dy}{dP} \quad (3)$$

#### 2.3.2. Identifiability analysis and correlations

After reducing the number of parameters either through the regularisation (as described in section 2.2) or utilizing sensitivities and a sequential removal method (as described in section 2.3.1) it is useful to consider the identifiability of the remaining parameters. Based on the available experimental data identifiability analysis asks which of those parameters can be uniquely identified. Those without a unique solution may have a range of acceptable values or it may not be possible to identify a reasonable value. Starting from the sensitivities it is possible to calculate the Fischer Information Matrix (FIM) which uses the local sensitivities around a fitted set of parameters ( $S_p$ ) together with the weighted variance ( $\hat{\Sigma}$ ).

$$FIM = \sum_{j=1}^{N_{exp}} \sum_{i=1}^{N_{sample}} \left( S_p(i, j) \right)^T \hat{\Sigma}^{-1} \left( S_p(i, j) \right) \quad (4)$$

The inverse of this FIM gives a lower bound for the covariance matrix which can be used to calculate the correlation matrix (Stoica and Ng, 1998; Baker et al., 2015). The following equations (5-8) are suggested in order to minimise correlations between parameters. From the correlation matrix ( $\rho$ ) it is possible to calculate an overall correlation coefficient as shown in equation 5. Subsequently from this correlation matrix the most highly correlated parameter (with the highest  $I_j$ ) is chosen as a pivot parameter and other parameters which are strongly correlated with this are defined by linear relations as in equation 6.

$$I_j = \sum_{i=1}^{N_{parameters}} \rho_{ij} \quad (5)$$

$$P_i = C_{ij} P_j \quad (6)$$

$$Z_1 = \min \left( \frac{y(C_{ij} P) - y(\hat{P})}{y(\hat{P})} \right)^2 \quad (7)$$

$$Z_2 = \min I_j(C_{ij}) \quad (8)$$

To reduce the magnitude of correlations present the correlation values can be optimised (subject to any bounds on the final parameter values from the model) by minimising  $Z_2$  as defined in equation 8. A second objective is given in equation 7 which is the deviation

in model output between using the original parameters and those modified by changing the correlation values. Hence, this is a multi-objective optimisation which should give a pareto curve of possible solutions. If parameters change without affecting model results this suggests they are not identifiable and can be eliminated or set to constant values.

### 3. Case studies

#### 3.1. Empirical models for biomass gasification

Biomass gasification models should predict the outlet gas composition based on a number of input variables. In particular the input biomass composition (*Ultimate analysis: %C, %H, %O, %Ash*), moisture content (*MC*), equivalence ratio (*ER*) and gasification temperature ( $T_{gas}$ ) are expected to affect the output gas composition. For the gasifier considered here additional inputs are also available including grate rotation speed (*Gr*), gas fan speed (*Fs*), Wet bulk biomass density (*bulk*) and biomass void percent (*void*).

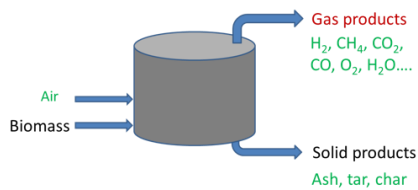


Figure 1. Block diagram representing a biomass gasifier.

Although more complex models are possible, for example based on knowledge of reaction kinetics, in this study simpler empirical models are considered relating the inputs with the outputs through linear and quadratic expressions. In this case the experimental data of Chee (1987) is used including 34 data points with inputs and outputs defined in Table 1. This is the same case investigated by Binns and Ayub (2021) who focused on finding the minimum cross-validation error model for each output. Here we consider the ranges of solutions which can be found with varying  $\lambda$ , the accuracy and simplicity of models obtained. For this purpose we consider the modelling of  $\text{CO}_2$  volume percentage in the produced gas with a range of possible solutions shown in table 1 and figure 2. As can be seen from figure 2 and table 2 the model accuracy varies with changing value of  $\lambda$ . For lower values all eleven parameters are non-zero but the cross-validation error is highest. The lowest cross-validation error is achieved in a linear model with 7 non-zero parameters and at high values of  $\lambda$  all the parameters are zero and the model reduces to a constant value with a single fixed value.

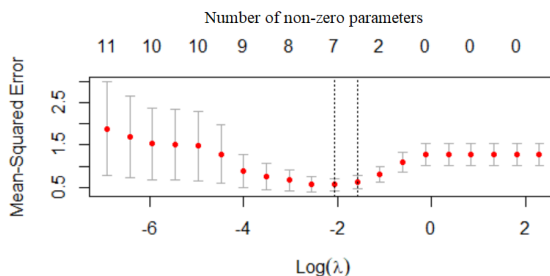


Figure 2. Cross-validation mean square error for  $\text{CO}_2$  using a linear model with varying  $\lambda$ . The number of non-zero parameters excludes a constant fixed parameter ( $\beta_0$ ).

Table 2. Parameter reduction using LASSO regularisation (Binns and Ayub, 2021)

$\lambda$ value	Cross validation error	Expression
$\lambda = 0.001$	1.8851	$CO_2(\%) = \beta_0 + \beta_1 Tgas + \beta_2 ER + \beta_3 MC + \beta_4 H + \beta_5 O + \beta_6 C + \beta_7 Ash + \beta_8 Gr + \beta_9 Fs + \beta_{10} Bulk + \beta_{11} void$
$\lambda = 0.2069$	0.6402	$CO_2(\%) = \beta_0 + \beta_2 ER + \beta_3 MC + \beta_5 O + \beta_7 Ash + \beta_8 Gr + \beta_{10} Bulk + \beta_{11} void$
$\lambda = 0.3360$	0.8053	$CO_2(\%) = \beta_0 + \beta_2 ER + \beta_3 MC$
$\lambda = 1.4384$	1.2909	$CO_2(\%) = \beta_0$

### 3.2. Population balance models for microalgae cultivation

The population balance model considered here is for the growth of the microalgae *Haematococcus pluvialis* from the model developed by Usai et al. (2022). This model describes the growth (G), birth (B), disappearance of mother cells (M) and cell lysis (D) giving the overall population balance in equation 9.

$$\frac{\partial \Psi_V}{\partial t} + G(v) = B(v) - M(v) - D(v) \quad (9)$$

where  $\Psi_V$  is the density distribution function of cells in the system.

The full model including expressions for each of the terms in equation 9 requires 34 parameters which are fitted to experimental data (Usai et al. 2022). It can be shown that applying sensitivity based parameter reduction 10 parameters can be eliminated with a tolerance of 0.05. Increasing the tolerance would allow more parameters to be eliminated (similar to the way increasing  $\lambda$  reduces the number of parameters).

Table 3. Parameter reduction using sensitivity approach (Usai et al., 2022b)

Tolerance	Parameters eliminated	Total parameters remaining
0.025	5	29
0.05	10	24
0.1	11	23

After the reduction of parameters, the sensitivity values can also be used to calculate the Fischer information matrix and subsequent covariance and correlation matrices. This identifies a number of strong correlations including existing relations between Monod kinetic parameters. If the objective function in equation 8 is minimised the magnitude of most correlations can be reduced by approximately 35%.

## 4. Discussion and conclusions

Two methods for model reduction based parameter estimation have been evaluated with two case studies and both are able to reduce complexity of the resulting models. A comparison of the two methods can be seen in table 4. The main advantage of the LASSO regularisation method is that it can start from a very complex model and directly reduce the number of parameters as part of the parameter fitting while the sensitivity and

identifiability approach generally starts from a full fitted complex model before reducing model complexity. However, the sensitivity and identifiability approach can more directly control the accuracy of the reduced model through the setting of a tolerance. It also offers more flexibility in terms of how to eliminate parameters, although possibly the LASSO regularisation could be modified so it can eliminate parameters in the same way. The main advantage of the sensitivity and identifiability methods are that they identify potential correlations between parameters and optimisation methods can be used to reduce the correlation values and potentially improve the identifiability of the model.

Table 4. Comparison of model reduction based parameter estimation methods

Issue	LASSO regularisation (Binns and Ayub, 2021)	Sensitivity and Identifiability analysis
Complex initial model	Can identify reduced models starting from a very complex model $\left( \begin{array}{l} N_{parameters} \gg N_{datapoints} \\ N_{reduced\ parameters} < N_{datapoints} \end{array} \right)$	Requires full potentially complex model with fitted parameters as starting point
Accuracy vs. simplicity control	Control using $\lambda$ factor (accuracy or model error must be calculated as an extra step)	Control using model tolerance
Parameter elimination	Set eliminated parameters to zero	Set eliminated parameters to either zero, one, or to specified values
Correlations and identifiability	Does not test for correlations. Reduces some statistical identifiability issues by reducing the number of parameters compared to the number of data points	Identifies potential correlations Can reduce correlation values to improve identifiability

## References

- S.M. Baker, C.H. Poskar, F. Schreiber, B.H. Junker, 2015, A unified framework for estimating parameters of kinetic biological models, *BMC Bioinformatics*, 16, 104
- M. Binns, H.M.U. Ayub, 2021, Model reduction applied to empirical models for biomass gasification in downdraft gasifiers, *Sustainability*, 13, 12191
- C.S. Chee, 1987, The air gasification of wood chips in a downdraft gasifier, Master's thesis, Kansas State University, Manhattan, USA.
- G. James, D. Witten, T. Hastie, R. Tibshirani, An introduction to statistical learning with applications in R, 2<sup>nd</sup> ed., Springer, New York, 2013, 59-126
- P. Stoica, B.C. Ng, 1998, On the Cramer-rao bound under parametric constraints, *IEEE Signal Processing Letters*, 5, 177-179.
- A. Usai, J.K. Pittman, C. Theodoropoulos, 2022, A multiscale modelling approach for haematococcus pluvialis cultivation under different environmental conditions, *Biotechnology Reports*, 36, e00771

# Improving Model Robustness with Transfer Learning for Product Property Models

Per Julian BECKER, Loic IAPTEFF, Benoit CELSE

*IFP Energies nouvelles, Rond-point de l'échangeur de Solaize, BP 3, 69360 Solaize, France*  
[per.becker@ifpen.fr](mailto:per.becker@ifpen.fr)

## Abstract

This paper describes the development of product property models for a new hydrocracking catalyst, developed by IFPEN. The naphtha density is used as an illustrative example. A Bayesian Transfer Learning approach is used to transfer the ordinary least squares (OLS) parameters from a source dataset consisting of 3980 points from industrial units operating previous (n-1) catalyst generations to a target dataset with 96 points from pilot plant tests with the new catalyst (n). Robustness of the new model is greatly improved because information from the larger dataset is included in the new model. The Transfer Learning approach is shown to give good results and lead to a more robust model with respect to feedstock descriptors.

**Keywords:** Transfer Learning, Modelling, Hydrocracking, Bayesian Statistics

## 1. Background

Accurate and robust predictive models are essential when commercializing a new chemical process in order to design the industrial units. Models can be derived from first principles or data driven. It is not always possible to do the former, which requires knowledge of all chemical reactions involved, as well as analytical methods for detailed characterization of the feedstocks and products. In this case empirical model must be developed, which require a much larger number of experimental data points in order to achieve the required precision and accuracy. Running pilot-scale experiments is rather expensive and time-consuming, which limits the amount of data which can be obtained in practice. The estimation of physical properties for petroleum refining products is a particularly challenging subject due to the complexity of the feedstocks, which are composed of hundreds of thousands of different hydrocarbon compounds.

### 1.1. Process Description

Hydroprocessing is a crucial component of a modern petroleum refinery (Jones 2006), which is used to convert heavy fractions obtained by distillation of crude oil (i.e. Vacuum Gas Oil – VGO) into more commercially valuable middle distillate (Diesel & Jet) or naphtha fractions over a fixed-bed catalyst under hydrogen pressure of 80 – 160 bar at temperatures from 360 – 410°C. Naturally occurring nitrogen and sulfur are removed in the Hydrotreatment (HDT) section, while long-chained hydrocarbons are broken up into smaller compounds in the Hydrocracking (HCK) section(s). The HDT section is necessary because organic nitrogen is a poison to zeolitic HCK catalysts. Ammonia (NH<sub>3</sub>) and hydrogen sulfide (H<sub>2</sub>S) gasses produced in the HDT section is carried over to the



HCK section, the  $\text{NH}_3$  inhibits HCK catalyst activity. The effluent of the reactor is split into the product cuts in a fractionator. It is possible to adjust the desired conversion and yield of product streams via selection of the catalyst(s) or operating conditions, this makes hydrocracking a very versatile process. HCK produces high-quality cracking products.

Higher rates of conversion can be obtained by recycling Unconverted Oil (UCO). In a two-stage process, the UCO-recycle stream is processed in a third HCK reactor, the effluents of the two HCK sections are blended before fractionation. Only trace amounts of  $\text{NH}_3$  are present in the second HCK reactor, which therefore requires much lower temperatures to achieve the desired conversion.

In maxi-middle distillate mode, the yields of Diesel and Jet cuts are favored, while maxi-naphtha mode aims to maximize naphtha production. Consequently some, or all, of the produced Jet and Diesel cuts are also recycled in this configuration.

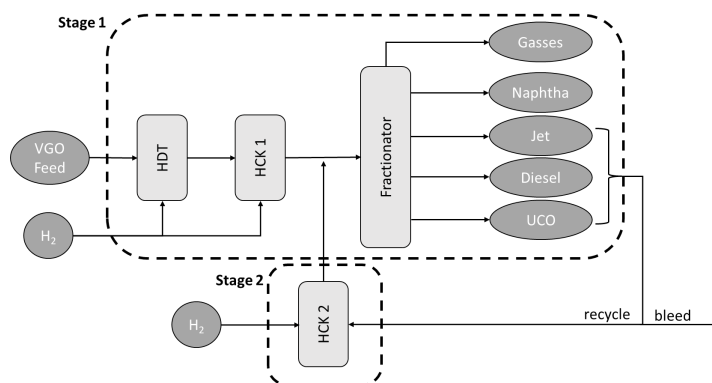


Figure 1 Schematic representation of the two-stage hydrocracking process in maxi-naphtha mode

### 1.2. Catalysts

IFPEN develops a wide range of HDT and HCK catalysts. The catalyst of generation n-1 is commercialized and used in single-stage maxi-middle distillate production and data from industrial units has been collected over the last 15 years. Catalyst of generation n has been developed for maxi-naphtha production in one- or two-stage configurations and is yet to be commercialized.

### 1.3. Product & Feedstock Properties

Petroleum feedstocks and products are complex mixtures of long-chained hydrocarbons, which can be roughly classified as paraffins (unbranched chains), iso-paraffins (branched chains), naphthenes (saturated cycles), and aromatics. The detailed characterization of VGO feedstocks, which are composed of 100,000s of individual species remains currently impossible. Petroleum cuts are characterized by Simulated Distillation (ASTM D2887), and density (spgr or  $d_{154}$ ), as well as the concentration of organic nitrogen and sulfur. Several product properties, such as RON/MON for naphtha or cold-flow properties for middle distillates are also measured – these properties are to be estimated from the macroscopic feedstock and process parameters.

The distillation curves of the products depend on the chosen cut points, which are adjusted to the requirements of each individual refinery, as well as the overlap between cuts, i.e.

efficiency of the fractionation column. The  $K_{\text{watson}}$  is a calculated feedstock descriptor, often used in petroleum refining, which combines the mean average boiling point (MeABP in °R) and specific gravity (spgr):  $K_{\text{watson}} = \text{MeABP}^{1/3} / \text{spgr}$ .

Accurate estimation of cut densities is particularly important because it is required to convert mass into volumetric flowrates, and because it is itself used as a descriptor in the models for other product properties.

## 2. Methodology

### 2.1. Bayesian Transfer Learning

The linear model for the target dataset is given below, where  $\beta_t$  is the vector of model parameters of size  $n_p$  and  $\mathbf{X}$  is the design matrix (a column filled with ones has been added to include the intercept).

$$y = \beta_t \mathbf{X} + \epsilon$$

For the transfer approach (Bouveyron 2010, Iapteff 2022, Robert 2007),  $\beta_t$  is considered to be a random variable of prior density  $\pi(\beta_t)$ . The Bayes theorem gives that the posterior of  $\beta_t$  with respect to the target observations  $\mathbf{X}_t$  and  $y_t$  is:

$$\pi(\beta_t | y_t) = \frac{\pi(\beta_t) f(y_t | \beta_t, \mathbf{X}_t)}{f(y_t | \mathbf{X}_t)}$$

An improvement of Zellner’s prior, also known as g-prior was proposed by L. Iapteff (2022), which acts on prior covariance in addition to prior distribution only:

$$\pi(\beta_t) = \mathcal{N}(\hat{\beta}_s, g \sigma_s^2 (\mathbf{X}_s^T \mathbf{X}_s)^{-1})$$

Where  $\hat{\beta}_s$  is the maximum likelihood estimator learned on only the source data. The mean of the corresponding posterior distribution is then:

$$\hat{\beta}_t = (\mathbf{X}_t^T \mathbf{X}_t + \sigma_t^2 g^{-1} \Sigma_s^{-1})^{-1} (\mathbf{X}_t^T y_t + \sigma_t^2 g^{-1} \Sigma_s^{-1} \hat{\beta}_s) \text{ with } \Sigma_s = \sigma_s^2 (\mathbf{X}_s^T \mathbf{X}_s)^{-1}$$

For large values of g the posterior mean tends to  $\hat{\beta}_s$ , for small values it tends to the prior mean ( $\hat{\beta}_t$ ). The difficulty lies in determining a suitable value for g. L. Iapteff (2022) showed that for normalized data a good estimate for the optima value of g can be defined as follows, in terms of  $\Sigma_s = \sigma_s^2 (\mathbf{X}_s^T \mathbf{X}_s)^{-1}$ :

$$g = \left( \frac{1}{n_p} \sum_{j=1}^{n_p} (\Sigma_s)_{j,j} \right)^{-1}$$

### 2.2. Feature Selection

A limited number of potential features is available due to the limitation described in 1.3. Hydrogen partial pressure and residue conversion (X370+) are retained as process descriptors. Nitrogen, Sulfur, density, and KWatson are retained as feedstock descriptors. Furthermore, the SimDis end-points (5% and 95%) were retained for both, the feedstock and naphtha cuts.

### 2.3. Data Normalization

The data is normalized according to the median for centering and interquartile range for reduction, based on the source training dataset (Iapteff 2022).

### 2.4. Process Simulation

In practice, the selected process and feedstock descriptors are always known with high confidence. The SimDis of the HCK effluent is estimated using a Continuous Lumping model and product cut SimDis is calculated using a model for fractionator efficiency (Kister 1992). The ensemble is implemented, together with the various product property correlations in a simulator developed in-house by IFPEN.

## 3. Datasets

### 3.1. Source Dataset

Follow-up data was obtained for 31 cycles from 15 industrial plants using HCK catalyst of generation n-1 in maxi-middle distillate, single stage mode, with or without recycle. A total of 3980 data points were retained for model development after expert-filtering and outlier detection (LOF), which was performed on a per-cycle basis. The entire industrial dataset was used to estimate the model parameters, a second dataset with 129 points obtained from IFPEN pilot plant tests was used for validation of the model. Performance indicators for the two datasets are shown in Table 1, the RMSE is in both cases below the acceptability criteria of  $\pm 0.005$  g/cm<sup>3</sup>.

Table 1 Indicators for source model

	<i>me</i>	<i>mae</i>	<i>rmse</i>	+/- 0.01 (%)	+/- 0.005 (%)	+/- 0.0025 (%)
<i>Train (Indus)</i>	0.0002	0.0030	0.0040	97.8	83.2	52.5
<i>Test (Pilot)</i>	0.0011	0.0034	0.0042	97.7	72.9	45.7

### 3.2. Target Dataset

The target dataset for the new catalyst contains 115 experimental points for stage 1 and 24 points for stage 2 from IFPEN pilot plant tests. Seven different feedstocks were available for testing in Stage 1, however, only two feeds were available for stage 2. The feeds were pre-treated in a separate pilot plant with HDT+HCK catalysts and tested in a single HCK reactor. This approach does not exactly reproduce the HCK 2 feeds in an actual recycle configuration but was considered sufficiently close to develop a first model. The principal weakness of this dataset is the lack in feedstock variability, whereas the feeds vary on a day-by-day basis during an industrial cycle.

#### 3.2.1. Gas-Liquid re-combination

An innovative analytical method based on recombination of gas and liquid analysis (chromatography), developed by IFPEN, was used to characterize the cuts below 175°C in terms of family (paraffins, iso-paraffins, naphthenes, and aromatics) and carbon number. Using this data, the distillation curves and densities for arbitrary naphtha cuts can be calculated. Three artificial cuts, 65-150°C, 80-150°C, and 80-175°C were thus generated, greatly increasing the number of points available for model development.

#### 3.2.2. Stage 1 – Changing the Reaction Coordinate

Residue conversion of the residue cut, with boiling points above 370°C (X370+) is used as reaction coordinate in maxi-middle distillate HCK, maxi-naphtha modem on the other hand aims to convert the middle distillate cuts. It is furthermore possible that the feed

contains significant amount of Gasoil (250-370°C), the X370+ is less meaningful in this case. The X175+ conversion is used instead. The transformation between X370+ and X370+ is essentially linear in nature.

3.2.3. Stage 2 – Changing the Model Perimeter

The model parameters for stage 1 and stage 2 are shown in Figure 1. The feedstock for stage 1 is the feed of the HDT reactor, which is known. The HDT effluent (i.e. HCK1 reactor feed) is rarely analyzed in industrial units, it is therefore not included in the model. For stage 2 pilot plant tests the HCK2 reactor feed, i.e., the 175+ cut of the blended HCK1 & HCK2 effluents is used in the model. This represents a significant deviation from the stage 1 case, as illustrated in Figure 2.

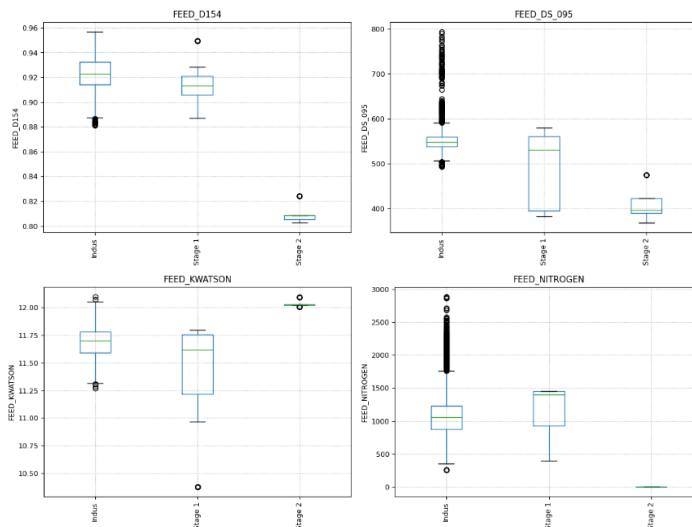


Figure 2 Feedstock characteristics in source (Indus) and target datasets for stage 1 and stage 2

4. Results & Discussion

The model indicators for Stage 1 and Stage 2 with and without transfer from the n-1 dataset are shown in Table 2 and Table 3 respectively.

Table 2 Performance indicators for Stage 1 model with and without transfer

	<i>me</i>	<i>mae</i>	<i>rmse</i>	+/- 0.01 (%)	+/- 0.005 (%)	+/- 0.0025 (%)
<i>Train without transfer</i>	-0.0008	0.0024	0.0035	97.6	86.0	68.4
<i>Test without transfer</i>	-0.0007	0.0022	0.0034	95.3	88.4	73.3
<i>Train with transfer</i>	-0.0006	0.0025	0.0037	97.9	84.5	64.4
<i>Test with transfer</i>	-0.0005	0.0025	0.0037	95.3	86.0	69.8

Table 3 Performance indicators for Stage 1 model with and without transfer

	<i>me</i>	<i>mae</i>	<i>rmse</i>	+/- 0.01 (%)	+/- 0.005 (%)	+/- 0.0025 (%)
<i>Train without transfer</i>	0.0011	0.0015	0.0019	100	100	80
<i>Test without transfer</i>	0.0007	0.0014	0.0017	100	100	90.5
<i>Train with transfer</i>	0.0011	0.0016	0.0021	100	97.3	74.7
<i>Test with transfer</i>	0.0008	0.0015	0.0019	100	100	81

For both datasets the results are well within the acceptable range of  $\pm 0.05 \text{ g/cm}^3$ . Model performance with transfer is equivalent to not using transfer learning, the values of the fitted parameter are however rather different. This suggests significant over-fitting when transfer learning is not applied.

An additional test was performed to validate the robustness of the two models. This time the pilot plant was operated in two-stage recycle mode half of the time and in single stage mode the other half. Two, previously unseen feedstocks were tested. The points were simulated using the in-house simulator (see 2.4). Figure 4 shows that the transfer modelling approach leads to more robust model. The model without transfer performs very poorly, confirming the over-fitting of the feedstock related features. The transferred model performs much better because information from the large industrial dataset is retained.

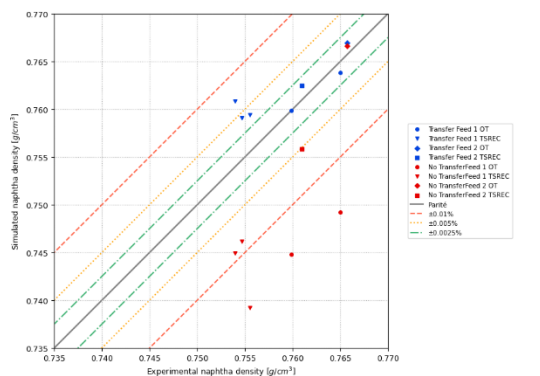


Figure 3 Parity graph for additional experimental points

## 5. Conclusion

Using transfer learning on the linear model, with an improved prior including the prior mean and covariance, was applied successfully on real-world datasets. Model robustness was shown to be improved when the target dataset is limited.

## References

- L. Iapteff, 2022, Transfer Learning for Smart Predictive Analytics, PhD thesis
- C. Bouveyron and J. Jacques, 2010, Adaptive linear models for regression: improving prediction when population has changed, *Pattern Recognition Letters*, pages 2237–2247.
- P.J. Becker, B. Celse, D. Guillaume, V. Costa, L. Bertier, E. Guillon, and G. Pirngruber, 2016, A continuous lumping model for hydrocracking on a zeolite catalysts: model development and parameter identification, *Fuel*, pages 73–82.
- H.Z. Kister, 1992, *Distillation Design*, McGraw-Hill, New York, USA, ISBN 0-07-034909-6.
- D.S. Jones, and P.P. and Pujado, 2006, *Handbook of petroleum processing*. Springer Science & Business Media.
- L. Iapteff, J. Jacques, M. Rolland, and B. Celse, 2021, Reducing the number of experiments required for modelling the hydrocracking process with kriging through bayesian transfer learning. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*.
- C.P. Robert et al., 2007, *The Bayesian choice: from decision-theoretic foundations to computational implementation*, volume 2. Springer.

# Modeling and simulation approach for modularized hydrogen electrolyzer plants

Kumar Rajan Gopa<sup>b</sup>, Isabell Viedt<sup>a</sup>, Lucien Beisswenger<sup>c</sup>, Leon Urbas<sup>a,b</sup>

<sup>a</sup>TU Dresden, Process Systems Engineering Group, 01069 Dresden, Germany

<sup>b</sup>TU Dresden, Chair of Process Control System, 01069 Dresden, Germany

<sup>c</sup>VDMA e.v., P2X4A, 60528 Frankfurt, Germany

[kumar.rajan\\_gopa@tu-dresden.de](mailto:kumar.rajan_gopa@tu-dresden.de)

## Abstract

The increasing importance of flexible and large-scale green hydrogen production necessitates the accompanying modeling and simulation of these processes to optimize plant flexibility. State-of-the-art monolithic modeling approaches do not address the flexibility and interoperability expected from modular plants, as shown by (Mädler *et al.*, 2022). Therefore, in this work, a new approach to the modeling of modular electrolyzers is proposed. Beginning with a steady-state equation system for the individual Process Equipment Assemblies (PEAs) it is later expanded to a detailed dynamic model. Instead of a single simulation flowsheet for the entire plant, this will result in a set of Functional Mockup Unit ready PEA models which can be combined depending on the plant assembly via co-simulation. This enables the user to generate models for various configurations and test the individual PEAs as well as the composite plant more efficiently. As a first step, a simple case study is presented where two steady state sub-models are developed and connected via co-simulation.

**Keywords:** Hydrogen Electrolysis, Modular plants, Simulation

## 1. Introduction

Alkaline water electrolysis (AEL) technology has been present for a long time. Most of the work related to the modeling of these electrolysis plants considered them to be monolithic and focused on the detailed modeling of the electrolyzer or on the entire plant, including the Balance of Plant (BoP). Ulleberg, (2003) presented a detailed parametric model that acts as a baseline for most of the works that followed. The closest work in alkaline electrolysis towards a modular design is done by Brée *et al.*, (2020), where the author defines the electrolysis model as a collection of individual half-cells which could be coupled to form the entire electrolyzer. But this doesn't account for the BoP.

As per VDI 2776 a modular plant is defined as an aggregation of individual process engineering assemblies (PEAs) that offers high flexibility and interoperability within the modules. Oppelt *et al.*, (2015) is one of the earlier works that point towards the use of simulation in the life cycle of a modular plant and describes the future scenarios involving model reuse and co-simulation. A design basis for such plants and their realization on a mini plant, pilot scale, and production scale is discussed by Eilermann *et al.*, (2018). This is clarified by Mädler, Viedt, *et al.*, (2022) in detail, explaining the requirements needed for a digital twin and elaborates the role of modeling and simulation in developing digital twins.

With the eventual goal of developing a digital twin, we try to contextualize this in the case of an alkaline water electrolysis plant, which could be modeled and simulated, taking into account the modular nature of the plant and propose a workflow that could be used for the modeling of different modular electrolysis assemblies. In the first section, the basic principles and functioning of the alkaline water electrolyzers is introduced along with an overview of the modular concept and how this is made part of the simulation process. In the next section a case study and its results are presented where the proposed workflow is applied to a couple of submodels. The paper ends with a conclusion pointing towards the next steps in the workflow.

## 2. Theory

Large scale and sustainable production of green hydrogen is gaining more and more importance and water electrolysis is one of the major technologies to achieve this goal. In project eModule of the hydrogen lead-project H<sub>2</sub>Giga, the focus is on developing modularized electrolysis plants that could be readily assembled and operated to achieve a high degree of flexibility and scaling up. Among the existing technologies, alkaline water electrolysis is the most mature and has been widely used commercially (Brée *et al.*, 2020). Water, being a stable compound, requires a lot of energy for its splitting. With electrolysis, this requires a high potential. In the case of alkaline water electrolysis, this is accomplished with the aid of an alkaline compound. This lowers the potential required for the breakup of water into hydronium and hydroxide ions. In the presence of an electric current hydrogen is produced at the cathode and oxygen is collected from the anode. Typically 20-40% KOH or NaOH solutions are used for electrolysis, and this is depicted in the below figure along with the reactions that occur during electrolysis (Shiva Kumar & Himabindu, 2019).

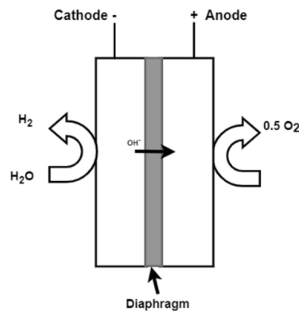
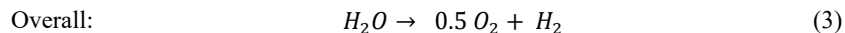
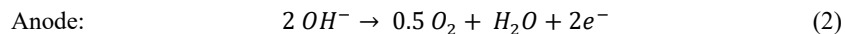
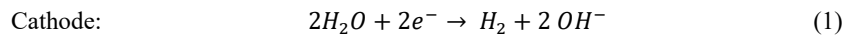


Figure 1: Alkaline Electrolysis (Shiva Kumar & Himabindu, 2019)



For increased flexibility and faster planning, modular approach to design and construction of electrolyzer-plants is assumed (Radatz *et al.*, 2019). With the modular approach, there will be a set of submodels that would have the functional properties corresponding to the individual and independent PEAs. A PEA or process engineering assembly consist of

process units that are largely independent, self-sufficient, and can perform a particular process step. It could include single or multiple functional equipment assemblies. These sub-models are brought together in a co-simulation framework where the various sub-models could be coupled together based on the required configuration. The individual PEAs could interact and communicate with each other exchanging information. The assembled model could later be simulated to obtain the overall simulation model of the plant. An overall process flow diagram for the alkaline electrolyzer along with the individual PEAs is shown in Figure 2.

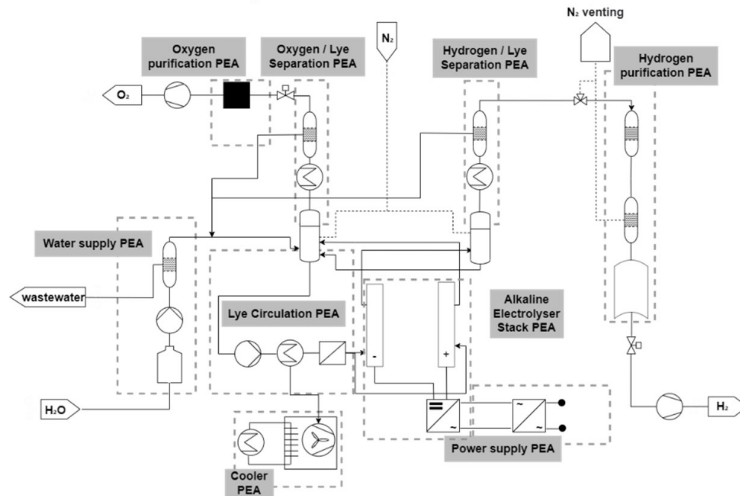


Figure 2: Process flow diagram for the AEL with individual PEAs, Courtesy of the VDMA Power-to-X for Applications platform (Beisswenger & Erdmann, 2022)

Compared to a continuous plant, where these are designated as individual units, here the distinction is made based on the function served. The PEAs present are:

1. Water supply – The function of this PEA is to supply water for the electrolyzer that meets the required specifications.
2. Lye circulation – This PEA assembly mixes the individual lye streams and makes up water, pressurizes the lye stream to the set pressure, heats it to the required temperature specification, and distributes the lye to the Electrolyzer.
3. AEL stack (Electrolyzer) – The water electrolysis occurs here generating hydrogen at the cathode and oxygen at the anode.
4. Cooler – Excess heat from the electrolyzer is removed
5. Power supply – Rectifier arrangement that supplies the necessary power for the electrolyzer
6. Hydrogen Separation - Separates the hydrogen from the gas-liquid lye stream coming from the electrolyzer PEA, cools the outgoing gas to room temperature, and eliminates any fine droplets present in the gas stream.
7. Oxygen separation – Similar to the hydrogen separator, oxygen is separated.
8. Hydrogen purification – removes any oxygen in the stream dries the stream and compress it to the required specification
9. Oxygen purification – Since oxygen is not considered further, this is sent to a combustion chamber and disposed of



Following the iterative modeling process described in Mädler *et al.*, (2021), the first step involves the creation of a simple model. Such a process provides a baseline model for comparison, easier debugging, and testing and, results in models with different levels of complexity. After gathering the necessary preliminary information like the flow diagrams, functional process description, etc. a sub model for each PEA is developed. Since the eModule project consists of several types of electrolyzer modules, individual PEAs are developed in such a way that they can be reused and integrated across multiple electrolyzer technologies with minimal modifications. To do this, each component is developed as an object with its characteristics that can be replicated and reused. They are converted into Functional Mock-up Units (FMU), assembled, and simulated. These steady state models would be used to check the basic performance of the electrolyzer assembly, and it also serve as a baseline with which the results of the more complicated configurations could be compared. In the final step, various tests are performed for the quality of the model (Viedt *et al.*, 2022).

In the second stage, the simple model is extended to include control strategies. Necessary modifications in the individual PEAs as well as the necessary control loops are made during this step. This stage allows for the modular automation of the PEAs. Again, the PEAs are assembled, simulated, and tested. In the final step, the equation system is expanded. If needed, additional equations and correlations for various parameters are supplemented including those for physical properties, thus resulting in a final detailed model that closely resemble a physical plant. Again, the control loops are tested, and the co-simulation is made. The approach is illustrated in the following diagram.

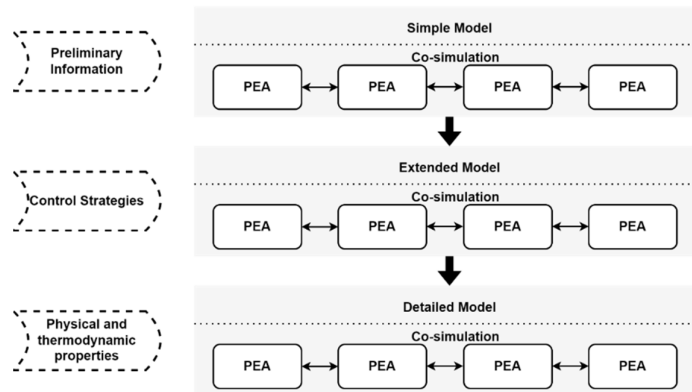


Figure 3: Workflow for Model Development

### 3. Case study

This is demonstrated by developing a co-simulation model for the PEA of a simple AEL sub-model and combining this with a lye separation sub model. In general, the starting point is a simple steady-state model for each individual PEA. In this phase the goal of the modeling process is to develop a minimal model that could be connected with other PEAs while outputting verified results with a certain degree of accuracy. For developing the initial model for AEL, the model equations presented by Ulleberg (2003) are used. This is a highly parametric model that has been experimentally validated. The main assumptions for this simple steady state model are:

- a. KOH is readily ionized, and water is replenished sufficiently to ignore the effect of OH- accumulation.
- b. All gases are in ideal condition and hydrogen is freely transported across the membrane (there is no resistance due to bubble formation, diffusion, etc.).
- c. A constant Faraday efficiency of 0.86 is assumed and excess heat is transferred to the cooler

For the cooler, the water inlet temperature is assumed constant at 20°C and exits at 30°C and the cooler is considered as an ideal heat exchanger with no pressure drop across it. Also, the physical properties are considered constant within each PEA. The individual PEAs were developed and after verifying them, they were exported as FMUs and simulated. Further details and the simulation models are provided as a GitHub repository (Rajan Gopa, 2022). Thus, totally independent submodels were successfully combined. Input conditions from Sánchez *et al.*, (2020) were used to validate results. For the simulation, 896 kg/h lye flow, Temperature, and pressure at 75 ° C and 7 bar and an electric power supply of 10 kW were used. The results from the co-simulation are:

Results	eModule model	Reference
Oxygen flow - cathode, kg/h	1.45	1.37
Hydrogen flow - anode, kg/h	0.18 (2.05 Nm <sup>3</sup> /h)	0.17
Temperature increase of Lye, °C	1.6	2.2
Current density, A/cm <sup>2</sup>	0.4721	0.42
Heat to cooler, W	1500	2392.8
Cooling water, kg/h	-	130

Table 1: Results of the simple model simulation and comparison to Sánchez *et al.* (2020)

The simple model adhered well to the steady state results from Sánchez *et al.*, (2020). The oxygen and hydrogen production are similar and the increase in value could be attributed to the higher current density for the co-simulation model. Only the excess heat that goes to the cooler is comparatively lower than that in the cited work. This could be due to the use of constant values for the physical properties and the underlying assumptions. This is also evident from the difference in the lye temperature increase. Cooling water consumption for this operating point was also obtained but this could not be validated with the referenced work. But the obtained value does correspond to a heat duty in the range of 1500 W. As mentioned before, the overall goal of the simple model is not to produce exact results but to provide a valid and reliable baseline that could be compared to, which the above results justify.

#### 4. Conclusion

In this work, modular workflow to the modeling and simulation of electrolyzers is discussed, and a simple model is implemented for two PEAs of the alkaline electrolyzer plant. These two sub-models are then exported as FMUs and simulated. The results from the co-simulation are compared to a similar work, thus validating this approach. Through this use case the viability of this method is demonstrated. Following this, the other PEAs will be modeled using the same approach to later simulate the entire electrolyzer plant. The deviation in the excess heat could either be due to the simplicity of the model or the

absence of any physical property data. This will be incorporated in the detailed models including their individual control strategies so that the final plant configuration could be selected from a combination of these PEAs and various test scenarios could be validated. Different scenarios involving the numbering up and interoperability of different electrolyzers from multiple stakeholders will be tested. To validate the proposed approach for the modular simulation models, the results from one of the co-simulation models will be compared to those obtained from a demonstrator plant.

### Acknowledgements

We would like to thank the German Federal Ministry of Education and Research and the Project Management Agency Jülich for their financial support within the framework of the eModule research project (FKZ 03HY116A-E) of the H2Giga lead platform.

### References

- Beisswenger, L., & Erdmann, D. (2022). Standardized flowsheet for the alkaline electrolysis. VDMA e.V.
- Brée, L. C., Wessling, M., & Mitsos, A. (2020). Modular modeling of electrochemical reactors: Comparison of CO<sub>2</sub>-electrolyzers. *Computers & Chemical Engineering*, 139, 106890. <https://doi.org/10.1016/j.compchemeng.2020.106890>
- Eilermann, M., Post, C., Radatz, H., Bramsiepe, C., & Schembecker, G. (2018). A general approach to module-based plant design. *Chemical Engineering Research and Design*, 137, 125–140. <https://doi.org/10.1016/j.cherd.2018.06.039>
- Mädler, J., Viedt, I., Lorenz, J., & Urbas, L. (2022). Requirements to a digital twin-centered concept for smart manufacturing in modular plants considering distributed knowledge. In *Computer Aided Chemical Engineering* (Vol. 49, pp. 1507–1512). Elsevier. <https://doi.org/10.1016/B978-0-323-85159-6.50251-7>
- Mädler, J., Viedt, I., & Urbas, L. (2021). Applying quality assurance concepts from software development to simulation model assessment in smart equipment. In *Computer Aided Chemical Engineering* (Vol. 50, pp. 813–818). Elsevier. <https://doi.org/10.1016/B978-0-323-88506-5.50127-3>
- Oppelt, M., Wolf, G., & Urbas, L. (2015). Towards an integrated use of simulation within the life-cycle of a process plant. 2015 IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA), 1–8. <https://doi.org/10.1109/ETFA.2015.7301521>
- Radatz, H., Schröder, M., Becker, C., Bramsiepe, C., & Schembecker, G. (2019a). Selection of equipment modules for a flexible modular production plant by a multi-objective evolutionary algorithm. *Computers & Chemical Engineering*, 123, 196–221. <https://doi.org/10.1016/j.compchemeng.2018.12.009>
- Rajan Gopa, K. (2022). AEL stack and cooler FMUs (Version 00) [MATLAB/Simulink]. Technische Universität Dresden/P2O Lab. [https://github.com/p2o-lab/AEL\\_stack\\_cooler](https://github.com/p2o-lab/AEL_stack_cooler)
- Sánchez, M., Amores, E., Abad, D., Rodríguez, L., & Clemente-Jul, C. (2020). Aspen Plus model of an alkaline electrolysis system for hydrogen production. *International Journal of Hydrogen Energy*, 45(7), 3916–3929. <https://doi.org/10.1016/j.ijhydene.2019.12.027>
- Shiva Kumar, S., & Himabindu, V. (2019). Hydrogen production by PEM water electrolysis – A review. *Materials Science for Energy Technologies*, 2(3), 442–454. <https://doi.org/10.1016/j.mset.2019.03.002>
- Ulleberg, O. (2003). Modeling of advanced alkaline electrolyzers: A system simulation approach. *International Journal of Hydrogen Energy*, 28(1), 21–33. [https://doi.org/10.1016/S0360-3199\(02\)00033-2](https://doi.org/10.1016/S0360-3199(02)00033-2)
- VDI 2776—Part 1: Process engineering plants; Modular plants. (2020).
- Viedt, I., Mädler, J., & Urbas, L. (2022). Quality assessment for dynamic, hybrid semi-parametric state observers. In *Computer Aided Chemical Engineering* (Vol. 51, pp. 1255–1260). Elsevier. <https://doi.org/10.1016/B978-0-323-95879-0.50210-1>

# Design and optimization of a shared heat exchanger network for an integrated rSOC system

Xinyi Wei<sup>a,b</sup>, Shivom Sharma<sup>a</sup>, Francois Marechal<sup>a</sup>, Jan Van Herle<sup>b</sup>,

<sup>a</sup>IPSE, EPFL Valais Wallis, 1950 Sion, Switzerland

<sup>b</sup>GEM, EPFL Valais Wallis, 1950 Sion, Switzerland

*xinyi.wei@epfl.ch*

## Abstract

To limit the global CO<sub>2</sub> emissions, installation of renewables is expected to increase significantly. This would require storage of renewable energy, possibly using an energy carrier such as hydrogen, methane, ammonia, etc. Reversible Solid Oxide Cell (rSOC) is a high temperature modern technology that can be operated in fuel cell or electrolysis mode. The Solid Oxide Electrolysis (SOEC) can be used to convert renewable electricity into hydrogen or carbon-based fuels (Power-to-Fuel). Conversely, Solid Oxide Fuel Cell (SOFC) has high overall efficiency for generating electricity from hydrocarbon fuels (Fuel-to-Power). The SOFC/SOEC technology is underdevelopment and being tested at several hundred kW scales. Generally, there are two different systems, SOFC and its balance of plant (BOP) system; SOE and methanation with their BOP system. The main challenge of rSOC system is to make both systems (SOFC and SOEC) highly integrated. Hence, it must utilize the maximum heat from heat sources and supply maximum heat to heat sinks. In this study, SOFC with oxy-combustion system and its reverse mode (SOEC) with methanation system have been optimized by minimizing the energy requirement and number of heat exchangers for individual system. The heat exchanger network synthesis problem is mixed integer linear programming problem, and it has been implemented in AMPL. For each heat exchanger (i.e., pair of streams), we have computed the heat exchanger area. This study evaluates the operating-flexibility of the shared HEN for rSOC-reformer/methanation system. The results show that heat exchangers can be shared between SOFC and SOEC modes, by manipulating the stream connections.

**Keywords:** Solid Oxide Cell, Heat Exchanger Network Synthesis, Renewable Energy Storage

## 1. Introduction

Today, the world's global primary energy supply is still heavily dependent on fossil fuels (up to 84.3% in 2019 [1]). It is well known that burning fossil fuels heavily contributes to climate change through the release of green-house gases (GHG) [2]. Furthermore, fossil fuels have a big impact on human health [3]. With the critical situation regarding climate change, there is a need to decarbonize economical activities with mitigation that occurs through different alternatives.

The scientific community is exploring several viable alternatives to replace the use of fossil fuels with renewables. However, variability, intermittency and storage are the main challenges for the renewable energy usage. Therefore, significant efforts have been invested on the development of efficient energy conversion technologies. Solid oxide fuel cell (SOFC) has attracted huge scientific attentions in the recent years. It is a modern and efficient power production technology that can reduce user's dependency on electricity-

grid. SOFC system can generate electricity using liquid [4] or gaseous fuels [5]. It can easily be adjusted for different industry power plant scales. Moreover, SOFC operates at high temperature (600 °C), and it cogenerates high-quality heat or steam, which can be used as a heat source within the system. As maximum fuel utilization for SOFC is around 85%, a burner is required for the combustion of unconverted fuels. The burner provides additional heat to the SOFC system, at the same time, it generates more CO<sub>2</sub>.

In order to meet “net zero” emission target for greenhouse gases in 2050, it is essential to introduce Carbon Capture and Storage (CCS) technology that can retrieve most of the produced CO<sub>2</sub> from emission intensive activities and store it permanently in nature (e.g., underground), leading to an almost carbon neutral activity. The stored CO<sub>2</sub> can also be used in methanation process, to convert H<sub>2</sub> into methane. In SOFC system, it is possible to perform oxy-combustion, by using stoichiometric amount of O<sub>2</sub> inside a catalytic burner. At the outlet of burner, CO<sub>2</sub> and H<sub>2</sub>O at high temperature are the only two products. After water condensation, CO<sub>2</sub> can be separated from the mixture automatically.

A reversible SOFC system or solid oxide electrolyzer cell (SOEC) has attracted huge attention for H<sub>2</sub> and fuel production. H<sub>2</sub> and O<sub>2</sub> are produced on the cathodic and anodic side of SOEC, respectively. H<sub>2</sub> can be stored and used in SOFC to produce electricity, whereas O<sub>2</sub> can also be stored and used in catalytic burner. SOEC with methanation is an alternative way of methane production, where excess renewable electricity is converted into methane. SOEC can be considered as a low-cost alternative to address the gap between hydrocarbon fuel/hydrogen supply and demand [6]. The reversibility of the solid oxide cells (rSOC) manages the variability of electrical grids by transforming and storing power through H<sub>2</sub>/methane when there is a surplus of electrical power or by generating electrical power through H<sub>2</sub>/methane when there is demand. A general system layout has been provided in

Figure 1.

For economic reasons, both systems shall share most of the components, and they should easily be switched. The heat exchange network (HEN) synthesis is very challenging for integrated rSOC system, due to two different operation mode of the system [7]. State of the art of HEN synthesis methodologies have been published over the past few decades, and sequential approach is commonly used [8][9]. In this study, a shared HEN has been designed and optimized for 10 kW SOFC system and 30 kW SOEC system.

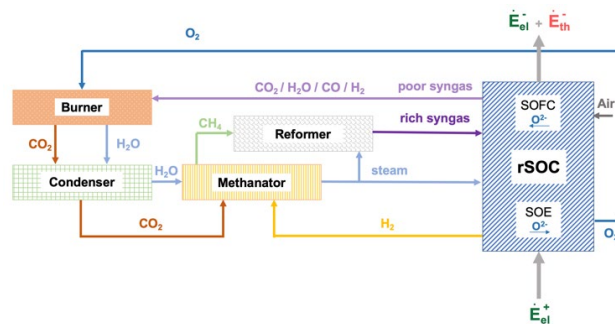


Figure 1. Global rSOC system layout

## 2. System and methodology description

### 2.1 SOFC with oxy-combustion

The fuel cell mode is operated when electricity demand cannot be fulfilled by the solar panels. In the fuel cell nominal mode, a mixture of methane and steam is fed into a reformer to convert part of it into syngas ( $H_2 + CO$ ). The resulting flow is then fed into the anode side of the fuel cell, where rest of the reforming takes place.  $H_2$  and  $CO$  are consumed inside the fuel cell, and steam and  $CO_2$  rich syngas comes out of fuel cell. On the cathode side, air is fed to supply oxygen, and oxygen depleted air is released into environment. The unconverted  $H_2$  and  $CO$  from fuel cell are combusted using air or pure oxygen. The oxygen combustion ensures absence of  $N_2$  in the exhaust gases, and recovery of product  $CO_2$  that can be used in SOEC/methanation,  $O_2$  has been assumed to be purchased from the market, but this configuration would allow to store and use oxygen produced from electrolyzer mode.

## 2.2 SOEC with methanation

Electrolyzer mode is operated when excess electricity has to be stored as methane. The excess electricity is fed into the electrolyzer, as well as steam on the cathode side. The cathode side outlet contains  $H_2$  and excess/unreacted steam. On the anode side, an air sweep is typically used to carry product  $O_2$ . Alternately,  $O_2$  can be used as sweep, and product  $O_2$  can be stored for further usage.  $H_2$  and  $CO_2$  are consumed inside the methanator to produce methane.

## 2.3 Heat exchanger network synthesis

The goal of the heat exchanger network synthesis formulation is to define process streams, objective function, heat balance equations, inequalities and potential constraints which will allow to obtain optimal amounts of hot and cold utility, the pairs of process streams which will exchange heat. Here, only essential details are presented, and detailed methodologies can be found in the literature [9].

- Overall stream heat balance

$$(T_i^i - T_i^o) \cdot Cp_i = \sum_{k \in ST} \sum_{j \in CS} Q_{ijk} + Q_{CUI} \quad i \in HS; \quad (T_j^o - T_j^i) \cdot Cp_j = \sum_{k \in ST} \sum_{j \in HS} Q_{ijk} + Q_{HUj} \quad j \in CS \quad (1)$$

- No splitting (avoid splitting of hot stream  $i$  and cold stream  $j$ )

$$\sum_{i \in HS} z_{ijk} \leq 1 \quad j \in CS, k \in ST; \quad \sum_{j \in CS} z_{ijk} \leq 1 \quad i \in HS, k \in ST \quad (2)$$

- Forbidden matches (avoid heat exchange between hot stream  $i$  and cold stream  $j$ )

$$z_{ijk} \leq 0 \quad k \in ST; \quad z_{CUI} \leq 0; \quad z_{HUj} \leq 0 \quad (3)$$

- Required matches (in the case of a required match, there are two options: data of heat exchange is removed from the streams input data, following constraints are activated)

$$\sum_{k \in ST} Q_{ijk} - Q_{min,ij} \geq 0; \quad Q_{CUI} - Q_{min,CUI} \geq 0; \quad Q_{HUj} - Q_{min,HUj} \geq 0 \quad (4)$$

- Restricted matches (the match is allowed but, in the case, where it exist the heat exchanged is subject to an upper bound)

$$\sum_{k \in ST} Q_{ijk} - Q_{max,ij} \leq 0; \quad Q_{CUI} - Q_{max,CUI} \leq 0; \quad Q_{HUj} - Q_{max,HUj} \leq 0 \quad (5)$$

- Objective functions: There are different options for the objective function. This study uses minimum heat requirement and minimum number of heat exchangers as objective functions.

$$\min \left( \sum_{i \in HS} Q_{CUI} + \sum_{i \in HS} Q_{HUj} \right); \quad \min \left( \sum_{i \in HS} \sum_{j \in CS} \sum_{k \in ST} z_{ijk} + \sum_{i \in HS} z_{CUI} + \sum_{j \in CS} z_{HUj} \right) \quad (6)$$

The HEN synthesis problem is mixed integer linear programming (MILP) problem, and it has been implemented in AMPL and solved by CPLEX solver. Cooling water (5 to 15 °C) is used as cold utility. Hot air is used as hot utility (800 to 1000 °C).

#### 2.4 Heat exchangers area estimation

In the first step, log-mean temperature difference is calculated. This study uses following values of overall heat transfer coefficients ( $U$ ,  $W/m^2/K$ ): 50 (air to air), 80 (gas to gas) and 160 (gas to water). The heat exchanger areas are computed using heat balance equation.

### 3. An integrated heat exchanger system design

Typical heat exchanger network of SOFC system has been studied in the previous literature [7]. SOFC system layout with HEN is shown in Figure 2. There are three heaters (H1, H2, H3) and four coolers (C1, C3, C4, C5). In order to improve the practical feasibility of HEN, HEN synthesis has been studied with some constraints. The evaporation of water occurs with direct injection of fuel in H1, therefore, the fuel and water streams are combined together, as shown in Figure 2. In practice, progressive water injection reduces heat requirement at constant temperature while providing a smooth temperature enthalpy profile. This aids in system heat integration. Moreover, in the previous studies, hot air has been used/split, to heat exchange with several streams. In reality, this design brings plant layout complexity, and it will also bring challenges in heat exchanger design, to avoid the potential leakage and explosion. Therefore, in this study, cold air heater (H3) has been restricted to perform heat exchange with hot air cooler (C3).

For SOFC system, the objective functions values for minimum energy requirement and number of heat exchangers are 2.01 kW (cold utility) and six, respectively. As expected, the exhaust gases from burner are used as the only heating source. The outlet stream from burner has been used to heat exchange with the stream before stack (C1-H2), and then it provides heat to the pre-reformer, finally, it passes through H1, to provide the heat for water heating, evaporation and superheating. The heat exchangers design specifications have been presented in Table 1 and Figure 4. There are mainly five heat exchangers: HEX 1, 2, 3, 4 and 5. HEX3 has the largest heat exchange area, where air flow rate is adjusted to balance the heat within the stack (exothermic reaction).

In SOEC system with methanator, the burner will be bypassed, as it is not necessary to burn part of cathode-off gas to provide heat to the system, and it decreases the energy efficiency.  $H_2$  produced from cathode side is pressurized and injected in methanator. Moreover, unlike SOFC exothermic operation mode, SOEC has an endothermic operating mode normally, therefore, additional electrical heating is required. For SOEC system, the objective functions values for minimum energy requirement and number of heat exchangers are 2.97 kW (hot and cold utilities) and six, respectively.

As shown in Figure 3, there are two heaters (H1', H3') and five coolers (C1', C2', C3', C4', C5'). Water flows through methanator as the first heat exchange step. Here, 10% heat loss has been assumed in methanator according to the experiment data. In the following step, steam exchanges heat with SOEC cathode off-gas (HEX6). The cathode off-gas is pressurized to 15 bar, and the temperature increases to 890 °C (maximum material resistance temperature). Finally, steam is heated up to 754 °C inside HEX7. Electrical heaters are used, to heat up anode inlet air from 760 to 800 °C, and steam from 754 to 800 °C. As shown in Table 1 and Figure 4, there are mainly five heat exchangers. Similar as SOFC system, the biggest heat exchanger is air to air heat exchanger (9.42 m<sup>2</sup>).

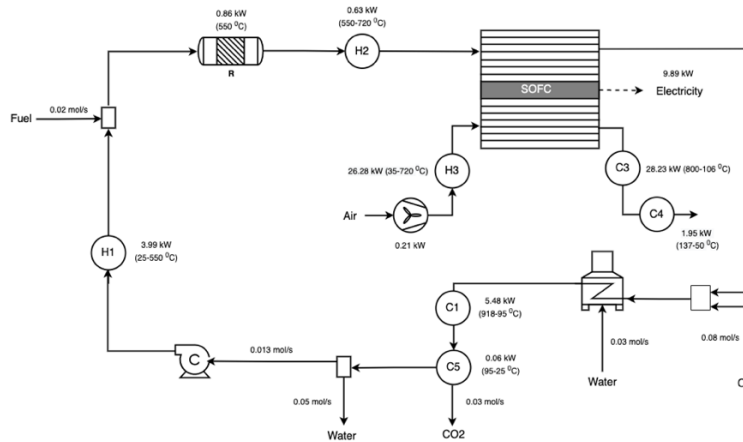


Figure 2. SOFC system with HEN design

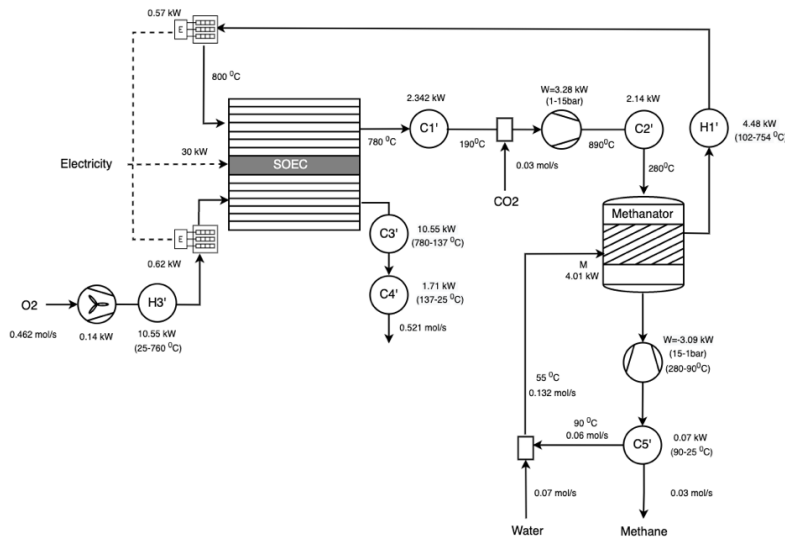


Figure 3. SOEC system with HEN design

By comparing duty and HEX area of two systems, HEX1, HEX2, HEX3, HEX4 and HEX5 from SOFC system can be reused as HEX6, HEX7, HEX9, HEX8 and HEX10 in SOEC system. As a conclusion, five shared heat exchangers are possible in rSOC system.

#### 4. Conclusion

This study presents a shared HEN design for rSOC system, comprising five heat exchangers, reformer and methanator. After studying heat integration, five heat exchangers can be shared. Air to air heat exchanger has largest heat exchanger area. The heat exchangers can be reused by adjusting pipelines connections. Moreover, methanator in SOEC system, has also been considered as the heat exchanger, this can be achieved by using adiabatic methanator, it benefits in waste heat usage and temperature maintainence. Additional electricity is required in electrical heaters and other components. Performing multi-period optimization for two systems simultaneously will be done in future work.



Table 1. SOFC and SOEC system key components specification

	Heater/Cooler	Tin (°C)	Thot (°C)	Duty (kW)	HEX area (m <sup>2</sup> )
<b>SOFC + oxy-combustion</b>					
HEX 1	C1	918	829	0.63	0.033
	H2	550	720		
C1 - R	C1	829	695	0.86	0.053
	R	550	550.1		
HEX2	C1	695	95	3.99	0.242
	H1	25	550		
HEX 3	C3	800	50	28.23	14.54
	H3	35	720		
HEX 4	C5	95	25	0.06	0.009
	CW	5	15		
HEX 5	C4	137	50	1.95	0.158
	CW	5	15		
<b>SOEC + methanation</b>					
H1' - M	M	280.1	280	4.01	0.250
	H1'	55	102		
HEX 6	C1'	780	190	2.34	0.061
	H1'	102	268		
HEX 7	C2'	890	280	2.14	0.524
	H1'	268	754		
HEX 8	C5'	90	25	0.07	0.011
	CW	5	15		
HEX 9	C3'	780	137	10.55	9.417
	H3'	25	760		
HEX 10	C4'	137	25	1.71	0.189
	CW	5	15		

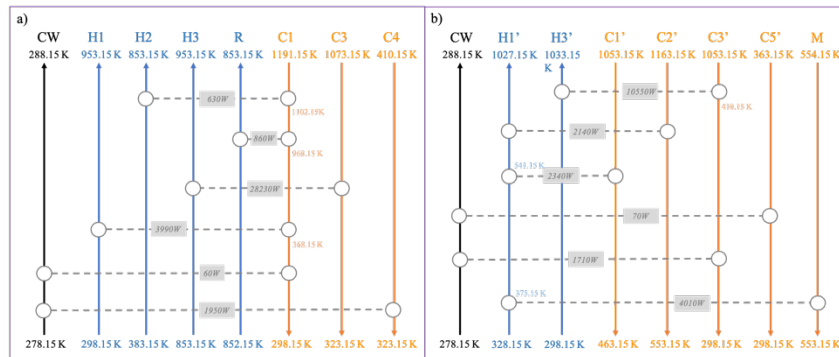


Figure 4. HEN for a) SOFC, b) SOEC systems

## References

- [1] Hannah Ritchie and Max Roser, "Energy Production and Consumption," *Our World in Data*, 2020. <https://ourworldindata.org/energy-production-consumption>
- [2] A. A. Lacis, G. A. Schmidt, D. Rind, and R. A. Ruedy, "Atmospheric CO<sub>2</sub>: Principal Control Knob Governing Earth's Temperature," *Science*, vol. 330, no. 6002, pp. 356–359, Oct. 2010, doi: 10.1126/science.1190653.
- [3] J. Lelieveld, K. Klingmüller, A. Pozzer, R. T. Burnett, A. Haines, and V. Ramanathan, "Effects of fossil fuel and total anthropogenic emission removal on public health and climate," *Proc. Natl. Acad. Sci.*, vol. 116, no. 15, pp. 7192–7197, Apr. 2019, doi: 10.1073/pnas.1819989116.
- [4] M. Santin, A. Traverso, and L. Magistri, "Liquid fuel utilization in SOFC hybrid systems," *Appl. Energy*, vol. 86, no. 10, pp. 2204–2212, Oct. 2009, doi: 10.1016/j.apenergy.2008.12.023.
- [5] Y. Kalinci and I. Dincer, "Analysis and performance assessment of NH<sub>3</sub> and H<sub>2</sub> fed SOFC with proton-conducting electrolyte," *Int. J. Hydrog. Energy*, vol. 43, no. 11, pp. 5795–5807, Mar. 2018, doi: 10.1016/j.ijhydene.2017.07.234.
- [6] P. Prabhakaran et al., "Cost optimisation and life cycle analysis of SOEC based Power to Gas systems used for seasonal energy storage in decentral systems," *J. Energy Storage*, vol. 26, p. 100987, Dec. 2019, doi: 10.1016/j.est.2019.100987.
- [7] F. Tanozzi, S. Sharma, F. Maréchal, and U. Desideri, "3D design and optimization of heat exchanger network for solid oxide fuel cell-gas turbine in hybrid electric vehicles," *Appl. Therm. Eng.*, vol. 163, p. 114310, Dec. 2019, doi: 10.1016/j.applthermaleng.2019.114310.
- [8] S. Sharma and F. Maréchal, "Robust Multi-Objective Optimization of Solid Oxide Fuel Cell–Gas Turbine Hybrid Cycle and Uncertainty Analysis," *J. Electrochem. Energy Convers. Storage*, vol. 15, no. 4, p. 041007, Nov. 2018, doi: 10.1115/1.4039944.
- [9] A. Mian, E. Martelli, and F. Maréchal, "Framework for the Multiperiod Sequential Synthesis of Heat Exchanger Networks with Selection, Design, and Scheduling of Multiple Utilities," *Ind. Eng. Chem. Res.*, vol. 55, no. 1, pp. 168–186, Jan. 2016, doi: 10.1021/acs.iecr.5b02104

# Multi-objective optimisation of the algae *Arthrospira platensis* under uncertainty: a Pareto ellipsoids approach

Wannes Mores, Satyajeet S. Bhonsale, Filip Logist, Jan F.M. Van Impe\*

*BioTeC+*, KU Leuven, Gebroeders De Smetstraat 1, Gent 9000, Belgium

\* [jan.vanimpe@kuleuven.be](mailto:jan.vanimpe@kuleuven.be)

## Abstract

Multi-objective optimisation allows a decision maker to evaluate trade-offs between objectives by supplying a set of optimal solutions, for example to find the balance between process performance and sustainability in process design. Accurate prediction of these objectives and, more importantly, safe operating conditions are needed. Thus, we need to account for uncertainty within the model. The Pareto ellipsoids approach ensures efficient use of computation while visualising the effect of parametric uncertainty. The novel Pareto ellipsoids-based algorithm is applied to the algae *Arthrospira platensis* and determines optimal light conditions for both biomass and product yield. Uncertain parameters are selected to investigate propagation of uncertainty from both the light intensity model and the Droop model. Finally, the obtained Pareto ellipsoids' accuracy is verified using a Monte Carlo approach.

**Keywords:** Multi-objective optimisation, Parametric uncertainty, Droop model

## 1. Introduction

Multi-objective dynamic optimisation is an important tool for (bio-)process design and operation (Logist et al., 2010). It allows a decision maker (DM) to evaluate different criteria, which are often of a conflicting nature. A typical example of this is the trade-off between an efficient, low-cost process and a high throughput process. To evaluate this trade-off, model-based multi-objective optimisation is often used. Here, a set of optimal solutions is generated, known as the Pareto front. With this set, a DM can decide where the point of operation lies. Pareto fronts can be generated with vectorisation (Bhaskar et al., 2000) and scalarisation approaches (Das and Dennis, 1998).

Model-based optimisation is a useful tool if the model represents the real process well. However, uncertainty on model parameters propagate through the model and influence the performance of the process. This not only leads to sub-optimal performance but can also cause critical errors in the process. Incorporating the effect of parametric uncertainty is important for safe and effective process design and operation (Vallerio et al., 2015).

Incorporating uncertainty into the optimisation is done in two ways. Robust optimisation accounts for the worst-case scenario and is a very conservative approach. Stochastic optimisation uses information on the distribution of the uncertainty and defines confidence on model outputs. Incorporating uncertainty into stochastic optimisation can be done in different ways, an overview is given in Nimmegeers et al. (2016).

In this work, stochastic methods are used to incorporate uncertainty. These methods lead to a significant increase in computational expense, it is therefore important to use

computationally efficient stochastic approaches and generate only significant Pareto points. Hashem et al. (2017) developed the D&C algorithm that finds relevant regions of the Pareto front with user-specified filter settings. In this work, the idea behind D&C is expanded by using information obtained from stochastic optimisation approaches. The uncertainty on model parameters will influence the performance regarding the objectives. In the case that the uncertainty on the objective is Gaussian, this can be shown by drawing ellipsoids around the Pareto points. These Pareto ellipsoids are then used to determine the significance of a Pareto point and aid in the spread of points on the Pareto front.

This novel Pareto ellipsoids-based optimisation algorithm is applied to the production of C-Phycocyanin through *Arthrospira platensis*. C-Phycocyanin is commonly used as a natural colorant in the food and cosmetic industry, reducing the need for synthetic colorants with possible toxic effects (Martelli et al., 2014). The use of *Arthrospira platensis* for production is promising as it yields high C-Phycocyanin concentration in cells (Xie et al., 2015). Del Rio Chanona et al. (2015) developed advanced dynamic models for this production process using experimental data from both batch and fed-batch processes. For this process, it was shown that light intensity plays a big role in C-Phycocyanin production. Higher light intensities promote biomass production but inhibit and even decrease intracellular C-Phycocyanin content. The substrate concentration promotes biomass growth but also suppresses C-Phycocyanin accumulation. In this work, a Droop model is used to determine optimal light conditions and substrate level for the trade-off between biomass production and C-Phycocyanin yield.

## 2. Methodology

In this section, the methodology for multi-objective optimisation under uncertainty is presented. The concept of Pareto ellipsoids is introduced, which is used to define a novel significance criterion in a Pareto ellipsoids-based algorithm. These methods are applied to the Droop model for *Arthrospira platensis*, described in the final part of this section.

### 2.1. Multi-objective dynamic optimisation

A multi-objective optimisation problem can generally be formulated as:

$$\min \mathbf{J} \quad (1)$$

subject to

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}, t) \quad (2)$$

$$\mathbf{c}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}, t) \leq 0 \quad (3)$$

Here,  $\mathbf{J}$  is the vector of objective functions to be evaluated,  $\mathbf{x}$  is the state vector,  $\mathbf{f}$  is the vector of state equations and  $\mathbf{c}$  is the constraint vector. The vector  $\mathbf{p}$  contains all the model parameters, including the uncertain parameters. It is assumed that the distribution is known for the uncertain parameters.

The propagation of this uncertainty is quantified using the sigma points method (Julier et al., 1996). This method relies on selecting sample points within the uncertainty distribution, which are propagated to the outputs of the model and combined to obtain the expected value  $E$  and variance  $V$  of model outputs. Other methodologies for uncertainty propagation exist, an overview is presented in Nimmegeers et al. (2016). The solutions of the multi-objective optimisation problem are obtained through the nominal boundary intersection (NBI) method, presented in Das and Dennis (1998).

### 2.2. Pareto ellipsoids

The uncertainty on the objective functions is used to define the concept of Pareto ellipsoids. If normally distributed uncertainty on the objective functions is assumed, the

uncertainty can be visualised using an ellipsoid. The center of the ellipsoid is the expected value  $E$ , and the radii can be defined using the covariance matrix  $\Sigma$  as follows:

$$(x - E[J])^T \Sigma_J (x - E[J]) = c_\alpha^2 \quad (4)$$

Where  $c_\alpha^2$  is the upper  $\alpha$ th percentile of a  $\chi^2$  distribution. The value of  $\alpha$  determines the level of confidence that the objective value is inside the ellipsoid.

### 2.3. Pareto ellipsoids significance criterion

Regions of the Pareto front with high uncertainty levels on objective performance are usually undesirable. A novel significance criterion is used to include this effect in the *a priori* smart filter. The Pareto ellipsoids criterion can be defined as follows:

A Pareto point is considered insignificant if the Pareto point falls within the Pareto ellipsoid of a neighboring Pareto point or if one of its neighboring Pareto points falls within the Pareto ellipsoid of the Pareto point.

For two Pareto points  $P1$  and  $P2$ , this can be checked by evaluating 2 inequalities:

$$(E[J_{P1}] - E[J_{P2}])^T \Sigma_J (E[J_{P1}] - E[J_{P2}]) \leq c_\alpha^2 \quad (5)$$

$$(E[J_{P2}] - E[J_{P1}])^T \Sigma_J (E[J_{P2}] - E[J_{P1}]) \leq c_\alpha^2 \quad (6)$$

### 2.4. Pareto ellipsoids-based algorithm

The underlying concept of the D&C algorithm with the information criterion from Hashem et al. (2017) can be extended with this novel criterion to reduce points in regions of high uncertainty. The information criterion settings are specified with  $\sigma_{low}^2, \sigma_{high}^2$ . The ellipsoid criterion is added when evaluating the information criterion. This requires the expected values and covariances of the objective functions, in this work obtained using the sigma points method. The ellipsoid criterion setting is specified using  $\alpha$ .

---

#### Algorithm 1 Pareto ellipsoid-based algorithm

---

**Input:** Specify  $\alpha, \sigma_{low}^2, \sigma_{high}^2$   
**Output:** Pareto (ellipsoid) set with adaptive resolution  $S$ .  
**Step 1:** Initialise solution set  $S = \{\}$ .  
**Step 2:** Construct weight cell  $C_{in}$  corresponding with the boundaries of the Pareto set  
**Step 3:** Initialise the weight cell  $C_w$ :  $C_w = C_{in}$   
**Step 4:** Start the recursive function for weight distribution  
**for**  $i = 1: n_w$  ( $n_w =$  number of entries in  $C_w$ )  
    **if**  $C_w \in \text{CHIM}$  **do**  
        Solve an NBI subproblem using  $C_w[i]$  to obtain Pareto point  $P_i$   
        Compute  $E[J_{P_i}]$  and  $V[J_{P_i}]$  (using sigma points)  
    **end for**  
**Step 5:** Construct a Pareto cell  $C_p = [P_1, \dots, P_{n_{Pareto}}]$ , with  $n_{Pareto} \leq n_w$   
**If** information criterion and Pareto ellipsoid criterion not activated **do**  
    Add  $C_p$  to  $S$   
    Divide: Construct daughter cells  $C_{d1}, \dots, C_{dn_w}$   
    Repeat **Step 4** with  $C_w = C_d \in \{C_{d1}, \dots, C_{dn_w}\}$   
**else Conquer:** Exit recursive call and stop exploring current segment.  
**Step 6:** Produce  $S$  when all recursive calls are ended  
**End**

---

### 2.5. Model description

The model used for optimisation is the Droop model for *Arthrospira platensis*. Droop models use intracellular quota to represent the nutrient storage inside the cell. In total, the model considers 4 states: biomass concentration  $c_X$ , nitrogen concentration  $c_N$ , intracellular normalised nitrogen quota  $c_q$ , and intracellular Phycocyanin quota  $c_{qc}$ . The model also incorporates the Aiba model, where illumination affects the phycocyanin production constant and specific growth rate. For brevity, the model equations are not included, a detailed description can be found in del Rio Chanona et al. (2015).

The light intensity throughout the reactor is determined using the Lambert-Beer law.

$$I(z) = I_0[\exp(-(\tau \cdot X + K_a) \cdot z) + \exp(-(\tau \cdot X + K_a) \cdot (L - z))] \quad (7)$$

Where  $I_0$  is the incident light intensity,  $\tau$  the cell absorption coefficient,  $K_a$  the bubble reflection coefficient, and  $z$  the distance from the light source. In this study, the photobioreactor is lit from both sides and has width  $L$ .

The light intensity plays a big role in the C-Phycocyanin production process as it promotes biomass growth and inhibits C-Phycocyanin production. To evaluate optimal lighting conditions for the process, two objectives are considered. These are maximisation of biomass growth ( $J_1$ ) and maximisation of C-Phycocyanin yield ( $J_2$ ). The control input for the optimisation is the incident light intensity  $I_0$ , and contributes to the contradicting aspect of the two objective functions as described previously. As optimisation is usually done by minimisation, the objective functions are negative resulting in maximisation.

$$J_1 = -c_{X,f} \quad (8)$$

$$J_2 = -\frac{c_{qcX,f}}{c_{N,0}} \quad (9)$$

Constraints are imposed on the model, which include minimal product return, a lower and upper bound on initial substrate concentration, and the initial state of  $c_X$ . The constraints are found in equations (9-11) The two parameters considered uncertain are chosen from both the Droop model equations and the nested Aiba model. These are the cell absorption coefficient  $\tau$  and the cell specific decay rate  $u_d$ . Both parameters are assumed to have a normally distributed uncertainty with a standard deviation of 10% of their nominal value.

$$c_X(0) = 0.1 \frac{g}{L} \quad (10)$$

$$c_N(0) \geq 1000 \frac{L}{mg} \quad (11)$$

$$c_{qcX}(f) \geq 100 \frac{mg}{L} \quad (12)$$

## 3. Results

The pareto ellipsoids-based algorithm is implemented in the POMODORO toolbox (Bhonsale et al., 2018), available at <https://cit.kuleuven.be/biotec/software/pomodoro>. It is a python-based toolkit for solving multi-objective optimal control problems. For optimisation, an  $\alpha$  of 70% is chosen, which means that the ellipsoids should have a 70% confidence that the true objective value lies within its boundaries. The settings for the information criterion  $\sigma_{low}^2, \sigma_{high}^2$  are chosen to be 0.001.

### 3.1. Pareto ellipsoids set

The multi-objective optimisation study results in a set of Pareto ellipsoids shown in Figure 1. Looking at the control profiles for the anchor points and a chosen trade-off solution, an early spike in light intensity is beneficial for initial biomass growth. Retaining the high light intensity results in efficient biomass production and limits the C-Phycocyanin production, which is seen in the maximised biomass control profile. The spike in control

profile is later for the maximised yield, which is explained through the fact that light intensity decreases the intracellular C-Phycocyanin content, and a later spike means less loss of product. The starting nitrate concentration is also lower for maximised yield. High nitrate concentrations suppress intracellular phycocyanin accumulation but promote biomass growth. The trade-off solution shows a mix of the two anchor points.

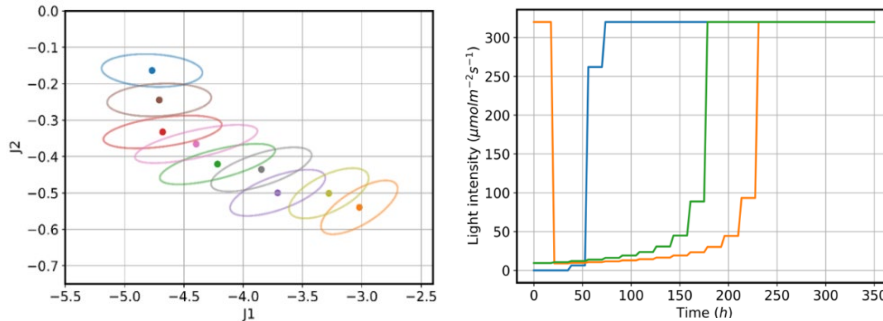


Figure 1: overview of the Pareto ellipsoids set (left) and selected control profiles (right)

Overall, the front of Pareto ellipsoids has relatively constant shape of ellipsoids. There is an increase in uncertainty in the middle of the trade-off zone between the anchor points. When minimising  $J_1$ , the ellipsoid seems relatively aligned with the objective axes, when moving towards minimising  $J_2$ , the ellipsoids have more rotation. Using this visualisation of uncertainty, a DM can improve his awareness of the uncertainty's effect and make a better decision regarding expected process performance and safety of operation.

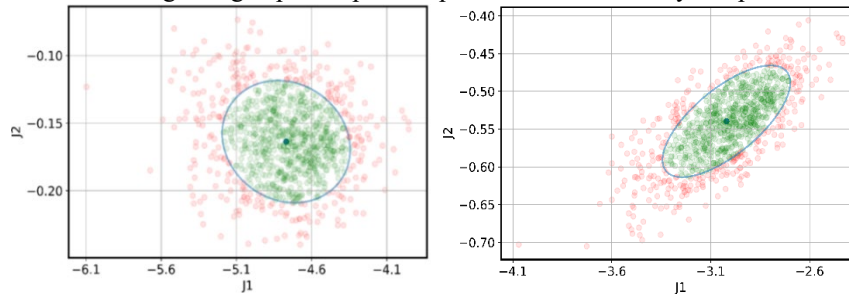


Figure 2: visualisation of objective values using Monte Carlo for both anchor points

### 3.2. Pareto ellipsoids verification

Monte Carlo simulation is used to verify the accuracy of ellipsoids. 1000 samples from the uncertain parameters' distribution are evaluated and shown in Figure 2. As the setting for the Pareto-ellipsoids based algorithm was set to  $\alpha = 70\%$ , the expectation is that 70% of all Monte Carlo iterations are within the boundary of the ellipsoid (green points).

Table 1: Sigma points accuracy through Monte Carlo simulation

Ellipsoid:	1	2	3	4	5	6	7	8	9
$\sigma_{J_1}^2$	97.6%	101.6%	100.2%	99.4%	101.9%	97.9%	98.0%	97.7%	97.9%
$\sigma_{J_2}^2$	98.2%	97.8%	97.2%	96.8%	96.4%	95.6%	96.1%	95.5%	96.0%
$\text{Cov}(J_1, J_2)$	103.7%	101.3%	99.7%	98.8%	108.7%	95.7%	94.9%	95.6%	95.8%
% green points	70.3%	69.3%	69.3%	70.1%	69.4%	68.8%	68.4%	68.5%	69.0%

Additionally, the covariance matrix obtained through sigma points can be compared to the covariance matrix from the Monte Carlo dataset to determine its accuracy. An overview of the covariance matrix accuracy is found in Table 1. Figure 2 shows the ellipsoid, which is using sigma points, and the Monte Carlo iterations as points.

Overall, the ellipsoids are very accurate. All ellipsoids contain between 68 – 71% of Monte Carlo iterations, which is very close to the desired 70%. The shape of the Monte Carlo dataset also seems to resemble the ellipsoid's shape. The covariance accuracy shows that the variance of objective 2 is underapproximated by around 3 – 4%. A more in-depth analysis is done for a range of different case studies in Mores et al. (2023).

#### 4. Conclusion

In this work, a novel Pareto ellipsoids-based algorithm for multi-objective optimisation is presented and applied to a case study of C-Phycocyanin production through *Arthrospira platensis*. The algorithm provides only relevant points, as defined by the user through settings  $\sigma_{low}^2$ ,  $\sigma_{high}^2$  and  $\alpha$ . This ensures efficient calculation of the Pareto set, which is important when implementing uncertainty propagation techniques. The Droop model as presented in del Rio Chanona et al. (2015) is used and two parameters are considered uncertain. Application of the novel algorithm results in a set of Pareto ellipsoids, which show how the uncertainty will propagate to the performance of the process. Early increase in light intensity is used to promote biomass growth, while delaying it leads to a more efficient C-Phycocyanin production. The ellipsoids' accuracy is evaluated using a Monte Carlo approach and shows that they represent the uncertainty accurately, allowing a DM to make a more informed decision on how to design and operate the process.

#### Acknowledgement

This research is co-funded by the Research Foundation Flanders (FWO) through the Senior Research Project G0B4121N ("Pharma4S") and the Strategic Basic Research Project S008121N ("BlueBioChain", H2020 ERA-NET Cofund on Blue Bioeconomy)

#### References

- Bhaskar et al. 2000, "Applications of multiobjective optimization in chemical engineering.", *Reviews in Chemical Engineering*, 16, 1
- Bhonsale et al. 2018, "Pomodoro: A novel toolkit for dynamic (multiobjective) optimization, and model based control and estimation.", *IFAC-PapersOnline*, 51(2), 719
- Das, I., & Dennis, J. 1998, "Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems.", *Siam Journal On Optimization*, 8, 631
- del Rio-Chanona et al. 2015, "Dynamic simulation and optimization for *Arthrospira platensis* growth and C-phycocyanin production.", *Industrial & Engineering Chemistry Research* 54.43, 10606
- Hashem et al. 2017, "A novel algorithm for fast representation of a Pareto front with adaptive resolution: Application to multi-objective optimization of a chemical reactor.", *Computers & Chemical Engineering*, 106, 544
- Julier et al. 1996, "A general method for approximating non-linear transformations of probability distributions.", Technical report, Robotics research group University of Oxford 1
- Logist et al. 2010, "Fast Pareto set generation for nonlinear optimal control problems with multiple objectives.", *Structural and Multidisciplinary Optimization* 42.4, 591
- Martelli et al. 2014, "Thermal stability improvement of blue colorant C-Phycocyanin from *Spirulina platensis* for food industry applications.", *Process Biochemistry*, 49.1,154
- Mores et al. 2023, "Multi-objective optimization under parametric uncertainty: A Pareto ellipsoids-based algorithm." *Computers & Chemical Engineering*, 169: 108099.
- Nimmegeers et al. 2016, "Dynamic optimization of biological networks under parametric uncertainty.", *BMC Systems Biology*, 10(1), 1
- Vallerio et al. 2015, "An interactive decision-support system for multi-objective optimization of nonlinear dynamic processes with uncertainty.", *Expert Systems with Applications* 42.21, 7710
- Xie et al. 2015, "Fed-batch strategy for enhancing cell growth and C-phycocyanin production of *Arthrospira (Spirulina) platensis* under phototrophic cultivation.", *Bioresource technology*, 180, 281

# Predicting the Evolution of Flammable Gases During Li-ion Battery Thermal Runaway Using Micro-Kinetic Modelling

Dr. Peter Bugryniec,<sup>a</sup> Dr. Sergio Vernuccio,<sup>a</sup> Prof. Solomon Brown<sup>a</sup>

<sup>a</sup>*Department of Chemical & Biological Engineering, The University of Sheffield, Sheffield, S1 3JD, UK  
[s.f.brown@sheffield.ac.uk](mailto:s.f.brown@sheffield.ac.uk)*

## Abstract

Li-ion batteries are a widely used electrochemical energy storage device. But, catastrophic failure via thermal runaway leads to great flammability and toxicity hazards. As such, there is a need to better understand the thermal runaway process. In doing so, reducing its occurrence and improving predictions of its hazards. To achieve this, we aim to develop a more detailed model of thermal runaway. This is based on fundamental reaction theory. Micro-kinetic modelling techniques are applied to predict the kinetic evolution of the reacting systems on a mechanistic level, based on a detailed analysis of the elementary reaction steps. Using this methodology, we simulate the thermal decomposition of dimethyl carbonate, as a model electrolyte solvent, and predict the product species present in the off-gas. We also investigate the impact of the temperature on the composition of the off-gas and the lower flammability limit. This demonstrates a method for predictive hazard assessments of Li-ion battery failure. For the DMC case study, we show that the LFL increases with increasing the operating temperature due to the large proportion of CO<sub>2</sub> generated. This effectively makes the off-gas safer in terms of explosion hazards. Further work will extend this methodology to construct the reaction systems for a complete Li-ion cell.

**Keywords:** Reaction network analysis, Hazard prediction, Dimethyl carbonate, Lower explosion limit, Thermal decomposition

## 1. Introduction

Li-ion batteries (LIBs) have been extensively adopted as energy storage devices. Hence, the safety of LIBs is of great importance. This is because under adverse conditions LIBs can undergo thermal runaway (TR) (Wang *et al.*, 2012) that occurs when the Li-ion cell's component materials exothermically decompose. This leads to an exponential increase in heat generation and cell temperature. Further, the decomposition process leads to the generation of a complex vapour-gas mixture including electrolyte vapours, as well as flammable, toxic, and poisonous gases (Fernandes *et al.*, 2018). The most notable toxic gas is HF. While CO<sub>2</sub> is not poisonous and can cause suffocation in large enough quantities. Flammable gasses also have the potential to cause explosions. LIBs are increasingly used in EVs, and domestic and commercial stationary energy storage. As a result, these hazards are present to more people and methods to improve LIB safety are essential.



To help achieve this goal LIB researchers and developers need to improve their ability to predict the complexities of TR. TR modelling has been used to predict thermal hazards of cells, some including flammability behaviour (Zhang *et al.*, 2022a). However, studies on the flammability of the off-gas are based on a given gas composition determined from experimental studies. This is because typical TR models that simulate the decomposition of the battery materials do not provide a detailed description of the reaction products (for example, Bugryniec *et al.*, 2020). Micro-kinetic modelling can overcome this limitation by including all reactants, e.g. Li-ion cell components, and all possible products, e.g. off-gasses (Bugryniec *et al.*, 2022) in the reaction network as well as each possible elementary step and reaction intermediate. Further, the kinetic parameters of these detailed models are typically estimated based on a solid theoretical basis and no a priori assumptions on the rate-determining steps are needed (Hermes *et al.*, 2019).

An important parameter to quantify the combustion risk is the lower flammability limit (LFL), or the lower explosion limit (LEL) (Ludwig, E., 1991). The LFL is the threshold percentage of gas that is required to be present in the air for it to ignite/explode. Below this threshold, the fuel conditions are too lean for combustion. Complimentary to this is the upper flammability limit (UFL) which indicates the highest fraction of flammable gas at which a flammable mixture is formed. In this work, we focus on the LFL as it is the start of the flammability hazard. The LFL can be determined experimentally or theoretically by the Le Chatelier formula which is discussed in detail in Section 2. The error calculating the LFL by Le Chatelier's formula is relatively small (<7%) (Li *et al.*, 2019). As such it is a widely accepted method in engineering applications.

The LFL has been calculated for the off-gas generated when experimentally abusing Li-ion cells. This can range from 4% to 50% depending on SOC and chemistry. Typically, the higher values of LFL relate to a higher fractional amount of CO<sub>2</sub> being generated (Chen *et al.*, 2020, Li *et al.*, 2019, Zhang *et al.*, 2022b). The LFLs of electrolyte solvents have also been reported. For dimethyl carbonate, diethyl carbonate and ethyl methyl carbonate the LFL are 3.2%-4.2%, 1.4%-1.6%, and 2.1% respectively (Henriksen *et al.*, 2019, IPCS INCHEM, 2000, IPCS INCHEM, 2005). However, the LFL of the off-gas from solvent decomposition has not been calculated, nor has the LFL of Li-ion cell off-gas been predicted. To the authors' knowledge, this is the first work on the prediction of the off-gas flammability behaviour on the topic of LIB safety.

Considering this, we aim to develop a TR model that predicts the generation and composition of LIB off-gas by applying micro-kinetic modelling methodologies. From this, the flammability (LFL) of the gas composition is assessed. In this work, we focus on one subcomponent of the cell, the electrolyte solvent dimethyl carbonate (DMC). For this component, we construct a reaction model in MATLAB® to simulate the generation of gas species under heating. With this, predictions are made on the gas composition during different stages of TR, from which changes in flammability hazard (i.e. LFL) are assessed.

## 2. Methodology

As stated previously, this work focuses on the reaction network for DMC thermal decomposition in an inert atmosphere. The reaction network applied is presented in Figure 1, reproduced from Fernandes *et al.*, 2019. The reaction network consists of 16 forward reactions and 12 species. The arrows in Figure 1 indicate the reactions proceeding in the forward direction. However, each elementary step included in the network was considered to be reversible. Within this, there are 3 possible pathways leading to the

decomposition of DMC: The direct decomposition of DMC to dimethyl ether, and the reaction of DMC with water or hydrogen.

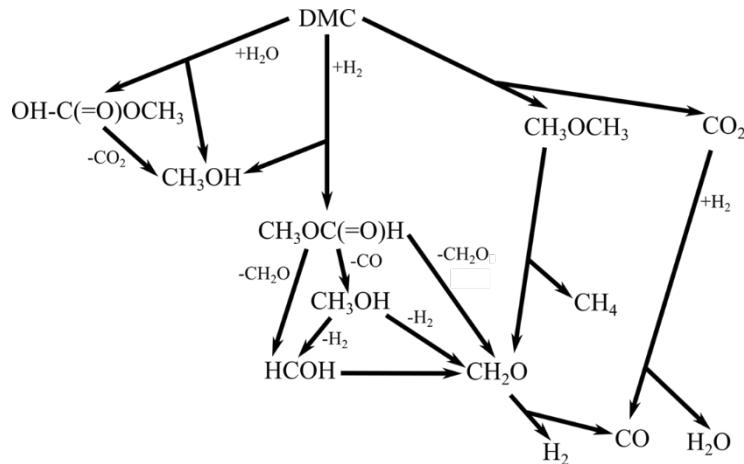


Figure 1. Reaction network for DMC thermal decomposition.

The micro-kinetic model consists of a system of 12 ordinary differential equations (ODEs) describing the change in concentration over time of the 12 species included in the mechanism. Each ODE is expressed as the algebraic sum of all relevant reaction rates ( $R_i$ ) corresponding to a given reaction step.  $R_i$  is governed by an Arrhenius equation of the form  $R_i = A_i e^{-E_{a,i}/RT} \prod_j C_j$ . In which  $A_i$  is the frequency factor,  $E_{a,i}$  is the activation energy,  $R$  is the gas constant,  $T$  is the temperature and  $C_j$  is the species concentration. The kinetic parameters of each rate-determining step are determined from transition state theory. The thermochemistry (of the reaction species, complexes, and transition state structures) was determined using density functional theory (DFT). This was carried out in Gaussian (Gaussian 16, 2019) with the method and basis set combination of B3LYP/6-311++G(d,p). The solution to the system of ODES is computed in MATLAB® (MATLAB, 2021).

The initial reactant system is assumed to be a mixture of only DMC and water with concentrations of 11.87 mole/L and 1.30E-3 mole/L, respectively. This is in line with the typical water content of commercially available DMC. The time-varying simulation assumes a constant linear temperature increase of 1°C/min, starting from 20°C and ending at 800°C. This is to approximate the common LIB abuse tests of adiabatic calorimetry. The predicted decomposition species are then used to calculate the temperature dependant LFL of the off gas.

As  $\text{CO}_2$  is expected to develop, the off-gas mixture can be considered dilute. The LFL for a dilute mixture,  $\text{LFL}_M$  (see Eq. 1), is calculated using the extended Le Chatelier's formula (Schroeder, 2016). The  $\text{LFL}_M$  is determined from the  $\text{LFL}'_i$ , the modified LFL of each species accounting for the diluting species, i.e.  $\text{CO}_2$ , (see Eq. 2). In which the  $\text{LFL}'_M$  is the LFL of only the flammable gas mixture (see Eq. 3). Further,  $a_i$  and  $B_k$  are the molar fraction of flammable gas  $i$  and molar fraction of inert gas  $k$ , respectively.  $\bar{K}$  is the average value of the nitrogen equivalency  $K_k$ , of the inert gases weighted according to their molar fractions. For pure  $\text{CO}_2$   $\bar{K} = 1.5$ .

$$LFL_M = \frac{100}{\sum_i^n \frac{a_i}{LFL'_i}} \quad (1)$$

$$LFL'_i = \frac{100 - LFL'_M - \left( (1 - \bar{K}) \frac{\sum_{p=1}^k B_k}{\sum_{i=1}^n a_i} \times LFL'_M \right)}{100 - LFL'_M} \times LFL_i \quad (2)$$

$$LFL'_M = \frac{100}{\sum_i^n \frac{a_i}{LFL_i}} \quad (3)$$

### 3. Results and Discussion

The predicted decomposition of DMC and generation of the off-gas species, along with the calculated LFL of the off-gas is presented in Figure 2. Only 8 species are presented in Figure 2 as the concentrations of the other 4 species ( $\text{H}_2\text{O}$ ,  $\text{OH-C(=O)OCH}_3$ ,  $\text{CH}_3\text{OC(=O)H}$ , and  $\text{HCOH}$ ) included in the reaction network do not reach a significant value. This is due to the low concentration of water in the initial solution which prevents  $\text{OH-C(=O)OCH}_3$  from being generated in high concentrations. On the other hand,  $\text{CH}_3\text{OC(=O)H}$  is not observed in high concentrations due to the number of favourable decomposition paths that this species can decompose by.

From Figure 2 it can be seen that DMC begins to decompose at  $450^\circ\text{C}$ . At this initial stage, DMC leads to  $\text{CH}_3\text{OCH}_3$  and  $\text{CO}_2$  generation only. At approximately  $550^\circ\text{C}$   $\text{CH}_3\text{OCH}_3$  starts to decompose leading to  $\text{CH}_4$  and  $\text{CH}_2\text{O}$  generation. At even greater temperatures,  $650^\circ\text{C}$ ,  $\text{CH}_2\text{O}$  decomposes to  $\text{H}_2$  and  $\text{CO}$ . At  $800^\circ\text{C}$  the gas produced predominantly consists of  $\text{CO}_2$  and  $\text{CH}_4$ , with  $\text{H}_2$  and  $\text{CO}$  in low quantities.

The  $LFL_M$  is calculated only considering the major species present. The LFLs of these individual species,  $LFL_i$ , are presented in Table 1. It can be seen from Figure 2 that  $LFL_M$  increases compared to the  $LFL_{\text{DMC}}$  as decomposition progresses. This is mainly driven by the diluting effect of  $\text{CO}_2$ . As even in the early stages of decomposition when  $\text{CH}_3\text{OCH}_3$  is present with a lower LFL than DMC the  $LFL_M$  still rises. This indicates that the greatest flammability risk when cell venting occurs (at  $\sim 200^\circ\text{C}$ ) would be from the solvent vapour (whilst not considering the electrode decomposition reactions). Further, we can see that generation of  $\text{CO}_2$  has a beneficial effect of increasing the  $LFL_M$ . However, a large quantity of  $\text{CO}_2$  presents a significant suffocation hazard.

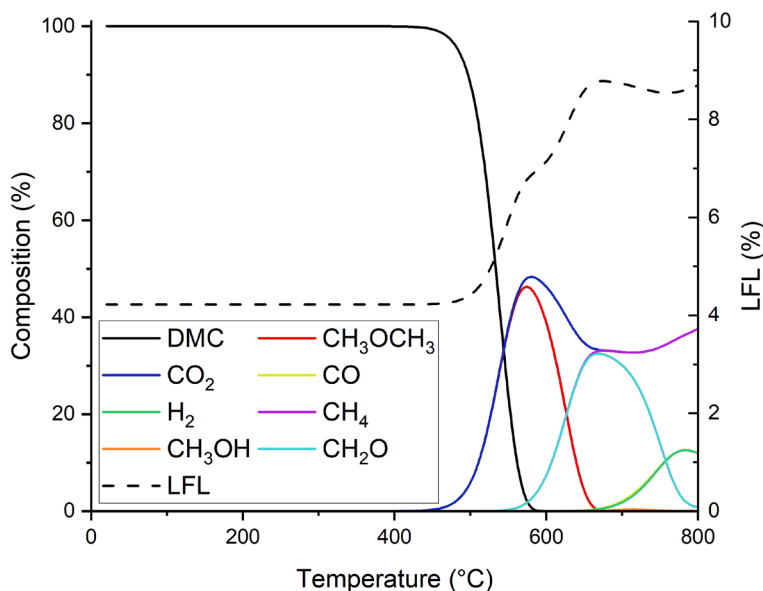


Figure 2. DMC decomposition and predicted off-gas species with calculated LFL of the mixture.

Table 1. LFL of major species present in the off-gas.

Species	LFL (%)	Source
DMC	4.20	(IPCS INCHEM, 2005)
CH <sub>3</sub> OCH <sub>3</sub>	3.35	(Zlochower, 2012)
CO <sub>2</sub>	n/a	(Chen et al., 2020)
CH <sub>3</sub> OH	6.00	(PubChem, Methanol)
H <sub>2</sub>	4.00	(Chen et al., 2020)
CH <sub>4</sub>	5.00	(Chen et al., 2020)
CH <sub>2</sub> O	7.00	(PubChem, Formaldehyde)
CO	12.50	(Chen et al., 2020)

#### 4. Conclusion

This work proves that micro-kinetic modelling can predict gas generation during LIB TR. We have shown the flammability risk can be evaluated using the extended Le Chatelier's formula. In doing so, the LFL of the off-gas is calculated. A large amount of CO<sub>2</sub> increases the LFL of the off-gas. This reduces the flammability hazard but increases the suffocation hazard. The composition of off-gas varies over the temperature range studied.

This work will be extended to full Li-ion cell systems to be practical for LIB developers. Including all electrolyte and electrode materials. Also, this method can be extended to calculate the ignition energy and the heat of reaction of the off-gas. Providing further predictive capabilities for risk assessment of LIB energy storage systems.

#### 5. Conflicts of Interest & Acknowledgements

There are no conflicts to declare. This work was supported by the Faraday Institution [grant number FIRG028].

## References

- P. Bugryniec, J. Davidson, S. Brown, 2020. Advanced abuse modelling of Li-ion cells – A novel description of cell pressurisation and simmering reactions. *Journal of Power Sources*, 474, 228396.
- P. Bugryniec, A. Yearlly, A. Jain, N. Price, S. Vernuccio, S. Brown, 2022. Gaussian-Process based inference of electrolyte decomposition reaction networks in Li-ion battery failure. *Computer Aided Chemical Engineering*, 51, 157-162.
- S. Chen, Z. Wang, J. Wang, X. Tong, W. Yan, 2020. Lower explosion limit of the vented gases from Li-ion batteries thermal runaway in high temperature condition. *Journal of Loss Prevention in the Process Industries*, 63, 103992.
- Y. Fernandes, A. Bry, S. de Persis, 2018. Identification and quantification of gases emitted during abuse tests by overcharge of a commercial Li-ion battery. *Journal of Power Sources*, 389, 106-119.
- Y. Fernandes, A. Bry, S. de Persis, 2019. Thermal degradation analyses of carbonate solvents used in Li-ion batteries. *Journal of Power Sources*, 414, 250-261.
- Gaussian 16, 2019. Revision C.01. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, *et al.*. Gaussian, Inc., Wallingford CT, 2019.
- M. Henriksen, K. Vaagsaether, J. Lundberg, S. Forseth, D. Bjerketvedt, 2019. Explosion characteristics for Li-ion battery electrolytes at elevated temperatures. *Journal of Hazardous Materials*, 371, 1-7.
- E. Hermes, A. Janes, J. Schmidt, 2019. Micki: A python-based object-orientaed microkinetic modeling code. *The Journal of Chemical Physics*, 151 (1), 014112.
- IPCS INCHEM, 2000. Diethyl carbonate. International Chemical Safety Card: 1022. <https://inchem.org/documents/icsc/icsc/eics1022.htm>
- IPCS INCHEM, 2005. Dimethyl carbonate. International Chemical Safety Card: 1080. <https://inchem.org/documents/icsc/icsc/eics1080.htm>
- W. Li, H. Weng, Y. Zhang, M. Ouyang, 2019. Flammability characteristics of the battery vent gas: A case of NCA and LFP lithium-ion batteries during external heating abuse. *Journal of Energy Storage*, 24, 100775.
- E. Ludwig, 1999. Process Safety and Pressure-Relieving Devices, in *Applied Process Design for Chemical & Petrochemical Plants*, 1, 485.
- MATLAB, 2021. version 9.10.0 (R2021a), Natick, Massachusetts: The MathWorks Inc.
- PubChem, Methanol. <https://pubchem.ncbi.nlm.nih.gov/compound/Methanol>
- PubChem, Formaldehyde. <https://pubchem.ncbi.nlm.nih.gov/compound/712>
- V, Schroeder, 2016. Calculation of Flammability and Lower Flammability Limits of Gas Mixtures for Classification Purposes. Bundesanstalt für Materialforschung und –prüfung (BAM), Berlin.
- Q. Wang, P. Ping, X. Zhao, G. Chu, J. Sun, C. Chen, 2012. Thermal runaway caused fire and explosion of lithium ion battery. *Journal of Power Sources*, 208, 210–224.
- Y. Zhang, E. Wang, C. Li, H. Wang, 2022a. 2D Combustion Modeling of Cell Venting Gas in a Lithium-Ion Battery Pack. *Energies*, 15, 15530.
- Q. Zhang, J. Niu, Z. Zhao, Q. Wang, 2022b. Research on the effect of thermal runaway gas components and explosion limits of lithium-ion batteries under different charge states. *Journal of Energy Storage*, 45, 103759.
- I. Zlochower, 2012. Experimental flammability limits and associated theoretical flame temperatures as a tool for predicting the temperature dependence of these limits. *J Loss Prev Process Ind*, 25(3), 555-560.

# Modeling of carbon offset networks for process systems to achieve net zero emissions

Soo Hyoung Choi

*Division of Chemical Engineering, Clean Energy Research Center, Jeonbuk National University, Jeonju, 54896, S. Korea*  
*soochoi@jbnu.ac.kr*

## Abstract

A carbon cycle model based method is proposed for performance evaluation of carbon offset projects that compensate industrial emissions. As a case study, carbon cycle impact assessment was conducted for five representative approaches to carbon neutrality. The result indicates that fossil to biomass fuel switching is doomed to a high concentration of atmospheric carbon dioxide and large scale deforestation, and that reduction of emissions is more effective than extension of sequestration from an environmental point of view. Finally, from a social point of view, optimization of carbon offset networks in terms of both carbon cycle dynamics and carbon market economics is suggested as future work.

**Keywords:** carbon neutrality, net zero emissions, carbon offset, carbon credit, carbon cycle impact assessment

## 1. Introduction

Carbon neutrality is an arithmetic concept that can be achieved by net zero emissions, which only guarantees a steady state for the carbon cycle (Manousiouthakis and Choi, 2021). If emissions and absorptions are equally large, the steady state concentration of carbon dioxide in the atmosphere can still be high even though carbon neutrality is attained. A typical example is harvested plant biomass, which inherently owes a huge carbon debt (Choi and Manousiouthakis, 2020) when combusted, and thus cannot be a proper alternative to fossil fuels (Manousiouthakis and Choi, 2021). Furthermore, emissions can be overcompensated by carbon offsets (Kotchen, 2009) such as renewable energy and energy efficiency projects that only indirectly reduce emissions. Therefore, carbon cycle impact assessment is necessary when a carbon offset project is considered as an approach to carbon neutrality.

Chemical plant industries, mostly carbon based, are on the way to carbon neutrality, reducing emissions by carbon capture and utilization (CCU), carbon capture and storage (CCS), and by growing and processing of biomass (Gabrielli et al., 2020), and earning or purchasing carbon credits (Blaufelder et al., 2021) that can be obtained by carrying out or investing in carbon offset projects such as forestry, renewable energy, energy conservation, and conversion of waste to energy (EIC, 2020). Among these projects, forestry is essential, because it directly controls atmospheric carbon dioxide. Forests should be protected because deforestation is equivalent to carbon emissions (Photopoulos, 2017), and one to one reforestation cannot fully recover the initial state of atmospheric carbon dioxide as long as deforestation continues (Choi and Manousiouthakis, 2020). Besides, as an alternative to forestry, direct air capture (DAC) is receiving attention (Chatelain, 2022) with skepticism on its efficiency (Day, 2021).

Carbon offsetting is criticized, especially strongly by environmentalists, because it does not directly reduce emissions (Ghussain, 2020). This issue could be moderated by corrective evaluation of carbon credits, reflecting the fact that less emissions are better than more sequestration. Therefore, carbon cycle impact assessment is suggested for emissions trading system (ETS) also. In this work, a model is proposed as shown in Fig. 1, as a tool for performance evaluation of carbon offset projects. The solid arrows represent main active flows involved in carbon cycle and carbon offsetting. The dashed arrows represent flows that are much less dynamic than the others. The model indicates that the total amount of carbon credits, i.e., allowances of carbon emissions, should be limited to the total capacity of sequestration by forests in order to achieve net zero emissions. Furthermore, the prices of carbon credits should be sufficiently high in order to drive the industry to substantially reduce emissions (Opanda, 2022).

## 2. Proposed Method

### 2.1. Modeling of Carbon Cycle

Carbon cycle impact assessment is required in order to rigorously evaluate carbon offset projects in terms of contribution to climate change mitigation. Proposed carbon cycle model equations are as follows:

$$\dot{C}_a = f_{rp}(C_p) + f_{rs}(C_s) - f_p(C_a, C_p, A_v) - f_a(C_a, C_{so}) + r_v + q(t) \tag{1}$$

$$\dot{C}_p = f_p(C_a, C_p, A_v) - f_{rp}(C_p) - f_l(C_p) - h(t) \tag{2}$$

$$\dot{C}_s = f_l(C_p) - f_{rs}(C_s) - f_t(C_s) \tag{3}$$

$$\dot{C}_{so} = f_a(C_a, C_{so}) + f_t(C_s) + f_{up}(C_{do}) - f_{dw}(C_{so}) \tag{4}$$

$$\dot{C}_{do} = f_{dw}(C_{so}) - f_{up}(C_{do}) - f_s(C_{do}) \tag{5}$$

$$\dot{C}_r = f_s(C_{do}) - r_v \tag{6}$$

$$\dot{A}_v = \left[ \int_0^t s(t - \tau) \dot{g}(\tau) d\tau - h(t) \right] \frac{A_v}{C_p} \tag{7}$$

where

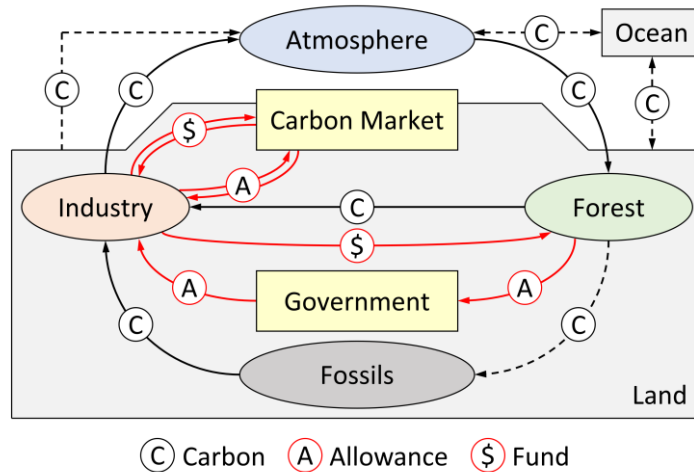


Figure 1. Carbon cycle and emissions trading system.

- $C_a$  = mass of carbon in the atmosphere, PgC  
 $C_p$  = mass of carbon in plants, PgC  
 $C_s$  = mass of carbon in soils, PgC  
 $C_{so}$  = mass of carbon in the surface ocean, PgC  
 $C_{do}$  = mass of carbon in the deep ocean, PgC  
 $C_r$  = mass of carbon in rocks, PgC  
 $A_v$  = normalized vegetated land area  
 $f_p$  = rate of photosynthesis, PgC/y  
 $f_{rp}$  = rate of plant respiration, PgC/y  
 $f_l$  = rate of litterfall, PgC/y  
 $f_{rs}$  = rate of soil respiration, PgC/y  
 $f_t$  = rate of transfer from soils to the surface ocean by the river flow, PgC/y  
 $f_a$  = net rate of absorption to the surface ocean, PgC/y  
 $f_{up}$  = rate of upwelling from the deep ocean, PgC/y  
 $f_{dw}$  = rate of down-welling from the surface ocean, PgC/y  
 $f_s$  = rate of sedimentation to the crust, PgC/y  
 $r_v$  = rate of volcano emission, 0.1 PgC/y  
 $q$  = rate of carbon emissions by fuel combustion, PgC/y  
 $h$  = rate of plant harvesting, PgC/y  
 $s$  = target rate of biomass production by planting seedlings, PgC/y  
 $g$  = normalized tree mass growth function,  $g(0) = 0$ ,  $g(\infty) = 1$

The unit PgC represents  $10^{15}$  grams of carbon, which is equivalent to GtC. The initial conditions at present are as follows:

$$\begin{aligned}
 C_a(0) &\approx 1.5C_a^\circ, C_p(0) \approx C_p^\circ, C_s(0) \approx C_s^\circ, C_{so}(0) \approx 1.02C_{so}^\circ, C_{do}(0) \approx C_{do}^\circ, \\
 C_r(0) &\approx C_r^\circ, A_v(0) \approx 0.75
 \end{aligned}$$

where the parameters with degree symbols represent reference values estimated for the time right before the industrial revolution. Detailed expressions of the functions in the above equations and model validation are referred to Choi and Manousiouthakis (2022).

## 2.2. Simulation of Carbon Offset Networks

Simultaneous implementation of multiple carbon offset projects can be viewed as a carbon offset network. The resulting rate of carbon emissions can be calculated as follows:

$$q(t) = \sum_{i=0}^n q_i(t) \tag{8}$$

where  $q_i < 0$  represents the reduced rate of emissions by the  $i$ -th project, which should be close to the annually earned carbon credit if properly evaluated. Rates of plant harvesting and planting seedlings can be calculated in the same way as follows:

$$h(t) = \sum_{i=0}^n h_i(t) \tag{9}$$



$$s(t) = \sum_{i=0}^n s_i(t) \tag{10}$$

Let us assume that harvested plant biomass is partially combusted as a fuel, and the rest is stored as a product. Then, carbon neutrality is achieved at  $t = t^*$  if the following condition is satisfied:

$$q(t) = h(t), t \geq t^* \tag{11}$$

Therefore, the above three functions  $q$ ,  $h$ , and  $s$  compose design variables for a carbon offset network to be optimized subject to the following constraints:

$$C_a \leq C_a^* \tag{12}$$

$$A_v \leq A_v^* \tag{13}$$

where  $C_a^*$  is set by the maximum allowable increase in the global temperature, and  $A_v^*$  by the maximum available land area for forestry.

### 3. Case Study

#### 3.1. Approaches to Carbon Neutrality

Consider a hypothetical task of achieving carbon neutrality in 25 years with initial emissions of 10 PgC/y, which approximately corresponds to the current situation. Five approaches are summarized in Table 1, and six cases are defined as follows:

Case 0. Fossil fuels

As a base case, it is assumed that fossil fuels are used as ever.

$$q(t) = 10, h(t) = s(t) = 0 \tag{14}$$

Case 1. Fossil to biomass with reforestation

Fossil fuels are gradually replaced by harvested plant biomass with 1:1 reforestation.


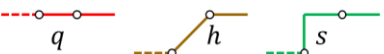
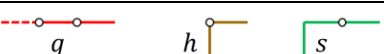
$$q(t) = 10, h(t) = s(t) = \begin{cases} 0.4t, & 0 \leq t < 25 \\ 10, & t \geq 25 \end{cases} \tag{15}$$

Case 2. Fossil to biomass with afforestation

Fossil fuels are gradually replaced by harvested plant biomass with 1:1 afforestation.

$$q(t) = 10, h(t) = \begin{cases} 0.4t, & 0 \leq t < 25 \\ 10, & t \geq 25 \end{cases}, s(t) = 10 \tag{16}$$

Table 1. Changes in design variables for different approaches to carbon neutrality

Case	Approach to carbon neutrality	Changes in design variables
1	Fossil to biomass with reforestation	
2	Fossil to biomass with afforestation	
3	Extension of sequestration	

4	Reduction of emissions	
5	Half reduction and half afforestation	

Case 3. Extension of sequestration

Fossil fuels are used while sequestration is extended by 1:1 afforestation.

$$q(t) = 10, h(t) = \begin{cases} 0, & 0 \leq t < 25 \\ 10, & t \geq 25 \end{cases}, s(t) = 10 \quad (17)$$

Case 4. Reduction of emissions

Fossil fuels are gradually replaced by non-carbon based energy.

$$q(t) = \begin{cases} 10 - 0.4t, & 0 \leq t < 25 \\ 0, & t \geq 25 \end{cases}, h(t) = s(t) = 0 \quad (18)$$

Case 5. Half reduction and half afforestation

Fossil fuels are gradually reduced to half, and sequestered by 1:1 afforestation.

$$q(t) = \begin{cases} 10 - 0.2t, & 0 \leq t < 25 \\ 5, & t \geq 25 \end{cases}, h(t) = \begin{cases} 0, & 0 \leq t < 25 \\ 5, & t \geq 25 \end{cases}, s(t) = 5 \quad (19)$$

3.2. Performance Evaluation

Carbon cycle impact assessment was carried out for the above six cases. The results are as shown in Fig. 2. Cases 1 and 2 indicate that fossil to biomass fuel switching is an improper approach to carbon neutrality because of huge carbon debts, i.e., the initial atmospheric carbon dioxide that exceeds that of fossil fuels, and large loss of forests. Cases 3 and 4 show that extension of sequestration cannot replace reduction of emissions in effect. Case 5 is a compromise between cases 3 and 4, which is considered to be a feasible approach whose impacts on the carbon cycle are acceptable.

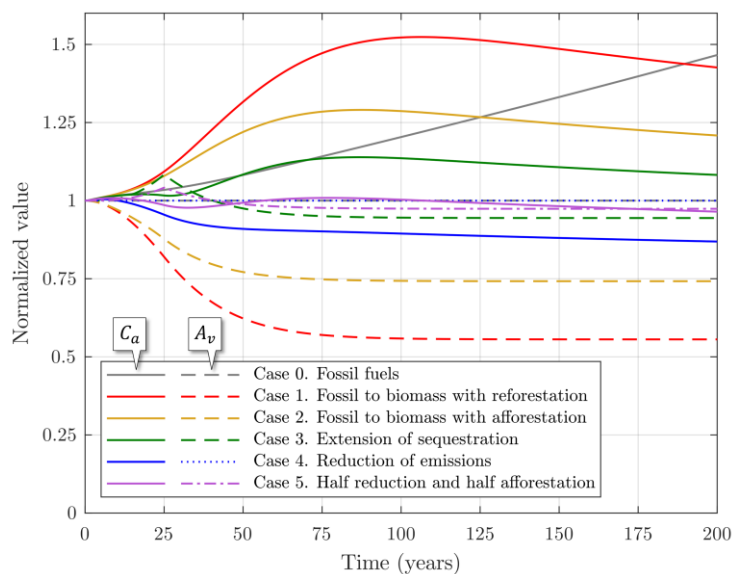


Figure 2. Predicted carbon cycle impacts of different approaches to carbon neutrality.

#### 4. Conclusions

The result of case study shows that net zero emissions can result in various levels of concentration of carbon dioxide in the atmosphere. Therefore, it is suggested that the concept of carbon neutrality should be more conservatively redefined in view of dynamic effects, and the credits for carbon offset projects should be more rigorously revaluated based on carbon cycle impact assessment. Then, the proposed method is expected to be able to contribute to optimal design of carbon offset networks in terms of not only carbon cycle dynamics, but also carbon market economics, for process systems to achieve net zero emissions within a given period of time.

#### Acknowledgements

This work was supported by National Research Foundation of Korea (NRF) funded by Ministry of Science and ICT (MSIT) (Grant No. 2022R1F1A1062901).

#### References

- C. Blaufelder, C. Levy, P. Mannion, and D. Pinner, 2021, A blueprint for scaling voluntary carbon markets to meet the climate challenge, <https://www.mckinsey.com/capabilities/sustainability/our-insights/a-blueprint-for-scaling-voluntary-carbon-markets-to-meet-the-climate-challenge>
- R. Chatelane, 2022, Companies set sights on removing troublesome CO<sub>2</sub> from air, <https://www.ny1.com/nyc/all-boroughs/news/2022/03/30/companies-set-sights-on-removing-troublesome-co2-from-air>
- S. H. Choi and V. I. Manousiouthakis, 2020, On the carbon cycle impact of combustion of harvested plant biomass vs. fossil carbon resources, *Computers and Chem. Eng.*, 140, 106942. <https://doi.org/10.1016/j.compchemeng.2020.106942>
- S. H. Choi and V. I. Manousiouthakis, 2022, Modeling the carbon cycle dynamics and the greenhouse effect, *IFAC-PapersOnLine*, 55, 7, 424-428. <https://doi.org/10.1016/j.ifacol.2022.07.480>
- L. Day, 2021, Carbon sequestration as a service doesn't quite add up, <https://hackaday.com/2021/10/15/carbon-sequestration-as-a-service-doesnt-quite-add-up/>
- EIC, 2020, 4 types of carbon offset projects, <https://www.eic.co.uk/4-types-of-carbon-offset-projects/>
- P. Gabrielli, M. Gazzani, and M. Mazzotti, 2020, The role of carbon capture and utilization, carbon capture and storage, and biomass to enable a net-zero-CO<sub>2</sub> emissions chemical industry, *Ind. Eng. Chem. Res.*, 59, 7033-7045. <https://doi.org/10.1021/acs.iecr.9b06579>
- A. A. Ghussain, 2020, The biggest problem with carbon offsetting is that it doesn't really work, <https://www.greenpeace.org.uk/news/the-biggest-problem-with-carbon-offsetting-is-that-it-doesnt-really-work/>
- M. J. Kotchen, 2009, Offsetting green guilt, *Stanford Social Innovation Review*, 7, 2, 26-31. <https://doi.org/10.48558/515j-d290>
- V. I. Manousiouthakis and S. H. Choi, 2021, A carbon cycle optimization method for fossil and biomass energy utilization, *Korean J. Chem. Eng.*, 38, 10, 2003-2008. <https://doi.org/10.1007/s11814-021-0899-9>
- S. Opanda, 2022, Carbon credit pricing chart: updated 2022, <https://8billiontrees.com/carbon-offsets-credits/new-buyers-market-guide/carbon-credit-pricing/>
- J. Photopoulos, 2017, Counting the carbon cost of forest destruction, <https://ec.europa.eu/research-and-innovation/en/horizon-magazine/counting-carbon-cost-forest-destruction>

# Identifying first-principles models for bubble column aeration using machine learning

Peter Jul-Rasmussen<sup>a</sup>, Arijit Chakraborty<sup>b</sup>, Venkat Venkatasubramanian<sup>b</sup>,  
Xiaodong Liang<sup>a</sup>, Jakob Kjøbsted Huusom<sup>a</sup>

<sup>a</sup>*Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søtofts Plads 228A, 2800 Kgs. Lyngby, Denmark*

<sup>b</sup>*Complex Resilient Intelligent Systems Laboratory, Department of Chemical Engineering, Columbia University, New York, NY 10027, USA*

## Abstract

Mass transfer of oxygen is investigated in this work using a pilot-scale bubble column unit with a two-fluid nozzle for aeration. First-principles models for the bubble column unit are identified by utilizing concepts in artificial intelligence (AI) and machine learning (ML), and applying the same to experimental data. By combining process knowledge with data-driven modeling, we discovered interpretable models for oxygen transport phenomena in bubble columns. By virtue of obtaining symbolic models, it is possible to perform post-hoc analyses on the same in order to gain physical insights into the mechanisms occurring in the system -- a convenience lost when using black-box models such as neural networks. This provides valuable understanding which can be applied when modeling more complex systems such as fermentation processes.

**Keywords:** Artificial intelligence; Machine Learning; Hybrid AI; Interpretable models; Mechanistic modeling

## 1. Introduction

With the advent of improved computing power, and the ability to generate and handle copious amounts of data, artificial intelligence (AI) has yielded successes in the domains of computer vision, game-playing, recommender systems, among other applications. It must be noted however, that such domains do not have guiding mathematical equations that determine their existence. In stark contrast, physicochemical phenomena, are governed by constitutive equations and conservation laws, among other functional relationships (Venkatasubramanian 2019). Further, knowledge about these phenomena is known in many cases, and can be leveraged to better describe the same. Consequently, when attempting to derive mathematical models of such systems using data-driven tools, it is imperative that one utilize any prior knowledge available about the system, and incorporate the same into the conventional data-driven modeling framework (Chakraborty et al. 2022). This hybrid approach results in simpler, more interpretable models, and can yield models that can provide mechanistic insights about the system.

The majority of industrial fermentation processes are aerobic i.e. they require oxygen to be supplied for the microbial metabolic processes to take place. Oxygen is supplied through aeration, in which air is bubbled through the fermentation broth, facilitating a mass transfer of oxygen from the bulk gas phase through the gas-liquid interface to the bulk liquid phase (Doran 2013). Typically modeling of the oxygen transfer in fermentation processes is based on the difference between the dissolved oxygen concentration at saturation and the dissolved oxygen concentration in the fermentation

broth as driving force, an overall liquid-phase mass transfer coefficient, and the surface area of the gas-liquid interface. Experimentally it is difficult to separate the overall liquid-phase mass transfer coefficient from the surface area of the gas-liquid interface, therefore these two terms are generally combined to a volumetric mass transfer coefficient, which can be determined through methods such as the sulphite oxidation technique and gassing-out techniques (Stanbury et al. 1995). Prediction of the volumetric mass transfer coefficient has typically relied on empirical correlations, while theoretical correlations have been introduced using Higbie's penetration theory and Kolmogoroff's theory of isotropic turbulence (Kawase et al. 1992).

In this work, we attempt to obtain an interpretable model of a bubble column aeration system. The data used for modeling has been experimentally generated under 6 different conditions. We propose a hierarchical modeling approach by broadly classifying the same into 2 operating regimes - with, and without addition of salt (here,  $\text{MgSO}_4$ ). This enables us to obtain high fidelity models accounting for the nonidealities that are manifested as a result of (or lack thereof) salt addition. In the remainder of the manuscript, we describe the bubble column aeration system in Section 2, and briefly mention the model discovery engine used to obtain the models in Section 3. We then discuss the results when compared to purely first-principles models, in Section 4. Finally, we conclude our findings and, highlight potential future directions to investigate.

## 2. Problem Description

The experimental set-up consists of a main column, a recycle loop for circulating liquid in the system, a compressed air supply, and a  $\text{H}_2\text{O}_2$  solution storage tank (Figure 1). Aeration in the bubble column is introduced using a two-fluid nozzle, in which air and recycled liquid are premixed before entering the column. The recycle flow in the system is enforced using a centrifugal pump, and a  $\text{H}_2\text{O}_2$  solution is fed to the system either in the bottom of the column, in the top of the column, or through the aeration nozzle using a variable speed dosing pump. Online sensor measurements are accessible through a Supervisory Control And Data Acquisition (SCADA) system and stored in a central data base. Online sensor measurements in the system includes: flow sensors for liquid recirculation ( $FI_1$ ) and compressed air ( $FI_2$ ), dissolved oxygen sensors at bottom ( $AI_1$ ) and top ( $AI_2$ ) of the column, temperature sensor ( $TI_1$ ), conductivity sensor ( $CT_1$ ), level sensor ( $LI_1$ ), and weight measurements of the  $\text{H}_2\text{O}_2$  solution storage tank ( $WT_1$ ). The system is used to investigate oxygen mass transfer and the influence of air flow, liquid recycle flow, and salts on the mass transfer. The hydrogenperoxide-catalase method is employed in the system, in which  $\text{H}_2\text{O}_2$  is decomposed to water and oxygen in the presence of an enzyme (Eq. 1).

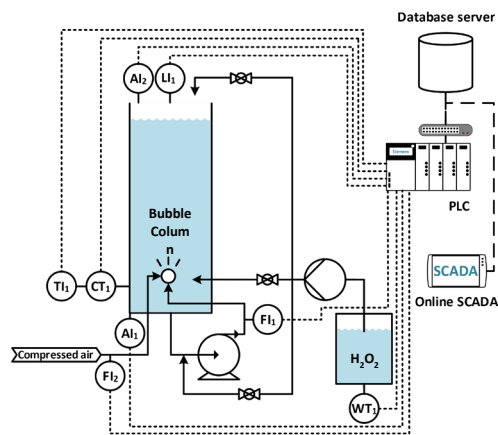


Figure 1 Experimental bubble column set-up with live measurements accessible through SCADA system.

The oxygen liberated in reaction (Eq. 1) will be dissolved in the liquid phase, effectively making the solution supersaturated. The oxygen transfer is thus going to be from the liquid phase to the gas phase. A model for the oxygen transfer in the bubble column is given in (Eq. 2) assuming that the conversion through reaction (Eq. 1) is instantaneous, irreversible, and that the column is ideally mixed.

$$\frac{dO}{dt} = \frac{r}{2} - K_L a(O - O^*) \tag{2}$$

In which  $O$  is the dissolved oxygen concentration,  $O^*$  is the dissolved oxygen concentration at saturation,  $K_L$  is the overall liquid-phase mass transfer coefficient, and  $a$  is the surface area of the gas/liquid interface.  $r = c_{H_2O_2}m/V$  is the rate of added  $H_2O_2$  per column volume,  $c_{H_2O_2}$  is the concentration of  $H_2O_2$ ,  $m$  is the volumetric flowrate of  $H_2O_2$ , and  $V$  is the volume of the column. Integrating (Eq. 2) from  $t_0$  to  $t$  using  $O(t_0) = O_0$  yields the analytical solution (Eq. 3).

$$O = O^* + \frac{c_{H_2O_2}}{2K_L aV}m - \frac{c_{H_2O_2}}{2K_L aV}me^{-K_L a(t-t_0)} - O^*e^{-K_L a(t-t_0)} + O_0e^{-K_L a(t-t_0)} \tag{3}$$

In the system, the mass transfer coefficient and the surface area of the gas/liquid interface is a function of the liquid recirculation flowrate ( $F_1$ ), the air flowrate ( $F_2$ ), and the presence of salts in the liquid. A total of six experimental runs were performed for the system using the experimental conditions given in Table (1).

Table 1 Experimental conditions used for the bubble column system

Experiment	$F_1$ (l/min)	$F_2$ (l/min)	Salt
1	0	300	None
2	0	600	None
3	100	300	None
4	100	600	None
5	0	300	MgSO <sub>4</sub>
6	100	300	MgSO <sub>4</sub>

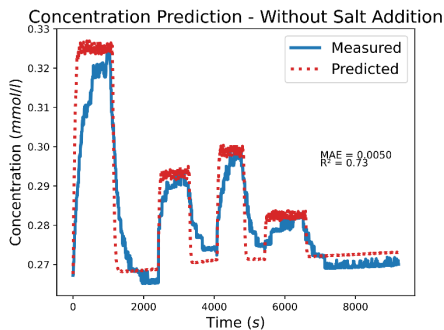


Figure 2 Model predictions using analytic model and measured oxygen concentrations for experiments in which no MgSO<sub>4</sub> was added.

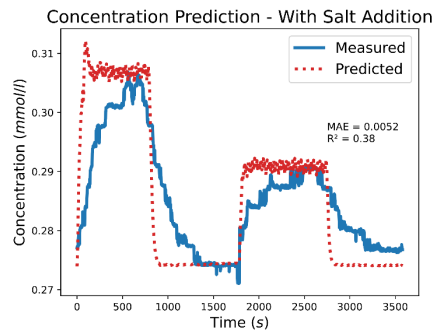


Figure 3 Model predictions using analytic model and measured oxygen concentrations for experiments in which MgSO<sub>4</sub> was present in the system.

By performing a post-hoc analysis of the data from the six experiments, steady-state  $K_L a$  values can be determined and introduced as a constant in (Eq. 3) to model the dissolved oxygen concentration for each of the six experiments.

Predictions of the experiments using the analytic model reveals that the full system dynamics are not captured in the deterministic model (Figure 2 and 3). This suggests that some of the assumptions in the deterministic model should be relaxed to account for the unmodelled phenomena. However, deriving a deterministic expression for the unmodelled phenomena is a challenge, as it includes complicated enzyme kinetics and mixing phenomena dependent on the configuration and geometry of the bubble column. Due to the complicated nature of the unmodelled phenomena in the system, it is considered to be unknown in this work.

### 3. Methods

The data obtained from the pilot-plant is used to discover interpretable models that bear some resemblance to the analytical solution, whilst also being able to account for the nonidealities in the system. A genetic algorithm-based approach - titled *AI-DARWIN* (Chakraborty et al. 2020, 2021) - was used for the same. This approach takes a list of permitted tractable function transformations for each of the inputs, and extends the concept of genetic algorithm (Koza 1995) to mathematical models, in an attempt to explore the possible search space of different possible models. The advantage of such an approach is that it searches through the space of potential models with only the user-permitted function transformations. This allows one to curb the complexity of the models obtained, and allows for inclusion of prior knowledge into the modeling exercise.

The bubble column is a complex dynamical system with a multitude of disparate operating regimes. Consequently, a hierarchical modeling approach was used due to the low variation in the experimental conditions. Two different models were determined: one for the system without  $\text{MgSO}_4$ , and one for the system containing  $\text{MgSO}_4$ .

### 4. Results

The models discovered using *AI-DARWIN* (Table 2 and 3) contain several features that are either a direct analogue, or are similar to features used in the analytic solution (Eq. 3).

Table 2 Models predicting oxygen concentration discovered using *AI-DARWIN* with standardized features

$\text{MgSO}_4$	<i>AI-DARWIN</i> output	Fitness (MAE and $R^2$ )
No	$O = \beta_0 + (\beta_1(-t)^2 + \beta_2 F_1 + \beta_3 F_2)m$ $+ \beta_4 m e^{-t} + \beta_5 (O^*)^2 e^{-t} + \beta_6 (-t)O^*$ $+ \beta_7 (-t)(O^*)^2 + \beta_8 F_2 O^* + \beta_9 F_2^2 (O^*)^2$ $+ \beta_{10} e^{-2t} + \beta_{11} e^{-t}$	0.0016 ( $R^2 = 0.97$ )
Yes	$O = \beta_0 + (\beta_1 m (O^*)^2 + \beta_2 m)m$ $+ \left( \frac{\beta_3 (-t)}{O^*} + \beta_4 (-t)^2 + \frac{\beta_5 (-t)^3}{O^*} \right) O^*$ $+ \beta_6 F_1 e^{F_1} + \beta_7 O^* + \beta_8 (O^*)^2 F_2^2$ $+ \beta_9 (O^*)^2 (-t)^2 + \beta_{10} e^{-t} + \beta_{11} e^m$	0.0011 ( $R^2 = 0.98$ )

Terms containing the features  $m$  and  $e^{-t}$  are found in all models, while the Taylor series expansion  $(-t + (-t)^2 + (-t)^3)$  is found as a feature in the model for experiments with  $\text{MgSO}_4$  discovered using *AI-DARWIN*. The intercept ( $\beta_0$ ) is approximately the same

as the initial value of  $O^*$  for both the discovered models. Additional terms correlating the features  $F_1$  and  $F_2$  to the output  $O$  appear in the discovered models, a correlation handled in the analytic solution by the  $K_L a$  values determined from post-hoc analysis. The extra terms found in the discovered models are accounting for the unknown phenomena in the bubble column.

Using the discovered models for predicting the concentration of dissolved oxygen in the bubble column it is found that the dynamic behavior in the system is significantly better represented compared to the analytic solution, as indicated by the improved fitness (Figure 4 and 5). The improved performance of the discovered models indicates that the unknown phenomena is successfully accounted for in the six different experiments in the bubble column system.

Table 2 Parameters for models predicting oxygen concentration discovered using *AI-DARWIN* with standardized features

MgSO <sub>4</sub>	No	Yes
$\beta_0$	0.26	0.29
$\beta_1$	-0.014	-0.000
$\beta_2$	-0.002	-0.004
$\beta_3$	-0.003	0.013
$\beta_4$	0.005	0.004
$\beta_5$	-0.015	-0.003
$\beta_6$	-0.004	-0.000
$\beta_7$	0.021	-0.009
$\beta_8$	-0.003	0.000
$\beta_9$	0.016	-0.001
$\beta_{10}$	-0.003	-0.003
$\beta_{11}$	0.017	0.003

Table 3 Features in deterministic model and selected features from the models discovered using *AI-DARWIN*

Model	Feature				
Deterministic	$O^*$	$m$	$me^{-(t-t_0)}$	$O^*e^{-(t-t_0)}$	$O_0e^{-(t-t_0)}$
<i>AI-DARWIN</i> w/o. MgSO <sub>4</sub>	$\beta_0$	$m$	$me^{-t}$	$(O^*)^2e^{-t}$	-
<i>AI-DARWIN</i> w. MgSO <sub>4</sub>	$\beta_0$	$m$	-	$O^*(-t + (-t)^2 + (-t)^3)$	-

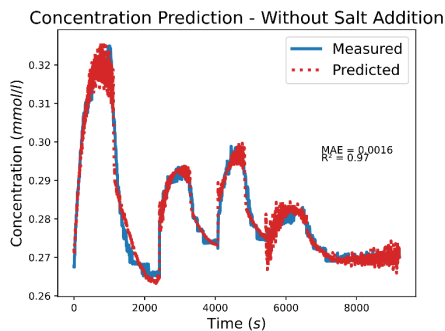


Figure 4 Predicted and measured oxygen concentrations for experiments 1, 2, 3, and 4 in which no MgSO<sub>4</sub> was added.

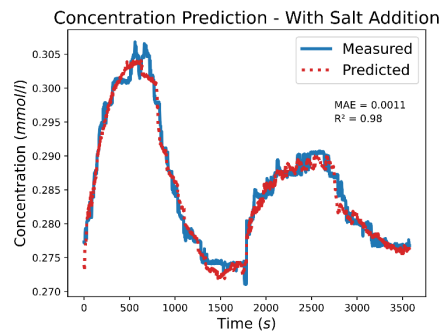


Figure 5 Predicted and measured oxygen concentrations for experiments 5 and 6 in which MgSO<sub>4</sub> was present in the system.

An interesting observation is the inclusion of certain terms which exist in the analytical solution, in the models predicted by *AI-DARWIN*. While some complex nonlinear parameterized features are represented through a Taylor Series expansion, other novel terms are included to account for nonidealities in the system. We encounter the interaction effects of the different input variables which yield more complex function



transformations, but result in better predictability. While it would be the desired goal to be able to decouple the effects of the ideal conditions as expressed by the analytical solution, and the nonidealities which are accounted for by virtue of the additional terms, it remains an ongoing challenge in the interpretability of ML models in the sciences and engineering.

## 5. Conclusion

A step forward in the direction of creating a digital-twin for a complex bubble column system has been made, by virtue of an interpretable modeling approach. Such a technique relied on the use of *a priori* knowledge of the system's dynamics, and was able to account for the analytical solution obtained from first-principles. Such a data-driven approach combined with first-principles permits interpretability, which is not possible in purely black-box models (Chakraborty et al. 2022). Due to the high cost of performing experiments, varied set of conditions for some input variables ( $F_1$  and  $F_2$ ) have not been considered in this work (at most 2 conditions were considered). Future steps include sampling of the data at different operating conditions using appropriate experimental design techniques. Incorporating this into the overall modeling framework can help us obtain more reliable and robust models. Certain key underlying assumptions of the first-principles-based modeling approach include the instantaneity of Reaction 1 and the bubble column being well-mixed -- both of which do not seem to hold upon performing the experiments. As a result, the  $K_L a$  values have been assumed to be constant at given conditions when modeling using first-principles. It would be worthwhile to explore the possible function form of the mass transfer coefficient under different operating regimes.

## Acknowledgements

This study was financially supported by the Sino-Danish Center for Education and Research (SDC).

## References

- A. Chakraborty, S. Serneels, H. Claussen, V. Venkatasubramanian, 2022. Hybrid ai models in chemical engineering—a purpose-driven perspective. *Computer Aided Chemical Engineering* 51, 1507–1512.
- A. Chakraborty, A. Sivaram, L. Samavedham, V. Venkatasubramanian, 2020. Mechanism discovery and model identification using genetic feature extraction and statistical testing. *Computers & Chemical Engineering* 140, 106900.
- A. Chakraborty, A. Sivaram, V. Venkatasubramanian, 2021. Ai-darwin: A first principles-based model discovery engine using machine learning. *Computers & Chemical Engineering* 154, 107470
- P. M. Doran, 2013, Chapter 10 - mass transfer, *Bioprocess Engineering Principles* (Second Edition), Academic Press, London, 379–444.
- Y. Kawase, B. Halard, M. Mooyoung, 1992. Liquid-phase mass transfer coefficients in bioreactors. *Biotechnology and Bioengineering* 39 (11), 1133–1140.
- J. R. Koza, 1995. Survey of genetic algorithms and genetic programming. In: *Wescon conference record*. Western Periodicals Company, pp. 589–594.
- P. F. Stanbury, A. Whitaker, S. J. Hall, 1995, Chapter 9 - aeration and agitation, *Principles of Fermentation Technology* (Second Edition), Pergamon, Amsterdam, 243–275.
- V. Venkatasubramanian, 2019. The promise of artificial intelligence in chemical engineering: Is it here, finally? *AIChE Journal* 65 (2), 466–478.

# Optimization framework based on a sensitivity analysis for the identification of the critical design variables

Maria de los A. Villarreal-de-Aquino,<sup>a</sup> Jaime D. Ponce-Rocha,<sup>b</sup> Eduardo S. Perez-Cisneros,<sup>c</sup> Verónica Rodríguez-López,<sup>d</sup> Edgar I. Murillo-Andrade,<sup>c</sup> Divanery Rodriguez-Gomez,<sup>e</sup> Ricardo Morales-Rodriguez<sup>a\*</sup>

<sup>a</sup>*Departamento de Ingeniería Química, Universidad de Guanajuato, Noria Alta S/N, Noria Alta, Guanajuato, Guanajuato, 36050, México. \*ricardo.morales@ugto.mx*

<sup>b</sup>*Dirección de Ingeniería y construcción de plantas, CIATEQ, A.C. Centro de Tecnología Avanzada, Av. del Retablo 150, Constituyentes Fovissste, Querétaro, Querétaro, 76150, México*

<sup>c</sup>*Departamento de Ingeniería de Procesos e Hidráulica, Universidad Autónoma Metropolitana – Iztapalapa, San Rafael Atlixco No. 186, Col. Vicentina, Iztapalapa, Ciudad de México, 09340, México*

<sup>d</sup>*Facultad de Farmacia, Universidad Autónoma del Estado de Morelos, Av. Universidad 1001 Chamilpa, Cuernavaca, Morelos, 62209, México*

<sup>e</sup>*Coordinación de Ingeniería Bioquímica, Instituto Tecnológico Superior de Irapuato, Carretera Irapuato-Silao Km 12.5, Irapuato, Guanajuato 36821, Guanajuato, México.*

## Abstract

In this work, a systematic framework was developed and implemented to optimize large scale processes, through a particular case study. In this process, 32 operating and design variables were found, which can be subject to modification using an optimization process. The identification of high impact variables was carried out through a sensitivity analysis using the method of standardized regression coefficients, together with the Latin hypercube method. Once these variables were identified, the optimization of the process was carried out for the maximization of profits together with the minimization of the operating costs using only the most important identified variables. The solution of the process model was done using Aspen Plus and the optimization was performed using the technique of genetic algorithms, which is available in MATLAB. The results illustrated that was possible to increase the profit by 5.02 % and decrease the energy cost 4.6 %. The framework allowed to reduce the computational time 34.2 % compared when all the manipulating variables were used in the optimization task.

**Keywords:** Latin hypercube method, sensitivity analysis, phytosterols production.

## 1. 1. Introduction

The development of new products and their manufacturing processes have demanded innovative manners to solve those challenges (Pistikopoulos et al., 2021). Process optimization has been employed as a paramount strategy to improve the performance in the development of new products and processes, through the introduction of new optimization strategies, advanced methods, combinations of computer-aided tools, etc.

Process optimization most of the time has been categorized as time consuming task, especially when metaheuristic methods are employed, such as genetic algorithms. The implementation of optimization in large and new processes could rise the complexity in the solution, since the number of variables to optimize could be large increasing the time to obtain the optimal solutions (Katok et al., 2021). In a previous work, Ponce-Rocha et al. (2022) proposed a rigorous process design optimization under uncertainty combining Aspen Plus-Matlab, using genetic algorithms (GA) obtaining remarkable results, and showing the functionality of the combination of the computer-aided tools. GA method uses a set of initial solutions (initial populations), randomly selected from the feasible solution space. The fitness function is evaluated for each solution, and the solutions are consequently ranked. Then, the population evolves through several operations, such as reproduction, crossover, and mutation to optimize the fitness function and obtain the final optimal solution. The process is repeated until some criteria are satisfied. This evolutionary algorithm is preferred to classical optimization approaches because it can handle the nonlinear, nonconvex, and nonsmooth optimization problem of the component sizing for the hybrid system (Katok et al., 2021); the nonconvexity of the problem makes it difficult for classical optimization methods to obtain a global optimum. On the other hand, GA globally searches the domain of possible solutions for an optimal solution. Therefore, the objective of this work is to present a systematic framework to optimize large processes, decreasing optimization time, maximizing profits, and decreasing operating costs, where the first step is the identification of the most significant variables of the process, followed by optimization using GA to find the best values of these variables using the combination of Aspen Plus and Matlab. The employed case study was the production of phytosterols and biodiesel from crude vegetable oils.

## 2. Methodology

The framework includes three main sections applied to one case study of relevant interest.

### 2.1. Case study

The production of phytosterols ( $\beta$ -sitosterol and stigmasterol), glycerol and clean biodiesel from crude vegetable oils was taken as a case study. The integrated process consists of two main sections as illustrated in Figure 1: A) Phytosterols production: the separation of triglycerides from phytosterols, phyto-glucoside compounds and fatty acids (FFA) is performed employing supercritical extraction using  $\text{CO}_2$  and a mixture of  $\text{CO}_2$  –  $\text{CH}_3\text{OH}$  as supercritical fluids and vacuum distillation processes. B) Biodiesel production: this is performed using a reactive distillation column to carry out the transesterification reactions of triglycerides to produce clean biodiesel and glycerol, followed by two distillation columns. It is possible to identify diverse important designs and operating variables in the process.

### 2.2. Generation of the process diagram in Aspen Plus and the connection with MATLAB.

Firstly, the process flow diagram is generated using a process simulator, in this case Aspen Plus can be used as simulation environment. Then, the connection between MATLAB® and Aspen plus® can be made using COM technology, which allows making the bridge between these two computer-aided tools. For the case study mentioned above, 32 variables were identified for case A-B and for case A only 13 variables were distinguished.

### 2.3. Sensitivity analysis.

As second step, a sensitivity analysis is performed where the Latin hypercube method can be used to obtain different possible scenarios and analysis the functionality of certain

variables using a Monte Carlo approach. One advantage of the Latin hypercube method is that guarantees that each of these components is represented in a completely stratified way (*i.e.*, the population is separated into exclusive segments). For sensitivity analysis, a random sampling of 1,000 samples were generated, which are subject to a 5 % variation for upper and lower limits. Then, the identification of the most significant variables of the system was performed using the standardized regression coefficient method.

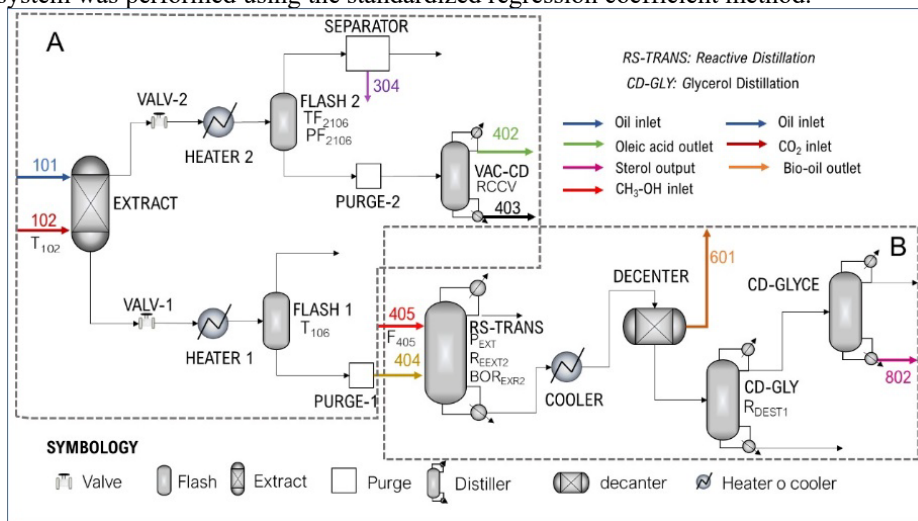


Figure 1. Flowsheet for the production process of phytosterols and biodiesel from palm oil.

#### 2.4. Optimization using multi-objective genetic algorithms.

The identified variables in the sensitivity analysis were used as decisions variables for the optimization step to improve the performance of the process. The optimization was carried out using the Genetic Algorithm Toolbox included in MATLAB. In a previous work Ponce-Rocha et al. (2022) found that the best working condition for genetic algorithms in MATLAB together with Aspen Plus was having a population of 100 and generations of 50.

### 3. Results

The framework was evaluated using the production of phytosterols and the process to produce phytosterols and biodiesel.

#### 3.1. Case study A: phytosterols production

The identifiability analysis largely depends on the interpretation of the sensitivity function of parameters and variables with respect to model outputs, usually corresponding to the measured variables such as operating costs or gross profit. The sensitivity analysis allowed to identify 6 variables with the highest impact in the operating cost, gross profit and net earnings. Table 1 illustrates the identified variables. For this case the confidence factor was 0.95-0.99 %. Figure 2 shows the significant variables for the operating cost, gross profit and the net earnings defined as the difference between the gross profit and operating costs. The results show that 3 model variables were especially important for the 3 output variables, 1 variable is significant for 2 output variables and 1 variable is important for the operating cost; the identified operating and design variables were employed in the optimization step.

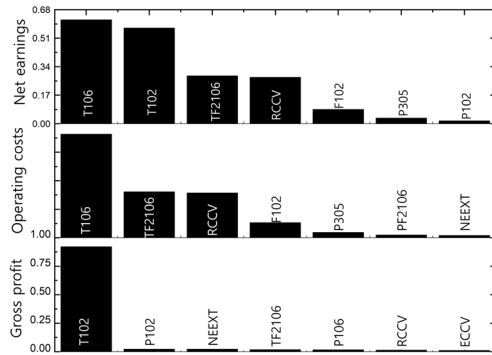


Figure 2. Variable significance ranking (only the first 7 highest ranking variables are shown) for case A model outputs based on the sensitivity measure of gross profit, operating costs and, net earnings.

Table 1. Results of Sensitivity analysis for case study A

Variable	Gross profit, %	Operating costs, %	Net earnings, %
F <sub>102</sub>	0.93	<b>11.6</b>	8.4
T <sub>102</sub>	<b>91.9</b>	1.0	<b>56.9</b>
T <sub>106</sub>	0.17	<b>79.8</b>	<b>61.9</b>
TF <sub>2106</sub>	1.72	<b>35.3</b>	<b>28.5</b>
RC <sub>CV</sub>	1.28	<b>34.5</b>	<b>27.6</b>
P <sub>102</sub>	<b>2.8</b>	0.5	1.6

### 3.2. Case study A-B: phytosterols and biodiesel production.

Figure 3 shows the identified variables with high impact in the output variables for process A-B. Table 2 illustrates the variables with high impact in the operating costs, gross profit and the net earnings (defined above) and, the percentage of relevance in process A-B.

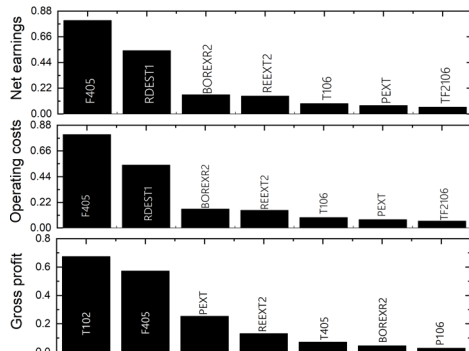


Figure 3. Variable significance ranking (only the first 7 highest ranking variables are shown) for case A-B model outputs based on the sensitivity measure of gross profit, operating costs and, net earnings.

Table 2. Results of sensitivity analysis for case study A-B

Variable	Gross profit, %	Operating cost, %	Net earnings, %
T <sub>102</sub>	<b>67.3</b>	0.14	0.41
F <sub>405</sub>	<b>57.2</b>	<b>79.9</b>	<b>79.8</b>
REXT <sub>2</sub>	<b>13.1</b>	<b>15.1</b>	<b>15.1</b>
BOR <sub>EXR2</sub>	4.4	<b>16.2</b>	<b>16.3</b>
P <sub>EXT</sub>	<b>25.2</b>	7.0	7.1
R <sub>Dest1</sub>	0.98	<b>53.9</b>	<b>53.9</b>

### 3.3. Results of the optimization for case study A

Once sensitivity measures have identified the significant variables in the process, the following step is to find out the optimal operating conditions with the aim of reducing the operating costs and increasing the net profit of the system. The variables related and identified with the gross profit were the temperature and pressure in the stream 102 (T<sub>102</sub>, P<sub>102</sub>, respectively); for the operating costs, the flowrate for stream 102 (F<sub>102</sub>), temperature of the flash unit 1 (T<sub>106</sub>), temperature for flash 2 (TF<sub>2106</sub>), reflux ratio (RC<sub>CV</sub>) were related and identified; a finally for the net earnings, the variables identified for the operating costs were also selected adding the temperature of the stream 102 (T<sub>102</sub>). The purple diamond (right and lower side of the plot) illustrates the value of the objective function for the base

case scenario (without optimization). The light blue triangles show the values of the objective function when the most important variables identified with the operating costs in the sensitivity analysis step, (see red and bold values in Table 1) were used as manipulated variables in the optimization task, the results showed a reduction of operating costs of about 4.5 % with respect to the base case, and a small increase in gross profit (about 3.4 %) as illustrated by the star with the same color. Similar case was observed when the optimization was performed using the significant variables for gross profits (see red and bold values in Table 1), here it was possible to see an increase of 4.3 % gross profit and only 3.0 % of reduction in the operating costs as illustrated by green star. In the case of net earnings, a considerable increase in profits was obtained (4.3 %) and a reduction of 4.6 % in operating costs was reached compared to the base case scenario as it is illustrated by the blue star. In the case of the optimization tasks using all the available manipulated variables (11 variables), it resulted in an increase of 3.7 % in the profit and a reduction of 2.1 % in the operating costs (see red star in Figure 4).

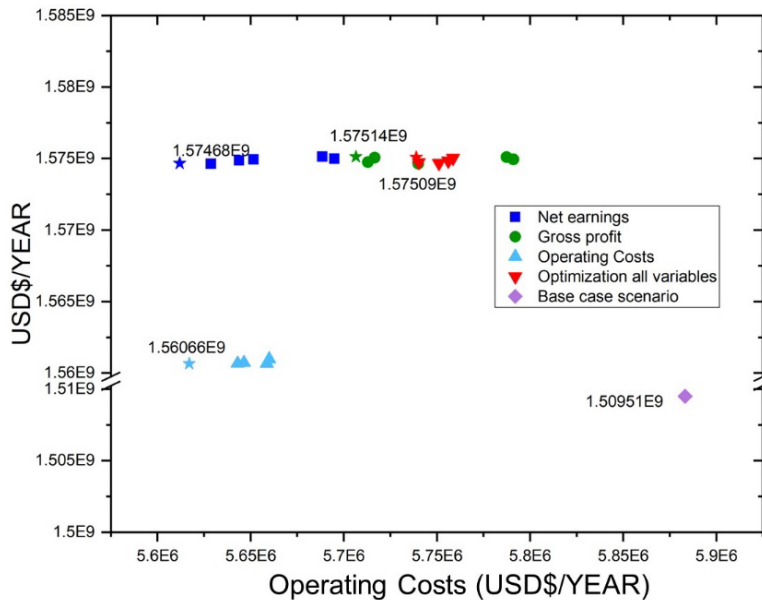


Figure 4. Pareto front to determine the optimal design case A.

### 3.4. Results of the optimization for case study A-B.

The results of case study A-B are shown in the Pareto front illustrated in Figure 5. The highest improvement was found when the identified variables using the net earnings (see red and bold values in Table 2) were used in the optimization step; the energy cost decreased by 4.4 % and the profit increased 5.02 % compared with base case scenario (purple diamond and blue star, respectively in Figure 5). When the optimization was performed using the 32 manipulated variables of the process A-B, an increase in profit of 3.02 % and a decrease in operating costs of 3.87 % (illustrated with the red star) were found compared with the base case scenario. The optimization results for the variables selected to reduce energy costs (see color red values in Table 2) provided good results because the reduction in operating costs was 4.2 %, but the increase in the profits was only 0.27 % (see light blue star). A similar case was observed when the variables with higher impact for the gross profit (see color red values in Table 2), where it was possible to increase 4.6 % the gross profit and decrease the energy consumption by 0.47 % (see green star).

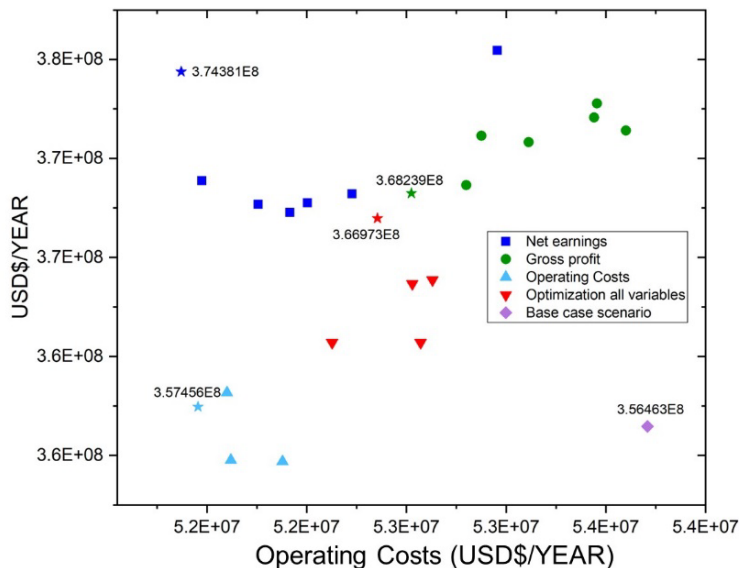


Figure 5. Pareto front to determine the optimal design case A-B.

During the optimization of the 32 variables, the optimization time was 10.56 hours, while the optimization time using only the variable identified in the sensitivity analysis was 6.95 hours, that is, the computational time was reduced 34.2 %, in addition with better results.

#### 4. Conclusion

A framework for a more efficient optimization task was developed. The sensitivity analysis revealed that based on the objectives (operating costs, gross profit and net earnings) only a few variables had an impact in the magnitude of the sensitivity measure. It was possible to observe that the identified variables were the most important in the optimization problem, allowing to reduce computing time and improving the optimal results. The sensitivity analysis generated for the variables provided an accurate picture in the selection of high impact variables with respect to the objective established for each process section and, this can be clearly seen in the optimization results for both case studies. The results also showed that was possible to find a better alternative design and operation of the plant in comparison to the base case scenario.

#### 5. Acknowledgments

M.A. Villarreal-de-Aquino acknowledges the National Council of Science and Technology (CONACyT) for the financial support on the development of this project.

#### 6. References

- S. Katoch, S. Chauhan, V. Kumar, 2021, A review on genetic algorithm: past, present, and future, *Multimedia Tools and Applications*, 80, 5, 8091-8126.
- E. N. Pistikopoulos, A. Barbosa-Povoa, J.H. Lee, R. Misener, A. Mitsos, G.V. Reklaitis, V Venkatasubramanian, F. You, R. Gani, 2021, Process systems engineering – The generation next?, *Computers & Chemical Engineering*, 147, 107252.
- J.D. Ponce-Rocha, M. Picón-Núñez, R. Morales-Rodríguez, 2022, A framework for optimal and flexible schemes design under uncertainty & sustainable aspects, *Computer Aided Chemical Engineering*, 51, 757-762

# Effect of coalescence models on the prediction of the separation of dispersed oil-water pipe flows

Nikola Evripidou,<sup>a</sup> Federico Galvanin,<sup>a</sup> Panagiota Angeli,<sup>a</sup>

<sup>a</sup> *Department of Chemical Engineering, University College London, London WC1E 7JE, UK*

## Abstract

The effect of coalescence models on the prediction of the separation of dispersed oil-water pipe flows was assessed using a one-dimensional mechanistic model. The mechanistic model predicts the formation and evolution of four characteristic layers along the pipe: a pure water layer at the bottom, a flotation/sedimentation layer, a dense-packed zone, and a pure oil layer on the top. It was shown that the film drainage coalescence model by Jeelani and Hartland (1994) that considers interfacial mobility produces good predictions at low mixture velocity, but it depends on the flowrate. The asymmetric film drainage coalescence model by Henschke et al. (2002) is independent of the mixture velocity and the dispersed-phase fraction, and produces reasonable predictions. There was small deviation between the model outputs in the presence of the four characteristic layers, but further investigation of the regions where a single dense-packed layer persists is required.

**Keywords:** coalescence; liquid-liquid; dispersion; separation; modelling

## 1. Introduction

Liquid-liquid pipe flows are common in the petroleum industry. Deepwater and marginal fields with lower volumes of reserve entail new development challenges, while heavy oils and mature wells require water flooding to enhance production. Increased volumes of water in the pipelines affect the flow, as they can alter the spatial configuration of the two immiscible phases. Considering the high cost associated with oil extraction and separation, as well as the increased demands for further offshore drilling, ensuring a successful and economical flow of oil-water mixtures is essential in optimising transportation and downstream separation of the extracted oil. Additionally, subsea separation facilities are used to reduce the cost and space requirements of remote deepwater operations. The oil-water flows in pipes are often in the dispersed pattern. Models that can predict flow pattern transitions in unstable dispersed pipe flows are essential during both design and operation of industrial facilities.

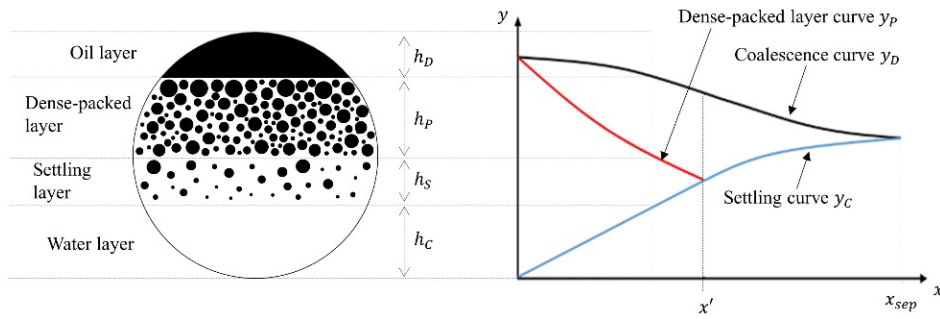
Henschke et al. (2002) developed a mechanistic model that predicts the evolution of heights of the characteristic layers that develop in separating batch dispersions as well as the average drop size. The model uses coalescence time correlations based on asymmetrical film drainage between the drops. Pereyra et al. (2013) attempted to extend Henschke's model to one-dimensional pipe flows by changing the time scale to a length scale. Evripidou et al. (2019) further modified the model to account for hindered settling of drops in dense dispersions. Another coalescence model was proposed by Jeelani and Hartland (1994) and is based on the interfacial mobility. In this work, we use the mechanistic model as presented in Evripidou et al. (2019). We consider both the



*asymmetric film drainage coalescence model* by Henschke et al. (2002) and the *interfacial mobility film drainage coalescence model* by Jeelani and Hartland (1994), to assess their ability to predict the formation and evolution of the characteristic layers in a pipe, for different oil-in-water dispersed flows.

## 2. Model description

Four characteristic layers may emerge in a separating dispersed pipe flow. In an oil-in-water dispersion, these are a pure water layer at the bottom, a settling layer (SL), a dense-packed layer (DPL), and a pure oil layer at the top. Figure 1 shows schematically these layers. The thickness of each layer depends on the drop settling (flotation/sedimentation) rate and the coalescence rate of drops, with their homophase. The drop size may also change through drop-drop coalescence.



**Figure 1: Diagram of the cross-sectional area of the pipe for an oil-in-water dispersion and diagram of the evolution of the characteristic layers.**

The mechanistic model is applicable to liquid-liquid pipe flows where the separation is primarily gravity-driven. A constant mixture velocity,  $u_M$ , is assumed for both phases – this allows velocity profiles to be ignored, which means that there is no exchange of momentum between layers. For the dispersed layers, monodisperse drop distributions are assumed for simplification, while interfacial tension is considered constant. Lastly, drop break-up and turbulence effects are neglected. In what follows, the model equations associated with coalescence are briefly discussed.

### 2.1. Coalescence analysis

#### 2.1.1. Drop-interface coalescence

Accumulation of drops near the top of the pipe results in coalescence and the formation of a pure oil layer of thickness  $h_D$ . Assuming a monodispersed DPL with drops of diameter  $d_{p,I}$ ,

$$\frac{dh_D}{dx} = \frac{2\varphi_I d_{p,I}}{3\tau_I u_M}, \quad (1)$$

where  $\tau_I$  is the drop-interface coalescence time. The oil fraction at the interface  $\varphi_I$  is taken as approximately equal to 1.

#### 2.1.2. Drop-drop coalescence

Drop-drop coalescence is considered only in the DPL. Assuming that in each step, all drops within the DPL are of equal size,  $d_{p,I}$ ,

$$\frac{d(d_{p,I})}{dx} = \frac{d_{p,I}}{6\tau_C u_M}. \quad (2)$$

where  $\tau_C$  is the drop-drop coalescence time.

### 2.1.3. Coalescence time

Two coalescence models are considered. These are outlined in table 1.  $d_{p,0}$  is the drop size at the inlet,  $\mu$  is viscosity,  $\rho$  is density,  $\sigma$  is the interfacial tension whereas the subscripts C and D refer to the continuous and dispersed phases respectively;  $g$  is the gravitational constant.

**Table 1: Coalescence models**

<i>Asymmetric film drainage model</i>	<i>Interfacial mobility film drainage model</i>
Drop-interface coalescence time: $\tau_I = \frac{(6\pi)^{\frac{7}{6}} \mu_C r_a^{\frac{7}{3}}}{4\sigma^{\frac{5}{6}} H^{\frac{1}{6}} r_{F,I} r_V^*} \quad (3)$	Drop-interface coalescence time: $\tau_I = \frac{\tau_{I,0} d_{p,I}}{h_P} \quad (9)$
Drop-drop coalescence time: $\tau_C = \frac{(6\pi)^{\frac{7}{6}} \mu_C r_a^{\frac{7}{3}}}{4\sigma^{\frac{5}{6}} H^{\frac{1}{6}} r_{F,C} r_V^*} \quad (4)$	Drop-drop coalescence time: $\tau_{I,0} = \frac{3\pi\mu_C r^4}{4(1+2m)F\delta_r^2} \quad (10)$
Drop-drop contact radius: $r_{F,C} = 0.3025 d_{p,I} \sqrt{1 - \frac{4.7}{La+4.7}} \quad (5)$	Modified coalescence equation: $\frac{dh_D}{dx} = \frac{2\phi_I h_P}{3\tau_{I,0} u_M} \quad (11)$
Drop-interface contact radius: $r_{F,I} = \sqrt{3} r_{F,C} \quad (6)$	Drop-interface film radius: $r = d_{p,0}^2 \sqrt{\frac{ \rho_C - \rho_D  g}{12\sigma}} \quad (12)$
Channel contour radius: $r_a = 0.5 d_{p,I} \left( 1 - \sqrt{1 - \frac{4.7}{La+4.7}} \right) \quad (7)$	Force due to gravity: $F = \frac{\pi d_{p,0}^3  \rho_C - \rho_D  g}{6} \quad (13)$
Modified Laplace number: $La = \left( \frac{ \rho_C - \rho_D  g}{\sigma} \right)^{0.6} h_P^{0.2} d_{p,I} \quad (8)$	Critical film thickness: $\delta_r = 0.267 \left( \frac{\pi r^4 H^2}{6\sigma F} \right)^{\frac{1}{7}} \quad (14)$

The left column of table 1 shows the *asymmetric film drainage coalescence model*, which depends on the deformation of the drops. Deformation increases with dense packed layer thickness below the drop considered. Two unknown parameters are present in this model: the Hamaker coefficient,  $H$ , and the asymmetry parameter,  $r_V^*$ .  $H$  is set to  $10^{-20}$  N m as proposed by Henschke et al. (2002) for all systems.  $r_V^*$  is system specific and can be obtained experimentally.

The right column of table 1 presents the *interfacial mobility film drainage coalescence model*. This coalescence model does not depend on drop size. Instead, the model allows equation (1) to be simplified into equation (11) through the use of equation (9), making  $h_D$  independent of  $d_{p,I}$ . The fitted parameter  $m$  is the interface mobility, i.e. the sum of the mobilities due to induced circulation in the adjacent phases and the interfacial tension gradient and is characteristic of each system. When  $m = 0$  the velocity at the interfaces on both sides of the draining film is 0, and the surfaces are deemed immobile; when  $m = 1.5$  the velocity at one of the interfaces is 0 while the velocity gradient at the other surface is 0. Under these conditions, film drainage, and thus the rate of coalescence, is extremely slow. Other values of  $m$  are also possible and correspond to different surface velocities and velocity gradients. Values of  $m$  larger than 1.5 correspond to more mobile interfaces.

### 3. Results and discussion

#### 3.1. Experimental methods

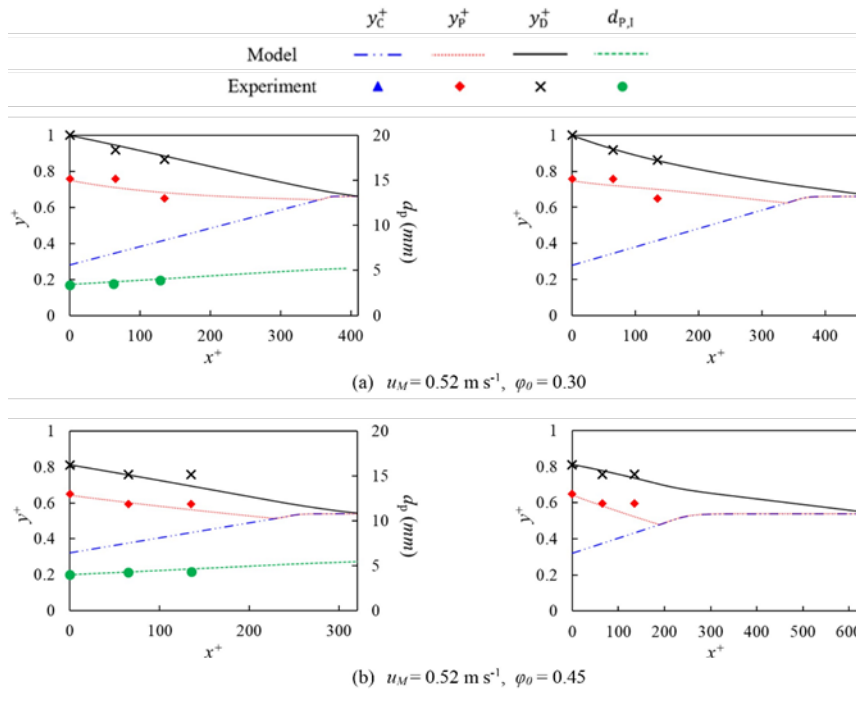
The experimental data used to assess the performance of this model were obtained by Voulgaropoulos (2018) in a two-phase liquid-liquid flow facility discussed in detail in Voulgaropoulos et al. (2016). In the experiments tap water and oil ( $828 \text{ kg m}^{-3}$ ,  $5.5 \text{ mPa s}$ ) were used as test fluids. The test section comprised of transparent acrylic pipes with an internal diameter of  $37 \text{ mm}$  and overall length of around  $8 \text{ m}$ . Partial dispersions of oil in water were generated at the inlet of the test section using a multi-nozzle mixer. High-speed imaging was employed at three locations along the spanwise dimension of the pipe to enable the identification of the flow patterns. A dual-conductance probe was implemented to measure the local volume fractions and the drop size distributions of the dispersions. Measurements were taken every  $2 \text{ mm}$ , spanning the whole pipe diameter.

#### 3.2. Results and discussion

Three case studies were investigated (c.f. Table 2). For each case, we solved the mechanistic model twice, implementing a different coalescence model each time, using gPROMS ModelBuilder at intervals of  $0.1 \text{ m}$  pipe length.  $C_h$  was taken as  $0.01$  as suggested by Evripidou et al. (2019) and  $r_V^*$  was set to  $0.007$  as suggested by Pereyra et al. (2013).  $m$  was fitted to experimental data obtained at  $u_M = 0.52 \text{ m s}^{-1}$  and was found to be  $360$ . The resulting flow profiles were nondimensionalized using the pipe diameter and are presented in Figures 2 and 3.

**Table 2: Inlet conditions of the experiments**

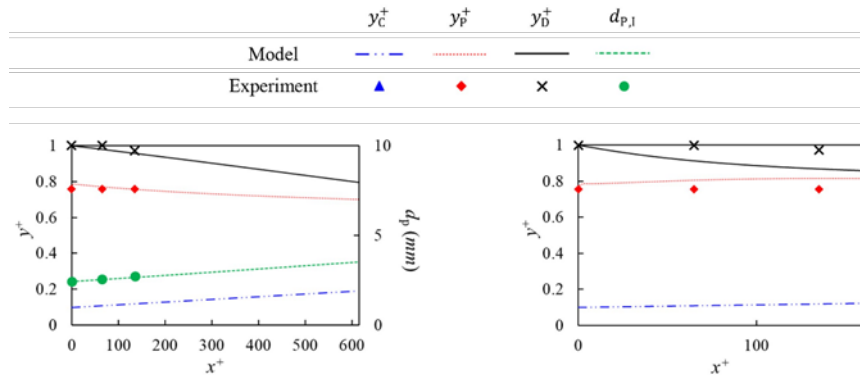
$u_M \text{ (m s}^{-1}\text{)}$	$\varphi_0$
<b>0.52</b>	0.30
	0.45
<b>1.04</b>	0.60



**Figure 2: Model predictions of the asymmetrical film drainage coalescence model (left) and the interfacial mobility film drainage coalescence model (right) for  $u_M = 0.52 \text{ m s}^{-1}$ .**

Figure 2 shows the two case studies at  $u_M = 0.52 \text{ m s}^{-1}$  and  $\varphi = 0.30$  and  $0.45$ . The two mixtures separate in a similar fashion and both coalescence models predict the separation with reasonable accuracy. The rate of drop-settling is large enough to deplete the SL first. On the contrary, the DPL persists throughout the pipe and coalescence controls the rate of separation.

Although both coalescence models show reasonable agreement with experimental data, the *interfacial mobility film drainage model*, where the coalescence rate is a function of the DPL thickness shows a better fit to the experiments. Despite that, the SL is depleted at similar axial lengths for both cases with deviations of 13% or less in  $x'^+$  between the predictions of the two models. The predictions of the interfacial mobility film drainage model for the total separation length are consistently larger than the predictions of the asymmetric film drainage model. Specifically, for the case with oil fraction of 0.30, the interfacial mobility film drainage model predicts an  $x_{sep}^+$  of 455, while the prediction of the asymmetric film drainage model is 9% less at 412. This difference is even larger for the case of oil fraction of 0.45. For this case, the asymmetric film drainage model predicts a separation length of 320, while the other coalescence model predicts  $x_{sep}^+ = 624$ , a value that is almost twice as large. The above observations suggest that the models behave in a similar manner at the pipe locations where all four characteristic layers are present, and any major deviations between the two arise past the point of depletion of the SL (i.e. at lengths  $x > x'$ ).



**Figure 3: Model predictions of the asymmetrical film drainage coalescence model (left) and the interfacial mobility film drainage coalescence model (right) for  $u_M = 1.04 \text{ m s}^{-1}$  and  $\varphi_0 = 0.60$ .**

Figure 3 presents the results obtained at a mixture velocity of  $1.04 \text{ m s}^{-1}$  and oil fraction of 0.60. The deviations in the predictions of the two models for this case study are significant. The coalescence model by Henschke et al. (2002) produces reasonable results with deviations of 2% or less to experimental measurements. On the other hand, the *interfacial mobility film drainage model* significantly overestimates the coalescence rate of drops with their homophase. As a result, the oil layer acquires a large thickness at the beginning of the pipe. This suggests that the mixture velocity may affect interfacial mobility and that  $m$  should be fitted for each mixture velocity.

Nevertheless, both models predict depletion of the DPL, which, according to Evripidou et al. (2022), occurs once the thickness of the DPL becomes smaller than the drop diameter along the interface. At this point, settling becomes the limiting separation mechanism and controls the rate of separation. Predictions of the complete flow profile

up to the point of total separation are not possible with the current model, as it is only applicable to regions where a DPL is present and coalescence controls the rate of separation.

#### 4. Conclusions

The paper presents a comparison between two coalescence models that can be used in mechanistic models of separating dispersed pipe flows. At the low mixture velocity, both coalescence models capture the drop-interface coalescence adequately. The *interfacial mobility film drainage model* shows better agreement with experimental data, hence may be preferred over the *asymmetric film drainage coalescence model*. The two models result in similar predictions for the length of depletion of the settling layer (SL), but the *interfacial mobility film drainage model* predicts significantly larger separation lengths  $x_{sep}$ . Hence, for the low mixture velocity, we concluded that the models behave similarly in pipe locations where all four layers are present, and any major deviations between the two models occur at pipe lengths greater than  $x'$  for which we have no experimental data. Therefore, further experimental studies, especially in the region  $x > x'$ , are needed to provide the necessary information to differentiate between the two coalescence models.

At the high mixture velocity, the *asymmetric film drainage coalescence model* shows good agreement with the experimental data. The *interfacial mobility film drainage model*, however, significantly overpredicts the rate of drop-interface coalescence. The observations suggests that the interfacial mobility coefficient  $m$  varies with  $u_M$ , but not with oil fraction  $\varphi$ . To the contrary,  $r_V^*$  is specific to the oil-water system but independent to both  $u_M$  and  $\varphi$ , hence the *asymmetric film drainage model* may be preferred for cases with variable flowrates. Nevertheless, both coalescence models predict the depletion of the DPL at the high mixture velocity. To account for the depletion of the DPL a different approach must be used such as that described in Evripidou et al. (2022).

#### Acknowledgements

NE would like to acknowledge Chevron Corporation and UCL for her PhD studentship.

#### References

- N. Evripidou, V. Voulgaropoulos, P. Angeli, 2019, Simplified mechanistic model for the separation of dispersed oil-water horizontal pipe flows, In BHR 19th International Conference on Multiphase Production Technology. OnePetro.
- N. Evripidou, C. Avila, P. Angeli, 2022, A mechanistic model for the prediction of flow pattern transitions during separation of liquid-liquid pipe flows, International Journal of Multiphase Flow, p.104172.
- M. Henschke, L.H. Schlieper, A. Pfennig, 2002, Determination of a coalescence parameter from batch-settling experiments, Chem Eng J 85, 369-378.
- S.A.K. Jeelani, S. Hartland, 1994, Effect of interfacial mobility on thin film drainage, J. Colloid Interface Sci., 164, 296-308.
- E. Pereyra, R.S. Mohan, O. Shoham, 2013, A simplified mechanistic model for an oil/water horizontal pipe separator, Oil and Gas Facilities, 2 (03): 40-46.
- V. Voulgaropoulos, L. Zhai, C. Ioannou, P. Angeli, 2016. Evolution of unstable liquid-liquid dispersions in horizontal pipes, 10th North American Conference on Multiphase Technology. BHR Group. Banff, 8-10 June.
- V. Voulgaropoulos, 2018, Dynamics of spatially evolving dispersed flows, PhD dissertation, University College London, London, UK.

# Learning interpretable multi-output models: Kaizen Programming based symbolic regression for estimating outlet concentrations of a splitter

Jimena Ferreira,<sup>a,b</sup> Martín Pedemonte,<sup>b</sup> Ana I. Torres,<sup>c\*</sup>

<sup>a</sup> *Instituto de Ingeniería Química, IQ, Facultad de Ingeniería, Universidad de la República, Montevideo, 11600, Uruguay*

<sup>b</sup> *Instituto de Computación, INCO, Facultad de Ingeniería, Universidad de la República, Montevideo, 11600, Uruguay*

<sup>c</sup> *Department of Chemical Engineering, Carnegie Mellon University, 5000 Forbes Ave., Doherty Hall, Pittsburgh, PA, 15213, USA*

\* [aitorres@cmu.edu](mailto:aitorres@cmu.edu)

## Abstract

Machine learning techniques for building surrogate models typically generate single-output expressions. However, Chemical Engineering processes are naturally multi-output, with several of the outputs depending on the same physico-chemical phenomena. This work is an extension of our previous work on Symbolic Regression aimed to learn interpretable models from data that reflect the shared physico-chemical phenomena as common bases in the derived expressions. To do so, we consider two multi-output algorithms: one based on multi-output linear squares regression, and the other one on an islands model. As case studies, we use real historical data from a C3/C4 splitter from an oil refinery. The results from the proposed multi-output algorithms are compared to those from the previous single-output algorithm and a multi-output Gaussian Process implementation. We conclude that both multi output strategies are competitive, and in some cases better, than the single-output one although within the same order of magnitude in the error in the validation set. In all cases the multi-output algorithms perform better than Gaussian Process.

**Keywords:** machine learning, surrogate modeling, symbolic regression, interpretable models, multi-output regression.

## 1. Introduction

Machine learning techniques for learning models from data are increasingly receiving attention within the Chemical Engineering community. As examples, the work of Eagle and Sahinidis (2022), developed a symbolic regression (SR) algorithm to obtain surrogate models for equations of state; and the work of Chakraborty et al. (2021), considers a machine learning technique to obtain interpretable models based on genetic feature extraction. Our own previous work in Ferreira et al., 2022a, also addresses the problem of learning an explicit interpretable model from data without a priori definition of the functional bases. The techniques used in all these examples are single output, meaning that if one wishes to derive expressions for multiple outputs, the algorithms must be repeatedly executed.

However, for many Chemical Engineering applications it is highly desirable to be able to learn models that result in common (interpretable) terms for those outputs that are known to depend on common phenomena. In Ferreira et al., 2022b, we proposed two multi output

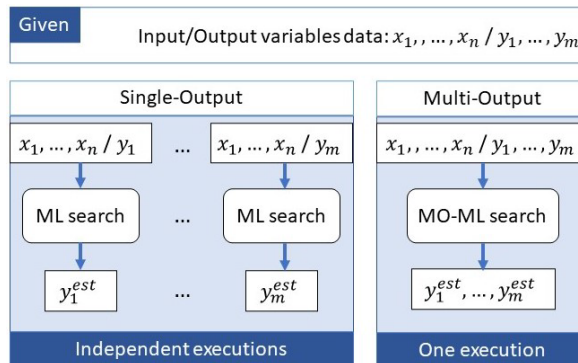


Figure 1 – Scheme of generation of surrogate models for several outputs with the same input data. Left: single-output algorithm executes to learn one model at a time. Right: the multi-output strategy learns all the models in the same run.

algorithms and tested them using benchmark functions or made-up data stemming from reaction engineering problems. In this work, we extend our previous contributions to learn multi-output interpretable models from real data from a distillation column that operates in an oil refinery. Specifically, we consider the data for the C3/C4 splitter in Ferreira et al. (2022a) and apply the two multi-output algorithms. The results from these, are compared with those from the previous single-output algorithm and a multi-output Gaussian Process implementation.

## 2. Algorithms

Figure 1 shows a scheme of the approaches considered in this work. On the left, we include the strategy used by most ML methods to solve multi-output surrogate modeling. This is, for each input-output dataset (represented by several vectors  $x$ , and one of the vectors  $y$ ), an individual execution is performed for each  $y_j$  to learn a model that relates it to the  $x$  variables. We named this strategy as the Single-Output (SO) approach. On the right, we consider the multi-output (MO) approaches where models for all  $y$  are learned in the same execution. Figure 2 shows the specific algorithms used in this work. In all cases, KP refers to the iterative algorithm that solves SR problems from input-output data. KP searches for models that are linear combinations of nonlinear expressions. These nonlinear expressions are the functional bases  $f_i^j$  that are created and modified by operations of Genetic Programming, and the coefficients  $\beta_i^j$  are calculated by Ordinary Least Squares (OLS). The details about KP have been published in Ferreira et al. (2022a). The first multi-output strategy,  $\text{MO}_{\text{MLR}}$ , performs a multi-output linear regression based on OLS in the step of calculation of the coefficients. This change forces term sharing between the  $y_j$  models at each iteration. The second multi-output strategy,  $\text{MO}_{\text{IM}}$ , is based on the islands model for parallelization of evolutionary algorithms.  $\text{MO}_{\text{IM}}$  alternates independent search steps with migration steps. In the independent search steps, each output variable  $y_j$  is learned as a SO strategy; in the migration step the functional bases of each output variable (that were searched in the independent executions) are shared with the others output variables. Thus, in the next independent search step, each independent search for  $y_j$  starts with all the functional bases of the other  $y$ . It is important to note that contrary to the other multi-output strategy,  $\text{MO}_{\text{IM}}$  does not force the final expressions of the models to all have the same functional bases but enhances the probability of that happening. More details on the MO algorithms can be found in Ferreira et al. (2022b).

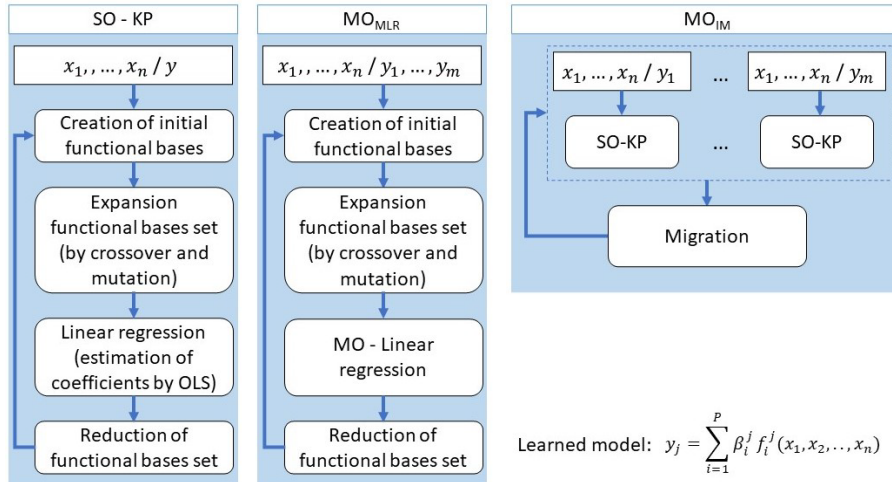


Figure 2 – Scheme of algorithms used in this work.

### 3. Case studies

As mentioned earlier, we applied the MO algorithms to the C3/C4 oil refinery splitter in Ferreira et al. (2022a). In such splitter, the inlet stream is mainly composed of C3 and C4 hydrocarbons, while other hydrocarbons with 1, 2, 5 and 6 carbons, as well as H<sub>2</sub> and H<sub>2</sub>S are the main minor components. The system is continuously monitored through eight online sensors that measure the feed, reboiler, head and distillate temperatures; the reflux, bottoms and distillate flows; and the bottom’s pressure. These 8 variables are the input variables to the models, i.e.  $x = [T_{feed}, F_{reflux}, F_{bottoms}, F_{distil}, T_{reb}, T_{head}, T_{distil}, P_{bottom}]$ .

The quality of the outlet streams is monitored through laboratory experiments (chromatography) that determine the composition of 19 species with daily or bi-daily frequency. As case studies we evaluate the development of models in two instances. One to obtain the models of the volumetric fraction of C4s in the distillate and the volumetric fraction of C3s in the bottoms (reboiler)  $y_{case-1} = [C4_{distillate}, C3_{reboiler}]$ . These fractions evaluate the quality of the outlet streams and depend on the temperature and pressure of the distillate/reboiler; we do not expect the expressions for the outlets to share these terms. In the second case study, we are interested in developing models for the volumetric fractions of the seven C4 hydrocarbons in the distillate  $y_{case-2} = [1,3\text{-butadiene}, 1\text{-butene}, \text{cis-2-butene}, \text{isobutane}, \text{isobutene}, \text{n-butane}, \text{trans-2-butene}]$ . This is the case where we expect the output variables to share many terms.

To compare the three algorithms (SO-KP, MO<sub>MLR</sub> and MO<sub>IM</sub>), we followed the statistical procedure in Alba and Luque (2005). This is, first, one hundred independent executions were performed for each algorithm; then, the Friedman test was used as an omnibus test to compare the RMSE distributions over validation sets. Finally, a pairwise Wilcoxon signed test to the RMSE distributions was performed in those cases in which differences among the distributions were found through the Friedman test. All the statistical tests were performed with a confidence level of 99%.



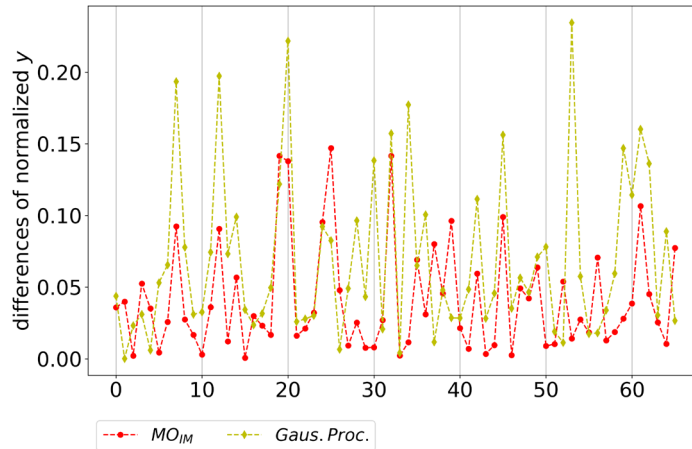


Figure 3 – Absolute differences of the estimation and data point by the best model with  $MO_{IM}$  and GP, over validation set.

## 4. Numerical results and discussion

### 4.1. Case study 1

Based on Friedman test we found that for  $y_1$  there is statistical difference between the distributions of the RMSE of the models obtained with the three algorithms ( $p_{value} = 4.81 \cdot 10^{-12}$ ), and for  $y_2$  there is not statistical difference ( $p_{value} = 0.0105$ ), with a median of RMSE between 0.062 and 0.067.

From the Wilcoxon signed rank test for  $y_1$ , we found that there is no statistical difference between the error's distributions of the models obtained by  $MO_{IM}$  and  $MO_{MRL}$  ( $p_{value} = 0.033$ ), whereas the models obtained by SO-KP had an overall lower median for RMSE ( $p_{value} = 2.51 \cdot 10^{-9}$  for  $MO_{IM}$  and SO-KP and  $p_{value} = 2.35 \cdot 10^{-7}$  for  $MO_{MRL}$  and SO-KP). The RMSEs are: 0.051, 0.056 and 0.057, respectively for SO-KP,  $MO_{MRL}$  and  $MO_{IM}$ .

Figure 3 presents the absolute difference between estimation and data over the validation set, for the second output variable by the best model obtained by  $MO_{IM}$ , and the one obtained using the multi-output Gaussian Process algorithm (GP, implemented in the *scikit-learn* package). From the figure, we can see that all estimations obtained with the KP-based algorithm are better than the ones obtained with GP. This is reinforced by the value of RMSE of the model obtained by GP: 0.088. It is worth commenting that all the values presented in these figures were normalized to preserve data confidentiality.

Table 1 shows the best model for each output variable obtained with each KP-based algorithm. It can be seen that the models obtained with  $MO_{IM}$  correctly capture the dependence of  $y_1$  on  $T_{head}$  and  $T_{distil}$ , and of  $y_2$  on  $T_{reb}$  and  $P_{bottom}$ . On the other hand, those obtained with SO-KP include almost all inlet variables in the mathematical expression. Thus, despite its slightly largest RMSE, we consider the  $MO_{IM}$  model to be superior.

### 4.2. Case study 2

In this case, based on Friedman test, we found that for  $y_1$  there is no statistical difference between the distributions of the RMSE of the models obtained with the three algorithms. For the rest of the output variables there exists statistical difference.

Table 1 – The best learned models for the second output variable with SO-KP, MO<sub>IM</sub> and MO<sub>MLR</sub> algorithms. The coefficients are presented as Greek letters in order to preserve confidentiality.

y <sub>1</sub>	SO-KP	$\beta_1^1 \frac{F_{reflux}^2}{(F_{reflux} + P_{bottom})T_{head}} + \beta_2^1 \frac{F_{reflux}F_{distil}}{T_{head}P_{bottom}} +$ $\beta_3^1 \frac{T_{reb}}{T_{head}P_{bottom}} + \beta_4^1 \frac{F_{reflux}T_{distil}}{T_{head}P_{bottom}} + \beta_5^1 \frac{F_{reflux}^2}{P_{bottom}} +$ $\beta_6^1 F_{reflux} + \beta_7^1 F_{bottoms} + \beta_8^1 F_{distil} + \beta_9^1 T_{feed} + \beta_{10}^1 T_{head} +$ $\beta_{11}^1 T_{feed}(F_{bottoms} + F_{distil}) + \beta_{12}^1 T_{distil} +$ $\beta_{13}^1 \left( \frac{T_{head}}{T_{reb}} \right)^{\left[ (P_{bottom} - F_{distil}) \frac{F_{reflux} - F_{bottoms}}{T_{distil} - P_{bottom}} \right]} +$ $\beta_{12}^1 (F_{bottoms} + F_{distil})F_{distil}$
	MO <sub>IM</sub>	$\gamma_1^1 F_{distil} + \gamma_2^1 P_{bottom} + \gamma_3^1 F_{reflux} + \gamma_4^1 T_{distil} + \gamma_5^1 T_{feed} +$ $\gamma_6^1 T_{reb} + \gamma_7^1 F_{reflux}F_{distil} + \gamma_8^1 F_{reflux}T_{feed} +$ $\gamma_9^1 (F_{reflux}^2 - T_{head} T_{distil}) + \gamma_{10}^1 T_{reb}T_{head}$
	MO <sub>MLR</sub>	$\delta_1^1 P_{bottom} + \delta_2^1 T_{head} + \delta_3^1 F_{distil} + \delta_4^1 F_{reflux} + \delta_5^1 T_{feed} +$ $\delta_6^1 F_{reflux}^2 + \delta_7^1 F_{reflux}T_{reb} + \delta_8^1 T_{reb} + \delta_9^1 \frac{T_{feed}F_{bottoms}T_{reb}}{T_{distil}}$
y <sub>2</sub>	SO-KP	$\beta_1^2 \left( \frac{F_{reflux} - F_{bottoms}}{F_{distil} + T_{reb}} \right) + \beta_2^2 F_{bottoms} + \beta_3^2 F_{reflux} +$ $\beta_4^2 F_{bottoms} + \beta_5^2 P_{bottom} + \beta_6^2 \frac{F_{bottoms} - F_{reflux}}{2F_{distil} - T_{distil}} + \beta_7^2 T_{reb} +$ $\beta_8^2 F_{distil}$
	MO <sub>IM</sub>	$\gamma_1^2 \frac{F_{bottoms}}{T_{reb}} + \gamma_2^2 F_{reflux} + \gamma_3^2 F_{bottoms} + \gamma_4^2 P_{bottom} + \gamma_5^2 T_{reb} +$ $\gamma_6^2 P_{bottom}^2 + \gamma_7^2 F_{reflux}T_{reb} + \gamma_8^2 F_{reflux}F_{distil}$
	MO <sub>MLR</sub>	$\delta_1^2 \left( \frac{T_{feed}}{F_{reflux} - T_{reb}} - \frac{P_{bottom}}{T_{head}} \right) + \frac{\delta_2^2 F_{distil}}{F_{reflux} - T_{reb}} +$ $\frac{\delta_3^2 P_{bottom}}{T_{feed}} + \delta_4^2 F_{reflux} + \delta_5^2 F_{bottoms} + \delta_6^2 P_{bottom} +$ $\delta_7^2 T_{reb} + \delta_8^2 T_{head} + \delta_9^2 (F_{reflux} - T_{feed})F_{bottoms}T_{distil} +$ $\delta_{10}^2 (F_{bottoms} - F_{reflux})T_{reb}$

For output variables y<sub>2</sub> to y<sub>7</sub>, the Wilcoxon signed rank test showed that there is not statistical difference for three output variables, MO<sub>IM</sub> has lower median RMSE than SO-KP for two output variables, and SO-KP lower than MO<sub>IM</sub>. in one case. Comparing SO-KP with MO<sub>MLR</sub>, SO-KP median RMSEs were lower than MO<sub>MLR</sub> in two cases. Finally, comparing MO<sub>MRL</sub> and MO<sub>IM</sub>, there was no statistical difference in two cases, MO<sub>IM</sub> had lower RMSEs in three cases, and MO<sub>MRL</sub> lower than MO<sub>IM</sub>, in one case.

Then, the models obtained by multi-output algorithms generally perform better in terms of RMSEs. MO models also showed better performance in terms of term sharing which is coherent with the fact that in this case, output variables are strongly related, and the multi-output learning is enriched by incorporation of data from all variables.

## 5. Conclusions

In this paper we tested the performance of two multi output (MO) strategies to learn an interpretable models (symbolic regression) where the bases functions were learned via genetic programming and the coefficients via least squares regression. The first MO strategy,  $MO_{MLR}$ , is based on the use of a multi-output linear regression instead of a single-output linear regression. The second algorithm,  $MO_{IM}$ , is based on the islands model for parallelization of evolutionary algorithms, in which steps of independent search alternate migration steps where the best functional bases are exchanged. The models obtained by the multi-output algorithms were compared with the single-output executions for each output variable, and a multi-output Gaussian Process algorithm.

Data from a real C3/C4 splitter from an oil refinery was used in the case studies. We considered a first case where the output variables were weakly related by the physico-chemical phenomena, and a second case, where the output variables were strongly related. We showed that in the first case study, the island based multi-output strategy  $MO_{IM}$  was able to identify the variables that were relevant for each output variable model, although having a slightly worse performance in terms of RMSE (in the validation set). In the second case study, multi-output strategies performed better than the single-output one, although within the same order of magnitude in RMSE. The three algorithms performed better than the multi output Gaussian Process one.

## Acknowledgements

The authors gratefully acknowledge the collaboration of ANCAP-Uruguay for providing the data used for the case study.

J. Ferreira thanks the Agencia Nacional de Investigación e Innovación of Uruguay for financial support for her graduate studies (Award No. POS\_NAC\_2018\_1\_152185).

## References

- E. Alba and G. Luque, Measuring the Performance of Parallel Metaheuristics. John Wiley & Sons, Ltd, 2005, ch. 2, pp. 43–62.
- A. Chakraborty, A. Sivaram and V. Venkatasubramanian, 2021, AI-DARWIN: A first principles-based model discovery engine using machine learning, *Computers & Chemical Engineering*, 54, 107470.
- M.R. Engle and N.V. Sahinidis, 2022, Deterministic symbolic regression with derivative information: General methodology and application to equations of state. *AIChe J.* 68 ( 6): e17457.
- J. Ferreira, A. I. Torres and M. Pedemonte, 2021, Towards a Multi-Output Kaizen Programming Algorithm, 2021 IEEE Latin American Conference on Computational Intelligence (LA-CCI), Temuco, Chile.
- J. Ferreira, M. Pedemonte, and A. I. Torres, 2022a, Development of a machine learning-based soft sensor for an oil refinery's distillation column. *Computers & Chemical Engineering*, 161, 107756.
- J. Ferreira, M. Pedemonte, and A. I. Torres, 2022b, A multi-output machine learning approach for generation of surrogate models in process engineering. *Computer Aided Chemical Engineering*, 49, 1771-1776.
- D.G. Krige, 1951, A statistical approach to some mine valuations and allied problems at the Witwatersrand. Master's thesis of the University of Witwatersrand.

# Decision regression for modelling of supply chain resilience in interdependent networks: LNG case

Adnan Al-Banna,<sup>a,b</sup> Brenno C. Menezes,<sup>b</sup> Mohammed Yaqot,<sup>b</sup> Jeffrey D. Kelly<sup>c</sup>

<sup>b</sup>*Division of Engineering Management and Decision Sciences, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar Foundation, Qatar*

<sup>b</sup>*Department of Logistics and Supply Chain, Milaha, Doha, Qatar*

<sup>c</sup>*Industrial Algorithms Ltd., 15 St. Andrews Road, Toronto M1P 4C3, Canada*  
[bmenezes@hbku.edu.qa](mailto:bmenezes@hbku.edu.qa)

## Abstract

As an application of advanced analytics (AA) in supply chains (SCs), to model supply chain resilience (SCR) of transactions, logistics, operations, etc., of such complex representation of networks, we propose a supervised machine learning approach as a predictive analytics decision regression modelling framework that uses a coefficient setup MIQP (mixed-integer quadratic programming) technique. This can determine optimisable surrogate models to correlate independent X variables (e.g., resistance and recovery of the SC resilience) to the dependent Y variable SCR considering a dynamic behaviour of the SC with lag- and dead-time. A novel methodology to quantify SCR, based on a tree of continuous x and binary y variables of resistance (avoidance and containment) and recovery (stabilisation and return), considers ad hoc relationships of x and y to be part of the SCR predictions in the machine learning MIQP identification method. Such SCR algebraic or analytical formulas obtained in this constrained decision regression approach (a type of predictive analytics) can be used in optimisation and control problems of prescriptive and detective analytics types of AA. The proposed model is applied in an oil and gas case pertaining to liquid natural gas (LNG) and its leaks known as boil-off gas (BOG). The vapors generated by the leaks or venting of BOG in the LNG supply chain reduces the SCR of this commodity. There are losses of materials, environmental impacts, reduction in the calorific value of the LNG to be re-gasified, potential safety issues, to name a few. This research aims to introduce a methodology to model and predict SCR in a general way (by a new design as example), but particularly, covers the digital transformation implementations that can potentially lead to an enhanced resilience in the supply chain of LNG towards the desired level of digital supply chain resilience (DSCR).

**Keywords:** supply chain, resilience, advanced analytics, surrogate modelling.

## 1. Introduction

Making determinations on complex design, operation, and control of today's industry may count on recent capabilities of advanced analytics (AA) that is widespread in machine learning techniques and big data novelties and applications. Particularly, supply chain and logistics organisations operate in a data-rich environment, managing huge amount of data about customers' demands, offerings of suppliers, previous and current orders, satisfaction levels and reviews, shipment movements, locations of containers, temperature, humidity, among others. Hence, many supply chain organisations have put data analytics at the top of their strategic priorities. This is driving interest in smart analytics techniques designed to help diagnose operational issues, optimise network

planning, and predict future scenarios. As a branch of the so-called smart manufacturing, AA can consider multiple scopes, scales, scenarios, and layers to model and solve decision-making problems as found in the engineering management and decision science literature. It moves from the mature or old-fashion descriptive analytics of past activities to the needed cognitive analytics with self-learning and adaptive structures in both manufacturing systems and supply chains (Menezes et al., 2019a).

The proposed SCR modelling in this work uses a methodology that considers a MIP-based machine learning modeling approach with the objective of identifying the causation and correlation among (1) two parent inputs (resistance and recovery) and (2) their four child inputs (avoidance and containment, and stabilisation and return) of the SCR from Melnyk et al. (2014) as seen in Figure 1. This can guide directions in the investments on new design and digital transformation building blocks for improvement of the liquid natural gas (LNG) supply chain resilience that is impacted by the boil-off gas (BOG) losses during the shipping and its adjacent areas (upstream and downstream).

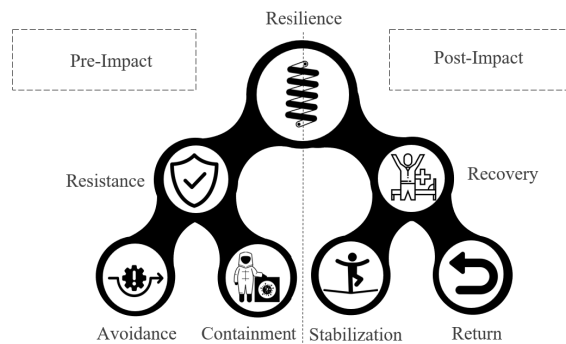


Figure 1: Tree of supply chain resilience (adapted from Melnyk et al., 2014).

## 2. Boil-off gas in the LNG supply chain

To be stocked and transported in the liquid state, natural gas (NG) requires very low temperatures, therefore it is very susceptible to heat invasion and other driven-forces and starts to boiling at  $-162\text{ }^{\circ}\text{C}$ . This vaporised LNG represents a substantial energy waste with a negative impact in economics, climate warming, and operational hazard. Henceforth, the LNG losses throughout the supply chain requires to be controlled and managed in the most efficient manner, avoiding the losses of BOG or containing it, if LNG boils. Also, if it is lost, measures to recovery the commodity to its initial product condition in terms of its properties and specifications (after the liquefaction plant).

Figure 2 illustrates potential deployment of internet-of-things (IoT), automated decision-making and controls, artificial intelligence (AI), etc., elements of the Industry 4.0 age on LNG vessels and their associated facilities. Besides the use of technology (more operational related) to increase SCR, other types of investments are related to changes in design, by changes in grassroots projects or retrofit of current assets. Among these investments, there are (1) new vessels propelled by LNG, (2) retrofit in current vessels to install liquefaction systems to liquefy the BOG; and (3) different storage tank inside the vessels considering the two main LNG vessel types, the MOSS and membrane. In the latter, as example, the membrane design is more resilient and advantageous due to many reasons, but mainly by its robust insulation, as it contains an amalgamation of multiple

layers of carefully selected materials that is split into two membranes (a primary and a secondary) that spans across almost one meter distance between the ship's inner hull and the LNG cargo.

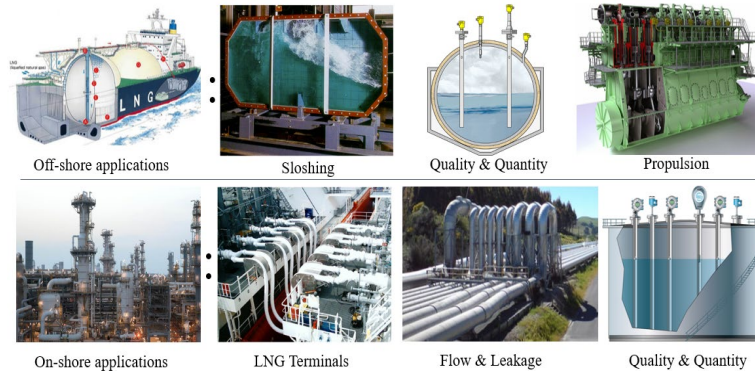


Figure 2: Potential investments to increase supply chain resilience.

### 3. LNG supply chain resilience stages

In the following, a detailed description of examples of resistance and recovery stages related to the BOG losses in the LNG supply chain is given.

#### 1) Resistance Stage

In the avoidance stage, supply chain resilience (SCR) can be associated to the supply chain vulnerability (SCV) for the BOG generation driven-forces. Some SCVs are the (a) fluctuation in the ambient temperature and tank pressure, and (b) sloshing (LNG fluid excessive movement and friction with its internal reservoir surfaces), which is influenced by excessive rolling of the ship during voyage. Besides, by augmenting the insulation between the ambient and the stored LNG layer in LNG ships is under constant development, and it is one of the main reasons the old MOSS type vessels are now progressively replaced with the new membrane type vessel. However, insulation enhancement is extensive work and requires excessive out of operations periods and expensive drydock repairs.

In the same context, controlling the tank pressure is a critical factor to avoid BOG generation, and this could be established with any sort of sensing-calculating-actuating (SCA) capability to proactively regulate the tank pressure (Menezes et al., 2019b). Moreover, route optimisation is an efficient methodology to main the ship voyage smooth and hence significantly reduce the time-period of the friction between the LNG and its tank surfaces, eventually reducing the BOG generation. Such an enhancement can be achieved via deploying artificial intelligence (AI) and big data (BD) enablers with the objective of calculating maritime records about waves, currents, and winds, among others to find optimised routes. From work experience, ship arriving late due to taking longer voyage route could have better economics compared to a ship that charted a shorter route but through rough waves that generated excessive BOG.

In the containment stage, the objective is to explore methodologies to contain and utilise the generated BOG within the LNG ship eco-system instead of losing it to the atmosphere.

Such enhancement can be via deploying re-liquefaction plants onboard the ship, which receive the generated BOG, re-liquefy it, and eventually direct it to the LNG storage tank. This non-I4.0 enhancement requires excessive time out of operations and drydock repairs due to the extensive pipe work involved. Another enhancement in this stage is directing the generated BOG towards the propulsion system of the ship, however limited percentage of LNG ships worldwide have this capability. Currently, there are two types of two stroke engines that deliver this flexible dual-fuel propulsion, that are MEGI (M-type, Electronically Controlled, Gas Injection), and XDF (low pressure 2-stroke engine). Most of the newly built LNG ships are either MEGI or XDF; however, despite some major marine engine manufacturers are starting to offer retrofitting ships solutions, this option may not be economically viable (Man Energy Solutions, 2022).

2) *Recovery Stage*: in this stage and its sub-stages of stabilisation and return, the objective is to achieve stabilisation and return post the generation of BOG, where it is not attainable to avoid, nor contain it. An envisaged enhancement in this stage is to deploy AI solutions with the objective to optimise receiving terminals pumps operation algorithms in a manner that assist in the recovery process. Similarly, another enhancement is injecting ethylene ( $\text{CH}_2=\text{CH}_2$ ) in the receiving terminals to enhance the calorific value and heat content of the delivered LNG during its re-gasification stage and injection in pipelines.

The aforementioned discussion is summarised in Table 1.

Table 1: BOG resilience tree.

Avoidance (On the ship), potential I4.0 enabler and desired function		
Tank temperature	Operational (I4.0)	Deploy
	Design (Retrofit)	Install enhanced insulation around the tank
Tank pressure	Operational (I4.0)	Deploy S-C-A to keep the pressure in the tank regulated
Optimise route	Operational (I4.0)	Deploy AI and BD to calculate maritime records about waves, currents, and winds, etc., and find optimised routes
Containment (On the ship), Potential I4.0 enabler and desired function		
Propulsion	Design (Grassroots)	Generated BOG fuelling propulsion systems
Re-liquefaction	Design (Retrofit)	Re-liquefy the BOG
Stabilisation (At receiving terminal), Potential I4.0 enabler and desired function		
Pumping at receiving terminals	Operational (I4.0)	Deploy AI to optimise operation of the pumps in receiving terminal (KBC, 2021)
Return (At the receiving terminal), Potential I4.0 enabler and desired function		
Injecting ethylene ( $\text{CH}_2=\text{CH}_2$ )	Operational	Inject ethane in the re-gasification stage in the receiving terminal to correct the calorific value and heat content of LNG

From the operational and design related projects as seen in Table 1, with I4.0 content or not, a plot can be sketched for the proposed resilience domains (resistance and recovery) that are function of the investments. This is shown in Figure 3, where two planes are highlighted to exemplify SCR investments more related to design and those with focus on operational investments (supposedly less expensive, but not such efficient in comparison with the design expenditures.

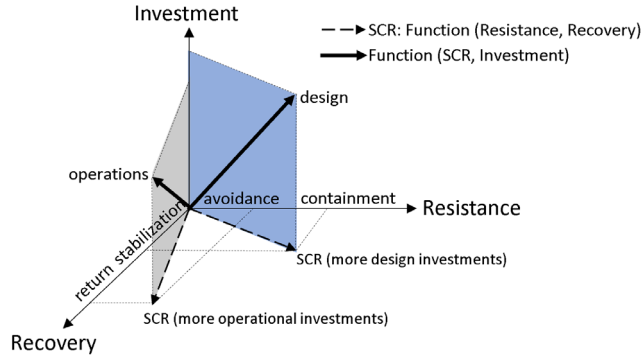


Figure 3. SCR investments in design and operations.

#### 4. Methodology and results

This research proposes a modeling approach for the supply chain resilience (SCR) concerning the BOG generation in the LNG supply chain by evaluating SCR in two distinct stages as shown in Figure 1: resistance and recovery, and their consequential sub-stages of: avoidance and containment (for resistance) and stabilisation and return (for recovery). The surrogates are built to represent dependent variables of interest  $Y$ , such as the SCR. Independent-dependent variations correlate as  $Y_i = f(X_j)$ . The input variables considered to form the surrogates represent the BOG resilience tree as in Table 1. When building surrogate models, it is important to consider some important aspects concerning the required data for training and testing, functional form of the surrogates and proper evaluation to verify their performance. A simplified framework methodology proposed by Franzoi et al. (2021a), as shown in Figure 4.

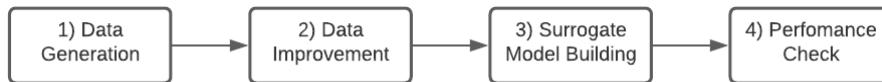


Figure 3: Framework for the surrogate model building strategy.

1) *Surrogate Model Building*: To model the SCR, bilinear surrogate models can be used, as shown in Equation (1) to calculate the dependent variable  $Y_i$  (see Franzoi et al., 2020).

$$Y_i = I_i + \sum_{j \in DV} b_{ij} X_j + \sum_{j \in IV} \sum_{k \in IV} c_{ijk} X_j X_k \quad \forall i \in DV, j \leq k, k \leq n \quad (1)$$

In which  $I_i$  is the intercept of each point  $i$  within dependent variables (DV), and  $b_{ij}$  and  $c_{ijk}$ , are the coefficients to be determined or estimated during the building process of the surrogate model by evaluating the accuracy of the input-output data to achieve the target of minimising the prediction error for each independent variable point (IV). This optimisation target is presented in Eq.(2) for minimising the least-squares error (LSE).

$$\text{Minimize} \sum_{p=1}^{n_{training}} (y_{ip} - Y_{ip})^2 \quad (2)$$

In which  $y_{ip}$  is the actual calculated using a simulation of the for the variable  $i$  at each point  $p$  within the training data set  $n_{training}$ , whereas  $Y_{ip}$  represents the calculated values using the surrogate model estimated coefficients for each independent variable  $i$  at each point  $p$ .



The dataset to construct the surrogate model is determined by concatenating the operational and design investments from Table 1. As example, if no installation of more efficient temperature insulation by using membrane technology instead of MOSS (independent variable  $X_2$ ), it is expected to have losses of the calorific value  $Y$  due to the no perfect control of the temperature (independent variable  $X_1$ ). The SCR modelling in this case is considered investment projects (operational- or design-related) that would avoid and contain the BOG or, if in the case of recovery, those measure that would stabilise or recovery to the initial stage of the calorific value  $Y$  in the beginning of the supply chain, from liquefaction plants to receiving terminals. Equation (3) is the surrogate for the calorific value  $Y$ . Cases of application of this method can be found in chemical engineering for applications on cutpoint optimization (Franzoi et al., 2020).and reactor systems where an adaptive sampling algorithm can converge to parameters of the surrogate formulas for ranged spaces (Franzoi et al., 2021b).

$$\begin{aligned}
 Y_{CV} = & -10X_1 - 5.1X_2 - 3.2X_3 - 10X_4 - 10X_5 + 7.4X_7 - 9.2X_1X_2 + 10X_1X_3 \\
 & - 10X_1X_4 - 10X_2X_3 + 10X_2X_4 + 7.3X_2X_7 - 10X_3X_4 + 10X_3X_5 \\
 & + 10X_3X_6 + 10X_3X_7 - 10X_4X_5 - 10X_4X_6 - 1.4X_4X_7 + 5.5X_5X_7
 \end{aligned} \quad (3)$$

## 5. Conclusion

The main purpose of this paper is to build surrogate models that can effectively determine digital supply chain resilience in considering the boil-off gas in the LNG supply chain. The surrogates can be built to represent the resistance and recovery measures to be managed in the proposed modeling. Considering the generic SCR modeling using MIP-based machine learning for the parent and child components as in the Figure 1 tree (parent: resistance and recovery, children: resistance, recovery, containment, stabilisation), a more complex combination will be proposed for the second-order or cross-product of the components. It will cover the following aspects of the CSR in the LNG supply chain such as (1) physical elements of the LNG shipping: pressure, temperature, BOG losses, etc., and (2) trustworthiness levels of the IoT considering communication, timeliness, etc.

## References

- R. E., Franzoi, B. C., Menezes, J. D., Kelly, J. A., Gut, I. E., Grossmann, 2020, Cutpoint temperature surrogate modeling for distillation yields and properties, *Industrial and Engineering Chemistry Research*, 59(41), 18616-18628.
- R. E., Franzoi, T., Ali, A., Al-Hammadi, B. C, Menezes, 2021a, Surrogate modeling approach for nonlinear blending processes. 1st International Conference on Emerging Smart Technologies and Applications, 1-8. IEEE.
- R. E., Franzoi, J. D., Kelly, B. C., Menezes, C. L. E., Swartz, 2021b, An adaptive sampling surrogate model building framework for the optimization of reaction systems, *Computers and Chemical Engineering*, 152, 107371.
- S. A. Melnyk, D. J. Closs, S. E. Griffis, C. W. Zobel, J. R. Macdonald, 2014, Understanding supply chain resilience. *Supply Chain Management Review*, 18, 34-41.
- B. C. Menezes, J. D. Kelly, A. G. Leal, G. C. Le Roux, 2019, Predictive, prescriptive and detective analytics for smart manufacturing in the information age, *IFAC-PapersOnline*, 52 (1), 568-573.
- B. C. Menezes, J. D. Kelly, A. G. Leal, 2019b, Identification and design of Industry 4.0 opportunities in manufacturing: examples from mature industries to laboratory level systems, *IFAC-PapersOnLine*, 52(13), 2494-2500.
- Man Energy Solutions, 2022, Retrofits for two-stroke marine engines, available on: <https://www.man-es.com/services/industries/marine/retrofit-upgrade/two-stroke-engines>, accessed on: 20 October 2022.
- KBC, a Yokogawa Company, 2021, available on: <https://www.kbc.global/events/transforming-decision-making-in-scheduling-of-terminal-operations>, accessed on 6/10/2022.

# Multi objective decision-making methodologies applied to GREENSCOPE sustainability indicators

Ricardo N. Dias<sup>a</sup>, Rui M. Filipe<sup>a,b</sup>, Henrique A. Matos<sup>a</sup>

<sup>a</sup>CERENA, Instituto Superior Técnico, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal

<sup>b</sup>Instituto Superior de Engenharia de Lisboa, Instituto Politécnico de Lisboa, R. Conselheiro Emídio Navarro 1, Lisboa 1959-007, Portugal

*ricardo.n.dias@tecnico.ulisboa.pt*

## Abstract

The current socio-economic and environmental concerns envisage entities to make sustainable investments and improvements. However, sustainable investment selection in a given industry might not be straightforward from the several technological alternatives available. Several methodologies exist to help in the alternatives assessment and help in the decision-making activity. GREENSCOPE (Ruiz-Mercado et al., 2012), used in this work, calculates 139 normalised sustainability indicators from four different classes: efficiency, environment, energy, and economy. Analysing such indicators for investment alternatives is a complex multi-objective decision-making process (MODM) for which many methodologies exist. GREENSCOPE suggests using the Additive Utility Method (AUM) (Smith & Ruiz-Mercado, 2014), which might have some limitations in identifying different alternatives. An alternative method is implemented in this work, the Analytic Hierarchy Process (AHP) (Saaty, 1978). Both methods were tested and tweaked, originating AHPmod and AUMmod. These were applied to three case studies.

**Keywords:** Sustainability, decision-making, multi-objective, sustainability indicators

## 1. Introduction

The newest IPCC report, released on April the 4th, 2022, focuses on the mitigation of climate change, calling for an improvement in sustainable development and an increase in efficiency (WGIII IPCC, 2022). Therefore, decision-making must take into consideration, not just economic sustainability but environmental sustainability as well. UN sustainable goals is a set of internationally recognized goals that call for sustainable development (UN General Assembly, 2015). The main challenge when evaluating the sustainability of a given project is the assessment of the environmental impacts since they might have a different importance in different regions. For a

region frequently affected by droughts, local industry's water consumption and efficiency might be of higher priority in the agenda against CO<sub>2</sub> emissions mitigation. Therefore, the decision-making process should always consider the socio-economic impacts (Ruiz-Mercado et al., 2012).

GREENSCOPE (GRNS) is a life cycle assessment methodology focused on gate-to-gate analysis that uses a total of 139 indicators. They are spread (not evenly) across 4 metrics: Mass Efficiency, Environment, Energy and Economy.

$$\%G = \frac{|x_i - x_{i,worst}|}{|x_{i,best} - x_{i,worst}|} \quad (1)$$

The normalization technique used on the GRNS methodology implies an indicator score between 0 % and 100 % (max-min normalisation). It makes for a straightforward interpretation of the results, with the user always aiming for a score as close to 100% as possible. Even though the maximum and minimum values assumed by the methodology for the normalisation step might have some degree of uncertainty and subjectivity, it is considered a well-based methodology. Despite the representation for the scores make for eye-appealing graphs, as per Figure 1, it is not easy to quantify the improvement from alternative to alternative.

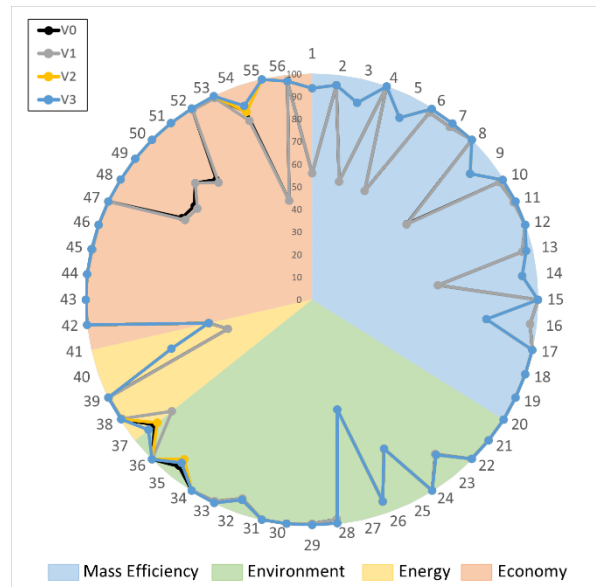


Figure 1 - Graphical representation of acetaldehyde case study GRNS scores.

This work's main goal is to establish a decision-making process based on the indicators calculated using GRNS. Even though the representation of the scores

in spider-like graphs (Figure 1) intends to help the decision-maker by showing that a given alternative with scores closer to the outside boundaries is better than other not as close, a quantification method is also needed.

Smith et al. (2013) use marginal rates of substitution (MRS) to determine the relative importance of any indicator, compared to the others, and AUM to decide upon the best alternative. However, the decision-making process is based on a limited number of indicators that are selected subjectively.

In this work, 2 different methods are implemented, AUM and AHP, together with 2 new variations of these methods. An analysis of the results obtained with each method is performed to decide on which is more suited to this type of multi-objective decision-making (MODM) problems.

## **2. Methodology**

In this section, MODM methodologies, such as AUM and AHP, are applied to GRNS indicators, to decide what is the best alternative for each case study.

### **2.1. Additive Utility Method (AUM)**

GRNS uses a scoring between 0 % and 100 % for each indicator, which may lead to accept that a higher sum of scores indicates the best alternative. However, the increase in one indicator score can come at the expense of another (Smith & Ruiz-Mercado, 2014). AUM is a method that uses the summation of indicator scores combined with MRS.

MRS uses the derivative of one indicator to another, with which it is possible to create an interaction matrix with size  $n$  (where  $n$  is the number of indicators considered for the analysis). From this matrix, one column, to be used as reference indicator, is selected and multiplied by the indicator scoring obtained from GRNS. The resulting vector values are summated, and a final value is obtained. This procedure, using the same interaction matrix from MRS, is repeated for each alternative. Comparison between alternatives is made using the calculated single scores for each alternative.

### **2.2. Analytic Hierarchy Process (AHP)**

AHP, a well-known Multi-Criteria Decision Making (MCDM) method (Saaty, 1994), proposes a systematic methodology to compare multiple possible choices structuring a problem in three parts: goal, possible solutions (alternatives), and criteria (indicators). To determine the best possible solution, one must assign priorities to the various criteria being considered. It can be different for any given problem; Saaty proposes a scale of relative importance from 1 to 9. Using this relative importance is possible to establish a

pairwise comparison matrix  $A$ , which is a  $n \times n$  reciprocal matrix (where  $n$  is the number of indicators taken into consideration).

For each indicator a matrix  $M$  must be calculated, which is a  $m \times m$  matrix (where  $m$  is the number of alternatives considered). And, like matrix  $A$ , it is a reciprocal matrix that is obtained by dividing the scores of every indicator by each other.

Vectors  $P$  and  $p$  (obtained from the geometric mean and normalisation of matrices  $A$  and  $M$ ) are then rearranged in a table, where their values are multiplied and summed to obtain a final score, see eq. (2).

$$Priority = \sum_{i=1}^n P''(i) \times p''_k(i) \quad (2)$$

### **2.3. Additive Utility Method modified (AUMmod)**

Since the AUM method uses a reference indicator arbitrarily chosen, whose selection may be arguable, a modified version that considers all the indicators is herein proposed. This method consists in using the vector  $P''$ , originated from the interaction matrix  $A$  on the AHP method, to balance the scores from GRNS, allowing for a single score.

### **2.4. Analytic Hierarchy Process modified (AHPmod)**

A proposed modification to the AHP method is also presented. An interaction matrix for every alternative with the values of every indicator was calculated, and the same mathematical steps of matrix line geometric mean and normalisation were applied. These two vectors,  $P''$  and new  $p$  are to be multiplied and summed, with the highest value being the best performing alternative.

### **2.5. Implementation**

The four methods previously outlined were implemented using MATLAB. A code was written to obtain the indicators scoring from the GRNS file and make the calculations. The implementation of AUM was straightforward since it had already been used for GRNS indicator scoring comparison.

For the AHP and AHPmod methods implementation, a systematic methodology to overcome the subjective score selection (1-9) was carried out. Therefore, matrix  $A$  was generated not based on the relative difference attributed by the user, but based on the delta calculated through the difference between maximum and minimum possible indicator values, that corresponds to the denominator of eq. (1).

The indicators used in every method were determined via a set of restrictions outlined to ensure calculations feasibility, since MRS uses delta divisions, which cannot be zero. The set of restrictions imposed on the indicator selection are defined in Table 1, as stages.

Table 1. Indicator selection, elimination stages with description.

Stage	Description
0	139 GRNS indicators
1	Remove $\nabla_i < 0.01$
2	Remove $\%G_i < 0.01$
3	Remove $P_i'' < 0.01$
4	Remove $P_i'' = \infty$
5	Remove $p_i'' = \infty$

In stages 4 and 5, the elimination of infinite values on the  $P''$  and  $p_i''$  vectors, is essential for the implementation of AHP and AUM methods. Stage 6 ensures that the same set of indicators is used on every alternative evaluation.

Finally, to facilitate the interpretation, a ranking was attributed to the scenarios evaluated, so if a scenario has a better final score than any other, it is ranked as the best alternative in a scale from 1<sup>st</sup> to 4<sup>th</sup>.

### 3. Case study

To carry out a sustainability analysis, there must be a reference scenario upon which the comparison can take place; otherwise, the score obtained can't be interpreted. In the context of the present work, 3 case studies were evaluated, for which 4 different scenarios were created. Since decision-making can only take place if there is more than one alternative to choose from, in Table 2 is possible to find a quick overlook of scenarios for the different case studies.

Table 2. Case studies and its scenarios.

Scenario	Description		
	Biodiesel plant	Acetaldehyde plant	Propylene Glycol plant
V0	Base scenario	Base scenario	Base scenario
V1	Lower price of feed streams	Energy integration (EI)	Energy integration (EI)
V2	Lower price of product streams	Recirculation, no EI	Recirculation, no EI
V3	Lower energy consumption (EI)	Recirculation with EI	Recirculation with EI

### 3.1. Biodiesel plant

The biodiesel plant was used as case study in GRNS (Ruiz-Mercado et al., 2013). Since the model for this case study was not available, modifications to the original biodiesel demonstration file (V0) were made, as follows: a decrease of 10 % in raw materials cost (V1), a 10 % reduction in products value (V2) and decrease of 10 % in energy consumption (V3). These scenarios were created to better understand the decision-making process based on the indicators calculated.

### 3.2. Acetaldehyde and propylene glycol plant

Acetaldehyde production is a very well-known academic case study, and so is the propylene glycol. These cases are particularly suited for a sustainability analysis since the assessment outcomes can be clearly validated. As for the biodiesel plant, several alternatives were created: base case scenario with no energy integration or material recirculation (V0), energy integration and no material recirculation (V1), recirculation without energy integration (V2) and material recirculation with energy integration (V3). The V3 alternative is expected to be more sustainable, with higher sustainable scores across the GRNS indicators.

## 4. Results

The methods previously described (see paragraphs 2.1 to 2.4) were applied to the GRNS scores obtained for every alternative of the case studies. Since the results obtained through the implementation of the MODM methods are not comparable and have very different orders of magnitude, the results obtained are presented in ranking order in Table 3.

Table 3. Rankings obtained to the alternatives of the case studies.

<b>Biodiesel plant case study</b>
-----------------------------------

	Expected	AUM	AUMmod	AHP	AHPmod
V0	3	4	4	2	<b>3</b>
V1	1	3	3	<b>1</b>	<b>1</b>
V2	4	1	1	3	<b>4</b>
V3	2	<b>2</b>	<b>2</b>	4	<b>2</b>
<b>Acetaldehyde plant case study</b>					
V0	4	4	4	2	3
V1	3	1	1	1	4
V2	2	<b>2</b>	<b>2</b>	3	<b>2</b>
V3	1	3	3	4	<b>1</b>
<b>Propylene glycol plant case study</b>					
V0	4	2	2	2	3
V1	3	1	1	1	2
V2	2	3	3	3	4
V3	1	4	4	4	<b>1</b>

From Table 3 the pattern set by not having material recirculation (V0 and V1) could not be identified by AHP, AUM and AUMmod methods as the worst alternatives. Moreover, those methods also failed in identifying the best and worst performing alternatives.

The AHPmod method ranked the alternatives in the expected order. In the biodiesel case, the ranking corresponds to the expected ranking set. For the acetaldehyde case, two main groups without material recirculation (V0 and V1) are identified as the worst performing alternatives, and ranked 3 and 4, respectively. The order in which V0 and V1 were ranked is not unexpected since energy integration represents capital investment increase, that should be carefully analyzed and justified. It can have a rippling effect across many sustainability indicators. V4 was considered the best-performing one, as expected, since it considers material recirculation and energy integration.

However, in the propylene glycol case, only the best-expected alternative was ranked as anticipated (using the AHPmod). This represents a good accuracy in



the AHPmod method for the best-performing cases, even though it can miss rank other alternatives that do not have a very different scoring.

### 5. Conclusion and future work

Decision-making using sustainability indicators can be a very subjective process. Many methodologies use weights that are determined subjectively to decide on the best and worst-performing alternatives. The work developed and presented in this paper shows that it is possible to use MODM methods to help in sustainability analysis. The proposed methods remove the need for weights and can, with a certain degree of uncertainty, identify a pattern of increasing sustainability. Even though AHP and AUM were not the best performing MODM methods, those are the base for the AHPmod method presented and proposed for further validation. Future work will use TOPSIS method for additional validation of the techniques presented, sensibility analysis of the indicator eliminations steps and Pareto analysis of the vector  $P''$  generated in the MRS reciprocal matrix.

### Acknowledgements

The present work was financed by the Portuguese Foundation for Technology and Science (FCT) PTDC/EAM-PEC/32342/2017 and CERENA strategic project FCT-UIDB/04028/2020. This support is gratefully appreciated.

### References

- Ruiz-Mercado, G. J., Gonzalez, M. A., & Smith, R. L. (2013). Sustainability indicators for chemical processes: III. biodiesel case study. *Industrial and Engineering Chemistry Research*, 52(20).  
<https://doi.org/10.1021/ie302804x>
- Ruiz-Mercado, G. J., Smith, R. L., & Gonzalez, M. A. (2012). Sustainability indicators for chemical processes: I. Taxonomy. *Industrial and Engineering Chemistry Research*, 51(5).  
<https://doi.org/10.1021/ie102116e>
- Saaty, T. L. (1994). How to make a decision: the analytic hierarchy process. *Interfaces*, 24(6), 19–43.
- Smith, R. L., & Ruiz-Mercado, G. J. (2014). A method for decision making using sustainability indicators. *Clean Technologies and Environmental Policy*, 16(4). <https://doi.org/10.1007/s10098-013-0684-5>

UN General Assembly. (2015). UN General Assembly, Transforming our world : the 2030 Agenda for Sustainable Development. In *Resolution adopted by the General Assembly on 25 September 2015* (Vol. 16301, Issue October).

WGIII IPCC. (2022). *IPCC, 2022: Climate Change 2022: Mitigation of Climate Change*.



# Process as a battery: Robust dynamic optimal operation of zeolite crystallization in a COBR with respect to the carbon footprint of electric power

Robin Semrau<sup>a</sup>, Sebastian Engell<sup>a</sup>

*<sup>a</sup>TU Dortmund University, August-Schmidt-Straße 1, 44227 Dortmund, Germany  
Robin.Semrau@tu-dortmund.de*

## Abstract

The process industries are increasingly electrifying their processes in order to reduce the carbon footprint by the use of “green” electric power. However, the supply of power from renewable sources and the prices of electric energy are volatile which creates an opportunity and a challenge. On the one hand, the volatility of the price creates an opportunity to reduce cost and at the same time reduce the CO<sub>2</sub>-footprint as the price and the availability of power from renewables are strongly correlated. On the other hand, the plants must be flexible enough to adapt to the price levels and the operational strategy must become more dynamic. This paper investigates the flexible operation of a continuous tubular reactor for the hydrothermal synthesis of zeolites with the goal to determine the economically optimal operating trajectory in the presence of parametric model-plant mismatch. The robust dynamic optimization strategy is presented and the benefits over the usage of the nominal model are evaluated.

**Keywords:** Demand response, Dynamic optimization, Robust optimization, Process Intensification, Continuous crystallization

## 1. Introduction

The industrial sector accounts for 38 % of the total global energy consumption. Therefore, besides increasing the energy efficiency of the production processes, a change of the primary energy sources from fossil to renewable sources in the industrial sector has to take place. In the chemical industry, the direct electrification of processes is one of the main contributors to this transition. The supply of electric power is fluctuating due to the increasing share of renewable power, while there is a (up to now mostly) steady demand for electric energy from the industrial sector. The synchronization of supply and demand via price signals is referred to as Demand Side Management or Demand Response. This adaptation has multiple benefits. The production company is able to use electric energy in periods of lower prices and therefore has an economic benefit. The grid operator on the other hand profits because supply and demand are balanced better. As a result, the share of fluctuating sources in the grid can be increased, while maintaining grid stability. So overall also the CO<sub>2</sub>-footprint is reduced. The procurement of electrical energy by large consumers is usually realized by longer-term contracts (e.g. time of use tariffs) and by the participation in different power markets, such as the day-ahead spot market the intraday market and the frequency reserves. The time procurement horizons of these markets range from the next day on the day-ahead spot market to 15 minutes before consumption on the intraday market to 30 seconds on the frequency reserves.

Demand Side Management strategies are frequently formulated as Mixed Integer Linear Problems (Zhang and Pinto, 2022, Dalle Ave et al., 2019), where the dynamics of the processes are reduced to minimum up and down times or ramp constraints on the change of the production levels, neglecting the real dynamics of the processes. However, for slow processes and fast load changes, these formulations are too coarse and dynamic models of the behavior of the plants have to be included. This has been investigated by several authors for a variety of examples, e.g. Air Separation Units (Pattison et al., 2016) and a continuous zeolite production plant (Semrau et al., 2022). The economic and environmental advantages of demand side management for the zeolite plant were investigated in detail in Semrau and Engell (2023), assuming a perfect plant model. In this work, the effect of stochastic parameter uncertainties on the optimal day-ahead commitment of the electric energy demand is studied. First, the plant and the optimization problem are described. Then the robust dynamic optimization problem is presented and the results are shown for historic data for the prices of electric energy. Finally, the difference to the usage of a nominal model is quantified.

## 2. Problem statement

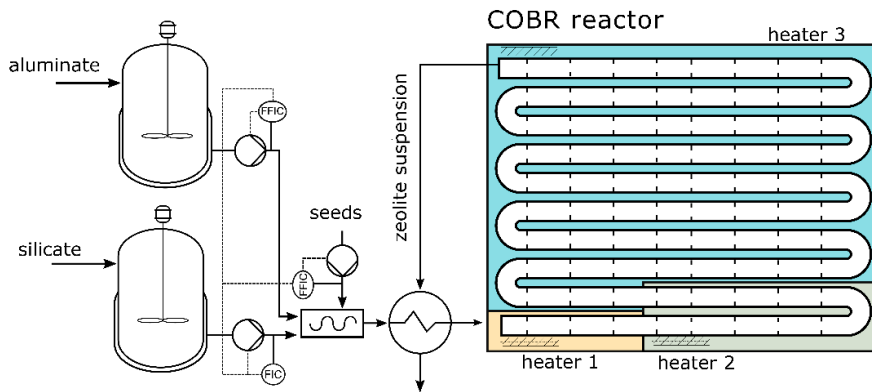


Figure 1: Scheme of the continuous zeolite production process

The process for the continuous production of NaX zeolites is shown in Figure 1. The silicate and the aluminate suspensions are mixed along with seed crystals and fed into the tubular reactor. The reactor is equipped with three heaters along the reactor. In the continuous oscillatory baffled reactor, the amorphous solids are converted to the desired crystalline zeolites. This takes place at elevated temperatures. The outlet stream is used to preheat the entering stream. The inputs of the system are the total flowrate and the heat duty of the three heaters. The zeolite crystal content in the outflow should be higher than 98%. The reaction temperature is limited to 403.5 K, due to the formation of side products. The production process is described in detail in Ramirez Mendoza et al. (2020). The reactor is modelled by the axial dispersion partial differential equation with reaction terms for the dissolution of the amorphous solid and the nucleation and growth of the zeolite particles. To solve the population balance of the zeolite particle formation, the method of moments is used. A general notation of the model is given by (1) with the states  $x$ , the inputs  $u$ , and the parameters  $p$ .

$$\dot{x} = f(x, u, p) \quad (1)$$

### 2.1.1. Parametric uncertainty

Models are only simplified description of the plant behavior. Especially, the nucleation and growth of the crystalline particles is a highly complex, stochastic process. Therefore, the prediction of the plant behavior is prone to errors. In this work, a parametric stochastic deviation from the nominal model prediction is assumed, which can be described by a stochastic differential equation (2) with a linear time invariant matrices  $A$ ,  $a$ ,  $B$  and the Wiener process  $dW$ . The evolution of mean value  $\bar{p}$  and the covariance matrix  $P$  can be described the differential equations (3),(4). The nucleation rate and the heat transfer coefficient of the preheater are assumed to be stochastic. The deterministic description of the system is replaced by a stochastic one.

$$dp(t) = (A p(t) + a)dt + BdW \quad (2)$$

$$\dot{P} = AP + PA^T + BB^T \quad (3)$$

$$\dot{\bar{p}}(t) = A\bar{p}(t) + a \quad (4)$$

$$d \begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} f(x, u, p) \\ Ap + a \end{pmatrix} dt + \begin{pmatrix} 0 \\ B \end{pmatrix} dW \quad (5)$$

### 2.1.2. Economic cost structure

In the considered case, the energy costs and the revenue are considered. Therefore, the costs can be described by the normalized usage of electric energy multiplied by the normalized price  $\gamma(t)$  subtracted by the relative throughput.

$$\phi(t) = \gamma(t) \frac{\sum_j^3 \dot{Q}_j}{\sum_j^3 \dot{Q}_{j,ref}} - \frac{F}{F_{ref}} \quad (6)$$

## 3. Method

In the following, the dynamic optimal schedule based on the day-ahead and the intraday electric energy price markets in the presence of stochastic parametric model-plant mismatch is presented.

### 3.1. The Unscented Transform

The Unscented Transform (Julier 2002) approximates the propagation of a multivariate Gaussian distribution  $\mathcal{N}(m, \Sigma)$  of dimension  $N$  through a nonlinear function  $g(m)$  with individual samples  $m$ , the sigma points. The expected value of the function  $g(m)$  can be approximated by a weighted sum of the sigma points. The individual weights  $w_i$  are chosen, such that the corresponding moments of the distribution are accurate to the third order.

$$m \sim \mathcal{N}(\bar{m}, \Sigma) \quad (7)$$

$$m_i = \bar{m} \pm \sqrt{(N+1)\Sigma_i} \quad (8)$$

$$\mathbb{E}(g(m)) \approx \sum_{i=0}^{2N} w_i g(m_i) \quad (9)$$

### 3.2. Robust dynamic optimization

The optimization problem under uncertainty is formulated in (10). The stage cost is the expected cost of electric energy. For the robustification of the solution, the constraints are satisfied for all sigma points. Similar approaches were used by e.g. Vallerio et.al.(2016). The optimization variables are the extended state vector  $X(t)$ , which consists of the states  $x_i$  for the corresponding sigma points  $p_i$ , and the input trajectory  $u(t)$  which is the same for all sigma points. The terminal cost is introduced to reduce the terminal sell-off effect, it consists of a linear term with  $\lambda_{ss}^{*,T}$  being the Lagrange multiplier at the optimal steady state (Zanon and Faulwasser, 2018). To further reduce the sell-off effect, the optimization problem is solved on two horizons, of which the second one is discarded, similar to model predictive control approaches. The constraints of the optimization are the dynamic model and the bounds on the process variables and the product quality, which need to be satisfied for all sigma points. Furthermore, the mean production is ensured.

$$\min_{x, u, E} \int_0^{t_f} \mathbb{E}(\phi(t)) dt + \lambda_{ss}^{*,T} X(t_f) \quad (6a)$$

$$s. t.: \dot{x} = f(x_i, u, p_i) \quad \forall i \in (0, 2N) \quad (6b)$$

$$x_i(0) = x_{0,i} \quad \forall i \in (0, 2N) \quad (6d)$$

$$0 \geq h(x_i, u, p_i) \quad \forall i \in (0, 2N) \quad (6e)$$

$$0.98 \leq X_i(t) \quad \forall i \in (0, 2N) \quad (6f)$$

$$F_{fix} = \frac{1}{t_{k+1} - t_k} \int_{t_k}^{t_{k+1}} F(t) dt \quad \forall k \in [0, 1] \quad \forall i \in (0, 2N) \quad (6g)$$

Time is discretized with the OCFE method using Radau polynomials of a second order. The optimization problem is implemented CasADi and solved with the NLP solver IPOPT. The length of the first horizon is chosen to be 24 h, the length of the second horizon is 12 h. The resulting nonlinear optimization problem has around 320.000 variables and is solved within 12h.

## 4. Results

For the selection of the sigma points, the mean and the covariance of the parameters are the steady state solution of the equation (2), (3). The resulting mean value of the parameter, is one, parameter covariance matrix is given by

$$P_{ss} = \text{diag}([65.53, 4.08]) 10^{-6}. \quad (7)$$

The tuning parameter of the Unscented Transform was chosen such that the isoprobability curve on which the Sigma Points lie, covers 95% of the distribution. The electric energy price trajectory for the day-ahead electric energy price is data from the EPEX Spot market from the 13.01.2020.

The results of the optimization are shown in Figure 2 for the nominal optimization, in which only the mean is considered and the robust optimization. The normalized electric energy price is shown in the bottom plot. The resulting optimal trajectories for the temperature (T) and crystallinity (X) at the outlet for the different sigma points.

Additionally, the throughput (F) and the heating powers for individual the heaters ( $\dot{Q}_j$ ) are displayed. The energy usage is shifted in both cases to the low energy price period, while the constraints are satisfied for all sigma points. It is clearly observable, that the inputs vary in a smaller range, and the height of input movement is smaller.

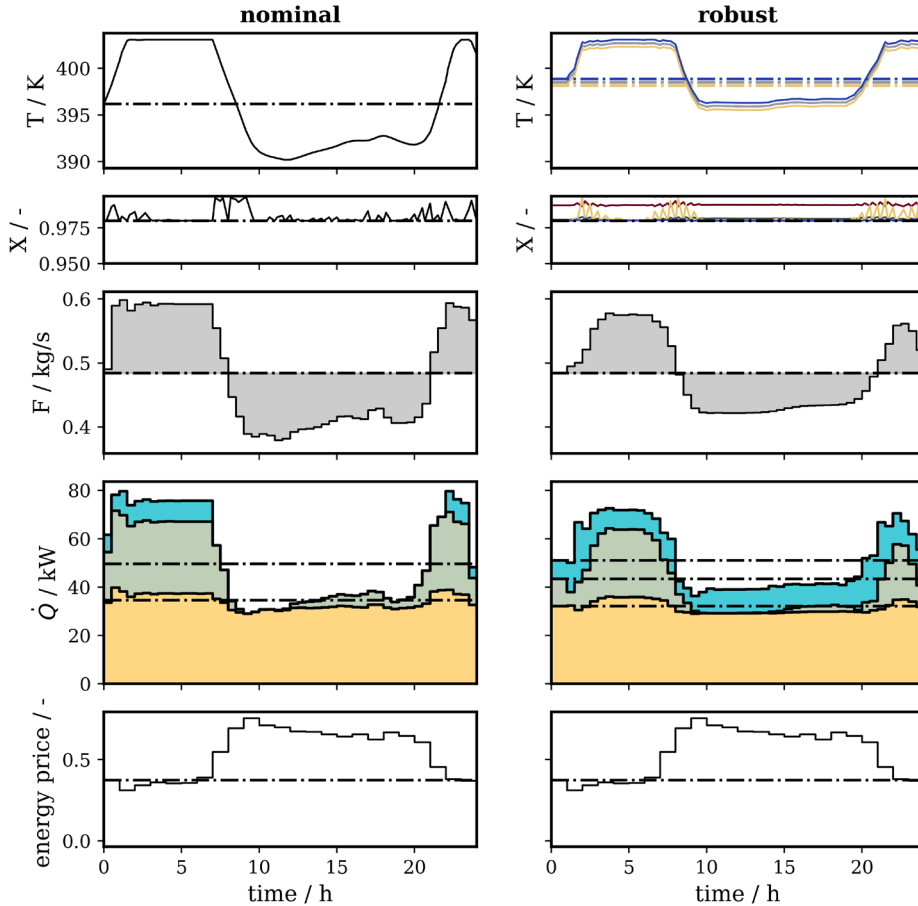


Figure 2: Optimal dynamic operation with energy price data for 13.01.2020 as the result of the robust optimization

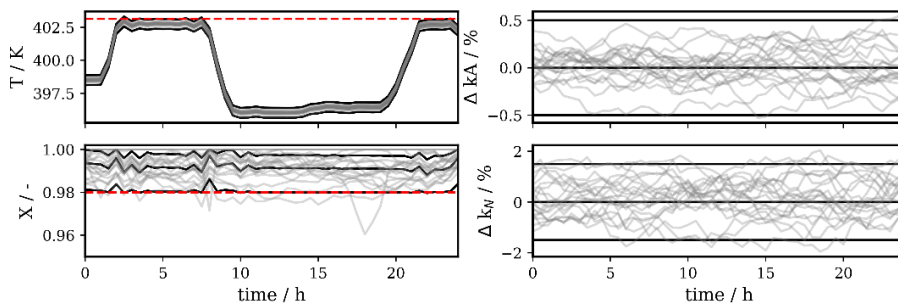


Figure 3: Application of the robust optimal trajectory to the stochastic system



To test the validity of the calculated solution the stochastic differential equation of the model (5) is numerically solved for 20 different realization of the noise the results are shown in Figure 3. In which, the outlet temperature, the crystallinity and the two varying parameters are shown. The grey lines represent the solution of the stochastic differential equation for different realizations, the constraints are marked in red and the considered sigma points in black lines. The robust optimal operation of the process limits the constraint violation to 3.67% instead of 33.26% in the nominal case. However, the economic energy cost savings decrease from 9.29% using the nominal trajectory to 4.13% for the robust optimization.

## 5. Conclusion

This work presents the extension of a moving horizon day-ahead optimization for the flexible usage of electric energy based on the day-ahead market by the inclusion of stochastic parametric model-plant mismatch. The problem statement and the robust optimization problem are introduced. The optimization problem is solved for a one-day operation using past energy price data. The benefit of using the optimization under uncertainty is demonstrated for multiple realization of the stochastic model description, demonstrating an improvement of the robustness with a negative impact on the calculated cost. Further work will address the estimator design to identify the model parameters and robust economic NMPC to be able to reduce the performance loss, while maintaining the robustness of the solution. Furthermore, the application of two stage stochastic programs will be exploited to optimally choose the energy purchase on the different markets.

## References

- Andersson, J.A.E., Gillis, J., Horn, G., Rawlings, J.B., Diehl, M., 2019. CasADi: a software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation* 11, 1–36.
- Dalle Ave, G., Harjunkski, I., Engell, S., 2019. A non-uniform grid approach for scheduling considering electricity load tracking and future load prediction. *Computers & Chemical Engineering* 129, 106506.
- Julier, S.J., 2002. The scaled unscented transformation, *Proceedings of 2002 American Control Conference*, IEEE, Anchorage, AK, USA, pp. 4555–4559 vol.6.
- Pattison, R.C., Touretzky, C.R., Johansson, T., Harjunkski, I., Baldea, M., 2016. Optimal Process Operations in Fast-Changing Electricity Markets: Framework for Scheduling with Low-Order Dynamic Models and an Air Separation Application. *Ind. Eng. Chem. Res.* 55, 4562–4584.
- Ramirez, H., Valdez, M., van Gerven, T., Lutz, C., 2020. Continuous flow synthesis of zeolite FAU in an oscillatory baffled reactor. *Journal of Advanced Manufacturing and Processing* 2.
- Semrau, R., Engell, S., 2023. Process as a battery: Electricity price aware optimal operation of zeolite crystallization in a continuous oscillatory baffled reactor. *Computers & Chemical Engineering* 108143. <https://doi.org/10.1016/j.compchemeng.2023.108143>
- Semrau, R., Yang, J., Engell, S., 2022. Process as a battery: Electricity price based optimal operation of zeolite crystallization in a COBR, in: *Computer Aided Chemical Engineering*. Elsevier, pp. 1165–1170
- Vallerio, M., Telen, D., Cabianca, L., Manenti, F., Impe, J.V., Logist, F., 2016. Robust multi-objective dynamic optimization of chemical processes using the Sigma Point method. *Chemical Engineering Science* 140, 201–216.
- Zanon, M., Faulwasser, T., 2018. Economic MPC without terminal constraints: Gradient-correcting end penalties enforce asymptotic stability. *Journal of Process Control* 63, 1–14.
- Zhang, Q., Pinto, J.M., 2022. Energy-aware enterprise-wide optimization and clean energy in the industrial gas industry. *Computers & Chemical Engineering* 165, 107927.

# Machine Learning models to optimize dairy cow culling: the case of the experimental farm UniLaSalle Beauvais

<sup>a,b</sup>Abdifatah OUSELEYE, <sup>a,c</sup> José RODRIGUEZ, <sup>a</sup> Hamilton ARAUJO \*

<sup>a</sup> Institut Polytechnique UniLaSalle, 19 rue Pierre Waguet, Beauvais 60000, France

<sup>b</sup> Université de Haute-Alsace, 2 rue des Frères Lumière, 68100 Mulhouse, France

<sup>c</sup> Université d'Artois, ULR7519, Beauvais 60026, France

\*[hamilton.araujo@unilasalle.fr](mailto:hamilton.araujo@unilasalle.fr)

## Abstract

In the process of milk production, dairy farmers are confronted with the management of their herd. Indeed, when a group of cattle is in the lactation phase, several genetic and phenotypic factors can impact the milk production (W. L. Weigel, 2017). That way, good herd management can increase production, lower costs, and at the same time reduce the carbon emissions. In this sense, one of the pillars of a good herd management is the renewal rate of the cows (voluntary culling). However, optimize the decision to cull cow is very important. For example, culling a cow with a good milk production potential or keeping a cow in the herd with a non-satisfactory production is not economically profitable. Regarding this, the objective of this work was to process the database of the experimental farm of the UniLaSalle polytechnic institute and try to understand, using machine learning models, the phenotypic and genetic parameters that played an important role in the longevity of cattle cow.

**Keywords:** cow culling; milk production; ai-based decision tool; machine learning; optimization

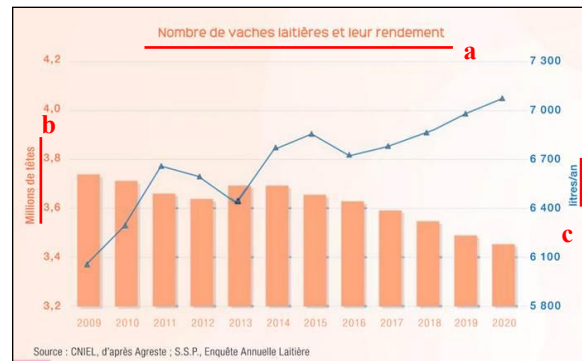
## 1. Introduction

The current global scenario is very unstable and uncertain. In particular, the tensions created by the war in Ukraine, the energy crisis, and the environmental crisis, among other implications, create for instance a pression on the price of raw materials. An important input for dairy cows is soybean, see in the Figure 1 the price evolution (from 208,51 US\$/ metric ton at the end of 2002 to 561,79 US\$ / metric ton at the end of 2022 or an increase of 169% in 10 years)



**Figure 1: U.S. Dollars per metric ton of soybeans in the last 10 years. Source: <https://fred.stlouisfed.org/series/PSOYBUSDQ> (International Monetary Fund)**

In this context, dairy farms face multiple challenges while remaining efficient and competitive. The trivial way to increase milk production on a farm would be to increase the size of the dairy herd. However, this approach will lead to economic and environmental consequences. In that sense, with the advancement of technology and genetic improvement of the herds, the trend observed in the field is the opposite: the number of dairy cattle decreases and productivity increases, as illustrated in Figure 2 (from 2009 to 2020 a 5,40% decrease in the number of cows and a 14,75% increase in milk production in France)



**Figure 2: Milk production per year from 2009 to 2020 vs number of cows in France. Source: CNIEL, Agreste. (a: Number of dairy cows and their yield, b: millions of cows and c: liters of milk per year)**

Knowing that the trend is to reduce the number of cattle and increase production, mastering herd management becomes very important. We should mention 3 aspects of herd management. First, if we keep productive cattle with good genetic and phenotypic parameters, the production tends to increase. Secondly, if we detect the least productive cattle, and the causes of this weak production, we can act to correct these malfunctions and reduce the costs of breeding. Finally, reduce the number of cows and keep the same production yield reduces the environmental impact and the economic. Accord to with (Lengagne, 2019) the average renewal rate of the herds is around 34% and the objective according to the Regional Chamber of Agriculture in 2022 (French Government) is to be less than 25%.

Milk production depends on several phenotypic and genetic factors. Phenotypic parameters are subject to external interferences (stochastic events) caused, as for example, by breeding conditions. For instance, diseases, artificial inseminations, or milk quality yields. These parameters can be explained, and they are influenced, by genetic indexes. In the sense that, the observations of reality are the expressions of the genetic indexes. Precisely, these genetic indexes tell us about the performance of the cows. On the field, the interactions experienced by animals in the environment will contribute positively or negatively to the milk production and it can disturb the veracity of these genetic variables.

On the other hand, with the advancement of data collection, processing and modeling, the use of phenotypic and genetic parameters is widely used to understand the milk production process (Weigel, 2017 and M. Lopez-Suarez et al, 2018). In that sense, our objective is to understand the longevity of a dairy cow. With genetic and phenotypic data can we explain the factors that have contributed, positively and negatively, in the longevity of a dairy cow?

## 2. The environment and the database

### 2.1. UniLaSalle's experimental farm

At UniLaSalle, a French engineering school in agriculture, based in Beauvais, northern region of France, "La Ferme du Bois" is an experimental farm used by the professors and the engineers-students to develop their projects. The farm is functional and productive to deserve the school's community.

### 2.2. The database

We have built up a database from our experimental farm with phenotypic data (milk production data, calving data, diseases data etc.) and genetic data (milk indexes, health index etc). See Figure 3. The phenotypic parameters are collected on the farm and referred in the database either by a farm employee or automatically by the milk robot. Genetic parameters are collected directly from the cattle by an external company and after referred it in the database.

*Example of variables in the database*

532 cows	calving_calving_interval
	diseases and diseases_types
	artificial_inseminations
	milk_protein_rate
	milk_fat_rate
	milk_production_rate
	number_of_calves
	longevity
	milk_production_index
	milk_quality_index
	general_synthesis_index

**Figure 3: some variables in the database.**

An important part of this work was cleaning up and extracting exploitable datasets for training machine learning models. In total we have 532 cattle already culled in the database, however if we would have, for instance, utilize 10 variables to evaluate their impact on the longevity we need enough culled cattle information to have a model.

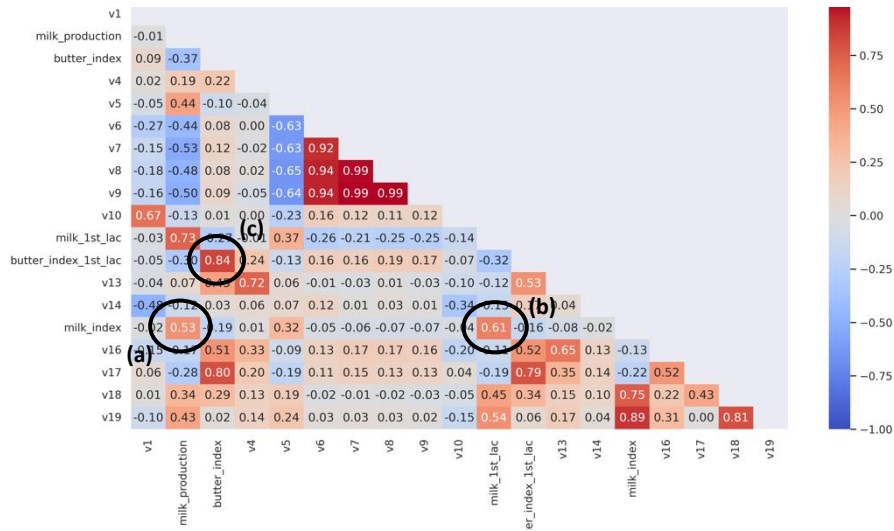
## 3. Methodology and models

In a first step, we have processed this database to verify the veracity of some genetic indexes and its influence in the phenotypic parameters. Secondly, we applied some machine learning models to understand the longevity of already reformed cattle referenced in the database.

For this, we used a python environment, in form of Jupyter Notebook, on Google Colab and we used the basic libraries for manipulation and visualization of tabular data (Pandas, Numpy, Seaborn and Matplotlib). For the predictions we used the models available in Scikit-Learn library (Boyu.Ji, 2022).

### 3.1. Genetic index phenotypic performance

Initially, we have looked for the veracity of some genetic indexes by making a comparison with the collected phenotypic data. In general, we have noticed 3 different instances between the indexes and the phenotypic parameters that we can exemplify with these 3 examples (see figure 4): The milk index (genetic) is 53% correlated with the milk production (phenotypic). Then the milk index (genetic) is 61% correlated with the first lactation milk production (phenotypic). Finally, the butter index is 84% correlated with the butter index of the first lactation (Gabriel M. Dallago et al., 2022).



**Figure 4: (a) milk\_production correlated 53% with the milk\_index (b) milk\_index correlated 61% with the milk production in the first calving (c) butter\_index correlated with 84% the butter\_index in the first lactation**

### 3.2. Machine Learning

We have deployed several supervised machine learning models, essentially linear regression, XGBoost and random forest to predict a phenotypic variable: longevity (A.Essl, 1998). For that, we standardized data by

$$\tilde{x} = \frac{x - \bar{x}}{\sigma}$$

where  $\bar{x}$  is the average of each variable and  $\sigma$  the standard deviation. To train our models we also used a 5-cross-validation test with parameters specific to each model. The data in each model was divided in 70% train et 30% test and the metric utilized to compare and to estimate the precision of the model was the  $R^2$ -metric. Regardless, we have also looked for the feature-importance of the of each prediction because it is important to understand the variables influence the longevity. In fact, initially all variables were considered and after some models were tested, we selected only the most pertinent variables with the objective of optimizing the models.

#### 4. Results

In some senses, we have succeeded in explaining the longevity of a dairy cow with the genetic and phenotypic data was available. The Figure 5 below shows a visual example of the prediction of one of the trained models. The example shows predictions for the longevity of at least 60 cattle. It means that the dataset used initially had 200 cows with enough information for the selected variables (more than 30 variables). Thus, the model was trained with 140 cows (70% training) and was tested with 60 cows (30% test) as shown in the Figure 5.

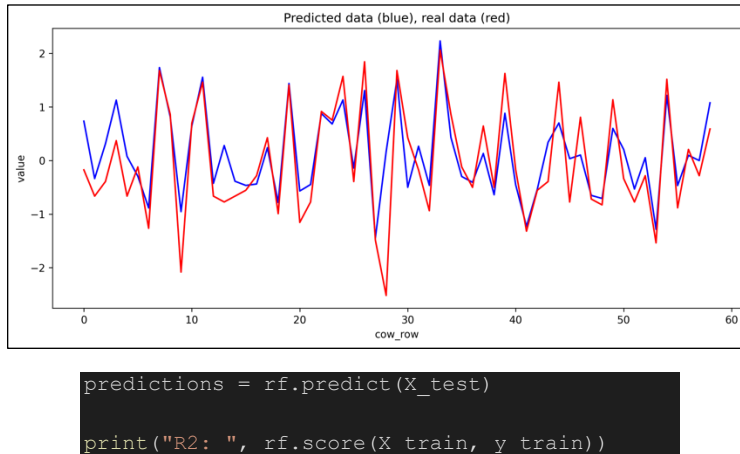


Figure 5: Random Forrest results: blue predicted values and red reality values

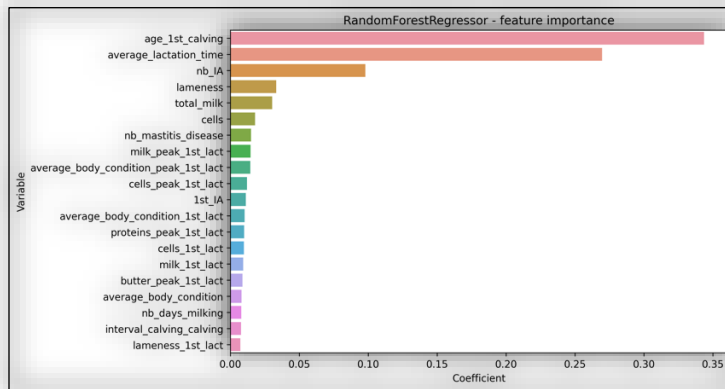


Figure 6: feature importance of the random forest model

In the figure 6 we can see that the age at first calving, the average of the duration of the lactation and the number of artificial inseminations are the most impacting variables concerning longevity. These factors must be evaluated and corrected by the cattle farmer.

#### 5. Discussion and outlook

In terms of improvement, a major difficulty in this work was the data preparation and processing phases. The database has an important number of missing values for several variables. In this case, we reduced the quantity of cows to have more information about

them in the models. Then, expanding our database is extremely important for the future. In addition, predicting longevity is important, but also linking milk production in the model is interesting. For instance, the following synthetic variable can be modeled by the parameters in our database.

$$\frac{m \times 300 \times (l - c_{age})}{l \times 365}$$

where  $m$  is the average of the dairy milk production,  $l$  is the longevity and  $c_{age}$  is the age at first calving. The value 300 is explained by the rest that the cow had after each calving, around two months.

Another axis of improvement is to find a way to use the information from the feature importance to help the cow farmers. It means, not only to have an estimate of the longevity of his cattle, but also what the breeder can act to improve the parameters impacting the longevity. Finally, training a clustering model to classify the cattle to be culled or not is also intended (Esther van der Heide, 2020).

The authors would like to mention that this work was carried out by the first author in the context of his master's thesis.

## References

- W. L. Weigel, 2017, A 100-Year Review: Methods and impact of genetic selection in dairy cattle—From daughter–dam comparisons to deep learning algorithms, *Journal of Dairy Science*, 100, 12, 10234-10250.
- M. Lopez-Suarez et al, 2018, Using Decision Trees to Extract Patterns for Dairy Culling Management. 14th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI), Rhodes, Greece. pp.231-239, <https://hal.inria.fr/hal-01821072>.
- E. M. M. van der Heide, R. F. Veerkamp, M. L. van Pelt, C. Kamphuis, B. J. Ducro, “Predicting survival in dairy cattle by combining genomic breeding values and phenotypic information”, Wageningen University & Research Animal Breeding and Genomics, PO Box 338, 6700 AH, Wageningen, the Netherlands, 2020
- A.Essl, “Longevity in dairy cattle breeding: a review”, Department of Livestock Sciences, University of Agricultural Sciences Vienna, Gregor-Mendel-Strabe 33, A-1180 Vienna, Austria, 1998
- Boyu.Ji, “A machine learning framework to predict the next month's daily milk yield, milk composition and milking frequency for cows in a robotic dairy farm”, 2022
- Gabriel M.Dallago, Roger I. Cue, Kevin M. Wade, René Lacroix, Elsa Vasseur, “Birth conditions affect the longevity of Holstein offspring”, Department of Animal Science, McGill University, Sainte-Anne-de-Bellevue, Quebec H9X 3V9, Canada , 2022.
- Lengagne L., Analyse du coût de renouvellement en élevage bovin lait : Analyse du renouvellement et de la carrière des vaches laitières, coût de production des génisses et identification de marges de progrès. *Sciences du Vivant [q-bio]*. 2019. [ffdumas-02649011](https://doi.org/10.1007/978-2-7005-6264-9_11)
- Ousleye A., Rodriguez J., Araujo H., Machine Learning models to optimize dairy cow culling: the case of the experimental farm UniLaSalle Beauvais, preprint, 2022.

# Analyzing the Large-Scale Supply of Low-Carbon Hydrogen in Germany

Paul Sizaire<sup>a</sup>, Emre Gençer<sup>a</sup>

<sup>a</sup>*MIT Energy Initiative, Massachusetts Institute of Technology, 400 Main Street, Cambridge, MA 02142, USA*

## Abstract

The increasing political momentum advocating for decarbonization efforts, in Europe and elsewhere, has led many governments to unveil national hydrogen strategies. Hydrogen is viewed as a potential enabler of deep decarbonization, notably in hard to abate sectors such as the industry. A linear programming model was developed to assess the procurement of low carbon hydrogen in Germany, a large industrial consumer.

Results show that decarbonizing the current industrial demand (55 TWh/yr) with domestic electrolytic hydrogen would result in a levelized cost of hydrogen totaling ~\$5kg/H<sub>2</sub>, requiring 22 GW of electrolyzer, 30 GW of solar and 15 GW of wind, highlighting the requirement to import. Retrofitting natural gas derived production will likely result in a carbon intensity exceeding low carbon requirements unless emissions from the natural gas supply chain are adequately addressed, engendering significant additional investments. Imports from neighboring countries such as Norway can significantly reduce costs and meet carbon intensity requirements if sourced correctly, leveraging the relatively short distance between countries using an undersea pipeline.

This model highlights the likely importance of imports in decarbonizing hydrogen procurement in Germany to address high costs of domestic electrolytic production and high carbon intensity of retrofitted natural gas derived production.

**Keywords:** hydrogen, supply chain, electrolyzer, linear programming, Germany

## 1. Introduction

Despite the ratification of the Paris agreement by 189 countries in 2015, carbon emissions have remained fairly constant and above 30 Gt per year (IEA, 2021). This has incited calls for further action, which resulted in several countries pledging to attain carbon neutrality by mid-century. An advocated decarbonization pathway promotes the reduction of carbon emissions from electricity generation and the electrification of several end-uses that currently rely on a variety of incumbent fuels. However, electrification may be impaired by feasibility or cost-effectiveness issues in several applications, notably in the industry sector. Among alternatives, a portfolio of low-carbon fuels such as hydrogen and biofuels are considered. Hydrogen's versatility is deemed to make it suitable to various end-uses such as industrial processes, long-haul transport, or renewable energy storage. This has incited several countries worldwide to release national hydrogen strategies, which highlights the potential future role of hydrogen in the decarbonization effort.

In Europe, the current largest consumer of hydrogen is Germany. It almost entirely consumes it in the chemical industry, predominantly for refining, followed by ammonia and methanol production. Germany produces most of its hydrogen, relying very little on imports, with an annual output reaching ~55 TWh. This production is dominated by a



single process – Steam Methane Reforming (SMR), which uses natural gas as a fuel, across 44 plants in the country. The predominance of hydrogen production from SMR is associated with substantial carbon emissions - this process emits  $\sim 10$  tCO<sub>2</sub>/tH<sub>2</sub> (IEA, 2019). To date, Germany possesses less than 200 MW of electrolytic capacity. To comply with the amendment to the Climate Change Act stating a net-zero target by 2045 (Bundestag, 2021), Germany has produced a national hydrogen strategy. The government only considers electrolytic hydrogen produced from renewable electricity to be sustainable in the long term. The strategy focuses on demand applications that are close to commercial viability and those which cannot be decarbonized in other ways, citing the steel and chemical industries. It expects 90-110 TWh of hydrogen demand in 2030. Setting a target of 5 GW of electrolyzers in 2030 aiming to produce 14 TWh of hydrogen using 20 TWh of renewable electricity, Germany is aware that this will not be sufficient to meet the entire demand. As such, Germany views imports as crucial.

The ambitious targets are likely to face several challenges. Current global electrolytic capacity is dwarfed by the sheer scale of envisioned deployment - reaching Germany's target corresponds to a 75% compound annual growth rate until 2030. Electrolytic hydrogen may be expensive in the short- to medium- term which may impair competitiveness in the industrial sector, current incumbent of hydrogen demand. Furthermore, this production requires low-carbon electricity to provide low-carbon hydrogen, triggering an extensive addition of renewable energy in the German electricity grid, which is currently characterized by a relatively large carbon intensity. The creation of a holistic supply chain will also require transmission pipelines and large-scale storage.

To provide insights regarding the least cost scenario for a developed hydrogen supply chain while addressing the various concerns aforementioned, an optimization model has been developed. Using a linear approach, a multi-nodal system representing Germany's regions optimizes for electricity generation, hydrogen generation, transmission, and storage. Hydrogen produced in Norway sent through a pipeline is evaluated as a method of import. Several power generation, hydrogen production, and storage options are considered. This model aims to provide insights regarding the development of a hydrogen supply chain realistically, by implementing several objectives laid out in the German hydrogen strategy.

## 2. Methodology & Case Study

### 2.1. Overview

The model presented in this report is a multi-nodal system that seeks to optimize for a hydrogen supply chain comprising power and hydrogen production, transmission, and storage. It is developed from a macroeconomic modeler's perspective, thus optimizing for social welfare by minimizing the overall costs. A perfect foresight is adopted using power generation data, which will be described subsequently.

Each node in the system represents a region of Germany. Within all nodes, capacities of renewable generation from solar and wind, electrolyzers, Auto Thermal Reforming (ATR) and SMR plants, and hydrogen and electricity storage are optimized for. Power generation is assumed to be wholly used for powering electrolyzers and compressors for hydrogen storage and transmission. A portion of the electricity that cannot be used is therefore curtailed, not sent into the grid. Most of the infrastructure considered in this analysis is greenfield. The cost of electricity omits any taxes, according to the stated

objective that will see the EEG surcharge non-applicable to electrolyzers (The Federal Government, 2020). The relative share of wind and solar resources to install is optimized for depending on the regions in which this production is located.

Transmission is optimized for both electricity and hydrogen. This model also performs a life cycle analysis, which is accounted for in the objective function via a carbon tax. It includes emissions for all infrastructure installed, which is assumed to linearly scale with increasing capacity. It notably includes carbon emissions not captured by CCS from ATR and SMR. Hydrogen storage is assumed to be provided by salt caverns. The technical potential of salt caverns has been evaluated and reported in several studies (Michalski et al., 2017; Welder et al., 2018). Hydrogen demand is assumed to come exclusively from industrial applications – as such, demand is considered to be flat throughout the year. Regional demand reflects their current relative hydrogen consumption.

## 2.2. Objective Function, Variables & Constraints

The objective function aims to minimize the annualized costs of the system, which comprises capital and operational expenditures of assets, fuel, and emissions costs. A set of electrical production infrastructure is denoted  $P_e$  and comprises solar PV, wind turbines, inverters, and electrical storage (Li-Ion batteries). Another set of hydrogen production/storage infrastructure is denoted  $P_h$  and comprises electrolyzers, ATR plants, storage and compressors, while  $P_{ng}$  contains SMR plants only. A set  $T$  denotes transmission, encompassing electrical cables as well as domestic and international hydrogen transmission. Finally, set  $E$  corresponds to electrolyzers, solar, and wind, while  $C$  includes SMR, ATR, and Norwegian hydrogen.  $Cap$  denotes the capacity of the assets,  $DF$  is a detour factor,  $AF$  is an annualized factor, and  $RCC$  is the retrofit cost. The resulting objective function is composed of several cost components, laid out below:

Capacity:	$CC = \sum_{p_e \in P_e} CapEx_{p_e} \times (AF_{p_e} + OpEx_{p_e}) \times \sum_{n \in N} Cap_{p_e, n} + \sum_{p_h \in P_h} CapEx_{p_h} \times (AF_{p_h} + OpEx_{p_h}) \times \sum_{n \in N} Cap_{p_h, n} + \sum_{t \in T} CapEx_t \times (AF_t + OpEx_t) \times \sum_{n_{in} \in N} \sum_{n_{out} \in N} Cap_{t, n_{in}, n_{out}} * DF_{n_{in}, n_{out}}$	(1)
Retrofit:	$RC = \sum_{p_{ng} \in P_{ng}} RCC_{p_{ng}} \times AF_{p_{ng}} \times \sum_{n \in N} Cap_{p_{ng}, n}$	(2)
Fuel Cost:	$FC = Production_{ng} * Cost_{ng} + Production_{Norway} h_2 \times Cost_{Norway} h_2$	(3)
Em. Costs:	$EC = \sum_{e \in E} Emissions_e \times Cap_e + \sum_{c \in C} Emissions_c \times Production_c$	(4)
<b>Objective:</b>	$OF = CC + RC + FC + EC$	(5)

The variables in the model relate to the capacities of all technologies, electricity flows, hydrogen flows, and hydrogen and electricity storage levels. The capacities are optimized for every node, while all other variables are additionally optimized hourly, enabling great insights into the temporal operation of the various flows within the supply chain. Three main constraints dictate the hourly production of electricity, associated flows, and electrolytic production ( $CF_{n,h}$  = hourly capacity factor at node  $n$  and time  $h$ ,  $F$  = flow,  $H$  = hours in a year,  $EE$  = electrolyzer efficiency):

Production:	$Gen_{elec, n, h} = Capacity_{wind, n} \times CF_{wind, n, h} + Capacity_{solar, n} \times CF_{solar, n, h} \quad \forall n \in N, h \in H$	(6)
flows:	$Gen_{elec, n, h} = F_{el \rightarrow etlz, n, h} + F_{el \rightarrow curtail, n, h} + F_{el \rightarrow storage, n, h} + \sum_{n_{out} \in N} F_{el \rightarrow transmission, n_{in}, n_{out}, h} \quad \forall n \in N, h \in H$	(7)
Electrolyzer:	$Gen_{n_2, n, h} = (F_{el \rightarrow etlz, n, h} + F_{el \rightarrow storage \rightarrow etlz, n, h} + \sum_{n_{in} \in N} F_{el \rightarrow transmission, n_{in}, n_{out}, h}) \times EE \quad \forall n \in N, h \in H$	(8)

Resulting electricity and hydrogen flows are subsequently used to model the operation of the supply chain, such as the compression, transmission, storage, and so on. They also set

the required capacity variables with a variation of the constraint  $flow_{n,h} \leq capacity_n \forall n \in N, h \in H$ .

### 2.3. Case Study

The system analyzed in this report is Germany, in which 12 regions are considered and modelled as nodes. The focus of this analysis consists of short- to medium-term decarbonization of hydrogen production – as such, most costs are not projections but current values. Similarly, demand was assumed to equal the current demand of 55 TWh/yr, kept instead of the projected 100 TWh/yr due to large uncertainties. This demand was spread across regions according to the industrial hubs using data from Welder et al. (2018). The cost parameters of the model consist of the CapEx, OpEx, and lifetime of all assets considered, gathered from public sources. Emission values were also obtained from publicly available sources. Hourly renewable production data for wind and solar were gathered from the RE-Europe dataset, which provides granular hourly capacity factors for both technologies (Jensen and Pinson, 2017). Finally, a land eligibility analysis was performed to define thresholds for solar, wind, and salt caverns availability regionally.

An initial scenario considered a theoretical edge case in which all production is ensured with domestic electrolyzers, to illustrate the potential required infrastructure deployment and the associated costs entailed. Another more realistic scenario caps the electrolytic production to 14 TWh/yr (as per the German strategy target for 2030), while the remainder of the production comes from either new ATR plants or retrofitted SMR plants. A final scenario allows imports to occur between Norway and Germany.

## 3. Results

### 3.1. First Scenario: Self-Sufficiency with Electrolytic Hydrogen Only

In this situation, all hydrogen production stems from electrolyzers. The resulting supply chain as well as the LOCH<sub>2</sub> are illustrated in Figure 1. The resulting LCOH<sub>2</sub> is 5.12 €/kgH<sub>2</sub>. The required renewable energy/electrolyzer capacity oversize ratio is ~2, resulting in a total of 11% of electricity curtailed. Total electrolytic capacity reaches 22 GW, which is more than twice as much as what Germany plans to build by 2040. By 2021, installed solar capacity totaled 58 GW and wind capacity reached 64 GW in Germany (IRENA, 2022). Such a system would therefore monopolize a significant portion of renewable capacity. This edge case highlights the requirement to either import or continue producing domestically using other processes.

Hydrogen transmission accounts for less than 1% of the total costs. While hydrogen storage negligibly participates in the levelized cost, the required capacity is significant. The technical potential, evaluated at ~200 TWh for Germany, heavily relies on assumptions – actual potential might be closer to the order of magnitude of the optimum capacity. The cost of storage is also notably complicated to quantify and bears significant uncertainties. Since the deployment of such assets is unlikely in the short-term, storage may have to be performed by tanks, which would drive up the levelized cost.

### 3.2. Addition of Natural Gas Derived Hydrogen with New Plants and Retrofits

A further self-sufficiency scenario was run with a cap on electrolytic hydrogen production set to 14 TWh/yr. The remainder of the demand, or 41 TWh/yr, is fulfilled by natural gas derived hydrogen production. Two cases were considered: all fossil production came from new ATR plants, or all fossil production came from SMR retrofits.

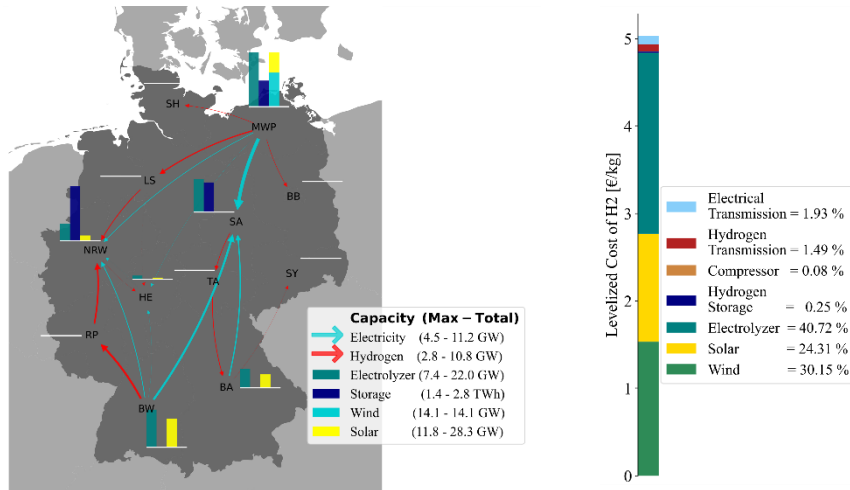


Figure 1: Supply chain required if Germany were to produce all its hydrogen with electrolyzers. “Max” represents maximum capacity between two regions (for electricity and hydrogen transmission) and within a region (for all other technologies), and “Total” represents overall capacity.

The resulting carbon intensities are laid out in the following table:

Table 1: Emission intensity for ATR (91% capture) and SMR (85%)

	ATR [kgCO <sub>2,eq</sub> /kgH <sub>2</sub> ]	SMR [kgCO <sub>2,eq</sub> /kgH <sub>2</sub> ]
Process (91% capture for ATR, 85% capture for SMR)	0.62	1.38
Indirect upstream (compression, transport infrastructure and flaring)	1.79	2.99
Electricity	1.44	1.77
Leaks and venting, production and transmission	1.35	2.25
<b>Total</b>	<b>5.20</b>	<b>8.38</b>
<b>Emissions per Year</b>	<b>6.39 MtCO<sub>2</sub>/yr</b>	<b>10.31 MtCO<sub>2</sub>/yr</b>

This assumes 9 gCO<sub>2,eq</sub>/MJ of natural gas for indirect upstream emissions (Bauer et al., 2021), 400 gCO<sub>2,eq</sub>/kWh intensity from the electricity grid, 1.5% methane leakage rate along the supply chain, a natural gas pipeline length of ~2000 km, and a 100 year global warming potential for methane. In contrast, electrolytic hydrogen has a carbon intensity of 0.8 kgCO<sub>2</sub>/kgH<sub>2</sub>. While direct emissions from the process are greatly reduced (down from ~10 kgCO<sub>2</sub>/kgH<sub>2</sub>), these results highlight the need to adequately reduce upstream emissions both from the natural gas supply chain as well as from the electricity.

### 3.3. Addition of Norwegian Hydrogen

This scenario considers imports of Norwegian hydrogen. This import is considered exogenously – hydrogen is assumed to be available at a constant rate and for a certain price (2€/khH<sub>2</sub>). Once again, electrolytic production is capped at 14 TWh/yr, implying 41 TWh/yr of imports. Imports occur through an undersea pipeline. The results show that a pipeline with a 4.7 GW capacity is required, dispatching hydrogen at a constant flow rate. An equal total capacity of domestic pipelines (~10.5 GW) is required in this scenario as would be in the 100% electrolysis case. Storage requirements are greatly decreased, highlighting the benefits of a steady source of hydrogen as opposed to an intermittent production.

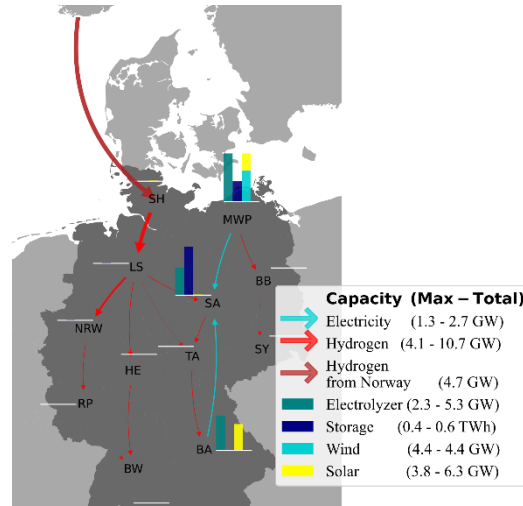


Figure 2: Hydrogen supply chain with 41 TWh/yr of imports from Norway.

Even though the levelized cost of electrolytic remains close to 5 €/kgH<sub>2</sub>, sourcing cheaper hydrogen from imports would in this case reduce the average levelized cost to 3.1 €/kgH<sub>2</sub>.

#### 4. Conclusion

This paper demonstrated that reaching hydrogen production self-sufficiency in a decarbonization context would face several limitations: cost hurdles, capacity limitation, both for electrolyzers and renewable technologies, and scale of natural gas plant retrofit required. Imports are therefore likely to become central to decarbonize procurement and should be further evaluated.

#### References

- Bauer, C., Treyer, K., Antonini, C., Bergerson, J., Gazzani, M., Gencer, E., Gibbins, J., Mazzotti, M., McCoy, S.T., McKenna, R., Pietzcker, R., Ravikumar, A.P., Romano, M.C., Ueckerdt, F., Vente, J., Spek, M. van der, 2021. On the climate impacts of blue hydrogen production. *Sustain. Energy Fuels* 6, 66–75.
- Bundestag, 2021. Climate Change Act - climate neutrality by 2045.
- IEA, 2021. Global Energy Review: CO<sub>2</sub> Emissions in 2020. International Energy Agency, Paris.
- IEA, 2019. The Future of Hydrogen - Analysis. International Energy Agency, Paris.
- IRENA, 2022. Renewable Capacity Statistics 2022. International Renewable Energy Agency.
- Jensen, T.V., Pinson, P., 2017. RE-Europe, a large-scale dataset for modeling a highly renewable European electricity system. *Sci. Data* 4, 170175.
- Michalski, J., Bünger, U., Crotogino, F., Donadei, S., Schneider, G.-S., Pregger, T., Cao, K.-K., Heide, D., 2017. Hydrogen generation by electrolysis and storage in salt caverns: Potentials, economics and systems aspects with regard to the German energy transition. *Int. J. Hydrog. Energy*, Special Issue on The 21st World Hydrogen Energy Conference (WHEC 2016), 13-16 June 2016, Zaragoza, Spain 42, 13427–13443.

The Federal Government, 2020. The National Hydrogen Strategy. Federal Ministry for Economic Affairs and Energy, Berlin.

Welder, L., Ryberg, D.S., Kotzur, L., Grube, T., Robinius, M., Stolten, D., 2018. Spatio-temporal optimization of a future energy system for power-to-hydrogen applications in Germany. *Energy* 158, 1130–1149.



# Equilibrium sensitivity analysis of Carbon dioxide hydrogenation to methanol reaction system and comparison with conversions achievable in a simplified packed bed reactor model.

Milton D. Cárdenas<sup>a</sup>, Camilo Rengifo<sup>a</sup>, Manuel Figueredo<sup>a</sup>

<sup>a</sup>*Facultad de Ingeniería, Universidad de la Sabana, Campus del Puente del Común, Km. 7, Autopista Norte de Bogotá. Chía, Cundinamarca, Colombia.*

Contacting email: [manuel.figueredo@unisabana.edu.co](mailto:manuel.figueredo@unisabana.edu.co)

## Abstract

An equilibrium analysis of the reaction system for the synthesis of methanol based on CO<sub>2</sub> & CO hydrogenation and Water-Gas Shift reaction was performed to define the conditions under which the reaction system favours methanol production. The results were compared with a sensitivity analysis in a simplified Tubular Packed Bed Reactor (TPBR) with commercial Cu/Zn/Al<sub>2</sub>O<sub>3</sub> catalytic bed. As a result, the highest feedstock conversion to methanol at equilibrium was found with increasing pressure, decreasing temperature, and setting the initial H<sub>2</sub>/CO<sub>2</sub> ratio to 2.55. The reactor model showed similar behaviour in pressure and H<sub>2</sub>/CO<sub>2</sub> inlet ratio but differed in the temperature effect on feedstock conversion to methanol. Differences between equilibrium analysis and final reactor conversions are briefly discussed.

**Keywords:** CO<sub>2</sub> hydrogenation, Equilibrium analysis, Tubular packed bed reactor (TPBR) simulation, Sensitivity analysis.

## 1. Introduction

Methanol is an important commodity commonly used in the chemical industry to produce diverse value-added chemicals. According to the analysis made by Sonthalia (2021), this molecule is the basis for synthesizing diverse important compounds such as alcohols, dimethyl ether (DME), formaldehyde, methyl tertiary butyl ether (MTBE), acetic acid, and methyl formate among others. This molecule is not only used not only as a feedstock but also as a fuel in internal combustion engines with great environmental advantages since it can be produced more easily from renewable energies than commonly used fuels (Verhelst et al., 2019). In the work made by Khojasteh-Salkuyeh (2021) are established some used methanol production pathways in the industry such as natural gas conventional reforming, dry-reforming, tri-reforming or direct carbon dioxide hydrogenation. Among them, implementing direct CO<sub>2</sub> hydrogenation can carry out some advantages compared with natural gas reforming (widely used production method), like providing a more controlled reaction process due to a lower temperature in the reaction and a favourable environmental analysis of net CO<sub>2</sub> emissions when using green H<sub>2</sub>. The direct CO<sub>2</sub> hydrogenation to methanol is a reaction system highly dependent on equilibrium conditions, thus, this study aims to state the implication of temperature, pressure, and inlet feedstock ratio on the methanol production from an equilibrium approach with actualized data and relate these results with a sensitivity analysis in a simplified tubular packed bed reactor (TPBR).

## 2. Model Formulation

### 2.1 Model for species concentration at equilibrium

The carbon dioxide hydrogenation to methanol is an exothermic reaction system  $\Delta H_r = -49.4 \text{ kJ/mol}$ . Defined by Graaf (1986) as:

Reaction	Name	Abbreviature	$\Delta H_{t@298K}$
$\text{CO} + 2\text{H}_2 \leftrightarrow \text{CH}_3\text{OH}$	CO Hydrogenation	$r_1$	-90.5 kJ/mol
$\text{CO}_2 + \text{H}_2 \leftrightarrow \text{CO} + \text{H}_2\text{O}$	Water Gas Shift (WGS)	$r_2$	41.1 kJ/mol
$\text{CO}_2 + 3\text{H}_2 \leftrightarrow \text{CH}_3\text{OH} + \text{H}_2\text{O}$	CO <sub>2</sub> Hydrogenation	$r_3$	-49.4 kJ/mol

Table 1 – Reaction system for the synthesis of methanol from carbon dioxide.



The molar fraction of a component in a gas system when the equilibrium is reached  $(y_i)_{eq}$  depends on the equilibrium constants  $(K_{eq,j})$ , the pressure (P) and temperature (T) according to the implicit relation in eq.1.

$$0 = K_{\phi,j}(T, P, y_i(\epsilon_j))_{eq} \cdot K_{y,j}(y_i(\epsilon_j))_{eq} \cdot P^{\alpha_j} - K_{eq,j}(T) \quad \text{eq.1}$$

The equilibrium constants can be calculated with a polynomial temperature-dependent expression as presented in eq.2, the constants for this equation were taken from Graaf (2016).

$$\ln K_{eq,j}(T) = (1/RT) \cdot (a_1 + a_2T + a_3T^2 + a_4T^3 + a_5T^4 + a_6T^5 + a_7T \ln T)_j \quad \text{eq.2}$$

The non-ideal behavior of gas species was corrected by calculating the fugacity coefficients  $(\hat{\phi}_i)$  from the modified Soave-Redlich-Kwong equation of state (SRK-EoS) proposed by Van Bennekom (2012) and replacing these values in eq.3.  $K_{y,j}$  was calculated with eq.4.

$$K_{\phi,j} = \prod_{i=1}^n (\hat{\phi}_i(T, P, y_i(\epsilon_j))_{eq})^{\alpha_{ij}} \quad \text{eq.3}$$

$$K_{y,j} = \prod_{i=1}^n (y_i(\epsilon_j))_{eq}^{\alpha_{ij}} \quad \text{eq.4}$$

Note that the composition is expressed as a function of the reaction extent  $y_i(\epsilon_j)$  and this is useful to avoid extra degrees of freedom since each variation in species concentration follows a stoichiometric relation. Eq.5-6.

$$y_i(\epsilon_j) = \frac{n_{i0} + \sum_{j=1}^k \alpha_{ij} \epsilon_j}{n_T} \quad \text{eq.5}$$

$$n_T(\epsilon_j) = n_0 + \sum_{i=1}^n \sum_{j=1}^k \alpha_{ij} \epsilon_j \quad \text{eq.6}$$

Finally, by finding reaction extents values that make true the equality of eq.1, is possible to calculate the equilibrium compositions of each specie when temperature, pressure, and initial compositions changes.

## 2.2 Model for the tubular packed bed reactor

It is important to highlight that all variables affecting the concentration at equilibrium in eq.1  $(T, P, y_i)$  can only be manipulated indirectly by affecting conveniently the reactor design variables and input conditions. To compare this ideal equilibrium approach to a more accurate analysis, it was proposed a simplified model for a tubular packed bed reactor (TPBR) and the mathematical model is summarized in Table 2.

<b>Mass Balance</b>	$\frac{dF_i}{dW} = r'_i = \sum_{j=1}^2 c_{ij} \cdot r'_j$	eq.7
<b>Energy Balance</b>	$\frac{dT_R}{dW} = \frac{UA_s(T_j - T_R) + \sum_{j=1}^2 -\Delta H_j \cdot r_j}{F_{tot} \cdot C_{p,m}}$	eq.8
	Co-Current Jacket Configuration $\frac{dT_j}{dW} = \frac{UA(T_R - T_j)}{\dot{m}_c C_p}$	eq.9
	Counter-Current Jacket Configuration $\frac{dT_j}{dW} = \frac{-UA(T_R - T_j)}{\dot{m}_j C_{pj}}$	eq.10
<b>Momentum Balance</b>	$\frac{dP_R}{dW} = -\frac{G}{\rho_g g_c D_p} \left( \frac{1-e}{e^3} \right) \left[ \frac{150(1-e)\mu}{D_p} + 1.75G \right] * \frac{1}{(1-e)A_c \rho_c n_t}$	eq.11

Table 2 – Mathematical model for a Tubular Packed Bed Reactor.

## 2.4 Model for properties estimation

The properties of the gas mixture (gas density, volumetric flow, heat capacity and viscosity) were calculated or estimated under ideal conditions considering the temperature and

pressure in which the reactor simulation was developed (175°C - 255 °C and 25 bar – 60 bar respectively).

### 2.5 Kinetic model

The kinetic model was taken from Graaf (1988) and is summarized in Table 3. This kinetic model is based on widely used Cu/Zn/Al<sub>2</sub>O<sub>3</sub> catalysts. This kinetic study implies no mass diffusion limitations inside the catalyst particle and no mass transfer limitation in the solid-gas interface.

$r_1$	=	$k_1 b_{CO} \cdot \frac{\left( P_{CO} P_{H_2}^{3/2} - \frac{P_{CH_3OH}}{P_{H_2}^{1/2} K_1} \right)}{\left( 1 + b_{CO} P_{CO} + b_{CO_2} P_{CO_2} \right) \cdot \left( P_{H_2}^{1/2} + b_{H_2O} / b_{H_2}^{1/2} P_{H_2O} \right)}$	eq.12
$r_2$	=	$k_2 b_{CO_2} \cdot \frac{P_{CO_2} P_{H_2} - P_{CO} P_{H_2O} / K_2}{\left( 1 + b_{CO} P_{CO} + b_{CO_2} P_{CO_2} \right) \cdot \left( P_{H_2}^{1/2} + b_{H_2O} / b_{H_2}^{1/2} P_{H_2O} \right)}$	eq.13
$r_3$	=	$k_3 b_{CO_2} \cdot \frac{\left( P_{CO} P_{H_2}^{3/2} - \frac{P_{CH_3OH} P_{H_2O}}{P_{H_2}^{3/2} K_3} \right)}{\left( 1 + b_{CO} P_{CO} + b_{CO_2} P_{CO_2} \right) \cdot \left( P_{H_2}^{1/2} + b_{H_2O} / b_{H_2}^{1/2} P_{H_2O} \right)}$	eq.14

Table 3 – Kinetic model of the reaction system for CO<sub>2</sub> hydrogenation to methanol.

### 3. Computational and simulation aspects

For solving the reaction extents that make true the equilibrium relation in eq.1, it was used a trust-region algorithm with finite central numerical evaluation provided by Matlab integrated function *fsolve*. In this numerical process is present a high instability since  $K_{\phi,j}(T, P, \epsilon_j)$  and  $K_{\gamma,j}(\epsilon_j)$  (eq.3-4) are defined by a product operator, and it is necessary to rearrange these equations to avoid singularities in the numerical evaluation. The reactor model was simulated with the Matlab function *ode45* and the optimization of both equilibrium and reactor was performed by *fmincon* Matlab function.

The variable to evaluate the effect of inlet conditions (temperature, pressure, and initial H<sub>2</sub>/CO<sub>2</sub> ratio) on methanol concentration when equilibrium is reached was the feedstock conversion to methanol ( $X_C$ ). The meaning of this response variable is the total conversion of hydrogen and carbon dioxide to methanol when the equilibrium is reached compared to the stoichiometric quantities (eq.15).

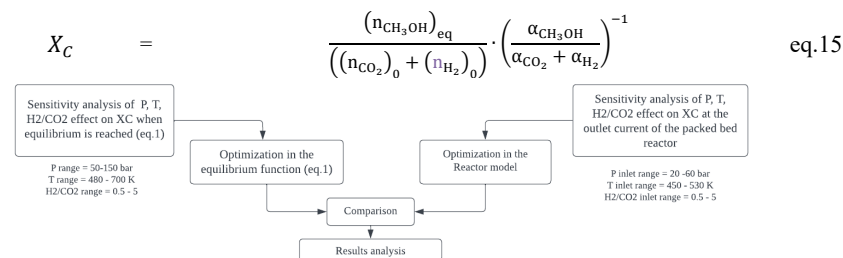


Figure 1 – General diagram to compare the effect of inlet temperature, pressure and H<sub>2</sub>/CO<sub>2</sub> on both current and equilibrium.

## 4. Results

### 4.1 Sensitivity analysis and optimization of feedstock conversion at equilibrium

The temperature and pressure effect on  $X_C$  were evaluated according to values in Table 4. As a result, it was evidenced a higher conversion of carbon dioxide and hydrogen to methanol at lower temperatures and higher pressures (Figure 2a). When evaluating the inlet H<sub>2</sub>/CO<sub>2</sub> ratio effect on  $X_C$ , it was evidenced a maximum value of feedstock conversion to methanol in a H<sub>2</sub>/CO<sub>2</sub> ratio equal to 2.55 (Figure 2b).

Constrained variables	Evaluation range		Optimal Values
	Min	Max	
Temperature	480 K	700 K	480 K
Pressure	50 bar	150 bar	150 bar
H <sub>2</sub> /CO <sub>2</sub> ratio	0.5	5	2.55
			$X_c = 0.79$

Table 4 – Optimal temperature, pressure and H<sub>2</sub>/CO<sub>2</sub> ratio to maximize feedstock conversion to methanol at equilibrium in the constrained numerical optimization of eq.1.

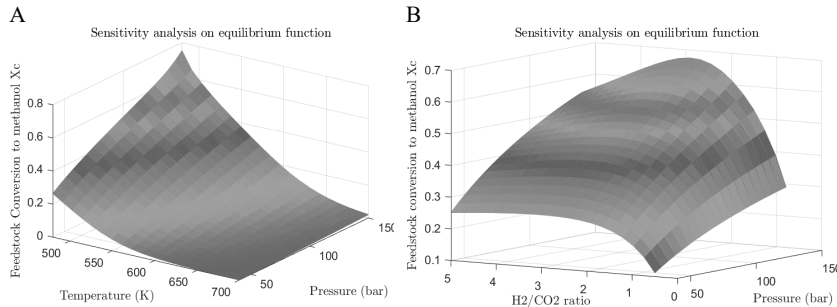


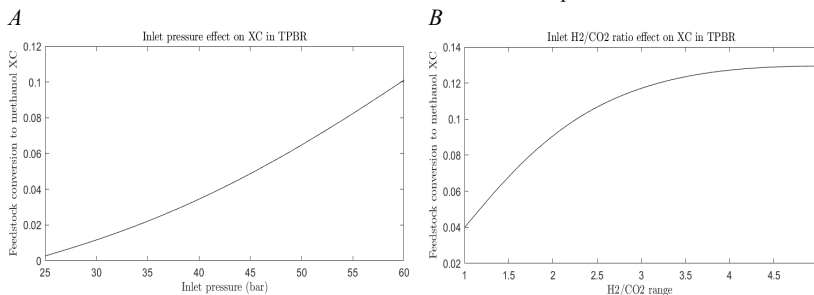
Figure 2 – Temperature, pressure and initial H<sub>2</sub>/CO<sub>2</sub> effect on feedstock conversion to methanol. a) T and P effect on  $X_c$ . b) P and inlet H<sub>2</sub>/CO<sub>2</sub> ratio effect on  $X_c$ .

#### 4.2 Sensitivity analysis and optimization of feedstock conversion in the TPBR

Only were evaluated temperature, pressure, and inlet H<sub>2</sub>/CO<sub>2</sub> effect, and other reactor-related variables was considered as parameters. When the sensitivity analysis was performed in the reactor model it was found a similar effect of inlet pressure and H<sub>2</sub>/CO<sub>2</sub> ratio on  $X_c$  as in equilibrium function (Figure 3a-b) but different in the temperature analysis (Figure 3c). An increment of the inlet temperature of feedstock current or in the liquid at the different jacket configurations increased  $X_c$  until limit value in which the trend was reversed and finally followed the same tendency as in equilibrium analysis. Note that the equilibrium is not reached in the TPBR at the end of the process and the temperature and pressure ranges differs in both equilibrium and reactor sensitivity analysis; this is why all values of  $X_c$  in the reactor model are lower than the direct analysis of the equilibrium expression (see Figure 3d). The optimization constrains of the TPBR are in Table 5.

	Evaluation range			Optimal Values
	Min	Max		
Temperature	Feedstock	480 K	530 K	508 K
	Jacket - Countercurrent	480 K	530 K	530 K
	Jacket - Co-current	480 K	530 K	530 K
Pressure	20 bar	60 bar		60 bar
H <sub>2</sub> /CO <sub>2</sub> ratio	0.5	5		4
				$X_c = 0.165$

Table 5 – Optimal temperature, pressure and H<sub>2</sub>/CO<sub>2</sub> ratio to maximize feedstock conversion to methanol at the outlet of the TPBR in the constrained numerical optimization of EDO in Table 2.



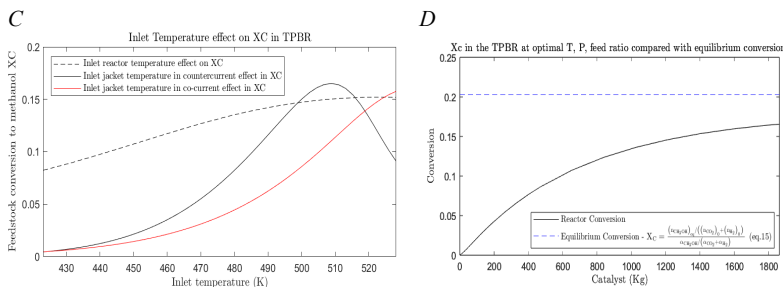


Figure 3 – a) Effect of inlet pressure on  $X_C$  in the TPBR. b) Effect of inlet  $H_2/CO_2$  on  $X_C$  the TPBR. c) Effect of inlet temperature (Inlet species temperature vs inlet Jacket temperature) on  $X_C$  in the TPBR. d) Direct comparison of equilibrium analysis and de conversions achievable in the proposed TPBR.

## 5. Comparison and brief results discussion

### 5.1 Temperature

Temperature affects mainly the equilibrium constants ( $K_{eq,j}$ ) and the fugacity coefficients of the components in the mixture ( $\hat{\phi}_i$ ). The evaluation of eq.1 in Figure 2a shows that lower temperatures favor methanol synthesis. As higher the temperature, higher the methanol oxidation to carbon dioxide and higher the production of undesirable products like CO and  $H_2O$  (WGS reaction). When the temperature was evaluated in the TPBR, the tendency was different, among others, because the activation energies in Graaf (1988) kinetic model. For instance, in Figure 4 the reaction rate of methanol is favored with temperature despite the different tendency in equilibrium sensitivity analysis. Then, it is necessary to optimize temperature to be high enough to reach a good react rate and be low enough to achieve good feedstock conversion to methanol. Finally, the optimal inlet temperature is different depending on the heat dissipation system properties throughout the catalytic bed weight. In the proposed model, the countercurrent configuration reached a higher feedstock conversion to methanol faster than co-current configuration.

### 5.2 Pressure and $H_2/CO_2$ ratio

The equilibrium composition of methanol increases with pressure (Figure 2, 3a). Pressure is directly related with concentration of moles in the gas phase and the species activity. When the  $H_2/CO_2$  ratio effect is analyzed (Figure 2b), the feedstock conversion to methanol reaches a maximum value when inlet feedstock ratio is 2.55. This value represents an optimal feedstock input and is a quantity close to the stoichiometric relation in the reaction.

## 6. Conclusions & further recommendations

As lower the values of temperature and higher the pressure in the gas-phase system, the feedstock conversion to methanol increases, but the temperature should be enough to promote the methanol reaction rate to achieve considerable feedstock conversion in the reactor geometry. When temperature raises, the system shifts towards carbon monoxide and water production. For a constrained optimization of equilibrium expression, the optimal values of temperature, pressure and  $H_2/CO_2$  ratio values were 480 K, 150 bar and 2.55 respectively. The temperature control in the TPBR model depends on the jacket fluid configuration and the characteristics of the heat dissipation system. For this specific set of conditions, it was found a jacket liquid temperature in countercurrent to maximize  $X_C$  of 508K, an inlet pressure of 60 bar and inlet  $H_2/CO_2$  ratio of 4. Also, is secure to affirm that

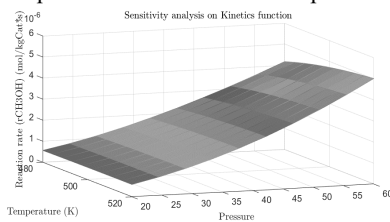


Figure 4 – Temperature and pressure effect on methanol reaction rate in Graaf (1988) kinetic model at the inlet conditions of the TPBR reactor.

an increase of the pressure is an excellent form to increase feedstock conversion to methanol due to its positive effect in both equilibrium expression and TPBR model analysis. The most substantial differences between the analysis of the equilibrium expression and the reactor model may lie in the fact that this reactor model is limited and cannot reach a closer conversion to equilibrium concentrations with the specified reactor design parameters. Another limitation is the small range of temperatures and pressures over which the reactor model is valid. Further analysis should robust the reactor model and compare all the temperatures and pressures ranges in the reactor with the equilibrium expression. But, with the basis relations founded in this simple reactor model is possible to understand the fundamentals of the reaction behavior and design more complex reactor models.

## 7. Notation table

$T_R$	Reactor Temperature (K)	$U$	Global Heat Transfer coefficient ( $J/m^2sK$ )
$T_j$	Jacket Temperature (K)	$A_s$	Heat Transfer Area ( $m^2/kgCat$ )
$P_R$	Reactor pressure (bar)	$\widehat{\Phi}_i$	Fugacity coefficient of the components in mixture (bar)
$F$	Total molar flow (mol/s)	$y_i$	Molar fraction in gas phase (-)
$F_i$	Species molar flow (mol/s)	$\epsilon_j$	Reaction extent
$H_2/CO_2$	Inlet Hydrogen vs Carbon dioxide relation (-)	$P, T$	Pressure (bar) – Temperature (K) for equilibrium function analysis.
$\rho_g$	Gas mixture density ( $kg/m^3$ )	$\alpha_{ij}$	Stoichiometric coefficients
$\mu_g$	Gas mixture viscosity ( $Pa \cdot s$ )	$\bar{\alpha}_j$	Sum of stoichiometric coefficients for $j$ reaction.
$G_g$	Gas mass Flux density ( $kg \cdot m^{-2} \cdot s^{-1}$ )	$K_\Phi$	Equilibrium constant based in fugacity coefficients
$C_{p_g}$	Molar gas mixture heat capacity ( $J/molK$ )	$K_y$	Equilibrium constant based on molar fractions
$C_{p_j}$	Molar Jacket liquid heat capacity ( $J/molK$ )	$K_p^0$	Equilibrium constant based in partial pressures
$e$	Catalyst void fraction (-)	$X_c$	Feedstock conversion to methanol (-)
$\rho_{cat}$	Catalyst density ( $kg_{cat}/m^3$ )	$\Delta H_{r,j}^0$	Enthalpy change of formation ( $J/mol$ )
$D_p$	Particle diameter (m)	$r'$	Reaction rate per weight of catalyst ( $mol/skg_{cat}$ )
$n_t$	Tubes number (-)	Subscripts	
$D_t$	Tubes diameter (m)	$i$	Refers to species $n = \{1,2,3,4,5\}$
$A_c$	Tube cross-sectional area ( $m^2$ )	$j$	Refers to reactions $k = \{1, 2, 3\}$
$g_c$	Conversion factor (SI=1)	$0$	Refers Initial state
$b$	Adsorption constant	$eq$	Refers to equilibrium

## References

- Graaf, G. H., Stamhuis, E. J., & Beenackersz, A. A. C. M. (1988). Kinetics Of Low-Pressure Methanol Synthesis. In *Chemical Engineering Science* (Vol. 43, Issue 12).
- Khojasteh-Salkuyeh, Y., Ashrafi, O., Mostafavi, E., & Navarri, P. (2021). CO2 utilization for methanol production; Part I: Process design and life cycle GHG assessment of different pathways. *Journal of CO2 Utilization*, 50. <https://doi.org/10.1016/j.jcou.2021.101608>
- Sonthalia, A., Kumar, N., Tomar, M., Edwin Geo, V., Thiyagarajan, S., & Pugazhendhi, A. (2021). Moving ahead from hydrogen to methanol economy: scope and challenges. *Clean Technologies and Environmental Policy*, 1, 1–25. <https://doi.org/10.1007/S10098-021-02193-X/TABLES/7>
- van Bennekom, J. G., Winkelman, J. G. M., Venderbosch, R. H., Nieland, S. D. G. B., & Heeres, H. J. (2012). Modeling and experimental studies on phase and chemical equilibria in high-pressure methanol synthesis. *Industrial and Engineering Chemistry Research*, 51(38), 12233–12243. <https://doi.org/10.1021/ie3017362>
- Verhelst, S., Turner, J. W., Sileghem, L., & Vancoillie, J. (2019). Methanol as a fuel for internal combustion engines. *Progress in Energy and Combustion Science*, 70, 43–88. <https://doi.org/10.1016/J.PECS.2018.10.001>

# Improving Transferable Force-Fields for Describing Crystal Structures Containing Hydrogen-Bonds

Benjamin I. Tan<sup>a</sup>, David H. Bowskill<sup>b</sup>, Adam Keates<sup>b</sup>, Constantinos C. Pantelides<sup>a</sup> and Claire S. Adjiman<sup>a</sup>

<sup>a</sup> *Department of Chemical Engineering, Sargent Centre for Process Systems Engineering and Institute for Molecular Science and Engineering, Imperial College London, London SW7 2AZ, United Kingdom*

<sup>b</sup> *Process Studies Group, Syngenta, Jealott's Hill International Research Centre, Bracknell, Berkshire, RG42 6EY, United Kingdom*

## Abstract

In Crystal Structure Prediction, transferable force-fields (FF) offer an efficient means to construct cheap surrogates of more accurate models. FF development, however, demands large volumes of high-quality reference data which are scarce in experimental databases. In this work, we use periodic DFT-D generated reference data to parameterize a FF apt for predicting properties of crystals in which molecules interact via hydrogen-bonds. Tang-Toennies damping is applied to the underlying Buckingham potential, and several fitting schemes are tested. Our FF achieves good agreement with the energy and geometry training data, for structures with and without hydrogen-bonds. When applied to a validation set, the quality of property predictions are largely preserved, evidence of FF transferability. Testing the popular FIT FF on our data fails to achieve the same accuracy, likely since it was derived from different reference data and lacks the damping function.

**Keywords:** Crystal Structure Prediction, Transferable Force-field, Parameter Estimation

## 1. Introduction

Many industrial process intermediates and commercial products exist in crystalline form. To predict the physicochemical properties of such materials requires knowledge of the three-dimensional arrangement of molecules within their crystal lattice. This, however, is complicated by the fact that most molecular crystals can exist in geometries other than their thermodynamically most stable form (Cruz-Cabeza *et al.*, 2015). Failure to identify these metastable polymorphs can have severe financial, legal, and social repercussions (Bučar *et al.*, 2015), especially since the interconversion to a more stable form is spontaneous yet highly unpredictable if the metastable form is kinetically trapped.

The central goal of Crystal Structure Prediction (CSP) is thus to characterize the structure and thermodynamic stability of all putative crystal forms from a description of only molecular connectivity (Pantelides *et al.*, 2014). Most CSP methodologies seek to accomplish this by minimizing the crystal static lattice-energy ( $U_{\text{latt}}$ ) relative to its geometry at 0 K and 0 Pa (Bowskill, 2021). By minimizing an ample number of crystalline geometries, an adequate survey of the crystal structure landscape may be achieved. Energetically low-lying minima, within 10 kJ/mol of the global minimum (Nyman & Day, 2015), are then regarded as potentially observable experimental forms.

The validity of any CSP study is thus dependent on both the breadth of the search and the accuracy of the energy model used. Hybrid energy models which use a combination of

*ab initio* calculations and empirical force-fields (FFs) offer a good balance between accuracy and computational cost. Whilst the *ab initio* components of these models have been studied extensively, the accompanying empirical FFs have not. Transferable FFs like FIT (Coombes *et al.*, 1996) and W99rev (Pyzer-Knapp *et al.*, 2016) are widely used despite deficiencies in the estimation of FF parameters (Gatsiou *et al.*, 2018). For instance, reliance on experimental heat of sublimation ( $\Delta H_{\text{sub}}$ ) data to provide “experimental” values of  $U_{\text{latt}}$  limits FF validity for two reasons. First, the relationship used to derive  $U_{\text{latt}}$  from  $\Delta H_{\text{sub}}$  is approximate. Second, experimental  $\Delta H_{\text{sub}}$  values are scarce, resulting in FF parameters biased towards geometry prediction. Finally, both FFs mentioned were derived using multipole (MP) descriptions of electrostatics. This may invalidate their use in the initial stages of CSP, with point-charge (PC) electrostatics.

In the larger context of CSP, dispersion-corrected periodic density-functional theory (DFT-D) is regarded as the gold-standard of accuracy (Brandenburg & Grimme, 2016) and is usually applied at end of CSP studies to refine results obtained with cheaper models. In the spirit of a blind CSP study, it is thus more sensible to design hybrid models that act as surrogates for the DFT-D landscape since the true experimental landscape remains unknown. For these reasons, in this work, we seek to develop a robust FF compatible with PC electrostatics, fit to high-quality DFT-D generated reference data. Details of the energy model and fitting procedure are described in Section 2. In Section 3, the performance of the newly parameterized FF is then explored.

## 2. Methodology

### 2.1. Lattice-Energy Model

Within our energy model,  $U_{\text{latt}}$  comprises intra- and intermolecular contributions,

$$U_{\text{latt}}(\mathbf{X}, \mathbf{p}) = \Delta U_{\text{intra}}(\mathbf{X}) + U_{\text{inter}}(\mathbf{X}, \mathbf{p}) = \Delta U_{\text{intra}}(\mathbf{X}) + U_{\text{elec}}(\mathbf{X}) + U_{\text{FF}}(\mathbf{X}, \mathbf{p}), \quad (1)$$

where  $\mathbf{X}$  is the vector describing the crystal geometry (cell lengths and angles, and atomic positions).  $U_{\text{latt}}$  contains the energy cost of distorting a molecule’s conformation relative to its gas-phase conformation ( $\Delta U_{\text{intra}}$ ), and stabilizing packing forces ( $U_{\text{inter}}$ ).  $U_{\text{inter}}$  comprises intermolecular charge-charge interactions ( $U_{\text{elec}}$ ), electron-correlation, and exchange-repulsion. The latter two effects are described by the pairwise summation of interactions between atoms  $i$  and  $j$  separated by distance  $r_{ij}$ , interacting via a Buckingham potential (BHP) with model parameters  $\mathbf{p} = \{A_{ij}, B_{ij}, C_{ij}\}$ ,

$$U_{\text{FF}}(\mathbf{X}, \mathbf{p}) = \sum_i \sum_j A_{ij} \cdot \exp\left(-\frac{r_{ij}}{B_{ij}}\right) - f_6(\beta_{ij}, r_{ij}) \cdot \frac{C_{ij}}{r_{ij}^6}, \quad (2)$$

$$f_6(\beta_{ij}, r_{ij}) = 1 - \left[ \sum_{k=0}^6 \frac{(\beta_{ij} \cdot r_{ij})^k}{k!} \right] \cdot \exp(-\beta_{ij} \cdot r_{ij}) \quad \text{where } \beta_{ij} = \frac{1}{B_{ij}}. \quad (3)$$

In its native form ( $f_6 = 1$ ) the BHP potential exhibits unphysical short-range attractions. When  $A_{ij}/B_{ij}$  is small and/or  $C_{ij}$  is large, this unphysical region extends to longer ranges and eventually the potential can become fully attractive. To mitigate this occurring in highly attractive hydrogen-bond (HB) interactions, the Tang-Toennies (Tang & Toennies, 1984) damping function ( $f_6$ ) is applied within our potential. Its damping factor ( $\beta_{ij}$ ) is correlated to the  $B_{ij}$  repulsion term to avoid overfitting. Finally,  $U_{\text{elec}}$  and  $\Delta U_{\text{intra}}$  are determined from *ab initio* calculations performed on the isolated molecules of interest, while the  $\mathbf{p}$  in  $U_{\text{FF}}$  are found empirically by fitting to crystal data. As a result,  $U_{\text{FF}}$  absorbs systematic errors from the *ab initio* components of the model (Pantelides *et al.*, 2014). For instance, the inductive effects molecules exert on each other in the crystal lattice are neglected by isolated molecule calculations and must be captured by  $U_{\text{FF}}$ .

## 2.2. Reference Data

DFT-D reference data are generated in the Vienna Ab initio Simulation Package (VASP v5.4.4, Kresse & Joubert, 1999) using the Tao-Perdew-Staroverov-Scuseria (Tao *et al.*, 2003) functional with D3 dispersion-correction (Grimme *et al.*, 2010). Experimental structures from the Cambridge Structural Database are first relaxed to obtain reference geometries. A single-point calculation is done on isolated molecule  $m$  at its crystalline conformation to obtain its monomer energy ( $U_{\text{iso},m}$ ). Using the number of molecules in the unit cell ( $Z$ ) and asymmetric unit cell ( $Z'$ ), the reference energy ( $U_{\text{inter,ref}}$ ) is found from

$$U_{\text{inter,ref}} = \frac{U_{\text{total,vasp}}}{Z} - \frac{\sum_{m=1}^{Z'} U_{\text{iso},m}}{Z'}. \quad (4)$$

The training set comprises 313 non-HB (NHB) crystals and 132 HB crystals, while an additional 28 structures (12 NHB and 16 HB) are used in the validation set. More details about the computational methods and training structures are described in Bowskill (2021).

## 2.3. Parameter Estimation Algorithm

The parameter estimation is conducted using CrystalEstimator (v1.4) as described in Bowskill (2021). During the parameter estimation, the following problem is solved,

$$\min_{\mathbf{p}} \frac{1}{2} \sum_{i=1}^{N_c} \left\{ \sum_{j=1}^{N_{X,i}} \frac{\omega_g [X_{ij}^{\text{ref}} - X_{ij}^*(\mathbf{p})]^2}{N_{X,i} \cdot [X_{ij}^{\text{ref}}]^2} + \sum_{j=1}^{N_{R,i}} \frac{\omega_g [R_{ij}^{\text{ref}} - R_{ij}^*(\mathbf{p})]^2}{3(N_{R,i} - 1)} + \frac{\omega_e [U_{\text{inter},i}^{\text{ref}} - U_{\text{inter},i}^*(\mathbf{p})]^2}{[U_{\text{inter},i}^{\text{ref}}]^2} \right\}, \quad (5)$$

where  $X_{ij}$  is the  $j^{\text{th}}$  lattice cell parameter (lengths and angles) for crystal  $i$ ,  $R_{ij}$  its  $j^{\text{th}}$  atomic fractional coordinate, and  $U_{\text{inter},i}$  its intermolecular energy respectively. The number of crystals is  $N_c$ , and the dimensions of  $X_i$  and  $R_i$  are crystal-specific ( $N_{X,i}$  and  $N_{R,i}$ ).  $\omega_g$  and  $\omega_e$  are the respective weights for the geometry and energy residuals. The superscripts ‘ref’ and ‘\*’ respectively indicate the reference values and the values calculated by minimizing  $U_{\text{latt}}$  at a given  $\mathbf{p}$ . The parameter estimation is thus a bilevel optimization since at each trial  $\mathbf{p}$ , all reference crystals must have their  $U_{\text{latt}}$  minimized first. In this minimization, molecules are kept rigid at their reference conformation ( $\Delta U_{\text{intra}} = 0$ ,  $U_{\text{latt}} = U_{\text{inter}}$ ) and  $U_{\text{elec}}$  is calculated from PCs derived in Gaussian09 using the HLYGAt scheme with the PBE0 functional and 6-311G(d,p) basis set. Due to the multiplicity of local minima on the  $\mathbf{p}$ -solution landscape, multi-start initialization using low-discrepancy Sobol’ sequences is used to generate the initial  $\mathbf{p}$  vectors. The Sobol’ point which yields the best objective function value (OF) is taken to be the final solution. At the solution, we evaluate the RMSD<sub>15</sub> value between optimized and reference structures, determined with the Cambridge Crystallographic Data Centre’s COMPACK tool (distance and angle tolerances are set to 0.4 and 40°), and deviations in  $U_{\text{inter}}$  ( $dU = U_{\text{inter}}^* - U_{\text{inter}}^{\text{ref}}$ ).

## 2.4. Lattice-Energy Model

To select a suitable number of parameters which capture the underlying physics without overfitting, several fitting schemes are tested. We opt to parameterize all heteroatomic interactions ( $i \neq j$ ) rather than using combining rules. Analysis of our training set and experimental crystal structures (Gavezzotti & Filippini, 1994) indicates that Hs bonded to O ( $H_o$ ) and N ( $H_n$ ) exhibit distinct HB geometries, hence we distinguish the two. It was previously asserted that polarized Hs would be stripped of electron density and exhibit no dispersive attractions ( $C_{ij} = 0$  for  $i = H_{o/n}$ ) (Pyzer-Knapp *et al.*, 2016). In practice, other parts of the model may not fully capture the HB effects and require the FF to compensate, hence, we test setting  $C_{H_o \text{ or } H_n/j} = 0$  and varying  $C_{H_o \text{ or } H_n/j}$ . Finally,  $A_{ij}$  and  $B_{ij}$  are highly correlated, thus it is common to fix  $B_{ij}$  and fit only  $A_{ij}$  and  $C_{ij}$ .  $B_{ij}$  were initially fixed to the FIT FF values, but subsequently some  $B_{H_o \text{ or } H_n/j}$  were also allowed to vary.



Table 1: Performance of each FF in describing the training set. Only FF1, FF2, and FF3 are damped. The average RMSD<sub>15</sub>, average  $dU$ , and average absolute  $dU$  ( $|dU|$ ) are presented for the entire training set (NHB+HB), NHB, and HB crystals. The number of structures which failed to give a valid RMSD<sub>15</sub> are given in parenthesis. Parameters are obtained either by combining rules (CR), varied in the optimization (VR), or fixed (FX).

FF	$i \neq j$	$B_{ij}$	$C_{ij}$	OF	Data	RMSD <sub>15</sub> (Å)	$dU$ (kJ/mol)	$ dU $ (kJ/mol)
FIT	CR	FX	VR	-	NHB+HB	0.2701 (-)	12.2450	14.1433
					NHB	0.2566 (-)	7.2671	9.6378
					HB	0.3021 (-)	24.0486	24.8268
FF1	VR	FX	0	0.692	NHB+HB	0.2406 (32)	0.8560	4.7473
					NHB	0.2168 (-)	-0.4978	3.6014
					HB	0.3154 (32)	4.0662	7.4643
FF2	VR	FX	VR	0.538	NHB+HB	0.2214 (1)	0.3874	4.1879
					NHB	0.2136 (-)	0.1417	3.4808
					HB	0.2401 (1)	0.9701	5.8645
FF3	VR	VR	VR	0.524	NHB+HB	0.2203 (-)	0.2888	4.0604
					NHB	0.2129 (-)	0.1730	3.4290
					HB	0.2377 (-)	0.5635	5.5573

### 3. Results

Three FFs are obtained using the fitting schemes described in Section 2. Their performance in describing the training set is compared against the FIT FF in Table 1. In some structures, failure to match 15 molecules results in no RMSD<sub>15</sub> being generated.

Over the training set, all three newly parameterized FFs present with similarly small errors in average  $dU$  ( $< 5$  kJ/mol) and  $|dU|$  ( $< 10$  kJ/mol). In a previous work which used the same training set (Bowskill, 2021), the average  $dU$  and  $|dU|$  of the HB structures were 22.31 kJ/mol and 22.48 kJ/mol respectively, indicating that the new FFs exhibit a marked reduction in the energy errors for these structures. In the previous work, energy data for the HB structures was omitted due to numerical issues associated with the optimization. The inclusion of the damping function in this work appears to have mitigated these problems. However, to match the HB energies well, the geometry predictions worsen significantly, particularly for the HB structures. Using FF1 and FF2, not only do the average RMSD<sub>15</sub> values for HB structures increase over the previous work (0.238 Å), but more critically, some structures fail to produce any RMSD<sub>15</sub> values at all.

This tradeoff between energy and geometry may be explained as follows. In FF1, fixing  $C_{ij} = 0$  for the H<sub>o/n</sub> interactions prevents these interactions from contributing to the deficit in  $U_{\text{inter}}$ . The only recourse is for the optimizer to set  $A_{ij}$  in these interactions to 0 (or close to 0), allowing greater short-range charge-charge interactions and other atoms to interact deeper within their energy-wells, at the expense of the fidelity to geometry. Allowing non-zero  $C_{ij}$  within FF2 does somewhat alleviate this issue, but there remain some interactions (e.g., O/H<sub>o</sub>) where  $A_{ij} = 0$ , resulting in fully attractive potentials. The absence of repulsive walls in FF1 and FF2 (Figure 1, left) permit unnaturally close contact distances that prevent the assessment of RMSD<sub>15</sub> and cast doubt on the transferability of these FFs. In contrast, by allowing some of the  $B_{ij}$  involving H<sub>o/n</sub> to vary (FF3), we attain potentials which *all* have repulsive walls. As the only ‘physically sensible’ FF, the transferability of FF3 is tested on an independent validation set. Overall, the performance of FF3 does not deteriorate radically from the training set (Table 2), with a similar quality of energy prediction. While the average RMSD<sub>15</sub> for HB structures increases, it is still within a cutoff that most CSP practitioners would consider acceptable ( $< 0.5$  Å).

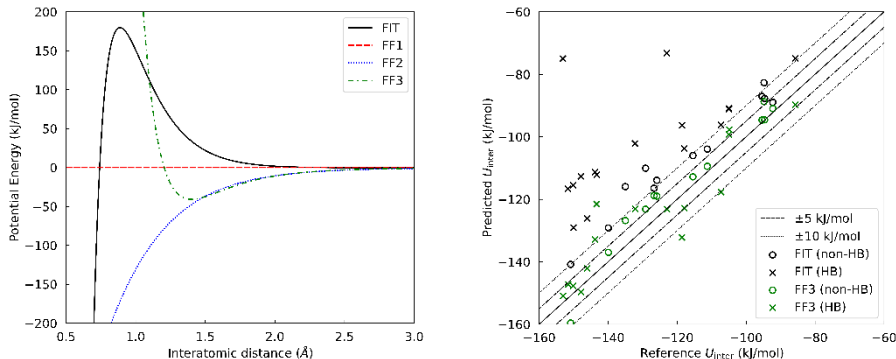


Figure 1: (Left) Comparison of the H/O potential for the four FFs. (Right) Parity plot of the validation set  $U_{inter}$ , comparing the performance using the FIT FF and FF3.

As mentioned before, the goal of modern CSP FFs is to act as proxies for the DFT-D landscape. To further benchmark our FFs, we compare against the FIT FF. Within both the training and validation set, whilst the FIT FF offers good geometry predictions, energies are substantially underbound, especially for HB structures. We note FIT is being used well beyond the data to which it was parameterized. Correcting from their  $\Delta H_{sub}$  to our DFT-D  $U_{inter,ref}$  is approximately given by  $-\Delta H_{sub} = U_{inter} + \Delta U_{intra} + 2RT$ . In deriving FIT, it was assumed that  $\Delta U_{intra} \approx 0$ , while our VASP calculations predict  $\Delta U_{intra}$  as large as 20 kJ/mol in HB structures even with small conformational changes. Moreover, our  $U_{inter,ref}$  are effectively at 0 K, while the room-temperature data used in FIT would result in a positive  $2RT$ , but this was ignored in the original fitting. Overall, the FIT reference energies underpredict the DFT-D  $U_{inter,ref}$ , explaining the large, positive average  $dU$  observed when applying FIT here. This is also apparent when one considers the shape of the potentials (Figure 1, left): FF3 has deeper energy-wells and shorter-range repulsive walls, while FIT has a repulsive wall at distances exceeding the typical HB lengths (O-H $\cdots$ O distances  $\approx 1.6$ -1.9 Å, Gavezzotti & Filippini, 1994), and shallower energy-wells.

#### 4. Conclusion

Electronic-structure methods like DFT-D offer our ‘best guess’ of the polymorphic landscape and are used in the final stages of a CSP study. With this in mind, we construct hybrid energy models to emulate the DFT-D landscape as opposed to the (unknown) experimental landscape. In parameterizing a FF using DFT-D generated training data, commonly used model-simplifications were found to yield unphysical results. By altering the functional form of the FF and avoiding setting model parameters to arbitrary values,

Table 2: Performance of FF3 and FIT in describing the validation set. The average  $RMSD_{15}$ , average  $dU$ , and average absolute  $dU$  ( $|dU|$ ) are presented for the full dataset, NHB, and HB crystals. All structures gave valid  $RMSD_{15}$  with both FFs. The change in each metric from the training set is given in parenthesis for FF3.

FF	Data	$RMSD_{15}$ (Å)	$dU$ (kJ/mol)	$ dU $ (kJ/mol)
FIT	NHB+HB	0.2811	20.8127	20.8127
	NHB	0.2097	10.7082	10.7082
	HB	0.3346	28.3911	28.3911
FF3	NHB+HB	0.2380 (+0.0177)	1.6665 (+1.3777)	6.3550 (+2.2946)
	NHB	0.1549 (-0.0580)	2.9774 (+2.8044)	4.4317 (+1.0027)
	HB	0.3003 (+0.0626)	0.6833 (+0.1198)	7.7975 (+2.2402)

we were able to obtain a FF that makes physical sense and is effective in describing crystal properties. In both the training and validation sets, our new FF has an average  $|dU| < 10$  kJ/mol, well within the threshold for identifying putative structures. Conversely, a FF like FIT that was parameterized to experimental data is unsuited to predict the DFT-D landscape. Ultimately, using computationally generated reference data will allow us to circumvent the limitations of scarce experimental data and explore new avenues in CSP FF development, such as the expansion of atom-typing.

## Acknowledgements

We are grateful to the UK Materials and Molecular Modelling Hub for computational resources, which is partially funded by EPSRC (EP/P020194/1 and EP/T022213/1). The authors also gratefully acknowledge funding provided from Syngenta, and computational resources provided by Imperial College London's High-Performance Computing Cluster.

## References

- D. H. Bowskill, 2021, Reliable and Efficient Parameter Estimation Methodologies for Crystal Structure Prediction, PhD Thesis, Imperial College London
- J. G. Brandenburg, S. Grimme, 2016, Organic crystal polymorphism: a benchmark for dispersion-corrected mean-field electronic structure methods, *Acta Crystallographica Section B, Structural Science, Crystal Engineering and Materials*, 72, 4, 502-513
- D. K. Bučar, R.W. Lancaster, J. Bernstein, 2015, Disappearing Polymorphs Revisited, *Angewandte Chemie International Edition*, 54, 24, 6972-6993
- D. S. Coombes, S. L. Price, D. J. Willock, M. Leslie, 1996, Role of Electrostatic Interactions in Determining the Crystal Structures of Polar Organic Molecules. A Distributed Multipole Study, *The Journal of Physical Chemistry*, 100, 18, 7352-7360
- A. J. Cruz-Cabeza, S.M. Reutzel-Edens, J. Bernstein, 2015, Facts and fictions about polymorphism, *Chemical Society Reviews*, 44, 23, 8619-8635
- C. A. Gatsiou, C. S. Adjiman, C. C. Pantelides, 2018, Repulsion–dispersion parameters for the modelling of organic molecular crystals containing N, O, S and Cl, *Faraday Discussions*, 211, 297-323
- A. Gavezzotti, G. Filippini, 1994, Geometry of the Intermolecular X-H...Y (X, Y = N, O) Hydrogen Bond and the Calibration of Empirical Hydrogen-Bond Potentials, *The Journal of Physical Chemistry*, 98, 18, 4831-4837
- S. Grimme, J. Antony, S. Ehrlich, H. Krieg, 2010, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, *The Journal of Chemical Physics*, 132, 15, 154104
- G. Kresse, D. Joubert, 1999, From ultrasoft pseudopotentials to the projector augmented-wave method, *Physics Review B*, 59, 3, 1758-1775
- J. Nyman, G. M. Day, 2015, Static and lattice vibrational energy differences between polymorphs, *Crystal Engineering Communications*, 17, 28, 5154-5165
- C. C. Pantelides, C. S. Adjiman, A.V. Kazantsev, 2014, General computational algorithms for ab initio crystal structure prediction for organic molecules, *Prediction and Calculation of Crystal Structures: Methods and Applications*, Springer, 25-58
- E. O. Pyzer-Knapp, H. P. G. Thompson, G. M. Day, 2016, An optimized intermolecular force field for hydrogen-bonded organic molecular crystals using atomic multipole electrostatics, *Acta Crystallographica Section B, Structural Science, Crystal Engineering and Materials*, 72, 4, 477-487
- K. T. Tang, P. J. Toennies, 1984, An improved simple model for the van der Waals potential based on universal damping functions for the dispersion coefficients, *The Journal of Chemical Physics*, 80, 8, 3726-3741
- J. Tao, J. P. Perdew, V. N. Staroverov, G. E. Scuseria, 2003, Climbing the Density Functional Ladder: Nonempirical Meta--Generalized Gradient Approximation Designed for Molecules and Solids, *Physics Review Letter*, 91, 14, 146401

# A data-driven uncertainty modelling and reduction approach for energy optimisation problems

Julien Vaes, Vassilis M. Charitopoulos

*Department of Chemical Engineering, The Sargent Centre for Process Systems Engineering, University College London, Torrington Place, London WC1E 7JE, UK.*

## Abstract

Taking uncertainty into account is crucial when making strategic decisions. To guard against the risk of adverse scenarios, traditional optimisation techniques incorporate uncertainty based on prior knowledge on its distribution. In this paper, we show how, based on limited historical data, we can generate from a low-dimensional space the underlying structure of uncertainty. To this end, we first exploit the correlation between the sources of uncertainty through a principal component analysis (PCA) to reduce dimensionality. Next, we perform clustering to reveal the typical uncertainty patterns, and finally we generate polyhedral uncertainty sets based on a kernel density estimation (KDE) of marginal probability functions.

**Keywords:** Polyhedral Uncertainty Set, Robust Optimisation (RO), PCA, Kernel density estimation (KDE), Dimensionality Reduction, Data scarcity.

## 1. Introduction

Optimisation under uncertainty aims to find optimal solutions given uncertain outcomes. This is typically done through methods like chance-constrained optimisation, robust optimisation, or stochastic optimisation, which require prior knowledge of an uncertainty set or probability distribution. Robust optimisation (RO) is for instance a suitable method in situations where insufficient historical data is available to accurately estimate distributions (Roald et al., 2023). In this method, the solution hedges against a range of possible realisations called uncertainty set with emphasis on the worst-case to guarantee feasibility of the policy derived. The definition of such uncertainty set is crucial: the larger it is, the less informative it becomes and the more conservative is the solution, which could render it impractical (Ben-Tal et al., 2009).

We mention here some methods and their limitation that have been recently proposed in the literature to generate uncertainty sets. First, Bianchi et al. (2018) employ statistical hypothesis tests to derive uncertainty sets' sizes. Nonetheless, these tests become impractical in cases of data scarcity. In Shang et al. (2017); Ning and You (2018), a single polyhedral uncertainty set (PUS) is obtained using respectively kernel learning (support vector clustering) or a principal component analysis (PCA) coupled with kernel smoothing. However, this assumes continuous data and cannot handle uncertainty formed as a collection of independent states. To circumvent this limitation, Ning and You (2017, 2019) define the uncertainty set as the union of PUS, which are generated using a Dirichlet process mixture model. However, the latent uncertainty used to generate the PUSs is of same dimension as the original uncertainty, leading to computational intractability for high dimensional uncertainty sets when used in RO (Ben-Tal et al. 2009).

There is consequently the need for a method that generates polyhedral uncertainty sets in the case of data sparsity where uncertainty is highly dimensional and consists of independent sets of realisations; this is the objective of the present paper.

The rest of this paper is structured as follows: In Section 2, we present the method which derives polyhedral uncertainty sets from a low-dimensional space based on historical data. We then show in Section 3 how it can be applied to the energy data in the United Kingdom. In Section 4 a discussion of the key findings is presented and finally in Section 5 conclusions on the main contributions of this work are drawn.

## 2. Method for deriving polyhedral uncertainty sets

To motivate the importance of building polyhedral uncertainty sets, we consider the following scenario-based linear adaptive robust optimisation (ARO) problem, which is recurrently used to formulate energy problems (e.g. unit commitment):

$$\underset{\mathbf{x}}{\text{minimise}} \mathbf{c}^T \mathbf{x} + \sum_{k \in \mathcal{K}} p_k \left( \max_{\mathbf{v}_k \in \mathcal{V}_k} \min_{\mathbf{y}_k} \mathbf{b}_k^T \mathbf{y}_k \right) \quad (1a)$$

$$\text{such that } \mathbf{A}\mathbf{x} \leq \mathbf{d} \text{ and } \forall k \in \mathcal{K}: \mathbf{h}_k - \mathbf{T}_k \mathbf{x} - \mathbf{M}_k \mathbf{v}_k \leq \mathbf{W}_k \mathbf{y}_k \quad (1b)$$

where  $\mathbf{x}$  and  $\mathbf{y}_k$  with  $k \in \mathcal{K}$  are respectively the first and second stage variables,  $\mathcal{K}$  is the set of distinct operation conditions/scenarios,  $\mathcal{V}_k$  is the uncertainty set associated to the uncertain parameters for scenario  $k \in \mathcal{K}$ , and  $p_k$  is its weight (or probability of occurrence). Such problem can be solved either in a monolithic fashion or by using variants of Benders' decomposition. If each set  $\mathcal{V}_k$  is a polyhedral uncertainty set, then each set can be expressed as linear constraints, which facilitates the resolution of the ARO problem by exploiting the dual formulation associated to a linear program (Birge and Louveaux, 2011). We now show how to derive these PUSs  $\mathcal{V}_k$  based on limited historical data and how to generate them from lower dimensional spaces. The method is here illustrated for energy uncertainty data but is not restricted to such application.

Let  $n \in \mathbb{N}$  be the number of uncertain attributes (e.g. electricity demand, temperature). Let  $\mathbf{v}^{(i)} \in \mathbb{R}^{24n}$ ,  $i \in \llbracket 1, m \rrbracket$ , be  $m$  historical data points, where the vector  $\mathbf{v}^{(i)}$  contains the standardised daily profiles (hourly values) of all attributes on a given day  $i \in \llbracket 1, m \rrbracket$ . We assume that the data is standardised in the sense that the sample mean and standard deviation of each component over all  $m$  data points is respectively 0 and 1. This allows for attributing an even weight to any uncertain component. Given a data point  $\mathbf{v}^{(i)} \in \mathbb{R}^{24n}$ , let  $\mathbf{V}^{(i)} \in \mathbb{R}^{n \times 24}$  be the matrix with the daily profiles of each attribute on its rows.

As an initial step, based a correlation analysis, we define a set  $\mathcal{A}$  of disjoint groups of attributes such that the data inside a group is strongly correlated while being weakly correlated with any other group. We then assume that these groups of attributes are independent and separately generate a polyhedral uncertainty set for each one of them. The (weak) correlation between the groups will partly be reinstated on a later stage when defining the scenarios. For any group of attributes  $a \in \mathcal{A}$ , we denote by  $\mathbf{v}_a^{(i)} \in \mathbb{R}^{24n_a}$  the data related to the  $n_a$  attributes associated to  $a$  for a day  $i \in \llbracket 1, m \rrbracket$ .

The following steps allow for deriving a PUS based on a lower dimensional space for each group  $a \in \mathcal{A}$ : First, we perform a principal component analysis (PCA) and express the data in the PCA basis: this allows for finding the directions along which the data has greatest variance. Mathematically, there exists an orthogonal matrix  $\mathbf{P}_a \in \mathbb{R}^{24n_a \times 24n_a}$

such that, any data point  $\mathbf{v}_a^{(i)}$ ,  $i \in \llbracket 1, m \rrbracket$ , expressed in the original basis can be expressed in the PCA basis using the linear transformation:  $\mathbf{w}_a^{(i)} = \mathbf{P}_a \mathbf{v}_a^{(i)}$ . As  $\mathbf{P}_a$  is orthogonal, the linear transformation from the PCA basis to the original basis is given by  $\mathbf{v}_a^{(i)} = \mathbf{P}_a^T \mathbf{w}_a^{(i)}$ . Given the data in the PCA basis, we perform dimensionality reduction by retaining only the first  $r_a$  components that explain the most the variability inside the data. As the data is strongly correlated inside any group  $a \in \mathcal{A}$ , we expect to have  $r_a \ll 24n_a$ . Let  $\bar{\mathbf{w}}_a^{(i)} \in \mathbb{R}^{r_a}$  denote the data points in the truncated PCA basis, where only the first  $r_a$  components are kept.

To define the typical operational conditions associated to a group  $a \in \mathcal{A}$  of attributes, we perform clustering (e.g. K-means) on the data points expressed in the truncated PCA basis. Clustering in the PCA basis rather than in the original basis allows to give relatively more importance to the directions explaining the most the variance and obtain more meaningful clusters (Ding and He, 2004). We denote by  $\mathcal{K}_a$  the set of clusters related to  $a \in \mathcal{A}$ , call each cluster  $k_a \in \mathcal{K}_a$  a *group-scenario* and define its probability of occurrence  $p_{k_a}$  as the proportion of data points attributed to this cluster.

We now follow a similar approach as Ning and You (2018) to construct a polyhedral uncertainty set for each group-scenario  $k_a \in \mathcal{K}_a$ . To this end, we estimate the marginal probability density function  $\hat{f}_{a,k_a,r}$ ,  $r \in \llbracket 1, r_a \rrbracket$ , along each truncated principal component direction (with a kernel smoothing method such as KDE for instance (Chen, 2017)). We denote by  $\hat{F}_{a,k_a,r}$  the associated cdf, and by  $\hat{F}_{a,k_a,r}^{-1}$  the associated quantile function. We now define the PUS related to the group of attributes  $a \in \mathcal{A}$  and the cluster  $k_a \in \mathcal{K}_a$ , which we call a *group-scenario PUS*, as follows:

$$\bar{\mathbf{w}}_{a,k_a}^{pol} := \left\{ \bar{\mathbf{w}} \in \mathbb{R}^{r_a} \left\{ \begin{array}{l} \mathbf{0} \leq \mathbf{z}^-, \mathbf{z}^+ \leq \mathbf{1} \\ \mathbf{1}^T (\mathbf{z}^- + \mathbf{z}^+) \leq \Phi_{a,k_a} \\ \forall i, j \in \llbracket 1, s_{a,k_a} \rrbracket: z_i^- + z_i^+ + z_j^- + z_j^+ \leq \Psi_{a,k_a} \\ \boldsymbol{\xi}_{a,k_a}^{lb} = [\hat{F}_{a,1,k_a}^{-1}(\alpha_{a,k_a}), \dots, \hat{F}_{a,r_a,k_a}^{-1}(\alpha_{a,k_a})]^T \\ \boldsymbol{\xi}_{a,k_a}^{ub} = [\hat{F}_{a,1,k_a}^{-1}(1 - \alpha_{a,k_a}), \dots, \hat{F}_{a,r_a,k_a}^{-1}(1 - \alpha_{a,k_a})]^T \\ \boldsymbol{\lambda} = \frac{1}{2} (\mathbf{z}^+ - \mathbf{z}^- + \mathbf{1}) \\ \bar{\mathbf{w}} = \boldsymbol{\xi}_{a,k_a}^{lb} \circ (\mathbf{1} - \boldsymbol{\lambda}) + \boldsymbol{\xi}_{a,k_a}^{ub} \circ \boldsymbol{\lambda} \end{array} \right. \right\} \quad (2)$$

where  $\mathbf{0}$  and  $\mathbf{1}$  are vectors of respectively zeros and ones of size  $r_a$ , where  $\circ$  defines the Hadamard product, and where vectors inequalities must be understood componentwise. This set is similar to the budgeted uncertainty set proposed by (Bertsimas and Sim, 2004), which enables to control the degree of conservatism: the larger the volume of the PUS, the more conservative is the resulting robust Optimisation Problem (1). The uncertainty set is parametrised by  $\alpha_{a,k_a}$ ,  $\Phi_{a,k_a}$  and  $\Psi_{a,k_a}$ . First,  $\alpha_{a,k_a}$  is used to exclude both tails of the marginal pdf along each principal component axis. Then,  $\Phi_{a,k_a}$  limits the cumulative dispersion from the nominal value along all retained PCA axes. Finally, the parameter  $\Psi_{a,k_a}$  additionally limits the pairwise dispersion along the first  $s_{a,k_a}$  PCA components. This parameter is important to further exclude unlikely data points from the uncertainty that would otherwise not be cut with the general budget constraint parametrised by  $\Phi_{a,k_a}$ .

We have so far generated a PUS for each group-scenario  $k_a \in \mathcal{K}_a$  for each  $a \in \mathcal{A}$ . Let  $\mathcal{K}^\times := \times_{a \in \mathcal{A}} \mathcal{K}_a$  denote the Cartesian product of the sets of group-scenarios. The probability  $\hat{p}_k$  associated to  $k \in \mathcal{K}^\times$  is estimated as the proportion of the data points such that the uncertainty part related to  $a$  is attributed to cluster  $k_a$  for all  $a \in \mathcal{A}$ . To take into

account the weak correlation between groups of attributes (so far assumed to be independent), we retain only the most probable combinations of group-scenarios  $k := \{k_a \in \mathcal{K}_a, a \in \mathcal{A}\} \in \mathcal{K}^\times$ , keeping those with associated probability greater than a threshold probability  $\tilde{p}$ . The more conservative we desire to be, the greater the acceptance probability threshold  $\tilde{p}$  should be. The set of scenarios  $\mathcal{K}$  is then constituted of the combinations in  $\mathcal{K}^\times$  satisfying the probability threshold  $\tilde{p}$ . The probability of occurrence  $p_k$  of each scenario  $k \in \mathcal{K}$  used in (1a) is then defined as the scaled probability when the scenarios that do not satisfy the acceptance threshold are excluded, *i.e.*  $p_k = \hat{p}_k / \sum_{k \in \mathcal{K}} \hat{p}_k$ . Finally, the PUS in the original basis related to a scenario  $k = \{k_a \in \mathcal{K}_a, a \in \mathcal{A}\} \in \mathcal{K}$  is then defined as follows:

$$\mathcal{V}_k := \left\{ \mathbf{v} \in \mathbb{R}^{24n} \mid \forall a \in \mathcal{A}: \begin{cases} \bar{\mathbf{w}}_a \in \bar{\mathcal{W}}_{a,k_a}^{pol} \\ \mathbf{v}_a = \mathbf{P}_a^T \begin{bmatrix} \bar{\mathbf{w}}_a \\ \mathbf{0}_{n_a-r_a} \end{bmatrix} \end{cases} \right\} \quad (3)$$

where  $\mathbf{0}_{n_a-r_a}$  is a vector of zeros of size  $n_a - r_a$ . The reduction in dimension appears clearly: the uncertain part  $\mathbf{v}_a \in \mathbb{R}^{24n_a}$  related to  $a \in \mathcal{A}$  is generated from elements of  $\bar{\mathcal{W}}_{a,k_a}^{pol}$ , which is in the lower-dimensional space  $\mathbb{R}^{r_a}$ .

In some applications (e.g. unit commitment), we are interested in taking into account the uncertainty over several days. Given the succession of  $N$  daily scenarios  $k_{1 \rightarrow N} := (k_1, \dots, k_N) \in \mathcal{K} \times \dots \times \mathcal{K}$ , we can define an associated PUS based on the definition of  $\mathcal{V}_k$  in (3) as follows:

$$\mathcal{V}_{k_{1 \rightarrow N}} := \left\{ \mathbf{v} \in \mathbb{R}^{24nN} \mid \begin{cases} \mathbf{v} = [\mathbf{v}_1, \dots, \mathbf{v}_N] \\ \forall i \in \llbracket 1, N \rrbracket: \mathbf{v}_i \in \mathcal{V}_{k_i} \\ \forall i \in \llbracket 1, N-1 \rrbracket: |(\mathbf{V}_{i+1}[:,1] - \mathbf{V}_i[:,24]) - \boldsymbol{\mu}_\Delta| \circ \frac{1}{\boldsymbol{\sigma}_\Delta} \leq c \end{cases} \right\} \quad (4)$$

where  $1/\boldsymbol{\sigma}_\Delta$  must be understood componentwise, where  $\mathbf{V}_i \in \mathbb{R}^{n \times 24}$  is the matrix representation of  $\mathbf{v}_i$ , and where  $\boldsymbol{\mu}_\Delta$  and  $\boldsymbol{\sigma}_\Delta$  are vectors in  $\mathbb{R}^n$  that correspond respectively to the sample mean and standard deviation of the step size between successive profiles for all  $n$  attributes. The third constraint in (4) enforces continuity between the daily profiles by limiting the step size with the parameter  $c > 0$ .

### 3. Case study: uncertainty related to energy data in the UK

We apply the method presented in the previous section to derive uncertainty sets for quantifiable sources of uncertainty related to power generation based on historical data for the UK in 2015. This includes national electricity demand, as well as regional data on gas demand, temperature, solar availability, offshore and onshore wind availability (we refer to the paper Charitopoulos et al. (2022) for more details). Each data

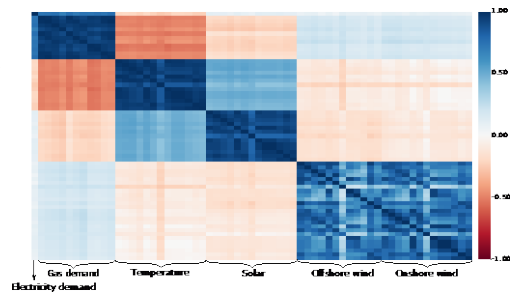


Figure 1: Correlation between all the attributes

point corresponds to the collection of hourly daily profiles of the different attributes. The size of the uncertainty space is therefore high dimensional, which motivates the need for a reduction technique to generate uncertainty sets. Based on the correlation plot (see Figure 1), three groups of attributes have strong intra-correlation and weak inter-

correlation: *seasonal* data (electricity/gas demand and temperature), *solar* data and *wind* data (off/onshore wind), i.e.  $\mathcal{A} = \{\textit{seasonal}, \textit{solar}, \textit{wind}\}$ .

Figure 3 illustrates for each month the proportions of days that are allocated to each cluster (4 clusters are generated for each group of attributes  $a \in \mathcal{A}$ ). The correlation between the solar and seasonality data is clearly apparent: the simultaneous occurrence of the blue wind cluster with the red solar cluster is for instance highly improbable. Now, if we focus on the clustering from April to October, we observe that the clusters proportions in the wind data are roughly the same for each month. This translates the idea that the seasonal data and wind data are independent for this period of the year. As a consequence of these observations, we do not define the scenarios as the Cartesian product of group-scenarios since all combinations are not as probable. Figure 2 depicts how the number of scenarios retained varies with the acceptance threshold  $\tilde{p}$ . It seems that adopting a probability threshold of  $\tilde{p} = 3\%$ , which results in keeping 14 scenarios, is in line with the trade-off between accuracy and computational tractability: a higher value would lead to too few scenarios (6 or less), while a lower value would lead to too many scenarios and therefore would not allow for the pruning of the least likely realisations.

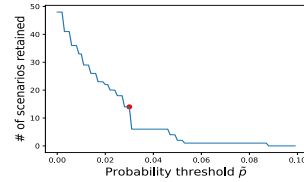


Figure 2: Number of scenarios given the probability threshold  $\tilde{p}$

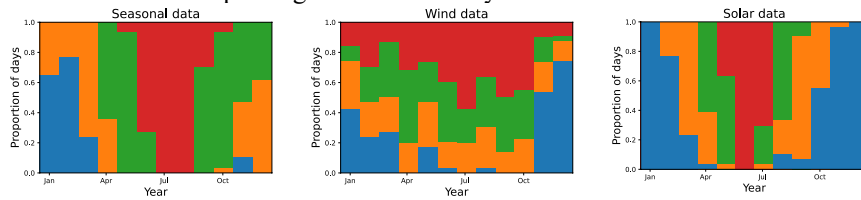


Figure 3: Proportion of days allocated to each cluster for all groups of attributes

Finally, Figure 4 depicts data generated with (3) for the parameters values:  $a = \textit{seasonal}$ ,  $k_a = \textit{blue}$ ,  $r_a = 24$ ,  $\alpha_{a,k_a} = 5\%$ ,  $\Phi_{a,k_a} = 5$ ,  $s_{a,k_a} = 24$  and  $\Psi_{a,k_a} = 1.5$ . We observe that the daily outliers in the historical data are ignored and, as desired, that the synthetic realisations are similar to the 90% of cases around the median. The size reduction is as follows:  $n_a = 40$  attributes are associated with the group  $a = \textit{seasonal}$ , which implies that the seasonal data is of size  $24n_a = 960$  and is generated on the basis of  $r_a = 24$  PCA components; this corresponds to a size reduction factor of 40. Figure 5 depicts realisations generated with (4) for the seasonal group of attributes for the succession of five days, where  $k_{1 \rightarrow 5} = \{\textit{blue}, \textit{blue}, \textit{orange}, \textit{orange}, \textit{orange}\}$  and where the continuity constraint parameter is equal to  $c = 2.5$ .

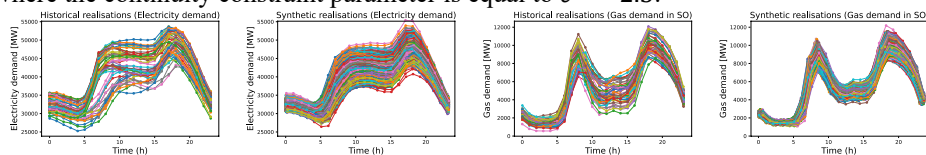


Figure 4: Comparison of historical and synthetic realisations for the blue cluster.



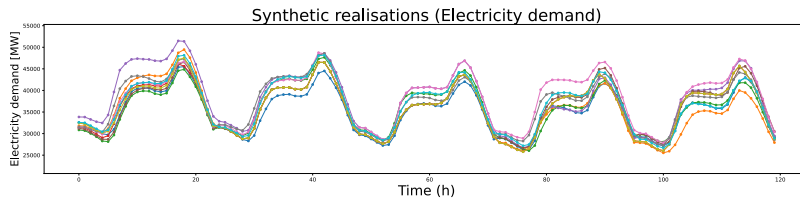


Figure 5: Synthetic realisations for  $k_{1 \rightarrow 5} = \{\text{blue, blue, orange, orange, orange}\}$ .

## 4. Results and discussion

We first point out why we did not start by clustering the data and then, for each cluster, generate a PUS, as it would implicitly take into account the (weak) correlation between attributes and avoid generating clusters based on a correlation analysis. The answer lies in two main points: Firstly, the scarcity of data. Indeed, in the case study, if we were to generate the 14 scenarios by clustering over the 365 historical data points, then each of them would have on average 26 data points. In contrast, in the method presented in this paper, only 4 clusters are generated per group of attributes. The quantiles of each cluster are then derived on the basis of around 90 data points. The PUS derived with our method are thus less subject to statistical estimate errors. Secondly, establishing groups of uncorrelated attributes allows for associating to each one of them a uncertainty dispersion budget  $\Phi_{a,k_a}$ . In this way, we do not constrain the realisations of independent attributes together.

## 5. Concluding remarks

In conclusion, the advantages of our method are the following: First, it considerably reduces the size of the original uncertainty set as it can be expressed as a linear transformation of a truncated basis using PCA. Secondly, the PUS generated are statistically meaningful and can exploit the possibly independence of uncertain attributes. Thirdly, it generates consistent uncertainty daily profiles for which the continuity from hour to hour is guaranteed. Finally, it derives polyhedral uncertainty sets that can then be used in a robust optimisation (RO) type of problem.

**Reproducibility:** All the graphs of this article can be reproduced with the code made available on [GitHub](#), where we have also shared the 2015 UK energy uncertainty dataset.

**Acknowledgements:** Financial support from the EPSRC (EP/T022930/1) is gratefully acknowledged.

## References

- A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. Robust Optimization. Princeton University Press, 2009. doi:10.1515/9781400831050.
- D. Bertsimas and M. Sim. The Price of Robustness. *Oper. Res.*, 52(1):35–53, 2004. doi:10.1287/opre.1030.0065.
- M. Bianchi, V. Gupta, and N. Kallus. Data-driven robust optimization. *Math. Program.*, 167(2):235–292, 2018. doi:10.1007/s10107-017-1125-8.
- J. R. Birge and F. Louveaux. Introduction to Stochastic Programming. Springer New York, 2011. doi:10.1007/978-1-4614-0237-4.
- V. M. Charitopoulos, M. Fajardy, C. K. Chyong, and D. Reiner. The case of 100% electrification of domestic heat in Great Britain. 2022. doi:10.17863/CAM.81913.
- Y. C. Chen. A tutorial on kernel density estimation and recent advances. *Biostat. Epidemiol.*, 1(1):161–187, 2017. doi:10.1080/24709360.2017.1396742.
- C. Ding and X. He. K-means clustering via principal component analysis. In *Twenty-first Int. Conf. Mach. Learn. - ICMML '04*, page 29. ACM Press, 2004. doi:10.1145/1015330.1015408.
- C. Ning and F. You. Data-driven adaptive nested robust optimization: General modeling framework and efficient computational algorithm for decision making under uncertainty. *AIChE J.*, 63(9):3790–3817, 2017. doi:10.1002/aic.15717.

- C. Ning and F. You. Data-driven decision making under uncertainty integrating robust optimization with principal component analysis and kernel smoothing methods. *Comput. Chem. Eng.*, 112:190–210, 2018. doi:10.1016/j.compchemeng.2018.02.007.
- C. Ning and F. You. Data-Driven Adaptive Robust Unit Commitment Under Wind Power Uncertainty: A Bayesian Nonparametric Approach. *IEEE Trans. Power Syst.*, 34(3):2409–2418, 2019. doi:10.1109/TPWRS.2019.2891057.
- L. A. Roald, D. Pozo, A. Papavasiliou, D. K. Molzahn, J. Kazempour, and A. Conejo. Power systems optimization under uncertainty: A review of methods and applications. *Electr. Power Syst. Res.*, 214:108725, 2023. doi:10.1016/j.epsr.2022.108725.
- C. Shang, X. Huang, and F. You. Data-driven robust optimization based on kernel learning. *Comput. Chem. Eng.*, 106:464–479, 2017. doi:10.1016/j.compchemeng.2017.07.004.



# Moisture and Throughput Control in an Integrated Pharmaceutical Purification Platform using PharmaPy

Inyoung Hur<sup>a</sup>, Daniel Casas-Orozco<sup>a</sup>, Gintaras Reklaitis<sup>a</sup>, Zoltan K. Nagy<sup>a\*</sup>

<sup>a</sup>*Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN 47907 USA.*

\* *Corresponding author email address: [zknagy@purdue.edu](mailto:zknagy@purdue.edu)*

## Abstract

A novel integrated continuous filter-drying carousel unit has as an enabling technology to handle filtration and drying of crystallization slurries within an end-to-end continuous pharmaceutical manufacturing train. However, due to the complexity of available contactless moisture sensing technologies, the challenges lie in developing an online monitoring system, which can accurately predict and control moisture content of the crystal product. In this work, we develop a hybrid monitoring framework to enhance the operation of this continuous filtration-drying unit by incorporating a knowledge-driven process model implemented in the object-oriented modeling platform, PharmaPy. The model is iteratively recalibrated from online measurements using a moving horizon state estimation methodology. The digital twin is shown to effectively track the offline measured moisture content. The integrated quality-by-control (QbC) framework achieves a substantial increase in process productivity and product quality.

**Keywords:** Continuous pharmaceutical manufacturing, Continuous filtration, Quality-by-Control, Digital Twin, Process Control

## 1. Introduction

Integrated filtration-drying has become an important emerging technology for continuous drug purification to produce high-purity small molecule oral drugs. [1] These steps serve as vital operations to achieve purity requirements of the final drug product [2]. Hence, process intensification (PI) of the filtration technology has attracted increased attention in the pharmaceutical sector due to its potential to achieve the critical quality attributes (CQAs) and target purity with improved economic and environmental outcomes [3]. The semi-continuous filtration carousel (CFC) has become a crucial technology by adopting separated small-scale batch process within continuous manufacturing (CM) mode. However, the CFC process is highly nonlinear due to; 1) the dynamic operation of the filtration-drying batch steps; 2) the uncertainty in the model parameters of these steps, and 3) the occurrence of unmodeled disturbances in the system, e.g., drifting critical process parameters (CPPs) due to fouling phenomena.

To overcome the challenges posed by input variability and model parametric uncertainty, quality-by-design (QbD) initiatives have been developed in the last years by the US Food and Drug Administration (FDA). QbD seeks to define the viable operating region within which the process can be operated while meeting target CQAs for given sets of critical material attributes (CMAs) and CPPs. With the increasing availability of process

analytical tools (PAT) and digital design tools, QbD has been used in various open-loop studies that create and use process digital twins to optimize process metrics (CQAs, costs, waste generation) with CPPs as decision variables [4]. For the pharmaceutical isolation step, crystal properties and final product impurities of regulated materials are the main CQAs. Since impurities remaining in the solid at the end of the drying step will be directed to the downstream process, failure to reach the drug's targeted purity will lead to off-spec products under the stringent regulatory requirements imposed on the CQAs [5]. To improve the estimation and monitoring of the CQAs, it is critical to incorporate the decision-making framework developed from the design space (DS) into the closed-loop control of the product quality to achieve a higher level of assurance.

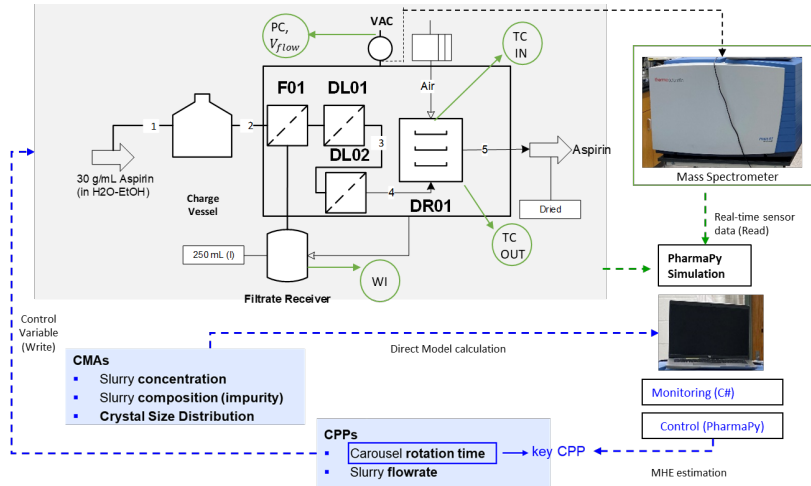
In this work, we develop a digital twin of the physical CFC system, which combines knowledge-driven model predictions calculated from dynamic flowsheet simulations with actual data retrieved in real-time from sensor measurements. The flowsheet model is simulated in the Python-based, object-oriented modeling platform PharmaPy [6]. The PharmaPy simulator is a first-principles model of the process, which includes dynamic mass, energy, and momentum balances for the multi-phase component system [7]. A moving horizon estimation (MHE) framework that relies on state estimation of average moisture content of active pharmaceutical ingredients (API) filtration cakes is applied to the flowsheet which simulates the pharmaceutical filtration-drying step of a real continuous drug substance isolation operation in CFC. This paper describes how the resulting digital twin accurately predicts the main CQAs, namely, cake moisture content (MC) and improves productivity, while minimizing plant-model mismatch.

## 2. Digital Twin Development on the Continuous Filtration Carousel (CFC)

### 2.1. Continuous CFC platform and Real-time Data Communication

In this study, we incorporate a prototype version of carousel designed and manufactured from the Alconbury Weston Ltd (UK) as a platform for performing continuous solid liquid separation of crystal slurries. The flowsheet of the process installed at Purdue University CryPTSys lab is shown in Fig 1. The main body of the carousel consists of 5 different stations in which the material goes through one of the filtration and drying steps. The slurry with API crystals is intermittently introduced into a charge vessel and transferred to the first station where filtration step occurs driven by the pressure gradient developed across the cake and filter mesh. Filtration is automatically followed by a deliquoring step (DL01-DL02) in station 3. Finally, the remaining mother liquor inside the cake is reduced by introducing a hot stream of air at DR01 (4<sup>th</sup> station). After the given rotation time, the final dried product is ejected from the system by a mechanical piston in 5<sup>th</sup> station.

Real-time data acquisition is enabled via OPC (OLE (Object Linking and Embedding) for Process Control) protocol, using the existing CFC process sensors and a mass spectrometer. A vacuum pump connected to the filtrate receiver creates a pressure gradient at the bottom of the processing stations, and pressure sensor (PC) measures the pressure drop in the system. Moreover, a flow sensor measures the flowrate of the gas coming from the DR01. The thermocouples (TCIN, TCOU) measure the temperature of the inlet and outlet of gas streams of the drying port respectively. A weighing scale, WI, monitors the mass of the filtrate collected in the filtrate receiver and a mass spectrometer is used to determine the composition of the drying gas. The data acquisition is implemented in the programmable logic control (PLC) of the unit by read/write functions and fed as the inputs to the PharmaPy simulator, which calculates the residual moisture content of the final dry product, acting as a soft sensor.



**Fig 1.** Schematic of CFC flowsheet and sensor network for a continuous purification process

## 2.2. Hybrid State Estimation Framework

A hybrid state estimation approach is presented to account for the plant-model mismatch under process disturbances. The purpose of the methodology is developing an online monitoring framework, which accurately captures the measurements from the physical sensor, hence it is used to construct controllers to control the MCs of the dry product by adapting cycle duration ( $t_{dry}$ ). First, the model calculates the process outputs by solving the mechanistic equations given real-time input information (Eq. 2). A moving horizon estimation (MHE) approach iterates this to solve optimization problem by adapting model parameters to the fluctuations of the CFC system (Eq. 1).

$$\min_{\theta_p} J(t, x, y, u, \theta_p) \quad (1)$$

$$s. t. \quad \dot{x} = PharmaPy(x, u, \theta_p), \quad x_0 = x(t_0) \quad (2)$$

$$y = h(x, u, \theta_p) + v \quad (3)$$

$$\theta_p^{lb} < \theta_p < \theta_p^{ub} \quad (4)$$

The variable  $J$  is an estimation function to determine the state disturbances,  $x = [T_{node}, x_{liq}, y_{gas}]$  is the state matrix comprised of vectors of the dynamic gas temperature, the species composition of the liquid phase, and the species composition of the gas phase from the drying port,  $y = [T_{out}^T, \sum_{i=0}^k \frac{dy_{gas}(t_i)}{dt}]$  is the measurement vector,  $u = [t_{dry}]$  is the process input,  $\theta_p = [S_{var}, k_y, h_T]$  is the vector of model parameters,  $PharmaPy$  represents the vector function of the dynamic state equations,  $h$  is the vector function of the measurement equations,  $v$  is measurement noise vector,  $k$  is the sampling instance within a cycle window.

To compensate for the performance degradation from the presence of process uncertainties, the digital twin process model should be continuously updated using the new measurements obtained at each batch cycle ( $k$ ). Here, the online state estimation framework is depicted in Fig 2. A MHE is exploited for state estimation in which the state variables are estimated by solving two consecutive optimization problems. First, the

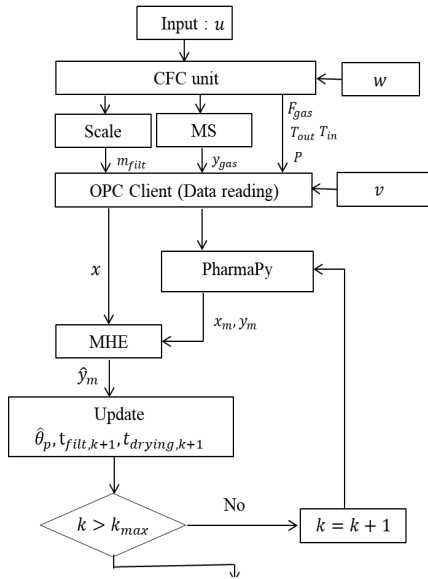


Fig 2. Block diagram of the framework

maximizing the correlation is to inform the model with an intermediate saturation value after the final deliquoring step based on different drying rate period MS data. The output of the decision variable is utilized to reinitialize the initial saturation of the deliquored

$$\min_{h_T, k_y} J = \sum_{t_i=(k-1)N_{past}}^{kN_{past}} (y_{t_i} - \hat{y}_{t_i})^T W_E (y_{t_i} - \hat{y}_{t_i}) \quad (6)$$

cake on the second problem (Eq. 6), and the problem is sequentially solved.

### 3. Result and Discussion

#### 3.1. Case study: CFC experiment on Aspirin-H<sub>2</sub>O/EtOH slurry

The monitoring framework is tested on a continuous purification process of Aspirin from solvent mixture of water-ethanol under the presence of disturbances, e.g., uncertain model parameters and control variables. The performance on the final MC prediction is examined by comparing with off-line measured MC of dry product from each batch cycle via a gravimetric method.

Fig 3 demonstrates the trajectory tracking control of the monitoring framework on the outlet temperature of the drying gas and 1<sup>st</sup> derivative of ethanol composition read from the MS (a) by adapting a set of uncertain model parameters per every cycle window (b). The soft sensor on MC of the dry product is validated against the off-line measured MC as shown in Fig 3-c. Overall, the soft sensor performs well in terms of following the trend of the measurements throughout the whole cycle.

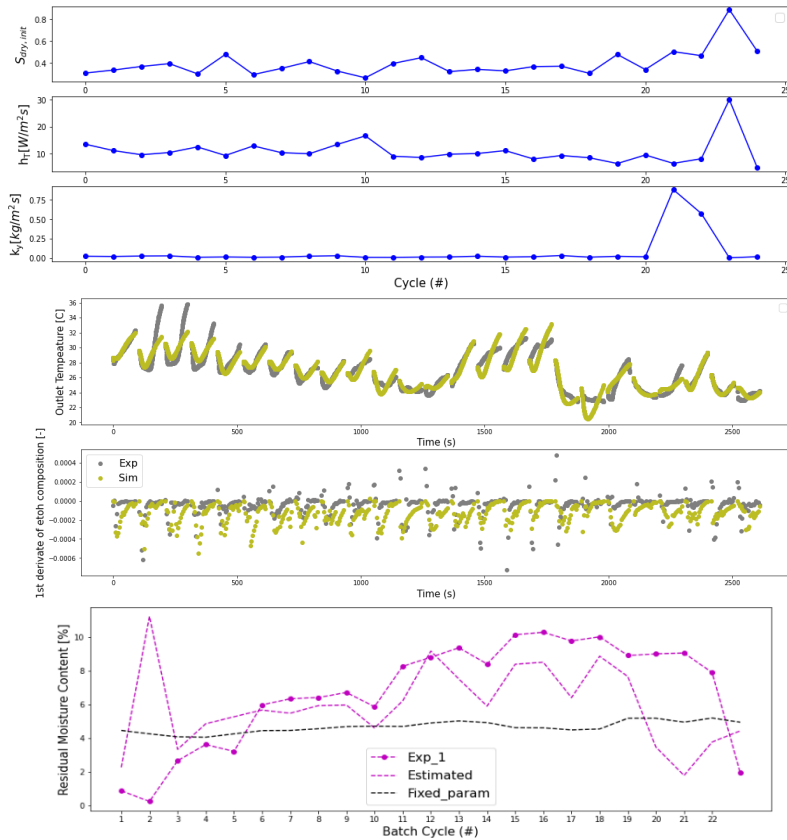
Moreover, the adaptive model parameter enables accurate capture of the relative performances of the subsequent drying batch compared to the fixed parameter model and shows good comparability to the offline measured residual MC. This proves our hybrid monitoring system allows the informative and efficient solution of state estimation problems to be real-time feasible within an online optimization framework.

distance between simulated  $f_{sim,t_i}$  and measured  $f_{p,t_i}$  gas composition is minimized by varying  $S_{var}$  (cake saturation after deliquoring) over a respective batch time frame (Eq. 5) from the best fitting line on the parity plot at  $i^{th}$  sampling point within the estimation horizon ( $N_{past}$ ) of the  $k^{th}$  cycle window.

$$\min_{S_{var}} J = \sum_{t_i=(k-1)N_{past}}^{kN_{past}} f_{p,t_i} - f_{sim,t_i} \quad (5)$$

Next, heat and mass transfer coefficients ( $h_T, k_y$ ) are estimated by solving Eq. 6.

The variable  $W$  is the weighting matrix,  $y$  is vector of measurements and  $\hat{y}$  is vector of predicted measurement. The first optimization problem (Eq. 5) is first solved to maximize the correlation of the dynamic trend of the EtOH composition in the gas phase. The purpose of



**Fig 3.** Result on online state estimation; top (a): decision variables from the problem; middle (b): comparison between prediction and measurement variables; bottom (c): comparison of moisture content (MC) from the model and measurement

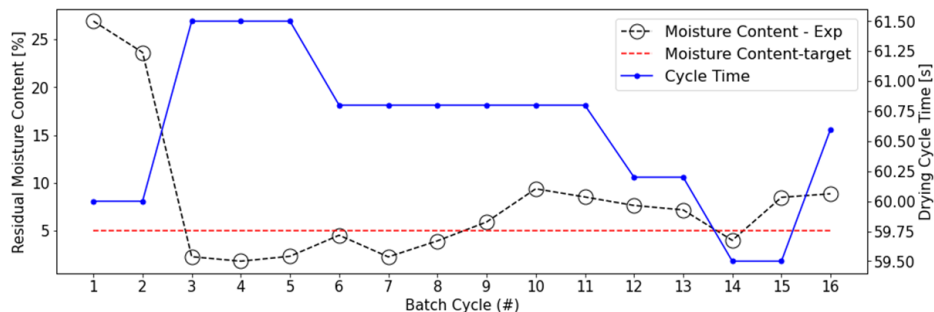
The monitoring framework is further coupled with a proportional-integral (PI) controller to update the control variables ( $t_{drying}$ ) based on the differences of the MC between prediction and the target MC (5%) (Eq. 8).

$$u(t) = u_{bias} + K_c e_i(t) + \frac{K_c}{\tau_I} \sum_{i=1}^k e_i(t) \Delta t \quad (7)$$

$$e_i(t) = \widehat{MC}_{est}(t_i) - MC_{target}, \quad MC_{target} = 5 \%$$

The initial  $t_{drying}$  of 60 s is chosen for the base control variables ( $u_{bias}$ ). The response of the control system to attain the target moisture content and resulting residual MC measurement is depicted in Fig 4. With initial cycle time, the batches produced in Cycle 1-2 was above 20% which do not meet the desired quality. The controller increases the cycle time to allow a longer operation based on the abnormal MC on the Cycle 3 and we see the MC to be regulated under the target MC. The controller starts decreasing  $t_{drying}$  by reducing the excessive operation duration, which is not needed anymore as batches produced in Cycle 3-14 are maintained either below or slightly above the target MC, hence effectively increasing the throughput. At Cycle 16, however, controller again increases the cycle time as the residual MC shows a consistent increment of its value due to the fouling of the process and simulator predicts that the CQA does not meet the quality threshold.





**Fig 4.** Controller responses and resulting measured MC of the dried product from experimentation

#### 4. Conclusions

A digital twin is developed and tested for the continuous filtration and drying operation of Aspirin and shown to produce sufficiently accurate estimates of the residual MC of the dried crystals with fluctuation in absolute errors within 4%. The estimated outputs are fed to a PI controller that adjusts MC and throughput of the dried product by manipulating drying cycle time. The result showcases a successful application of a real-time hybrid monitoring framework that offers the advantages of both knowledge-driven process modeling and optimization that rely on real-time data acquisition and analysis of the continuous filtration and drying process.

#### 5. Acknowledgments

This project was supported by the United States Food and Drug Administration through grant U01FD006738. The views expressed by the authors do not necessarily reflect official policies of the Department of Health and Human Services; nor does any mention of trade names, commercial practices, or organization imply endorsement by the United States Government.

#### References

- [1] Price, C. J et al., (2020). CHAPTER 13. Continuous Isolation of Active Pharmaceutical Ingredients. In *The Handbook of Continuous Crystallization*. The Royal Society of Chemistry
- [2] ICH Quality Implementation Working Group. Point to Consider: ICH-endorsed Guide for ICH Q8/Q9/Q10 Implementation. 2011.
- [3] Van Gerven, T. (2014). In *Workshop on Process Intensification 2014*. The Royal Society of Chemistry.
- [4] Su, Q et al., (2019). Data reconciliation in the Quality-by-Design (QbD) implementation of pharmaceutical continuous tablet manufacturing. *International Journal of Pharmaceutics*, 563(April), 259–272
- [5] International Council for Harmonisation (2016), Impurities: guideline for residual solvents Q3C(R6)
- [6] Casas-Orozco, D et al., (2021) PharmaPy: An object-oriented tool for the development of hybrid pharmaceutical flowsheets. *Computers and Chemical Engineering*, 153, 107408.
- [7] Hur, I et al., *Dynamic Flowsheet Simulation and Application of Soft Sensors on an Intensified and Integrated Purification Step for Pharmaceutical Upstream Manufacturing*, AIChE Annual Meeting, 2021

# Parameter estimation and dynamic optimization of an industrial fed-batch reactor

Jan G. Rittig,<sup>a</sup> Jan C. Schulze,<sup>a</sup> Lars Henrichfreise,<sup>a</sup> Sebastian Recker,<sup>b</sup> Rolf Feller,<sup>b</sup> Alexander Mitsos,<sup>a,c,d</sup> Adel Mhamdi,<sup>a,\*</sup>

<sup>a</sup>*RWTH Aachen University, Process Systems Engineering (AVT.SVT), Aachen, Germany*

<sup>b</sup>*LANXESS Deutschland GmbH, Cologne, Germany*

<sup>c</sup>*Forschungszentrum Jülich GmbH, Institute for Energy and Climate Research IEK-10: Energy Systems Engineering, Jülich, Germany*

<sup>d</sup>*JARA-ENERGY, Aachen, Germany*

*\*Corresponding author, email: [adel.mhamdi@avt.rwth-aachen.de](mailto:adel.mhamdi@avt.rwth-aachen.de)*

## Abstract

Modeling and optimization of fed-batch reactors with several multi-step reaction pathways is challenging due to the nonlinear dynamic system behavior and large number of kinetic parameters. We showcase the model-based optimization of an industrial (20 m<sup>3</sup>) fed-batch reactor by using our open-source dynamic optimization software DyOS. First, we build a detailed mechanistic model of the fed-batch reactor. Second, we conduct parameter estimation of the mechanistic model with 25 states and 44 fitting parameters using historic time-series industrial production data. Third, we perform dynamic multi-stage optimization and identify optimal feeding profiles for the operation, targeting improvements in economic profit over the established experience-based production routine. We demonstrate substantial economic improvement: The optimized production recipe can save up to 10% of raw material at the same yield of main product. Our findings underline the strong capabilities of model-based process optimization and its application to industrial challenges in process design and operation.

**Keywords:** Dynamic process optimization, Fed-batch reactor, Industrial application, Reaction optimization.

## 1. Introduction

The determination of feeding strategies is of paramount importance in the design and control of fed-batch reactors. Finding optimal feeding strategies is typically based on the dynamic optimization of a reactor model. However, modeling reaction processes with several multi-step reaction pathways can be challenging due to nonlinear system behavior that needs to be expressed by many equations and parameters, and a lack of mechanistic knowledge about the considered reaction, e.g., reaction kinetics and corresponding parameters. When experimental data of the reactor is available, kinetic parameters can be estimated using dynamic optimization. Subsequently, dynamic optimization can be employed for determining optimal feeding profiles with respect to a defined objective, e.g., reaction yield, product quality, and economic targets. Dynamic optimization of fed-batch reactors has been performed successfully for several applications including the production of (bio-)polymers (Zavala et al, 2005; Lopez et al., 2010, Jung et al. 2015), yeast (Hjersted & Henson, 2006) and drugs (Banga et al., 2005). We have also recently optimized fed-batch reactors for microgel (Jung et al., 2019) and polymer (Faust et al.

2021) production. To this end, we have employed our dynamic optimization software DyOS (Caspari et al., 2019), which is an open-source framework for the optimization of large-scale differential algebraic equation (DAE) systems.

Herein, we use DyOS for the optimization of feeding strategies for a fed-batch reactor of industrial scale. Specifically, we conduct three steps with the overall goal to identify optimal feeding strategies: (i) mechanistic modeling of the fed-batch reactor with the reaction kinetics, (ii) estimation of kinetic reaction and heat transfer parameters, and (iii) dynamic multi-stage optimization of raw material feeding profiles. In the first step, we develop together with the plant operators a mechanistic reactor model in the Modelica modeling language (Modelica, 2022). The parameter estimation in the second step includes the collection of historic industrial production data of the currently established experienced-based feeding profiles that we then use to fit model parameters with DyOS. In the third step, we start from the established feeding profiles and apply DyOS to identify economically optimized feeding strategies. Our in-silico results suggest that the optimized feeding strategies can lead to substantial economic improvements and savings in raw materials.

Due to confidentiality reasons, we cannot provide detailed insights on the model, the corresponding data, and the literature sources we used to collect information about the considered reaction process. We thus focus on the methodological approach of using dynamic optimization for fitting model parameters and optimizing feeding profiles. In the following, we schematically describe the considered fed-batch reactor, explain the three steps (i – iii) in detail, and present the modeling and optimization results.

## 2. Fed-Batch Reactor Model Development

We consider an industrial stirred-tank fed-batch reactor producing one main product in a multi-step reaction. The reactor has a volume of 20 m<sup>3</sup>. A schematic illustration of the reactor is shown in Figure 1. The inputs to the reactor are composed of four raw materials (RM1, RM2, RM3, RM4), water (H<sub>2</sub>O), and a catalyst (CAT). All input components are fed in liquid form, either in pure species or in aqueous solutions. RM2, H<sub>2</sub>O, and CAT are filled into the reactor before the reactions start. When RM1 and RM3 are introduced to the reactor, several multi-step reactions involving multiple intermediate products (IPs) are initiated. RM1 and RM3 are dynamically fed during the reaction process and can be used to influence the reaction rates. The production goal is the main product (MP). Additionally, several undesired byproducts (BPs) occur in the reaction competing with the MP production. The reactor temperature during operation is controlled with cooling water (CW) that flows through the cooling jacket of the reactor and enters with ambient temperature. To stop the reaction process an inhibitor (RM4) is fed to the reactor. After feeding RM4, the species concentrations remain approximately constant and the reactor can be drained.

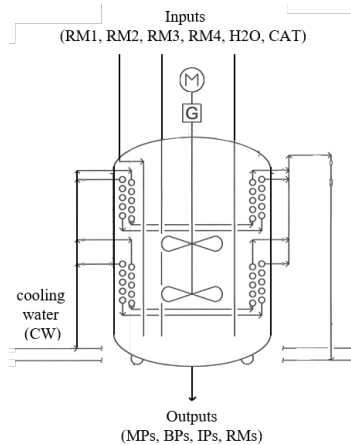


Figure 1. Schematic illustration of fed-batch reactor with input, output, and cooling water flows.

### 2.1. Mechanistic Modeling

We develop a dynamic model of the reactor and reaction system. The model includes 25 differential equations with 44 fitting parameters corresponding to the mass balance with the reaction kinetics and to the energy balance. The reaction kinetics involve 21 reactions for which reaction constants are calculated using the Arrhenius equation, i.e., two parameters in the form of the pre-exponential factor and the activation energy. Note that 3 side reactions have the same rate constant such that the reaction kinetics effectively have 38 fitting parameters in total. The energy balance accounts for the enthalpy of reactions, enthalpy flows, and cooling heat transfer rate affecting the temperature of the reaction mixture. Five reaction enthalpy parameters need to be estimated as several side reactions are modeled by the same enthalpy parameter. For the cooling, the heat transfer coefficient represents another fitting parameter, whereas we calculate the heat exchange surface and the cooling-water hold up based on the reactor geometry. For all other model parameters, e.g., as part of density and heat capacity correlations, we use standard values from the literature. In addition, we neglect other energy flows, e.g., heat loss over the reactor surface. Next to the differential equations, several algebraic equations, e.g., for pH, density, heat capacity, and yield calculation, are part of the model.

For modeling the dynamic input streams, we discretize the feed flows of RM1, RM3, RM4, and CW in one-minute intervals. Since RM2, H<sub>2</sub>O, and CAT are loaded to the reactor without any main reactions occurring, we use the respective total amount loaded to the reactor as initial values.

### 2.2. Implementation

We implement the reactor model in the modeling language Modelica (Modelica, 2022). Further, we apply our dynamic optimization framework DyOS (Caspari et al., 2019) with integrator NIXE (Hannemann et al., 2010) and NLP solver SNOPT (Gill et al., 2005). We use the Python interface of DyOS to implement and run the case study in Python. We use a Microsoft Windows Server with an Intel(R) Xeon(R) E5-2630 v2 processor running at 2.6 GHz and 128 GB RAM; simulating the model with given parameters took ca. 2 sec.

## 3. Parameter estimation

The model constitutes overall 44 fitting parameters that we initialize by data obtained from laboratory experiments and literature. However, as the simulation results when using these parametric values did not match the measured temperature profiles and final concentrations available from historical reactor data, we performed parameter estimation. To this end, we measured concentrations of some main components and the reactor temperature profile for one fed-batch run. In Figure 2, the feeding strategy for this run for (a) RM1, RM3, RM4, and (b) CW is illustrated. Subsequently, we used this experimental data for fitting the parameters of the model.

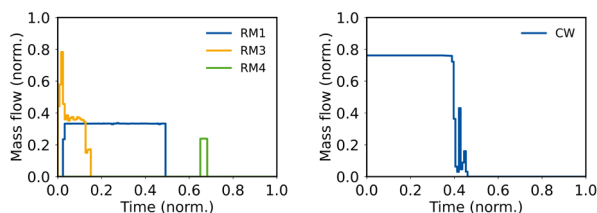


Figure 2. Feeding profiles of one fed-batch run for (a) RM1, RM3, RM4 and (b) CW.

In the parameter estimation process, we utilized DyOS to fit the 44 model parameters with the objective to minimize the mean squared error (MSE) of the model deviation to

the measured time-series of the concentration profiles and the reactor temperature profile. We solve the optimization problem using direct single shooting with optimality and feasibility tolerances of  $10^{-3}$ . We apply absolute and relative integrator tolerances of  $10^{-4}$ . Figure 3 shows the simulation results with the initial and fitted parameters in comparison to the experimental data for the temperature profile and the concentration profiles of the MP and RM2. In contrast to the initial model, the results from the fitted model match the temperature profile with slight deviations. The fitted model also predicts the concentration measurements of MP and RM2 with high accuracy. We stress that due to the large model size and large number of parameters, we did not perform a parameter identifiability study. Other parameter values might also yield a good fit but were not explored since we find a sufficient match of the experimental data using the fitted model.

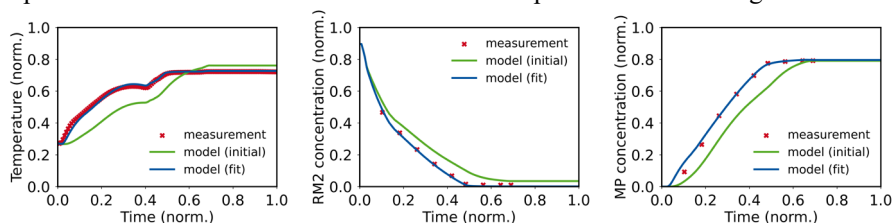


Figure 3. Comparison of the experimental results from batch run used for fitting with the simulation results obtained with the initial and the fitted model: Temperature (left) and concentrations for RM2 (middle) and MP (right).

For validation of the fitted model parameters, we conducted another experimental run of the industrial reactor at similar conditions. Figure 4 illustrates the temperature profiles and concentration measurements against the simulation results of the fitted model. We observe that the model matches the profiles closely. Some deviations regarding the MP concentration are, however, noticeable.

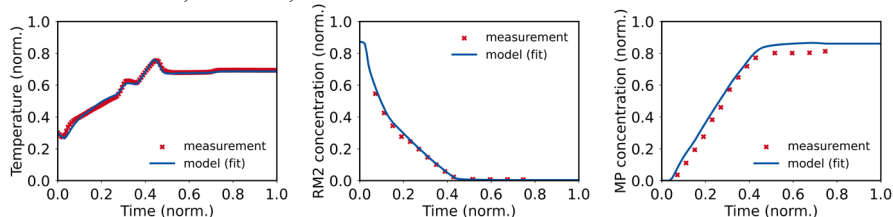


Figure 4. Comparison of the experimental results from validation batch run with the simulation results obtained with the fitted model: Temperature (left) and concentrations for RM2 (middle) and MP (right).

We also tested the model for other batch runs for which historical data on the temperature profile and the end concentrations but not all initial conditions, e.g., temperatures and impurities of raw material feed flows, was available. Here, we also observe a generally good match of the fitted model results and the experimental data, but also find some deviations, especially in the concentrations of some byproducts. We attribute the deviations to uncertainty factors in the initial conditions including BP impurities in the feed flows, limitations of the mechanistic model in capturing all characteristics of the multi-step reaction, and presumably slight overfitting of the parameters. In comparison to the initial model, the parameter fitting improves the prediction accuracy.

Overall, we conclude that the fitted model correctly captures the trends of temperature profiles and concentration measurements. In the next step, we will thus use the model with the fitted parameters for optimization of the feed profiles.

#### 4. Dynamic optimization of the reactor feeding profiles

In the optimization of the feeding profiles, we aim at identifying changes in the currently established feeding strategy to produce the same amount of MP while using fewer raw materials per batch. As a starting point for the optimization, we herein consider the feeding profiles from the fed-batch run for which we collected concentration measurements and fitted the model parameters. The objective of the optimization is to decrease the total costs of RM1 and RM3 fed to the reactor, which we formulate as

$$\begin{aligned}
 & \min_{\dot{m}_i(t)} \quad m_{RM1} \cdot c_{RM1} + m_{RM3} \cdot c_{RM3} \\
 & \text{s. t.} \quad m_{MP,cur}(t_{end}) \leq m_{MP}(t_{end}), \\
 & \quad T_{R,min} \leq T_R \leq T_{R,max}, \\
 & \quad V_{L,min} \leq V_L \leq V_{L,max}, \\
 & \quad 0 \leq \dot{m}_i(t) \leq \dot{m}_{i,max}, \\
 & \quad c_{BP,j} \leq c_{BP,j,max}, \\
 & \quad i \in \{RM1, RM3, CW\}, j \in BPs,
 \end{aligned} \tag{1}$$

where  $c_{RM1}$  and  $c_{RM3}$  are cost factors and the mass flows,  $\dot{m}_i(t)$ , of the raw materials RM1 and RM3 and the CW are the control variables, i.e., degrees of freedom. To enforce a minimal MP production, we add a constraint for the final produced amount of MP,  $m_{MP}(t_{end})$ , to at least matching the currently established strategy (cur),  $m_{MP,cur}(t_{end})$ . We consider additional constraints that include limits on the reactor temperature,  $T_R$ , the reactor level  $V_L$ , the mass flows,  $\dot{m}_i(t)$ , and concentrations of BPs,  $c_{BP,j}$ . The upper and lower limits on the reactor temperature and level arise from safety considerations, the upper limits on the mass flows are based on plant equipment limits, and the upper limit on the BPs' concentrations results from process downstream quality limits. Furthermore, the process control system only allows for a maximum number of four step changes in the mass flow of RM1 which we formulated as a multi-stage problem with free stage duration. For solving the optimization problem (Eq. 1), we apply DyOS with the same integrator and optimizer settings as for parameter estimation (cf. Section 3).

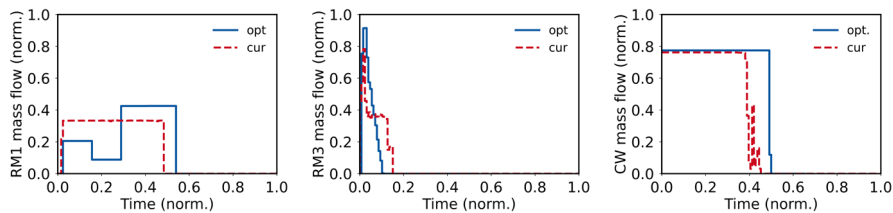


Figure 5: Optimized vs. currently established feeding strategy for RM1, RM3, and CW.

Figure 5 illustrates the currently established (dashed line) and optimized (full line) feed. According to the optimization results, RM1 should be fed in three steps with different flow rates instead one constant rate. The optimized RM3 feed is at a higher rate from the start of the batch and is then reduced almost linearly instead of in three steps. The cooling water feed is also at the maximum possible rate and changes only slightly compared to the established routine; specifically, the reactor is cooled longer. With the optimized feeding strategy, we find the produced MP amount being at the lower bound, i.e., matching the MP amount produced with the established feed profiles, which is expected as it is enforced by a constraint in the problem formulation. At the same time, the amounts of RM1 and RM3 can be reduced by 6% and 17%, respectively. This corresponds to about 10% cost savings as the same MP amount is produced by using less raw material. The main reasons for the reduced amount of required raw materials are higher main product

yield, less byproduct production, and less overfeeding of RM3. We are currently working on the experimental validation of these results.

## 5. Conclusions

We present the dynamic modeling and optimization of an industrial 20 m<sup>3</sup> fed-batch reactor. We first develop a mechanistic reactor model and then collect experimental data on the reaction process and perform parameter estimation using our in-house dynamic optimization software DyOS. Based on the developed model, we apply DyOS to identify promising feeding strategies for which the model suggests improved reaction yield and higher economic profit by 10 %. The experimental validation of the identified feeding strategies is planned as the next step with the overall goal to establish a new feeding routine for the industrial operation of the reactor.

Future work could include further model validation and potentially improvements of the model, e.g., by collecting additional experimental data that can be used for parameter fitting. Next to model refinements, more fine-grained feed profiles could be investigated to incentivize changes to the current process control system limitations. Furthermore, the presented method could be transferred to other processes with fed-batch reactors.

**Acknowledgments:** This work was funded by LANXESS Deutschland GmbH. The authors also gratefully acknowledge the financial support of the Kopernikus project SynErgie 2 by the German Federal Ministry of Education and Research (BMBF) and the project supervision by the project management organization Projektträger Jülich (PtJ). This project was further supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 466417970 - within the Priority Programme 'SPP 2331: Machine Learning in Chemical Engineering'.

## References

- J. R. Banga, E. Balsa-Canto, C. G. Moles, & A. A. Alonso, 2005. Dynamic optimization of bioprocesses: Efficient and robust numerical strategies. *J. Biotechnol.*, 117(4), 407-419.
- A. Caspari, A. M. Bremen, J. M. M. Faust, F. Jung, C. D. Kappatou, S. Sass, Y. Vaupel, R. Hannemann-Tamás, A. Mhamdi, & A. Mitsos, 2019. Dyos-a framework for optimization of large-scale differential algebraic equation systems. *Comput. Aided Chem. Eng.*, 46, 619-624.
- J. M. M. Faust, S. Hamzehlou, J. R. Leiza, J. M. Asua, A. Mhamdi, & A. Mitsos, 2021. Closed-loop in-silico control of a two-stage emulsion polymerization to obtain desired particle morphologies. *Chem. Eng. J.*, 414, 128808.
- P. E. Gill, W. Murray, & M. A. Saunders, 2005. SNOPT: An SQP algorithm for large-scale constrained optimization. *SIAM review*, 47(1), 99-131.
- R. Hannemann, W. Marquardt, U. Naumann, & B. Gendler, 2010. Discrete first-and second-order adjoints and automatic differentiation for the sensitivity analysis of dynamic models. *Procedia Comput. Sci.*, 1(1), 297-305.
- J. L. Hjersted & M.A Henson, 2006. Optimization of fed-batch *Saccharomyces cerevisiae* fermentation using dynamic flux balance models. *Biotechnol. Prog.*, 22(5), 1239-1248.
- F. Jung, F. A. Janssen, A. Caspari, H. Spütz, L. Kröger, K. Leonhard, A. Mhamdi, & A. Mitsos, 2019. Dynamic optimization of a fed-batch microgel synthesis. *IFAC-PapersOnLine*, 52(1), 394-399.
- T. Y. Jung, Y. Nie, J. H. Lee, & L. T. Biegler, 2015. Model-based on-line optimization framework for semi-batch polymerization reactors. *IFAC-PapersOnLine*, 48(8), 164-169.
- J. A. López, V. Bucalá, & M. A. Villar (2010). Application of dynamic optimization techniques for poly ( $\beta$ -hydroxybutyrate) production in a fed-batch bioreactor. *Ind. Eng. Chem. Res.*, 49(4), 1762-1769.
- Modelica, 2022. <https://modelica.org/> (accessed on 31/10/2022).
- V. M. Zavala, A. Flores-Tlacuahuac, & E. Vivaldo-Lima, 2005. Dynamic optimization of a semi-batch reactor for polyurethane production. *Chemical Engineering Science*, 60(11), 3061-3079.

33<sup>rd</sup> EUROPEAN SYMPOSIUM ON  
COMPUTER AIDED PROCESS  
ENGINEERING

VOLUME 2





**COMPUTER-AIDED CHEMICAL ENGINEERING, 52**

**33<sup>rd</sup> EUROPEAN SYMPOSIUM ON  
COMPUTER AIDED PROCESS  
ENGINEERING**

**VOLUME 2**

*Edited by*

**Prof. Antonios C. Kokossis**  
*National Technical University of Athens  
School of Chemical Engineering  
Athens, Greece*

**Prof. Michael C. Georgiadis**  
*Aristotle University of Thessaloniki  
School of Engineering  
Department of Chemical Engineering  
Greece*

**Prof. Efstratios Pistikopoulos**  
*Chemical Engineering  
Texas A&M Energy Institute  
TX, USA*



**ELSEVIER**

**Amsterdam – Boston – Heidelberg – London – New York – Oxford  
Paris – San Diego – San Francisco – Singapore – Sydney – Tokyo**

Elsevier  
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands  
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK  
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, USA

Copyright © 2023 Elsevier B.V. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: [www.elsevier.com/permissions](http://www.elsevier.com/permissions).

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

#### Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

#### British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

#### Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

ISBN (Volume 2): 978-0-443-23554-2

ISBN (Set) : 978-0-443-15274-0

ISSN: 1570-7946

For information on all Elsevier publications visit our website at <https://www.elsevier.com/>

		<b>Working together to grow libraries in developing countries</b>
<a href="http://www.elsevier.com">www.elsevier.com</a> • <a href="http://www.bookaid.org">www.bookaid.org</a>		

*Publisher:* Candice Janco  
*Acquisition Editor:* Anita Koch  
*Editorial Project Manager:* Lena Sparks  
*Production Project Manager:* Paul Prasad Chandramohan  
*Designer:* Greg Harris

Typeset by STRAIVE

# Contents

- 188. Optimization of large-scale Direct Air Capture (DAC) Model using SCR algorithm**  
*So-mang Kim, Syed Ali Zaryab, Salar Fakhraddinfakhriazar, Emanuele Martelli, Grégoire Léonard* 1179
- 189. Simulation of biodiesel production from algae**  
*Suresh Kumar Jayaraman* 1185
- 190. Different methods of model-based process optimization presented on real world examples from water industry**  
*Dr. Ewa Bozek, Dr. Corinna Busse, Delia Gietmann, Dr. Bernd-Markus Pfeiffer* 1191
- 191. A Reduced Order Model for the Prediction of the Dynamics of a Methane Reactor**  
*Enrico A. Cutillo, Erasmo Mancusi, Katarzyna Bizon, Piero Bareschino, Gaetano Continillo* 1197
- 192. A Novel Approach for Cost-effective and Sustainable Capacity Expansion Utilizing Process Intensification in Chemical Process Industries**  
*Chinmoy B. Mukta, Selen Cremaschi, Mario R. Eden* 1203
- 193. Determining the performance and network properties of petrochemical clusters**  
*Michael D. Tan, Paola Ibarra-Gonzalez, Igor Nikolic, Andrea Ramirez* 1209
- 194. Quality assessment of partial models for co-simulation of modular electrolysis plants**  
*Isabell Viedt, Kumar Rajan Gopa, Jonathan Mädler, Leon Urbas* 1215
- 195. Mechanistic modeling of industrial fermentation processes for antibiotic production**  
*Atli F. Magnússon, Stuart M. Stocks, Jari P. Pajander, Gürkan Sin* 1221
- 196. A new framework and online solution engines for multiparametric Model Predictive Control**  
*Diogo A.C. Narciso, Dustin Kenefake, Sahithi Srijana Akundi, F.G. Martins and Efstratios N. Pistikopoulos* 1227

- 197. Comparison of regression techniques for generating surrogate models to predict the thermodynamic behavior of biomass gasification systems**  
*Meire Ellen Gorete Ribeiro Domingos, Daniel Flórez-Orrego, Julia Granacher, Marie Jones, Moisés Teles dos Santos, François Maréchal* **1233**
- 198. Global optimization of symbolic surrogate process models based on Bayesian learning**  
*Tim Forster, Daniel Vázquez, Gonzalo Guillén-Gosálbez* **1239**
- 199. Molecular Property Targeting for Optimal Solvent Design in Extractive Distillation Processes**  
*Zihao Wang, Teng Zhou, Kai Sundmacher* **1245**
- 200. A Bayesian-based screening framework for optimal development of safe-by-design nanomaterials**  
*Kostas Blekos, Effie Marcoulaki* **1251**
- 201. Artificial neural network modelling for the prediction of the deactivation of CaO-based adsorbents in the calcium looping process for CO<sub>2</sub> capture**  
*Rui M. Filipe, Sérgio González-De-La-Cruz, Adrian Bonilla-Petriciolet, Carla C. Pinheiro* **1257**
- 202. Dynamic operation for the effective use of green hydrogen in Power-to-X value chains**  
*Michael Mock, Hannes Lange, Isabell Viedt, Kumar Rajan Gopa, Jonathan Mädler, Leon Urbas* **1265**
- 203. Evaluation of two-stage stochastic programming applied to the optimization of crude oil operations scheduling**  
*Tomas Garcia Garcia-Verdier, Gloria Gutierrez, Carlos Mendez, Cesar de Prada* **1271**
- 204. Generalized First-Principles Suite for Dynamic Modeling of Circular Solid Waste Thermal Treatments: a Sewage Sludge Oxy-Gasification Case Study**  
*Francesco Negri, Anna Nova, Andrea Galeazzi, Francesco Gallo, Flavio Manenti* **1277**
- 205. A Discrete-Event Simulation Approach to the Design and Planning of Biomass Supply Chains considering Technological Learning**  
*Helena Paulo, Miguel Vieira, Bruno S. Gonçalves, Susana Relvas, Tânia Pinto-Varela and Ana P. Barbosa-Povoa* **1283**

- 206. Data-driven modeling to predict the rate of Boil-off Gas (BOG) generation in an industrial LNG storage tank**  
*Suraj Prakash Singh, Rajagopalan Srinivasan, I. A. Karimi* 1291
- 207. Synthesis of process configurations and solvent blend for CO<sub>2</sub> capture from CCGT power plants**  
*Nahyan Arshad, Ahmed Al Hajaj* 1299
- 208. PREDICTION OF BIOMEDICAL WASTE GENERATION IN SANITARY EMERGENCIES FOR URBAN REGIONS USING MULTIVARIATE RECURRENT NEURAL NETWORKS**  
*Nicolas Galvan-Alvarez, David Rojas-Casadiego, Viatcheslav Kafarov, David Romo-Bucheli* 1305
- 209. Hybrid Modeling and Multi-Fidelity Approaches for Data-Driven Branch-and-Bound Optimization**  
*Suryateja Ravutla, Jianyuan Zhai, Fani Boukouvala* 1311
- 210. Intensified Process for the Production of Biojet Fuel from Mexican Biomass**  
*Rosa Angelica Morales-Gutierrez, Eduardo Sánchez-Ramírez, Brenda Huerta-Rosas, Juan José Quiroz-Ramírez, Gabriel Contreras-Zarazua, Juan Gabriel Segovia-Hernández* 1317
- 211. Metrics for Evaluating Machine Learning Models Prediction Accuracy and Uncertainty**  
*Yushi Deng, Mario R. Eden, Selen Cremaschi* 1323
- 212. Optimization of Process Families for Deployment of Carbon Capture Processes using Machine Learning Surrogates**  
*Georgia Stinchfield, Bashar L. Ammari, Joshua C. Morgan, John D. Siirola, Miguel Zamarripa, Carl D. Laird* 1329
- 213. Global Optimal Explainable Models for Biorefining**  
*Jiayang Ren, Kaixun Hua, Heather Trajano, Yankai Cao* 1337
- 214. NUMERICAL SIMULATION OF THE L-V EQUILIBRIUM WITHIN A STAGE IN A DISTILLATION COLUMN USING CFD**  
*Perla G. Canchola-López, Ariadna E. Vázquez-Hernández, Jazmín Cortez-González, Rodolfo Murrieta-Dueñas, Roberto Gutiérrez-Guerra, Carlos E. Alvarado-Rodríguez* 1345
- 215. MONITORING OF THE DRILLING REGION IN OIL WELLS USING A CONVOLUTIONAL NEURAL NETWORK**  
*Caroline Dias Grossi, Vinicius P. Barbosa, Rubens Gedraite, Maurício B. de Souza Jr., Cláudia Mirian Scheid, Luís Américo Calçada, Luiz Augusto da Cruz Meleiro* 1351

<b>216. An inverse optimization approach to decision-focused learning</b> <i>Rishabh Gupta, Qi Zhang</i>	<b>1357</b>
<b>217. Multi-scale modeling and techno-economic analysis of biogas catalytic reforming for hydrogen &amp; power production with CO<sub>2</sub> capture feature</b> <i>Alessandra-Diana Selejan, Simion Dragan, Ana-Maria Cormos, Mihaela Dragan, Calin-Cristian Cormos</i>	<b>1365</b>
<b>218 Modeling hierarchical systems via nested generalized disjunctive programming</b> <i>Hector D. Perez, Ignacio E. Grossmann</i>	<b>1371</b>
<b>219. Fusion and integrated correction of chemometrics and machine learning models based on data reconciliation</b> <i>Pál P. Hanzelik, Alex Kummer,  Ipkovich, J Abonyi</i>	<b>1377</b>
<b>220. Model-based Optimisation of Regional Nutrient Flow and Recovery for Resource and Environmental Sustainability</b> <i>Purusothmn Nair S Bhasker Nair, Nan-Hua Nadja Yang, Wei Zhang, Kok Siew Ng, Aidong Yang</i>	<b>1383</b>
<b>221. Improving results of a Continuous Fluidized Bed Process for the Separation of Enantiomers by applying mathematical optimization</b> <i>Nadiia Huskova, Jonathan G, Michael Mangold, Heike Lorenz, Andreas Seidel-Morgenstern</i>	<b>1389</b>
<b>222. Machine-learning-based optimization of operating conditions of naphtha cracking furnace to maximize plant profit</b> <i>Chonghyo Joo, Hyukwon Kwon, Junghwan Kim, Hyungtae Cho, and Jaewon Lee</i>	<b>1395</b>
<b>223. Methanol synthesis and hematite reduction using waste tires</b> <i>Athi-enkosi Mavukwana, Celestin Baraka Sempuga</i>	<b>1401</b>
<b>224. Unlocking the potentials of integrating direct air capture with HVAC system</b> <i>Yasser M. ABDULLATIF, Odi Fawwaz Alrebei, Ahmed Sodiq, Tareq AL-ANSARI, and Abdulkarem I. Amhamed</i>	<b>1407</b>
<b>225. Model development of amine scrubbing and CO<sub>2</sub> liquefaction process using waste LNG cold energy</b> <i>Jonghun Lim, Yurim Kim, Yup Yoo, Hyungtae Cho, Jaewon Lee, Jinwoo Park Junghwan Kim</i>	<b>1413</b>
<b>226. A parametric analysis of the co-gasification of biomass and plastic waste using Aspen Plus</b> <i>Nomadlozi L. Khumalo, Bilal Patel</i>	<b>1419</b>

- 227. Food security in an oligopolistic EWF nexus system: a cooperative vs a non-cooperative case**  
*Sarah Namany, Maryam Haji, Mohammad Alherbawi, Tareq Al-Ansari* 1425
- 228. Multiple-output Gaussian Process based Global Sensitivity Analysis for the cost-performance of electric vehicle in the United Kingdom**  
*Min Tao, Jude O. Ejeh, Robert A. Milton, Joan Cordiner, Solomon F. Brown* 1431
- 229. Extended Multiple-Curve Resolution framework for the calibration of first-principles models**  
*Daniel Casas-Orozco, Jaron Mackey, Ilke Akturk, Gintaras V. Reklaitis, Zoltan K. Nagy* 1437
- 230. Energy analysis of CO<sub>2</sub> capture by flash and distillation**  
*Giorgia De Guido* 1443
- 231. Feature Embedding of Molecular Dynamics-Based Descriptors for Modeling Electrochemical Separation Processes**  
*H. K. Gallage Dona, T. Olayiwola, L. A. Briceno-Mena, C. G. Arges, R. Kumar, J. A. Romagnoli* 1449
- 232. An NLP-based framework for extracting the catalysts involved in Hydrogen production from scientific literature**  
*Avan Kumar, Hariprasad Kodamana* 1455
- 233. Simulation of low-temperature district heat networks from mine water energy**  
*Thomas Cowley, Timothy Hutty, Solomon Brown* 1461
- 234. In-silico investigation of microalgae culture performance in lab and pilot scale: The trade-off between reactor size and productivity**  
*Christos Chatzidoukas, Vasileios Parisis* 1467
- 235. Process modelling of the hydrothermal liquefaction of oil-palm waste for biocrude and hydrochar production**  
*Muhammad Shahbaz, Mohammad Alherbawi, Prakash Parthasarathy, Gordon McKay, Tareq Al-Ansari* 1473
- 236. Heuristic approaches to solving multistage stochastic programs with type II endogenous uncertainty**  
*Yasuhiro Shoji, Selen Cremaschi* 1481



- 237. Hybrid modeling of the catalytic CO<sub>2</sub> methanation using process data and process knowledge**  
*Luisa Peterson, Jens Bremer, Kai Sundmacher* **1487**
- 238. Dynamic Optimization of Active Pharmaceutical Ingredient (Semi-)Batch Crystallization using Population Balance Modelling**  
*Gustavo L. Quilló, Jan F.M. Van Impe, Alain Collas, Christos Xiouras, Satyajeet S. Bhonsale* **1493**
- 239. A Systematic Framework for Iterative Model-Based Experimental Design of Batch and Continuous Crystallization Systems**  
*Hemalatha Kilari, Yash Barhate, Yung-Shun Kang, Zoltan K. Nagy* **1499**
- 240. Evaluation of Factors Affecting Novel Technology Development using Mathematical Programming**  
*Pooja Zen Santhamoorthy, Selen Cremaschi* **1505**
- T2: Control, scheduling, and operability at the process and enterprise-level**
- 241. An Encrypted MPC Framework for Security to Cyber-Attacks**  
*Atharva V. Suryavanshi, Aisha Alnajdi, Mohammed S. Alhajeri, Fahim Abdullah, Panagiotis D. Christofides* **1511**
- 242. Machine learning-based product concentration estimation, real-time optimization, and multivariable control of an experimental electrochemical reactor**  
*Berkay Çitmacı, Junwei Luo, Joon Baek Jang, Carlos Morales-Guio, and Panagiotis D. Christofides* **1517**
- 243. Symbolic regression based interpretable data-driven soft-sensor for process quality control**  
*Harry Kay, Sam Kay, Max Mowbray, Amanda Lane, Cesar Mendoza, Philip Martin, Dongda Zhang* **1523**
- 244. Oligomerization Using Reactive Distillation, Design and Model Predictive Control**  
*Gabriel Contreras-Zarazúa, Juan José Quiroz Ramírez, Eduardo Sánchez-Ramírez, Esteban Abelardo Hernández-Vargas, Juan Gabriel Segovia-Hernández* **1529**
- 245. Fault Diagnosis with GAT and PageRank: From Local Attention to Global Attention**  
*Deyang Wu, Jinsong Zhao* **1535**

- 246. Model predictive control of compact combined cycles in offshore power plants integrating a wind farm**  
*Kang Qiu, Leif E. Andersson, Cristina Zotică, Adriana Reyes-Lúa, Rubén M. Montañés, Adriaen Verheyleweghen, Valentin Chabaud, Til Kristian Vrana* **1543**
- 247. Application of Bayesian Optimization in HME Batch Concentration Process**  
*Chong Liu, Cheng Ji, Chengyu Han, Chenxi Gu, Jindong Dai, Wei Sun, Jingde Wang* **1549**
- 248. Symbolic regression-based method for developing a physics-informed surrogate model for a manufacturing process**  
*Utsav Awasthi, and George M. Bollas* **1557**
- 249. Demand Response in Microgrids with Attention-Based Deep Reinforcement Learning**  
*Jiahua Xie, Akshay Ajagekar, Fengqi You* **1563**
- 250. Multi-Kernel Canonical Variate Analysis with Bayesian Optimized Kernel Designs for Nonlinear System Identification**  
*Jan Vincent Madayag, Karl Ezra Pilario* **1571**
- 251. Modular Development of Condition Monitoring Systems for the Tennessee Eastman Process**  
*Rexonni B. Lagare, Marcial Gonzalez, Zoltan K. Nagy, Gintaras V. Reklaitis* **1577**
- 252. Data-driven predictive model for irrigation management in greenhouses under CO<sub>2</sub> enrichment and high solar radiation**  
*Ikhlas Ghiat, Rajesh Govindan, Tareq Al-Ansari* **1583**
- 253. Greenhouse temperature regulation in the presence of uncertainties using data-driven robust model predictive control**  
*Farhat Mahmood, Rajesh Govindan, Amine Bermak, David Yang, Tareq Al-Ansari* **1589**
- 254. A model predictive control framework for the production of functional, safe and sustainable nanomaterials**  
*Argyri Kardamaki, Athanassios Nikolakopoulos, Mihalis Kavousanakis, Philip Doganis, Haralambos Sarimveis* **1595**
- 255. Multi-period optimisation of oligopolies with contracts: A cooperative approach to customer fairness**  
*Asimina Marousi, Jose M. Pinto, Lazaros G. Papageorgiou, Vassilis M. Charitopoulos* **1601**

<b>256. Artificial Neural Network-Based Real-Time PID Controller Tuning</b> <i>Tate Bestwick, Kyle V. Camarda</i>	<b>1607</b>
<b>257. State of the Art Flexibility Analysis for Natural Gas Monetization Production Processes Under Uncertainties</b> <i>Noor Yusuf, Ahmed AlNouss and Tareq Al-Ansari</i>	<b>1613</b>
<b>258. Online Monitoring and Robust, Reliable Fault Detection of Chemical Process Systems</b> <i>Zheyu Jiang</i>	<b>1621</b>
<b>259. Perspectives on New Trends of Statistical Process Monitoring for Industrial Process Safety</b> <i>Cheng Ji, Jingde Wang, Wei Sun</i>	<b>1627</b>
<b>260. Aeration Optimization and Control for Wastewater Treatment Processes</b> <i>Ioana Naşcu, Wenli Du, Ioan Naşcu</i>	<b>1635</b>
<b>261. Safe deployment of reinforcement learning using deterministic optimization over neural networks</b> <i>Radu-Alexandru Burtea and Calvin Tsay</i>	<b>1641</b>
<b>262. Distributional Constrained Reinforcement Learning for Supply Chain Optimization</b> <i>Jaime Sabal Bermúdez, Antonio del Rio Chanona, and Calvin Tsay</i>	<b>1647</b>
<b>263. Machine learning-based decomposition for complex supply chains</b> <i>Niki Triantafyllou, Stavros Papaiakovou, Andrea Bernardi, Matthew Lakelin, Nilay Shah, Antonis Kokossis, Maria M. Papathanasiou</i>	<b>1653</b>
<b>264. Optimal energy management in greenhouses using distributed hybrid DRL-MPC framework</b> <i>Benjamin Decardi-Nelson, Fengqi You</i>	<b>1659</b>
<b>265. Multi-Agent Reinforcement Learning and RL-Based Adaptive PID Control of Crystallization Processes</b> <i>Qingbo Meng, Paul Danny Anandan, Chris D. Rielly, Brahim Benyahia</i>	<b>1665</b>
<b>266. Globally Convergent Composite-Step Trust-Region Framework for Model-Based Real-Time Optimization</b> <i>Duo Zhang, Xiang Li, Kexin Wang, Zhijiang Shao</i>	<b>1671</b>

- 267. Fault Detection and Diagnosis for Chemical Processes based on Deep Neural Networks with Continuous Wavelet Transform**  
*Chinatsu Ukawa, Yoshiyuki Yamashita* 1679
- 268. Anomaly detection in chemical processes with semantic knowledge graphs: an approach to reduce cause-effect diagrams**  
*Nazanin Hamedi, Ilhan Mutlu, Fatima Rani, Leon Urbas* 1685
- 269. Performance analysis of three-phase fluidized bed absorber for CO<sub>2</sub> capture industrial application**  
*Flavia-Maria Ilea, Ana-Maria Cormos, Simion Dragan, Calin-Cristian Cormos* 1691
- 270. Decentralized control for optimal operation under changing active constraints**  
*Lucas F. Bernardino, Sigurd Skogestad* 1697
- 271. Development of a Deep Learning-based Schedule-Aware Controller: Toward the Integration of Scheduling and Control**  
*M. Abou El Qassime, A. Shokry, A. Espuña, E. Moulines* 1703
- 272. Robust (explicit) optimization and control via Mixed Integer Programming**  
*Nikolaos A. Diangelakis, Iosif Pappas, Efstratios N. Pistikopoulos* 1709
- 273. A process data prediction method for chemical process based on the frozen pretrained transformer model**  
*Yiming Bai, Jinsong Zhao* 1715
- 274. Sensor bias detection, isolation, and estimation in a subsea pump system**  
*Rafael D. Oliveira, Halvor A. Krog, Galo A. C. Le Roux, Johannes Jäschke* 1723
- 275. Fusion of pupil and gaze-based features to estimate cognitive workload of control room operators**  
*Mohd Umair Iqbal, Babji Srinivasan, Rajagopalan Srinivasan* 1729
- 276. Quantification of Process Operability Using Flexibility Analysis**  
*Fei Zhao, Chentao Mei, Lingyu Zhu, Xi Chen* 1735
- 277. Coupling of support vector machine and acoustic models to characterize the droplet size distribution of emulsions using ultrasonic techniques**  
*Samuel V. Saraiva, Carlos A. M. Silva, Darlan Bonetti, Ana M. F. Fileti, e Flávio V. Silva* 1743

- 278. An Approach to Dependable Hybrid Modeling with Application to an Industrial Distillation Column**  
*Mohamed Elsheikh, Yak Ortmanns, Felix Hecht, Volker Roßmann, Stefan Krämer, Sebastian Engell* 1751
- 279. Integrating mineral mining and metallurgical supply chains: a qualogistics approach**  
*Mohammed Yaqot, Brenno C. Menezes* 1757
- 280. A simple two-parameter steady-state detection algorithm: Concept and experimental validation**  
*Evren Mert Turan, Johannes Jäschke* 1763
- 281. Residence Time Distribution characterization in a Continuous Manufacturing tableting line using PCA and PLS-DA modeling**  
*Pau Lapidra Carrasquer, Satyajeet S. Bhonsale, Liang Li, Jan F.M. Van Impe* 1769
- 282. Optimal power distribution in a P2A plant**  
*Joachim Weel Rosbo, Tobias K. S. Ritschel, Steen Hørsholt, Anker D. Jensen, John Bagterp Jørgensen, Jakob K. Huusom* 1775
- 283. Blend-scheduling optimisation for continuous and batch mixtures: modelling and solving algorithms**  
*Mahmoud A. Ahmednooh, Brenno C. Menezes, Mohammed Yaqot, Jeffrey D. Kelly* 1781
- 284. A machine learning dynamic modelling scheme for wastewater treatment plants using cooperative particle swarm optimization and neural networks**  
*Teo Protoulis, Ioannis Kalogeropoulos, Ioannis Kordatos, Haralambos Sarimveis, Alex Alexandridis* 1787
- 285. Can electrical utilities play a decisive role in generating knowledge for better use of energy in small and medium-sized industries?**  
*José Luis Suarez Castaneda, Mélissa Lemire, Marc-André Richard* 1793
- 286. Hybrid data-driven and first principles monitoring applied to the Tennessee Eastman process**  
*Eduardo Iraola, José M. Nougués, Antonio Del Rio Chanona, Lluís Batet, Luis Sedano* 1801
- 287. Data-Driven Process Monitoring for Knowledge Discovery: Local and Global Structures**  
*Estelle E. Seghers, José A. Romagnoli* 1807

- 288. Fault detection in high-pressure pumps in Low-Density Polyethylene Autoclave Reactors**  
*Maria Giuliana F. Torraga, Tahyna B. Fontoura, Rodrigo M. Lima, Luciana A. da Silva, Tiago da Silva Osório* **1815**
- 289. Using ultrasonic and process data to identify surging and gas lock phenomena during ESP operation**  
*Tiago F. Souza, Cáo C. S. Araújo, Maurício M. F. Figueiredo, Ana M. F. Fileti* **1821**
- 290. A data-driven predictive model of gas flow rate from ultrasound and process variables using artificial neural networks**  
*Cáo C. S. Araújo Tiago F. Souza, Maurício M. F. Figueiredo, Ana M. F. Fileti* **1827**
- 291. An NN-NMPC for Controlling a Crystallization Process in the Saturated and Undersaturated Zones**  
*Fernando Arrais R. D. Lima, Marcellus G. F. de Moraes, Amaro G. Barreto, Argimiro R. Secchi, Maurício B. de Souza Jr.* **1833**
- 292. A novel neural network bounds-tightening procedure for multiparametric programming and control**  
*Dustin Kenefake, Rahul Kakaodkar, Moustafa Ali, Efstratios N. Pistikopoulos* **1839**
- 293. Multistage Economic NMPC for Gas Pipeline Networks with Uncertainty**  
*Sakshi Naik, Robert Parker, Lorenz T. Biegler* **1845**
- 294. A new time-bucket MILP formulation for optimal lot-sizing and scheduling of an industrial make-and-fill process**  
*Roderich Wallrath, Florian Seanner, Matthias Lampe, Meik Franke* **1851**
- 295. Study of Different Formulations for the Multiperiod Blending Problem Applied to Lithium Recovery from Produced Water**  
*Daniel Ovalle, Joshua L. Pulsipher, Camilo Gomez, Jorge M. Gomez, Carl D. Laird, Markus G. Drouven, Ignacio E. Grossmann* **1859**
- 296. State and Parameter Estimation in Dynamic Real-time Optimization with Closed-Loop Prediction**  
*Jose Matias, Christopher L.E. Swartz* **1865**
- 297. Unsupervised anomaly detection model for diesel off-spec color change triggered by flooding**  
*A. Eren Vedin, Sadık Odemis, Aysegul Sener, Gizem Kayar, Mammad Aliyev* **1871**

- 298. Artificial neural networks-based identification of the WWTP DO sensor types of faults**  
*Norbert-Botond Mihály, Alexandra-Veronica Luca, Vasile Mircea Cristea* **1877**
- 299. Energy Out-of-distribution Based Fault Detection of Multivariate Time-series Data**  
*Umang Goswami, Jyoti Rani, Deepak Kumar, Hariprasad Kodamana Manojkumar Ramteke* **1883**
- 300. Development of a Centralized Classifier for Decentralized Decision Making**  
*Marco S. Reis, Eugeniu Strelet, Joel Sansana, Margarida J. Quina, Licínio M. Gando-Ferreira, Tiago J. Rato* **1889**
- 301. Fault detection using Fourier neural operator**  
*Jyoti Rani, Tapas Tripura, Umang Goswami, Hariprasad Kodamana, Souvik Chakraborty* **1895**
- 302. Proximal policy optimization for the control of mAB production**  
*Nikita Gupta, Shikhar Anand, Deepak Kumar, Manojkumar Ramteke, Hariprasad Kodamana* **1901**
- T3: Safe and sustainable products by design**
- 303. Process design and economic evaluation of biomass-based negative emission technologies**  
*Wei Wu, Rasa Supankanok, Walairat Chandra-Ambhorn, Nattapat Pongboriboon* **1907**
- 304. Design and intensification of the production process of propylene glycol as a high value-added glycerol derivative**  
*Jahaziel Alberto Sánchez-Gómez, Fernando Israel Gómez-Castro, Salvador Hernández* **1913**
- 305. Optimal Design of Heat Integrated Hybrid Dividing Wall Columns by Vapor Recompression**  
*Dian Ning Chia, Fanyi Duanmu, Eva Sorensen* **1919**
- 306. Efficient use of energy in distillation sequences**  
*José A. Caballero, Juan A. Labarta, Laureano Jiménez* **1925**

- 307. The use of Design Thinking in the innovation of artificial connected reefs**  
*Justine Michel, Manon Arnaud, Adrien Roudaut, Albane Tenneson, Lucie Mace, April Schnaffner, Mame Cheikh Sow, Nicolas Chaton, Nader Wehbe, Paul-Emeric Roger, Anne Asensio, Arnaud Coutu* **1931**
- 308. Substitution of hexane in vegetable oil extraction using Computer Aided Molecular Design**  
*Mohamad Nehmeh, Ivonne Rodriguez-Donis, Vincent Gerbaud, and Sophie Thiebaud-Roux* **1937**
- 309. Innovative process for manufacturing pharmaceutical mini-tablets using 3D printing**  
*Varun Sundarkumar, Wanning Wang, Zoltan K. Nagy, Gintaras Reklaitis* **1943**
- 310. Process simulation coupled with life cycle assessment of an alternative polyamide precursor**  
*Andrea Mio, Maurizio Fermeglia* **1949**
- 311. Connectivity Matrix-based Descriptors with Deep Learning for Estimation of Pure Component Properties**  
*Qiong Pan, Xiaolei Fan, Jie Li* **1955**
- 312. Novel biorefinery ideas for conversion of biomass to biofuel**  
*Niloufar Ghavami, Karhan Özdenkçi, Simeone Chianese, Dino Musmarra, Cataldo De Blasio* **1961**
- 313. Waste integration for sustainable operation of WWTP's**  
*Elena C. Blanco, Mariano Martín, Pastora Vega* **1967**
- 314. Multiscale design methodology for flow reactors supporting rapid prototyping**  
*Tomoyui Taguchi, Shigeru Kado, Toshiyuki Watanabe, Yoshiyuki Yamashita* **1973**
- 315. Computer-aided multi-scale simulation of MOF-based membrane separation for CO<sub>2</sub> removal via integration of molecular and process simulation through machine learning**  
*Xi Cheng, Yangyanbing Liao, Xiaolei Fan, Jie Li* **1979**
- 316. Decision making software for cosmetic product design based on an ontology**  
*Alex Gabriel, Juliana Serna, Valentin Plantard-Wahl, Antoine Le Jemtel, Vincent Boly, Véronique Falk* **1985**



- 317. Evaluation of amines for the removal of acid gases from the liquid hydrocarbons**  
*Siddig S. Khalafalla, Umer Zahid* **1991**
- 318. Embedding Operating Flexibility in Process Design**  
*Steven Sachio, Cleo Kontoravdi, Maria M. Papathanasiou* **1997**
- 319. Transfer learning for process design with reinforcement learning**  
*Qinghe Gao, Haoyu Yang, Shachi M. Shanbhag, Artur M. Schweidtmann* **2003**
- 320. Data augmentation for machine learning of chemical process flowsheets**  
*Lukas Schulze Balhorn, Edwin Hirtreiter, Lynn Luderer, Artur M. Schweidtmann* **2009**
- 321. Systematic synthesis of pathways for hydrogenation and hydrogenolysis of acetylene over catalytic surfaces**  
*Manjusha C. Padole, Swayam P. Misra, V. Sai Phani Kumar, Parag A. Deshpande* **2015**
- 322. New Methodology Integrating Tray Efficiency Predictions with Internals Design for Valve-Tray Distillation Columns**  
*Chenguang Zhu, Megan Jobson, Nan Zhang* **2021**
- 323. Process Design and Intensification of Circulating Catalytic Fluidized Bed Membrane Reactor for Oxidative Coupling of Methane**  
*Moustafa Ali, Yuhe Tian, Dustin Kenefake, Efstratios N. Pistikopoulos* **2029**
- 324. Solvent pre-selection for extractive distillation using Gibbs-Helmholtz Graph Neural Networks**  
*Edgar Ivan Sanchez Medina, Kai Sundmacher* **2035**
- 325. Dhrushti-AI: A multi-screen multi-user eye-tracking system to understand the cognitive behavior of humans in process industries**  
*Thasnimol Valuthottiyil Shajahan, Rahul Madbhavi, Mohammed Aatif Shahab, Babji Srinivasan, Rajagopalan Srinivasan* **2041**
- 326. Digital Twin- A System for Testing and Training**  
*Michael Schueler, Tanja Mehling* **2047**

- 327. Process Validation in Modular Plants considering the Scale-Up and Prior Knowledge**  
*J. Mädler, A. Koch, A. Bamberg, L. Urbas* **2055**
- 328. Digital Twins for Scale-Up in Modular Plants: Requirements, Concept, and Roadmap**  
*A. Koch, J. Mädler, A. Bamberg, L. Urbas* **2061**
- 329. Modeling and Simulation of a PSA Process for CO<sub>2</sub>/Syngas Separation**  
*Magno Fonseca Santos, Stefano Ferrari Interlenghi, Antonio Esio Bresciani, Newton Libanio Ferreira, Gabriel Soares Bassani, Rita Maria de Brito Alves* **2067**
- 330. Simultaneous design of integrated cyanobacteria-based biorefinery and its heat exchanger network**  
*Matías Ramos, Romina Lasry Testa, Fernando Ramos, Vanina Estrada, Maria Soledad Diaz* **2073**
- 331. Sustainable Conceptual Design of a Hydrogen Production via Natural Gas pyrolysis**  
*Patience B. Shamaki, Galo A.C Le Roux* **2079**
- 332. The influence of biomass characteristics and their uncertainties on the production of sustainable aviation fuel**  
*Moaaz Shehab, Diego Freire Ordóñez, Mai Bui, Kai Moshhammer, Edwin Zondervan* **2087**
- 333. Translation Ontology of optimal Decision Making along a Distributed Production Chain by Example**  
*Peter Klein, Heinz A. Preisig, Natalia Konchakova* **2093**
- 334. The Integration of Drug Substance and Drug Product Manufacturing Models: The Missing Link for Model-based End-to-End Process Development**  
*Charalampos Christodoulou, Samir Diab, Gabriele Bano, Magdalini Aroniada, Neil Hodnett, Simeone Zomer* **2099**
- 335. Direct Compression Flowsheet Modelling to Assess Tablet Manufacturability and *In Vitro* Dissolution**  
*Magdalini Aroniada, Charalampos Christodoulou, Houda Khaled* **2105**
- 336. Enhanced hot-liquid water pretreatment of biomass with recovery and valorization of side products**  
*Tamara Janković, Adrie J. J. Straathof, Anton A. Kiss* **2111**

- 337. Development of deep learning framework to predict physicochemical properties for Ionic liquids**  
*Sadah Mohammed, Fadwa Eljack, Saad Al-Sobhi, Monzure-Khoda Kazi* 2117
- T4: Green and sustainable processes for the circular economy**
- 338. Economic optimization of the Northern Italian supply chain for residual plastic packaging waste treatment**  
*Fabio Cieno, Daniel Cristiu, Federico d'Amore, Fabrizio Bezzo* 2123
- 339. Optimisation modeling for decision support in the industrial circular economy activities**  
*K. G. Stylianopoulou, E. M. Kondili, C. M. Papapostolou, J. K. Kaldellis* 2129
- 340. Model-Based Design of Experiments for the identification of microalgae growth models with limiting nutrients**  
*Alberto Saccardo, Beatriz Felices-Rando, Eleonora Sforza, Fabrizio Bezzo* 2135
- 341. Enabling circularity: Analysis of factors influencing MSW sorting behaviour in Central Norway**  
*Tuva Grytli, Cansu Birgen* 2141
- 342. A Critical Look at Waste Composition Analyses: Challenges and Opportunities**  
*Tuva Grytli, Cansu Birgen* 2147
- 343. Waste valorisation within the Energy-Water-Food Nexus: A hybrid techno-geospatial optimisation approach**  
*Mohammad Alherbawi, Sarah Namany, Maryam Haji, Gordon McKay, Tareq Al-Ansari* 2153
- 344. Levelised Energy-Water-Food (LEWF) Indicator for Biomass Assessment and Optimisation**  
*Mohammad Alherbawi, Ahmed AlNouss, Gordon McKay, Tareq Al-Ansari* 2159
- 345. The controlled environment agriculture: a sustainable agrifood production paradigm empowered by systems engineering**  
*Liang Wang, Guoping Lian, Zoe Harris, Mark Horler, Yang Wang, Tao Chen* 2165

- 346. CO<sub>2</sub> Capture by Amine-Functionalized Magnesium Oxide: Experimental and Modelling Studies**  
*Sagheer A. Onaizi* 2171
- 347. Development of ‘Green’ LNG through a CO<sub>2</sub> allocation procedure within a CO<sub>2</sub> utilisation network in Qatar**  
*Razan Sawaly, Mohammad Alherbawi, Tareq Alansari* 2177
- 348 Process Design and Bayesian Optimization of 5-Hydroxymethylfurfural Hydrodeoxygenation**  
*Yuqing Luo, Zhaoxing Wang, Prahalad Srinivasan, Dionisios G. Vlachos, Marianthi Ierapetritou* 2183
- 349. Investigation of biomass blending ratios for optimal biochar’s soil application**  
*Farah Obar, Mohammad Alherbawi, Gordon Mckay, Tareq Al-Ansari* 2191
- 350. Methodology for estimating a country’s production of biomethane from agricultural residues: a multiscale and holistic approach**  
*Manuel Taifouris, Mariano Martin* 2197
- 351. Energy management for controlled environment agriculture based on physics informed neural networks and adaptive linearization based data-driven robust model predictive control with AI**  
*Guoqing Hu, Fengqi You* 2203
- 352. Life Cycle Assessment on Chemical Recycling-aided Cascaded Use of Polypropylene Plastics**  
*Xiang Zhao, Fengqi You* 2209
- 353. Enviro-economic analysis of tandem and direct processes for ethylene electrosynthesis**  
*Grazia Leonzio, Benoît Chachuat, Nilay Shah* 2215
- 354. Modeling and Experimental Validation of Poly(3-Hydroxybutyrate-co-3-Hydroxyvalerate) Chain Length Distribution**  
*Stefanie Duvigneau, Annette Wilisch-Neumann, Robert Dürr, Achim Kienle* 2221
- 355. Predictive Modeling and scale-up of Wet Oxidation for Hydrothermal Liquefaction Process Water treatment**  
*Carolin Eva Schuck, Thomas Schäfer, Konstantinos Anastasakis* 2227

- 356. Process Feasibility for Converting the Waste Plastics to Hydrogen and Methanol**  
*Ali A. Al-Qadri, and Usama Ahmed* 2233
- 357. Process development for efficient hydrogen transportation and analyses: technical, economic, and environmental perspectives**  
*Byeongchan Ahn, J. Jay Liu, Wangyun Won* 2239
- 358. Integration of metabolic models in biorefinery designs using superstructure optimisation**  
*Lucas Van der Hauwaert, Alberte Regueira, Miguel Mauricio-Iglesias* 2247
- 359. Environmental Impacts of Bio-derived Silicon: Uncertainty in the Benefit of Industrial Transition**  
*Ethan Errington, Miao Guo, Jerry Heng* 2253
- 360. Technoeconomic and Life Cycle Assessment of an mRNA Vaccine Integrated Manufacturing Plant**  
*Shang Gao, Brahim Benyahia* 2259
- 361. Obtaining biofuels from agroindustrial waste of *Fragaria spp* and *Mangifera indica spp***  
*Ernesto Barrera-Reyes, Luis Alberto Bretado-Aragón, Guillermo Adolfo Anaya-Ruiz, Pascual Eduardo Murillo-Alvarado, José Luis Baltazar-Álvarez, José Joel Román-Godínez and Gabriela Guadalupe Esquivel-Barajas* 2265
- 362. Life Cycle Assessment of Cellulose Nanocrystals Production in Sugarcane Biorefineries**  
*Gustavo Batista, Ana Carolina Borges Silva, Cristiane Sanchez Farinas, Antonio José Gonçalves da Cruz* 2271
- 363. Decarbonization of energy-intensive industries: sustainable implementation of CO<sub>2</sub> recycling within the industrial symbiosis**  
*Marta Rumayor, Javier Fernández-González, Antonio Domínguez-Ramos, Angel Irabien* 2277
- 364. Gray and hybrid green ammonia price sensitivity to market fluctuations: the Russia-Ukraine war case**  
*Andrea Isella, Alberto Lista, Gabriele Colombo, Raffaele Ostuni, Davide Manca* 2283
- 365. Paving the way for the integration of synthesis, assessment, and design tools within an ontological framework**  
*Adrián Pacheco-López, Kristiano Prifti, Flavio Manenti, Ana Somoza-Tornos, Moisés Graells, Antonio Espuña* 2289

- 366. Pyrolysis of livestock manures: Optimal operating conditions and feedstock blending ratios**  
*Prakash Parthasarathy, Mohammad Alherbawi, Snigdhendubala Pradhan, Gordon McKay, Tareq Al-Ansari* 2295
- 367. Investigating the feasibility of biomass integrated gasification combined cycle (BIGCC) as a power alternative in Gaza Strip**  
*Rawan Shihada, Rania Salem, Amanda Munia, Nashat Naim, Mohammad Alherbawi* 2301
- 368. The Infinity Reactor: A new conceptual design for a more cost-efficient CO<sub>2</sub> to methanol route**  
*Hilbert Keestra, Edwin Zondervan, Wim Brilman* 2307
- 369. Optimal design of a microbial electrolysis cell for biohydrogen production**  
*Uriel Roberto Pedroza-Medina, Luis Felipe Cházaro-Ruiz, Ricardo Femat, Alicia Román-Martínez* 2315
- 370. Simulation and Exergoeconomic Analysis of Brewer's Spent Grains convective drying process**  
*Mero-Benavides M., Enriquez-Posligua J., Barcia-Quimi A.F., Tinoco-Caicedo Diana L* 2321
- 371. Incorporating negative emissions technologies with policy instruments for net-zero emissions**  
*Elizabeth J. Abraham, Patrick Linke, Dhabia M. Al-Mohannadi* 2329
- 372. Waste valorization of non-commercialized edible mushrooms**  
*Brenda I. Lazaro-Molina, Teresa Lopez-Arenas* 2337
- 373. Enviro-economic assessment of sustainable aviation fuel production from direct CO<sub>2</sub> hydrogenation**  
*Andrea Bernardi, Daniel Bagan Casan, Andrew Symes, Benoit Chachuat* 2343
- 374. Ammonia powered solid oxide fuel cells for general aviation propulsion systems: challenges and opportunities**  
*Giampiero Di Legge, Seyed Ali Nabavi, Lorenzo Mazzei, Riccardo Da Soghe, Cosimo Bianchini, Amirpiran Amiri, Theoklis Nikolaidis, Soheil Jafari* 2349



# Optimization of large-scale Direct Air Capture (DAC) Model using SCR algorithm

So-mang Kim,<sup>a</sup> Syed Ali Zaryab,<sup>b</sup> Salar Fakhraddinifakhriazar,<sup>a</sup> Emanuele Martelli,<sup>b</sup> Grégoire Léonard<sup>a</sup>

<sup>a</sup>*Chemical Engineering, University of Liège, B6a Sart-Tilman, 4000, Liège, Belgium*

<sup>b</sup>*Politecnico di Milano, Department of Energy, Via Lambruschini 4, Milan 20156, Italy*

## Abstract

This paper describes optimization studies of a large-scale fixed bed reactor model using Aspen Adsorption cycle simulations. The model uses Lewatit® VP OC 1065 amine-functionalized adsorbents and temperature vacuum swing adsorption (TVSA) cycles to capture CO<sub>2</sub> from the ambient air. Building a comprehensive direct air capture (DAC) model to optimize process design and cost is crucial to assess the feasibility of DAC technologies and therefore, a novel surrogate-based derivative free global optimization algorithm, referred to as SCR is implemented to evaluate a large-scale DAC system. Lastly, in order to achieve zero emissions and assess the viability of deploying large-scale DAC systems, assessments involving emission factors of various energy sources in different countries' electricity energy grid systems are studied. The optimization study showed a reduction in capture cost by 45% compared to the base case.

**Keywords:** Direct air capture, TVSA, process modelling, surrogate-based optimization

## 1. Introduction

Negative emission technologies (NETs) with a cumulative CO<sub>2</sub> capture of 450 to 1000 Gt are necessary to meet the strict 1.5 °C temperature increase target between 2010 and 2100 (Rogelj et al., 2015). One of the potential NETs is Direct Air Capture (DAC), and the likelihood of technological progress could significantly reduce DAC's energy usage, making it a practical long-term mitigation strategy (Realmonte et al., 2020). DAC was first proposed by Lackner and coauthors in 1999 to reduce climate change and the captured CO<sub>2</sub> may be used in various ways, including enhanced oil recovery, construction materials, and production of goods like chemicals, fuels, cement, etc. The CO<sub>2</sub> can also be permanently stored in deep geological formations (IEA, 2022).

Negative emission technologies (NETs) with a cumulative CO<sub>2</sub> capture of 450 to 1000 Gt are necessary to meet the strict 1.5 °C temperature increase target between 2010 and 2100 (Rogelj et al., 2015). One of the potential NETs is Direct Air Capture (DAC), and the likelihood of technological progress could significantly reduce DAC's energy usage, making it a practical long-term mitigation strategy (Realmonte et al., 2020). DAC was first proposed by Lackner and coauthors in 1999 to reduce climate change and the captured CO<sub>2</sub> may be used in various ways, including enhanced oil recovery, construction materials, and production of goods like chemicals, fuels, cement, etc. The CO<sub>2</sub> can also be permanently stored in deep geological formations (IEA, 2022).

There are several DAC techniques, including membrane, chemical, and electro-swing adsorption (ESA) methods (IEA, 2022, Voskian and Hatton, 2019, Fujikawa et al., 2021). The chemical methods can be further classified as a liquid solvent and solid sorbent



techniques which are the furthest technologies in development. In the case of the liquid solvent approach, an aqueous KOH solution is used as the capture medium and a calcium caustic loop is used as the recovery mechanism at high temperatures between 300-900°C (Keith et al., 2018). Solid sorbents technologies usually employ amine groups on the internal surface of the sorbent to react with CO<sub>2</sub> from the ambient air at low pressure, followed by regeneration at medium temperature of 80 – 120°C (McQueen et al., 2021). There are 18 operating DAC plants that capture almost 0.01 MtCO<sub>2</sub> annually according to the IEA report (2022) however, the majority of current DAC plants operate on a small scale for testing and demonstration purposes. Also, few studies for DAC have been conducted on process scales (Schellevis et al., 2021). Therefore, in order to address the aforementioned gaps, this paper presents a large-scale fixed bed reactor system of Kim and Léonard (2022), to capture CO<sub>2</sub> from ambient air via Temperature Vacuum Swing Adsorption (TVSA) using Lewatit® VP OC 1065 amine-functionalized adsorbents. Moreover, important operating parameters impacting the energy and cost of DAC systems are identified and CO<sub>2</sub> capture cost is optimized via SCR algorithm. Lastly, the viability of deploying DAC systems based on various electricity grid systems in different countries is explored to achieve net zero emissions.

## 2. Modelling

A large-scale one-dimensional fixed bed reactor model for gas adsorption (at 400 ppm CO<sub>2</sub>) presented in the work of Kim and Léonard (2022b) is studied further to find feasible operating conditions under TVSA cyclic operation (adsorption, evacuation, desorption and cooling). The large-scale model was scaled up in the aforementioned work from a kg-scale DAC model where this small-scaled model was validated with experimental results presented in the works of Bos et al. (2019) and Yu et al. (2017). The validations of equilibrium loadings of CO<sub>2</sub> onto the solid sorbent, breakthrough curves and model assumptions are presented in Kim and Léonard (2022a&b).

## 3. Optimization Methodology

In this work, the DAC unit is being optimized using a recently published optimization algorithm called SCR which was presented by Zaryab et. al. (2022). SCR is a surrogate - based derivative free optimization algorithm; this algorithm was chosen because the simulation model of DAC has a very high computational cost, but the use of a surrogate model helps reduce it. SCR considers the DAC simulation code as a black box function and creates Kriging-based surrogate models of this function and its constraints. SCR then finds the region where the global minimum of this surrogate model lies and then calls the black box function within this region to look for the optimum. Using this optimization strategy helps to greatly reduce the number of functional evaluations required to reach the optimum thus saving significant time. For the optimization a total of five optimization variables have been selected, these include: the adsorption time ( $t_{ads}$ ), depressurization time ( $t_{eva}$ ), desorption time ( $t_{des}$ ), cooling time ( $t_{cool}$ ) and the vacuum pressure ( $P_L$ ). In the case of the desorption temperature ( $T_{des}$ ), Aspen Adsorption model was not robust enough to allow  $T_{des}$  to be a part of the optimization variables within the time limitation and therefore, two optimization runs were carried out at two fixed desorption temperatures ( $T_{des}$ ) of 352.15 K and 360.15 K. The goal of the optimization is to minimize the specific cost of the DAC unit (€/ton<sub>CO2</sub>). The constraints of the optimization are that the recovery should be greater than 90% and the purity should be greater than 85%.

## 4. Results and Discussions

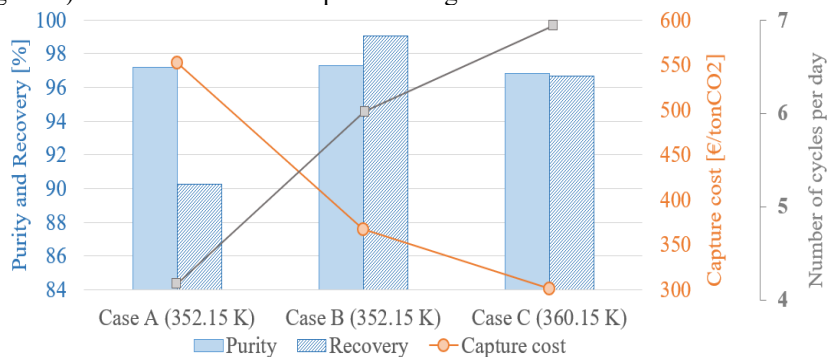
### 4.1. Optimization of large-scale DAC systems

A benchmark study of Climeworks Hinwil plant (Climeworks, 2017) capturing 2460 kg/day with 18 DAC units was selected in the work of Kim and Léonard (2022b), where a single large-scale DAC unit model was developed. This model could capture 127 kg of CO<sub>2</sub> per day via 4 cycles and this led to a total capture of 2286 kg/day using 18 DAC units which is similar to the Climeworks Hinwil plant's capture scale. Based on this result, a CO<sub>2</sub> capture cost of 620.7 €/tonCO<sub>2</sub> was reported in the aforementioned work. In the current paper, the DAC costs of Kim and Léonard (2022b) are re-evaluated with updated equipment and utility costs (case A) and then optimized using SCR algorithm of Zaryab et al. (2022) at two  $T_{des}$  of 352.15 K (case B) and 360.15 K (case C). Case A's revised capture cost was found to be 552.78 €/tonCO<sub>2</sub> and then TVSA step times ( $t_{ads}$ ,  $t_{eva}$ ,  $t_{de}$ ,  $t_{cool}$ ) and  $P_{vac}$  were optimized in the cases B and C using SCR by minimizing the capture cost (€/tonCO<sub>2</sub>). Variables before and after optimization are presented in Table 1 below.

**Table 1.** Summary of variables before and after optimization

	Before optimization		After optimization	
	Case A	Case B	Case C	
$T_{des}$ [K]	352.15	352.15	360.15	
$P_{vac}$ [bar]	0.33	0.37	0.46	
$t_{ads}$ [s]	6500	6467	6664	
$t_{eva}$ [s]	30	72	56	
$t_{des}$ [s]	6700	7539	5561	
$t_{cool}$ [s]	6300	1535	1000	
Captured CO <sub>2</sub> per cycle [kg/cycle]	31.66	34.57	34.75	
Capture cost [€, 2021/tonCO <sub>2</sub> ]	552.78	367.23	301.35	

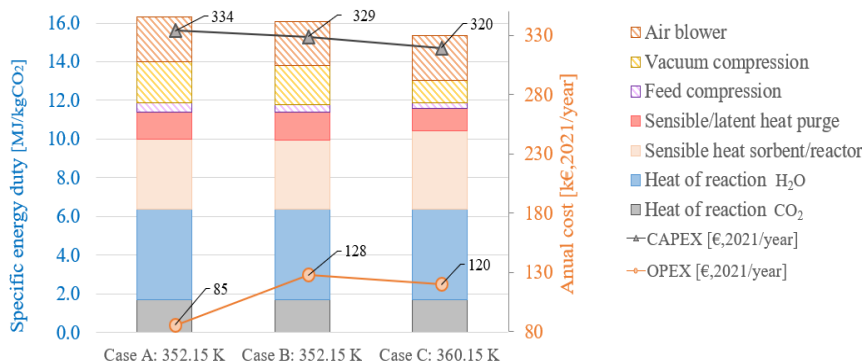
It was found in both cases B and C that  $t_{cool}$  has been shortened significantly. Before optimization, the adsorption bed temperature was allowed to cool down to ambient temperature while in the optimized cases, due to the shortened cooling time, the bed temperature is not fully cooled down to its initial temperature. However, as air is fed into the bed during the adsorption step, the bed rapidly cools down to its initial state. This may seem highly ideal but the temperature profile of the simulation can be further refined through experimental data to validate the cooling steps. Based on the simulation results, the drastic decrease in cooling time allows more cycles to be achieved per day (grey line) and this allows more CO<sub>2</sub> to be captured which reduces capture cost per ton of CO<sub>2</sub> (orange line) for Case B and C as depicted in Figure 1.



**Figure 1.** Purity and recovery of the system with capture cost and number of daily cycles

The purity of the different cases stayed around 97% while recoveries of case B increased to 99% from 90% (case A). This is due to an increase in desorption time in case B which allows improved utilization of the adsorption bed. Case C shows desorption at a higher temperature of 360.15 K with longer  $t_{ads}$  (6664 s) and the shortest  $t_{des}$  of 5561 s among cases. Although the recovery may be slightly lower (97%) than case B (99%), this indicates that it is more advantageous to operate at a higher  $T_{des}$  as even at a shorter desorption time, a large recovery per cycle can be achieved.

Specific energy duties of the cases are also calculated and presented in Figure 2 with the annualized CAPEX and OPEX. The electrical energy includes feed (air) compression, vacuum compression of purge streams and air blowers (in hatched column graph sections) while the thermal energy comprises the heat of reaction of  $\text{CO}_2/\text{H}_2\text{O}$ , sensible and latent heat of purge streams as well as sensible heat of sorbents and reactors (in solid column graph sections). Note that water adsorption isotherm is not included in this study and therefore, a theoretical heat of reaction of  $\text{H}_2\text{O}$  (the energy required for  $\text{H}_2\text{O}$  desorption) is calculated using the vaporization enthalpy (Veneman et al., 2015).



**Figure 2.** Specific energy duties (column graphs) with annualized costs (line graphs)

There are small changes in specific energy duties in the case of A and B (before and after optimization at 352.15 K respectively). For the optimized results at 360.15 K (case C), due to its higher  $T_{des}$ , there is an increase in sensible heat of sorbent/reactor but operating at higher  $T_{des}$  is beneficial since this reduces compression duties and sensible/latent heat of purge streams (less steam is required) which makes the overall specific energy duty to decrease. Climeworks process needs between 1.8 and 2.6 MJ/kg of electrical energy and between 5.4 and 11.9 MJ/kg of thermal energy (Deutz and Bardow, 2021), where the minimum value represents the target value for the future and the maximum value depicts the energy requirements at present. For all cases, the thermal energy is in the agreement with the reported range in literature while the electrical energies including air blowing obtained in this study are in the same magnitude but higher in both cases B and C. However, in case C, an electrical duty of 3.76 MJ/kg was estimated which is 31% higher than the upper boundary from the literature. More detailed optimization can be performed to further reduce the electrical duties.

In terms of economics, the annualized CAPEX throughout the cases showed a small change since the dimensions of the columns and the air feed flow rates are kept constant. However, due to the increase in the number of cycles per day in cases B and C, higher operating costs are expected. Nevertheless, the capture costs after the optimization showed a reduction of 33.6% and 45.5% for case B and C respectively as compared to case A since more  $\text{CO}_2$  can be captured with a small increase in costs.

#### 4.2. Achieving net-zero emission with DAC

As it is known, the energy source for DAC systems is a key factor in achieving negative emissions and, low-cost and low-carbon resources are crucial to obtain this goal of the DAC process. To point out this issue, the assessment of integrating different renewable resources into DAC systems has been suggested in the literature (Zolfaghari et al., 2022). Moreover, using currently available electricity sources for DAC is considered the most cost-effective approach for this technology (Hanna et al., 2021). One way to decrease the amount of emitted CO<sub>2</sub> by the DAC process itself is the utilization of electricity not only for electrical equipment (e.g. fans and pumps) but also for heat generation. Therefore, in this section, the total energy requirement, which is acquired through the optimization of the DAC model, is used to calculate the amount of CO<sub>2</sub> produced from DAC operations. The energy mix of electricity grids presented in Figure 3 (Global change data lab, 2022) and the emission factor of different sources from Bruckner et al. (2014) are used to assess the possibility of achieving negative emission in the case of a fully electrified DAC process.

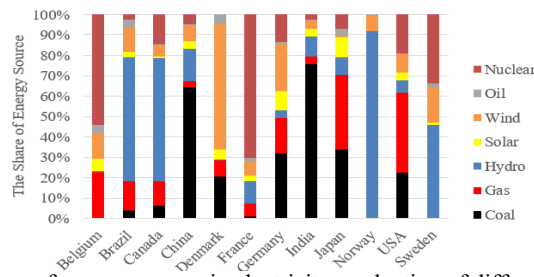


Figure 3. Share of energy sources in electricity production of different countries.

As it is depicted in Figure 4, due to the lower CO<sub>2</sub> emission from electricity grid systems in countries with a higher share of renewables or nuclear energy, the usage of grid electricity for DAC is effective to achieve negative emissions in countries like Norway, Sweden and France, producing about 0.15-0.3 kg<sub>emitted,CO2</sub>/kg<sub>captured,CO2</sub>. On the other hand, the application of grid electricity for carbon removal from ambient air in countries that have a higher share of fossil fuels in power productions (e.g. China, India, and the USA) results in producing more carbon dioxide than captured. Nevertheless, to be able to compare DAC processes with other capture technologies and minimize the side effects of these systems, a detailed Life Cycle Assessment (LCA) is essential.

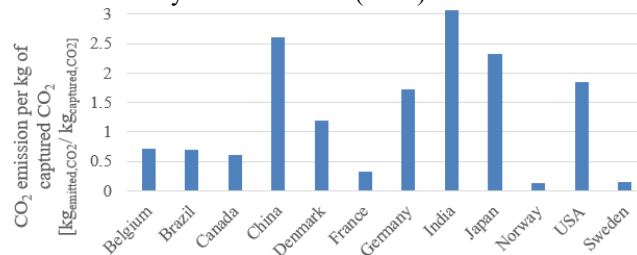


Figure 4. The amount of produced CO<sub>2</sub> through DAC process.

### 5. Conclusion and perspectives

In this paper, a large-scale DAC model of Kim and Léonard (2022b) is studied. The capture cost (€,2021/tonCO<sub>2</sub>) of DAC systems was minimized by varying cyclic operating parameters (adsorption time, evacuation time, desorption time, cooling time and vacuum pressure) and using the SCR algorithm of Zaryab et. al. (2022) at two fixed

desorption temperatures of 352.15 K and 360.15 K. The optimized results showed that there are desirable step times which can yield a higher capture rate and there is a significant reduction in the cooling time (assuming ideal cooling profiles) allowing more CO<sub>2</sub> to be captured per day, increasing its performance. In terms of the specific energy duties and CAPEX of the systems, small changes were observed before and after the optimization but larger operating costs were observed due to more cycles per day. The increase in the number of cycles per day and the higher cyclic capture rate [kg/cycle] however greatly decreased the capture cost by 33.6% and 45.5% for optimized case B and C respectively. The thermal energy duties of the considered cases agree with Climeworks' benchmark unit. The higher desorption temperature reduces electrical duties but further studies involving different bed configurations, dimensions, different sorbents and air/purge flows are required to further reduce the electrical (as well as thermal) energies of the current model. In next steps, more robust optimization studies including desorption temperature, steam flow rates and cooling cost would be needed.

Also, a location for implementing a fully electrified DAC system is important in achieving net-zero emissions since depending on the country's electricity grid systems, the same DAC system may or may not be a negative emission option. However, a detailed LCA is required to obtain more insights into the relevance and potential deployment of DAC technology. Lastly, the co-adsorption of CO<sub>2</sub> and H<sub>2</sub>O, not considered in the present study, can be included to fully assess the impacts of humidity on the performances of DAC units. Overall, this study shows a complex model with a high computational cost can benefit from SCR algorithm which is a surrogate-based derivative free optimization method. The results from the optimization also suggest the potential for DAC improvement as well as its promising potential as a negative emission technology.

## Acknowledgement

The authors are grateful to the Belgian Federal Public Service Economy and to the Belgian Energy Transition Fund which support this research within the project PROCURA.

## References

- M. J. Bos, S. Pietersen and D.W.F. Brilman, 2019, *Chem. Eng. Sci.* X, 2, p.100020.  
 T. Bruckner et al., 2014. In *Clim. Change*, (pp. 1329-1356). Cambridge University Press.  
 Climeworks, 2017, Climeworks makes history with world's first commercial DAC plant  
 S. Deutz and A. Bardow, 2021, *Nat. Energy* 6, 203–213.  
 S. Fujikawa, R. Selyanchyn and T. Kunitake, 2021, *Polym. J.*, 53(1), pp.111-119.  
 Global change data lab, 2022, Published online at OurWorldInData.org.  
 R. Hanna, A. Abdulla, Y. Xu and D.G. Victor, 2021, *Nat. Commun.*, 12(1), 1-13.  
 IEA, 2022, Direct Air Capture, IEA, Paris <https://www.iea.org/reports/direct-air-capture>  
 D.W. Keith, G. Holmes, D. St.Angelo, K. Heidel, Joule, 2018. 2(8): p. 1573-1594.  
 S. Kim and G. Léonard, 2022a, *Comput. Aided Chem. Eng.* (Vol. 51, pp. 265-270). Elsevier.  
 S. Kim and G. Léonard, 2022b, SSRN, GHGT-16.  
 K. Lackner, H. J. Ziock and P. Grimes, 1999, (No. LA-UR-99-583), LANL  
 N. McQueen et al., 2021, *Prog. Energy*, 3(3), p.032001.  
 G. Realmonte et al., 2019, *Nat. Commun.*, 10.  
 J. Rogelj et al., 2015, *Nat. Clim. Change*, 5(6), pp.519-527.  
 H.M. Schellevis et al., 2021, *Int. J. Greenh. Gas Control*, 110, p.103431.  
 Q. Yu and D.W.F. Brilman, 2017, *Energy Procedia*, 114, pp.6102-6114.  
 R. Veneman et al., 2015, *Int. J. Greenh. Gas Control*, 41, 268–275  
 S. Voskian and T.A. Hatton, 2019, *Energy Environ. Science*. 12 3530–47  
 S. Zaryab, A. Manno, E. Martelli, 2022, *Comput. Aided Chem. Eng.* 51, 1213-1218, Elsevier.  
 Z. Zolfaghari, A. Aslani, A. Moshari, and M. Malekli, 2022, *Int. J. Energy. Res.*, 46(1), 383-396.

# Simulation of biodiesel production from algae

Suresh Kumar Jayaraman,

*AVEVA Group Plc, Lake Forest, California 92630, United States of America*

## Abstract

Increasing oil demand and depletion of fossil fuel acts as a motivator for alternate fuel research. Biofuels, derived from renewable resources, have the potential to replace their petroleum counterparts, providing a domestic, and carbon-neutral fuel. This study analyzes a large-scale alga dewatering technique and production of biodiesel from algae oil. The dewatering stage extracts algae from water with liquefied propane or butane, which is part of a refrigeration process, is found to be cost effective compared to other dewatering methods. Dry algae are processed into algal oil, and the oil is then trans esterified to Fatty Acid Methyl Ester (FAME), a common biodiesel. The glycerol byproduct is converted to methanol, a more profitable product. This article involves developing a detailed steady state simulation of production of biodiesel from algal oil using first principles unit-op models and rigorous thermodynamics on a next generation commercial simulation tool, AVEVA Process Simulation.

**Keywords:** dewatering, algae oil, biodiesel, process simulation

## 1 Introduction

Biofuels, derived from renewable resources, have the potential to replace their petroleum counterparts, providing a domestic, and carbon-neutral fuel. Common sources of biofuel include sunflower, safflower, canola, soy, and corn. In order to address the issue of food crops being used for fuel production, researchers have turned their focus to non-food resources like algae<sup>1</sup>. Algae are ubiquitous photosynthetic organisms, ranging from single-celled forms to complex multicellular colonies<sup>2</sup>. Algae require sunlight, carbon dioxide, and nutrients like nitrates and phosphates to grow.

The major four processing steps in the production of biofuels are cultivation, Dewatering, Algae extraction, and Biodiesel production. Nutrients are the most cost-involving requirements of algae growth, so optimizing the nutrients requirement can reduce the associated production cost<sup>3</sup>. Out of various algae species, members of the marine green algae such as *Nannochloropsis* are the best species for biodiesel production, because of high lipid content and biomass productivity<sup>4</sup>. The fully grown concentration of algae in an aqueous culture medium is usually 0.1 w/w%, and this concentration is enhanced up to 20 w/w% through a traditional method like drying using heat or electricity<sup>5</sup>. Given the high latent heat of water, algae dewatering by evaporation is an energy intense process<sup>6</sup>, alternative approaches are necessary to improve the energy efficiency of algae dewatering process. Mechanical disruption techniques and hexane extraction are often used to extract algae oil from algae biomass<sup>7</sup>. The algae oil is mostly composed of triglycerides, which is converted

to biodiesel by the transesterification reaction<sup>8</sup>. This paper proposes a novel dewatering technique using propane as solvent, as part of a plant-wide refrigeration system.

## 2 Process Simulation

Process simulations are used to assess the commercial feasibilities of the proposed processes. The Next generation process simulation software, AVEVA™ Process Simulation developed by AVEVA, is used in this work. Microalgae consists of carbohydrates, proteins, and lipids. Here, algae are assumed to be of 50 % lipids, 25 % proteins, 25 % carbohydrates. Glucose is used to model carbohydrates, and poly lysine is used to model protein, and both are simulated as solid components. For simplicity just triolein ( $C_{57}H_{104}O_6$ ) is used to represent lipid content of the algae in the simulation of dewatering process<sup>2</sup>. Triolein (glycerol esterified with a free fatty acid called oleic acid commonly found in various oil sources) is used to represent triglycerides and Methyl Oleate (MOLEATE) is used to represent FAME. All the components are added from DIPPR (Design Institute for Physical Properties) data bank. Non-Random Two-Liquid (NRTL) and Ideal Gas Law thermodynamic properties are used in this simulation. Missing thermodynamic properties are entered manually in the fluid.

## 3 Algae-dewatering process:

Algae-dewatering has been identified as one of the bottlenecks limiting the overall energy efficiency of algal biofuel technology<sup>9</sup>, which currently inhibits the large-scale processing of algae into useable products. Centrifuges are usually used to concentrate the algae and remove lipids, but this processing step is energy intensive and is currently an area of significant research. Algae biomass production rates can vary based on the species chosen, cultivation technique, nutrients, and pond conditions. It is suggested that algal biomass production of  $30 \text{ g m}^{-2} \text{ d}^{-1}$  ( $109.6 \text{ t ha}^{-1} \text{ y}^{-1}$ ) is possible<sup>10</sup>. The cultivation land of 1,500 acres, wet algae flowrate of 760,000 Kg/hr, and 1 w/w % of algae concentration is assumed in the simulation. Figure 1 shows the algae-dewatering process using liquid propane to reduce the moisture content from 99 % to 0.02%. This process stands out because a chemical manufacturer (such as fertilizer plant) generates waste (such as N,  $\text{CO}_2$  and heat) and needs refrigeration. The proposed propane extraction idea uses the dewatering energy to supplement the plant refrigeration need and integrating the algae production with the plant converts waste to needed raw material. The condenser is used in the process to cool the compressed propane gas to liquid propane. Since propane and water are immiscible, they can be separated by decanting. Then propane is separated from algae biomass using flash separators

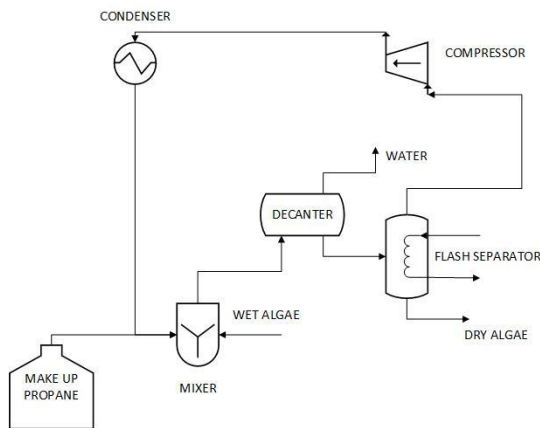


Figure 1: Algae dewatering process

and recycled to the fresh feed. The water from the decanter can be recycled back to the pond, thereby minimizing the wastage of water. In this process the water is removed, not by heating or drying but by a much less energy intensive process. The dried algae from this process are sent to an extraction unit to extract algae oil. The propane is evaporated in the flash separator, leaving the algal paste to be removed mechanically. The gaseous propane is compressed, constituting a refrigeration cycle, thereby supplementing an energy need for an adjacent plant. Butane can be used as an alternative solvent choice in the dewatering process. Though Coefficient of performance (COP) of butane is better than propane, the process economics suggest that propane dewatering technique would be a cost-effective way to dewater algae. However, the interaction of algae with propane, and ways to remove algal paste from the separator remain unknown.

#### 4 Algae Oil Extraction Process:

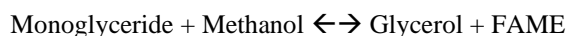
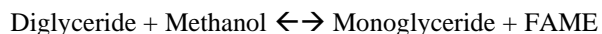
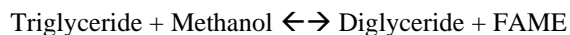
After dewatering, the algae biomass undergoes a pre-treatment process to enhance the efficiency of lipid extraction process. The pre-treatment is either performed in a single step or multiple steps. If pre-treatment is not performed, the algae biomass concentrate is directly processed for lipid extraction. The algae biomass concentrate is either exposed to a cell disruption technique or dried and then milled into fine powders<sup>11</sup>. The pre-treated algae (disrupted concentrate or dried powder) is then exposed to an eluting extraction solvent which extracts the lipids out of the cellular matrixes<sup>11</sup>. Due to limited data on oil extraction process, a simple conversion model is used in this simulation study to convert 50% of the dry algae to algae oil/lipid. Under nutrient stress during the cultivation stage, the increase in total lipid primarily consists of triglycerides<sup>12 13</sup>.

#### 5 Biodiesel Production Process:

Component information and thermodynamic models are set up and the process steady-state simulation is simulated in a next generation simulation tool, AVEVA Process Simulation. Process flowsheet of production of biodiesel is shown in Figure 2. The oil content of algae is assumed to be 50% in the simulation. The Biodiesel production process in this work consists of following sections:

##### 5.1 Transesterification reactions

The most common way to produce biodiesel is by stepwise transesterification, which is a catalyzed chemical reaction between algae oil and an alcohol to yield fatty acid alkyl esters (biodiesel) and glycerol as shown below.



When triglycerides react with alcohol (e.g., methanol), the three fatty acid chains are released from the glycerol skeleton and combine with the alcohol to yield fatty acid alkyl esters<sup>10</sup> and glycerol as a by-product. Methanol is the most commonly used alcohol and Transesterification



reactions can be alkali-catalyzed, acid-catalyzed, or enzyme-catalyzed<sup>10</sup>. Based on several studies in the literature<sup>10, 14</sup>, the transesterification reaction is carried out at a temperature about 60°C and a 6:1 molar ratio of methanol to algae oil is fed to the Plugflow reactor, which helps in shifting the equilibrium to the product side. The reaction submodel is developed to define the reaction kinetics of the transesterification reactions<sup>15</sup>. Fresh make up methanol, recycled methanol, and sodium hydroxide with a concentration of 1.0 wt% of the algae oil are mixed prior to being pumped into stoichiometric reactor by a pump. For the study, 97% conversion of triglycerides in transesterification reaction is assumed in the reactor.

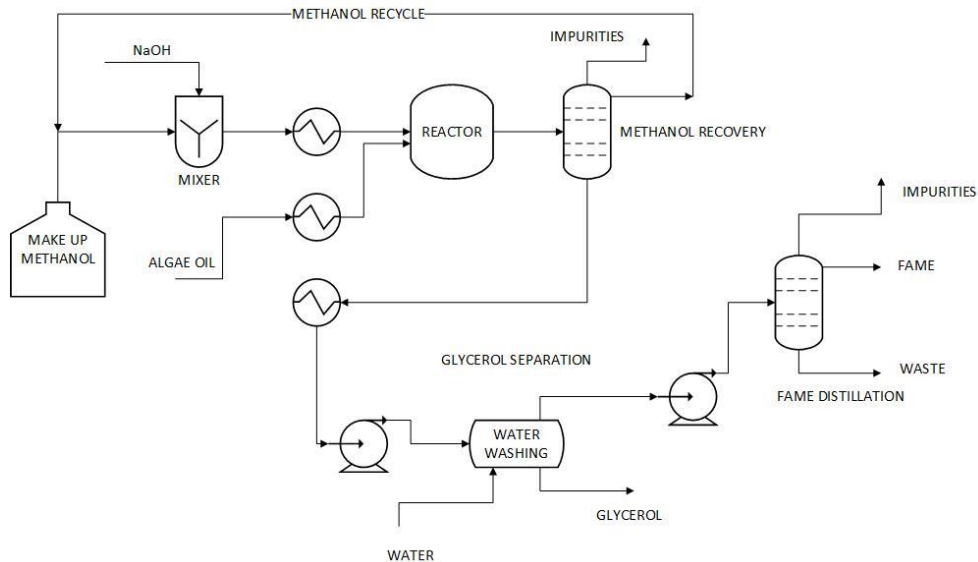


Figure 2: Biodiesel production process

### 5.2 Methanol recovery

The distillation column used for MeOH recovery has four theoretical stages, a reflux ratio of 2, total condenser, and kettle reboiler for a good separation of methanol from other components. Pressure in the range of 20 kPa is used to maintain the temperature of bottom stream under 150°C. Methanol is collected in the distillate stream with a purity of 96%. Methanol from distillate stream is recycled to the fresh make up methanol and then fed to the stoichiometric reactor. Bottom stream is sent to decanter to separate glycerol and FAME. Carbon-dioxide and oxygen are removed as impurities in this section.

### 5.3 FAME and glycerol separation

FAME is separated from glycerol, methanol and catalyst using a water washing column in this study. The FAME is separated from glycerol, methanol, and catalyst by adding 110 kg/hr water. The unreacted triglycerides, methanol and water in the top stream are less than 6 % and glycerol is collected in the bottom stream. Feed water amount is determined by performing the sensitivity analysis to achieve complete separation between FAME and glycerol.

#### 5.4 FAME purification

Biodiesel purity of 99.65% is required to meet ASTM D 6751 of biodiesel specification <sup>22a</sup>. To obtain a final biodiesel product adhering to specifications, FAME distillation column with four theoretical stages, a reflux ratio of 1, total condenser, and kettle reboiler is used. Distillation column is operated under vacuum to keep temperatures low to avoid degradation of the FAME. Methanol, water, and unreacted oil are collected in the column overhead using a total condenser. FAME is obtained as distillate with the purity of 99.8 %.

#### 5.5 Methanol production from glycerol

For every 9 kg of biodiesel produced, about 1 kg of a crude glycerol is formed as a by-product. The crude natural glycerol could be converted to a useful chemical, to increase the profitability of the biodiesel production plant. Glycerol can be converted to methanol, which can be used as fresh make up solvent to the transesterification reaction. Glycerol is converted to methanol and CO<sub>2</sub> as shown below



Glycerol from the water washing column is sent to the conversion reactor along with water, and 99% of glycerol is assumed to be converted to methanol and CO<sub>2</sub>. The reaction takes place at 470 °C and 5 MPa to produce methanol and CO<sub>2</sub>. Separation of methanol

and CO<sub>2</sub> is done in a distillation column with 5 theoretical stages and a reflux ratio of 1.5. Pure CO<sub>2</sub> is collected as a distillate, containing 99% of the total CO<sub>2</sub> and pure methanol is collected as a bottom product, containing 95% of the total methanol. Combining this process along with the biodiesel

production, methanol can be used as fresh make up solvent for the transesterification reaction and CO<sub>2</sub> can be used as a nutrient for the growth of algae.

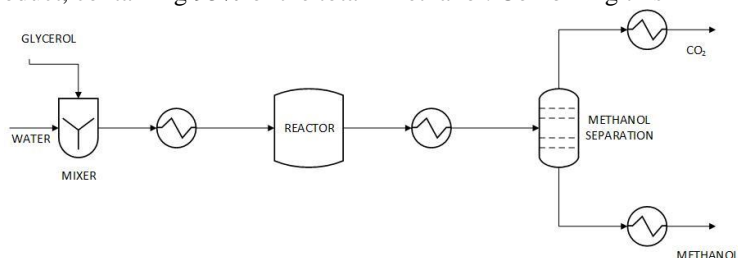


Figure 3: Methanol production process

## 6 Conclusion:

In this work, a new dewatering technique using propane is discussed and simulated using AVEVA Process Simulation. This work simulated a biodiesel production plant using the simulation tool to produce 33000 tons of biodiesel per year, with algae species of 50% oil content. The results show that the production unit is profitable only with 50% and 70% oil content, suggesting that the choice of algae species plays a vital role in the production of biodiesel. Algae species with 70% oil content is not achievable, so this work concludes that

the production plant is profitable with algae species of 50% oil content. So future research should focus on cultivation and pond construction to make this biodiesel production more profitable. Since the production of glycerol exceeds the world demand, the byproduct glycerol is converted to methanol and CO<sub>2</sub>, which can be used as make up solvents and nutrients for algae respectively. However, the present study shows that the biodiesel can be produced at a reasonable cost, thereby making algae a potential source for biofuels.

## References:

1. Chisti, Y., Biodiesel from microalgae. *Biotechnology Advances* **2007**, *25* (3), 294-306.
2. Brennan, L.; Owende, P., Biofuels from microalgae—A review of technologies for production, processing, and extractions of biofuels and co-products. *Renewable and Sustainable Energy Reviews* **2010**, *14* (2), 557-577.
3. Blair, M. F.; Kokabian, B.; Gude, V. G., Light and growth medium effect on *Chlorella vulgaris* biomass production. *Journal of Environmental Chemical Engineering* **2014**, *2* (1), 665-674.
4. Sani, Y. M.; Daud, W. M. A. W.; Abdul Aziz, A. R., Solid acid-catalyzed biodiesel production from microalgal oil—The dual advantage. *Journal of Environmental Chemical Engineering* **2013**, *1* (3), 113-121.
5. Lardon, L.; Hélias, A.; Sialve, B.; Steyer, J.-P.; Bernard, O., Life-Cycle Assessment of Biodiesel Production from Microalgae. *Environmental Science & Technology* **2009**, *43* (17), 6475-6481.
6. Shao, P.; Darcovich, K.; McCracken, T.; Ordorica-Garcia, G.; Reith, M.; O'Leary, S., Algae-dewatering using rotary drum vacuum filters: Process modeling, simulation and techno-economics. *Chemical Engineering Journal* **2015**, *268* (0), 67-75.
7. Geciova, J.; Bury, D.; Jelen, P., Methods for disruption of microbial cells for potential use in the dairy industry—a review. *International Dairy Journal* **2002**, *12* (6), 541-553.
8. Vyas, A. P.; Verma, J. L.; Subrahmanyam, N., A review on FAME production processes. *Fuel* **2010**, *89* (1), 1-9.
9. Rawat, I.; Ranjith Kumar, R.; Mutanda, T.; Bux, F., Biodiesel from microalgae: A critical evaluation from laboratory to large scale production. *Applied Energy* **2013**, *103* (0), 444-467.
10. Benemann, J. R.; Oswald, W. J. *Systems and economic analysis of microalgae ponds for conversion of CO<sub>2</sub> to biomass. Final report.*
11. Silva, C.; Soliman, E.; Cameron, G.; Fabiano, L. A.; Seider, W. D.; Dunlop, E. H.; Coaldrake, A. K., Commercial-Scale Biodiesel Production from Algae. *Industrial & Engineering Chemistry Research* **2014**, *53* (13), 5311-5324.
12. Halim, R.; Danquah, M. K.; Webley, P. A., Extraction of oil from microalgae for biodiesel production: A review. *Biotechnology Advances* **2012**, *30* (3), 709-732.
13. Hu, Q.; Sommerfeld, M.; Jarvis, E.; Ghirardi, M.; Posewitz, M.; Seibert, M.; Darzins, A., Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *The Plant Journal* **2008**, *54* (4), 621-639.
14. (a) Pankaj Kumar, M. R. S., Kiran Toppo, Physico-Chemical Characterization of Algal oil: a Potential Biofuel. *Asian Journal of Experimental Biology* **2011**, *2* (3); (b) Demirbas, A., Progress and recent trends in biodiesel fuels. *Energy Conversion and Management* **2009**, *50* (1), 14-34.
15. Integrated Process Modeling and Product Design of Biodiesel Manufacturing, Ai-Fu Chang and Y. A. Liu, *Industrial & Engineering Chemistry Research* **2010** *49* (3), 1197-1213

# Different methods of model-based process optimization presented on real world examples from water industry

Dr. Ewa Bozek<sup>a</sup>, Dr. Corinna Busse<sup>a</sup>, Delia Gietmann<sup>a</sup>, Dr. Bernd-Markus Pfeiffer<sup>a</sup>

<sup>a</sup>*Siemens AG, Siemenspromenade 3, Erlangen 91058, Germany*  
*ewa.bozek@siemens.com*

## Abstract

With the digitalization of industry and further development of computational methods, different ways to control and optimize production processes are possible. Depending on which automation level (i.e. process, field or control level) is chosen for optimization, a variety of operator decision support systems exists. Using the example of water and wastewater industry this contribution explores existing optimization methods: high-fidelity first-principle, data-driven and linear dynamic black-box models. All described solutions can connect online to distributed control systems, but use case, engineering and deployment differ significantly for each of these applications.

**Keywords:** process simulation, optimization, model-predictive control

## 1. Introduction

Saving resources, securing sustainable supply, lowering maintenance and increasing market volatility are some of many challenges that the water industry is facing. A variety of model-based methods for process optimization exists to advice on optimal operating setpoints in an increasingly complex environment while meeting internal and external constraints. The following methods are described below in more details:

- Process level: high-fidelity first-principles models, optimization problem can be formulated in an arbitrary way, additional value provided by e.g. soft sensors, steady state or dynamic optimization for a specified time horizon (gPROMS & gDAP example: reverse osmosis desalination plant)
- Field level: Data-driven models, optimization problem is predefined: minimize switching operations or energy costs for a time horizon of 24 h. (SIWA Optim example: drinking water supply network)
- Control level: linear dynamic black-box models, dynamic optimization problem is essentially to minimize future deviations from specified setpoints in a specified prediction horizon for any multivariable feedback control problem (PCS 7 embedded MPC (Model Predictive Control) Example: wastewater treatment plant).

## 2. Real-time optimization based on first principles model

First principles models rely on fundamental laws of physics, chemistry and biology. The knowledge captured in these models is used since a long time for offline simulation and optimization. Using such models for online applications is a fairly new use case.

### 2.1. General Application Use Case

The provision of desalinated water by reverse osmosis is an energy intensive task. Due to the high osmotic pressure of seawater, significant energy is required to generate water of

a certain quality (salt content). The energy consumption of the process is influenced by various decisions operators take based on process experience. To support operators' decisions, real-time optimization (RTO) based on first principles models is explored. In RTO, the optimum values of setpoints for underlying controllers are re-calculated periodically (e.g., every hour or every day). These repetitive calculations involve solving a constrained steady-state optimization problem. Therefore, a steady-state process model, a performance index (e.g., energy consumption) and constraints are required.

## 2.2. Process Model

The process model of the reverse osmosis section of a desalination plant was developed in gPROMS PROCESS using the library for membranes including a detailed description of the mass transfer (see [1.]). The model also includes dynamics like the gradually increasing fouling layer on the membranes. Overall, in gPROMS a differential-algebraic system of equations (DAE) is created, which can be used for a variety of use cases: simulation, global system analysis, soft-sensing and optimization.

### 2.2.1. Optimization problem

The optimization problem was configured within gPROMS PROCESS itself. A major advantage of a first principles model in the context of optimization is the flexibility to define an optimization problem: any calculated variable can be chosen as objective function, e.g., summation of all energy consumers. Similarly, all assigned variables can be decision variables and constraints can be specified according to process requirements. gPROMS also allows the use of discrete decision variables leading to mixed-integer optimization problems. Finally, constraints like product quality can be configured based on all calculated variables. The time horizon can be set-up as steady-state or dynamic.

### 2.2.2. Engineering Workflow

The engineering workflow of an RTO application from a first principles model follows four steps: 1. Development of a model based on design data like tank size, pump curves, membrane specifics etc. 2. Model validation based on historical process data to determine unknown parameters like mass transfer coefficients or fouling functions. 3. The calibrated model are then used to determine optimal operating points offline. 4. If the robustness of the model is proven, the optimization task can be transferred into an online application.

### 2.2.3. Deployment for Online Operation

For the online deployment the gPROMS digital application platform gDAP is used. The core component is the computational module required for the optimization based on the offline optimization task. Additionally, online connections are configured like e.g., OPC-UA, SQL Server, Historian. Measurement data coming from the real world like flowrates, pressure, quality measurements are validated before using in the computational module. Optimization results are then validated as well before sending them back to the real world.

### 2.2.4. Example from Water Industry

A high-fidelity process model including a detailed model of the membranes used for the separation of retentate (high salt loading) and permeate (lower salt content) with a detailed description of the mass transfer incl. fouling function depending on the load is developed in gPROMS PROCESS [1.]. This model was used for offline optimization studies to determine the optimization potential in typical operation conditions, as described by [1.]. As objective function, the energy consumption of a defined time horizon is chosen. As decision variables the following assigned variables are chosen: split fraction between 2<sup>nd</sup> stage and product tank, setpoints for the flowrate of the booster pumps in the 1<sup>st</sup> stage and setpoints for the flowrate of the high-pressure pump. As constraints, salt content in the permeate and overall production capacity are considered. Based on the offline

optimization study, a significant potential for energy savings is determined. The developed set up is ready for online deployment with the dedicated framework gDAP.

### **3. Demand driven Optimization**

The economic operation of pumps is one of the biggest challenges in the operation of water networks due to different energy tariffs and the increasing number of renewable energies. Since pumps are the largest energy consumer in the network they should be operated in a smooth and efficient way. This is a special use case of RTO, that can be solved by a Siemens software SIWA Optim dedicated to this optimization task.

#### *3.1. General Application Use Case*

The goal of RTO for drinking water networks is calculating a schedule for the most efficient pump operation, ensuring water supply at all times, taking all important parameters such as tank level, energy prices, consumer demands or pump availabilities into consideration. SIWA Optim assists in daily operation by continuously calculating schedules in individually set time intervals or supporting maintenance and operation planning with manually started what-if scenarios and calculations [2.].

#### *3.2. Process Model*

The water network is modeled in a flowsheet with realistic tank sizes and linearized pump models describing the relationship between transported volumes and required energy in the given environment. Model parameters are derived from process data. The simulation uses a linear, flow-based model (in contrast to a hydraulic model that consider pressure calculation). Besides filling of tanks, no other dynamic effects are described in the model.

##### *3.2.1. Optimization problem*

The optimization problem is solved by a Mixed Integer Linear Program Solver (MILP). It can be executed in a manual or cyclic way. In a manual mode it is possible to change the parameters and study the system in more detail, mainly for operation or maintenance purpose. Cyclic calculations are running autonomously in the background in a defined time interval between 1-24 h. Pump status and tank levels are updated before every cyclic optimization. Short intervals between the optimizations can increase operation safety since unexpected events like damaged pumps or low tank levels can be recognized sooner and taken into consideration for the next pump schedule. The calculated pump schedule can also be sent back to the DCS where a dedicated function block is receiving it and transfers the flow setpoints per station or pump to the control unit of the pump group.

##### *3.2.2. Engineering Workflow*

The water network is modeled as a flow based, abstract distribution network. It includes consumer forecast demands, water suppliers feeding fixed quantities of water or providing water in a specific range. Tanks with minimum and maximum level restrictions, pumping stations with limited allowed combinatorics of pumps as well as energy suppliers which calculate the working price for the specific stations and limit power consumption are also included into the model.

This data is used to calculate an optimal schedule for the upcoming 24 hours. Data like pump status and current tank level can be imported from the control system via an OPC UA connection, which allows reading independently of control system provider. The model is fine-tuned by validating the results with the real plant data.

##### *3.2.3. Deployment*

The application is running in Linux dockers within a virtual machine hosted by Hyper-V Manager. The user interface is accessible via Google Chrome. The network engineering for the application is done in an MS Visio based engineering tool.

## 4. DCS embedded Model Predictive Control (MPC)

An optimization problem for a dynamic process model on a finite time horizon, where the performance index includes deviations of model output variables from specified setpoints (or ranges) can be classified as an MPC problem, if it is solved repeatedly in a fixed sample time and the time horizon is shifted accordingly (“moving horizon”).

### 4.1. General Application Use Case

The general use case is multivariable feedback control, namely model predictive control. In some cases, the aim is to reduce the demand for manual operator interactions and increase the degree of automation. In most cases the goal is to optimize process operation with respect to resource consumption, throughput or yield. This is achieved in two steps: 1. Reduce variances of process variables by improved feedback control. 2. This allows to push process operation towards constraints without the risk to violate them. Constraint pushing means to exploit the full potential of the process. As an example from water industry, aeration control in a waste water treatment plant will be described.

### 4.2. Process Model

All MPC algorithms exploit the IMC principle (Internal Model Control [3.]): a dynamic process model is not only used for offline controller design but constitutes an integral part of the online control algorithm, where it is used to predict future process behavior along a specified finite prediction horizon. Most industrial applications of MPC rely on linear dynamic process models that are identified from active experiments (like step testing) in the plant. The typical models are multivariable FIR (Finite Impulse Response) or FSR (Finite Step Response) models as they allow simulation of model response with minimal computing effort. Nonlinear MPC is out of the scope here but described elsewhere [4.].

#### 4.2.1. Optimization problem

The second central idea of MPC is to formulate the feedback control problem as an optimization problem: minimize the squared sum of “control deviations” and other “costs” of control like MV (Manipulated Variables) moves over a finite time horizon. The specific formulation of the objective function for optimization offers many degrees of freedom for controller design. Control deviations are differences between predicted process outputs and future setpoints. Different controlled variables (CVs) can be assigned weights in the performance index to reflect their relative importance to process operation. Move penalties for manipulated variables in the objective function restrict controller moves to be more or less aggressive. Hard constraints like MV limits and soft constraints for model output variables are considered in the optimization problem. The time horizon for optimization is user-defined and in typical applications set for a few hours. Special version of Quadratic Programming solvers are applied inside linear MPC algorithms.

#### 4.2.2. Engineering Workflow

Linear MPC models are typically not derived from first principles but from process data. Active experiments in the plant like step tests around the typical plant operating point are required to create process data with a sufficient information content on process dynamics. From such data, the linear dynamic process model is created automatically, using one of the existing MPC software packages. Simulations of process model are compared to real process data to verify model quality. Definition of optimization problem reduces to specification of MPC design parameters like priority of controlled variables and move penalties for manipulated variables. Offline simulations of closed MPC control loop with linear dynamic process model inside MPC software allow fine-tuning of parameters.

#### 4.2.3. Deployment for Online Operation

There are two deployment types of MPC in context of DCS in process industry: 1. deployment on a Windows PC interfaced to DCS via OPC. 2. DCS embedded MPC, where the online control algorithm is a function block similar to a PID controller. This is currently only available from two DCS vendors: Emerson DeltaV and Simatic PCS 7.

#### 4.2.4. Example from Water Industry

An aeration control in a wastewater treatment plant is reviewed: [5.], [6.]. Closed-loop control of the biological operations in the activated sludge tank is a big challenge for sewage treatment plant automation, because aeration is the most sensitive and energy-consuming part of the WWTP.  $O_2$ -concentration in aeration tank is controlled by PID controllers which manipulate air blower speed. However, the decisive KPIs for water cleaning performance are  $NH_4$  and  $NO_x$  outlet concentrations. Instead of leaving those KPIs to the plant operator, closed loop feedback control by MPC can be established. Manipulated variables of MPC are setpoints for existing PID controllers for recirculation rate and  $O_2$ -concentration. Fluctuations in WWTP feed (feed rate and pollution load) are a challenge for aeration control. Thus, they are included as measurable disturbance variables in the MPC concept, improving prediction quality and providing better compensation of the disturbances effects on  $NH_4$  and  $NO_x$  concentration. WWTP plant operation by MPC assures the fulfillment of the legal specifications for purified wastewater. Test runs in a large WWTP operated by Hansewasser in Germany confirm that energy savings of 3-10% compared to manual operation with constant  $O_2$ -setpoint and heuristic reflux can be achieved.

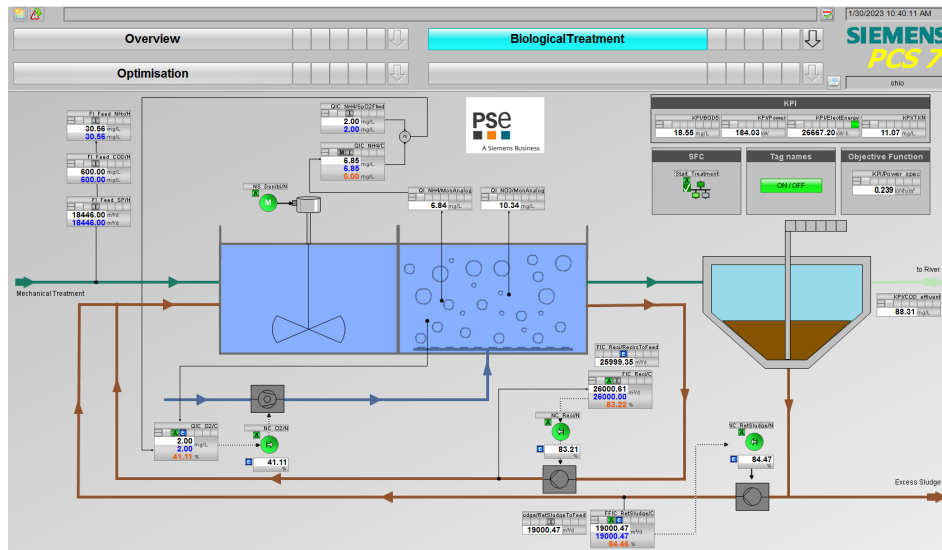


Figure 1: Simatic PCS 7 Unit template for aeration control in wastewater treatment plant [7.]

## 5. Conclusion

If a dedicated software solution for a given optimization problem (like SIWA Optim) is available in context of existing automation hardware and software architecture of a process plant, it is always recommended to exploit this offering. Otherwise, a very general approach like RTO can be applied. In some cases the term “optimization” in requirement specifications is not used in the narrow sense of mathematics. It has to be carefully



checked if there is actually a feedback control problem in the plant, or rather potential for RTO. The typical cycle time for feedback control in process plants is seconds or minutes, whereas steady state RTO is calculated daily or hourly. In general, apply feedback control if target values (or target ranges) for the most important process variables are known and can be used as setpoints for MPC. Apply RTO if the focus is not on stabilizing plant operation at specified setpoints, but rather driving the plant in a desired direction of optimization, e.g. minimize energy consumption. The decision support workflow is shown in Figure 2. It is recommended to discuss it with the experts, to consider all factors.

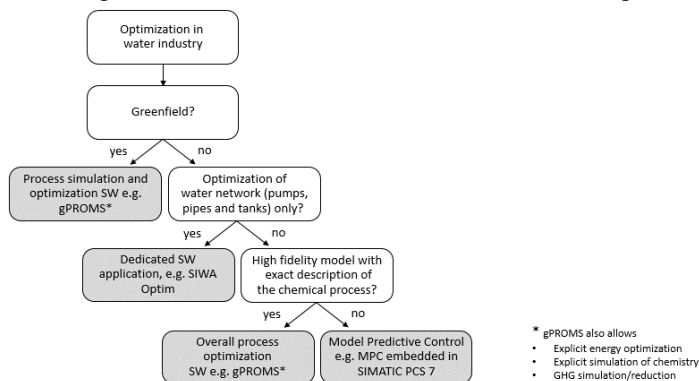


Figure 2: Decision support workflow for optimization in the water industry.

## References

- [1.] Pfeiffer, B-M., Lawal, A., Oppelt, M., Ruminski, L., Lade, M.: Plantwide Dynamic Optimization of Seawater Desalination with a Digital Process Twin. Industry Talks, Automation 2020, Baden-Baden (virtual), Jun. 2020.
- [2.] Siemens AG, Flyer: SIWA Optim – Operate your water supply in a resource-saving and efficient way, 2018. <https://assets.new.siemens.com/siemens/assets/api/uuid:c2bfa03-0f72-468c-b03e-b9d92698203c/flyer-siwa-optim-en.pdf>
- [3.] Garcia, C.E., Prett, M., Morari, M. (1989): Model predictive control: theory and practice – a survey. Automatica 25 Nr. 3, pp. 335-348.
- [4.] Pfeiffer, B-M., Oppelt, M., Leingang, Ch., Pantelides, C., Pereira, F.: Nonlinear Model Predictive Control based on Existing Mechanistic Models of Polymerisation Reactors. IFAC World Congress 2020, Berlin (virtual), Jul. 2020.
- [5.] Oppermann, J., Radtke, N., Haber-Quebe, J., Pirsing, A., Pfeiffer, B-M., Labisch, D.: Einsatz von Advanced Process Control auf kommunalen Kläranlagen. 12. Aachener Tagung Wassertechnologie, Okt. 2017, Aachen.
- [6.] Siemens AG, Whitepaper: Optimization of Sewage Treatment Plants by Advanced Process Control, March 2015. <https://silo.tips/download/optimization-of-sewage-treatment-plants-by-advanced-process-control>
- [7.] Simatic PCS 7 Water Unit Template - Control of Biological Stage of a Wastewater Treatment Plant with upstream Denitrification. <https://support.industry.siemens.com/cs/de/en/view/109485916>

# A Reduced Order Model for the Prediction of the Dynamics of a Methane Reactor

Enrico A. Cutillo,<sup>a\*</sup> Erasmo Mancusi,<sup>a</sup> Katarzyna Bizon,<sup>b</sup> Piero Bareschino,<sup>a,c</sup>  
Gaetano Continillo,<sup>a</sup>

<sup>a</sup>*Dipartimento di Ingegneria, Università degli Studi del Sannio, Benevento 82100, Italy*

<sup>b</sup>*Faculty of Chemical Engineering and Technology, Cracow University of Technology,  
ul. Warszawska 24, 31-155 Kraków, Poland*

<sup>c</sup>*Istituto di Scienze e Tecnologie per l'Energia e la Mobilità Sostenibili - Consiglio  
Nazionale delle Ricerche, Italy*

\*[cutillo@unisannio.it](mailto:cutillo@unisannio.it)

## Abstract

The present work studies fixed bed methanation reactor for the upgrading of biogas, which is expected to provide chemical energy storage for non-programmable sources. The dynamics of such systems can exhibit complex behaviors, that could make the numerical solution computationally expensive. In this work the influence of both the inlet temperature and the recycle ratio on the reactor dynamics are studied employing an empirically based reduced-order model allowing a significant reduction of computational time. Moreover, an assessment of the accuracy of the proposed reduced-order model enables its use within real-time control applications, or in speeding up the resolution of time-consuming optimization.

**Keywords:** reduced models, POD, methanation, step response, Power-to-Methane.

## 1. Introduction

Numerical simulation of high-fidelity models of complex systems can take hours or even days. Some applications, such as parametric studies or optimization, often require thousands of those simulations, and for this reason, overall CPU time rapidly increases. Anyhow, simulations enable to optimize process parameters, to reduce costs, or maximize the productivity of the studied process (Bizon and Continillo, 2021) without the need for time-consuming and expensive experimental work but, most importantly, permit to predict unwanted behaviors and adopt adequate control strategies (Bareschino et al., 2022). In this work, the studied process is related to sustainable energy conversion and reduction of CO<sub>2</sub> emissions, which are challenges that cannot be procrastinated any further, in line with the core principles of the European Green Deal and Power-to-Gas concept. Particularly, the production of biofuels as green energy resources is very promising (Mancusi et al., 2021). Among biofuels, biogas from the anaerobic digestion of organic waste stands out as an attractive way to reduce landfilling and to produce energy. However, the CH<sub>4</sub> content in biogas usually never reaches values above 70%, so it needs to be purified (removal of trace components) and improved (removal of CO<sub>2</sub>) before it can be used. One of the most widely studied upgrading processes involves the methanation of CO<sub>2</sub> (Bareschino et al., 2021). Recently, modeling studies on the methanation process have addressed specific aspects such as the dynamical characterization of methanation reactors (Zimmermann et al., 2022). In this work, the

simulations of a dynamical system that describes the methanation of biogas in an adiabatic fixed-bed reactor with a nickel-based catalyst and a recycle loop are conducted by applying classical proper orthogonal decomposition (POD) for the model-order reduction. This methodology enables the use of the previously derived robust numerical model (Bareschino et al., 2022) for more computationally intensive applications. POD is a popular model reduction technique that, combined with Gal rkin projection, is widely applied in the field of chemical engineering and many other fields, to alleviate the computational expense required for the solution of very high-dimensional systems (Sahyoun and Djouadi, 2013). Both methanation (Bremer et al., 2017) and fixed-bed reactors with recycle have been simulated successfully with this methodology (Bizon and Continillo, 2012).

To understand if the proposed POD/Gal rkin-based reduced-order model (ROM) can correctly predict the dynamics of the methanation reactor, in this work a step response analysis is performed in a region of the space of the parameters where periodic solutions are encountered, and in its vicinity. Temperature oscillations inside the reactor must generally be avoided, as they can lead to permanent deactivation of the catalyst. Outlet concentration variations are also generally inconvenient. The ROM developed here can be useful if incorporated into the design of a control system.

## 2. Mathematical model and computational methods

### 2.1. Mathematical model of an adiabatic fixed bed reactor for biogas upgrading

The methanation reactor investigated is a fixed-bed adiabatic reactor with mass recycle. The catalyst considered is a nickel-based catalyst and the kinetic model employed is the one proposed by Xu and Froment (1989). The chosen model is one-dimensional pseudo-homogeneous with axial dispersion (R nsch et al., 2016). Gas recycle is necessary to keep the temperature inside the reactor within an acceptable limit, determined by catalyst deactivation and carbon formation. The model is mathematically formalized in species and heat balances, Eq.(1) and (2) respectively.

$$\varepsilon_g \frac{\partial c_i}{\partial t} = D_{ax} \frac{\partial^2 c_i}{\partial z^2} - u_{sg} \frac{\partial c_i}{\partial z} - (1 - \varepsilon_g) \rho_c r_i \quad (1)$$

$$\left( \varepsilon_g \rho_g c_{pg} + (1 - \varepsilon_g) c_{pc} \right) \frac{\partial T}{\partial t} = \lambda_{eff} \frac{\partial^2 T}{\partial z^2} - u_{sg} \rho_g c_{pg} \frac{\partial T}{\partial z} - \rho_c \sum_{j=1}^3 \Delta H_{R_j} R_j \quad (2)$$

where  $z$  is the axial position along the reactor,  $\varepsilon_g$  represents the bed porosity,  $c$  ( $\text{kmol} \cdot \text{m}^{-3}$ ) the gas concentration,  $\rho_c$  ( $\text{kg} \cdot \text{m}^{-3}$ ) the packed-bed density, and  $r_i$  ( $\text{kmol} \cdot \text{kgcat}^{-1} \cdot \text{s}^{-1}$ ) the rate of consumption or formation of the  $i^{\text{th}}$  species ( $i = \text{CH}_4, \text{CO}, \text{CO}_2, \text{H}_2, \text{H}_2\text{O}$ ); the latter is determined by summing up the reaction rates of that species in all the reactions,  $R_j$  (Xu and Froment, 1989) according to the stoichiometric coefficient, as follows:

$$r_i = \sum_{j=1}^3 \nu_{i,j} R_j . \quad (3)$$

The boundary conditions are:

$$c_i(0, t) = c_{i,in}(t); \quad T(0, t) = T_{in} \quad (4)$$

where

$$c_{i,in}(0, t) = \frac{c_{i,0}}{1 + R} + R \frac{c_{i,out}(t)}{1 + R} \quad (5)$$

In Eq.(5)  $R$  represents the recycle ratio. To solve the above system of partial differential equations, the method of lines was initially employed: spatial derivatives were approximated using finite differences over a uniform grid of 200 discrete nodes, yielding a system of 1200 ordinary differential equations (ODE) that was solved with MATLAB ode15s, a solver for stiff ODEs. To simplify the description of the reduction method, and to efficiently code the system of ODEs in MATLAB, the following matrix notation is employed:

$$\frac{d}{dt}\mathbf{c}_i = \mathbf{A}_c \mathbf{c}_i + \mathbf{B}_c \mathbf{c}_i \odot \mathbf{u} + \mathbf{F}_c(\mathbf{c}_i, \mathbf{T}) \quad (6)$$

$$\frac{d}{dt}\mathbf{T} = (\mathbf{A}_T \mathbf{T} + \mathbf{B}_T \mathbf{T} \odot \mathbf{u}) \odot \frac{1}{\varepsilon_g \rho_g(\mathbf{c}_i, \mathbf{T}) c_{pg} + \rho_c c_{pc} (1 - \varepsilon_g)} + \mathbf{F}_T(\mathbf{c}_i, \mathbf{T}) \quad (7)$$

where  $\mathbf{A}$  and  $\mathbf{B}$  are the coefficient matrices discretizing spatial derivatives which include also the terms resulting from the outlet boundary conditions, whereas the nonlinear functions  $\mathbf{F}_c$  and  $\mathbf{F}_T$  incorporate the inlet boundary conditions (Eq. (4) and Eq. (5)). The discretized system above is hereafter referred to as the full-order model (FOM).

### 2.2. Reduced order model formulation.

The POD method allows to extract a set of empirical basis functions from observations originating from experiments or numerical simulation (Holmes et al. 1996). The observations, often referred to as snapshots, are collected in the snapshots matrix  $\mathbf{Y}$ . The choice of the most relevant snapshots is often driven by the knowledge of the spatiotemporal complexity of the system and its variability when changing key parameters of the model. In this work, the most straightforward strategy (uniform sampling) is employed (Bizon et al., 2012). After collecting the snapshots, the POD basis is determined by solving the eigenvalue problem:

$$\mathbf{C}\Phi = \Lambda\Phi \quad (8)$$

In Eq. (8)  $\mathbf{C}$  is the autocorrelation matrix defined as  $\mathbf{Y}\mathbf{Y}^T/M$ , where  $M$  is the number of observations used for the construction of the basis, while  $\Phi$  and  $\Lambda$  are, respectively, the matrix containing basis functions and the relevant eigenvalues. The magnitude of each eigenvalue indicates the contribution of the corresponding POD mode to the ‘‘information’’ content of the original data. Using the POD basis, the generic state variable vector  $\mathbf{y}$  can be expressed in truncated form as:

$$\mathbf{y} \approx \tilde{\mathbf{y}} = \sum_{k=1}^K \phi_k x_k = \Phi_K \mathbf{x}_K \quad (9)$$

where  $K \leq N$  is the dimension of the reduced model, and  $\Phi_K \in \mathbf{R}^{N \times K}$  is a matrix composed of the first  $K$  columns of matrix  $\Phi$ . In Eq. (9), the unknowns are the  $K$  time-dependent coefficients  $\mathbf{x}_K$ . They can be determined by resolving the following ROM obtained after substitution of Eq. (9) into Eq. (6)-(7), followed by Gal erkin projection onto  $\Phi_K$ :

$$\frac{d}{dt}\mathbf{x}_K = \Phi_K^T \mathbf{A} \Phi_K \mathbf{x}_K + \Phi_K^T \mathbf{B} \Phi_K \mathbf{x}_K \odot \mathbf{u} + \Phi_K^T \mathbf{F}(\Phi_K \mathbf{x}_K) \quad (10)$$

The state variables vector has been divided into six sets of variables, associated with the species concentrations and temperature. Each one of these sets has been reduced with its own POD basis.

## 3. Results and discussion

The values of the main operating conditions and parameters characterizing the reactor geometry used in both FOM and ROM simulations are reported in Table 1. Fig. 1 reports

a schematic of the system (a) and the grid of parameters explored (b). Before starting the construction and validation of the reduced-order model, several simulations of FOM were performed for the different values of the three parameters analyzed in this work, namely:  $T_{in}$  – the inlet temperature assumed to calculate the steady state,  $\Delta T_{in}$  – the amplitude of the step forcing, and  $R$  – the value of the recycle ratio used for both steady-state and step response calculations.

Table 1. Main model parameters

Parameter	Value	Parameter	Value
$L_r$ [m]	7	$d_r$ [m]	1.2
$P$ [atm]	20	$\rho_c$ [ $\text{kg}\cdot\text{m}^{-3}$ ]	2350
$c_c$ [ $\text{J}\cdot\text{kg}^{-1}\cdot\text{K}^{-1}$ ]	1107	$y_{\text{CH}_4,in}$	0.271
$Q_{in}$ [ $\text{Nm}^3\cdot\text{s}^{-1}$ ]	1	$y_{\text{CO}_2,in}$	0.146
$D_{ax}$ [ $\text{m}^2\cdot\text{s}^{-1}$ ]	$10^{-2}$	$y_{\text{H}_2,in}$	0.583

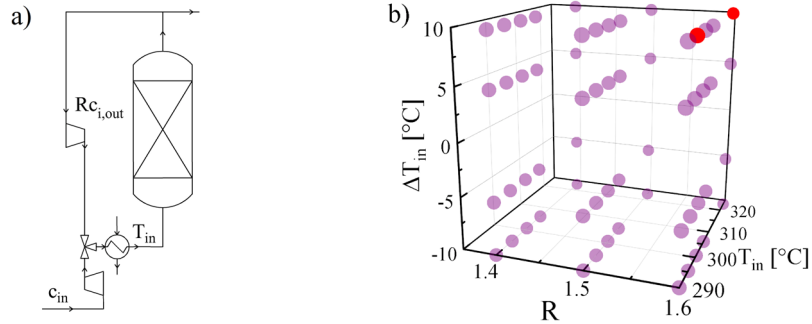


Figure 1. Scheme of methanation reactor (a) and values of the parameters used in simulations performed with the FOM (b). The two darker dots correspond to the set of parameter values employed in the construction of the ROM.

The simulations were conducted in two steps: first, the steady state (ss) solution is determined for different values of  $T_{in}$  and  $R$ ; then, a step change  $\Delta T_{in}$  of the inlet temperature is introduced, and the response of the reactor is simulated. Different values of  $\Delta T_{in}$  are considered. The whole set of FOM simulations performed is represented inside the space of the studied parameters (Fig. 1b). The snapshots matrix is built with the spatial profiles of the step response of the FOM simulations, sampled uniformly in time. To build the snapshots matrix, different sampling strategies in the parameter space have been evaluated, but only the one that has given the best results in terms of ROM accuracy is reported in this work. Particularly, the snapshots matrix is built using only two FOM step response solutions. Those solutions were obtained for the values of the parameters highlighted in Fig. 1b, that is  $T_{in} = [295.3, 323]$  °C,  $\Delta T_{in} = 10$  °C, and  $R = 1.6$ . After computing the set of basis functions, the truncation order was set to  $K = 35$  for each state variable. Temperature time series and spatial profiles obtained using FOM and ROM are shown in Fig. 2(a) and (b), respectively. Before the inlet temperature step change, the system was simulated only using FOM, once the step change is imposed both FOM and ROM solutions are determined and reported in Figs. (2). This is because the ROM was built to reproduce only the step response of the system. As an interesting case study, to assess the ROM accuracy, an inlet temperature step change was applied to the operating conditions for which the oscillatory regime (or) was encountered. The parameters used in

the simulation are  $T_{in}=306\text{ }^{\circ}\text{C}$  and  $R=1.4$ . As reported in a previous study (Bareschino et al., 2022), this set of parameters is inside a region of periodic regimes. To avoid these undesired behaviors, the inlet temperature can be reduced, for instance, by  $10\text{ }^{\circ}\text{C}$ . Forcing the system with a step of  $\Delta T_{in}=-10\text{ }^{\circ}\text{C}$  results in a steady-state solution. As can be seen from the overlapping curves shown in Fig. 2a, the ROM predicts even the steepest transient oscillations quite accurately.

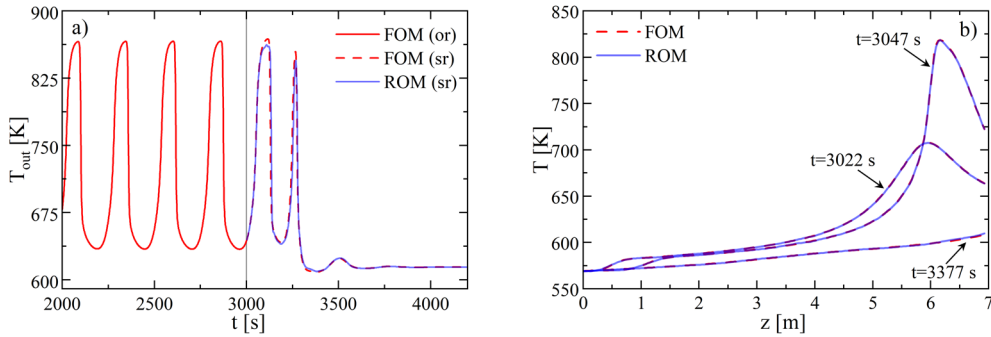


Figure 2. Time series (a) of the outlet temperature as computed with FOM before the step forcing (in the legend, or: oscillatory regime), and with both FOM and ROM after the step forcing (in the legend, sr: step response); (b) spatial profile of the temperature at different time instants confirming the accuracy of the ROM solution inside the reactor.

The choice of using only 35 POD modes was made to achieve a sufficient reduction of CPU time and to maintain suitable accuracy of the obtained ROM. The resulting ROM has 210 degrees of freedom, which is one order of magnitude lower than the corresponding FOM. The simulation of the step response toward a stable steady state is less computationally expensive, and in this case, the speed up (the ratio of CPU time of ROM simulation over FOM simulation) of the analyzed ROM is 15. For the chosen approximation, the computed relative root mean squared error (RMSE) between FOM and ROM (calculated for the solutions described in Fig. 2) is of the order of  $10^{-3}$  with a maximum value of the error in correspondence of the peak value of  $T$  (at 3130 s) of  $1.9 \cdot 10^{-2}$  and a steady state error of  $6.5 \cdot 10^{-5}$ . Although the ROM was built without taking into account periodic regime solutions, it can predict such regimes accurately.

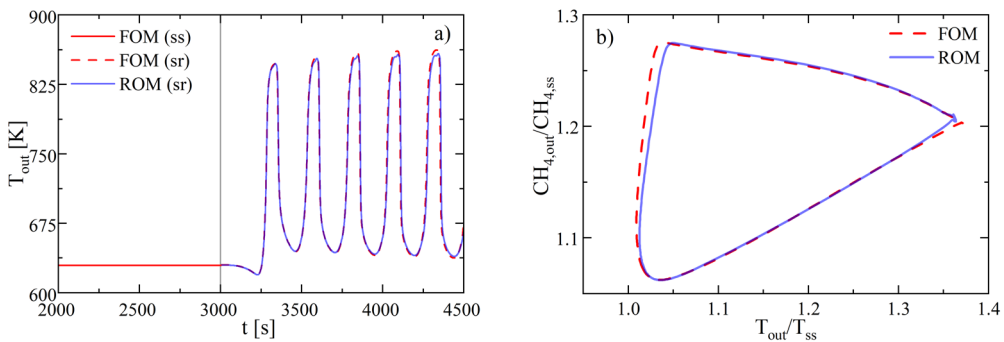


Figure 3. Time series as computed with FOM before the step forcing (in the legend, ss: steady state), and with both FOM and ROM after the step forcing – after the vertical line – (a); and phase space to highlight the accuracy of the ROM predicting the oscillatory regime (b).

This is demonstrated in Fig. 3: as it can be seen, if the system is forced to go from a stable steady state (ss) regime towards an oscillatory regime, the ROM correctly predicts periodic oscillations. The parameter values used in the simulation are  $T_{in}=300.7$ ,  $R=1.6$ ,  $\Delta T_{in}=10$  °C. In Fig. 3b, to highlight the accuracy of the ROM, the limit cycle obtained using both models are reported. In this case, the simulation is more expensive in terms of CPU time both for FOM and ROM because the solution is characterized by steep gradients that slow the convergence of the employed stiff-stable solver. In this case, the ROM achieves a speed up of 5 with a computed RMSE of  $4 \cdot 10^{-3}$ .

## Conclusions

The simulation of the complex system that describes the methanation of biogas in a fixed bed reactor with mass recycle has been conducted by means of a purposely developed efficient reduced-order model, and its application is critically discussed. For the case analyzed in this study, it has been shown that sufficiently reliable POD basis can be extracted using two solutions obtained from FOM simulations. Despite the nonlinearities, which might have made ROM perform worse due to the need of recomputing the nonlinear terms in the FOM space at each step, the chosen approach resulted in a significant reduction in calculation time, with speedup between 5 and 15 for periodic and stationary regimes, respectively. The dimension of the system has been reduced by one order of magnitude, preserving accuracy both of stationary and periodic solutions. Having assessed the performance of the proposed ROM, a more efficient optimization or a real-time control application can be developed. A considerably larger benefit in terms of reduction of CPU time is expected if the methodology is applied to reduce more complex models such as those considering the radial distribution of temperature and concentration.

## References

- P. Bareschino, A.E. Cuttillo, C. Tregambi, F. Pepe, G. Continillo, E. Mancusi, 2022, Periodic Oscillations in Methane Reactor: Effects of the Main Operating Parameters, *Comput. Aided Chem. Eng.* 51 1–6.
- P. Bareschino, E. Mancusi, C. Tregambi, F. Pepe, M. Urciuolo, P. Brachi, G. Ruoppolo, 2021, Integration of biomasses gasification and renewable-energies-driven water electrolysis for methane production, *Energy*, 230, 120863
- K. Bizon, G. Continillo, 2021, Efficient optimization of a multifunctional catalytic fixed-bed reactor via reduced-order modelling approach, *Chem. Eng. Res. Des.* 165 214–229.
- K. Bizon, G. Continillo, M. Berezowski, J. Smuła-Ostaszewska, Optimal model reduction by empirical spectral methods via sampling of chaotic orbits, 2012, *Physica D*, 241, 1441-1449.
- K. Bizon, G. Continillo, 2012, Reduced order modelling of chemical reactors with recycle by means of POD-penalty method, *Comput. Chem. Eng.* 39 22–32.
- J. Bremer, P. Goyal, L. Feng, P. Benner, K. Sundmacher, 2017, POD-DEIM for efficient reduction of a dynamic 2D catalytic reactor model, *Comput. Chem. Eng.* 106 777–784.
- E. Mancusi, P. Bareschino, P. Brachi, A. Coppola, G. Ruoppolo, M. Urciuolo, F. Pepe, 2021, Feasibility of an integrated biomass-based CLC combustion and a renewable-energy-based methanol production systems, *Renew. Energy*. 179 29–36.
- S. Rönsch, J. Schneider, S. Matthischke, M. Schlüter, M. Götz, J. Lefebvre, P. Prabhakaran, S. Bajohr, 2016, Review on methanation - From fundamentals to current projects, *Fuel*. 166 276–296.
- S. Sahyoun, S. Djouadi, 2013, Local Proper Orthogonal Decomposition based on space vectors clustering, 2013 3rd Int. Conf. Syst. Control. ICSC 2013. 665–670.
- J. Xu, G.F. Froment, 1989, Methane steam reforming, methanation and water-gas shift: I. Intrinsic kinetics, *AIChE J.* 35 88–96.
- R. T. Zimmermann, J. Bremer, K. Sundmacher, 2022, Load-flexible fixed-bed reactors by multi-period design optimization, *Chem. Eng. J.* 428 130771.

# A Novel Approach for Cost-effective and Sustainable Capacity Expansion Utilizing Process Intensification in Chemical Process Industries

Chinmoy B. Mukta, Selen Cremaschi, Mario R. Eden\*

*Department of Chemical Engineering, Auburn University, Auburn, AL 36849, USA*

*\*edenmar@auburn.edu*

## Abstract

Sustainable and cost-effective capacity expansion in chemical process industries is essential. In this work, a systematic capacity expansion procedure is developed based on reactor intensification, which can help reliably formulate the expansion horizon and investment plans to be determined from a large solution space. Using the developed approach, the case study showed the impact of ethylene oxide reactor intensification on various capacity expansion modes. With a base case of 100 kton/yr, the Brownfield process is the cheapest capacity expansion option up to 112 kton/yr. Subsequently, the retrofitted mode yielded a minimum cost until a capacity of 157 kton/yr, and above that, we would need green field capacity expansion. On the other hand, in terms of utility carbon emission, Greenfield was the lowest carbon producer among the investment alternatives.

**Keywords:** Process Intensification; Capacity expansion; Greenfield process; Brownfield process; Process Retrofitting

## 1. Introduction

Widespread use of modern amenities by human society requires continued expansion of the capacity of the chemical process industry. As the commodity industry contributes about 32% of US emissions stemming from energy usage (Brueske *et al.*, 2015), the question remains: What are the most cost-effective and sustainable ways to expand the capacity of chemical processes? In this regard, various process intensification technologies might play a pivotal role in sustainable capacity expansions. Process intensification can provide innovative approaches to process and equipment design, which can bring significant benefits in process efficiency and lower capital and operating expenses (Demirel *et al.*, 2019).

New reactor intensification technologies have shown significant advancement, such as enhanced heat transfer reactors (Sheng *et al.*, 2011; Rye *et al.*, 2017). Reactor intensification has a profound impact on downstream separation and as such the overall process. We analyzed the keywords of 1800 research journal papers highlighting “Process Intensification”. A high number of journal papers indicate significant interest in process intensification and their distribution in various fields is shown in Figure 1 by the size of the circle representing the relative research concentration in each field. That also emphasizes the need for considering separation process and capability upon reactor intensification.



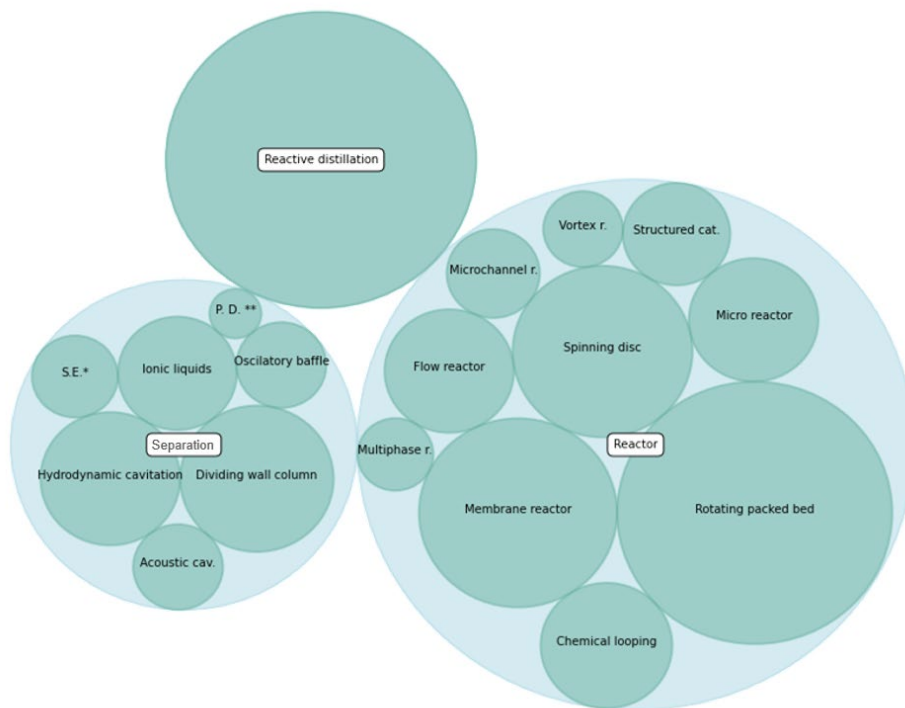


Figure 1. Research areas within Process Intensification (\*Supercritical extraction; \*\*Precipitation disk; r.: reactors.)

Traditional ways of capacity expansion mostly hinge on ultimately building new plants (Greenfield). Increasing the capacity of existing plants (Brownfield) through retrofitting sometimes occurs without effectively considering alternative technologies. We have shown the potential capacity expansion option advantages in Figure 2. First, it can be assumed that potential process intensification options can be readily applied to the green field mode in all parts of the process as we are building the process plant from the ground up. For retrofit, the process intensification options can only be applied to a particular (smaller) section of the plant that we are retrofitting, as we might be using rest of the existing process flowsheet and equipment configuration. For Brownfield, assuming no retrofitting is done in this mode, the scope of process intensification gets reduced even further. It is also evident that for Greenfield the execution time and capital investment might be higher than the Brownfield and retrofitted processes.

It is important to find a suitable intensification technology candidate which can capture the effects of the different capacity expansion modes that we have discussed. Most of the intensification technologies might only affect certain unit operations rather than the overall process. Among the intensification technologies discussed, reactor intensification may have a profound impact on both feed conditioning and downstream separations. Micro fibrous Entrapped Catalyst (MFEC), a reactor intensification technique, can be an ideal candidate in this case (Cheng *et al.*, 2022). MFEC is a microfibrinous metal mesh of either Nickel, steel, or Copper. Moreover, MFEC can remove the reaction heats fast from the catalyst-filled tubes through a conductive heat transfer mechanism instead of

traditional convective heat transfer. Moreover, the smaller particles used in MFEC-based reactors, in comparison to larger particles (>1mm) in the traditional packed bed reactor, provide better control of the reactor temperature and thus quench the heat. MFEC has the potential to reduce the intra-particle and inter-particle transport resistances considerably. All these properties of MFEC are advantageous and affect the downstream process and overall capacity expansion plans.

## 2. Methodology

This section expands on cost calculation methodologies that we discussed in our previous work for a different capacity expansion scenario (Mukta *et al.*, 2022). In this work we are applying the methodology to reactor intensification by Microfibrous Entrapped Catalyst (MFEC), which is a highly conductive catalyst support that intensifies reactors through conductive rather than convective heat transport and large surface area. This technology is an ideal candidate to test our methodology for the impact of reactor intensification on various capacity expansion modes as MFEC intensified reactors have the advantage of positively impacting reaction rates, mass transfer, and heat transfer inside the reactor. This can significantly affect the downstream separation process i.e. capacity expansion plan.

### 2.1. Capacity Expansion Options

Greenfield cost calculation followed the conventional process plant installation procedure. Different formulations considering the lack of land development requirements, module cost, and utility requirements should be considered for Brownfield and retrofitting cases.

Table 1: Capacity expansion options

<b>Greenfield</b>	<b>Retrofit 1</b>	<b>Retrofit 2</b>	<b>Brownfield</b>
A new plant on a newly developed land	Replacing reactor or other equipment	Replacing reactor	Replacing the catalyst in the existing reactor with MFEC
Capital cost calculations based on grass root equipment cost	Capital cost calculations based on modular equipment cost		
New utility systems	Existing utility systems and new utility systems can be installed	Existing utility systems	

### 2.2. Systematic Procedure for Capacity Expansion

It is important to utilize existing infrastructure from both economic and sustainability perspectives. A capacity expansion plan in process industries considering existing facility utilization entails large number of decisions. Identification of the optimum capacity expansion may require significant computational resources. A systematic procedure is necessary to improve the tractability of this potentially large problem. That is why in addition to traditional Greenfield, we analyzed and compared process feasibility and economic performance of other viable capacity expansion modes such as Brownfield and retrofitted processes. The developed systematic procedure is illustrated in Figure 2, where a step-by-step approach is used to selectively reduce the solution space as we expand the

capacity of process plants. First, we determine process flowsheet alternatives of green field mode upon reactor intensification using ProCAFD (Tula *et al.*, 2019). Subsequently, the Brownfield alternative is analyzed to see if equipment configurations can be optimized further to achieve capacity expansion. Next, the retrofitted process is analyzed as it has a significantly higher number of decision variables/constraints compared to the Greenfield and Brownfield options.

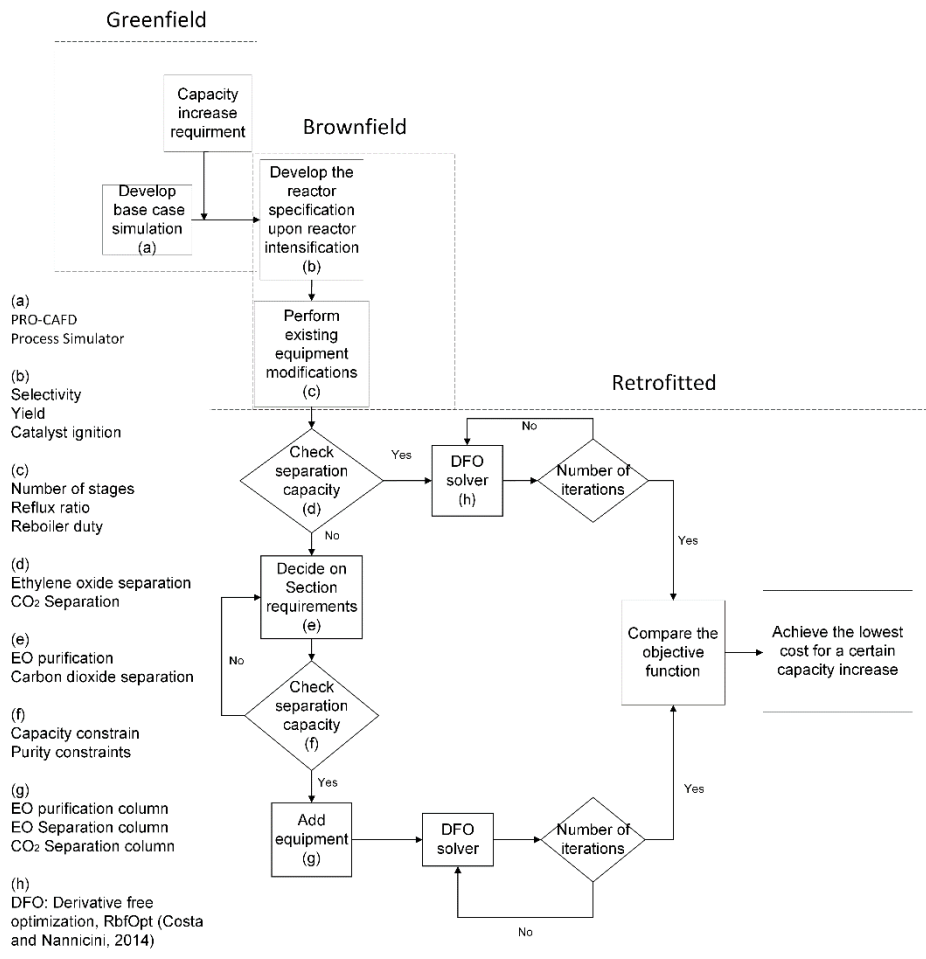


Figure 2. Flowchart showing capacity expansion procedure.

### 3. Case Study

A baseline simulation model of the conventional Ethylene oxide (EO) process has been developed using Aspen Plus<sup>TM</sup> and was described in our previous work (Mukta *et al.*, 2020). Next, standard coefficient-based costing equations were used to calculate capital and operating costs. Then the cost was annualized to form our objective function for the derivative-free optimization algorithm.

Brownfield process achieved the highest cost reduction, but above 112 kton/yr, the separation capacity was insufficient to handle excess reactor capacity. So the framework suggests the next favorable process is the retrofitted process. The retrofitted process-1 achieved a considerable decrease in production cost, but above 130 kton/yr, production costs increased dramatically with additional columns for ethylene oxide separation. Greenfield process expansion becomes cost-effective only after 160 kton/yr. The Greenfield process achieved the lowest carbon emission. The retrofitted process had a moderate increase in carbon emissions. The Brownfield process showed significant increase in carbon emissions upon an increase in capacity as the separation column operates at a higher capacity than it was designed for thus resulting in suboptimal conditions. It can be seen from the utility carbon emission graph that the Greenfield mode produces the minimum utility carbon among all capacity expansion scenarios that we checked. Carbon emission due to construction was not considered for the new equipment for retrofitted process or Greenfield, which might affect the overall carbon emissions of the process plant.

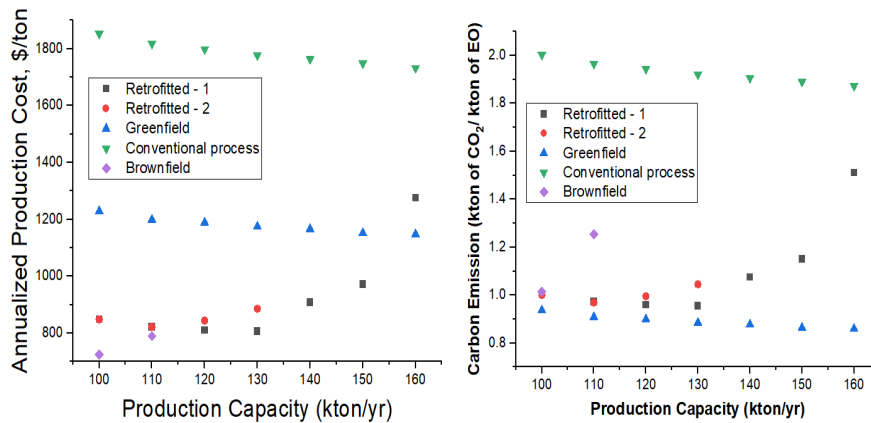


Figure 3: Techno-economic capacity expansion results and associated utility carbon emission.

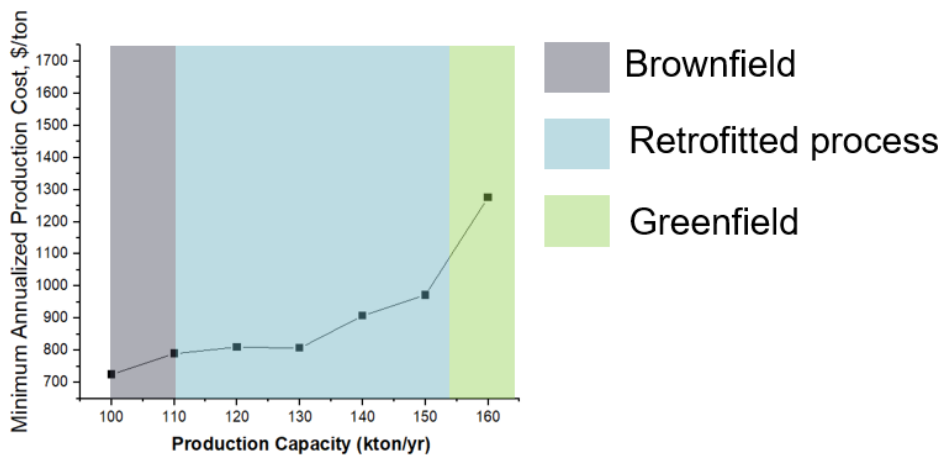


Figure 4: Various capacity expansion modes regions.

## Conclusion

A systematic procedure of capacity expansion upon reactor intensification for ethylene oxide production has been developed that might be applicable to other chemical processes. The Brownfield process is found to be the cheapest capacity expansion option up to 112 kton/yr. Subsequently, the Retrofitted mode yielded a minimum cost until 157 kton/yr capacity, and above that, we would need Greenfield expansion. But in terms of utility carbon emission, Greenfield was the lowest carbon producer among the investment alternatives.

## References

- Brueske, S., Kramer, C., & Fisher, A. (2015). Bandwidth Study on Energy Use and Potential Energy Saving Opportunities in US Chemical Manufacturing (No. DOE/EE-1229). Energetics.
- Cheng, X., Yang, H., & Tatarchuk, B. J. (2016). Microfibrous entrapped hybrid iron-based catalysts for Fischer–Tropsch synthesis. *Catalysis Today*, 273, 62-71.
- Costa, A., Nannicini, G., 2014. RBFOpt : an open-source library for black-box optimization with costly function evaluations. *Optim.* online no.4538.
- Demirel, S. E., Li, J., & Hasan, M. F. (2019). Systematic process intensification. *Current Opinion in Chemical Engineering*, 25, 108-113.
- Lutze, P., Babi, D. K., Woodley, J. M., & Gani, R. (2013). Phenomena-based methodology for process synthesis incorporating process intensification. *Industrial & Engineering Chemistry Research*, 52(22), 7127-714
- Mukta, C. B., Rayaprolu, N. R., Cremaschi, S., Eden, M. R., & Tatarchuk, B. J. (2022). Techno-Economic Study of Intensified Ethylene Oxide Production Using High Thermal Conductivity Microfibrous Entrapped Catalyst. In *Computer Aided Chemical Engineering* (Vol. 49, pp. 697-702). Elsevier.
- Ryu, J. H., Lee, K. Y., La, H., Kim, H. J., Yang, J. I., & Jung, H. (2007). Ni catalyst wash-coated on metal monolith with enhanced heat-transfer capability for steam reforming. *Journal of Power Sources*, 171(2), 499-505.
- Sheng, M., Yang, H., Cahela, D. R., & Tatarchuk, B. J. (2011). Novel catalyst structures with enhanced heat transfer characteristics. *Journal of Catalysis*, 281(2), 254-262.
- Tula, A. K., Eden, M. R., & Gani, R. (2019). ProCAFD: Computer-aided Tool for Sustainable Process Synthesis, Intensification and Hybrid solutions. In *Computer Aided Chemical Engineering* (Vol. 46, pp. 481-486). Elsevier.

# Determining the performance and network properties of petrochemical clusters

Michael D. Tan,<sup>a\*</sup> Paola Ibarra-Gonzalez,<sup>a</sup> Igor Nikolic,<sup>b</sup> Andrea Ramirez<sup>a</sup>

<sup>a</sup> *Department of Engineering Systems and Services, Faculty of Technology, Policy and Management, Delft University of Technology, Jaffalaan 5, 2628BX Delft, Netherlands*

<sup>b</sup> *Department of Multi-Actor Systems, Faculty of Technology, Policy and Management, Delft University of Technology, Jaffalaan 5, 2628BX Delft, Netherlands*

[m.d.tan@tudelft.nl](mailto:m.d.tan@tudelft.nl)

## Abstract

The reliance of the petrochemical industry on fossil-based sources will need to be reduced by the introduction of Alternative carbon sources (ACS). Introducing ACS in a petrochemical cluster will require existing processes to be modified or replaced, potentially affecting other chemical processes within the cluster due to existing material and energy interconnections. Therefore, it is important to understand the current level of interconnections, functioning, and performance of the petrochemical cluster before introducing ACS. In this work, a representative cluster model based on the petrochemical cluster of the Port of Rotterdam was developed and considered as a case study. This model was analyzed using complex network analysis and environmental and technical key performance indicators. The selected key performance indicators (KPIs) provide insight into the performance of a petrochemical cluster, while the network properties give an understanding of the exchange of material and energy in an industrial cluster.

**Keywords:** petrochemical cluster, key performance indicators, complex network properties, material transition, alternative carbon sources

## 1. Introduction

The petrochemical industry is reliant on fossil-based sources as an energy and carbon source, and therefore it will have to transition to alternative carbon sources to reach the CO<sub>2</sub> reduction goals. The use of new carbon feedstocks (CO<sub>2</sub>, biomass, waste) will require the modification or removal of existing chemical processes. Given the high level of interconnections in existing industrial clusters, these changes are likely to impact existing material and energy connections between processes and among companies in petrochemical clusters. Before these impacts can be assessed, it is needed first to evaluate the complexity and interdependency of petrochemical clusters in terms of material and energy connections and quantify their performance so they can be used as a point of departure.

Network analysis is a method used to study complex systems such as industrial clusters. It allows the interdependencies of processes or companies within a cluster to be assessed by quantifying the exchange of material, energy, or knowledge between processes and or companies. Prior studies (Domenech and Davies, 2011; Song et al., 2018) focused on the occurrence of a link between processes and companies, and assumed that each link is equally important. However, processes are diverse, not only in terms of products but also

in terms of production capacities, and the magnitude of their incoming and outgoing materials and energy flows will also vary as a result. Therefore, it is important that the relative importance of the interconnections is identified and included in the analysis.

In addition to understanding the interdependencies of processes within a petrochemical cluster, additional metrics are required to determine the environmental and technical performance of a petrochemical cluster. Key performance indicators (KPIs) that have been used to assess the performance of industrial clusters are, for instance, CO<sub>2</sub> emissions (Yu et al., 2015) and energy consumption (Sokka et al., 2011).

In this work, a new framework for assessing the complex network properties and performance of a petrochemical cluster is presented. The framework was applied to a case study based on the petrochemical cluster in the Port of Rotterdam (PoR).

## 2. Petrochemical Cluster Model Framework

The methodological framework used to characterize and evaluate the performance of the representative petrochemical cluster (RPC) is presented in Figure 1. The first step is to select the petrochemical processes and utility units that are part of the RPC. For instance, processes producing chemical building blocks (CBBs), intermediate chemicals (ICs), and end-of-value chain chemicals (EVCs) were selected based on the PoR cluster. Additionally, utility production processes that provide the required steam, electricity, and auxiliary chemicals for the chemical processes need to be selected.

In the next step, detailed process models based on publicly available data were built for each selected chemical and utility generation process using Aspen Plus. Each process model was modeled according to the production capacity and process technology of its counterpart in the PoR cluster. These process models provided detailed material and energy balances for each process and utility generation unit. Furthermore, the bare equipment costs and land footprint were determined for each process using Aspen Process Economics Analyzer.

In the third step, the material and energy connections between the processes in the RPC were mapped. This was done by collecting the results of the models and publicly available data on infrastructure and connections of the PoR and using it to match the material and energy requirements within the processes in the RPC. The resulting mapping of exchanges contains all connections between the processes within the cluster and the processes' connections to the outside world. For example, natural gas imported into the cluster, wastewater sent to wastewater treatment plants, and emissions emitted to the environment.

Next, detailed material and energy stream data from the Aspen Plus process models is automatically extracted using an in-house developed Python module. The material and

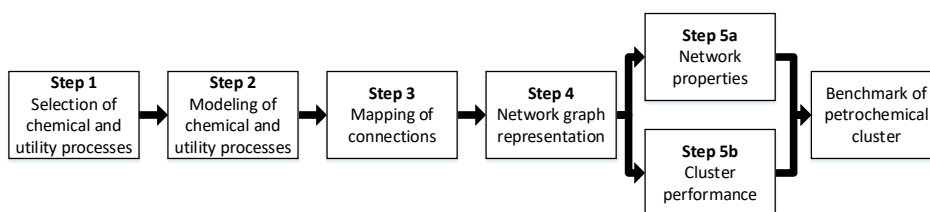


Figure 1: Methodological framework for modeling and assessing a petrochemical cluster

energy data is used to construct a complex network representation in Python using Py3plex (Škrlj et al., 2019). This complex network representation uses graph theory to describe the petrochemical cluster, where the material and energy connections are represented as links, and the chemical and utility generation processes are depicted as nodes. In this work, we developed a multiplex graph of the cluster, where each layer represents a different type of interaction between the nodes. For instance, the first layer represents all the material interactions, and the second layer depicts all the energy exchanges between processes. The only type of interlayer connections are with the counterpart of each node present in the other layers.

Py3plex is based on the complex network analysis Python module NetworkX by Hagberg et al. (2008), thereby allowing the complex network of the graph to be calculated. The number of connections a node has  $k_i$ , also known as the degree of node  $i$ , was determined by:

$$k_i = \sum_{j \in N} a_{ij} \quad (1)$$

where  $N$  is the number of nodes, and  $a_{ij}$  is zero if there is no direct link between nodes  $i$  and  $j$  and one if there is. Alternatively, the degree centrality  $C_{D,i}$  was calculated by:

$$C_{D,i} = \frac{\sum_{j \in N} a_{ij}}{N-1} \quad (2)$$

The degree centrality only considers whether a connection between processes is present, and the magnitude of the material or energy exchange is not considered in its calculation. It determines the importance of a process in the petrochemical cluster by considering the number of connections it has to other processes in the cluster. To consider the magnitude of the material or energy exchange, the strength of a node  $s_i$  was calculated by:

$$s_i = \sum_{j \in N} w_{ij} \quad (3)$$

where  $w_{ij}$  is the weight of the link between nodes  $i$  and  $j$ . For instance, in the context of the transformation of petrochemical clusters, the mass of carbon in a link could be considered to calculate the weight of a link as it allows the most important nodes in terms of carbon flows between processes to be identified.

The final step consisted of determining the environmental and technical performance of a petrochemical cluster. The graph representation and list of connections were used to calculate the key performance indicators carbon efficiency, CO<sub>2</sub> emissions, and total energy usage within the boundaries of the cluster. The carbon efficiency  $\eta_{Carbon}$  of the cluster is calculated by:

$$\eta_{Carbon} = \frac{\sum m_{Carbon,p}^{Product}}{\sum m_{Carbon,f}^{Feed}} \quad (4)$$

where  $m_{Carbon,p}^{Product}$  is the mass of carbon present in stream  $p$  leaving the petrochemical cluster, while  $m_{Carbon,f}^{Feed}$  is the total mass of carbon present in the material feed stream  $f$  entering the cluster. The total CO<sub>2</sub> emissions of the petrochemical cluster are determined by the mass of CO<sub>2</sub> present in the streams being emitted to the environment:



$$m_{CO_2} = \sum_{i=1}^{N_{streams}} m_{i,CO_2}^{Environment} \quad (5)$$

Where,  $m_{i,CO_2}^{Environment}$  is the mass flow rate of CO<sub>2</sub> in stream  $i$  emitted to the environment, and  $N_{streams}$  is the total amount of streams emitted to the environment.

As described before, the framework presented was implemented using an RPC based on the PoR cluster, and the results of its implementation are discussed in the next section.

### 3. Results and Discussion

Nine utility generation processes and 33 petrochemical processes producing 52 chemicals were selected and modeled in Aspen Plus. Based on these models, a complete mapping of all the material and energy streams of the cluster was created and used to construct a complex network representation of the petrochemical cluster. A two-layer multiplex graph was created containing 49 nodes, where the first layer represented the material exchanges and contained 64 links, while the second layer represented the exchange of energy and contained 48 links. In figure 2, a multiplex graph representation of a small section of the cluster consisting of four chemical processes and three utility generation units is shown.

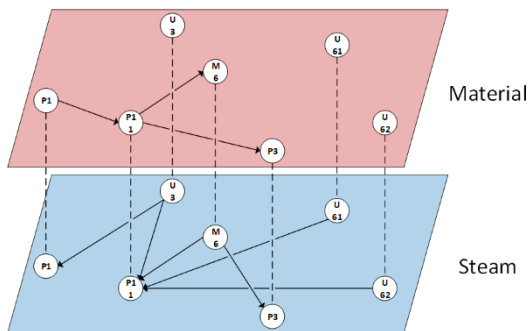


Figure 2: Multiplex graph representation.

The degree centrality and strength of the nodes of the material layer are presented in Table 1. It shows that the EDC/VCM plant and the SMR plant are the most interconnected processes of the cluster, followed by the aromatics and olefins plants. The EDC/VCM plant is part of the chlorine sub-cluster, which contains many interconnections. The SMR plant supplies H<sub>2</sub>, which is used as auxiliary material in the production of several chemicals, while the aromatics and olefins plant produce the CBBs that are either directly or indirectly used by the other processes within the cluster. Therefore a high level of interconnectivity is expected for these processes. When considering the strength of the nodes, the aromatics plant appears as the most critical process in terms of carbon flows, followed by the ethylbenzene (EB) and olefins plants, respectively.

Table 1: Degree centrality and strength of the most interconnected processes on the material layer

Process	Degree centrality	Process	Strength (ktonne of carbon per year)
Ethylene dichloride/Vinyl chloride monomer (EDC/VCM)	0,19	Aromatics	1757,19
SMR	0,19	Ethylbenzene (EB)	1430,39
Olefins	0,17	Olefins	1110,50
Aromatics	0,17	Propylene oxide/Tert-butyl alcohol (PO/TBA)	978,34
Chlorine	0,13	Propylene oxide/Styrene monomer (PO/SM)	866,78

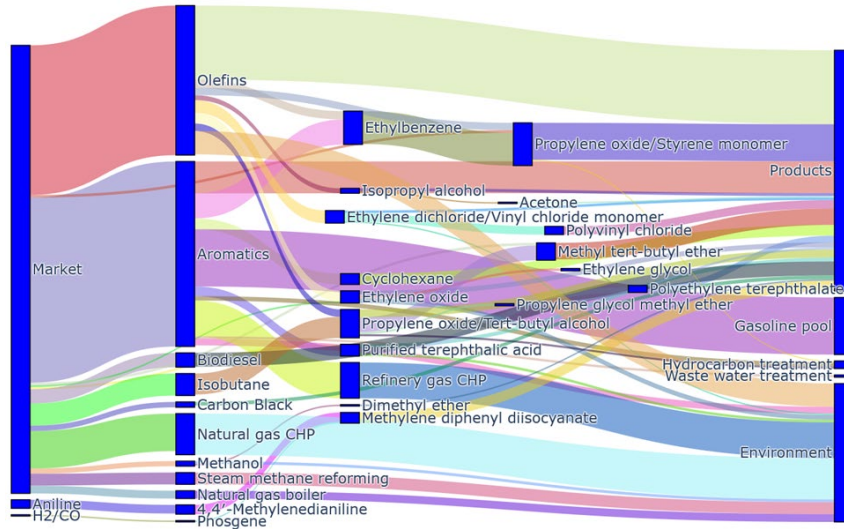


Figure 3: Sankey diagram of all the carbon mass flows in the petrochemical cluster.

In terms of performance, carbon mass flows of the petrochemical cluster are shown in a Sankey diagram in Figure 3. In this diagram, all the chemical processes, the mass flows of carbon between the processes in the cluster, and the carbon flows from and to the outside world are shown. The market represents any material flows being imported into the cluster, while the products represent all material being exported out of the cluster, and the environment represents any CO<sub>2</sub> directly being emitted to the environment. The petrochemical cluster has an overall carbon efficiency of 58.5%. Nearly all the fossil-based carbon feedstock imported into the cluster is sent to the aromatics and olefins plants, with 44.3% and 35.9% of the carbon feedstock, respectively. These plants transform the carbon feedstock into CBB, and these are distributed across the different value chains that make up the cluster. Therefore, replacing these fossil-based processes with alternative carbon source processes such as CO<sub>2</sub>, biomass, or waste will most likely have the most significant impact on the transformation of the petrochemical industry. Additionally, not all chemicals produced by the olefins plant and the aromatics plants are used by the downstream chemical processes in the petrochemical cluster and instead exported out of the cluster. Compared to the strength of the nodes, the importance of the olefins and aromatics is more clearly defined in the Sankey diagram. This is due to the

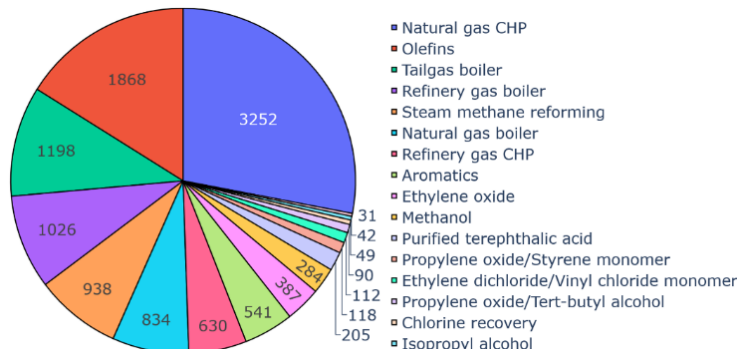


Figure 4: Pie chart of the CO<sub>2</sub> emissions of the chemical and utility generation processes in the cluster in ktonne of CO<sub>2</sub> per year.

manner the strength of a node is calculated, with only links between processes inside the cluster being considered and not the stream entering or leaving the cluster.

In this case study, the petrochemical cluster emits 11.647 Ktonne of CO<sub>2</sub> every year. The distribution of direct CO<sub>2</sub> emission for all chemical and utility processes emitting at least 25 ktonne per year is shown in Figure 4. A major part of all CO<sub>2</sub> emissions in the cluster is from the utility units that provide the required steam, electricity, and auxiliary chemicals by the chemical processes. The olefins plant has the highest direct CO<sub>2</sub> emissions, as it burns byproducts to provide the heat required for the process. However, new alternative carbon source processes could result in a shift of CO<sub>2</sub> from inside the cluster to outside the cluster. Therefore, an LCA cradle-to-gate approach should be implemented. This will be investigated in further research.

#### 4. Conclusions

The transformation of a petrochemical cluster to more sustainable carbon sources can impact the exchange of material and energy between processes and the performance of the cluster. Thus, network properties and performance indicators are required to assess the current petrochemical cluster and future configurations. In this work, a framework was presented for modeling a petrochemical cluster from the bottom up based on the cluster in the Port of Rotterdam. This model was analyzed using complex network analysis and environmental and technical key performance indicators. In future work, the network properties and performance of a modified cluster, including alternative carbon source processes, will be compared with the benchmark developed in this work, allowing the potential impacts of the transformation of petrochemical cluster to be identified.

#### Acknowledgements

This publication is part of the project Unravelling the impacts of using alternative raw materials in industrial clusters (with project number VI.C.183.010 of the research programme Vici DO which is (partly) financed by the Dutch Research Council (NWO)).

#### References

- Domenech, T., Davies, M., 2011. Structure and morphology of industrial symbiosis networks: The case of Kalundborg. *Procedia - Soc. Behav. Sci.* 10, 79–89. <https://doi.org/10.1016/j.sbspro.2011.01.011>
- Hagberg, A.A., Schult, D.A., Swart, P.J., 2008. Exploring network structure, dynamics, and function using NetworkX. *7th Python Sci. Conf. (SciPy 2008)* 11–15.
- Škrlić, B., Kralj, J., Lavrač, N., 2019. Py3plex toolkit for visualization and analysis of multilayer networks. *Appl. Netw. Sci.* 4, 94. <https://doi.org/10.1007/s41109-019-0203-7>
- Sokka, L., Pakarinen, S., Melanen, M., 2011. Industrial symbiosis contributing to more sustainable energy use - An example from the forest industry in Kymenlaakso, Finland. *J. Clean. Prod.* 19, 285–293. <https://doi.org/10.1016/j.jclepro.2009.08.014>
- Song, X., Geng, Y., Dong, H., Chen, W., 2018. Social network analysis on industrial symbiosis: A case of Gujiao eco-industrial park. *J. Clean. Prod.* 193, 414–423. <https://doi.org/10.1016/j.jclepro.2018.05.058>
- Yu, F., Han, F., Cui, Z., 2015. Reducing carbon emissions through industrial symbiosis: A case study of a large enterprise group in China. *J. Clean. Prod.* 103, 811–818. <https://doi.org/10.1016/j.jclepro.2014.05.038>

# Quality assessment of partial models for co-simulation of modular electrolysis plants

Isabell Viedt<sup>a</sup>, Kumar Rajan Gopa<sup>a</sup>, Jonathan Mädler<sup>a</sup>, Leon Urbas<sup>a,b</sup>

<sup>a</sup>*TU Dresden, Process Systems Engineering Group, 01069 Dresden, Germany*

<sup>b</sup>*TU Dresden, Chair of Process Control System, 01069 Dresden, Germany*

## Abstract

Co-simulation of modular process plants combines partial models for each Process Equipment Assembly (PEA) from different manufacturers. Considering the manufacturer's intellectual property (IP), the co-simulation of the partial models must be realized via a standardized interface using Functional Mock-up Units (FMU). This approach also requires a flexible approach to model quality assessment. In this paper, a workflow for the quality assurance of modularized simulations models for co-simulation is presented via the example of a modular electrolysis plant. The first step is an analysis of the required partial models to derive requirement specifications, the resulting quality models, and the test specifications. These specifications lay the groundwork for the test implementation for the modular electrolysis plant. This use case addresses eighteen partial models of varying modeling depth, including models provided solely as a FMU.

**Keywords:** Model Quality Assurance, Modular Process Plants, Modular Electrolysis.

## 1. Introduction

Modular process plants according to VDI 2776 consist of individual Process Equipment Assemblies (PEA) from different manufacturers (VDI, 2020). With these PEAs envisioned to be equipped with simulations models and a digital twin (Mädler et al., 2022), the simulation of the entire modular process plants must also combine partial models from the individual manufacturers which are potentially confidential. Integrated into varying composite models, co-simulation of different plant configurations can not only be used to determine the best plant setup, but also to test out different scale-up scenarios. As most manufacturers are assumed to be reluctant to share their internal model structure because of IP concerns (Mädler et al., 2022), the co-simulation of the partial models must be realized via a standardized interface, the Functional Mock-up Interface (FMI), using Functional Mock-up Units (FMUs).

Since green hydrogen production will play an important role in the decarbonization of the process industry, scale-up of the hydrogen production is necessary (Buttler & Spliethoff, 2018; IEA, 2021). One approach is the transfer of the concepts from VDI 2776 to electrolysis plants (VDI, 2020). Simulation models will be vital in the testing of the necessary scale-up concepts. In this paper, the workflow for quality assurance of simulations models for co-simulation for modular hydrogen production is presented. The case study consists of eighteen partial models of varying modeling depth and approach, including models provided solely as a FMU. The models are implemented in MATLAB/Simulink and tested via the test-driven modeling workflow.

The remainder of this paper is structured as follows. Section 2 introduces current methods for quality assessment and the methodology for model quality assurance approach

regarding modular plants. In section 3 the case study for the quality assurance of a modular electrolysis plant is presented. The first step is an analysis of the required partial models to derive the requirement specifications, the resulting quality models, and the test specifications. With this, necessary quality factors and criteria such as interoperability, flexibility, and security are discussed as well as functional suitability and reliability. These specifications lay the groundwork for the test implementation and modeling of the electrolysis sample plant. Section 4 provides a conclusion and further research potential.

## 2. Methodology for model quality assurance

### 2.1. Quality assessment of simulation models

In literature, different strategies for the quality assessment of simulation models exist (Murray-Smith, 2015). The most widely known and used strategies are the methods for verification and validation (V&V) by Balci and Sargent (Sargent & Balci, 2017). Although they provide a wide spectrum of assessment methods, the main focus remains on the models' accuracy. Within the meta analysis of Sargent and Balci (2017) it was found that many simulation studies do not consequently apply V&V methods or even disregard them altogether. Mädler et al. (2021) and Viedt et al. (2022) discuss the applicability of quality assurance methods from software development to simulation models. For this approach, quality assessment strategies from software development such as test-driven development and quality models to define quality features in form of FCM (factor–criteria–metrics) models are applied to ensure test-driven modeling.

### 2.2. Quality assurance methodology for modular process plant simulation models

With the increasing importance of model-based approaches for plant optimization and scaling concepts, the quality of simulation models also gains importance. Especially for large-scale co-simulations, where partial models from different manufacturers are combined to form a process simulation of the entire modular plant, quality aspects for the interaction of those models must be considered more directly (Viedt et al., 2022). To ensure high quality co-simulation, quality assurance moves into focus. Thus, quality needs to be a driving factor for the implementation of the models. A shift from 'tagged-on' quality assessment to a quality driven-modeling approach is necessary.

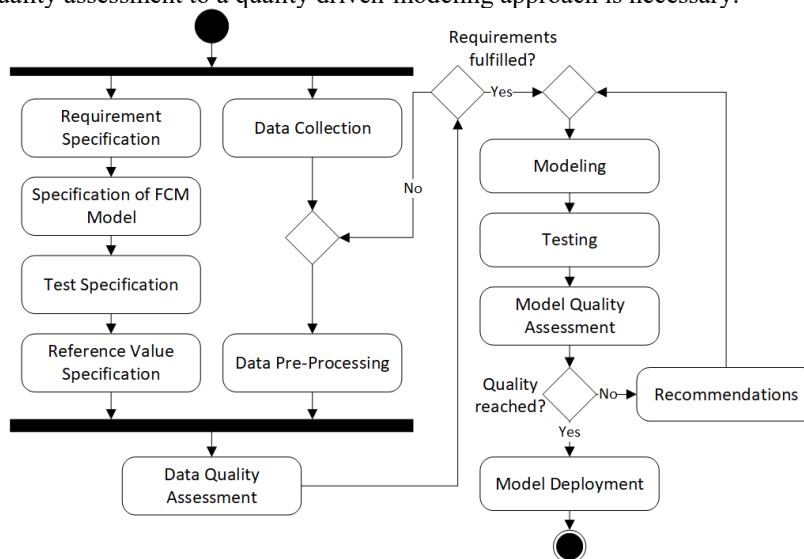


Figure 1: Quality assurance workflow for simulation models

As a result, the pre-definition of requirements, quality features, and tests prior to the start of the actual modeling is of paramount importance. With this, the authors propose a workflow for the quality assurance of simulation models which is shown in Figure 1. The workflow for model quality assurance considers two distinct quality iteration loop: data quality assurance and model quality assurance. Data quality assurance will be necessary in almost all modeling cases. This is because even if data is not used during the actual implementation of the model for a data-driven modeling approach, it will at least be utilized for model validation as well as for parameter identification (Mädler et al., 2022).

### 3. Case study modular electrolysis plant

#### 3.1. Modular electrolysis sample plant

In the research project eModule, standardized plant concepts consisting of PEAs were developed for all three major electrolysis technologies: alkaline electrolysis (AEL), proton exchange membrane electrolysis (PEMEL) and high-temperature electrolysis (HTEL) (Buttler & Spliethoff, 2018). Fig. 2 shows the exemplary flowsheet for the AEL.

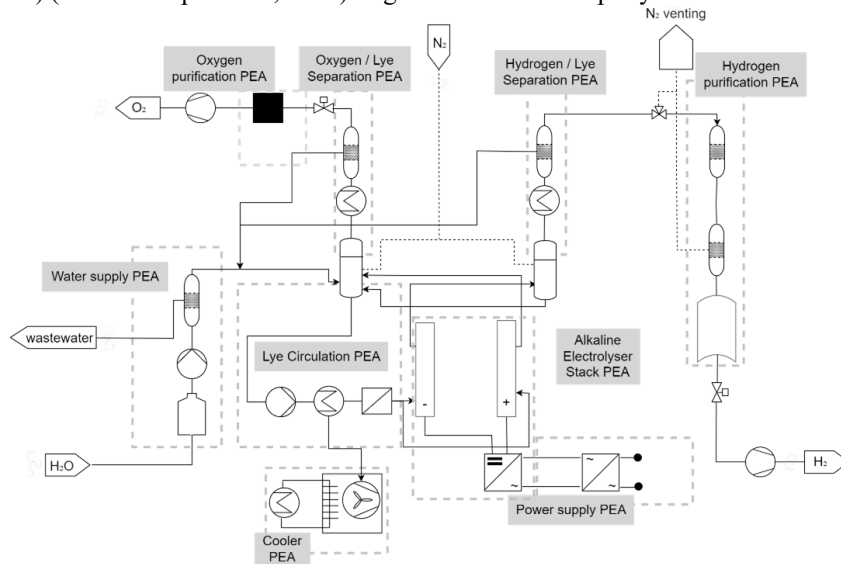


Figure 2: Flowsheet for the AEL - Courtesy of the VDMA Power-to-X for Applications platform (Beisswenger and Erdmann, 2022).

#### 3.2. Partial models and requirements definition

The first step in the test-driven modeling workflow is the specification of the model goal and scope. For the modular electrolysis plant shown in Fig. 2 for the AEL, eight independent partial models were defined: water supply, cooler, lye circulation, AEL stack, oxygen/lye separation, hydrogen/lye separation, oxygen purification and hydrogen purification. In accordance with the project scope, the power supply was not implemented as an individual partial model. For all three considered technologies, AEL, PEMEL, and HTEL, eighteen partial models are defined as necessary according to the requirements specification. This specification defines the necessary functions e.g. interfaces, components, and control functions of each individual model as well as the required modeling depth and approach e.g. a data-driven approach. Once the requirement specification for a model is derived from the modeling and simulation goal, a corresponding FCM model is derived to structure the quality features of the models.

Every metric is allocated a tag that is later used to assign the necessary test specification during the test execution. This tag consists of the tag for the partial model, e.g. WS for water supply or H2LS for the hydrogen/lye separation, and the tag for the specific metric in the FCM model. An excerpt is shown in Table 1.

Table 1: Excerpt of the FCM model for the water supply partial model.

Factor	Criterion	Metric	Tag
Functional Suitability	Functional Correctness	Target conductivity	WS-F1C1M1
	Functional Appropriateness	Degradation rate	WS-F1C2M1
		Conversion efficiency	WS-F1C2M2
	Functional Completeness	Implemented components	WS-F1C3M1
Implemented control functions		WS-F1C3M2	
Reliability	Fault Tolerance	Fault injection tolerance	WS-F2C1M1
Compatibility	Interoperability	Interface standard adherence	WS-F3C1M1

### 3.3. Test specification

The next step in the test-driven modeling workflow is the test specification which also includes the corresponding reference values for the tests. These test specifications are compiled from the aggregated information from the requirements specification, process flow diagram (PFD), the piping and instrumentation diagram (P&ID), the formalized process description, and manufacturer information. Test specification in an early model development phase allows a more comprehensive testing approach, which also makes it easier to detect missing information at an earlier development stage. An excerpt of an exemplary test specification for the AEL stack model is shown in Table 2.

Table 2: Excerpt of an exemplary test specification for the AEL stack partial model.

Tag	Test
AEL-F1C1M1	If the AEL is supplied with power, H <sub>2</sub> O is converted into H <sub>2</sub> and O <sub>2</sub> .
AEL-F1C2M2	If the AEL is not supplied with power, H <sub>2</sub> O is not converted into H <sub>2</sub> and O <sub>2</sub> .
AEL-F3C1M1	The model must adhere to the FMI standard.

### 3.4. Modeling and FMU preparation

In the modeling step the pre-defined partial models are implemented in the chosen modeling environment, here MATLAB/Simulink 2022b. In Fig. 3 the model for the hydrogen/lye separation is shown in a first iteration. In the next step, FMUs must be created for all implemented models. The corresponding FMUs will have the same inlets and outlets as the non-FMU model shown in Fig. 3.

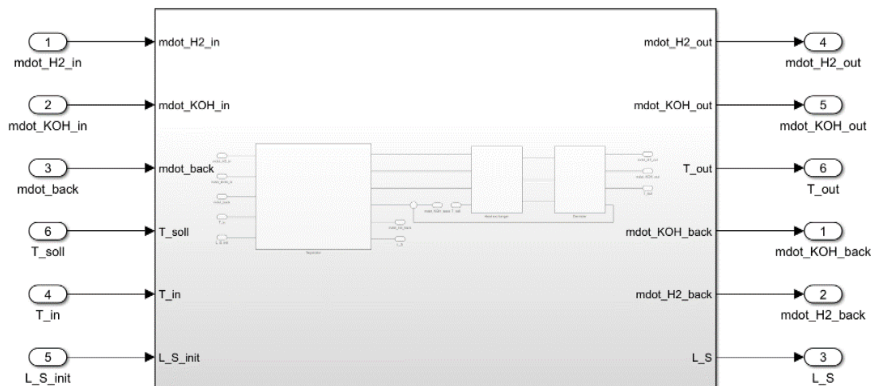


Figure 3: Partial model for the hydrogen/lye separation for AEL.

To enable co-simulation at a later stage, the model and system configurations, which describe the connections between the individual FMUs and the system structure as a whole, must be created. These configurations are compiled from the aggregated information from the requirements specification, block flow diagram (BFD), process flow diagram (PFD), piping and instrumentation diagram (P&ID), and manufacturer information.

### 3.5. Testing and model quality assessment

Once the first modeling iteration step is concluded, the models are tested against the requirements specification and the test specification. For this case study, the tests are implemented and automated via the Simulink Test Manager for both the models implemented in MATLAB/Simulink and the provided FMUs. Since the FMUs represent the partial models provided by the different PEA manufacturers, the internal structure and modeling approach is not known. Therefore, to ensure trust in them, the testing of these partial models must be even more rigorous with more tests and a higher test coverage.

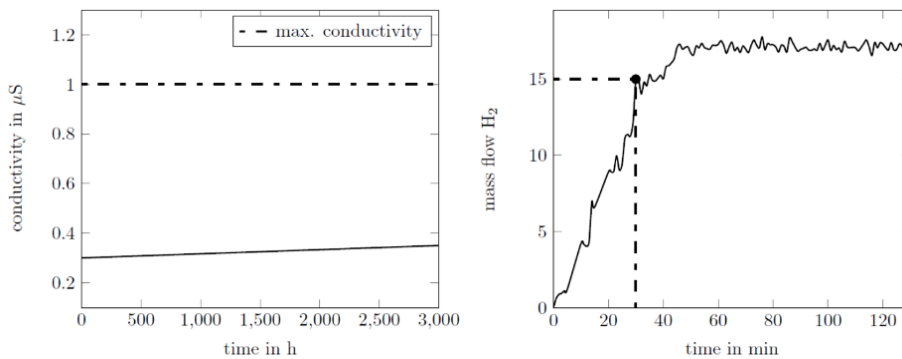


Figure 4: Excerpt of the test results for water supply (left) and hydrogen/lye separation (right).

For the modular electrolysis case study all partial models are tested against their test specification and the requirements specification. Figure 4 shows an excerpt of these test results for the tests tagged WS-F1C1M1 for the water supply and H2LS-F1C1M2 for the hydrogen/lye separation. Both partial models reach the required quality for the two shown metrics of target conductivity and target hydrogen production and can therefore be deployed for the intended purpose. If a partial model does not reach the desired quality, as per the quality assurance workflow (cf. Table 1) another modeling iteration must follow to adjust the model to improve quality. One example of this is the behavior of the hydrogen/lye separation model. In the first modeling iteration, the hydrogen purity metric, H2LS-F1C1M1, did not meet the specification. A corresponding recommendation was the adjustment of the assumptions for vapor pressure and gas solubility. The requirements specification and the FCM model also include further quality factors such as compatibility and efficiency, which are only tested qualitatively. For co-simulation in particular, quality features which do not describe the explicit model behavior are of paramount importance as runtime efficiency and model compatibility will enable successful execution.

## 4. Conclusion and outlook

With this case study it was shown that the presented quality assurance workflow allows for an easy integration of additional requirements and tests to cover multiple modular plant configurations while enabling test-driven modeling and co-simulation. Especially for a large number of models, this approach allows for a test-driven and structured



assurance process which enables better result tracking. Since the partial models are compiled into a co-simulation framework, quality features outside of the models behavior must also be considered to assure successful model interaction. One future challenge will be the rigorous assessment of FMUs and composite FMUs. Since the person conducting quality assurance is not the model builder, acceptable quality of the individual models does not automatically lead to sufficient quality of the composite models if the modeling approach etc. is not known. Therefore, there is a need for an exchange of context information between manufacturer and operator of the co-simulation to ensure compatibility in addition to standard adherence like the modeling goal and some information about the modeling approach (Mädler et al., 2022).

Another challenge for the quality assurance of the co-simulation and the model process will be the definition and creation of the co-simulation configuration. As of now, the model assurance workflow does not include the quality assessment of the configurations content besides standard adherence itself.

## Acknowledgements

We would like to thank the German Federal Ministry of Education and Research and the Project Management Agency Jülich for their financial support within the framework of the eModule research project (FKZ 03HY116A) of the H2Giga lead platform.

## References

- Beisswenger, L., Erdmann, D., 2022. Standardized flowsheet for the alkaline electrolysis.
- Buttler, A., Spliethoff, H., 2018. Current status of water electrolysis for energy storage, grid balancing and sector coupling via power-to-gas and power-to-liquids: A review. *Renewable and Sustainable Energy Reviews* 82, 2440–2454. <https://doi.org/10.1016/j.rser.2017.09.003>
- International Energy Agency (IEA), 2021. *Global Hydrogen Review 2021*. OECD. <https://doi.org/10.1787/39351842-en>
- Mädler, J., Viedt, I., Lorenz, J., Urbas, L., 2022. Requirements to a digital twin-centered concept for smart manufacturing in modular plants considering distributed knowledge, in: Yamashita, Y., Kano, M. (Eds.), *Computer Aided Chemical Engineering, 14 International Symposium on Process Systems Engineering*. Elsevier, pp. 1507–1512. <https://doi.org/10.1016/B978-0-323-85159-6.50251-7>
- Mädler, J., Viedt, I., Urbas, L., 2021. Applying quality assurance concepts from software development to simulation model assessment in smart equipment, in: Türkay, M., Gani, R. (Eds.), *Computer Aided Chemical Engineering*. Elsevier, pp. 813–818. <https://doi.org/10.1016/B978-0-323-88506-5.50127-3>
- Murray-Smith, D.J., 2015. *Testing and Validation of Computer Simulation Models, Simulation Foundations, Methods and Applications*. Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-319-15099-4>
- Sargent, R.G., 2010. Verification and validation of simulation models, in: *Proceedings of the 2010 Winter Simulation Conference*. Presented at the 2010 Winter Simulation Conference - (WSC 2010), IEEE, Baltimore, MD, USA, pp. 166–183. <https://doi.org/10.1109/WSC.2010.5679166>
- VDI 2776-1:2020, 2020. *Verfahrenstechnische Anlagen - Modulare Anlagen - Grundlagen und Planung modularer Anlagen*.
- Viedt, I., Mädler, J., Lorenz, J., Urbas, L., 2022a. Requirements for the quality assessment of virtual commissioning models for modular process plants, in: *Computer Aided Chemical Engineering*. Elsevier, pp. 805–810. <https://doi.org/10.1016/B978-0-323-85159-6.50134-2>
- Viedt, I., Mädler, J., Urbas, L., 2022b. Quality assessment for dynamic, hybrid semi-parametric state observers, in: *Computer Aided Chemical Engineering*. Elsevier, pp. 1255–1260. <https://doi.org/10.1016/B978-0-323-95879-0.50210-1>

# Mechanistic modeling of industrial fermentation processes for antibiotic production

Atli F. Magnússon,<sup>a</sup> Stuart M. Stocks,<sup>b</sup> Jari P. Pajander,<sup>b</sup> Gürkan Sin,<sup>a</sup>

<sup>a</sup>*Technical University of Denmark, Søtofts Plads, 227, 2800 Kgs. Lyngby, Denmark*

<sup>b</sup>*LEO Pharma A/S, Industriparken 55, 2750 Ballerup, Denmark*  
*afmag@kt.dtu.dk*

## Abstract

An unstructured mechanistic model is proposed to describe the industrial-scale production of Fusidic Acid in fed-batch cultivations. The model accounts for differences in dead and viable biomass and the effect of the primary carbon source and oxygen on cell growth and production. The model parameter is calibrated, and performance is tested using experimental data obtained from an operating industrial production in Denmark. The model predicted the main product concentration with a relative mean error of 7%. These successful implementations open up opportunities for soft sensor implementations for key state variables but also set a good foundation for further model extensions, such as hybrid modeling.

**Keywords:** Fermentation, Process Modelling, Pharmaceuticals

## 1. Introduction

Industrial microbiology has been one of the most common methods for the mass reproduction of antibiotics since the widescale production of penicillin. In a very similar sense, Fusidic Acid is a secondary metabolite that is commercially produced using a filamentous microorganism. At the same time, there is extensive literature on the modeling of penicillin production with varying degrees of complexity (Goldrick et al. 2015), but there is no published research on the mechanistic modeling of the Fucidin process. The bioprocess industry has seen rapid advancements toward digitalization. This has led to increasing interest in industrial applications of digital twins. These technologies require good-quality process models to act as the foundation. Mechanistic models are the gold standard as they incorporate current knowledge into a first principles mathematical description of the system of interest (Gernaey et al. 2010). While more challenging to develop, these model types are notably beneficial for application in the fermentation industry since they can be better extrapolated to new scenarios than machine learning algorithms and artificial intelligence.

A mechanistic model describes the dynamic behavior of a system with a series of mathematical formulas, typically ordinary differential equations (ODEs). These models are based on prior knowledge of the system phenomena like mass, energy, and moment balances.

The mechanistic model of Bajpai and Reuss (Bajpai and Reuss 1980) is a good starting point when designing a model structure for a poorly researched organism. It's unstructured, has low complexity, few parameters, and can be calibrated based on

experimentally measured growth profiles. Furthermore, it was made to model the secondary metabolites of a filamentous organism and showed excellent agreement with experimental data. Extensions and changes will be proposed based on observed phenomena on industrial and lab scales to improve predictive qualities and applications. This study aims to design and evaluate a mechanistic model to quantitatively describe the behavior of an industrial-scale Fusidic Acid process. To this end, a mechanistic model structure is inspired by unstructured penicillin models with extensions or modifications to account for observable phenomena in the production setting and during lab experiments performed at LEO Pharma A/S. Data from industrial-scale fed-batch fermentation is collected to calibrate and validate model parameters and structure.

## 2. Materials and Methods

A model is developed to describe a filamentous fermentation fed-batch process operated at LEO Pharma A/S. The traditional fed-batch equation is utilized for the overall concentration balance with an extension to account for the effects of evaporation.

$$\frac{dC_i}{dt} = q_i + \frac{E}{M}C_i + \frac{F_{feed}}{M}(C_{i,f} - C_i) + k_l a_i(C_i^* - C_i)$$

Where the first term  $q_i$  denotes the biochemical kinetic rate of component  $i$ , the second term accounts for increased concentration due to mass loss from water evaporation and offgas balance, where  $E$  describes the broth mass change due to evaporation and offgas balance  $E = F_{evap} + F_{OUR} - F_{CER}$ . It is assumed that all relevant components are nonvolatile and thus are not present in the offgas. The third term describes the dilution due to feed where  $F_{feed}$  is the feed rate,  $M$  is the broth weight, and  $C_{i,f}$  is the concentration of component  $i$  in the feed. The final term describes the mass transfer of component  $i$  from the gas phase to the liquid phase, where  $k_l a_i$  is the mass transfer coefficient of component  $i$  and  $C_i^*$  is the equilibrium solubility of component  $i$ . For simplifications, it is assumed that  $k_l a$  is 0 for all components except oxygen.

As the model is unstructured, it is assumed that all the biochemical kinetics can be explained with a specific growth rate term, which can be either positive, indicating growth/production, or negative, indicating consumption.

$$\mathbf{q} = \boldsymbol{\mu}X$$

Where  $X$  is the concentration of viable biomass and  $\boldsymbol{\mu}$  is a vector containing the specific rates of each species calculated with an unstructured model equation structure. Note that  $X$  is viable biomass and not the total dry weight, which accounts for the accumulation of cell debris and precipitated main product.

$$X_{TDW} = X + X_D + P_{precipitated}$$

The Contois model kinetics is utilized for growth due to the excellent agreement with experimental data for multiple microbial systems (Alqahtani et al. 2015). The overall system of equations describing the biochemical model of the main components are

$$\mu_X = \mu_{X,max} \frac{S}{K_{SX}X + S} \frac{DO}{K_{OX}X + DO} \left( 1 - \exp\left(\frac{-t}{t_{lag}}\right) \right) - k_d$$

$$\mu_P = \mu_{P,max} \frac{S}{K_{SP}X + S}$$

$$\mu_S = \frac{\mu_X}{Y_{SX}} + \frac{\mu_P}{Y_{SP}} + m_S \frac{S}{K_{SS}X + S}$$

The changes in broth mass are modelled by considering all components entering and leaving the fermenter

$$\frac{dM}{dt} = F_{feed} - F_{evap} + F_{OUR} - F_{CER}$$

Feedrate and offgas measurements are available. To estimate the evaporation, we consider the amount of water vapors in the process air and offgas.

$$F_{evap} = Q_{air} \left( \frac{\phi_{out} p_{H_2O,out}^*}{RT_{out}} - \frac{\phi_{in} p_{H_2O,in}^*}{RT_{in}} \right)$$

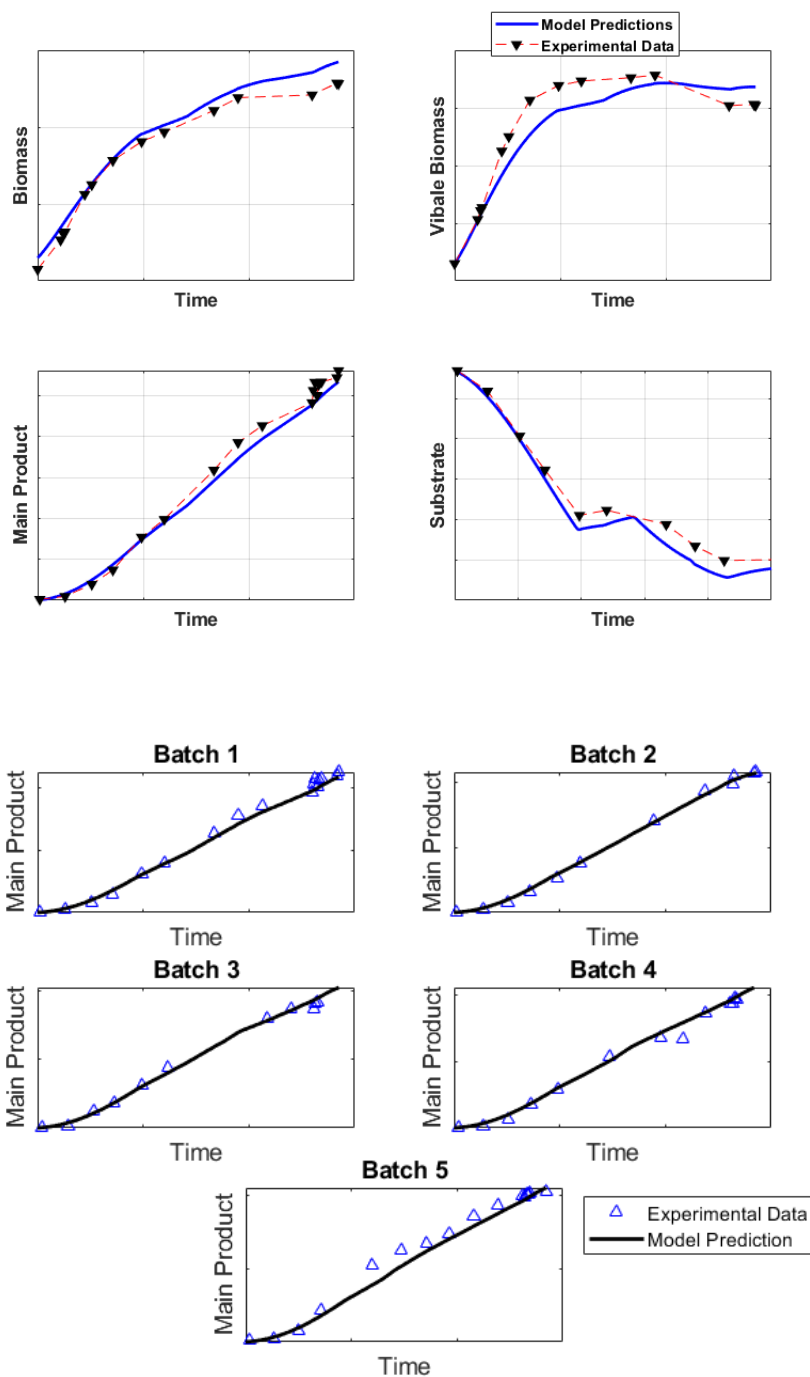
Where  $\phi$  is the relative humidity and  $p_{H_2O}^*$  is the saturation pressure of water in the air. Relative humidity of process air is available from batch measurements, and it is assumed that the offgas is fully saturated with water, i.e.,  $\phi_{out} = 1$ . The saturation pressure of water can be estimated using a steam table or the Antoine equation.

The model equation structure is implemented into MATLAB R2021B and solved using the stiff solver *ode15s*. Measured online data for Dissolved Oxygen (*DO*), pH and offgas  $F_{OUR}$  and  $F_{CER}$  are used as inputs when model is calibrated. The model contains 11 growth parameters that need to be estimated, no previous nominal values exist for the fungus and are all estimated simultaneously using *lsqnonlin* function in MATLAB.

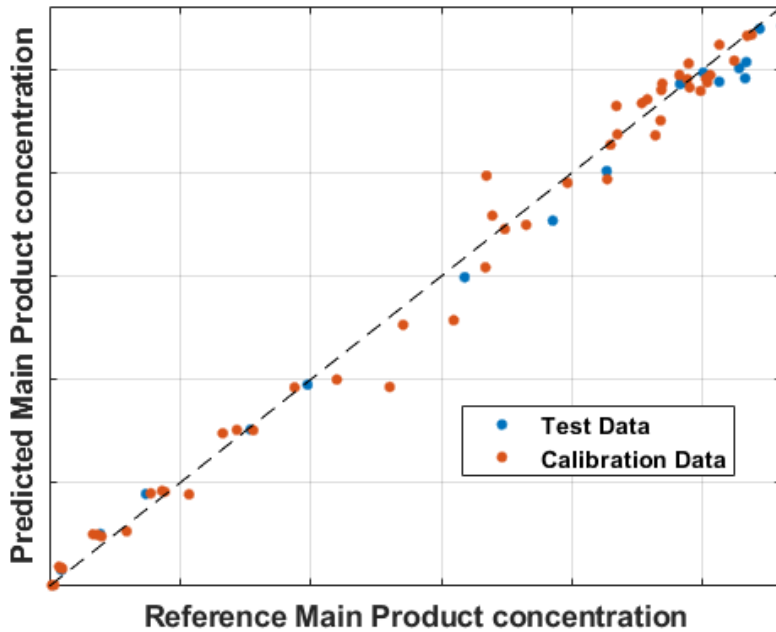
### 3. Results

Experimental data collected from sampling industrial production are used for parameter estimation. An initial fit for the validation batch can be seen in Figure 1. The prior estimation showed a generally satisfactory fit with experimental data and can accurately

describe the concentration profiles of viable biomass and main product. It also captured the concentration of the primary carbon source reasonably.



However, for engineering purposes, we are primarily interested in the main product concentration, which is the main economic driver of the entire process. Therefore, the comparison of measured and simulated growth profiles of all sampled batches is shown in Figure 2. The overall fit to the main product considering all experimental values is depicted in figure 3 with an overall prediction error of 6.6%. This Model shows excellent fits for the main product.



Model quality was assessed with the root mean sum of squared errors (RMSSE) but is reported here as a scaled percentage deviation from the experimental measured mean value to preserve confidentiality of process scales of industrial sponsor.

$$RMSSE(\%) = 100 * \frac{\sqrt{\frac{1}{n} \sum_{i=1}^N (y_{meas,i} - y_{pred,i})^2}}{\bar{y}}$$

Process Variable	RMSSE (%)	R <sup>2</sup>
Total Dry Weight	12.75	0.98
Viable Biomass	12.99	0.91
Main Product	6.60	0.99
Substrate	19.52	0.98

Initial parameter estimation shows that the Model has the most trouble predicting main carbon source concentrations. If all measurements are included, there is an error of approximately 20%. It’s hypothesized that the complex media blend gives rise to multiple available carbon sources. At the same time, the Model focuses only on a singular primary carbon source. This may lead to different consumption rates depending on the variation

in the initial media blending. Both biomass measurements show a similar error of approximately 13% which is considered a good fit. Similarly, with the substrate predictions, it may be possible to improve fits by considering the effects of potential alternative carbon sources, but this is a highly complicated data collection and modeling process. One of the main drawbacks of using the current iteration for batch planning or as the basis for digital twins is that the model cannot run as an independent simulation. This is because the model relies on Oxygen Uptake and Carbon Evolution data to estimate evaporation accurately. Without this information, any optimization that relies on model outputs will probably underestimate final product concentrations while overestimating the current broth weight. In the current state, the mechanistic model uses batch measurements to calculate estimates of key process variables that describe the batch states that are not observable meaning it has potential applications in soft-sensor technology.

#### 4. Conclusions

A mechanistic model structure was proposed to describe an antibiotic production process of a novel filamentous fungi strain. The model can be used to simulate growth profiles on an industrial scale. It can simulate biomass growth, main product synthesis, and substrate uptake. Due to the lack of sophisticated hardware sensor technology, these key process variables are not commonly measured online. The predictions are supported by readily available online measurements such as offgas mass spectrometry to evaluate carbon evolution and oxygen uptake. Future work will use this model as a foundation for integrating machine learning techniques to expand the model capabilities to obtain readily available predictions of the missing key state variables that went unexplained in this work. We also wish to study the model's reliability in the framework of good modeling practices (Sin *et al.* 2009) by statistically analyzing the effects of uncertainty of the model structure and parameters.

#### References

- S. Goldrick, A. Ştefan, D. Lovett, 2015, The development of an industrial-scale fed-batch fermentation simulation, *Journal of Biotechnology*, 193, 70-92
- K. V. Gernaey, A.E. Lantz, P. Tufvesson, J. M. Woodley, G. Sin, 2010, Application of Mechanistic Models to Fermentation and Biocatalysis for next-Generation Processes, *Trends in Biotechnology*, 28, 7, 346-354
- R. K. Bajpai, M. Reuss, 1980, Mechanistic Model for Penicillin Production, *Journal of chemical technology and biotechnology*, 30, 6, 332-344
- R. T. Alqahtani, M. I. Nelson, A. L. Worthy, 2015, A biological treatment of industrial wastewaters: Contois kinetics, *Anziam Journal*, 55, 4, 397-415
- G. Sin, K. V. Gernaey, A. Elliasson Lantz, 2009, Good Modeling Practice for PAT Applications: Propagation of Input Uncertainty and Sensitivity Analysis, *Biotechnology Progress*, 25, 4, 1043-1053

# A new framework and online solution engines for multiparametric Model Predictive Control

Diogo A.C. Narciso<sup>a,b\*</sup>, Dustin Kenefake<sup>c,d</sup>, Sahithi Srijana Akundi<sup>c,d</sup>, F.G. Martins<sup>a,b</sup> and Efstratios N. Pistikopoulos<sup>c,d</sup>

<sup>a</sup>*LEPABE - Laboratory for Process Engineering, Environment, Biotechnology and Energy, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal*

<sup>b</sup>*ALiCE - Associate Laboratory in Chemical Engineering, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal*

<sup>c</sup>*Texas A&M Energy Institute, Texas A&M University, College Station, TX 77843, USA*

<sup>d</sup>*Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77843, USA*

*dnarciso@fe.up.pt*

## Abstract

In this paper, we combine the multiparametric programming approach proposed in Pistikopoulos et al., 2002 for quadratic Model Predictive Control (mp-QP/MPC) problems, with the solution strategy proposed in Narciso et al. 2022 for mp-QP problems. The problem statements in these two contributions are fully compatible, and we show firstly how they can be combined to deliver critical regions and the corresponding control laws of MPC problems in the offline stage. Then, a new online architecture is proposed for the calculation of optimal control actions. To this end we explore the convenient solution structure presented in Narciso et al. 2022 with the specific aim to calculate them as fast as possible via parallel processing, and where a restricted collection of critical regions is kept and continuously updated in time.

**Keywords:** Multiparametric programming, Quadratic Programming, Model Predictive Control.

## 1. Introduction

Model Predictive Control (MPC) enables an optimization-based approach to control problems, whereby a cost function driving the system's states to a predefined set point and a set of operational constraints are defined (Pistikopoulos et al. 2020). As the size of MPC problems increase, so does the total time required to solve them. This is particularly critical when the total times required to solve MPC problems exceed the required response times, which renders the optimization-based route ineffective at best (Pappas et al. 2020).

Multiparametric programming (mp) has been shown to mitigate this limitation (Pistikopoulos et al., 2002): the expensive optimization calculations are moved to the *offline* stage, where a collection of explicit critical regions and matching control laws are obtained, and thus requiring a set of inexpensive function evaluations during the *online* calculation of the optimal control actions. The drawback of mp is that as MPC problems grow very large, so do the corresponding explicit solutions, which in turn generally slow down the calculation of optimal controls in the online stage.



Several strategies have been proposed to deliver faster online calculations for MPC problems (Pappas et al. 2020). In this work, we focus on the class of mp-QP/MPC (quadratic programming) problems presented in Pistikopoulos et al., 2002, and use the mp-QP algorithm presented Narciso et al. 2022 to deliver critical regions and control laws in compact form during the offline stage. We explore this convenient solution structure and propose a new architecture for the fast calculation of controls during the online stage.

## 2. Offline calculation of optimal control laws

We consider the following class of QP/MPC problems, where the state, output, and control vectors are denoted as  $x \in \mathbb{R}^n$ ,  $y \in \mathbb{R}^p$  and  $u \in \mathbb{R}^m$  (Pistikopoulos et al., 2002):

$$\begin{aligned}
 \min_U \quad & x'_{t+N_y|t} P x_{t+N_y|t} + \sum_{k=0}^{N_y-1} [x'_{t+k|t} Q x_{t+k|t} + u'_{t+k} R u_{t+k}] \\
 \text{s. t.} \quad & y_{\min} \leq y_{t+k|t} \leq y_{\max}, k = 1, \dots, N_c \\
 & u_{\min} \leq u_{t+k|t} \leq u_{\max}, k = 1, \dots, N_c \\
 & x_{t|t} = x(t) \\
 & x_{t+k+1|t} = A x_{t+k|t} + B u_{t+k}, k \geq 0 \\
 & y_{t+k|t} = C x_{t+k|t}, k \geq 0 \\
 & u_{t+k|t} = K x_{t+k|t}, N_u \leq k \leq N_y
 \end{aligned} \tag{Eq. 1}$$

where  $U = [u'_t, u'_{t+1}, \dots, u'_{t+N_u-1}]' \in \mathbb{R}^s$ , with  $s = mN_u$ .  $P = P' \geq 0$ ,  $Q = Q' \geq 0$  and  $R \geq 0$  are constant cost matrices.  $A$ ,  $B$  and  $C$  are constant matrices defining the state-space model, and subscripts min and max denote the lower and upper bounds, respectively, on  $y$  and  $u$ .  $K$  is a constant feedback gain matrix, and  $N_y \geq N_u$ . Using  $x_{t+k|t} = A^k x(t) + \sum_{j=0}^{k-1} A^j B u_{t+k-1-j}$ , an equivalent QP/MPC is defined:

$$\begin{aligned}
 \min_U \quad & 1/2 U' H U + x(t)' F U + 1/2 x(t)' Y x(t) \\
 \text{s. t.} \quad & G U \leq W + E x(t)
 \end{aligned} \tag{Eq. 2}$$

where  $H \geq 0$ , and all additional matrices are obtained from Equation 1 and  $x_{t+k|t} = A^k x(t) + \sum_{j=0}^{k-1} A^j B u_{t+k-1-j}$ . Using the variable transform,  $\gamma = U + H^{-1} F' x(t)$ , and defining the allowable space of state variables (space of independent parameters), a mp-QP/MPC problem is obtained accordingly:

$$\begin{aligned}
 \min_{\gamma} \quad & 1/2 \gamma' Q \gamma \\
 \text{s. t.} \quad & G \gamma \leq W + S x(t) \\
 & P_A x(t) \leq P_b
 \end{aligned} \tag{Eq. 3}$$

where  $\gamma \in \mathbb{R}^s$  and  $S = E + G H^{-1} F'$ . Multiple routes are available to solve Eq. 3, namely via the explicit calculation of critical regions of the state space and the corresponding control laws (Pappas et al. 2020). In this work we consider the solution strategy for mp-QP problems presented in Narciso et al. 2022, where a final recast step is required. Note that a new transformed vector of parameters  $z = S x(t)$  is defined, and no bounds are enforced on the corresponding space of parameters  $Z$ :

$$\begin{aligned} \min_{\gamma} \quad & 1/2 \gamma' Q \gamma \\ \text{s. t.} \quad & G\gamma \leq W + z \end{aligned} \quad (\text{Eq. 4})$$

In this convenient format, it is possible to deliver the map of all critical regions (Eq. 5) and the corresponding optimal control laws (Eq. 6) in compact form for all active sets, as follows:

$$z = z^* + V_i^z \left( (\alpha^{full} - \alpha) l \right) + V_a^z(\alpha l), \alpha \in \{0,1\}, l \geq 0 \quad (\text{Eq. 5})$$

$$\gamma^{opt} = \gamma^* + V^\gamma(\alpha l) \quad (\text{Eq. 6})$$

where  $\alpha$  is the binary active set vector of Eq. 4 with size  $q = (p + m)N_c$ , and such that 0/1 denote inactive/active constraints, respectively.  $\alpha^{full}$  denotes the full active ( $\alpha^{full} = [1, \dots, 1]'$ ). Optimizers are denoted as  $\gamma^{opt}$ , and  $l$  is a non-negative multiplier vector of size  $q$ . All vectors ( $z^*$ ,  $\gamma^*$ ) and matrices ( $V_i^z$ ,  $V_a^z$ ,  $V^\gamma$ ) in Eqs 5 and 6 are obtained from the Karush-Kuhn-Tucker (KKT) conditions on all the  $q$  single-constraint active sets.

An algorithm and discussion on the significance of these vectors/matrices is presented in Narciso et al. 2022. For any given active set  $\alpha$ , it suffices to set it directly in Eqs. 5 and 6 to obtain the corresponding critical region in the transformed parameter space and optimal control law, respectively; the explicit calculation of critical regions may also optionally be enabled, where  $V_{\{\alpha\}}^z$  includes the relevant column vectors from  $V_i^z$  and  $V_a^z$  for any  $\alpha$ :

$$z = z^* + V_{\{\alpha\}}^z l, l \geq 0 \quad (\text{Eq. 7})$$

### 3. Online calculation of optimal control actions

To obtain the optimal control actions for any  $x(t)$ , 4 steps are required:

- 1) Compute  $z^{input} = Sx(t)$ ;
- 2) Solve Eq. 5 for  $z = z^{input}$  to obtain  $(\alpha, l)$ ;
- 3) Compute  $\gamma^{opt}$  from Eq. 6;
- 4) Compute the first set of control actions ( $u_t$ ) from  $U = \gamma^{opt} - H^{-1}F'x(t)$ .

Steps 1, 3 and 4 are inexpensive matrix/vector operations. Step 2, on the other hand, is a less trivial problem: when solving Eq. 5 for a given  $z$ , such statement defines in fact a Linear Complementarity Problem (LCP) (Murty and Yu, 2010). Using one of the state-of-the-art LCP solvers is the obvious choice for this step, but in fact the convenient structure of Eq. 5 enables other alternative routes for its solution in the context of mp-QP/MPC, namely via: (i) the explicit calculation of critical regions/control laws (as is in fact the common practice in the field), or (ii) recast Eq. 5 as an NLP or MILP problem.

All routes above are equivalent with respect to the solutions obtained from them. In a control context, however, not only are we concerned with finding the optimal control actions, but also to deliver them as fast as possible to avoid any undesirable delays and the consequent loss of control efficiency (Pistikopoulos et al. 2020). We propose the following procedure to calculate  $(\alpha, l)$  via step 2 as depicted in Figure 1.

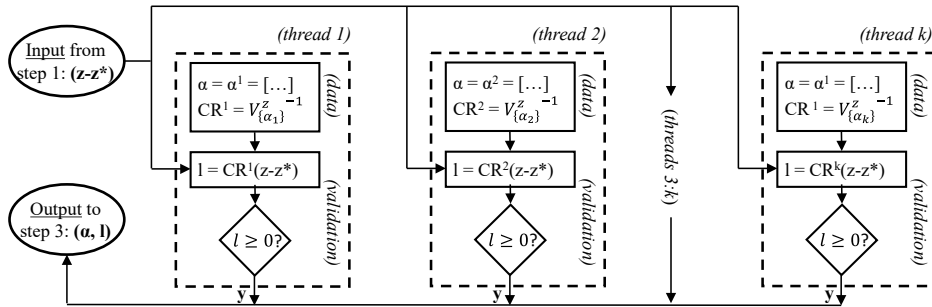


Figure 1: Online calculation of  $(\alpha, l)$  via parallel processing of all critical regions.

From Eq. 7, critical regions are obtained for all  $k \leq 2^q$  relevant active sets of Eq. 4 as  $V_{\{\alpha\}}^z^{-1}$ , which we refer to as the critical region *identifier* for  $\alpha$ . In this convenient solution format, it suffices to compute  $l = V_{\{\alpha\}}^z^{-1}(z - z^*)$ , and check which critical region satisfies  $l \geq 0$ , to identify  $\alpha$  for any given  $z = z^{input}$ . Since all identifiers are tested in *parallel* in Figure 1, this amounts in fact (and in theory) to a very fast calculation of  $(\alpha, l)$ .

In MPC problems containing many constraints, the scheme depicted in Figure 1 may become prohibitive, since the online controller requires that up to a total of  $2^q$  threads are created to store and process all active sets. A modified architecture is proposed where  $k \ll 2^q$ ; this avoids the large memory/processing requirements, while preserving the speed and effectiveness of the online controller. This architecture is shown below:

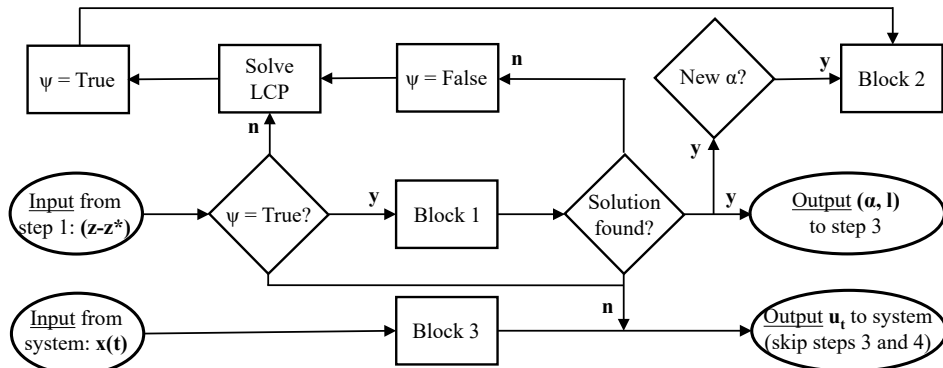


Figure 2: Online control architecture (step 2).

The middle row of tasks in Figure 2 captures the *nominal* mode of operation, where the pair  $(\alpha, l)$  is calculated via Block 1: it shares the architecture depicted in Figure 1, but where a much smaller number of *candidate* active sets and their identifiers are kept in memory (details below). Whenever a new  $\alpha$  is found between consecutive sampling times, an update of the  $k$  threads is executed in Block 2 to ensure the candidate active sets and their identifiers remain relevant as the vector of states is steered towards the set-point.  $\psi$  is a Boolean variable providing information if the current set of identifiers kept in memory are up to date and managing the calculation of controls accordingly.

Since only a small fraction of all active sets is kept in memory, it is possible that the matching active set for a given  $x(t)$  is not found for a particular  $t = t$ . In this case, a *remedy* control action is implemented via Block 3, using a conventional feedback controller (e.g. PID). When this occurs, Block 1 becomes unavailable until the optimal solution is obtained, which requires the solution of a more expensive LCP problem.

The structure of active sets and their identifiers kept in the  $k$  threads are as follows. As a minimum, a single thread must be defined: if an active set  $\alpha$  is the optimal solution at  $t = t$ , the unique defined thread includes  $\alpha$  and the matching identifier at  $t = t + 1$ . Using a single thread/active set, however, would most likely frequently trigger solving the LCP problem and using the feedback controller, as it is expected that the optimal active set does not remain constant. To strengthen the predictive capability of Block 1, additional threads are defined: this includes all active sets matching  $\alpha$ , except for up to  $r$  coordinates, where  $r$  is a design parameter. The total number of threads is given by the number of combinations  $C(q, 0) + C(q, 1) + \dots + C(q, r) = k$ , where  $k \geq 1$  (when  $r = 1$ ) and  $k \leq 2^q$  (when  $r = q$ ). When a new  $\alpha$  is found, Block 2 updates the full set of active sets in all threads, taking  $\alpha$  as a reference and updating selectively their coordinates:

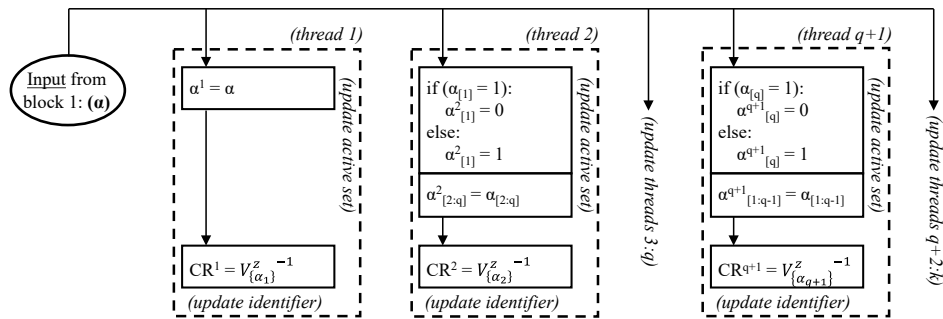


Figure 3: Threads' update of in step 2 (Block 2).

Parameter  $r$  defines a *compromise* between the computational complexity (number of active sets and identifiers kept in memory) and the predictive capability of Block 1. In stable systems, it is expected that the vector of states undergoes a smooth trajectory, and where the optimal active set frequently remains constant between consecutive times, or where a restricted number of constraints change status. Therefore, a careful tuning of  $r$  facilitates an efficient route for step 2, with a minimal fraction of time spent on solving LCPs. The most frequent or nominal control tasks are highlighted below:

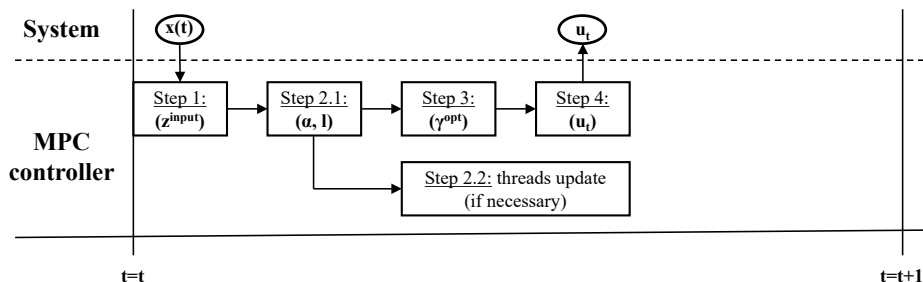


Figure 4: Control tasks in the nominal operation mode for any given  $t = t$ .

Steps 1-4 are all executed in sequence. Step 2 is broken down in two stages where steps 2.1 and 2.2 correspond to Blocks 1 and 2, respectively. Note that the threads update via step 2.2 is launched in parallel (when necessary) with steps 3 and 4 to speed up this task.

#### 4. Discussion and future work

We have shown in Section 2 that combining the mp-QP/MPC problem formulation in Pistikopoulos et al. 2002 with the mp-QP solution strategy in Narciso et al. 2022 is a trivial task, to deliver critical regions and control laws in compact form.

While this is a significant gain from the offline perspective, even more important is the ability to calculate optimal control actions very fast. With this goal in mind, a new architecture for online calculations was proposed. It makes use of the convenient structure of Eq. 5, to keep a restricted collection of active sets and their critical region identifiers. Parallel computing is used to compute the optimal control actions from the states vector via a sequence of inexpensive matrix/vector multiplications.

Several improvements to this architecture have been identified for future work, and which we discuss briefly here. Firstly, we remark that this architecture includes no provisions to deal with infeasible solutions: failing to obtain a solution from Block 1 is not necessarily a result of the identifiers stored in memory not being sufficient, but simply because no feasible solution exists for the MPC problem with a given  $x(t)$ . Additionally, to the explicit calculation of identifiers, the framework in Narciso et al. 2022 also allows the calculation of infeasible regions, which can be used for a more efficient control strategy.

A maximum of  $2^q$  critical region identifiers may be obtained from Eq. 5 since the parameter space ( $Z$ ) for  $x(t)$  is defined unbounded. This choice is made for the generality of the solution obtained in Narciso et al. 2022. In the present context, where the explicit identifiers are kept in memory, an additional check of  $Z$  may simplify the online control problem: checking for instance which constraints are always inactive or pre-calculating subspaces in  $Z$  where a constraint can be active reduces the online complexity.

More broadly, this research aims at implementing the fastest online strategy for MPC problems. Additional developments on the solution of LCPs in this specific context is of great significance. Two LCP solvers are currently under development for this class of problem. An assessment of the performance of all the viable routes on computing control actions for MPC problems of varying sizes is a key research objective. It is expected that this work can facilitate establishing several guidelines for practical applications.

#### References

- E. N. Pistikopoulos, N. A. Diangelakis and R. Oberdieck, 2020, Multi-parametric Optimization and Control, ISBN: 978-1-119-26518-4.
- I. Pappas, D. Kenefake, B. Burnak, S. Avraamidou, H. S. Ganesh, J. Katz, N. A. Diangelakis and E. N. Pistikopoulos, 2021, Multiparametric Programming in Process Systems Engineering: Recent Developments and Path Forward, *Front. Chem. Eng.*, 21.
- E. N. Pistikopoulos, V. Dua, N. Bozinis, A. Bemporad, M. Morari, 2002, On-line optimization via off-line parametric optimization tools, *Comput. Chem. Eng.*, 26, 175-185.
- D. A. C. Narciso, I. Pappas, F. G. Martins and E. N. Pistikopoulos, 2022, A new solution strategy for multiparametric quadratic programming, *Comput. Chem. Eng.*, 164, 107882.
- K. Murty, F. Yu, 2010. Linear Complementarity, *Linear and Non Linear Programming*. [http://www-personal.umich.edu/murty/books/linear complementarity webbook/](http://www-personal.umich.edu/murty/books/linear%20complementarity%20webbook/).

# Comparison of regression techniques for generating surrogate models to predict the thermodynamic behavior of biomass gasification systems

Meire Ellen Gorete Ribeiro Domingos,<sup>a,b</sup> Daniel Flórez-Orrego,<sup>b</sup> Julia Granacher,<sup>b</sup> Marie Jones,<sup>b</sup> Moisés Teles dos Santos,<sup>a</sup> François Maréchal<sup>b</sup>

<sup>a</sup>*Polytechnic School, University of Sao Paulo, Department of Chemical Engineering, Sao Paulo, Brazil.*

<sup>b</sup>*École Polytechnique Fédérale de Lausanne, Valais, Sion, Switzerland*  
*meire.ribeirodomingos@epfl.ch*

## Abstract

Biomass resources can play an important role in the energy transition, being a dynamic feedstock that can be transformed either into solid, liquid or gaseous fuels. Biomass gasification is a versatile way to convert waste into energy. In this work, the modeling and simulation of two different gasification processes using wood and black liquor as feedstock are performed using Aspen Plus®. The surrogate models for these biomass-based gasification systems are generated considering different techniques (e.g. artificial neural networks, random forest, and Gaussian process regression) using an Active Learning Artificial Intelligence approach. These techniques are compared in terms of their capabilities for predicting the thermodynamic behavior of the gasification systems for the different biomass resources. As a result, the surrogate models developed were able to estimate the process design and operating conditions, and the Gaussian process regression outperformed the artificial neural networks and random forest techniques. The generated models could be helpful to be further used for replacing the simulation systems in other applications, such as multi-objective optimization, at expense of lower computational requirements.

**Keywords:** Surrogate models, Biomass, Gasification, Artificial intelligence.

## 1. Introduction

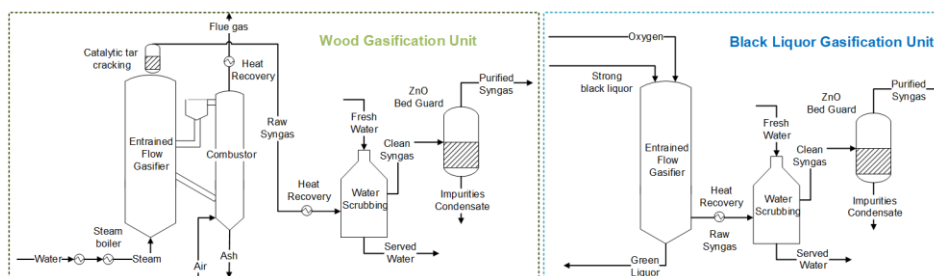
The gasification process poses modeling complexities involving several stages that occur simultaneously, such as drying process, and pyrolysis, gasification and combustion reactions. These reactions are highly endothermic and the energy balance becomes a function of the combustion and the reduction steps. In addition, the production rate, composition and properties of the syngas produced strongly vary according to the biomass properties, the gasification agent used (type and flow), the equivalence ratio, the type of reactor used (fixed bed, fluidized bed or indirect gasifier), the operating conditions (temperature, pressure, residence time) and the downstream processes considered for gas cleaning and conditioning. Numerous models have been reported, aiming at understanding and describing the behavior of such systems, using commercial simulation software. However, if integrating these complex processes into industrial systems and

performing optimization on the system level, the direct simulation approach entails high computational expenses. Surrogate models emerge as a robust solution that may help accelerate the convergence of the optimization problem (Granacher et al., 2021). This work proposes the generation of surrogate models for two gasification systems, considering wood and black liquor as feedstock, using different regression techniques applying the Active Learning Artificial Intelligence (ALAI) approach (Granacher et al., 2021). The models are compared to demonstrate their capability in predicting the processes behavior.

## 2. Methodology

### 2.1. Modeling and simulation of the gasification systems

Figure 1 illustrates the wood and black liquor gasification setups based on (Domingos et al., 2022, Flórez-Orrego et al., 2019). The simulations are performed in the Aspen Plus® software (Aspentech, 2015), using the Peng-Robinson EoS with Boston-Mathias modifications as thermodynamic model.



**Figure 1.** Wood and black liquor gasification systems.

The ultimate mass-based wood composition is set to 48.2% C, 5.8% H, 45% O, 0.2% N, 0.02% S and 0.8% ashes, whereas the mass-based proximate analysis is considered as 10% moisture (after drying), 16.4% fixed carbon, 82.8% volatiles, and ash in balance (Telmo et al., 2010). The Battelle Columbus Laboratory (BCL) indirect gasifier operating at atmospheric pressure and using steam as gasification medium is used for modeling the wood gasification system shown in Fig. 1. This system avoids the dilution with nitrogen of the syngas produced, as the combustion and gasification processes occur in a separate double column system (Kinchin, 2009). After the produced syngas leaves the gasifier, a thermal catalytic cracking of the produced tar is performed.

The black liquor ultimate composition (mass) is set as 29.86% C, 3.27% H, 29.05% O, 0.1% N, 4.09% S, 0.90% Cl and 32.73% ashes, whereas proximate analysis (mass) is considered as 15% moisture (after drying), 10.21% fixed carbon, 57.06% volatiles, and ash in balance. The pressurized entrained flow high temperature black liquor gasification (PEHT-BLG) available in the market as Chemrec® gasifier (Consonni et al., 2009) is used for modeling the black liquor gasification system. This technology has the advantage of including tar-free and uniform temperature gasification at turbulent conditions which results in a high conversion efficiency and high reaction rate. In the gasification step, the carbonaceous materials in the BL are converted into syngas rich in CO, H<sub>2</sub>, H<sub>2</sub>O and CO<sub>2</sub> using oxygen as agent, while a smelt containing mainly Na<sub>2</sub>CO<sub>3</sub> and Na<sub>2</sub>S is formed in the bottom when a quench cooler separates the regenerated salts that will form the green liquor (Darmawan et al., 2017).

In both systems, a fraction of the char produced in the pyrolysis step is combusted to supply the heat required by the endothermic drying, pyrolysis and reduction reactions.

After the produced syngas leaves the gasifier, it is cooled down and scrubbed with water, in order to remove the impurities that may affect the downstream equipment.

### 2.2. Surrogate model design

The algorithm used to generate and test the surrogate models considering the ALAI approach (Granacher et al., 2021) is presented in Fig. 2. The simulation systems developed in Aspen Plus® (Aspentech, 2015) are used to retrieve the dataset comprising simulation inputs and outputs, taking into account the decision variables selected. The quasi-random Latin-Hypercube sampling (McKay et al., 1979) algorithm is applied for sampling the design space  $D$ . The dataset is then used to generate and train the surrogate models, where the uncertainty with which a prediction is made is quantified. The active learning is responsible for continuously enriching the database, and new data points are selected from a pool of unlabeled data based on the predicted uncertainty. This strategy results in a fast improvement of the surrogates' quality. When the desired quality is met following the defined performance metrics, the surrogate model is saved (Granacher et al., 2021).

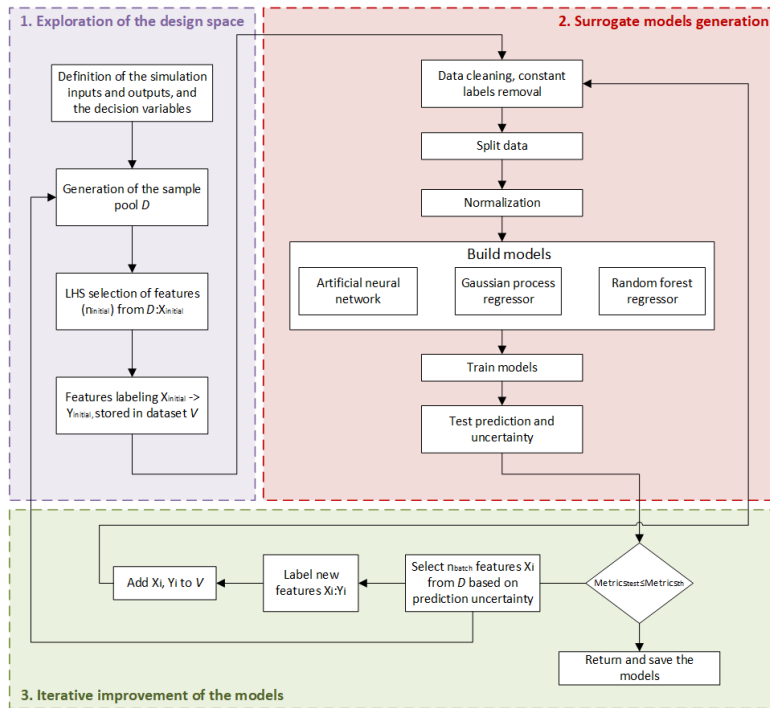


Figure 2. Algorithm for generation of the surrogate models, adapted from (Granacher et al., 2021).

In this work, three regression techniques, namely Gaussian process regression, Artificial neural networks (ANN) and Random Forest regression, are assessed and compared to generate surrogate models for the wood and black liquor gasification systems. In brief, the Gaussian process regression is a nonparametric, Bayesian approach to regression that makes predictions with uncertainty (Rasmussen and Williams, 2006). This method calculates the probability distribution over all admissible functions that fit the data instead of returning only exact predictions for each. The ANN is inspired by the biological nervous system and it consists of several layers connected by neurons that perform the input data transformations aiming to obtain the required outputs (Silva et al., 2016).



Finally, Random Forest regression is a supervised learning algorithm that combines predictions from multiple algorithms to make a more accurate prediction compared to a single model (Breiman, 2001).

For each gasification system, three parameters were chosen to be varied based on the operating conditions reported in the literature for these setups (Larsson et al., 2006, Puig-Arnabat et al., 2012). The parameters and their respective range are reported in Table 1. All the important simulation inputs, outputs, and operating conditions were set to be extracted to compose the dataset for building the surrogate modes.

**Table 1.** Operating conditions varied for wood and black liquor gasification systems.

Wood gasification		Black liquor gasification	
Parameter	Range	Parameter	Range
Gasification temperature (°C)	[760,960]	Gasification temperature (°C)	[900,1100]
Air preheating (°C)	[300,500]	Gasification pressure (bar)	[25,32]
Steam-to-biomass ratio	[0.75±30%]	Oxygen-to-fuel ratio	[0.2,0.5]

In order to measure the quality of the surrogate model, a random set of additional points is generated for testing. The new simulation points are compared against the surrogate model's response using error-based performance metrics. The mean squared error (MSE) (Eq.1) measures the average squared difference between the estimated values coming from the response of the surrogate models and the actual simulation value.

$$MSE = \frac{\sum_{i=1}^n (x_{simul,i} - x_{surr,i})^2}{n} \quad (1)$$

where  $x_{simul,i}$  is the simulation output and  $x_{surr,i}$  is the surrogate response for each parameter  $i$ , and  $n$  is the number of points in the data set. In order to facilitate the comparison between the dataset with different scales, the simulation data is normalized.

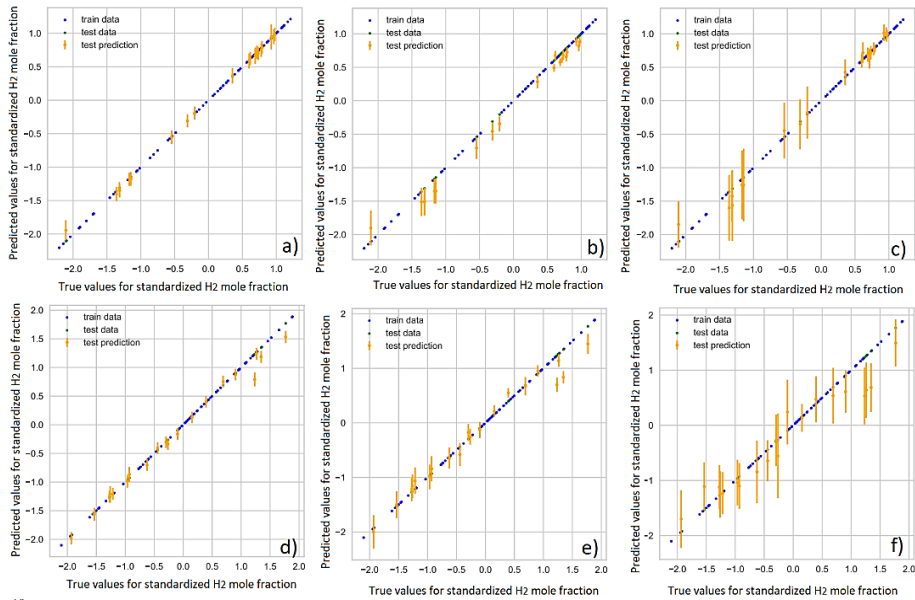
### 3. Results and discussion

Table 2 presents the mean squared errors in absolute dimensions for the surrogate models generated considering the three regression techniques. It can be seen that the Gaussian Process regression outperformed the artificial neural networks and the random forest approaches, since it presented the lowest MSE.

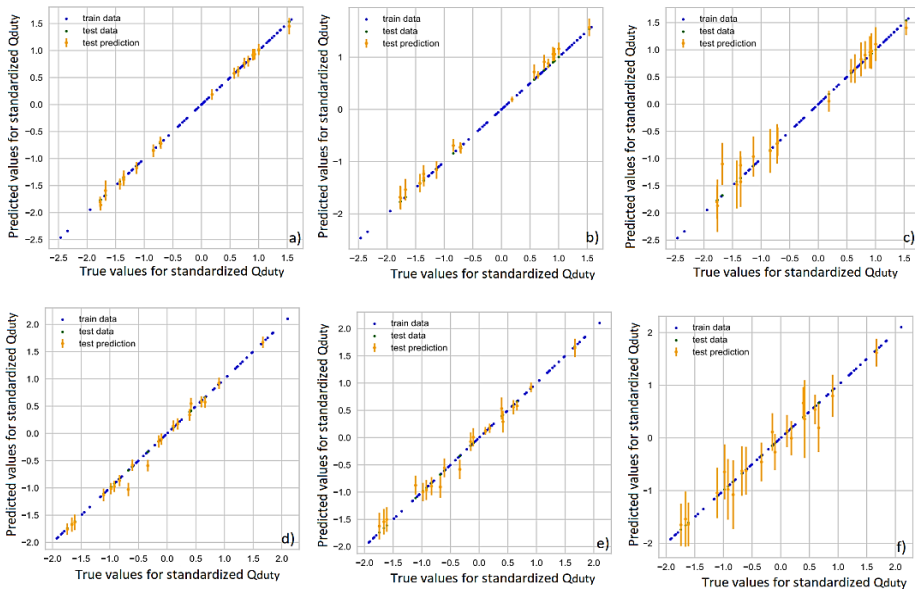
**Table 2.** Mean squared error (MSE<sub>test</sub>) considering 100 data points, absolute dimensions.

Regression model	MSE <sub>test</sub> - wood gasification	MSE <sub>test</sub> - black liquor gasification
Gaussian Process regression	0.02459	0.01498
ANN	0.02696	0.06296
Random forest	0.03061	0.05464

In Fig. 3, the predictions are displayed against the true values, with their respective uncertainty (see the orange bars in Fig. 3), for the hydrogen mole fraction in the purified syngas stream for the two gasification systems. The results demonstrate that the Gaussian process regression technique resulted in lower uncertainty, related to the prediction capability of the surrogate model considering the syngas composition. The same behavior was observed for the prediction of the heat exchanger duty related to the 'Raw Syngas' stream compared to the true values retrieved from the simulation (see Fig. 4).



**Figure 3.** Test predictions and uncertainty considering 100 data points for hydrogen mole fraction in the syngas from wood gasification using a) Gaussian regression, b) ANN, c) random forest, and from black liquor gasification using d) Gaussian regression, e) ANN, f) random forest.



**Figure 4.** Test predictions and uncertainty considering 100 data points for the heat recovery heat exchanger in the ‘Raw Syngas’ stream for wood gasification using a) Gaussian regression, b) ANN, c) random forest, and for black liquor gasification using d) Gaussian regression, e) ANN, f) random forest.

The comparison between the predictions of the surrogate models against the simulation results were also performed for the other components of the streams as well as all other process design and operating conditions. The results pointed towards a good capability of

the surrogate models in predicting the process design and operating conditions for the studied gasification systems.

#### 4. Conclusion

In this work, three regression techniques were compared to generate surrogate models applying the ALAI approach for wood and black liquor gasification systems simulated in the Aspen Plus software. The Gaussian process regression outperformed the artificial neural networks and random forest techniques for both analyzed processes, achieving lower mean squared error and associated uncertainties. Overall, the surrogate models were able to accurately predict the thermodynamic performance and the high nonlinear behavior associated to both gasification systems. Additionally, the derived surrogate models are computationally less expensive and can be further used to explore the tradeoffs between conflicting objectives, such as operating costs and syngas composition.

#### 5. Acknowledgements

MD and MJ thank European Union's Horizon Europe Research and Innovation programme under Grant Agreement No. 101084288 and No. 945363.

#### References

- Aspentech. 2015. Aspen Plus V8.8. Bedford, United States: Aspen technology Inc.
- J. Granacher, I. D. Kantor, F. Maréchal. 2021. Increasing Superstructure Optimization Capacity Through Self-Learning Surrogate Models. *Frontiers in Chemical Engineering*, 3.
- M. E. G. R. Domingos, D. Florez-Orrego, M. Teles dos Santos, S. de Oliveira Jr., F. Maréchal. 2022. Techno-economic and environmental analysis of methanol and dimethyl ether production from syngas in a kraft pulp process. *Computers & Chemical Engineering*, 163, 107810.
- D. Florez-Orrego, F. Maréchal, S. de Oliveira Jr. 2019. Comparative exergy and economic assessment of fossil and biomass-based routes for ammonia production. *Energy Conversion and Management*, 194, 22-36.
- C. Telmo, J. Lousada, N. Moreira. 2010. Proximate analysis, backwards stepwise regression between gross calorific value, ultimate and chemical analysis of wood. *Bioresource technology*, 101, 3808-3815.
- C. M. Kinchin, R. L. Bain. 2009. Hydrogen Production from Biomass via Indirect Gasification: The Impact of NREL Process Development Unit Gasifier Correlations, Technical Report NREL/TP-510-44868, May 2009. National Renewable Energy Laboratory.
- S. Consonni, R. E. Katofsky, E. D. Larson. 2009. A gasification-based biorefinery for the pulp and paper industry. *Chemical Engineering Research and Design*, 87, 1293-1317.
- A. Darmawan, F. Hardi, K. Yoshikawa, M. Aziz, K. Tokimatsu. 2017. Enhanced Process Integration of Entrained Flow Gasification and Combined Cycle: Modeling and Simulation Using Aspen Plus. *Energy Procedia*, 105, 303-308.
- McKay, M. D. et al. (1979). Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* 21, 239-245.
- C. E. Rasmussen, C. Williams, *Gaussian Processes for Machine Learning*. 2006. MIT Press. ISBN 026218253X.
- I. N. Silva, D. Hernane Spatti, R. Andrade Flauzino, L. Liboni, S. dos Reis Alves. 2016. *Artificial Neural Networks*. New York, NY: SpringerBerlin Heidelberg.
- L. Breiman. (2001). Random Forests. *Machine Learn.* 45, 5-32.
- A. Larsson, A. Nordin, R. Backman, B. Warnqvist, G. Eriksson. 2006. Influence of black liquor variability, combustion, and gasification process variables and inaccuracies in thermochemical data on equilibrium modeling results. *Energy and Fuels*, 20, 359-363.
- M. Puig-Arnavat, J. C. Bruno, A. Coronas. 2012. Modified thermodynamic equilibrium model for biomass gasification: a study of the influence of operating conditions. *Energy & Fuels*, 26, 1385-1394.

# Global optimization of symbolic surrogate process models based on Bayesian learning

Tim Forster<sup>a</sup>, Daniel Vázquez<sup>a</sup>, Gonzalo Guillén-Gosálbez<sup>a,\*</sup>

<sup>a</sup> *Department of Chemistry and Applied Biosciences, Institute for Chemical and Bioengineering, ETH Zurich, Vladimir-Prelog-Weg 1, 8093 Zurich, Switzerland*

\* [gonzalo.guillen.gosalbez@chem.ethz.ch](mailto:gonzalo.guillen.gosalbez@chem.ethz.ch)

## Abstract

In this work, we address the global optimization of process surrogates using Bayesian symbolic regression and deterministic global optimization algorithms. In contrast to other surrogates of process models that are hard to (globally) optimize, e.g., artificial neural networks or Gaussian processes, symbolic regression leads to a closed-form mathematical expression describing the observed data that can subsequently be globally optimized using off-the-shelf deterministic solvers. After providing an introductory example, we show the capabilities of our approach in the optimization of a methanol production plant. We further discuss the model accuracy, CPU times for model building and optimization, and outline the advantages and limitations of the proposed strategy.

**Keywords:** Global optimization, Symbolic Regression, Surrogate Modelling.

## 1. Introduction

Optimizing process flowsheets is a challenging problem in Process Systems Engineering, for which several approaches have been proposed to date. Process optimization often relies on mechanistic models based on first principles (Haydary, 2019) and closed-form expressions that enable the direct application of deterministic optimization algorithms, including global optimization methods. However, building such first principles is a challenging task, potentially leading to convergence issues during their optimization. Alternatively, when there is limited knowledge about the process, data-driven models can be the method of choice to build a mathematical expression that can be used for optimization purposes. Specifically, state-of-the-art machine learning algorithms (i.e., neural networks or support vector machines) can be applied to build a data-driven process model. However, optimizing classical machine learning models to global optimality is not trivial due to their intrinsic complexity and nonlinearities (Mitsos et al., 2009; Schweidtmann and Mitsos, 2019). Recently, symbolic regression algorithms have emerged to build algebraic models from data without assuming any specific structure (Cozad et al., 2014; Guimerà et al., 2020). This work explores the use of symbolic regression to build analytical surrogates of process flowsheets that can then be optimized with state-of-the-art global optimization (GO) algorithms. A priori assumptions about the mathematical structure of the model are avoided, and well-established global optimization algorithms are applied to the algebraic surrogate without the need for tailored strategies. Notably, after building a well-fitting analytical model using Bayesian symbolic regression, BARON (Sahinidis, 1996) is used to identify its global optimum. We deploy the proposed approach in an introductory example. Additionally, we then

demonstrate the performance of the approach with a complex flowsheet of an industrial process.

## 2. Methodology

### 2.1. Problem Statement

In this work, we consider a given process in steady state and a set of input variables  $x_i$ , with  $i \in I$  being the set of dependent variables, i.e., degrees of freedom, of the flowsheet. A target objective is defined by  $f(x)$ , which can be, for example, the total cost of a plant or the conversion of a reactor. In the first step, a suitable surrogate model  $F(x)$  that approximates the studied process target  $f(x)$  appropriately is identified. Second, a model-based global optimization is performed to find the global optimum  $x^*$ , as follows:

$$\begin{aligned}
 F^* &= \min_x F(x) \\
 \text{s. t. } \quad &g_n(x) \leq 0, \forall n \in N \\
 &h_m(x) = 0, \forall m \in M \\
 &\underline{x} \leq x \leq \bar{x} \\
 &x \in \mathbb{R}^{|I|}
 \end{aligned} \tag{1}$$

Where  $F^*$  corresponds to the model-based optimum of the process approximation  $F(x)$ , and  $g_n(x)$  and  $h_m(x)$  represent inequality and equality constraints, respectively. Herein, we aim to identify an accurate surrogate model without assuming a pre-defined model structure. In addition, we apply state-of-the-art deterministic GO solvers to model (1). Due to the sake of simplicity, the purpose of this work is a proof-of-concept, rather than a performance comparison.

### 2.2. Model Building

We seek to identify an accurate surrogate process model  $F(x)$  that maps the input data  $x$  to the corresponding objectives  $f(x)$ . In addition, to apply deterministic optimization approaches,  $F(x)$  is required to be in an algebraic form. To this end, in this work, we use the algorithm developed by Guimerà et al. (2020), the Bayesian machine scientist (BMS). It implements symbolic regression that uses a Markov chain Monte Carlo (MCMC) approach to explore the space  $E$  of closed-form mathematical expressions  $z_e, e \in E$ . During this explorative search, the algorithm minimizes the description length  $\mathcal{L}(z_e)$ , which can be approximated by the Bayesian Information Criterion (BIC) and an initial prior knowledge about a given expression  $z_e$ , where Bayesian learning is incorporated:

$$\mathcal{L}(z_e) \approx \frac{BIC(z_e)}{2} - \log(p(z_e)) \tag{2}$$

For simplicity, only the approximation of the description length is given here. For a detailed derivation, the reader is referred to the original work by Guimerà et al. (2020).

#### 2.2.1. Model-Based Global Optimization

With the obtained algebraic surrogate model  $F(x)$  at hand, the model-based optimization problem in (1) can be solved to identify the optimal decision variables values  $x^*$ , providing the best value of the surrogate  $F^* = F(x^*)$  and which can be evaluated in the original process model as  $f^* = f(x^*)$ . This observed response  $f^*$  can finally be compared

to the model-based optimum  $F^*$  to measure the mismatch between the surrogate and the process model in the optimum of the surrogate  $x^*$ .

### 3. Case Studies

#### 3.1. Software Implementation

The calculations were carried out on an AMD Ryzen 5 3600 CPU (introductory example) and on an Intel®Core™ i7-8700 CPU (methanol production plant), both with 16 GB of RAM. Python v3.9 (introductory example) and v3.8 (methanol production plant) were used to implement the algorithm. The flowsheet was simulated in Aspen HYSYS v11. Python and HYSYS were connected through a COM interface. The algorithm provided by Guimerà et al. (2020) was used to train the BMS. The obtained algebraic expression was implemented using the General Algebraic Modeling System (GAMS Software GmbH, 2022) v40.2 and optimized using BARON 20.4 (Sahinidis, 1996).

#### 3.2. Introductory Example

We illustrate the approach using first a simple two-dimensional unconstrained global optimization problem, the Rosenbrock function (Rosenbrock, 1960):

$$f(x) = (x_1 - a)^2 + b(x_2 - x_1^2)^2 \tag{3}$$

There are two dependent variables  $x_1$  and  $x_2$ , and two parameters  $a$  and  $b$ . For the subsequent illustrations, parameters  $a$  and  $b$  are chosen to be 1 and 10, respectively. A graphical representation of the resulting contour plot bound to  $\underline{x} = [-2, -1]$  and  $\bar{x} = [2, 3]$ , is given in Figure 1.

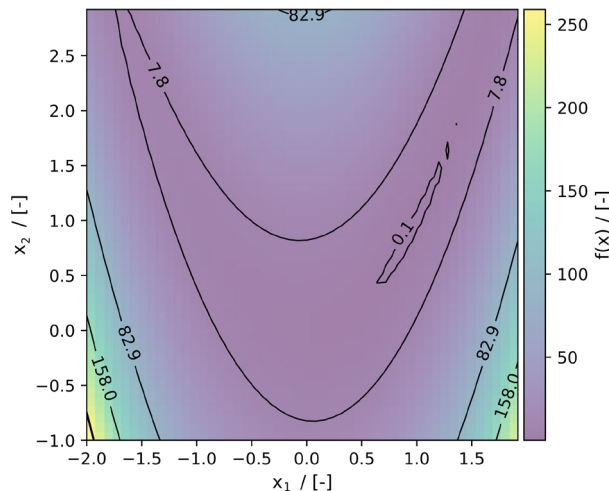


Figure 1: Contour of the Rosenbrock function. The color gradient refers to the value of  $f(x)$ . The contour levels are indicated by the black lines with the corresponding values of  $f(x)$ .

To evaluate  $f(x)$  according to (3), 225 equally distant data points were generated within the bounds. The global optimum (minimization) is given by  $f^* = 0$  at  $x^* = [1, 1]$ . The obtained samples ( $N = 225$ ) were randomly split into training ( $N_{tr} = 157$ ) and test set ( $N_{te} = 68$ ). Details of the BMS training are shown in Table 1.

### 3.3. Methanol Production Plant

A methanol (MeOH) production plant model is investigated next (Figure 2), which was adapted from Vázquez and Guillén-Gosálbez (2021). The feed consists of carbon dioxide and hydrogen, whose pressure and temperature are adjusted to the desired conditions before being sent to a plug-flow reactor. The reactor outlet is sent through two flash drums and into a distillation column. Methanol is collected in the distillate, while water is the main product at the bottom. The vapor streams of the flash drums are partly recycled (from the first flash) and purged (from both drums). For a detailed description, the reader is referred to the work by Vázquez and Guillén-Gosálbez (2021).

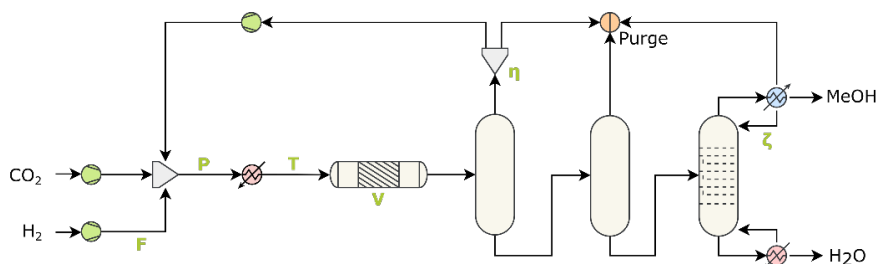


Figure 2: Simplified schematic representation of the MeOH plant under study adapted from Vázquez and Guillén-Gosálbez (2021), with the decision variables highlighted in green.

The degrees of freedom  $x = [T, P, \eta, V, F, \zeta]$  to be optimized (in green in Figure 2) consist of the reaction temperature ( $T$ ), the reaction pressure ( $P$ ), the purge ratio of the splitter ( $\eta$ ), the reactor volume ( $V$ ), the hydrogen flow rate ( $F$ ), and the reflux ratio of the distillation column ( $\zeta$ ). The objective is to minimize the unitary cost ( $UC$ ) of methanol. The carbon dioxide inlet flow is fixed. Vázquez and Guillén-Gosálbez (2021) provide the exact calculation of  $UC$  and a more detailed description of the simulation.

To determine the optimum heat exchanger network (HEN), and, therefore, the cost of the HEN, the mixed-integer nonlinear programming model developed by Yee and Grossmann (1990) (SYNHEAT) is solved for each sampling point. The obtained samples ( $N = 1000$ ) were randomly split into training ( $N_{tr} = 800$ ) and test set ( $N_{te} = 200$ ). The bounds indicated in Table 1 were used to generate these samples. The applied settings and hyperparameters for the BMS training are shown in Table 1.

Table 1: The chosen BMS hyperparameters (MCMC steps and allowed operations) are given, whereas otherwise, the default settings chosen by the authors (Guimerà et al., 2020) were applied. Lastly, used relative optimality gap  $\epsilon_R$  of BARON is indicated for both examples.

Case Study	BMS MCMC steps	Allowed Operations	Optimizer Settings
Rosenbrock	$5 \cdot 10^3$	$\exp(x), \log(x), x^2, x^3, \sqrt{x}$ ,	$\epsilon_R = 0$
Methanol	$20 \cdot 10^3$	$+, -, \div, x^a$ , and $ x $ .	

## 4. Results and Discussion

### 4.1. Introductory Example

A summary of the results is given in Table 2. The BMS could identify the true equation, given by the following expression:

$$F(x) = (x_1 - \theta_1)^2 + \theta_2(x_2 - x_1^2)^2 \tag{4}$$

The simultaneously identified parameters were found to be  $\theta = [10, 1]$ . The model training required 788 s. These results can also be visually compared by considering the true contour plot in Figure 1, and the predictions given in Figure 3 (b).

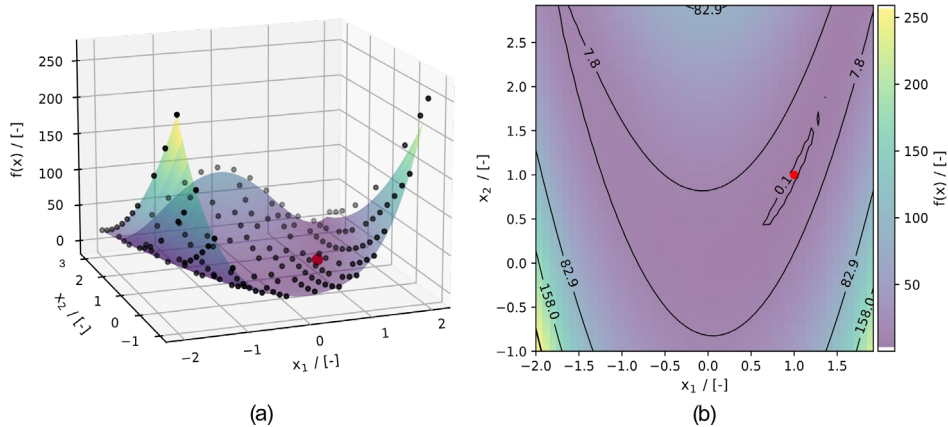


Figure 3: (a) The training data (dark dots) is shown together with the model predictions (colored surface). The identified optimum is depicted as a red dot. (b) The contour of the BMS model predictions. The identified model-based global optimum  $F^*$  is again shown as a red dot.

Figure 3 (a) provides the training data together with the model predictions, where Figure 3 (b) represents the obtained contour plot of the model predictions. Optimizing the equation in GAMS leads to the optimum solution given by  $F^* = 0$  with  $x^* = [1, 1]$ . The optimization with BARON was carried out in 0.2 s. Implementing  $x^*$  into  $f(x)$  results in  $f^* = 0$ , which is also the true global optimum of the Rosenbrock function.

Table 2: Training and optimization results summary for the introductory example. The CPU times for training and optimization (minimization) are shown together with the calculated root mean squared error (RMSE) for the training and testing set. In addition, the obtained optimum  $x^*$  with the corresponding solution  $F^*$  is shown. Lastly, the process output  $f^*$  is shown for the corresponding optimum  $x^*$ .

CPU Training	RMSE	CPU Optimization	$x^*$	$F^*$	$f^*$
788 s	0 [-] (Train) 0 [-] (Test)	0.2 s	$x_1 = 1$ [-] $x_2 = 1$ [-]	0	0

#### 4.2. Methanol Production Plant

The goal of this six-dimensional case study is to minimize the unitary cost, which is a nonlinear programming problem (NLP). After training the BMS, an algebraic model could be obtained and then optimized in GAMS. The training and optimization steps took 40\*600 s and 1 s, respectively. Inserting the identified model-based optimum  $x^*$  into HYSYS again led to an optimum solution  $f^* = 0.73$  \$/kg, which was very close to the prediction made with the algebraic surrogate at its optimum of  $F^* = 0.72$  \$/kg. The training time of the BMS is substantial (>10 hours), yet this allows for a fast and easy optimization. However, more advanced symbolic regression algorithms could reduce the



computational time of the training phase. Moreover, an additional advantage of having an algebraic surrogate expression at hand is that it could be used for calculations that need to solve optimization models iteratively, such as in multi-objective optimization or in real-time optimization, since the optimization step is greatly simplified using the algebraic surrogate.

## 5. Conclusions

In this work, we explored the use of algebraic surrogate process models constructed using symbolic regression. Accurate process models were built using symbolic regression based on Bayesian learning and then solved in GAMS using an off-the-shelf global optimization algorithm (BARON). We showed the capabilities of the approach in two examples. The high computational training times for symbolic regression still represent a computational bottleneck, yet they might be reduced with the future emergence of faster regression algorithms. Moreover, analytical surrogate process models could be used to simplify the calculations in applications requiring solving multiple optimization problems iteratively.

## References

- A. Cozad, N.V. Sahinidis, D.C. Miller, 2014, Learning surrogate models for simulation-based optimization, *AIChE Journal*, 60, 2211–2227.
- GAMS Software GmbH, 2022, GAMS Development Corporation. General Algebraic Modeling System (GAMS) Release 40.2.0, 2021.
- R. Guimerà, I. Reichardt, A. Aguilar-Mogas, F.A. Massucci, M. Miranda, J. Pallarès, M. Sales-Pardo, 2020, A Bayesian machine scientist to aid in the solution of challenging scientific problems, *Science Advances*, 6.
- J. Haydary, 2019, *Chemical Process Design and Simulation: Aspen Plus and Aspen Hysys Applications*, Wiley, Haydary2019.
- A. Mitsos, B. Chachuat, P.I. Barton, 2009, McCormick-Based Relaxations of Algorithms, *SIAM Journal on Optimization*, 20, 573–601.
- H.H. Rosenbrock, 1960, An Automatic Method for Finding the Greatest or Least Value of a Function, *The Computer Journal*, 3, 175–184.
- N.V. Sahinidis, 1996, BARON: A general purpose global optimization software package, *Journal of Global Optimization*, 8, 201–205.
- A.M. Schweidtmann, A. Mitsos, 2019, Deterministic Global Optimization with Artificial Neural Networks Embedded, *Journal of Optimization Theory and Applications*, 180, 925–948.
- D. Vázquez, G. Guillén-Gosálbez, 2021, Process design within planetary boundaries: Application to CO<sub>2</sub> based methanol production, *Chemical Engineering Science*, 246, 116891.
- T.F. Yee, I.E. Grossmann, 1990, Simultaneous optimization models for heat integration—II. Heat exchanger network synthesis, *Computers & Chemical Engineering*, 14, 1165–1184.

# Molecular Property Targeting for Optimal Solvent Design in Extractive Distillation Processes

Zihao Wang,<sup>a</sup> Teng Zhou,<sup>b,#</sup> Kai Sundmacher<sup>a,c,\*</sup>

<sup>a</sup> Department for Process Systems Engineering, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, D-39106 Magdeburg, Germany

<sup>b</sup> Sustainable Energy and Environment Thrust, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China

<sup>c</sup> Chair of Process Systems Engineering, Otto-von-Guericke University Magdeburg, Universitätsplatz 2, D-39106 Magdeburg, Germany

\* [sundmacher@mpi-magdeburg.mpg.de](mailto:sundmacher@mpi-magdeburg.mpg.de) (K. Sundmacher), # [tengzhou@ust.hk](mailto:tengzhou@ust.hk) (T. Zhou)

## Abstract

In extractive distillation (ED), a suitable solvent is added to increase the relative volatility of azeotropic or close-boiling mixtures, thereby facilitating the separation. The viability and efficiency of this separation are largely dependent on the selection of solvent. In this work, solvents are optimally designed to improve the performance of ED processes by a molecular property targeting method. Specifically, optimal solvents are identified by directly targeting desirable molecular properties obtained from process optimization. First, data-driven process models are established to estimate key performance indicators of the ED process with the most important process-relevant physical properties of the solvents. Based on the established models, multi-objective optimization is then performed to maximize the product purity while minimizing energy cost, through which optimal molecular properties are obtained. Hypothetical target molecules featuring the desirable properties are thereby generated. Subsequently, several real solvents that approximate the optimal property profiles of the hypothetical molecules are selected from a large solvent database. In a final assessment step, the performance of these real solvent candidates is evaluated via detailed process simulations. The proposed molecular-property-targeting solvent design methodology is demonstrated by the separation of 1-butene and 1,3-butadiene. For this separation task, three real solvents are finally identified, allowing for a higher product purity and lower energy consumption compared to the benchmark solvent n-methyl-2-pyrrolidone (NMP).

**Keywords:** solvent design, artificial neural network, data-driven modeling, continuous molecular targeting, extractive distillation

## 1. Introduction

Extractive distillation (ED) processes enable the separation of azeotropic or close-boiling mixtures, where a suitable solvent is introduced to increase the relative volatility of the mixture (Gerbaud et al., 2019). The viability and efficiency of this separation largely depends on the selection of the solvent. In the past, most solvents used in chemical industries were primarily chosen based on chemical knowledge and have not been systematically designed (Chemmanattuvalappil, 2020). They may not be the real optimal candidates to meet the separation requirements. Therefore, discovering better alternatives to replace current solvents following a systematic solvent screening or design strategy can improve the performance of the ED processes.

Computer-aided molecular design (CAMD) provides a promising route to systematically design solvents fulfilling a set of target molecular attributes or process performance indicators (Gani, 2004). Taking the advantages of modern molecular property models and process models, CAMD methods have been extensively used to design solvents for various applications such as absorption (Burger et al., 2015; Zhang et al., 2021; Zhou et al., 2021), extraction (Song et al., 2018; Ten et al., 2021), and reaction (Struebing et al., 2013; Zhou et al., 2015; Zhou et al., 2016). So far, the most commonly used approach for representing molecular structures in CAMD is the decomposition into functional groups (Marrero & Gani, 2001). By this method, solvents are represented by discrete variables (e.g., the number of molecular building blocks) which are optimized under certain molecular structural constraints. To circumvent discrete molecular decisions in CAMD, Bardow et al. (2010) proposed a continuous-molecular-targeting (CoMT) approach by considering a continuous molecular structure space in terms of molecular property-related parameters. The optimal CAMD solution is then used to identify existing molecules, the properties of which approximate the optimal parameter values best. The CoMT approach has been applied to design solvents for CO<sub>2</sub> absorption (Stavrou et al., 2014) and working fluids for the organic Rankine cycle (Schilling et al., 2017).

In this contribution, the CoMT approach is adopted for the optimal design of solvents in ED processes. To achieve higher product purity and lower energy consumption, solvents are designed by directly targeting desirable solvent properties obtained from process optimization. First, data-driven process models are established to estimate key performance indicators of the ED process from the most important process-relevant physical properties of the solvent. Subsequently, solvent design is performed in two steps, namely molecular property targeting and molecular mapping. In the first step, the optimal molecular parameters are obtained by model-based optimization. Thereby, hypothetical target molecules featuring the desirable molecular properties are generated. In the subsequent step, real solvents that approximate the optimal property profiles of the hypothetical molecules are selected from a large solvent database. Finally, the performance of these solvent candidates is evaluated via detailed process simulations. The proposed molecular-property-targeting solvent design methodology is illustrated using an industrially relevant case, i.e., the separation of the close-boiling mixture 1-butene/1,3-butadiene (C<sub>4</sub>H<sub>8</sub>/C<sub>4</sub>H<sub>6</sub>).

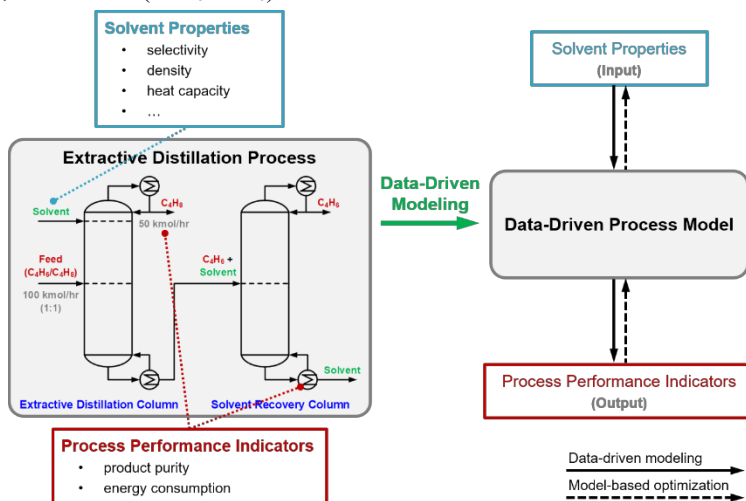


Figure 1. Data-driven modeling for the ED process.

## 2. Data-Driven Process Modeling

Process models are required to guide the design of solvents that meet the separation requirements. In a data-driven manner, process models are built where solvents described by molecular properties are directly linked to their corresponding process performance (Figure 1). Such models can direct the design of solvents by identifying the optimal values of molecular properties to maximize the process performance. We perform the solvent design under the operating conditions pre-determined by process optimization using *n*-methyl-2-pyrrolidone (NMP) as the solvent. Thus, the present work can be considered as the search for potentially better alternatives (e.g., showing lower energy consumption in extractive distillation) to replace the industrially used solvent NMP.

The dataset for data-driven modeling contains input-output pairs for 126 solvents. The inputs are five process-relevant physical properties of the solvent, including selectivity of 1-butene/1,3-butadiene at the infinite dilution condition, molecular weight, density, molar heat capacity, and viscosity. The outputs are the key performance indicators of the ED column, including product purity ( $C_4H_8$  purity in the distillate) and energy consumption (described by the reboiler heat duty for simplicity). The dataset is generated through rigorous process simulation in Aspen Plus based on the UNIFAC thermodynamic model. It is randomly split into two sets, i.e., a training set (80%) for model development and a test set (20%) for model evaluation. Feature scaling is adopted to standardize the model input.

Feedforward neural network (FNN), the most straightforward type of artificial neural networks, is used to build the data-driven process model using PyTorch (Paszke et al., 2019). To constrain model complexity and reduce overfitting, the FNN has up to two hidden layers with a maximum of four neurons in each layer during the hyperparameter optimization. Different types of non-linear activation functions are considered.

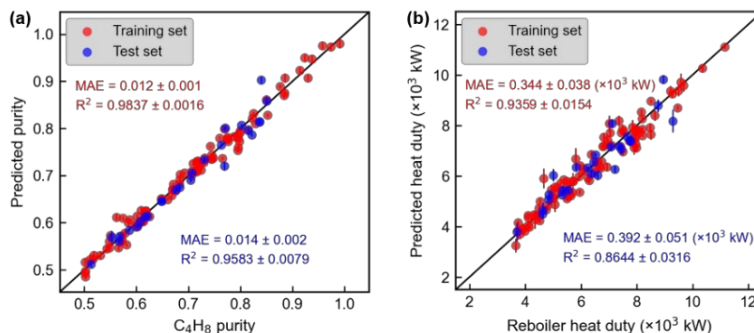


Figure 2. Performance of the data-driven models in predicting (a)  $C_4H_8$  purity and (b) reboiler heat duty. The corresponding process performance for each solvent is predicted using five models derived from the five-fold cross-validation, and the point and error bar show the average and standard deviation of these predictions, respectively.

After identifying the optimal FNN configurations (hyperparameters) using five-fold cross-validation, the models are evaluated on the training and test sets. As they present satisfactory accuracy in the prediction of  $C_4H_8$  purity and heat duty (Figure 2), the data-driven models are subsequently used for solvent design.

## 3. Molecular Property Targeting

Based on the data-driven models, a multi-objective optimization problem aiming at maximizing the  $C_4H_8$  purity, while minimizing the heat duty is formulated as follows:

$$\begin{aligned} & \min_{x \in X} (1 - f_1(x), f_2(x)) \\ \text{s.t.} \quad & X_i^L \leq X_i \leq X_i^U \\ & X_i \in \mathbb{R} \end{aligned}$$

where  $f_1(x)$  and  $f_2(x)$  are the data-driven models for the estimation of  $C_4H_8$  purity and reboiler heat duty, respectively.  $X$  is the molecular property space, in which each variable is standardized individually.  $L$  and  $U$  denotes the lower and upper bounds. Box constraints are set on each variable based on the corresponding minimum and maximum values from the dataset.

The multi-objective optimization problem is solved using a genetic algorithm implemented in Pymoo (Blank & Deb, 2020). It converges in 100 generations, obtaining a set of Pareto-optimal solutions (i.e., hypothetical target solvent molecules). The hypothetical molecules are considered to be the optimal solutions found in the design space closest to the ideal point. Figure 3 depicts the objective function values for the hypothetical target molecules featuring optimal molecular properties.

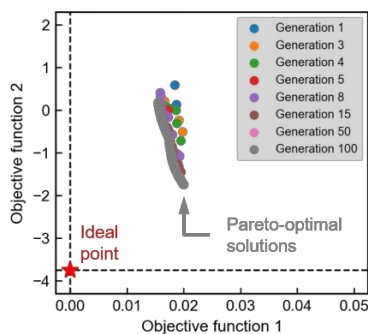


Figure 3. Multi-objective optimization visualized in the space of objective functions. Standardization is applied to the heat duty (objective function 2) in the optimization so that the magnitudes of two objective functions are comparable.

#### 4. Molecular Mapping

The industrially used solvent NMP is set as the benchmark for comparison. Our target is to find better solvents that allow for a higher  $C_4H_8$  purity and a lower heat duty under the given operating conditions. Figure 4a shows the estimated process performance for hypothetical molecules (in gray) and benchmark solvent NMP (in green). The hypothetical molecules are closer to the ideal point than NMP, demonstrating the success of the molecular property targeting step.

In the molecular mapping step, the hypothetical target molecules obtained by molecular property targeting are mapped into real solvents. The molecular mapping is performed by searching a large database consisting of 1,259 real solvents. A preliminary criterion to find the real solvents closest to the hypothetical target molecules is based on the Euclidean distance in the molecular property space. Thereby, optimal real solvents that approximate the optimal property values are identified.

Nineteen solvent candidates are obtained from the molecular mapping. Their estimated process performance is also presented in Figure 4a. It is observed that two solvent candidates are closer to the ideal point than the hypothetical molecules. Therefore, they could in principle show better process performance than the hypothetical molecules.

A detailed simulation of the ED process on the 19 solvent candidates proves that 16 of them are technically viable to achieve the separation of  $C_4H_8/C_4H_6$ . Among them, nine

solvents present both a lower heat duty and a lower  $C_4H_8$  purity compared to NMP under the same operating conditions. Besides, three solvents, methyl cyanoacetate, glutaronitrile, and 1,4-dicyano-2-butene (blue dots in the green area of Figure 4b), are better alternatives to the benchmark solvent NMP, because they allow for a higher product purity and lower energy consumption under the specified operating conditions. Moreover, their process performance can be further improved by performing rigorous process optimization.

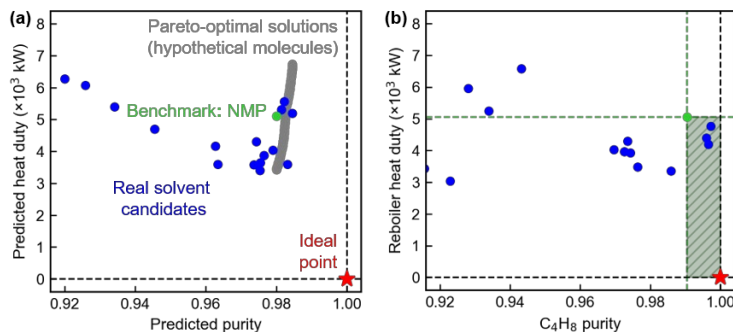


Figure 4. (a) Process performance estimated by the data-driven models for hypothetical molecules, NMP, and real solvent candidates identified by molecular mapping. (b) Process performance evaluated via rigorous simulation of the ED process. Vertical and horizontal lines in green represent the  $C_4H_8$  purity and heat duty of the benchmark process.

For the three solvent candidates, their molecular properties and corresponding process performance are summarized in Table 1. These solvent candidates have similar molecular properties (except viscosity) to the benchmark solvent NMP. A correlation analysis performed on the dataset indicates that a solvent with a higher selectivity at infinite dilution could lead to a higher  $C_4H_8$  purity. This can also be inferred from Table 1. All the three solvents have higher selectivity and higher  $C_4H_8$  purity than NMP. Therefore, the infinite dilution selectivity of the solvent may have played a vital role in designing solvents for a better separation of  $C_4H_8$  and  $C_4H_6$ .

Table 1. Molecular properties and the corresponding process performance of the solvents.

Solvent	NMP	Methyl cyanoacetate	Glutaronitrile	1,4-Dicyano-2-butene
CAS number	872-50-4	105-34-0	544-13-8	18715-38-3
Molecular formula	$C_5H_9NO$	$C_4H_5NO_2$	$C_5H_6N_2$	$C_6H_6N_2$
Selectivity at infinite dilution	1.642	1.753	1.666	1.689
Molecular weight (g/mol)	99.13	99.09	94.12	106.13
Density ( $kg/m^3$ )	1027	1117	983	1002
Molar heat capacity (J/mol·K)	161.7	192.5	191.0	200.9
Viscosity (cP)	1.89	2.82	6.17	7.13
$C_4H_8$ purity	0.9904	0.9965	0.9960	0.9972
Reboiler heat duty (kW)	5059	4198	4394	4761

## 5. Conclusion

This work presents a molecular property targeting approach for the design of solvents in ED processes. Based on data-driven process models, optimal molecular properties are obtained, thereby generating a set of hypothetical target molecules. Nineteen real solvent candidates are identified from a large solvent database by molecular mapping. Their performance on the separation of 1-butene and 1,3-butadiene is evaluated via detailed

process simulations. It turns out that three solvents allow for higher product purity and lower energy consumption under the given operating conditions, compared to the benchmark solvent NMP. Further process optimization is necessary to find the optimal ED process for these solvent candidates respectively. Future work will also account for the integrated design of the solvent and the ED process incorporating the molecular property targeting approach, where environmental, health, and safety impacts of solvents can be considered.

## Acknowledgement

This research work is supported by the International Max Planck Research School for Advanced Methods in Process and Systems Engineering (IMPRS ProEng), Magdeburg, Germany.

## References

- A. Bardow, K. Steur, J. Gross, 2010, Continuous-molecular targeting for integrated solvent and process design, *Industrial & Engineering Chemistry Research*, 49, 6, 2834-2840.
- J. Blank, K. Deb, 2020, Pymoo: Multi-objective optimization in python, *IEEE Access*, 8, 89497-89509.
- J. Burger, V. Papaioannou, S. Gopinath, G. Jackson, A. Galindo, C.S. Adjiman, 2015, A hierarchical method to integrated solvent and process design of physical CO<sub>2</sub> absorption using the SAFT- $\gamma$  Mie approach, *AIChE Journal*, 61, 10, 3249-3269.
- N.G. Chemmangattualappil, 2020, Development of solvent design methodologies using computer-aided molecular design tools, *Current Opinion in Chemical Engineering*, 27, 51-59.
- R. Gani, 2004, Computer-aided methods and tools for chemical product design, *Chemical Engineering Research and Design*, 82, 11, 1494-1504.
- V. Gerbaud, I. Rodriguez-Donis, L. Hegely, P. Lang, F. Denes, X. You, 2019, Review of extractive distillation. Process design, operation, optimization and control, *Chemical Engineering Research and Design*, 141, 229-271.
- J. Marrero, R. Gani, 2001, Group-contribution based estimation of pure component properties, *Fluid Phase Equilibria*, 183, 183-208.
- A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, 2019, Pytorch: An imperative style, high-performance deep learning library, *Advances in Neural Information Processing Systems*, 32.
- J. Schilling, M. Lampe, J. Gross, A. Bardow, 2017, 1-stage CoMT-CAMD: An approach for integrated design of ORC process and working fluid using PC-SAFT, *Chemical Engineering Science*, 159, 217-230.
- Z. Song, C. Zhang, Z. Qi, T. Zhou, K. Sundmacher, 2018, Computer-aided design of ionic liquids as solvents for extractive desulfurization, *AIChE Journal*, 64, 3, 1013-1025.
- M. Stavrou, M. Lampe, A. Bardow, J. Gross, 2014, Continuous molecular targeting-computer-aided molecular design (CoMT-CAMD) for simultaneous process and solvent design for CO<sub>2</sub> capture, *Industrial & Engineering Chemistry Research*, 53, 46, 18029-18041.
- H. Struebing, Z. Ganase, P.G. Karamertzanis, E. Sioukrou, P. Haycock, P.M. Piccione, A. Armstrong, A. Galindo, C.S. Adjiman, 2013, Computer-aided molecular design of solvents for accelerated reaction kinetics, *Nature Chemistry*, 5, 11, 952-957.
- J.Y. Ten, Z.H. Liew, X.Y. Oh, M.H. Hassim, N. Chemmangattualappil, 2021, Computer-aided molecular design of optimal sustainable solvent for liquid-liquid extraction, *Process Integration and Optimization for Sustainability*, 5, 2, 269-284.
- X. Zhang, X. Ding, Z. Song, T. Zhou, K. Sundmacher, 2021, Integrated ionic liquid and rate-based absorption process design for gas separation: Global optimization using hybrid models, *AIChE Journal*, 67, 10, e17340.
- T. Zhou, K. McBride, X. Zhang, Z. Qi, K. Sundmacher, 2015, Integrated solvent and process design exemplified for a Diels-Alder reaction, *AIChE Journal*, 61, 1, 147-158.
- T. Zhou, H. Shi, X. Ding, Y. Zhou, 2021, Thermodynamic modeling and rational design of ionic liquids for pre-combustion carbon capture, *Chemical Engineering Science*, 229, 116076.
- T. Zhou, J. Wang, K. McBride, K. Sundmacher, 2016, Optimal design of solvents for extractive reaction processes, *AIChE Journal*, 62, 9, 3238-3249.

# A Bayesian-based screening framework for optimal development of safe-by-design nanomaterials

Kostas Blekos, Effie Marcoulaki

*System Reliability and Industrial Safety Laboratory, National Centre for Scientific Research "Demokritos", P.O. Box 60037, Agia Paraskevi 15310, Greece*  
[emarcoulaki@ipta.demokritos.gr](mailto:emarcoulaki@ipta.demokritos.gr)

## Abstract

In this article we propose a novel screening approach for the development of new nanomaterials (NM) which are safe(r)-by-design. An integral part of our approach involves the use of recently proposed notation standards and extensions, which facilitate the representation and processing of data related to NMs. In particular, we utilize a general extendable representation for NMs which is built in line with InChI extensions for NMs. This representation enables us to link the composition, structure, and experimental conditions of a NM to its toxicity profile, using descriptor information obtained from available datasets. The processed data are then complemented with predictions from QSAR models and fed to a Bayesian-based procedure to assess the safety of new NM designs. This comprehensive approach can be used to build a local screening tool to search for less hazardous NM configurations.

**Keywords:** nanomaterial, toxicity, safe-by-design, Bayesian, nanoinformatics.

## 1. Introduction

During the last few decades, nanotechnology and nanomaterial (NM) applications have emerged as key enabling technologies for the global industry, since the unique properties of NMs and nano-enabled products allow for a very wide range of applications with significant potential. The risks of engineered nanomaterials (ENM) are largely unknown, however their size and composition make them inherently threatening to the human health and the environment. In effect, WHO (2017) published precautionary guidelines for their handling and use, and OECD developed appropriate Test Guidelines to identify, assess and manage potential emerging risks related to NMs throughout the development chain (OECD, 2020) towards a Safe(r) Innovation Approach (Shandilya et al., 2020).

Early identification, assessment and management of the potential health, safety and environmental risks related to ENMs throughout their life-cycle is extremely important. Risk assessment of ENMs is challenging because of the diversity in their physical and chemical composition, their surface functionalization and structural properties, and the variety of alternative configurations. Significant public and private investment over the past decades enabled the development of advanced experimental methods for generation of extensive nanotoxicity data, and supported the development of informatics tools for data management / processing and for the prediction of physicochemical, exposure and (eco)toxicological properties of novel NMs to enable in silico risk assessment (Marcoulaki et al., 2021). Among these, the NanoInformaTIX H2020 project develops a database comprising the significant amounts of available data, and a set of tools for modeling materials, exposure and dose-response. These tools will become available on a user-friendly web-based platform for ENM risk management, to support sustainable manufacturing of ENM-based products.



The development of optimization tools for designing ENMs with defined properties is a very new research field. Preliminary works applied Generative Adversarial Networks (GAN) to optimize functionality (So and Rho, 2019; Kim et al., 2020). This work proposes an incremental procedure to develop a screening framework for safe-by-design ENMs. We start from novel representation schemes, build extensions to improve the data analysis potential, propose an assessment procedure using similarity functions built on top of the representation. All these finally enable the generation and assessment of an ensemble of alternative ENM designs, starting from a given ENM design.

## 2. Representation methodology

In a recent work we proposed a general, extendable representation for NMs from the point of view of computational applications and we set qualifications for a computational-friendly representation (Blekos et al., 2022). Our proposal was in line with the recently proposed InChI for NMs (Lynch et al., 2020) and its realization greatly facilitates similarity assessments between components which, in turn, is a crucial part of calculating NM affinities. Our work identified key challenges for a computational-friendly NM representation and proposed strategies to address them. Through the resulting representation qualifications, we created a proposal of 5 extensions for the InChI for NMs that aimed to address the identified challenges.

These extensions enable the computationally-friendly representation of important and commonly found features of NMs. They are designed to facilitate standardization, curation, and comparison of data on NMs. This can help improve the sharing and management of data on NMs, and can assist researchers in identifying potential safety concerns and making informed decisions about the use of these materials. The extensions are: a) an extension to describe distributions, which can be applied anywhere a real number needs to be represented; b) an extension to describe shapes; c) an extension for the representation of mixtures and interfaces using logical operators that apply to other layers already defined, building on the mixture-InChI; d) an extension for structured comments that allows for consistent and uniform capturing of auxiliary information and facilitates database curation; and e) an extension for the definition of macros and shortcuts that allows for the creation of reusable, parametrizable, and human-readable components.

## 3. Data preprocessing

The representation and extensions described in Section 2 provide a consistent and comprehensive framework for describing and analyzing the properties of NMs. In particular, the proposed extensions for distributions and shape/morphology, provide new capabilities for data-driven tools, and enables safety assessment, which is crucial for the design of materials with improved safety profiles.

### 3.1. Data preprocessing

Starting from a given dataset, the proposed screening framework uses a bottom-up and a top-down stage to generate two outputs that will be fed to the next stages. The two data preprocessing stages are to be conducted simultaneously and they are complementary, as shown in Figure 1 and described below.

In the *top-down stage*, we access available NM databases to obtain descriptors that provide information about the composition, structure, and other characteristics of NMs. Descriptor information is taken from available datasets, linking composition, structure and experimental conditions information to the NM toxicity profile. These descriptors can be further analyzed and encoded through QSAR and/or other similar models, to explore the properties of NMs and identify safety and toxicity profiles.

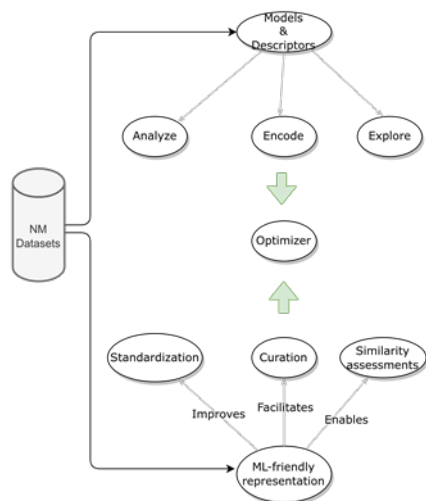


Figure 1: Two stage data preprocessing

In the *bottom-up stage*, we convert the NM data into the computational-friendly representation outlined in Section 2. The representation and the proposed extensions provide a comprehensive and consistent framework for describing and analyzing the properties of NMs, including their chemical composition, morphology, and interactions with biological systems. This allows the more detailed and accurate evaluation of the comparative properties (e.g. toxicity) of NMs.

Additionally, the proposed extensions to the representation, such as the ability to describe probability distribution functions (pdfs) and bypass categorical shape descriptors, provide new capabilities for analyzing and improving the safety of NMs. For example, the extension for describing distributions can be used to better characterize

the distribution of particles within a NM, which can have important implications for its safety and toxicity. Similarly, the extension for bypassing categorical shape descriptors allows for more flexible and comprehensive analysis of the toxicity of the NMs, which can be useful for identifying novel materials with improved safety profiles.

### 3.2. Similarity and toxicity matrices

In the next stage, the outputs of data preprocessing are used to create a matrix of toxicity values, with each toxicity model providing a separate value for each NM (Figure 2). We can additionally conduct similarity assessments based on the proposed representation model. This provides information on the similarity between different NMs, based e.g. on their chemical composition, structure, zeta potential and other characteristics etc. (Figure 3). The obtained information is used to create a matrix of similarity values for each NM and each axis of comparison (Figure 4).

## 4. Bayesian-based screening

The toxicity and similarity matrices of Section 3 are utilized as input data to create and train a Bayesian network (BN). The BN is a probabilistic graphical model that represents the conditional dependencies between different variables in a system, such as the likelihood of a given NM being toxic given its similarity to other NMs. Once the BN is obtained, it can be used to perform probabilistic inference and assess the toxicity of NMs options which are not included in the dataset.

As shown in Figure 4, the BN constitutes the core function of a local search procedure towards less toxic NMs. Note that, in principle we can consider functionality as an additional objective. Otherwise, we assume that small changes to the design descriptors are unlikely to significantly affect the functionality of the NM.

The search starts from an existing NM with known safety characteristics and the desired functionality for a particular application. A local search is performed in the vicinity of similar designs, to identify design options featuring lower toxicity. The search uses the similarity functions and the Bayesian network, to alter the NM descriptors and evaluate the new toxicity, respectively. Iterative application of this procedure on the new NMs, progressively generates design alternatives with the desired safety profile.

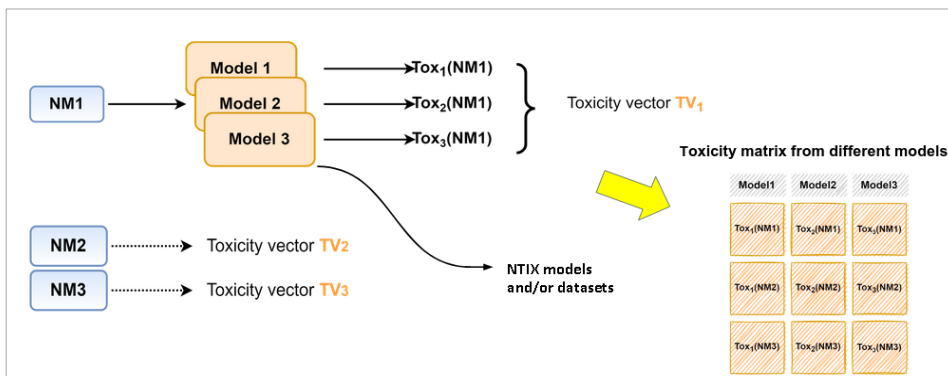


Figure 2: Toxicity matrix using externally available QSAR models

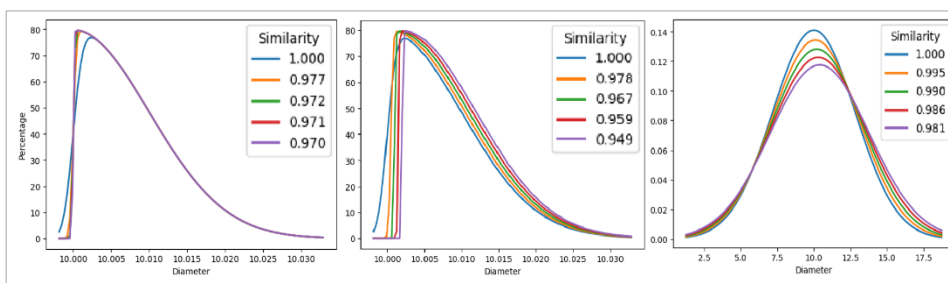
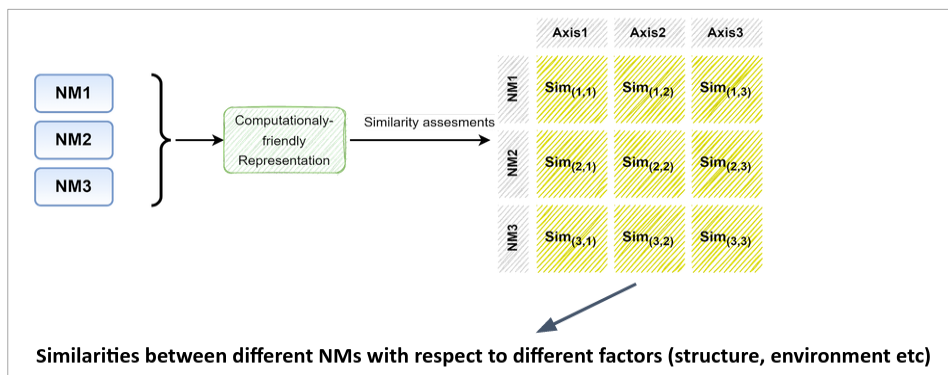


Figure 3: Results of similarity calculations for different pdfs of the NM major diameter



Similarities between different NMs with respect to different factors (structure, environment etc)

Figure 4: Similarity matrix obtained from the computational-friendly representation

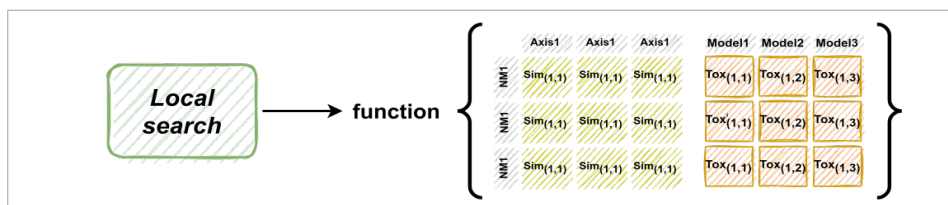


Figure 5: Bayesian-based screening using the Toxicity and Similarity matrices as inputs.

The success of the proposed framework relies heavily on the availability, the accuracy and the range of descriptor information to generate reliable inputs for the BN. The data can come from relevant datasets and/or from QSAR predictions as described in section 2.

QSAR models allow for cost-effective and efficient assessment of the potential risks of NMs without the need for extensive in vivo experiments (Huang et al., 2021). Latest trends use machine learning (ML) algorithms to augment or develop QSARs (Furxhi et al., 2020; Balraadjsing et al., 2022; Singh et al., 2020). There are however many challenges in the development and use of these models for nano-toxicity predictions. A major challenge is the limited number of experimental data, making it difficult to train models that require large and diverse datasets to achieve acceptable prediction accuracy. There is also need to connect and harmonize different data sets, for example by using read across and other methods of filling data gaps. Finally, the complexity and diversity of NMs hinder the identification of the most relevant structural and/or physicochemical properties that affect their toxicity. The above can lead to the development of ML models that have low predictive power and/or a limited domain of applicability. On a different perspective, all the predictive toxicology models should be validated in a manner that is acceptable to regulatory agencies. This requires the development of transparent and reproducible validation procedures, to ensure that these models are suited to support regulatory decision-making in the field of nanotoxicity (Winkler, 2020).

To avoid these concerns, we are currently implementing a simplified version of the proposed framework, relying only on an extended and well curated dataset. This version will soon be integrated on the NanoInformaTIX project platform. Advanced QSARs are also under development in NanoInformaTIX and will be available on the platform, along with data sources and other models to enable risk assessment and property predictions. These will enable to fully implement the proposed SbD framework in the future.

## **5. Discussion and conclusions**

The physical, chemical and biological behavior of the materials at nanoscale is very different from their conventional behavior at macroscale, due to a major shift in the dominant phenomena. This creates enormous opportunities towards novel advanced solutions in a wide range of application areas. It also raises concerns on possible risks on the human health and safety and the environment, during the entire life cycle of ENMs and nano-enabled products.

The above concerns need to be taken into account ex-ante, during the early stages of material development, under the safe-by-design (SbD) principles. This work proposes a framework for SbD, using a **general and extendable representation for NMs** suitable for computational applications, and a **data-driven assessment method using Bayesian networks**. These are herein the **essential components for the systematic and efficient screening for new NM designs with improved safety profiles**.

The representation is based on the InChI for NMs, and facilitates the comparison of NM components, which is crucial for calculating NM affinities. A set of extensions to NM InChI aim to address key challenges in computational-friendly NM representation, including the ability to describe statistical distributions and bypass categorical shape descriptors typically used to describe NMs (Blekos et al., 2022). Note that, the authors are part of the NM InChI Working Group and many aspects of these extensions will be included in the proposal for a new InChI standard (InChI Trust, 2022).

The screening framework proposed here depends on the availability of many and reliable data and appropriate QSAR models to fill gaps in nanotoxicity predictions. The platform developed under the NanoInformaTIX H2020 project will contain a variety of

suitable models integrated into workflows for human and environmental risk assessment, providing the perfect playground to evaluate the proposed approach. A simplified version of the proposed framework, without dependency on predictions from external models, will soon be available at the NanoInformaTIX platform.

The proposed framework has the potential to be used in other or extended applications, including the design of new NMs with improved sustainability and functionality profiles. By taking into account these additional factors, we can better formulate the NM design problem and consider a variety of NMs, their production processes and applications.

## Funding

This work is funded via the EU H2020 project “NanoInformaTIX: Development and Implementation of a Sustainable Modelling Platform for NanoInformatics” (grant agreement No. 814426).

## References

- Balraadsing, S., Peijnenburg, W.J.G.M., Vijver, M.G., 2022. Exploring the potential of in silico machine learning tools for the prediction of acute *Daphnia magna* nanotoxicity. *Chemosphere* 307, 135930.
- Blekos, K., Chairetakis, K., Lynch, I., Marcoulaki, E., 2022. Principles for nanomaterial representations to facilitate machine processing and cooperation with nanoinformatics tools, *Journal of Cheminformatics*, accepted
- Furxhi, I., Murphy, F., Mullins, M., Arvanitis, A., Poland, C.A., 2020. Practices and Trends of Machine Learning Application in Nanotoxicology. *Nanomaterials* 10, 116.
- Huang, H.-J., Lee, Y.-H., Hsu, Y.-H., Liao, C.-T., Lin, Y.-F., Chiu, H.-W., 2021. Current Strategies in Assessment of Nanotoxicity: Alternatives to In Vivo Animal Testing. *International Journal of Molecular Sciences* 22, 4216.
- InChI Trust, 2022. Nanomaterials Working Group. <https://www.inchi-trust.org/nanomaterials/>
- Kim, I., Viswanathan, K., Kasi, G., Sadeghi, K., Thanakkasaranee, S., Seo, J., 2020. Preparation and characterization of positively surface charged zinc oxide nanoparticles against bacterial pathogens. *Microb Pathog* 149, 104290.
- Lynch, I., Afantitis, A., Exner, T., Himly, M., Lobaskin, V., Doganis, P., et al., 2020. Can an InChI for nano address the need for a simplified representation of complex nanomaterials across experimental and nanoinformatics studies?. *Nanomaterials*, 10(12), p.2493.
- Marcoulaki, E., de Ipiña, J.M.L., Vercauteren, S., Bouillard, J., Himly, M., Lynch, I., et al., 2021. Blueprint for a self-sustained European Centre for service provision in safe and sustainable innovation for nanotechnology. *NanoImpact*, 23, p.100337.
- NanoInformaTIX: Development and Implementation of a Sustainable Modelling Platform for NanoInformatics (H2020-NMBP-TO-IND <https://cordis.europa.eu/project/id/814426>)
- OECD. Moving Towards a Safe(r) Innovation Approach (SIA) for more sustainable nanomaterials and nano-enabled products, ENV/JM/MONO(2020)36/REV1.
- Shandilya, N., Marcoulaki, E., Barruetabena, L., Llopis, I.R., Noorlander, C., Jiménez, A.S., et al., 2020. Perspective on a risk-based roadmap towards the implementation of the safe innovation approach for industry. *NanoImpact*, 20, p.100258.
- Singh, A.V., Ansari, M.H.D., Rosenkranz, D., Maharjan, R.S., Kriegel, F.L., Gandhi, K., Kanase, A., Singh, R., Laux, P., Luch, A., 2020. Artificial Intelligence and Machine Learning in Computational Nanotoxicology: Unlocking and Empowering Nanomedicine. *Advanced Healthcare Materials* 9, 1901862.
- So, S., Rho, J., 2019. Designing nanophotonic structures using conditional deep convolutional generative adversarial networks. *Nanophotonics* 8, 1255–1261.
- WHO guidelines on protecting workers from potential risks of manufactured nanomaterials, 2017
- Winkler, D.A., 2020. Role of Artificial Intelligence and Machine Learning in Nanosafety. *Small* 16, 2001883.

# Artificial neural network modelling for the prediction of the deactivation of CaO-based adsorbents in the calcium looping process for CO<sub>2</sub> capture

Rui M. Filipe,<sup>a,b</sup> Sérgio González-De-La-Cruz,<sup>c</sup> Adrian Bonilla-Petriciolet,<sup>c</sup>  
Carla C. Pinheiro<sup>d,e</sup>

<sup>a</sup>*Instituto Superior de Engenharia de Lisboa, Instituto Politécnico de Lisboa, R. Conselheiro Emídio Navarro 1, 1959-007 Lisboa, Portugal*

<sup>b</sup>*Centro de Recursos Naturais e Ambiente, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal*

<sup>c</sup>*Instituto Tecnológico de Aguascalientes, Aguascalientes 20256, Mexico*

<sup>d</sup>*Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal*

<sup>e</sup>*Centro de Química Estrutural, Institute of Molecular Sciences, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal*

## Abstract

Calcium-looping (CaL) is a promising cyclic process for CO<sub>2</sub> capture based on the reversible chemical reaction between CaO-based adsorbents and CO<sub>2</sub> to form CaCO<sub>3</sub>. Currently, the main CaL challenge is the decay of the adsorbents carrying capacity with increasing number of cycles of calcination-carbonation due to the particles sintering and pores blockage.

In this work, different types of CaO-based adsorbents are investigated in a cyclic laboratory scale CaL unit for CO<sub>2</sub> capture using a real flue gas from a cement plant and a synthetic gas mixture, both with same CO<sub>2</sub> content.

An artificial neural network (ANN) model is developed to predict the adsorbents' deactivation along the cycles, based on their composition and on the type of gas used. A good agreement between the experimental and predicted values was obtained, demonstrating that ANN can be successfully used to predict the adsorbents' deactivation based on the materials composition and type of gas.

**Keywords:** Calcium looping, CaO-based adsorbents, CaO deactivation, neural networks.

## 1. Introduction

The calcium looping (CaL) process is a promising CO<sub>2</sub> capture technology for industrial plants, based on the reversible chemical reaction between CaO and CO<sub>2</sub> to form CaCO<sub>3</sub>, which has received great attention in recent years. CaL process is based on the cyclic calcination-carbonation of CaO-based adsorbents, usually natural limestones (CaCO<sub>3</sub>), to capture CO<sub>2</sub> from a flue gas. The adsorbents can be reused for several cycles, but their CO<sub>2</sub> carrying capacity decreases along the cycles due to the particles sintering and pores blockage (Abreu et al. 2020; Teixeira et al., 2020). CaL is currently a potential sustainable option for emissions' reduction in the cement industry because the purged deactivated CaO-based materials can be conveniently reused and integrated in the clinker production.

Despite the synergy between the CaL process and the cement production plants, the decay of the adsorbents' CO<sub>2</sub> carrying capacity is a problem that needs to be tackled for improving the operation and efficiency of the CaL process. A fresh adsorbent make-up stream may be used to maintain the adsorbent activity and the desired CO<sub>2</sub> capture level. The degree of deactivation of each type of adsorbent material depends, among other factors, on its chemical composition as well as on the flue gas composition, and it may be assessed in the laboratory through a sequence of several calcination-carbonation cycles. A tool to predict the deactivation along the cycles is thus urgent and valuable for the CaL process at industrial scale, as it would avoid the laboratory testing burden and it would also allow the estimation of the flowrate of a purge stream of the deactivated CaO-based adsorbent, as well as a make-up stream of the fresh adsorbent.

CaO-based adsorbents are usually non expensive materials such as natural geological or waste-based materials precursors. In this work, the use of different samples of both types of CaO-based materials in a cyclic laboratory scale CaL unit for CO<sub>2</sub> capture (Marques et al., 2022) is investigated, using a real flue gas from a cement plant and a synthetic gas mixture, both with the same CO<sub>2</sub> content.

An artificial neural network (ANN) is developed to obtain a model that predicts the adsorbents' deactivation along the cycles, based on their composition and on the type of gas used.

## 2. Methodology

### 2.1. Adsorbent deactivation data

The data reported by Marques et al. (2022) for six different samples of Ca-adsorbents comprising four natural limestones (LA, LB, LC, and LD) and two waste marble powders (WMA, and WMB), was used in this work. The oxide content of the adsorbents is shown in Table 1. The samples were tested in a fluidized bed reactor using two different gases for the carbonation step, a real flue gas (FG) from a cement plant chimney (pressurized to a bottle) and a synthetic gas (SG) mixture of nitrogen and air, that replicates a real carbonation atmosphere, both with 15 % of CO<sub>2</sub>. The calcination was always performed with a SG mixture (70 % CO<sub>2</sub> balanced in air) to replicate realistic calcination conditions. The chemical composition of the gases and a simplified diagram of the experimental set-up may be found in Marques et al. (2022).

The deactivation of the adsorbent was evaluated using the amount of CO<sub>2</sub> captured by unit mass of adsorbent as indicator, the carrying capacity (CC), over ten cycles of calcination and carbonation. The characterization of the gases used, and the experimental procedure details are given by Marques et al. (2022) and will not be reproduced here.

Table 1. Adsorbents chemical composition (wt. %, adapted from Marques et al. (2022))

	SiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	Fe <sub>2</sub> O <sub>3</sub>	CaO	MgO	K <sub>2</sub> O
LA	9.00	3.69	1.72	45.52	0.99	0.72
LB	14.78	3.94	1.80	40.57	2.02	0.92
LC	0.49	0.26	0.11	54.90	0.46	0.06
LD	2.31	0.85	0.60	45.89	0.78	0.12

WMA	0.60	0.02	0.05	55.33	0.25	0.02
WMB	3.38	0.60	0.28	52.22	0.79	0.30

A set of 120 experimental results is available. After a careful analysis of the data, two points were considered as outliers and removed from the data set, resulting in a total of 118 points available for training, testing and validation of the ANN.

## 2.2. Model development

This work aims at developing an ANN to correlate the adsorbent characteristics and its working conditions with the capacity of capturing CO<sub>2</sub>. Eight inputs were selected for the ANN: adsorbent oxide composition (SiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, Fe<sub>2</sub>O<sub>3</sub>, CaO, MgO, and K<sub>2</sub>O), number of the cycle, and type of gas used. The output of the ANN is the carrying capacity.

A feedforward backpropagation neural network model using Levenberg-Marquardt backpropagation and a hyperbolic tangent sigmoid activation function were used. The data was divided for training (70 %), testing (15 %), and validation (15 %) using interleaved indices, with the same sets being used to train, test, and validate each network. As there are no fixed rules for the selection of the ANN size and configuration, a preliminary analysis was performed to identify the simplest architecture giving adequate results. Different ANN architectures were created, by varying the number of neurons and hidden layers. The number of internal layers tested was 1 and 2, with the number of neurons varying between 4 and 20 for layer 1, and between 1 and 3 for layer 2.

The different architectures were tested using MATLAB Deep Learning Toolbox. To increase the probabilities of achieving better performance the model was trained 500 times for each architecture, and the performance evaluated using mean squared error (MSE) as criteria. The best performing case for each architecture was then used to evaluate the influence of the number of neurons and layers in ANN performance, while also checking for an adequate description of the process. The criteria used to select the final ANN architecture was the minimization of the number of layers and neurons, while having a low value of MSE and providing a proper description of the process.

Statistical parameters were calculated to assess ANN model predictions, and included the coefficient of correlation (R) and the fitting error ( $e_i$  (%)) for each data point  $i$ , defined as:

$$e_i = \frac{(y_i^p - y_i)}{y_i} 100 \quad (1)$$

where  $y_i^p$  is the value predicted by the ANN model and  $y_i$  the experimental value.

## 3. Results and discussion

### 3.1. ANN architecture

Several tests were made to identify the proper architecture to be used, varying the number of layers, and the number of neurons in each layer. Although different results were obtained depending on the conditions used to train the ANN, as the number of times the network is trained, or the sets used for each task, for example, only a small improvement was obtained when using two layers instead of one. Furthermore, no significative improvement was observed in the performance indicator (MSE) when the number of neurons is increased over a certain number.

As an illustrative example, Figure 1 depicts the variation of the minimum value of MSE obtained for each ANN architecture, with a varying number of neurons and internal



layers. The small improvements observed, particularly for more than nine neurons, indicate that a simpler ANN, with only one layer, can still provide adequate results, while requiring a minimum number of parameters to be estimated. In view of this, an architecture with only one layer and nine neurons was selected for this problem, resulting in an ANN with eight inputs, nine neurons and one output (8:9:1).

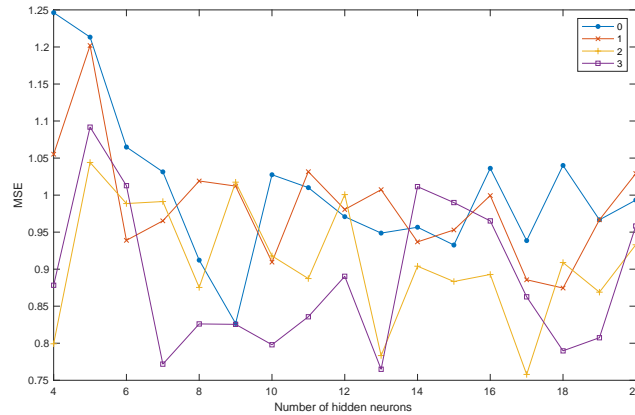


Figure 1. Mean squared error for different ANN architecture and number of hidden neurons in the first layer (the legend indicates the number of neurons in the second layer).

### 3.2. ANN performance

Figure 2 shows the fitting performance for the three subsets used, individually and for all the data used, for the architecture 8:9:1. An overall good behavior is observed, with R values greater than 0.99 for all the cases. The histogram in Figure 3 shows a normal distribution, with most of the estimated points having a deviation lower than 0.34. The calculated overall deviations are comprised in the interval  $[-3.2, 2.8]$  and the error ( $e_i$ ) in the interval  $[-18, 14]$  %.

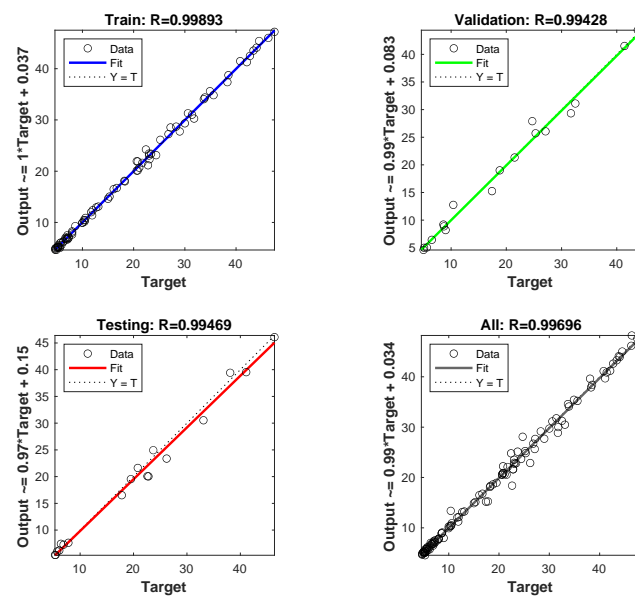


Figure 2. Fitting performance for the sets used.

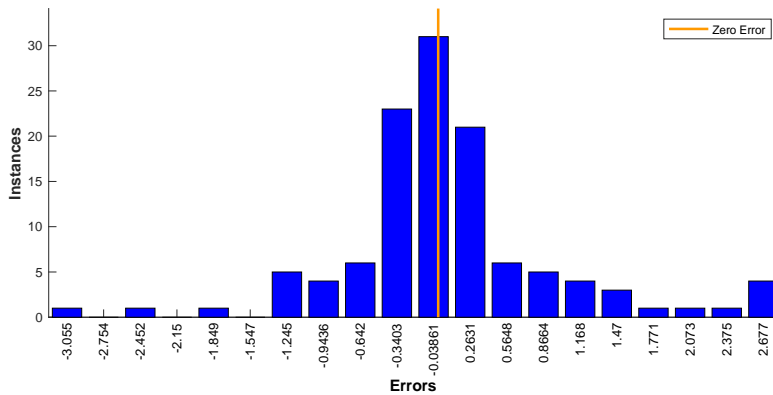


Figure 3. Histogram of error obtained in fitting.

Figure 4 and Figure 5 depict the evolution of the adsorbents carrying capacity along the carbonation-calcination cycles. As expected, a decay over the cycles is observed, which is justified by adsorbent particles sintering and pores blockage. It is also observed that each adsorbent has different initial and final carrying capacities, and deactivation profile depending on the adsorbent composition and type of gas used.

The figures show a good agreement between the target measured values ( $y$ ) and the predicted values ( $y^p$ ) for most of the cases, with no apparent signs of overfitting. Some of the major discrepancies observed may be justified by the inclusion of some experimental data that would be considered as outliers, in a stricter data validation process.

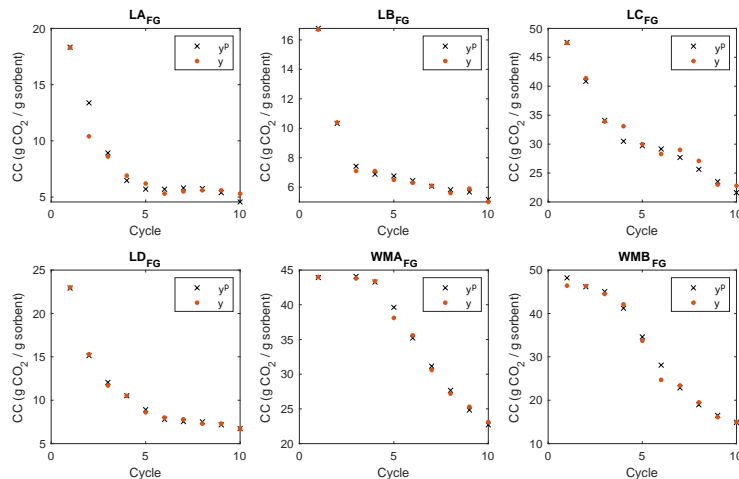


Figure 4. Experimental ( $y$ ) and predicted ( $y^p$ ) carrying capacity (CC) for each sorbent as a function of the number of cycles using real flue gas.

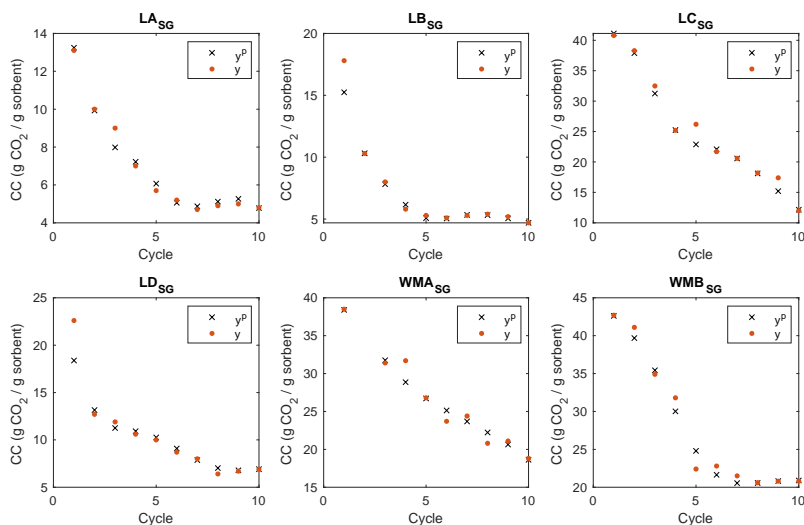


Figure 5. Experimental ( $y$ ) and predicted ( $y^p$ ) carrying capacity (CC) for each sorbent as a function of the number of cycles using synthetic gas.

#### 4. Conclusions

An ANN was developed to predict CaO-based adsorbents deactivation along the Ca-looping cycles. A study on the architecture of the ANN was performed, with no significant improvements observed above a minimum of nine neurons, even when more complex configurations were used. A good agreement between the experimental and the predicted values of the adsorbents' CO<sub>2</sub> carrying capacity was obtained for the tested materials, demonstrating that the ANN is sensible to the input variables, and can differentiate among samples composition, type of gas used, and number of cycles applied to each sorbent. The ANN model developed is a powerful tool, that can now be used to assist the selection of new adsorbents by predicting the behavior over several cycles. It may also provide useful information for models used in process design, optimization, and operation. Future work will investigate how extra information supplied to the ANN affects performance.

#### Acknowledgments

The financial support from FCT under projects UIDB/00100/2020, UIDP/00100/2020, UIDB/04028/2020, UIDP/04028/2020, and PTDC/EAM-PEC/32342/2017, and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 778168, is gratefully acknowledged.

#### References

- M. Abreu, P. Teixeira, R. Filipe, L. Domingues, C.I.C. Pinheiro, H. Matos, 2020, Modeling the deactivation of CaO-based sorbents during multiple Ca-looping cycles for CO<sub>2</sub> post-combustion capture, *Computers & Chemical Engineering*, 134, 106679.

- L.M. Marques, S.M. Mota, P. Teixeira, C.I.C. Pinheiro, H.A. Matos, (2022), Ca-looping process using wastes of marble powders and limestones for CO<sub>2</sub> capture from real flue gas in the cement industry, *Journal of CO<sub>2</sub> Utilization*, (submitted).
- P. Teixeira, I. Mohamed, A. Fernandes, J. Silva, F. Ribeiro, C.I.C. Pinheiro, 2020, Enhancement of sintering resistance of CaO-based sorbents using industrial waste resources for Ca-looping in the cement industry, *Separation and Purification Technology*, 235, 116190.



# Dynamic operation for the effective use of green hydrogen in Power-to-X value chains

Michael Mock,<sup>a</sup> Hannes Lange,<sup>a</sup> Isabell Viedt,<sup>a</sup> Kumar Rajan Gopa,<sup>a</sup> Jonathan Mädler,<sup>a</sup> Leon Urbasa

<sup>a</sup> *TU Dresden, Chair of Process Control System and Process Systems Engineering Group, Helmholtzstr. 14, 01069 Dresden, Germany*  
*michael.mock@tu-dresden.de*

## Abstract

In the upcoming decades, the integration of renewable energies into our continuous process routes, by producing and processing green hydrogen, is of paramount importance to combat climate change. To include green hydrogen in conventional process routes large hydrogen storages or dynamic process control strategies for hydrogen processing are needed. In this work a dynamic process control strategy for a Power-to-Ammonia (PtA) value chain is studied. A PtA process plant and respective process control strategies to adjust the plant load are conceptualized. To study the feasibility, a Haber-Bosch reactor system is modeled and implemented in MATLAB 2022b. Simulation case studies are carried out to emulate load changes. The simulation shows that the conceptualized process enables an intermittent production of ammonia based on green hydrogen. Further studies with more detailed models and refined process control strategies are planned for a detailed analysis to optimize the efficiency and flexibility of the production.

**Keywords:** Power-to-X, Power-to-Ammonia, Requirements Analysis, Dynamic Simulation, Renewable Energies

## 1. Introduction

The integration of renewable energies, like wind energy or photovoltaics, into our existing chemical production lines is necessary to combat climate change and reduce our economic dependency on fossil resources (IEA, 2021). Power-to-X (PtX) technologies offer the possibility to integrate fluctuating energy into our value chains, by producing and processing green hydrogen. To integrate intermittent energy in conventional production lines large electricity storages or large hydrogen storages are necessary to decouple the subsystems. Since both electrical storage and hydrogen storage are expensive and exhibit increasing losses over time, flexible PtX value chains through dynamic process control is a promising approach towards integrating renewable energies in the chemical industry (Moradi, 2019). The aim of this work is to elaborate a plant and control concept for the dynamic control of a PtX value chain. The implementation and simulation of the developed concepts will be used to study whether the PtX value chain can adapt to the load changes of sustainable energy sources without large-scale electricity or large-scale intermediate hydrogen storage. Chapter 2 will shortly describe the methodology used in this work. The developed requirements, concepts and simulation will be specified in chapter 3. The conclusion of the work can be found in chapter 4.

## 2. Methodology

The plant and process control concept are developed for a selected PtX use case, namely Power-to-Ammonia (PtA), based on the available literature and a preceding requirements analysis. The concept primarily includes conventional applications to study the possibility to use established technologies for PtX value chains. Subsequently, simplified concepts are implemented and simulated in MATLAB 2022b. In this work, the focus is on the simulation of a Haber-Bosch reactor system. Necessary parameters for the reactors are taken from Khademi et al. (2017) to create models for dynamic tubular Haber-Bosch reactors. The resulting PDE system of the reactors models are discretized to derive an ODE system and are solved using the ode23t solver.

Case studies are simulated in which different load change rates are simulated and the dynamic response of the reactor system is considered. The simulation results, such as temperature and concentration profiles and molar flow rates, are validated by those from literature, but due to missing data, validation with real plant data is omitted. The time-dependent temperature and concentration profiles under load variations are evaluated to assess whether dynamic operation is feasible with the input of sustainable energy. The focus here is on whether the plant concept can maintain the permitted operating parameters even under fluctuating input conditions and different load change rates. In later iterations the complete PtA value chain with added electrolyzers, a small-scale intermediate hydrogen tank and auxiliary components shall be modeled and simulated. The determined range of load change rates can subsequently be used in the dynamic PtA value chain.

## 3. Methodology

Since PtX technologies include a wide range of production routes a specific value chain must be chosen. Ammonia production accounts for about three-quarters of the hydrogen consumption in chemical production lines (IEA, 2021). Therefore, the Power-to-Ammonia (PtA) value chain is chosen as an example to conceptualize and evaluate the flexible operation of PtX technologies through dynamic process control.

### 3.1. Use-case for main energy source, secondary energy source and plant location

The process plant design and process control strategies have to fit the chosen use-case of the PtA value chain. In the first step a use-case is defined, considering the main energy source, a potential secondary energy source and the location of the plant.

Wind energy is chosen as the main energy source, since it is the most widely used energy source for PtX applications and offers a high number of nominal operational hours with 1400-5000 h/a (Chehade, 2019). A grid connection is chosen as a secondary energy source. This enables the use of a stand-by operation, which is necessary for safe and economical operation with fluctuating energy supply without large electricity or hydrogen storage. The grid connection makes the use of a stand-by operation during times of low wind energy availability possible. The main disadvantage is that this can lead to the use of electricity from fossil resources to maintain the stand-by operation.

Finally, the process plant location is chosen to be next to the energy provider. While a positioning next to ammonia consumers or existing ammonia production facilities offer advantages like existing infrastructure, ammonia is far more suitable for long distance transport than hydrogen (IEA, 2021).

### 3.2. Plant requirements for a dynamic Power-to-Ammonia value chain

The plant, especially the electrolyzer, must be able to adjust to the load gradient of a wind turbine in order to respond to incoming energy fluctuations. It must have a large partial

load range in which the plant can be safely operated efficiently and economically. To minimize the potential use of fossil fuels, the energy consumption in stand-by operation should be as low as possible. The generated heat should be integrated in the plant. The CAPEX should be kept as low as possible, as this is a significant cost factor for the ammonia produced in small- and medium-scale plants (Chehade, 2021). The hydrogen tank necessary to act as a buffer for the different system dynamics should be as small as possible to reduce CAPEX and storage losses.

### 3.3. Process control requirements for a dynamic Power-to-Ammonia chain

The operating conditions have to be kept within the permitted value range to prevent the degradation of plant equipment and catalyst materials. The maximum warm start-up time for minimal operation of the ammonia synthesis should be less than 1 hour. If the available energy falls below minimal load, the plant transitions to stand-by operation automatically and holds the operating conditions with inert gas. The stand-by operation can be upheld for a maximum of 20 hours. Energy forecast data is used to decide, whether the plant will be kept in stand-by operation or is shut down.

### 3.4. Concept of the Power-to-Ammonia value chain using dynamic process control

The PtA plant and their respective process control strategies for the developed use-case are conceptualized, modeled and simulated. The goal is to study the plant and the respective process control concept in terms of flexibility and efficiency.

#### 3.4.1. The Power-to-Ammonia plant concept

In simplified form, the plant concept consists of a technology selection for a water electrolysis, a nitrogen recovery and a reactor concept for the Haber-Bosch synthesis for ammonia production. The concept for the process control strategies consists of strategies to hold the operating conditions in the permitted value ranges as well as strategies to change the plant load during operation.

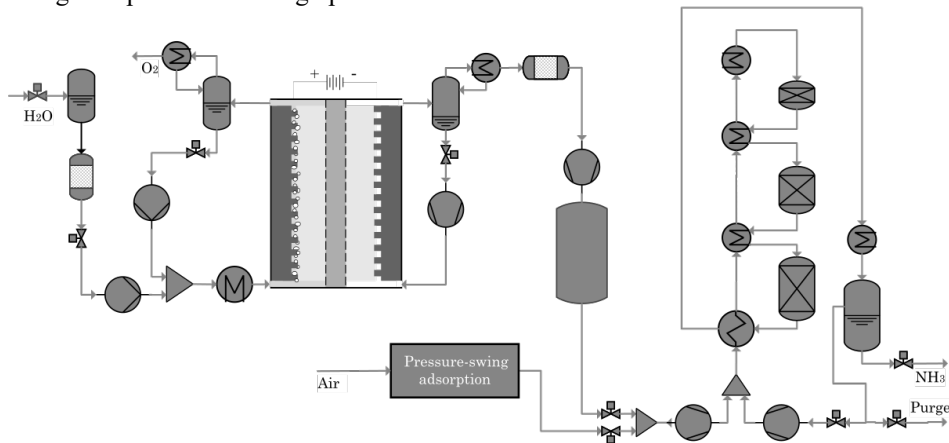


Figure 1: Simplified process flow diagram of the PtA plant concept

The Polymer Membrane Electrolysis (PEMEL) was chosen for the hydrogen production. The advantage of the PEMEL is the flexibility of the plant operation. Both the cold start with 10–15 min and the warm start with <10 s allow the PEMEL to start up quickly. A heating curve of 0.3–1 K/min can be achieved (Buttler, 2018; Hermesmann, 2021). PEMEL systems achieve load gradients of 10–90 %/s with load flexibility of 0–160 % if the power electronics are also designed for it (Grigoriev, 2020; Milanzi, 2018; IEA, 2019). Likewise, PEMELs are more compact than other electrolysis systems for the same hydrogen output (Buttler, 2018). On the other hand, the PEM has higher requirements for



water purity and is also more expensive to procure (1390 – 2320 €/kWel) than an alkaline electrolysis (AEL) with a procurement price of 620 – 1200 €/kWel (Buttler, 2018; Hermesmann, 2021).

For nitrogen production a pressure swing adsorption is used. A minimum load range of 30 %, fast start-up behavior and efficient dynamic operation makes the compact PSA attractive for operating small- and medium-scale PtA systems (Bañares Alcántara; 2015). The ammonia synthesis takes place within a reactor cascade of three tubular reactors with interstage cooling. The Haber-Bosch process is catalyzed through an iron based catalyst. While a ruthenium based catalyst is better suited for a dynamic process due to the lower operating pressure and temperature, an iron based catalyst is chosen due to the lower cost (Chehade, 2021).

### 3.4.2. The Power-to-Ammonia dynamic process control concept

The process control focuses on the control strategy for the reactor load, because the reactors are the production speed determining components. Nitrogen production is not considered in detail and a constant nitrogen supply is assumed.

The electrolyzers are started up within 10 minutes and begin to feed the produced hydrogen into the buffer tank. Temperature in the reactor system is continuously ramped up with 20 K/h to the operating conditions of about 400 °C and the reactor is pressurized to 250 bar while the electrolyzers are starting up and already producing. The produced hydrogen is temporarily stored in the buffer tank until the operating conditions of the ammonia synthesis with 690 K reactor inlet temperature and 286 bar reactor inlet pressure are reached. The load of the Haber-Bosch reactors is controlled by adjusting the hydrogen input, since hydrogen production accounts for about 85 – 95 % of the total energy consumption (Khademi, 2017). With the assumed constant nitrogen supply a feedforward control system is used, which controls the hydrogen input to the reactor system based on the available power, the fill level of the hydrogen buffer tank and the hydrogen content in the recirculation. Adjusting the stoichiometry of the reaction to the load change through hydrogen input has been established as the most efficient method to control the Haber Bosch reaction (Khademi, 2017). The nominal ramp-up rate of the hydrogen input is set to 50 % of nominal load per hour (Viedt, 2022). The reactor pressure is maintained by the recycle compressor. In stand-by operation the temperature is maintained by the heating element upstream of the reactor.

### 3.5. Simulation of the Power-to-Ammonia value chain

The concept is modeled and simulated in MATLAB 2022b. Due to the fast dynamics of the PEM electrolysis, the electrolyzer models are simplified by assuming constant material parameters and quasi-stationary conditions in the cells, while neglecting the influence of water vaporization, gas cross-permeation and the heat capacity of the membrane and electrodes. The reactors are modeled as dynamic tubular flow reactors. Assuming constant material parameters, the one-dimensional material and energy balances of the adiabatic reactors result in:

$$\frac{dc_i}{dt} = v \frac{dc_i}{dz} + c_i \frac{dv}{dz} + R_i \quad (1)$$

*for i = H<sub>2</sub>, N<sub>2</sub>, NH<sub>3</sub>*

$$\frac{dc_{Ar}}{dt} = v \frac{dc_{Ar}}{dz} + c_{Ar} \frac{dv}{dz} \quad (2)$$

$$\frac{dT_R}{dt} = v \frac{dT_R}{dz} + T_R \frac{dv}{dz} + \frac{r(-\Delta h_R)}{\rho c_p} \quad (3)$$

The reactors are discretized in equidistant  $z$ -locations to convert the PDE system into an ODE system. A simplified convergence study shows that the temporal behavior changes negligibly with a number above 40 equidistant discretization points. The load change rate is varied from 50 500 % per hour to study the influence of the load change rate on the dynamic operation of the Haber-Bosch reactors.

### 3.5.1. Simulation results

The simulation of the reactor system provides results for the temperature and concentration curves in the Haber-Bosch reactors. For model validation, the local concentration and temperature profiles of the steady-state operating points were compared with the steady-state simulation of different Haber Bosch reactor systems by Khademi et al. (2017). The continuous operation delivers satisfactory results in this respect. The dynamic response of temperature and concentrations is difficult to compare with other literature sources due to low availability of data and different design and model assumptions. However, due to the simplifying assumptions, it is expected, that the simulated plant has a significantly faster dynamic response than real plants or more complex simulations. Figure 2 shows the temperature curve and ammonia concentration curve for 50, 200 and 500 % load change rate per hour of the first reactor during transition from zero load to nominal load.

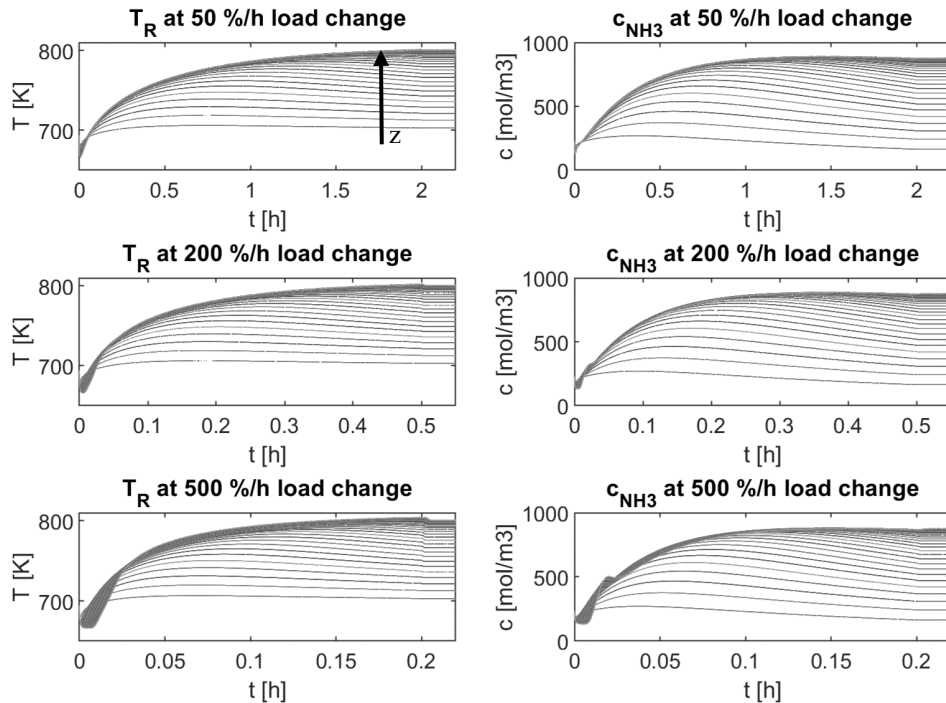


Figure 2: Temperature profile and ammonia concentration profile of the first reactor during transition from nominal load to minimal load with 50, 200 and 500 %/h load change rate

It can be seen, that the maximum reactor temperature slightly increases with increasing load change rate. At load change rates above 200 % per hour the temperature peak in the reactor exceeds the allowed maximum operation temperature of 805 K.

#### 4. Conclusion

In summary, the operation of conventional Haber-Bosch reactors for PtA value chains can be done with load change rates of up to 200 % of nominal load per hour. This information will be used in further simulations of the entire plant and the dynamic control strategies to enable the integration of sustainable energy sources into chemical production chains even without large-scale energy storage or intermediate hydrogen storage. The simplified simulation study showed that the dynamic operation of Haber-Bosch reactors is possible without exceeding or falling below the permitted temperature range. The simulation study is thus only applicable for a limited value range. The assumptions of adiabatic operation and the neglect of loss terms and thermal masses in the plant simulation provide faster simulated system dynamics than reality and need to be included in more advanced simulations. Complex models must be created and real plant data must be used for simulation validation to provide a more reliable simulation.

Another important aspect is the choice of concept. The choice of an iron-based catalyst for ammonia synthesis represents the conventional technology well, but is not well suited for dynamic operation, due to the high operating temperature and pressure required. The potential use of new catalyst and reactor systems, such as ruthenium-based catalysts and adsorption reactors would increase the usability of the plant concept for dynamic operation, however at the cost of an increased CAPEX and a more complex control system.

#### Acknowledgements:

We would like to thank the German Federal Ministry of Education and Research and the Project Management Agency Jülich for their financial support within the framework of the eModule research project (FKZ 03HY116A) of the H2Giga lead platform.

#### References

- R. Bañares-Alcántara, 2015, Analysis of islanded ammonia-based energy storage systems, University of Oxford
- A. Buttler, 2018, Current status of water electrolysis for energy storage, grid balancing and sector coupling via power-to-gas and power-to-liquids: A review, *Renewable and Sustainable Energy Reviews* 82
- G. Chehade, 2021, Progress in green ammonia production as potential carbon-free fuel, *Fuel* 299
- D.S. Falcão, 2020, A review on PEM electrolyzer modelling: Guidelines for beginners, *Journal of Cleaner Production* 261
- R. Fang, 2019, Control strategy of electrolyzer in a wind-hydrogen system considering the constraints of switching times, *International Journal of Hydrogen Energy*
- M. Hermesmann, 2021, Promising pathways: The geographic and energetic potential of power-to-x technologies based on regeneratively obtained hydrogen, *Renewable and Sustainable Energy Reviews* 138
- International Energy Agency, 2019, The future of hydrogen
- International Energy Agency, 2021, Global Hydrogen Review 2021
- M.H. Khademi, 2017, Comparison between three types of ammonia synthesis reactor configurations in terms of cooling methods, *Chemical Engineering Research and Design* 128
- H. Kim, 2013, One-dimensional dynamic modeling of a high-pressure water electrolysis system for hydrogen production, *International Journal of Hydrogen Energy* 38
- S. Milanzi, 2018, Technischer Stand und Flexibilität des Power-to-Gas-Verfahrens
- R. Moradi, 2019, Hydrogen storage and delivery: Review of the state of the art technologies and risk and reliability analysis, *International Journal of Hydrogen Energy* 44.23
- I. Viedt, 2022, Anforderungen aus der dynamischen Nutzung von grünem Wasserstoff in Power-to-X-Wertschöpfungsketten, *Automation* 2022

# Evaluation of two-stage stochastic programming applied to the optimization of crude oil operations scheduling

Tomas Garcia Garcia-Verdier <sup>a,b</sup>, Gloria Gutierrez <sup>a,b</sup>, Carlos Mendez <sup>c</sup>, Cesar de Prada <sup>a,b</sup>

<sup>a</sup>*Dpt. of Systems Engineering and Automatic Control, University of Valladolid, c/ Prado de la Magdalena 3-5, Sede Mergelina EII, 47011 Valladolid, Spain*

<sup>b</sup>*Institute of Sustainable Processes, Dr. Mergelina s/n, 47011 Valladolid, Spainw*

<sup>c</sup>*Center for Advanced Process Systems Engineering (CAPSE), INTEC (UNL - CONICET), Industrial Engineering Dpt. (FIQ-UNL), Güemes 3450, 3000 Santa Fe, Argentina*

## Abstract

This paper addresses the optimization of crude oil operations scheduling in a marine-access refinery, considering uncertainty in the arrival date of the ships that supply the crudes. Furthermore, we evaluate the performance of a two-stage stochastic mixed-integer nonlinear programming (MINLP) model based on continuous-time representation. For this purpose, we calculate the Expected Value of Perfect Information (EVPI) and the Value of the Stochastic Solution (VSS), which allow us to assess and compare the solution of the stochastic model against solutions obtained from deterministic models. Finally, we consider the Conditional Value-at-Risk (CVaR) measure as the objective function and evaluate the solutions obtained for different risk levels.

**Keywords:** Stochastic optimization; Continuous-time representation; Crude oil scheduling; Expected Value of Perfect Information; Conditional Value-at-Risk.

## 1. Introduction

In this paper, the problem under study is the optimization of crude oil operations scheduling in a refinery with marine access. In addition, we address two different topics related by means of stochastic programming.

First, we evaluate the solution obtained from a two-stage stochastic MINLP model based on continuous-time representation by comparing it with the one obtained from a deterministic model. We are interested in finding out if the deterministic solution is close to the stochastic optimum or, on the contrary, if it is worth using stochastic optimization in spite of the increase in computational effort. Another question that arises is how much we are willing to pay to obtain more information about the uncertainties. To answer these inquiries, there are two measures: the value of the stochastic solution (VSS) and the expected value of perfect information (EVPI), which are explained throughout the article. Second, we analyze the result of including risk management, that is, the effect of deviations from average values on the quality of the solution. For this purpose, we employ

the Conditional Value-at-Risk (CVaR) measure as the objective function and evaluate the solutions obtained for different risk levels.

The rest of the paper is structured as follows. The definition of the problem is given in Section 2. The mathematical formulation is described in Section 3. The concepts EVPI and VSS are defined in Section 4. The risk management method is described in Section 5. The solution strategy is mentioned in Section 6. Then, problem instances and computational results are reported in Section 7. Finally, conclusions are drawn in Section 8.

## 2. Problem definition

As stated before, we tackle the optimization of crude oil operations scheduling in a refinery, which is supplied with crude oil by ships. Before beginning with the description of the problem, it is convenient to clarify the meaning of "crude operations" and "operations scheduling".

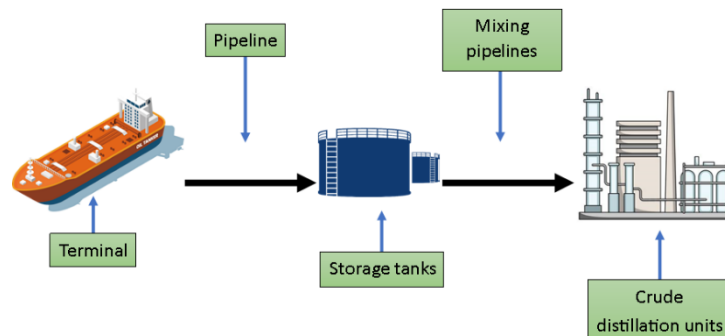
The term "crude operations" refers to the operations involved in the supply, storage, and processing of crude oil. Then, when we talk about operations scheduling, we refer to the process of allocating resources to operations and sequencing their execution to comply with the production plan.

Considering these definitions, we can determine that the optimization of crude operations scheduling consists of deciding the best way to operate the terminal-refinery system, considering the management of the arrival and unloading of vessels, crude oil inventory in tanks, and loading of CDUs. Moreover, in this case, the crude oil supply availability is subject to uncertainty due to the impact of weather conditions on ship arrival dates.

Figure 1 shows a typical scheme of a refinery with a marine terminal. The crude oil is received through this terminal and stored in the refinery tanks; these two areas (terminal and tanks) are connected by a pipeline. Then, from the stored crude, the feed blends are formed, which must meet certain quality specifications. Finally, the blends obtained are loaded into the crude distillation units (CDUs) through a mixing pipeline system.

In this article, we address the problem by means of a two-stage stochastic programming with recourse model (Birge & Louveaux, 2011). In these types of models, we must define two kinds of decision variables: first-stage variables ("here and now") which have to be implemented now and affect all future decisions, and second-stage variables that would be implemented later on when more information about the process becomes available (recourse, "wait and see" variables).

For this case, the first-stage variables are: allocation of tanks to CDUs, total volumes transferred to CDUs, and start-time, end-time, and duration of slots. With respect to the second-stage ones, the following were defined: variables related to the activities carried out at the marine terminal, inventory level in tanks, and amount of each type of crude oil



transferred from tanks to CDUs. Finally, the uncertainty in the supply is represented by a discrete set of scenarios that contemplate different arrival dates.

### 3. Model formulation

We use the model developed in (García-Verdier et al., 2022), which consists of a two-stage stochastic MINLP model with continuous-time representation. In this model, the first and second-stage variables are solved jointly, i.e., the deterministic equivalent program is solved. In order not to exceed the allowed page limit, we do not show the complete mathematical formulation.

A relevant feature of this model is that the precedence between ships does not depend on the order of the elements of the set “Vessels” (Cerdá et al., 2015), nor is there a pre-assignment of slots for each ship (Reddy et al., 2004). For this case, it has been implemented the concept of predefined precedence, which consists in considering the set of slots preordered, and then, the optimization algorithm is in charge of assigning each ship to some of those slots (Gómez Palacín, 2020; Palacín et al., 2019).

#### 3.1. Objective function

To define the objective function, the cost associated with each scenario is calculated by using (1). The first summation comprises the costs due to the difference between processed volume and required demand in each CDU and constitutes the first-stage cost. The second summation involves the costs due to demurrage and departure tardiness of vessels and represents the second-stage cost.

$$z e_e = \sum_u (COP_u * op_u + CSP_u * sp_u) + \sum_b (CDMG_b * dm g_{b,e} + CTDN_b * tdn_{b,e}) \quad \forall e \in E \quad (1)$$

Where,

$COP_u$  = cost due to positive difference between processed volume and required demand by a CDU.

$CSP_u$  = cost due to negative difference between processed volume and required demand by a CDU.

$CDMG_b$  = demurrage or sea waiting cost.

$CTDN_b$  = departure tardiness cost.

$op_u$  = positive difference between processed volume and required demand by CDU u.

$sp_u$  = negative difference between processed volume and required demand by CDU u.

$dm g_{b,e}$  = demurrage of vessel b under scenario e.

$tdn_{b,e}$  = departure tardiness of vessel b under scenario e.

Finally, the objective function consists of minimizing the cost associated with the first stage and the expected cost of the second stage, considering all scenarios (2). The parameter  $\pi_e$  represents the probability of occurrence of scenario e.

$$MIN \sum_e \pi_e * z e_e \quad (2)$$

### 4. Assessment of the stochastic solution

As mentioned above, one of the objectives of the paper is to evaluate whether two-stage stochastic programming offers any advantage over simpler deterministic approaches. To this end, (Birge & Louveaux, 2011) proposed the value of the stochastic solution (VSS)

and the expected value of perfect information (EVPI) but before defining both measures it is necessary to explain some concepts.

First, the two-stage stochastic programming model is known as the recourse problem (RP), and its solution as the here-and-now solution. Second, the expected value (EV) problem corresponds to the model obtained by substituting all random variables by their expected values, and the EEV value refers to the expected result of fixing the first-stage variables according to the solution of the EV problem. Finally, if we have perfect information, we can apply the optimal solution corresponding to each scenario. Then, the expected long-term cost, known as the "wait and see" (WS) solution, will be equal to the sum of the costs of each scenario weighted by their probabilities of occurrence. In summary, the WS value represents the expected long-term cost of using the optimal solution for each scenario.

Once these concepts are established, we continue with the definition of the value of the stochastic solution (VSS) and the expected value of perfect information (EVPI). The VSS (3) quantifies the improvement obtained in the objective function when considering the randomness of the uncertainty (RP), versus its weighted average (EEV). The EVPI (4) measures the maximum amount that the decision maker would be willing to pay in exchange for complete and accurate information about the future, i.e., for perfect information.

$$VSS = EEV - RP \quad (3)$$

$$EVPI = RP - WS \quad (4)$$

## 5. Risk management

The approach described above does not evaluate the risk associated with the objective function, i.e., it is risk-neutral. The purpose of this approach is to minimize the expected value in the long run without considering the probability distribution of the objective function. However, it is often important to consider this distribution to reduce the risk that the solution obtained takes extreme values in the most unfavorable scenarios. For this purpose, there are two popular risk measures, Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR). On the one hand, VaR at confidence level  $1-\alpha$  determines the minimum value  $\omega^*$  such that the probability of the objective function taking a value less than  $\omega^*$  is greater than  $1-\alpha$ . On the other hand, CVaR at confidence level  $1-\alpha$  represents the average value of the tail of the distribution, above  $VaR_{1-\alpha}$ . The latter is more useful in optimization problems since it is simple to calculate, and is convex.

The stochastic programming model is reformulated by incorporating constraints (5) and (6), where  $\phi_e$  is an auxiliary variable to evaluate CVaR and the scalar  $\alpha$  corresponds to the CVaR significance level.

$$ze_e - var \leq \phi_e \quad \forall e \in E \quad (5)$$

$$cvar = var + (1/\alpha) * \left( \sum_e \pi_e * \phi_e \right) \quad (6)$$

Finally, the new objective function (7) minimizes CVaR for a given confidence level.

$$MIN \quad cvar \quad (7)$$

## 6. Solution strategy

The stochastic programming problem is solved from its deterministic equivalent, following a solution strategy. The solution strategy for the two-stage stochastic MINLP model consists of two steps. A more detailed description of the strategy is given in (García-Verdier et al., 2022).

### 7. Results

In this section, the resolution of an example is carried out. It consists of a 120-hour scheduling horizon, 5 storage tanks, 2 crude distillation units, and 5 classes of crude characterized by a single property. The arrival of 2 ships is expected. The arrival dates and probabilities for each scenario are detailed in Table 1. The expected departure date is 12 hours after the arrival. The demand for CDU 1 is 100,000 m<sup>3</sup> and for CDU 2 is 65,000 m<sup>3</sup>. The example has been solved using GAMS 39.2.1 software, OsiGurobi for MILPs, and CONOPT 4.19 for NLPs on a computer with Intel Core i7-10510U 2.30 GHz processor and 16 GB RAM.

Table 1. Arrival times and probabilities.

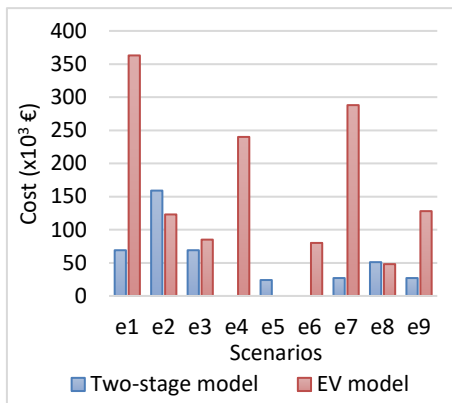
Scenarios	Probabilities	Arrival time (h)	
		Ship 1	Ship 2
1	0.01	10	40
2	0.05	50	40
3	0.01	90	40
4	0.18	10	70
5	0.5	50	70
6	0.18	90	70
7	0.01	10	100
8	0.05	50	100
9	0.01	90	100

#### 7.1. EVPI and VSS

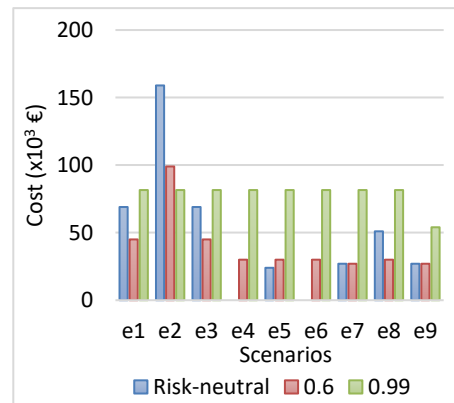
The values obtained for RP, EEV, and WS are shown in Table 2. The RP problem involves 9,345 continuous variables, 1,272 binary variables, and 19,700 constraints; it was solved in 51.58 seconds. In addition, EVPI and VSS values are presented in the same table to analyze the effect of considering uncertainty. Figure 2 (a) compares the costs of each scenario for both the two-stage stochastic programming model and the deterministic EV model.

Table 2. Expected costs associated with proposed models, and values of EVPI and VSS.

RP	WS	EEV	EVPI	VSS
(x10 <sup>3</sup> €)				
24.42	0.74	74.79	23.68	50.37



(a) Two-stage and EV



(b) Risk-neutral, CVaR<sub>0.6</sub>, and CVaR<sub>0.99</sub>

Figure 2. Cost of scenarios.



## 7.2. CVaR

The results obtained for the two-stage stochastic programming model with risk management are shown below. Table 3 shows the VaR and CVaR values at confidence level 0.99 and 0.6. In addition, the expected costs and resolution time for both cases, including the risk-neutral approach, are shown. Figure 2 (b) compares the costs of each scenario in each of the previously mentioned approaches.

Table 3. VaR and CVaR values with confidence levels of 99% and 60%.

$1-\alpha$	VaR $_{1-\alpha}$ ( $\times 10^3 \text{€}$ )	CVaR $_{1-\alpha}$ ( $\times 10^3 \text{€}$ )	Expected cost ( $\times 10^3 \text{€}$ )	Time (s)
0.99	81.56	81.56	81.28	93.67
0.6	30	39.38	33.69	131.06
Risk-neutral (RP)	NA	NA	24.42	51.58

## 8. Conclusions

If we have access to perfect information, the RP solution will improve, and we would be willing to pay up to 23.68 k€.

The VSS indicates that it is worth using the two-stage stochastic optimization since the expected cost, when using the mean values of the uncertain parameters, increases by 50.37 k€.

Based on Figure 2 (a), we can see that, although there are some scenarios where the EV solution is better than the RP solution, the performance of the RP solution is superior. There are scenarios where the EV solution incurs very high costs (e1, e4, and e7). Therefore, it is preferable to apply the RP solution.

Finally, when including risk in the two-stage stochastic model, it is observed that the expected cost increases with increasing risk aversion. However, from Figure 2 (b) we notice that the solutions obtained by considering risk are more "stable" than the risk-neutral solution. Specifically, in scenario 2, the cost of the risk-neutral solution is twice the cost of CVaR at confidence level 0.99.

## Acknowledgements

Financial support received from the Spanish Government with projects a-CIDiT (PID2021-123654OB-C31) and InCo4In (PGC 2018-099312-B-C31), and from European Social Fund.

## References

- Birge, J. R., & Louveaux, F. (2011). Introduction to Stochastic Programming. In Springer Series in Operations Research and Financial Engineering.
- Cerdá, J., Pautasso, P. C., & Cafaro, D. C. (2015). Efficient approach for scheduling crude oil operations in marine- Access refineries. *Industrial and Engineering Chemistry Research*, 54(33), 8219–8238.
- García-Verdier, T. G., Gutiérrez, G., Méndez, C., Palacín, C. G., & de Prada, C. (2022). Minimizing risk in the scheduling of crudes in an oil refinery. *IFAC-PapersOnLine*, 55(7), 809–814.
- Gómez Palacín, C. (2020). Efficient scheduling of batch processes in continuous processing lines. Universidad de Valladolid.
- Palacín, C. G., Méndez, C. A., & de Prada, C. (2019). Slots start-up synchronization with shared resources dependency. *Chemical Engineering Transactions*, 74, 1321–1326.
- Reddy, P. C. P., Karimi, I. A., & Srinivasan, R. (2004). A new continuous-time formulation for scheduling crude oil operations. *Chemical Engineering Science*, 59(6), 1325–1341.

# Generalized First-Principles Suite for Dynamic Modeling of Circular Solid Waste Thermal Treatments: a Sewage Sludge Oxy-Gasification Case Study

Francesco Negri<sup>a,c</sup>, Anna Nova<sup>a,b</sup>, Andrea Galeazzi<sup>a,b</sup>, Francesco Gallo<sup>c</sup>, Flavio Manenti<sup>a,b,\*</sup>

<sup>a</sup>*Politecnico di Milano, CMIC Dept. "Giulio Natta", Piazza Leonardo da Vinci 32, Milan 20133, Italy*

<sup>b</sup>*Consorzio Interuniversitario Nazionale per la Scienza e Tecnologia dei Materiali, Via Giusti 9, Firenze 50121, Italy*

<sup>c</sup>*Itelyum Regeneration spa, Via Tavernelle 19, Pieve Fissiraga 26854, Lodi, Italy*

\**flavio.manenti@polimi.it*

## Abstract

It is estimated that 2000 Mt/y of Municipal Solid Waste (MSW) are currently produced worldwide. A strategy for disposing of the non-recyclable fraction of MSW through chemical recycling based on thermal treatments is proposed, leading to the conversion of waste into high-value chemicals. A first-principles suite for the dynamic modeling of thermal conversion processes has been developed to reliably predict the product distribution starting from several types of waste. The GasDS suite models pyrolysis, gas-solid interactions, and gas-phase reactions by dividing the domain into finite volumes, with the possibility to choose among different fluid-dynamic configurations for a given unit. A case study based on a steam-moderated oxy-gasification of sewage-sludge-derived hydrochar is presented in this work. Detailed results showing the syngas productivity and the H<sub>2</sub>/CO molar ratio as functions of the Equivalence Ratio (ER) and the Oxygen content of the gasifying medium (O<sub>2</sub> mol%) are provided.

**Keywords:** Circularity, First-principles, GasDS, Municipal Solid Waste, Sewage Sludge

## 1. Introduction

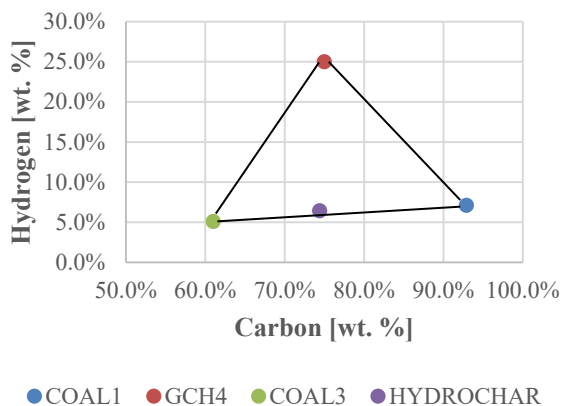
Solid waste management is a crucial aspect of every socio-economic system worldwide. The option of conventional recycling is crucial, since it is the one that allows to accelerate the transition towards a circular economy in which waste is transformed back into a new product. In the last years, the topic of chemical recycling has been growing in importance, since it allows to valorize the non-recyclable fraction of waste which would be otherwise incinerated and/or landfilled. Chemical recycling has been shown to have a great potential in reducing CO<sub>2</sub> emissions from the treatment of residual municipal solid waste compared with incineration, especially when the electric grid has a high percentage of renewables (Voss et al., 2021). Some chemical recycling routes start from biological processes and are then followed by appropriate chemical syntheses, such as the production of biofuels from organic-waste-derived biogas (Fedeli et al., 2022; Negri et al., 2022a). Other pathways may start from the thermal degradation of solid waste and then lead to the

synthesis of biofuels and biosolvents. Some examples of said processes are based on air-based gasification (Prifti et al., 2021a), and oxy-gasification (Negri et al., 2022b; Prifti et al., 2021b). Chemical recycling processes based on thermal degradation of solid matter need a proper description of this reactive step. The most common approach typically involves a thermodynamic analysis, which has several limitations since it cannot analyze in detail complex phenomena such as chemical kinetics, intra-particle and inter-particle mass and heat transfer. Moreover, it cannot analyze the feasibility of the start-up policies proposed for the unit. This work shows an innovative approach to this problem that exploits the GasDS package, a generalized first-principle suite for the dynamic modeling of gas-solid reactive systems, already validated for biomasses (Ranzi et al., 2014), and coal (Corbetta et al., 2015). The structure of the suite is shown in its entirety, along with a case study that shows a detailed analysis of a gasification chamber working with sewage-sludge-derived hydrochar, thus opening the possibility for design and optimization of innovative thermochemical recycling processes.

## 2. Methods

### 2.1. Waste Characterization

The GasDS suite includes a rich and expandable database. Lignin, cellulose, and hemicellulose can be used to describe organic wastes, while several types of coal and tar can be used to describe char-like substances. These species are characterized in detail inside the GasDS database, which contains basic thermochemistry data and properties related to mass and energy transfer. Any general waste can be expressed as a linear combination of said species to perform a simulation. Higher precision can be obtained by entering new tailored species in the database. For example, this is the case for plastic waste, where it is advisable to add entries related to pure polymers like PE, PP, PS, which are the main components of the solid matrix. This paper analyzes the gasification of hydrochar, a waste obtained from the hydrothermal carbonization of sewage sludge. It is possible to express the composition of hydrochar as a linear combination of two types of coal, given its similarity with this substance (Knötig et al., 2021). The sample is almost a perfect linear combination of the species named COAL1 and COAL3, but it is necessary to add a dummy species to close the atomic balance on carbon and hydrogen, here represented by methane and named GCH4. Of course, this provides accurate results only if the final amount of the dummy species is relatively small on a dry-and-ash-free basis like in this case, as shown in Figure 1.



**Figure 1.** Visual representation of the linear combination for the characterization of hydrochar

### 2.2. Detailed Kinetic and Mass/Energy Transfer Modeling

The kinetic model implemented in the GasDS suite describes the pyrolysis of the solid matrix, heterogeneous gas-solid interactions, and secondary gas-phase reactions. Pyrolysis models from the literature are already implemented for biomass (Ranzi et al., 2008) and coal (Vascellari et al., 2015), and it is possible to add new reaction sets with the aim of analyzing different type of feedstocks. For example, one may include models relative to the most common polymers (PE, PP, PS, etc.) to describe the pyrolysis of plastic waste (Kaminsky et al., 2004). Gas-solid heterogeneous interactions make use of kinetic models related to the gasification and combustion of char generated by the previous pyrolysis step (Groeneveld and van Swaaij, 1980). The heterogeneous gas-solid system is modeled in detail for what concerns intra-particle mass and energy transfer taking place in parallel with the reactions. The migration of the generated gaseous species to the homogeneous gas phase is modeled by considering appropriate inter-particle transport phenomena. Finally, the homogeneous gas phase is modeled with a detailed kinetic scheme that describes the secondary gas-phase reactions occurring in said environment (Ranzi et al., 2012). The multi-scale approach used to model the reactive system is based on two crucial assumptions: both the solid particle and the reaction environment are discretized into a finite number of volumes, which are then modeled by employing zero-dimensional dynamic mass and energy balances relative to the  $i^{\text{th}}$  component and the  $j^{\text{th}}$  sector. The equations are shown in Table 1 (Corbetta et al., 2015; Ranzi et al., 2014).

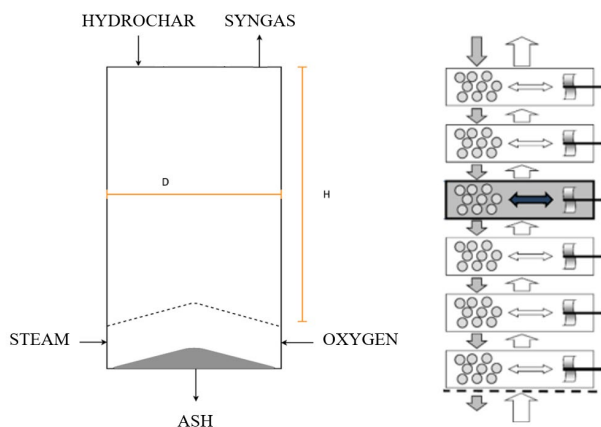
**Table 1.** Balance equations used in the GasDS suite (Corbetta et al., 2015; Ranzi et al., 2014)

EQUATION	DESCRIPTION	(#)
$\frac{dm_{j,i}}{dt} = V_j R_{j,i}$	Mass balance of the solid phase	(1)
$\frac{dm_{j,i}}{dt} = J_{j-1,i} S_{j-1} - J_{j,i} S_j + V_j R_{j,i}$	Mass balance of the gas phase in the solid particle	(2)
$\frac{d \sum_{i=1}^{NCP} m_{j,i} h_{j,i}}{dt} = J C_{j-1} S_{j-1} - J C_j S_j + S_{j-1} \sum_{i=1}^{NCG} J_{j-1,i} h_{j-1,i} - S_j \sum_{i=1}^{NCG} J_{j,i} h_{j,i} + V_j H R_j$	Energy balance of the solid particle	(3)
$J_{j,i} = -D_{j,i}^{eff} MW_i \left. \frac{dc_{j,i}}{dr} \right _{r_j} - \frac{D_{a,j} dP_j}{\mu_j} \left. \frac{dP_j}{dr} \right _{r_j} c_{j,i} MW_i$	Diffusive mass and heat flux in the particle	(4)
$J C_j = -k_j^{eff} \left. \frac{dT_j}{dr} \right _{r_j}$	Heat flux in the particle by conduction	(5)
$\frac{dg_i}{dt} = G_{in,i} - G_{out,i} + J_{N,i} S_N \eta + V_R R_{g,i}$	Gas-phase mass balance for the reactor	(6)
$\frac{d \sum_{i=1}^{NCG} g_i h_{g,i}}{dt} = \sum_{i=1}^{NCG} G_{in,i} h_{g,in,i} - \sum_{i=1}^{NCG} G_{out,i} h_{g,i} + \sum_{i=1}^{NCG} J_{N,i} h_{N,i} S_N \eta + h_{ext} (T_N - T^{bulk}) S_N \eta + V_R H R_g$	Gas-phase energy balance for the reactor	(7)

The discretization of the domain makes the GasDS suite extremely flexible and adaptable to several system configurations. Dividing the particle in sectors is useful to describe thermally thick particles, which are defined as those having a significant temperature gradient between the core and the surface, a situation that may show up whenever the

particle cannot be reduced to a small enough size (Ranzi et al., 2014). The discretization of the reaction chamber allows to describe a lot of reactor morphologies. A single reaction volume can be used to describe fluidized-bed systems, in which the approximation of perfect mixing typically holds. More than one volume may be used to describe fixed/moving bed configurations, both in co-current and counter-current. The gasification chamber modeled in this work is a fixed-bed, countercurrent, cylindrical gasifier that processes 100 kg/h of hydrochar by using a binary H<sub>2</sub>O/O<sub>2</sub> mixture as gasifying agent. The unit is shown in Figure 2 and the following specifications are implemented:

- Reactor height/diameter: 2 m / 1 m
- Temperature of gasification agent at start-up/steady-state: 930°C / 230°C
- Particle average diameter: 0.01 m
- Number of discrete volumes for the analysis of the solid particles/gasifier: 2/3



**Figure 2.** Simplified scheme of the gasification chamber modeled in GasDS

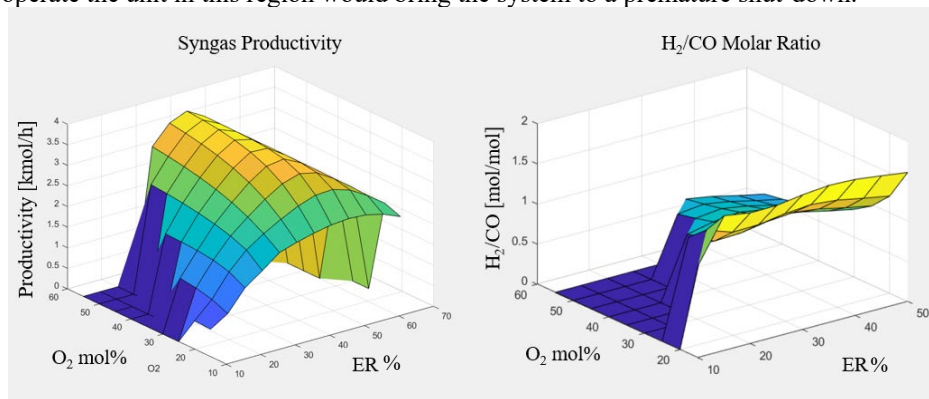
### 3. Results and Discussion

The system described beforehand still has two degrees of freedom that will be the pivots of the detailed analysis. The first one is the *Equivalence Ratio* (ER), which is here defined as per equation 8, with  $ER < 1$  for fuel-lean mixtures, and  $ER > 1$  for fuel-rich ones.

$$ER = \frac{O_2|_{REAL}}{O_2|_{STOICH}} \quad (8)$$

The second degree of freedom of the system is the oxygen content of the binary gasifying mixture, which is here expressed on a molar basis. It is thus possible to evaluate the behavior of the gasification chamber in a two-dimensional domain identified by these two independent variables. Figure 3 shows the first KPIs of the analysis, the syngas productivity on the left (accounting only for CO and H<sub>2</sub>), and the H<sub>2</sub>/CO molar ratio on the right. The productivity shows an optimal value of ER for any given O<sub>2</sub> mol%, with a monotonically increasing behaviour as the Oxygen content increases. Looking at the H<sub>2</sub>/CO ratio, the value remains relatively constant with respect to ER at a given O<sub>2</sub> mol%, and the curve is monotonically decreasing as the Oxygen content increases. Discovering this kind of behaviour is crucial for a proper optimization of the operating conditions of the gasifier. For example, if the user aims at producing an hydrogen-rich molecule such as methanol and/or dimethyl ether starting from this syngas, it will be necessary to find

an operating point which is a trade-off between productivity and  $H_2/CO$  ratio, which have opposite behaviour with respect to the Oxygen content of the gasifying mixture (Negri et al., 2022b). Moreover, the GasDS suite allows to highlight deviations from engineering “best practices”, which is a typical situation when dealing with unconventional feedstocks such as solid waste. Specifically, the gasification of carbonaceous material typically uses an ER of around 0.30 – 0.33 as rule of thumb for optimal operations (Jangsawang et al., 2015); however, this study highlights that the maximum syngas productivity is systematically achieved for higher values of ER, around 0.35 – 0.45. Finally, GasDS shows an interesting behaviour of the gasifier for low values of ER, which would be lost in a thermodynamic analysis. For these extremely fuel-lean mixtures, the system cannot start-up properly with the procedure described in the previous paragraph. The system is too fuel-poor to sustain the reaction environment during the switch from the hot gasifying agent used for starting-up the chamber, to the colder steady-state current. Trying to operate the unit in this region would bring the system to a premature shut-down.



**Figure 3.** Syngas productivity in kmol/h (left) and  $H_2/CO$  molar ratio (right)

#### 4. Conclusions

A generalized first-principles suite for the dynamic modeling of thermal treatments of carbonaceous feedstocks has been introduced. The steps of a GasDS simulation have been shown in detail, starting from the characterization of the raw material to be processed, continuing with the balance equations that make up the core of the suite, and finishing with the set-up of a case study based on a steam-moderated oxy-gasification of hydrochar. The results include a detailed analysis of the syngas productivity and the  $H_2/CO$  ratio, which are fundamental KPIs for the steps following the thermal degradation. These parameters are shown as a function of the Equivalence Ratio (ER) and the oxygen content of the gasifying agent ( $O_2$  mol%). The GasDS suite allows to understand functional dependencies between the analyzed variables, deviations from engineering “best practices” shown by the system due to the unconventional feedstock, and even unfeasible operating conditions which would lead to a premature shut-down of the system. Thus, the suite can be used to model and optimize thermo-chemical waste recycling processes.

#### Acknowledgments

The authors acknowledge HBI Group for providing data on hydrochar composition.

## References

- Corbetta, M., Bassani, A., Manenti, F., Pirola, C., Maggio, E., Pettinau, A., Deiana, P., Pierucci, S., Ranzi, E., 2015. Multi-scale kinetic modeling and experimental investigation of syngas production from coal gasification in updraft gasifiers. *Energy and Fuels* 29, 3972–3984. <https://doi.org/10.1021/acs.energyfuels.5b00648>
- Fedeli, M., Negri, F., Manenti, F., 2022. Biogas to advanced biofuels: Techno-economic analysis of one-step dimethyl ether synthesis. *J. Clean. Prod.* 376. <https://doi.org/10.1016/j.jclepro.2022.134076>
- Groeneveld, M.J., van Swaaij, W.P.M., 1980. 39 Gasification of char particles with CO<sub>2</sub> AND H<sub>2</sub>O. *Chem. Eng. Sci.* 35, 307–313. [https://doi.org/https://doi.org/10.1016/0009-2509\(80\)80101-1](https://doi.org/https://doi.org/10.1016/0009-2509(80)80101-1)
- Jangsawang, W., Laohalidanond, K., Kerdsuwan, S., 2015. Optimum Equivalence Ratio of Biomass Gasification Process Based on Thermodynamic Equilibrium Model. *Energy Procedia* 79, 520–527. <https://doi.org/https://doi.org/10.1016/j.egypro.2015.11.528>
- Kaminsky, W., Predel, M., Sadiki, A., 2004. Feedstock recycling of polymers by pyrolysis in a fluidised bed. *Polym. Degrad. Stab.* 85, 1045–1050. <https://doi.org/https://doi.org/10.1016/j.polymdegradstab.2003.05.002>
- Knötig, P., Etzold, H., Wirth, B., 2021. Model-based evaluation of hydrothermal treatment for the energy efficient dewatering and drying of sewage sludge. *Processes* 9. <https://doi.org/10.3390/pr9081346>
- Negri, F., Fedeli, M., Barbieri, M., Manenti, F., 2022a. A versatile modular plant for converting biogas into advanced biofuels. *Invent. Discl.* 2, 100008. <https://doi.org/https://doi.org/10.1016/j.inv.2022.100008>
- Negri, F., Nova, A., Basso, D., Manenti, F., 2022b. Techno-Economic Analysis of Dimethyl Ether Biofuel Production Plant Based on Sewage Sludge Gasification. *Chem. Eng. Trans.* 94, 1087–1092. <https://doi.org/10.3303/CET2294181>
- Prifti, K., Basso, D., Pavanetto, R., Manenti, F., 2021a. Improving hydrothermal carbonization (HTC) processes by hydrochar gasification. *Chem. Eng. Trans.* 86, 925–930. <https://doi.org/10.3303/CET2186155>
- Prifti, K., Galeazzi, A., Margarita, I., Papale, A., Miele, S., Bargiacchi, E., Barbieri, M., Petea, M., Manenti, F., 2021b. Converting end-of-life plastic waste into methanol: The gasiforming™ process as new, efficient and circular pathway. *Environ. Eng. Manag. J.* 20, 1629–1636. <https://doi.org/10.30638/eemj.2021.151>
- Ranzi, E., Corbetta, M., Manenti, F., Pierucci, S., 2014. Kinetic modeling of the thermal degradation and combustion of biomass. *Chem. Eng. Sci.* 110, 2–12. <https://doi.org/10.1016/j.ces.2013.08.014>
- Ranzi, E., Cuoci, A., Faravelli, T., Frassoldati, A., Migliavacca, G., Pierucci, S., Sommariva, S., 2008. Chemical Kinetics of Biomass Pyrolysis. *Energy & Fuels* 22, 4292–4300. <https://doi.org/10.1021/ef800551t>
- Ranzi, E., Frassoldati, A., Grana, R., Cuoci, A., Faravelli, T., Kelley, A.P., Law, C.K., 2012. Hierarchical and comparative kinetic modeling of laminar flame speeds of hydrocarbon and oxygenated fuels. *Prog. Energy Combust. Sci.* 38, 468–501. <https://doi.org/https://doi.org/10.1016/j.pecs.2012.03.004>
- Vascellari, M., Roberts, D.G., Hla, S.S., Harris, D.J., Hasse, C., 2015. From laboratory-scale experiments to industrial-scale CFD simulations of entrained flow coal gasification. *Fuel* 152, 58–73. <https://doi.org/https://doi.org/10.1016/j.fuel.2015.01.038>
- Voss, R., Lee, R.P., Seidl, L., Keller, F., Fröhling, M., 2021. Global warming potential and economic performance of gasification-based chemical recycling and incineration pathways for residual municipal solid waste treatment in Germany. *Waste Manag.* 134, 206–219. <https://doi.org/https://doi.org/10.1016/j.wasman.2021.07.040>

# A Discrete-Event Simulation Approach to the Design and Planning of Biomass Supply Chains considering Technological Learning

Helena Paulo<sup>a,b</sup>, Miguel Vieira<sup>a,c,d</sup>, Bruno S. Gonçalves<sup>e</sup>, Susana Relvas<sup>a</sup>, Tânia Pinto-Varela<sup>a</sup> and Ana P. Barbosa-Povoa<sup>a</sup>

<sup>a</sup>*CEG-IST, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal*

<sup>b</sup>*ISEL - Instituto Superior de Engenharia de Lisboa, IPL, Lisboa, Portugal*

<sup>c</sup>*EIGeS, Universidade Lusófona, Lisboa, Portugal*

<sup>d</sup>*Univ Coimbra, CEMMPRE, Dept Mech Engn, Coimbra, Portugal*

<sup>e</sup>*ESTG, Politécnico de Leiria, Leiria, Portugal*

## Abstract

The present work intends to contribute to the supply chain design and planning, by addressing the technological learning due to non-mature technologies at biorefineries. The methodological approach combines optimization with discrete-event simulation, assessing its investment risk through the Conditional Value at Risk. Through the optimization model, the optimal strategic variables for the biomass supply chain under deterministic conditions are obtained. In contrast, the simulation model allows to analyze the technological uncertainty supported by the learning curve concept to evaluate the technologies' performance. The results highlight the advantage of exploring this combined approach to consider variability parameters for strategic and tactical solutions and the particular influence of technological learning in supply chain design and planning.

**Keywords:** Learning curves, Discrete-event simulation, Biomass, Supply chain design.

## 1. Introduction

Most decisions related to the design and planning of supply chains take place in an increasingly dynamic environment where variability is present at multiple decisional levels. The impact of uncertainty in decision-making has long been recognized, although modelling and solution approaches still call for improvements. For example, the complexity of strategic decisions is greatly influenced by the uncertainty on availability, quality, product demand, and process operations. Moreover, considering technological process parameters represent an additional challenge in their design and planning to ensure long-term financial viability. The general learning curve concept can be set to model the uncertain technological performance in adopting new technology. The mathematical relation reflects the technical and economic progress of the technology as users gain experience with time (Weiss et al., 2010). It is assumed that the learning factor can be translated into the technological conversion of the main process (e.g. the transesterification conversion factor from biomass to biodiesel), which becomes an uncertain parameter. In this study, we propose the development of an approach capable of modelling the impact of an expected technological development based on the learning curve concept. The approach is applied to the design and planning of integrated biorefineries supply chains. Integrated biorefineries are a promising alternative to bioenergy and biomaterials production, becoming an emerging industry where the technology performance involving biomass conversion at an industrial scale has yet to be



fully known. Therefore, the design and planning of integrated biorefineries involve a significant degree of uncertainty associated with the low maturity of technological learning, with correlated influence in the processing capacity and bioproducts availability or demand. In this scope of multilevel decision-support, discrete-event simulation (DES) can provide a useful approach to model uncertainty and evaluate a solution for the supply chain under different scenarios of the network.

The Portuguese case of a biomass supply chain is used as a real case study to illustrate the application of the proposed methodology. By combining optimization-based simulation, a detailed DES model allows to incorporate of the dynamics and uncertain conditions in the supply chain structure, evaluating a solution with the optimal number of integrated biorefineries, their location, capacities, and technology, while considering the assessment of their financial viability using Conditional Value-at-Risk (CVaR).

## 2. Methodology

The problem scope consists of assessing the biomass-to-biofuels supply chain design and planning through measuring the financial risk, with particular relevance to the technological uncertainty at integrated biorefineries. The learning curve is introduced to model the process conversion efficiency and its impact on the overall supply chain decisions. The methodology proposes an evaluation approach combining an optimization model which generates an initial deterministic solution, then simulated under the conditions of a detailed DES model with variable technological learning. Considering this overall perspective based on strategic long-term investment, the Conditional Value at Risk (CVaR) is used as a risk measure to evaluate the likelihood that a specific loss will exceed a certain value at risk. The hybridization of optimization with simulation excels in the advantages of having an analytical model that formulates a solution under ideal conditions, with a more precise simulation model that mimics how the system reacts under realistic outcomes and hard-to-model constraints (Figueira et al., 2014). The approach allows to explore alternative network solutions with reduced computational hindrance by accounting for the variability provided by uncertain operational parameters.

### 2.1. Supply chain simulation-optimization modeling

The approach combines an optimization model and a simulation model. A Mixed Integer Linear Programming model, based on Paulo et al. (2015), is implemented in GAMS to consider the supply chain superstructure with all possible locations of biomass sources (with the corresponding availability per type), biorefineries (set of available technologies and capacities), demand centers and transportation modes with the distances between all points. The objective is to minimize supply chain total cost, given by Eq.1, identifying the number, location, technology, and capacity of biorefineries to install, with the corresponding amounts produced and demand centers to serve, as well as the amount of biomass type collected at each source and transportation flows.

$$\min Total Cost = \sum (Biomass Cost + Fixed/Variable Operating Cost at Biorefiner + Transportation Costs + Annualized Investment Cost) \quad (1)$$

The DES model aims to represent the real-world environment (as is the Portuguese case) to evaluate the optimized solution under variable conditions. Through the SIMIO platform, the model computes multiple output scenario statistics involving the occurrence of events (e.g. a new bioproduct order) using probability distributions of parameters. The main library of objects available in the SIMIO software is used, which, in general, considers model entities (e.g., products) processed at servers (e.g., biorefinery), and then departs from the system. Suitable to the case level of detail, the focus is set on the impact

of the technological learning curve in the design capacity of the facilities to accommodate an expected increase in bioproduct demand. For simplification, the availability of biomass is considered non-limiting. Additional uncertainty in overall cost and prices forecast are introduced as sources of variability, being considered 1000 replications to guarantee statistical significance. The interaction of the two models aim to evaluate a supply chain configuration with a risk measure of the techno-economic investment viability. To account for the time value of money, the simulation model computes the net present value (NPV) of cashflow operations (Eq.2), for  $n$  time periods and  $CF_k$  as the annualized value of cashflows at time period  $k$  with  $i$  discount rate.

$$NPV = \sum_{k=0}^n \frac{CF_k}{(1+i)^k} \quad (2)$$

### 2.2. Learning Curves and Technological Conversion Efficiency

Empirical assessments can demonstrate that costs decrease with the gain of experience from the use of a particular technology, which usually translates in a performance improvement of the production process. There are different types of learning curves: the one-factor learning curve represents learning by accumulated production; the two-factor formulates learning by doing and by searching effects; and other multi-factor learning curves are explained by other factors. Although a detailed characterization of the learning process seems appealing, multiple factors are often difficult to quantify, therefore the one-factor learning curve is chosen for the development of the present work. Mapping the integrated biorefineries and the information on their specifications allow to find the trend of the process conversion efficiency for a given biomass-to-biofuel technology. And by defining these equations, it is possible to calculate values of conversion that reflects the learning-by-doing performance of each technology over the time horizon, which is expected to stabilize once reached a plateau (Weiss et al., 2010). The learning curve effect on costs can be generalized as a power law function given by (Eq.3), with  $CC_{cum}$  the unit cost of production in monetary units,  $CP$  the cumulated production amounts, and  $CC_{ref}$  and  $CP_{ref}$  the initial conditions at an arbitrary starting point.

$$CC_{cum} = CC_{ref} \left( \frac{CP}{CP_{ref}} \right)^{-\varepsilon} \quad (3)$$

The factor  $(CP/CP_{ref})^{-\varepsilon}$  represents the decrease ratio in the unit cost of production, with  $\varepsilon$  the learning coefficient related to the technology. Considering the conversion efficiency  $\mu$  as the ratio of the amount of bioproduct by the amount of biomass feedstock  $CB$  used to process it, the equation can be rewritten as (Eq.4).

$$CC_{cum} = CC_{ref} \left( \frac{\mu \cdot CB}{\mu_{ref} \cdot CB_{ref}} \right)^{-\varepsilon} \quad (4)$$

From (Eq.4), the relationship between the technological conversion efficiency and the learning-by-doing performance can be obtained. As higher the values of conversion, higher the accumulated production quantities obtained from the same feedstock, and therefore, lower unit production costs.

### 2.3. Financial risk assessment

Attracting investment to the biomass-based industry represents a substantial challenge, whereas making the supply chain efficient under uncertain technological process parameters increases the impact in its design and planning to assure a long-term viability. The assessment of uncertainty appeals for risk management studies and the research

community has been demonstrated their importance. The CVaR is a widely used risk metric to quantify the tail risk, where conclusions can be drawn to either downside losses or upside gains for an industrial investment, as demonstrated by Vieira et al. (2021). Given a discretized probability distribution function of NPV values, CVaR measures the weighted average outcome of far-edge tail events beyond the VaR for a specified confidence level ( $\alpha$ ). The VaR for downside losses is set as the outcome value for the  $(1-\alpha)$ -quantile of the distribution. The CVaR is given in (Eq.5), multiplying the probability of scenario  $s$  occurrence  $\Pi_s$  by the positive variable  $\delta_s$  calculated in (Eq.6), which measures the positive difference between the VaR and the NPV of each shortfall scenarios.

$$CVaR = VaR - \frac{1}{1-\alpha} \sum_s \Pi_s \delta_s \quad (5)$$

$$\delta_s = VaR - NPV_s, \delta_s \geq 0 \quad (6)$$

### 3. Case Study and Results

The described methodology has been applied to an illustrative case study of biomass supply chain of bioethanol and biodiesel production in Portugal. Supply chain characteristics and data considered in this study is summarized on Fig.1, considering two technologies – fermentation for bioethanol and transesterification for biodiesel.

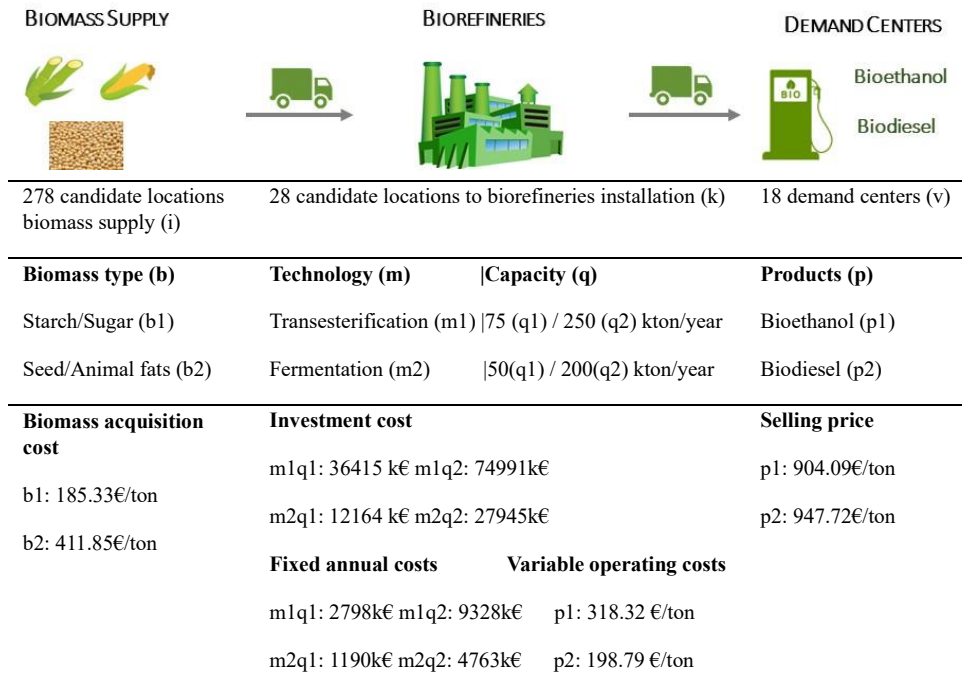


Figure 1. Case study data

The initial solution provided by the optimization model defines the supply chain superstructure, identifying 4 biorefineries to install. Lisboa, Pombal and Vila Real are the selected locations to install biorefineries using fermentation technology, with a processing capacity of 250 000 tons and 75 000 tons of bioethanol, respectively.

Santarém is selected to install a transesterification biorefinery of 50 000 tons of production capacity for biodiesel. The optimal solution also defines the biomass flows from each collecting point to each biorefinery, as well as the product flows to each demand center. This information is used to compile the simulation model details. With regard to the learning curve integration, it is considered that the two technologies selected have dissimilar characteristics in its learning-by-doing progress. From the work by Seabra (2021) based on real data, the equations used to calculate the decreasing costs for each level  $n$  of accumulated production for the fermentation ( $m1$ ) and transesterification ( $m2$ ) processes were, respectively (Eq.7) and (Eq.8).

$$CC_{n\ m1} = 534.43 (CP_{n\ m1}/10000)^{-0.3219} \tag{7}$$

$$CC_{n\ m2} = 282.10 (CP_{n\ m2}/10000)^{-0.1520} \tag{8}$$

The derived conversion efficiency trends are translated by (Eq.9) and (Eq.10), whereas the values pondered for the first time period were set as  $\mu_{fermentation}=25.58\%$  and  $\mu_{transesterification}=27.51\%$ .

$$\mu_{fermentation} = 0.055(t - 1) + 0.2558 \quad \forall t > 1 \tag{9}$$

$$\mu_{transesterification} = 0.035(t - 1) + 0.2751 \quad \forall t > 1 \tag{10}$$

It should be noted that the biomass conversion efficiency progress is expected to stabilize, which the study by Sharew et al. (2022) suggested the interval [46%-56%] for these technologies, so it is assumed a plateau value of 51%. This is added to the characteristic of each corresponding biorefinery in the model, which computes the statistical simulation for a time horizon of  $n=10$  years, and the NPV considering the defined parameters distributions, such as cost and price variability, with a discount factor  $i$  of 10%. Assuming the scenario where the transition to biofuels is fostered, the increase in the demand follows a uniform distribution between 2-4%. The results displayed in Fig. 2 demonstrates the output simulation replications, represented by a probability distribution, which enable the calculation of the CVaR by discretizing the scenarios under the VaR for a confidence level of 95% (13,6.10<sup>6</sup>€). The value of CVaR means that the proposed supply chain configuration and scenario conditions guarantees at least a NPV of 12,1.10<sup>6</sup>€ with a 95% of probability.

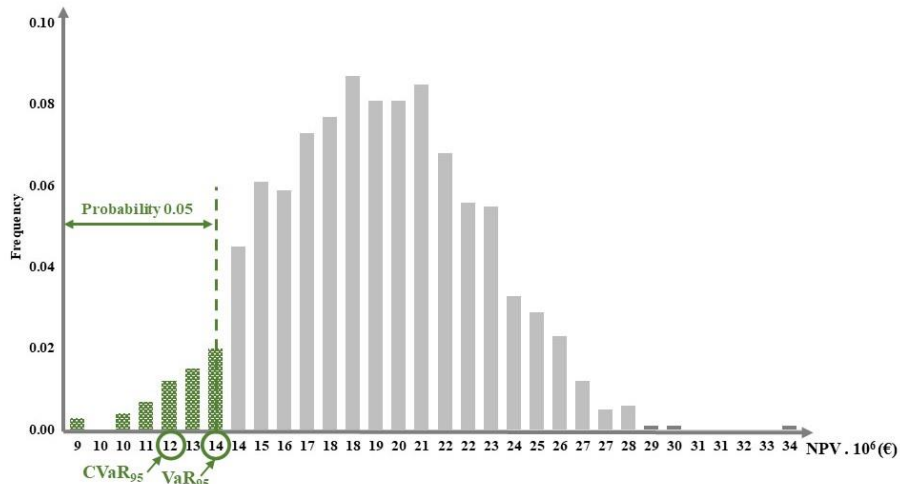


Figure 2 – Risk evaluation results: NPV VaR and CVaR representation

With Fig. 3 it is possible to analyze the learning curve effect in the increase of overall capacity of the biorefineries, as well as its utilized capacity while accommodating the demand increase. For example, Lisboa and Vila Real show a higher utilization of its installed capacity, while the remaining recommend to evaluate an alternative scenario of a reduced capacity or network configuration. Likewise, sensitivity analysis to learning curve progress can be simulated, for example the CVaR obtained, for a 95% confidence level, with an increase of 0,5% in the learning rate (faster learning) rises to 20,1.10<sup>6</sup>€, while a decrease of 0,5% (slower learning) generates a probable downside loss of 3,9.10<sup>6</sup>€.

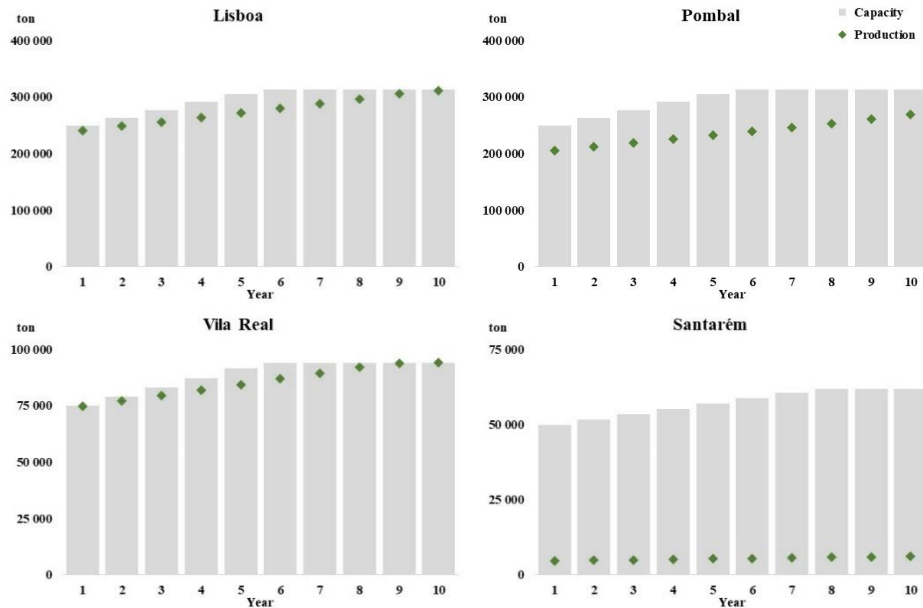


Figure 3 – Biorefineries installed capacity (bars) and production (dots) by year

#### 4. Conclusions

This work explores the effect of learning-by-doing in non-mature technologies over the biomass supply chain performance. The methodology measures the investor risk by combining simulation and optimization techniques, allowing to explore an optimal-based solution through the simulation of variable dynamic conditions, parameters distributions, and technological learning conversion. By using CVaR to measure NPV as an investment risk, the decision maker is able to have more robust information on its viability, as well as to evaluate alternative scenarios. The results demonstrate the impact of different learning rates and uphold the simulation of alternative configurations to the supply chain design, pursuing the assessment of the multiple uncertainty factors affecting these decisions.

#### Acknowledgments

UE/FEDER through COMPETE and FCT projects UIDB/00097/2020 and UIDB/00285/2020.

#### References

- M. Weiss, et al., 2010, A review of experience curve analyses for energy demand technologies. *Technological Forecasting and Social Change*, 77(3), 411–428.
- G. Figueira. and B. Almada-Lobo, 2014, Hybrid simulation–optimization methods: A taxonomy and discussion. *Simulation Modelling Practice and Theory*, 46:118-134.
- H. Paulo et al., 2015, Supply chain optimization of residual forestry biomass for bioenergy production: The case study of Portugal. *Biomass and Bioenergy*, 83, 245-256

- M. Seabra, 2021, Modelling uncertainty in the technologies' conversion efficiency for the design and planning of biomass supply chains, Master in Industrial Engineering and Management, Master Thesis, Instituto Superior Técnico, Universidade de Lisboa.
- M. Vieira et al., 2021, Assessment of financial risk in the design and scheduling of multipurpose plants under demand uncertainty. *International Journal Production Research*, 59(20), 6125–6145
- S. Sharew et al., 2022, Alternative Energy Potential and Conversion Efficiency of Biomass into Target Biofuels: A Case Study in Ethiopian Sugar Industry Wonji-Shoa. *Biomass*, 2(4), 279–298.



# Data-driven modeling to predict the rate of Boil-off Gas (BOG) generation in an industrial LNG storage tank

Suraj Prakash Singh,<sup>a</sup> Rajagopalan Srinivasan,<sup>a,b\*</sup> I. A. Karimi<sup>c</sup>

<sup>a</sup> *Department of Chemical Engineering, Indian Institute of Technology Madras, Chennai 600036, India*

<sup>b</sup> *American Express Lab for Data Analytics, Risk and Technology, Indian Institute of Technology Madras, Chennai 600036, India*

<sup>c</sup> *Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore*

\*[raj@iitm.ac.in](mailto:raj@iitm.ac.in)

## Abstract

Natural gas is stored as Liquefied Natural Gas (LNG) under cryogenic conditions in a storage tank. The tank is highly insulated; nevertheless, there is heat ingress from the surrounding, causing the generation of boil-off-gas (BOG). BOG generation over-pressurizes the tank and can lead to tank failure, so proper BOG management is critical to plant safety. This paper seeks to develop a data-driven model for a real industrial LNG terminal. Modeling was performed on historical real-time LNG data using different machine learning algorithms - Linear regression, Random Forest, and XGBoost. The performance of algorithms is analyzed based on  $R^2$ , Mean-Absolute-Error (MAE %), and Root-mean-square error (RMSE %) values. The critical input features are calculated based on the Shapley additive explanation method (SHAP value); less important features were removed to decrease the model complexity. Our studies show that Random Forest outperforms the other two algorithms in terms of accuracy. The developed model can help plant operators make decisions quickly with better confidence.

**Keywords:** Data-driven modeling, Liquefied Natural Gas, Boil-Off-Gas, BOG management, Machine learning.

## 1. Introduction

Natural gas (NG) is considered one of the cleanest fossil fuels, with a variety of end-uses, including fertilizer production, power generation, along with industrial and residential heating. With the global annual consumption of NG being 4.04 Tm<sup>3</sup> (Statistical Review of World Energy, 2021), it is likely to become the second most widely used fossil fuel by 2035. It is globally traded in its liquefied form (LNG) via specialized carriers that transport LNG from suppliers to receiving terminals, where it is unloaded into storage tanks. Storing LNG under cryogenic conditions enables safe, convenient, and cost-effective distribution. However, continuous heat ingress from the surrounding environment causes the vaporization of LNG, known as boil-off gas (BOG). When LNG of two different densities is stored in the same storage tank, rather than forming a homogeneously mixed liquid, it can result in stratification. This is because when the different layers are mixed, a significant amount of BOG is generated due to interface disruptions, and the tank pressure builds up. This heterogeneous mixing phenomenon is termed rollover. In this work, the broad objective is to prevent or alleviate the adverse consequences of rollover.



Rollover is a common problem across storage units in LNG supplier facilities, FSRUs, LNG carriers, and regasification terminals, with a higher frequency of occurrence at the latter (Energy report, 2018). Rollover can have catastrophic repercussions; for example, the incidents at La Spezia (Italy, 1971) and Partington (the UK, 1993) are prominent examples of how BOG generation due to rollover can have devastating impacts on the regasification terminal, causing significant economic losses. According to the GIIGNL 2019, more than 24 rollover incidents have been reported over the last 60 years. Such events are continuously increasing due to the ever-increasing use and trade of LNG. The amount of BOG generation during a rollover is enormous; nevertheless, the mechanism is, until now, not fully understood. Therefore, it is essential to develop a model to predict the rate of BOG generation in the LNG regasification terminal.

The basics of rollover and the rate of BOG generation have been the subject of several studies. These have a theoretical, experimental, or numerical foundation. Theoretical studies concentrated on how many intermediate stages were assumed to better comprehend rollover events, such as two-stage (Heestand, 1983), three-stage (Koyama, 2007), and four-stage (Li, 2015). Experimental studies focused on the effects of LNG composition, initial concentration difference (Munakata, 1994), visual observations of boundary layers, and formation and decomposition of evolved LNG layers (Bates, 1997). Numerical studies involved computations based on different geometries, modes, and angles of LNG filling (Hubert, 2019; Saleem, 2020) that can include initial conditions and boundary conditions (Zakaria, 2017), interface instability (Wei, 2018), among other factors. However, all of them are computationally expensive and difficult to model. Making industrial decisions based on these models during critical situations is difficult. We aim to address this issue in the current work. In this paper, we develop a data-driven model based on collected historical data to predict the rate of BOG generation in an LNG tank in a regasification terminal. The modelling was performed by using a simple linear model (Linear regression (LR)) as well as more complex nonlinear models (Random Forest (RF) and XGBoost). Further, the performance of all three models is analyzed to select the best model. Finally, the model is used to perform various case studies based on industrial problem scenarios.

## **2. Process description**

In the LNG regasification terminal, LNG is imported through ships and unloaded into the highly insulated storage tank. This stored LNG is continuously supplied to the customers, and whenever the tank level reaches the minimum value, a new ship arrives for fresh LNG unloading. The LNG in the tank is stored nearly at its saturation temperature under atmospheric pressure. The continuous heat ingress from the surroundings to the tanks and pipeline increases the temperature of LNG. Due to this, LNG evaporates and forms BOG, which leads to pressure build-up in the tank. For the safety of the tank and regasification terminal, the generated BOG is continuously removed from the tank through a compressor. The generated BOG contains nearly 99% methane, which has the highest calorific value, and its removal reduces the overall calorific value of LNG. The BOG can be managed in various ways within the plant, such as (1) directly supplied to the customer, (2) recondensed back to the tank, and (3) flared in of the plant. These methods are either not economically suitable or permissible due to environmental concerns. So, it is crucial to understand the behavior and BOG generation rate under various conditions and operate the plant so as to reduce BOG formation. A data-driven model is developed for this purpose.

### *2.1. Data collection*

Industries continuously monitor the pressure inside the tank and accordingly estimate the BOG generation. For process control and operations, various other variables like temperature and density of LNG, tank vapor pressure, LNG level, compressors operation, and LNG outflow are also continuously monitored. These data are continuously monitored and saved on an hourly basis between two consecutive LNG ship unloading.

This research aims to predict the rate of BOG generation under various conditions and under the influence of various parameters. Thus, an accurate relationship has to be established between the input and output variables. The model input variables are LNG level ( $x_1$ ), LNG density at the top ( $x_2$ ), LNG density at the bottom ( $x_3$ ), LNG temperature at the top ( $x_4$ ), LNG temperature at the bottom ( $x_5$ ), vapor pressure ( $x_6$ ), LNG outlet flowrate ( $x_7$ ) and operation mode of two compressors ( $x_8$  and  $x_9$ ). The model output variable is the rate of generation of BOG ( $y$ ) in the LNG storage tank.

In our work, we have used historical real-time data of all the input and output variables collected over nearly one month (934 datasets) for model development. These data are collected every second but saved at every hour interval (sampling time) for future study purposes. The data consists of various missing values and outliers, so proper data pre-processing is necessary. In our case, as the number of missing values (1.60%) and outliers (0.96 %) are small, we removed those samples from the datasets. The whole dataset of every month is segregated into different segments based on the mode of operation in the regasification terminal. The first mode considered is the unloading operation, where LNG is unloaded from an LNG ship to the storage tank. The second mode is normal operations when there is no input to the tank, but LNG is sent to the regasification process, and the gas is sent-out to customers. In this paper, we consider the latter, i.e., the normal operations mode. The data from normal operations is split into training and testing data sets in the ratio of 4:1. The training datasets are used to develop various models based on different machine learning algorithms.

## **3. Results and discussion**

The approach considered to establish a relationship between input and output features includes (1) data collection, (2) data pre-processing, (3) designing a model, (4) training with suitable algorithms, and (5) model validation. Initially, the nine input features mentioned above were used to design the model without considering the effect of time (i.e., a steady-state model was developed). We start from the simple LR model and then compare its performance against two nonlinear models – RF and XGBoost. The model's performance was further enhanced by removing outliers. Finally, a complete comparison of all three models with and without outliers is considered in the analysis, and results are reported accordingly in the results and discussion section.

### *3.1. Linear Regression*

In this section, the model is trained using multivariate linear regression, and the results are analyzed. The model was initially trained by considering the outliers and then checked the performance based on both training and testing datasets. The results clearly show that the presence of outliers in the training and testing datasets decreases the model's accuracy. The accuracy of the model was further improved by removing outliers from both training (0.64 %) and testing (0.32 %) datasets. The performance of both results is compared by calculating the R-Square value ( $R^2$ ), Mean-Absolute-Error (MAE %), and Root-mean-square error (RMSE %). After removing the outliers, the  $R^2$  value improved from 0.84 to

0.95 in training and from 0.82 to 0.96 in the testing datasets. Similarly, MAE (%) enhanced from 3.66% to 2.21% in training and 4.23% to 2.29% in testing datasets, and RMSE (%) improved from 5.5% to 3% in training and 6.3% to 2.9% in testing datasets. The mode of compressor operation, tank pressure, bottom layer temperature and LNG level are considered as important variables in the linear regression modelling.

In order to improve accuracy, two nonlinear machine learning algorithms Random Forest (RF) and XGBoost were also studied.

### 3.2. Random Forest

In this section, the model is trained and tested based on Random Forest (nonlinear algorithm). The value of  $R^2$  remains 0.99 in both cases in training datasets, while it increases from 0.88 to 0.99 in testing datasets. Similarly, the value of MAE (%) and RMSE (%) remain the same at 0.49% and 1.23% with the training datasets, while it improves from 1.29% to 0.65% and 5.03% to 1.02% in testing datasets, respectively. The model indicates mode of compressor operation, LNG level and LNG density plays crucial role in the generation of BOG in the storage tank.

### 3.3. XGBoost

The performance of XGBoost algorithm is also good but lesser than the RF algorithm. The value of  $R^2$  remains 0.98, which is unchanged with or without considering the outliers in training datasets, while it improves from 0.88 to 0.99 in testing datasets. Similarly, the MAE (%) and RMSE (%) value remains the same at 1.15% and 1.83% in training datasets, while it improves from 1.68% to 1% and 5.82% to 1.32% in the testing dataset, respectively. This algorithm suggested mode of compressor operation, LNG level, LNG outflow and LNG temperature are important input variables.

**Table 1.** Comparison of the performance of LR, RF and XGBoost, based on  $R^2$ , RMSE (%) and MAE (%)

Models		$R^2$		RMSE (%)		MAE (%)	
		Training datasets	Testing datasets	Training datasets	Testing datasets	Training datasets	Testing datasets
Linear regression	With outliers	0.84	0.82	5.50	6.30	3.66	4.23
	Without outliers	0.95	0.96	3.00	2.90	2.21	2.29
Random Forest	With outliers	0.99	0.88	1.23	5.03	0.49	1.29
	Without outliers	0.99	0.99	1.23	1.02	0.49	0.65
XGBoost	With outliers	0.98	0.88	1.83	5.82	1.15	1.68
	Without outliers	0.98	0.99	1.83	1.32	1.15	1.00

The comparison of all three models revealed that the mode of compressor operation has a significant impact on the rate of BOG generation. The compressor operation at lower capacity increases vapor pressure inside the tank, resulting in a lower rate of BOG generation. Another important variable suggested by the models are LNG level in the

tank. The LNG level provides direct information about the volume of LNG in the storage tank. Thus, higher volume results more generation of BOG.

### 3.4. Performance of models

The performance of the three models (LR, RF and XGBoost) are compared on training and testing data sets with and without considering the outliers. The performance of the Linear regression model is the least as compared to the other two nonlinear models. Among the two nonlinear models, the RF has better performance than XGBoost. The performance of all three models can easily be visualized in Fig. 1 and table 1 based on  $R^2$ , MAE and RMSE values.

### 3.5. Sensitivity analysis

The sensitivity analysis of the model helps to calculate the contribution of each individual feature in the output features. Initially, all the measurable input features (9 features) are used for modelling, but it may be possible that the model is independent of some of the input features. These extra independent features increase the complexity of the model and are responsible for both overfit and increased simulation time for large datasets. To avoid these possibilities, a sensitivity analysis was performed and analyzed the features really crucial for the modelling by using Shapley additive explanation method (SHAP value). By considering all nine input features, the SHAP value is calculated and reported in Fig. 2. (a). The lowest SHAP value input features have to be removed, and checked the performance of the model. Similarly, one by one, input features have to be removed and reviewed the performance of the model until a significant decline is observed.

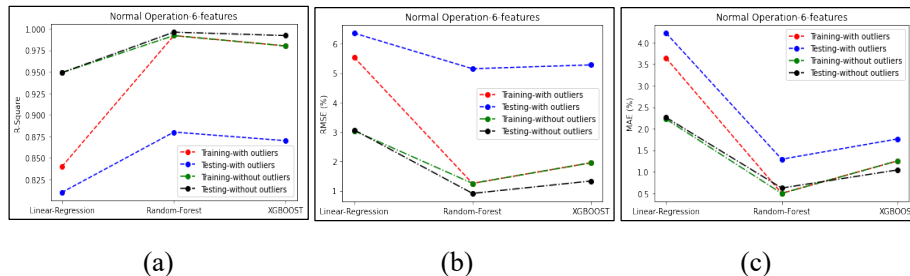


Fig. 1. Comparison of the performance of LR, RF and XGBoost with 6 input features, based on (a)  $R^2$ , (b) RMSE (%) and (c) MAE (%) with and without considering the outliers in both training and testing dataset

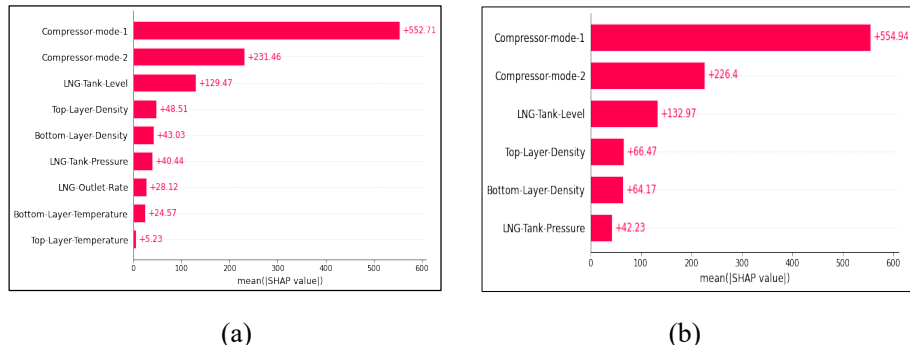


Fig. 2. Comparison of the mean (SHAP) value by considering (a) 9 input features, and (b) 6 input features in the model

In this modelling, we would not have observed a significant change in the model's performance after removing the three lowest SHAP value (Fig. 2 (b)) input features (top-temperature, bottom-temperature and LNG flowrate). Thus, six input features have to be considered in the model and the performance reported in Fig. 1 (a), (b), and (c).

#### 4. Conclusions

A data-driven model is used to predict the rate of boil-off-gas (BOR) generation during normal operation in an industrial LNG storage tank based on different linear and nonlinear machine learning (ML) algorithms. The three different ML algorithms (Linear regression, Random Forest and XGBOOST) are used to model the data by considering nine input features and checking the accuracy of the model. The performance of all three algorithms is analyzed based on  $R^2$ , MAE (%) and RMSE (%) values. Out of all three algorithms, the performance of Random Forest is considered to be the best, with  $R^2$  value of 0.99 in training datasets with and without outliers. The performance of  $R^2$  in testing datasets improve from 0.88 with outliers to 0.99 without outliers.

Further, sensitivity analysis is performed using the SHAP value method to find the important input features for the model. The input features with lesser SHAP values are removed one by one from the model and again checked the performance of the model. After removing top layer temperature, bottom layer temperature and LNG outlet flow rate, the performance of the model remains almost the same in both the training and testing datasets. This analysis indicates that the remaining six input features – compressor operation of 1 and 2, level of LNG, bottom LNG density, top LNG density and tank vapor pressure – play the most important role in BOG generation in the industrial LNG storage tank. Specifically, the mode of compressor operation and the LNG level are the most significant features. While the resulting model has good overall performance, it has several limitations. The dynamic behavior of the system, i.e., the temporal dimension is not considered in this model. The model also does not cover other operating modes such as LNG unloading, compressor maintenance, and special events like rollover. These will be addressed in our future work. Also, other input features like the effect of ambient temperature, and dimensions of the LNG tank would be incorporated into the model.

#### Acknowledgments

The authors gratefully acknowledge financial support from the Indian Institute of Technology Madras and the Ministry of Human Resource Development, Government of India, under the Scheme for Promotion of Academic and Research Collaboration (SPARC).

#### References

- GIIGNL 2019 Annual Report - <https://giignl.org/document/giignl-2019-annual-report>  
<https://www.energy.gov/sites/prod/files/2018/11/f57/draft-eis-0531-port-delfin-ling-app-r-2016-07.pdf>
- Heestand, J.; Shipman, C.W.; Meader, J.W., 1983, A predictive model for rollover in stratified LNG tanks. *AIChE J.* 29, 199–207.
- Bates, S.; Morrison, D.S., 1997, Modeling the behaviour of stratified liquid natural gas in storage tanks: A study of the rollover phenomenon. *Int. J. Heat Mass Transf.*, 40, 1875–1884.
- Saleem, A.; Farooq, S.; Karimi, I.A.; Banerjee, R., 2020, CFD Analysis of Stratification and Rollover Phenomena in an Industrial-Scale LNG Storage Tank. *Ind. Eng. Chem. Res.*, 59, 14126–14144.

Zakaria, Z.; Kamarulzaman, K.; Samsuri, A., 2017, Rollover phenomena in liquefied natural gas storage: Analysis on heat and pressure distribution through CFD simulation. *Int. J. Innov. Eng. Technol.*, 8, 392–400



# Synthesis of process configurations and solvent blend for CO<sub>2</sub> capture from CCGT power plants

Nahyan Arshad,<sup>a,b</sup> Ahmed Al Hajaj,<sup>a,b</sup>

<sup>a</sup>*Department of Chemical Engineering, Khalifa University of Science and Technology, Abu Dhabi, UAE*

<sup>b</sup>*Research and Innovation Center of CO<sub>2</sub> and Hydrogen, Abu Dhabi, UAE*

## Abstract

Post-combustion CO<sub>2</sub> capture (PCCC) from flue gas of power plants is a promising route to mitigate the climate change in the near term. However, high energy consumption remains the major challenge of MEA based PCCC process. This study aims to investigate the potential of a hybrid approach by replacing MEA with an energy efficient MDEA/PZ blend and applying process configurations to have the maximum reduction in the regeneration energy. Rigorous rate-based model of a MDEA/PZ based PCCC process integrated with a 750 MW CCGT power plant was developed in Aspen Plus. In addition, the study examined the synergistic effects of combining three process configurations in a single flowsheet on the regeneration energy. The energy analysis results show that AIC+RSS+LVC is the energy-optimal combination as it provides the lowest regeneration energy (2.12 GJ/t<sub>CO2</sub>). However, considering the significant electricity consumption by the lean vapor compressor, the equivalent energy results highlight that AIC+RSS+IHS is the optimum combination as it results in lowest equivalent energy (0.985 GJ/t<sub>CO2</sub>) corresponding to an overall energy saving of 8.96%.

**Keywords:** Post-combustion CO<sub>2</sub> capture; MDEA/PZ; Process configurations; Energy savings

## 1. Introduction

The significant use of fossil fuels for energy generation has resulted in high concentration of CO<sub>2</sub> in the atmosphere which primarily contributes to the global climate change. Combined Cycle Gas Turbine (CCGT) power plants are rapidly adopted worldwide for energy generation compared to coal power plants due to higher energy efficiency, lower CO<sub>2</sub> emissions rate and flexible power output; hence, they can be used in conjunction with variable renewable energy sources (Cheng et al.,2022). One of the key pathways to minimize the anthropogenic CO<sub>2</sub> emissions in the short term is Post Combustion CO<sub>2</sub> Capture (PCCC) using amines due to its technological maturity and large-scale availability. Monoethanolamine (MEA) is the benchmark solvent in PCCC applications due to its low cost, low volatility, low viscosity, and high absorption kinetics at low CO<sub>2</sub> concentrations as in the case of flue gases from power plants. In the recent years, research experts have proposed various solvents with optimal properties for CO<sub>2</sub> capture such as ammonia, ionic liquids, deep eutectic solvents and water lean solvents. However, the commercial applications of these solvents are still immature and require further testing to evaluate their techno-economic feasibility at industrial scale (Li et al.,2016). One of the major limitations of MEA based PCCC process is the significant thermal energy requirement for CO<sub>2</sub> desorption which accounts for majority of the energy consumption



and operational costs of the process (Li et al.,2016). Other limitations of MEA include high corrosion and degradation which increases the solvent makeup costs.

The two practical ways of reducing this high energy demand are the development of energy efficient solvents and synthesis of process configurations. Methyldiethanolamine (MDEA) is a potential solvent to replace MEA as it has attractive properties for CO<sub>2</sub> capture such as higher CO<sub>2</sub> absorption capacity, lower regeneration energy and degradation rate compared to MEA (Ibrahim et al.,2014). The major drawback of MDEA is the low CO<sub>2</sub> absorption rate which increases the capital costs associated with height of the absorber column. This limitation of MDEA can be improved by blending it with an absorption kinetics promoter. Among the amine family, Piperazine (PZ) has the highest CO<sub>2</sub> reaction kinetics as it rapidly forms carbamates with CO<sub>2</sub>; hence it can be considered as an efficient absorption kinetics promoter in tertiary amines such as MDEA (Ibrahim et al.,2014). The experimental study by Bishnoi and Rochelle showed the absorption rate of MDEA/PZ blend is comparatively higher than MDEA/MEA and MDEA/DEA blends (2002). Khan et al. (2020) evaluated the CO<sub>2</sub> absorption rate and regeneration energy of MDEA/PZ at varying concentrations of PZ. The Aspen Plus results of the study revealed that the optimum concentration ratio of MDEA/PZ is 30/20 wt%. The other approach to reduce the high energy demand is process configurations which are modifications in PCC plants to reduce the parasitic energy consumption of the capture process. The configurations evaluated in this paper include Absorber Inter Cooling (AIC), Rich Solvent Preheating (RSP), Rich Solvent Split (RSS), Rich Solvent Flash (RSF), Lean Vapor Compression (LVC), Rich Vapor Compression (RVC) and Inter Heated Stripper (IHS).

Majority of the studies on process configurations were based on MEA and very few studies have examined the potential of MDEA/PZ and configurations on the energy requirement of PCCC process. Therefore, this paper focuses on a combined approach of replacing MEA with MDEA/PZ blend and applying process configurations to maximize the reduction in the regeneration energy. In addition, previous work on process configurations have considered few individual configurations and single combination of three configurations due to convergence challenges. Hence, there is a need to examine several individual configurations and multiple combinations of three configurations in a single flowsheet to demonstrate the cumulative effect of combining three configurations on the regeneration energy of MDEA/PZ process. This work aims to evaluate the performance of seven process configurations and their combinations on the regeneration energy of MDEA/PZ based PCCC process integrated with a 750 MW CCGT power plant. Rate-based model was developed followed by a detailed energy analysis to identify the optimum combination interms of overall energy savings.

## **2. Methodology**

### *2.1 Baseline CO<sub>2</sub> Capture Process*

A conventional CO<sub>2</sub> capture process is shown below in Fig 1. The hot flue gas (106°C) composed of 74.32% N<sub>2</sub>, 12.09% O<sub>2</sub>, 8.67% H<sub>2</sub>O, 4.04% CO<sub>2</sub> and 0.88% Ar from a 750 MW CCGT power plant was pre-cooled to 40°C in a direct contact cooling (DCC) column. The cooled flue gas then enters the bottom of the absorber column with lean solvent fed from top of the absorber to allow efficient chemical absorption of CO<sub>2</sub> from flue gas into lean solvent. The capture rate is set at 90% after which the lean flue gas leaves from top of absorber to a waterwash unit to condense any escaped solvent vapor. The rich solvent from absorber is preheated in a Lean Rich Heat Exchanger (LRHX) followed by the desorption of CO<sub>2</sub> at elevated temperatures in a stripper column. The hot lean solvent

from the reboiler is pre-cooled in the LRHX after which it is recycled back to the top of absorber. The desorbed CO<sub>2</sub> is cooled in an overhead condenser and sent to a compression train which compresses the CO<sub>2</sub> product to 152 bars for storage purposes.

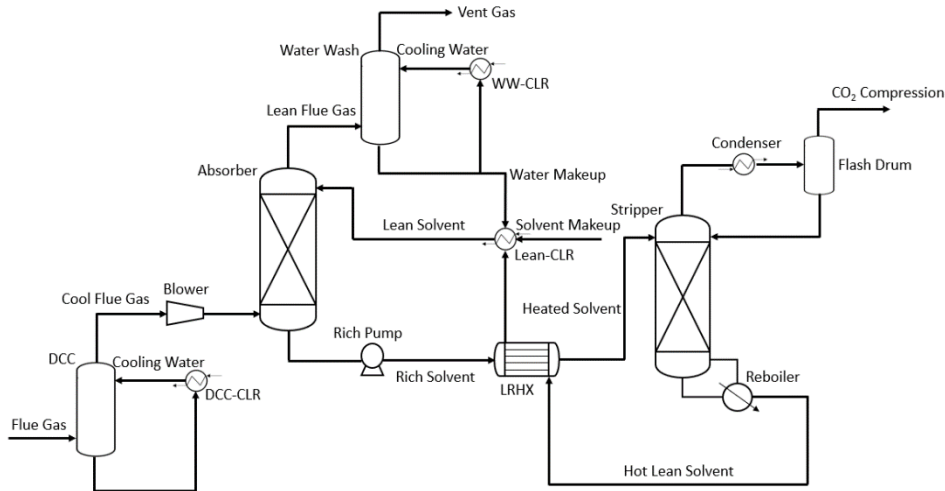


Figure 1: Conventional CO<sub>2</sub> Capture Process

### 2.2 Thermodynamic and Rate Based Model

The process model for the MDEA/PZ based CO<sub>2</sub> capture was developed in Aspen Plus V10 software. The liquid properties of the ionic species were computed by the Electrolyte Non Random Two Liquid (ENRTL) method while the vapor properties of the species were computed using Redlich Kwong (RK) Equation of State. The equilibrium and kinetic reactions of the MDEA-PZ-CO<sub>2</sub>-H<sub>2</sub>O system were the same as in the inbuilt MDEA/PZ CO<sub>2</sub> capture model in Aspen Plus.

The absorber and stripper columns were modeled using the rate-based modeling as it considers mass transfer resistance between phases and reaction kinetics assuming equilibrium only at the fluid interface. Hence, rate based modeling provide more reliable results for column design compared to equilibrium based modeling. The initial data for rate-based modeling such as packing specifications, film resistance, mass and heat transfer correlation methods, etc was obtained from the work of Li et al. (2016).

### 2.3 Equivalent Energy

Equivalent energy is a standard KPI which unifies the thermal and electrical energy consumptions to make a reasonable comparison of overall energy consumption of conventional process with process configurations such as LVC and RVC consuming additional electrical energy. The equation to determine equivalent energy is shown below where the first part represents the equivalent electrical penalty of thermal energy while the second part represents the electrical energy consumed by compressors and pumps in the process.

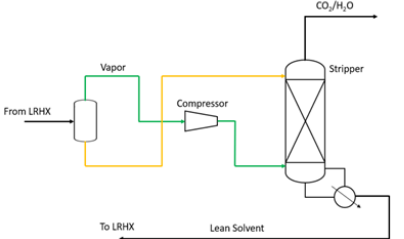
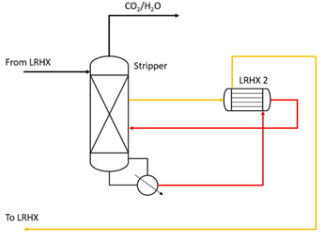
$$E_{eq} = 0.75Q_{reb} \left( \frac{T_{reb} + 10K - T_{sink}}{T_{reb} + 10K} \right) + E_{elec}$$

### 2.4 Process Configurations

The descriptions and schematics of process configurations evaluated in this study are summarized in Table 1:

Table 1: Description and Schematics of Process Configurations

Configuration	Description	Schematic
Absorber Inter Cooling	Extraction and cooling of a portion of semi-rich solvent to enhance the driving force for CO <sub>2</sub> absorption	
Rich Solvent Preheating	Additional heating of the preheated rich solvent after LRHX to lower the amount of energy required to heat the solvent	
Rich Solvent Split	Splitting of rich solvent into cold stream which enters from top of stripper and heated stream which enters from middle to allow pre-stripping of CO <sub>2</sub> from hot vapors	
Rich Solvent Flash	Flashing of the heated rich solvent after which vapor is sent to overhead condenser to reduce the heat required for CO <sub>2</sub> desorption	
Lean Vapor Compression	Flashing of the lean solvent followed by the compression and reinjection of vapor into stripper to minimize steam requirement in the reboiler	

<p>Rich Vapor Compression</p>	<p>Flashing of the rich solvent followed by the compression and reinjection of vapor into stripper to lower reboiler's steam demand</p>	
<p>Inter Heated Stripper</p>	<p>Withdrawing and heating of semi-lean solvent using heat from hot lean solvent from reboiler to reduce sensible heat required in the stripper</p>	

### 3. Results and Discussion

The summary of regeneration energy, equivalent energy and percentage energy savings for the individual and combined process configurations are presented below in Table 2:

Table 2: Summary of Energy Analysis Results

Configuration	Regeneration Energy (GJ/tCO <sub>2</sub> )	Equivalent Energy (GJ/tCO <sub>2</sub> )	% Savings in Energy
Base Case	2.87	1.082	-
AIC	2.58	1.046	3.33
RSP	2.97	1.102	-1.85
RSS	2.76	1.070	1.12
RSF	2.87	1.083	-0.10
LVC	2.38	1.053	2.68
RVC	3.02	1.135	-4.90
IHS	2.72	1.058	2.22
AIC+RSS+LVC	2.12	0.996	7.95
AIC+RSS+IHS	2.30	0.985	8.96
AIC+IHS+LVC	2.18	1.002	7.40
RSS+IHS+LVC	2.22	1.011	6.56

#### 3.1 Discussion of energy analysis results

The results presented in the above table demonstrate that AIC provided the highest savings in equivalent energy (3.33%) as it enhances the absorption capacity of MDEA/PZ blend, thus reducing the regeneration energy. RSP provided negative energy savings (-1.85%) due to consideration of additional heat consumption by the preheater in the overall energy analysis. The integration of waste heat from flue gas cooling process for the rich

MDEA/PZ preheating process is not possible due to significant difference in specific heat capacities of flue gas and rich MDEA/PZ. RSS resulted in an energy savings of 1.12% as there is less pre-stripping of CO<sub>2</sub> by hot vapors from heated fraction of amine, hence having less impact on the driving force for CO<sub>2</sub> desorption. RSF had negligible energy savings (0.10%) as the reduction in regeneration energy was almost equivalent to an increase in the sensible heat demand in the stripper due to flashing the heated rich solvent. LVC and RVC have similar mechanisms; however, the results demonstrate that LVC provided positive energy savings (2.68%) while RVC resulted in negative energy savings (-4.90%). This is because the electrical energy consumption by rich amine compressor and pump is relatively higher than the thermal energy reduction by RVC. IHS provided an energy savings of 2.22% as the heating of semi lean solvent elevates the temperature of the stripper which enhances the CO<sub>2</sub> desorption kinetics, thus lowering the overall energy requirement.

Various combinations of three configurations were proposed, so the advantages of each configuration can be combined to have the maximum reduction in the regeneration energy. The energy analysis results show that all combinations resulted in positive energy savings with AIC+RSS+LVC providing the lowest regeneration energy (2.12 GJ/t<sub>CO2</sub>). However, LVC consumes substantial amount of electrical energy which reflects in the equivalent energy results concluding that AIC+RSS+IHS as the optimal combination in terms of overall energy savings (8.96%). In comparison to MEA base case (1.202 GJ/t<sub>CO2</sub>) present in the literature, the hybrid case (MDEA/PZ and AIC+RSS+IHS) provides 18.2% savings in the equivalent energy.

#### 4. Conclusion and Future Work

This study investigated the performance of seven process configurations and their combinations applied to MDEA/PZ based PCC process attached with 750 MW CCGT power plant. In addition, the study demonstrated that combining three configurations showed higher energy savings than individual configurations due to combined advantages of each configuration. The results of the energy analysis conclude that AIC+RSS+IHS is the optimum combination based on the equivalent energy savings. The future work should include a detailed techno-economic analysis to gain deeper insights on the tradeoffs between energy savings and higher capture costs of different configurations in order to identify optimum combination in terms of energy and cost savings.

#### References

- S. Bishnoi and G. Rochelle, 2002, Thermodynamics of Piperazine/Methyldiethanolamine/Water/Carbon Dioxide, *Industrial & Engineering Chemistry Research*, 41, 3, 604-612
- F. Cheng, N. Patankar, S. Chakrabarti and J. Jenkins, 2022, Modeling the operational flexibility of natural gas combined cycle power plants coupled with flexible carbon capture and storage via solvent storage and flexible regeneration, *International Journal of Greenhouse Gas Control*, 118, 103686
- A. Ibrahim, F. Ashour, A. Ghallab and M. Ali, 2014, Effects of piperazine on carbon dioxide removal from natural gas using aqueous methyl diethanol amine, *Journal of Natural Gas Science and Engineering*, 21, 894-899
- B. Khan, A. Ullah, M. Saleem, A. Khan, M. Faiq and M. Haris, 2020, Energy Minimization in Piperazine Promoted MDEA-Based CO<sub>2</sub> Capture Process, *Sustainability*, 12, 20, 8524
- K. Li, A. Cousins, H. Yu, P. Feron, M. Tade, W. Luo and J. Chen, 2016, Systematic study of aqueous monoethanolamine-based CO<sub>2</sub> capture process: model development and process improvement, *Energy Science & Engineering*, 4, 1, 23-39

# PREDICTION OF BIOMEDICAL WASTE GENERATION IN SANITARY EMERGENCIES FOR URBAN REGIONS USING MULTIVARIATE RECURRENT NEURAL NETWORKS

Nicolas Galvan-Alvarez<sup>a</sup>, David Rojas-Casadiago<sup>a</sup>, Viatcheslav Kafarov<sup>a</sup>,  
David Romo-Bucheli<sup>a</sup>

<sup>a</sup> *Universidad Industrial de Santander, Bucaramanga, Santander, Colombia.*

## Abstract

Biomedical waste (BMW) generation is severely affected by generalized sanitary emergencies such as epidemics, as shown recently during the COVID-19 pandemic. These sanitary emergencies often increase plastic use in personal protection items, single-use plastics, and other healthcare elements. This increase might surpass the capacity of the waste management mechanism of a specific region, leading to a potential increase in its population health risks. Predicting the trends of BMW generation is not straightforward because it depends on several variables associated with the local health system and the health emergency status. However, a substantial amount of work has been done in epidemics modelling. Our main hypothesis is that BMW generation is strongly associated with sanitary emergencies dynamics. We propose a simulation framework that uses historical data from an ongoing sanitary emergency to build a model that can predict BMW generation trends in urban regions of developing countries.

**Keywords:** Biomedical waste, Simulation, Epidemics, Neural networks, Developing countries.

## 1. Introduction

Biomedical waste (BMW) is defined as any waste produced during the diagnosis, treatment, or immunization of human or animal research activities or in the production or testing of biological or in health camps. Recently disposal of BMW has become a major problem all over the world due to the influence of health emergencies such as epidemics on the generation of this type of waste. Specifically, the COVID-19 pandemic has exacerbated the problem of BMW management in several regions around the world. For instance, the world health organization (WHO) estimates that between March 2020 and November 2021, approximately 87,000 tons were purchased and are presumed to have ended up as waste (WHO, 2022). Also, according to press releases in March 2021, the generation of medical waste (infectious and non-infectious) increased significantly (+370%) in Hubei Province, with a high proportion of plastics. Additionally, from January 20 to March 31, 2020, accumulated medical waste across China was estimated at 207 kt, and in Wuhan medical waste increased from a normal level of 40 tons per day (t/d) to a peak of 240 t/d, exceeding the maximum incineration capacity in that city of 49 t/d (Klemeš, 2020). The purpose of this work is to develop a prediction model to estimate

the amount of BMW during different scenarios associated with sanitary emergencies, such as the COVID-19 pandemic. The main motivation of this type of simulation models would be to inform the government institutions in the selection of strategies for BMW management and avoid potential health risks.

## 2. Related Work

Several strategies have been used to simulate waste generation and treatment. Kannangara et al. carried out a simulation study for waste management. The purpose was to develop models for accurate prediction of municipal solid waste generation and diversion from demographic and socioeconomic variables (Kannangara, 2018). Additionally, in May 2021, an article was published by Vu et al. in which machine learning models are used for modelling municipal solid waste disposal rates during COVID-19 (Vu, 2021). On the other hand, Yu et al. proposed a reverse logistics network strategy to address the drastic increase in healthcare waste generation during the pandemic (Yu, 2020). This reverse logistics network aims to optimize decisions to establish temporary waste management facilities and transportation strategies at different stages. Similarly, in 2022 Chowdhury et al. carried out a simulation study by calculating the amount of waste generated per patient. A full estimate of the healthcare waste generated from March 2020 to May 2021, during the COVID-19 pandemic in the period, depended on whether the patient is infected, isolated, deceased, ICU or in quarantine (Chowdhury, 2022). The main contribution of this work is the development of a simulation strategy aiming to predict the amount of BMW in a specific geographical region based on epidemics variables. The simulation integrates a “Susceptible Exposed Infectious Recovered Deceased” (SEIR-D) compartmental model and a multivariate RNN (LSTM) model. The integrated strategy leverages the SEIRD model, broadly known in epidemics modelling, to analyze different sanitary emergency scenarios (pessimistic, neutral, and optimistic). The SEIR-D model generates synthetic data associated with the number of cases and deaths. These variables are then fed to the LSTM to simulate BMW generation rates. The simulation tool allows to study the influence of the sanitary emergency dynamics and its environmental impact.

## 3. Data Collection

### 3.1. India's COVID-19 dynamics and India Biomedical Waste Dataset

For this study, a dataset was constructed with daily values of BMW generation for five states of India. Monthly reports available from the central pollution control board (CPCB) of India's ministry of environment, forest, and climate change, with measurements of BMW during COVID-19 emergency (from June 2020 to December 2021) were retrieved (CPCB, 2022). For the COVID-19 dynamics, we added daily data on infected and deceased individuals for each of the states and territorial unions of India. The data used come from daily worldwide reports from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (Dong, 2020). Additionally, data associated with changes in residential mobility, workplaces, and public transportation from Google's Local Mobility reports were also included (Google, 2021).

### 3.2. COVID-19 Colombian Dataset and AMB Biomedical Waste Dataset

Data on COVID-19 cases and deaths in Bucaramanga Metropolitan Area (AMB in Spanish), were extracted from a dataset designed by the Colombian government, which includes COVID-19 cases and deaths for each municipality in Colombia. On the other hand, BMW data was obtained from a waste collection company in AMB. Moreover, data associated with changes in residential mobility, workplaces, and public transportation from Google's Local Mobility reports on COVID-19 were also included (Google, 2021).

### 4. Methods

The simulation model consists of two integrated components: a health emergency dynamic modelling component (COVID-19) and a BMW generation estimation component (Figure 1a).

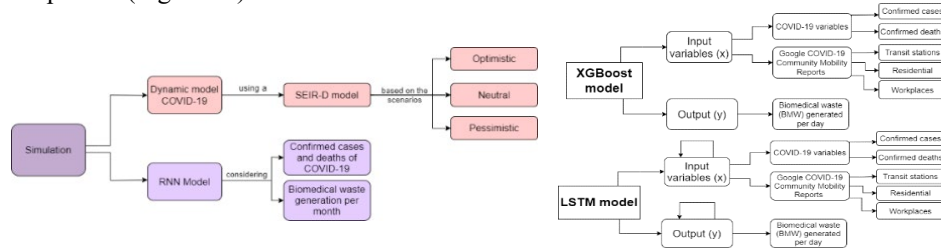


Figure 1. Left (a): Schematic of the model. Upper right (b): Inputs and outputs of the XGBoost model. Predictions for x days consider only the input variables of these x days. Down right (c): Inputs and outputs of the LSTM model. The network previous state is used to make a new prediction.

#### 4.1. XGBoost Model

For the BMW estimation component, the XGBoost algorithm (Figure 1b) was implemented. The model is fed with variables such as confirmed COVID-19 cases and deaths per day, and mobility indices (residential, workplaces and transit stations). The task is to predict the BMW. It should be clarified that the model predicts data for the same day on which it is predicted, i.e., the data of the 5 variables mentioned for a certain day are entered, and the amount of BMW is predicted for that same day.

#### 4.2. Multivariate LSTM Model

Subsequently, to develop a much more complete model, a recurrent neural network (RNN) was implemented for the prediction of such BMW, which is a type of neural network that is naturally adapted to the processing of time series data and other sequential data. To predict biomedical residuals based on COVID and mobility variables, a type of RNN called “Long short-term memory” (LSTM) was used (Figure 2). Introduced by Cho et al. in 2014, LSTM aims to solve the leakage gradient problem that comes with a standard recurrent neural network (Cho, 2014).

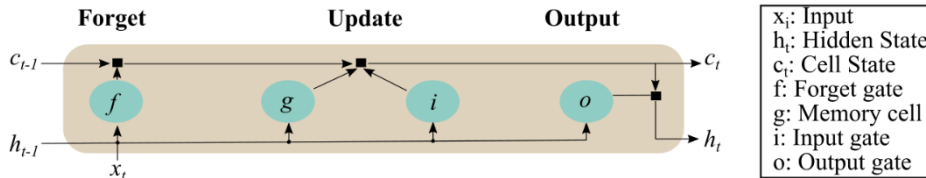


Figure 2. Structure of the LSTM cell (Adapted from Matlab, 2023).

Similar to the XGBoost model, the LSTM (Figure 1c) is fed with variables such as confirmed cases, deaths per day, and some mobility indices calculated during the pandemic (residential, workplaces, and transit stations). The model aims to predict BMW in a specific time window. However, the main difference with respect to the XGBoost is that it takes 7 days of input as a reference for the initialization of the model, and from then on it is fed back with the data it is predicting. This behavior is more convenient, since it allows us to make longer predictions with a smaller amount of data.

Several LSTM models for each region were trained with data from: new COVID-19 cases and deaths, mobility indexes, and amount of BMW from June 11st 2020 to June 29th, 2021. First, data is normalized by the region population. Then a windowing



process is performed, which consists of dividing the data into windows of 7 days. The LSTM model is trained to predict the amount of BMW on the 8th day.

#### 4.3. Epidemics Model (SEIR-D)

For the dynamics component (COVID-19) a SEIR-D compartmental epidemiological model was implemented, which is applied for the mathematical modelling of infectious diseases and allows dividing the individuals of a population  $N$  into five compartments: Susceptible (S), Exposed (E), Infected (I), Recovered (R) and Deceased (D), which are governed by the following set of ordinary differential equations:

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\beta SI}{N} & \frac{dE}{dt} &= \frac{\beta SI}{N} - \sigma E & \frac{dD}{dt} &= \mu I \\ \frac{dI}{dt} &= \sigma E - \gamma I - \mu I & \frac{dR}{dt} &= \gamma I & N &= S + E + I + R + D \end{aligned}$$

Equation 2.

Where  $\beta$  is the rate at which infected individuals interact with other infected individuals,  $\sigma$  is the estimated incubation period of the disease,  $\gamma$  is the duration of illness or recovery rate,  $\mu$  is the mortality rate due to the disease, and  $R_0$  is the transmission of the disease or basic reproduction number. Considering the variability in the trends and the different COVID-19 dynamics, pessimistic, neutral, and optimistic scenarios were proposed based on historical data from the AMB. The parameters for each of the scenarios were defined by optimizing the least squares of the residual data between the initial simulations and the historical data depending on whether there was a high number of cases (pessimistic scenario), a medium number of cases (neutral scenario), or a low number of cases (optimistic scenario). Table 1 shows the optimized parameters for each scenario.

Table 1. Parameters defined for the scenarios.

Scenario	$R_0$	$\beta$	$\sigma$	$\gamma$	$\mu$
Optimistic	1.4	0.121	0.001	0.083	2.21E-10
Neutral	1.99	0.173	0.002	0.043	1.39E-04
Pessimistic	1.07	6.755	0.815	6.274	-2.28E-04

#### 4.4. Performance metrics for predictive models

The mean squared error (MSE) was used to evaluate the models. MSE represents the mean squared difference between the estimated values and the true value, and can be expressed as  $MSE = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2$ , where  $x_i$  are the estimated values, and  $y_i$  are the actual values of the corresponding vectors.

## 5. Results

This section will explain the results by analyzing the predictive capability of the LSTM model and subsequently integrating it with the full model.

### 5.1. Results of multivariate predictive models

When evaluating the prediction of BMW in the AMB, the XGBoost obtained a mean square error of 0.2551 for the 3 months predicted, obtaining the results presented in Figure 2a. As can be seen, the performance of this model is poor, which may be since in this prediction period there is a drop in BMW never seen before in the training window. On the other hand, Table 3 shows the results of the LSTM model based on its MSE. For the study, 5 states of India whose population numbers are similar to the AMB were taken. As can be seen, the best and worst performing region was Nagaland and the AMB respectively. The associated error for the Nagaland state is extremely low because in the predicted period its BMW and COVID variables remained somewhat constant

(probably related to bad BMW data collection). On the other hand, the main cause of the poor results in the AMB region is that the model has not seen such a steep drop in the amount of BMW during the training phase, so it is very difficult to predict.

Table 2. MSE results for the LSTM Model.

Region	window-size		
	7	14	28
Puducherry	0.025368	0.140410	0.152633
Goa	0.030012	0.025989	0.007791
Manipur	0.018458	0.012636	0.016037
Nagaland	5.403496E-05	8.69878E-06	2.12439E-07
Mizoram	0.001639	0.000436	0.000371

More specifically, the results obtained in the AMB case study are presented in Figure 2b, obtaining a mean square error of 0.1186 tons on average for the 3 months predicted, using the model with a 7-day window.

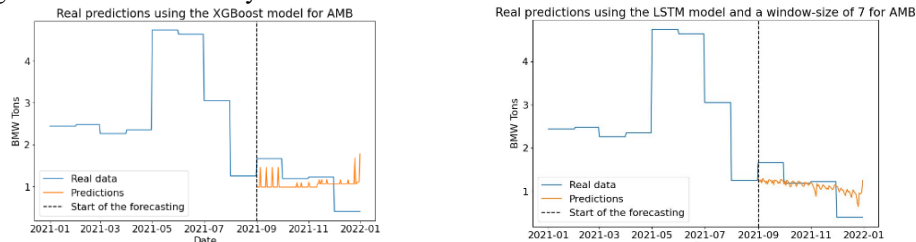


Figure 3. Right (a): LSTM model results for the AMB. Left (b): XGBoost model results for the AMB.

### 5.2. SEIR-D integrated model

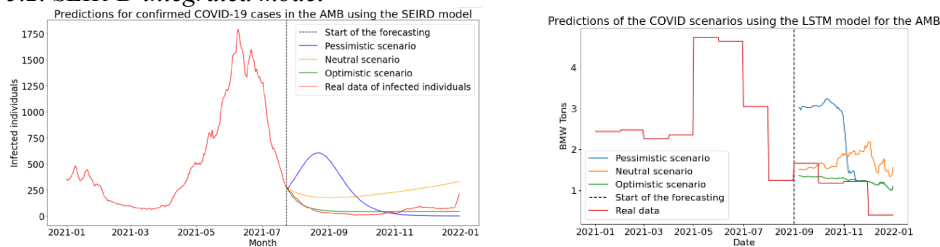


Figure 4. Left (a): SEIR-D model predictions for confirmed COVID-19 deaths in the AMB. Right (b): Results of the LSTM model for simulating BMW generation for the three COVID-19 scenarios in the AMB.

Once the parameters of Table 1 were defined, predictions were made with the SEIR-D model, and the results collected (Figure 3a). The pessimistic scenario has a peak of infected persons in September 2021. The cases, and deaths of the neutral scenario increase over time, and both variables remain constant for the optimistic scenario. The cases and deaths predicted by the SEIR-D model were used as input for the LSTM network (Figure 1c), and for the mobility indicators the actual data was used. Figure 3b shows the predictions for the proposed scenarios: the pessimistic scenario produces the highest amount of BMW at the peak of COVID cases in September 2021. The neutral scenario has the second largest BMW increase. Finally, for the optimistic scenario, the output is fairly constant and has the lowest amount of BMW. The simulation results are difficult to evaluate, but the general tendency of each scenario is coherent. Additionally, it is important to note that the closest scenario to the actual COVID19 data (optimistic scenario), also obtained the smallest error in terms of BMW generation (Table 3).

Table 3. MSE results for the simulated scenarios in the integrated model.

Scenario	COVID-19 cases	BMW
Pessimistic	67513.363322	1.516234
Neutral	29823.396586	0.700147
Optimistic	<b>1013.564280</b>	<b>0.172478</b>

## 6. Conclusions

In this work, a widely used epidemic model for modelling the transmission of diseases (SEIR-D) was successfully developed as an initial stage to simulate the process of plastic BMW generation. The SEIR-D model simulates different scenarios (pessimistic, neutral, and optimistic) for the development of a health emergency. The output of the SEIR-D compartmental model is then fed to multivariate recurrent models that can estimate BMW generation rates that are consistent with historical values observed in a specific geographic region. Simulation results demonstrate that the recurrent neural networks successfully capture the characteristics associated with the healthcare system of a particular location during a health emergency. We have integrated data from different domains (i.e, epidemiological (number of cases and deaths) and mobility indexes (based on Google mobility report)) to estimate the biomedical waste generation via a recurrent neural network. The model trained on historical data, could be useful afterwards to simulate different developing scenarios for sanitary emergencies. The simulation outputs are consistent with expected tendencies, showing that non-linear recurrent models using multimodal information might become an important modelling approach for complex scenarios.

### Acknowledgments

Acknowledgements to the Universidad Industrial de Santander, the Vice Rector's Office for Research and Extension (VIE) and to the internal call for research "UIS contributions to the COVID-19 pandemic".

### References

- Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. ArXiv:1406.1078 [Cs, Stat]. <http://arxiv.org/abs/1406.1078>
- COVID-19 biomedical waste management status. "CPCB | Central Pollution Control Board." <https://cpceb.nic.in/covid-waste-management/> (accessed Jan. 20, 2022).
- E. Dong, H. Du, and L. Gardner, "An interactive web-based dashboard to track COVID-19 in real time," *The Lancet Infectious Diseases*, vol. 20, no. 5, pp. 533–534, May 2020, doi: 10.1016/S1473-3099(20)30120-1.
- Google, "COVID-19 Community Mobility Report". <https://www.google.com/covid19/mobility?hl=es> (accessed Jun. 21, 2021).
- H. L. Vu, K. T. W. Ng, A. Richter, N. Karimi, and G. Kabir, "Modeling of municipal waste disposal rates during COVID-19 using separated waste fraction models," *Science of The Total Environment*, vol. 789, p. 148024, Oct. 2021, doi: 10.1016/j.scitotenv.2021.148024.
- H. Yu, X. Sun, W. D. Solvang, and X. Zhao, "Reverse Logistics Network Design for Effective Management of Medical Waste in Epidemic Outbreaks: Insights from the Coronavirus Disease 2019 (COVID-19) Outbreak in Wuhan (China)," *IJERPH*, vol. 17, no. 5, p. 1770, Mar. 2020, doi: 10.3390/ijerph17051770.
- J. J. Klemeš, Y. V. Fan, R. R. Tan, and P. Jiang, "Minimising the present and future plastic waste, energy and environmental footprints related to COVID-19," *Renewable and Sustainable Energy Reviews*, vol. 127, p. 109883, Jul. 2020, doi: 10.1016/j.rser.2020.109883.
- Matlab, "Long Short-Term Memory (LSTM) Networks". <https://es.mathworks.com/discovery/lstm.html> (accessed Feb. 6, 2023).
- M. Kannangara, R. Dua, L. Ahmadi, and F. Bensebaa, "Modeling and prediction of regional municipal solid waste generation and diversion in Canada using machine learning approaches," *Waste Management*, vol. 74, pp. 3–15, Apr. 2018, doi: 10.1016/j.wasman.2017.11.057.
- T. Chowdhury, H. Chowdhury, M. S. Rahman, N. Hossain, A. Ahmed, and S. M. Sait, "Estimation of the healthcare waste generation during COVID-19 pandemic in Bangladesh," *Science of The Total Environment*, vol. 811, p. 152295, Mar. 2022, doi: 10.1016/j.scitotenv.2021.152295.
- World Health Organization, "Tonnes of COVID-19 health care waste expose urgent need to improve waste management systems." <https://www.who.int/news/item/01-02-2022-tonnes-of-covid-19-health-care-waste-expose-urgent-need-to-improve-waste-management-systems> (accessed Nov. 15, 2022).

# Hybrid Modeling and Multi-Fidelity Approaches for Data-Driven Branch-and-Bound Optimization

Suryateja Ravutla<sup>a</sup>, Jianyuan Zhai<sup>b</sup>, Fani Boukouvala<sup>a</sup>

<sup>a</sup>*Department of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA*

<sup>b</sup>*Engineering & Data Sciences, Cargill Inc, Shanghai, 200031, China*

Email: [sravutla3@gatech.edu](mailto:sravutla3@gatech.edu)

## Abstract

High-Fidelity (HF) simulations are essential in quantitative analysis and decision making in engineering. In cases where explicit equations and/or derivatives are unavailable, or in the form of intractable nonlinear formulations, simulation-based optimization methods are used. We recently proposed a data-driven equivalent of spatial branch-and-bound that constructs underestimators of high-fidelity simulation data. Within this framework, low-fidelity surrogate data can also be used to inform underestimators. In this work, we utilize the recent advances in hybrid multifidelity surrogate modeling techniques to improve the validity of our underestimators, which leads to better bounds and incumbent optima with lower sampling requirements. Specifically, we show that by modeling the error between the high-fidelity and low-fidelity data, the surrogates learn more about the underlying function with less sampling requirements.

**Keywords:** Data-driven Optimization, Branch-and-bound, Convex underestimators, Multifidelity surrogate models, neural networks, support vector regression, hybrid modeling.

## 1. Introduction

In many engineering fields, it is desirable to simulate complex processes and use these quantifications of system performance for decision-making. Many such cases exist for chemical engineers, including molecular simulations, flowsheet simulations, computational fluid dynamic models, agent-based models, and more. These models are referred to as High-Fidelity (HF) simulations. But often these HF function evaluations are computationally expensive, and this implies that the number of function evaluations is limited by cost or time. Adding to this difficulty, in most cases, the objective functions and the constraints are only available as black-box evaluation function outputs (Fisher, Watson et al. 2020, van de Berg, Savage et al. 2022, Zhai and Boukouvala 2022). As a result, optimization of such systems becomes increasingly prohibitive. The absence of information on the system calls in for derivative-free optimization (DFO) or simulation-based optimization techniques. DFO techniques can be broadly classified into Sampling-based and Model-based methods. Numerous algorithms have been proposed in the recent years (Rios and Sahinidis 2013) for both cases. Sampling-based methods rely on comparing the function values directly and utilize this information to further sample adaptively. On the other hand, Model-based methods rely on constructing Machine Learning (ML) surrogate models as approximations of HF simulations with the aim to expedite the optimization process (Kim and Boukouvala 2020, Li, Dong et al. 2021). A disadvantage of using sampling-based methods is that they require too many samples to

infer the optimum solution, while in surrogate model-based approaches, constructing an accurate surrogate can be challenging and computationally expensive. A solution that we have recently proposed combines advantages of both the methods in the form of a data-driven equivalent of the spatial branch-and-bound algorithm (DDSBB) (Zhai and Boukouvala 2022). The key idea of DDSBB is based on constructing convex underestimators of simulated data. A schematic of spatial branch-and-bound and its data-driven equivalent are shown in Figure 1.

These underestimators serve as relaxations and are convex, so they can be efficiently optimized, circumventing the task of directly optimizing nonconvex fitted surrogates. Samples drawn from the HF simulation serve as upper bounds (UB) of the global optimum and the minimum of the convex underestimator serve as lower bounds (LB). The search space is then progressively partitioned by using branching, node selection and pruning rules and adaptively sampling in the non-pruned subspaces. To build underestimators, using  $N$  samples, the underlying formulation is shown in Equation 1.

$$\begin{aligned}
 & \min_{\mathbf{a}, \mathbf{b}, c} \sum_i^N (f(\mathbf{x}_i) - f_{lb}(\mathbf{x}_i)) \\
 & \text{s.t. } f(\mathbf{x}_i) - f_{lb}(\mathbf{x}_i) \geq 0 \quad \forall i = 1 \text{ to } N \\
 & f_{lb}(\mathbf{x}_i) = \mathbf{a}\mathbf{x}_i^2 + \mathbf{b}\mathbf{x}_i + c \quad \forall i = 1 \text{ to } N \\
 & \mathbf{a} \geq 0, \quad \mathbf{a}, \mathbf{b} \in \mathbb{R}^D, \quad c \in \mathbb{R}
 \end{aligned} \tag{1}$$

In contrast to the conventional sample-based and surrogate-based approaches, DDSBB employs some of the positive aspects of both. Like the sample-based approaches, it adaptively samples in the search space and uses this information to improve the bounding of the original function, by constructing the relaxations and pruning the subspaces that are not promising. It also utilizes surrogates to construct Low fidelity (LF) data but does not directly optimize the surrogates or rely on a single surrogate prediction. This LF data can be utilized along with the HF data to build the convex relaxations. In our recent work (Zhai and Boukouvala 2022), we have shown that by jointly using LF and HF data (multifidelity MF), we can optimize a higher fraction of benchmark problems

While the MF approach employed in DDSBB showed promising performance, a fraction of benchmark studies was still not optimized given limits on sampling requirements. In this current work, we improve the performance of our framework by incorporating more advanced hybrid multifidelity modeling techniques. Specifically, using the same amount of HF data as before, we attempt to learn more about our underlying black-box problem by modeling the error between the HF and LF data. The hypothesis is that this will overall improve our underestimators, and consequently the overall efficiency of the DDSBB

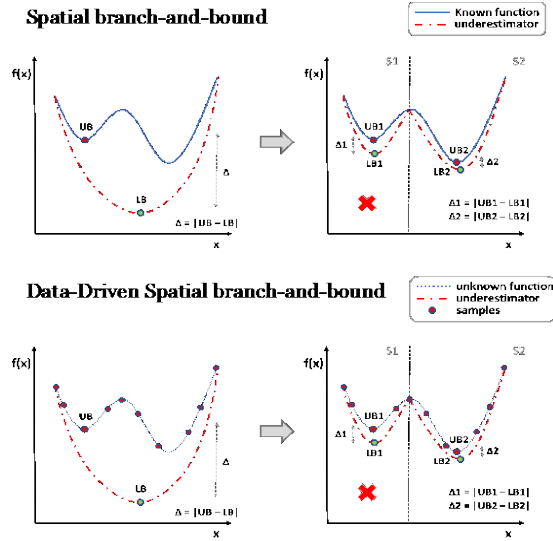


Figure 1: Deterministic Spatial Branch-and-Bound, and Data-Driven Spatial Branch-and-Bound, and the process of branching and bounding in both variants

approach, without increasing sampling requirements. Recent studies show that there is an increasing amount of focus in constructing surrogate models that combine data with different fidelity (Meng and Karniadakis 2020, Bradley, Kim et al. 2022). These Multifidelity Surrogate Models (MFSMs) exploit the relation between LF data and HF data. In this study we integrate these MFSMs into the DDSBB architecture and quantify their performance by benchmarking the method on 2-10 dimensional black-box optimization problems.

## 2. Overview and explored methods

### 2.1.1. Multifidelity approach in DDSBB

The overview of DDSBB is shown in Figure 2. Initially, an input design is generated in the variable search domain by employing Latin Hypercube Sampling (LHS) to generate HF samples. This HF samples are then used for constructing quadratic underestimators by employing the formulation shown in Equation (1). This is shown in the figure using a solid red line. Subsequently, a set of branching, node selection and pruning rules are used to adaptively add samples in the non-pruned subspaces until convergence. Alternatively, as shown in red dotted line, if the MF approach is selected, these HF samples are used to generate LF surrogate models and LF samples. These LF samples are then used along with the HF samples to construct the underestimators. Currently, DDSBB has the capability to use Support Vector Regression (SVR), Neural Networks (NN) and Gaussian Process Regression (GPR) as the surrogate options. For the rest of the study, we utilize SVR as the surrogate option.

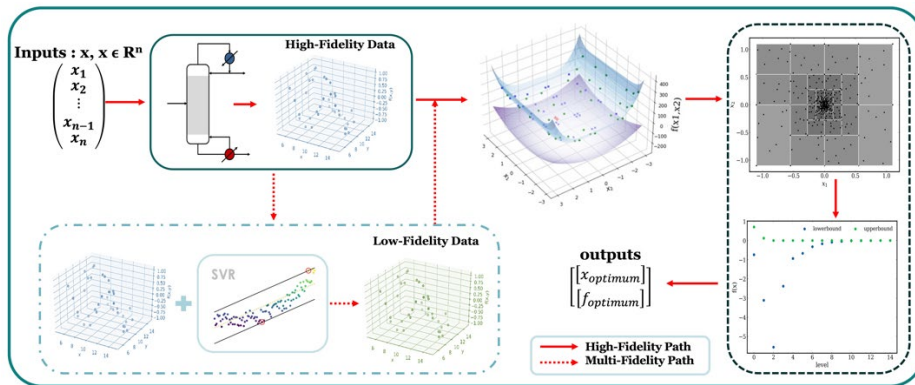


Figure 2: Overview of DDSBB. High-fidelity approach is shown in solid red line and multi-fidelity approach is shown in dotted red line.

### 2.1.2. Multifidelity Surrogate models to improve LF data

There is an inherent trade-off between the number of HF samples that are used to construct LF surrogates, and the accuracy of the constructed surrogates. Also, adding the LF samples makes the convex underestimators more conservative with respect to bounding the function. We have found that in certain cases this leads to improving their validity, and as a result it leads to locating the global optimum of challenging benchmarks (Zhai and Boukouvala 2022). At the same time, LF surrogate predictions can make the underestimators overly conservative, thus leading to large sampling requirements for convergence. In this work, our hypothesis is that, MFSMs can be used to improve the accuracy of the LF data, which leads to improvement in underestimator validity. A

widely used correlation to build MFSMs is:  $y_H = \rho(x)y_L + \delta(x)$  where  $y_L, y_H$  represent the low and high-fidelity data respectively,  $\rho(x)$  is multiplicative correlation surrogate and  $\delta(x)$  is the additive surrogate. In a more general way, we can re-write it as  $y_H = F(x, y_L)$ . To establish a correlation between the HF and LF data, one would need a HF and a LF model to generate that data (Meng and Karniadakis 2020, Bradley, Kim et al. 2022). In case an LF model is available, it can be directly used. In cases where the LF model is not available, we propose a framework (Workflow 1) shown below to create a LF model using a fraction of available data.

---

### Workflow 1: Constructing MFSMs

---

Let the data set  $[\mathbf{x}_{HF}, \mathbf{y}_{HF}]_{tot}$  represent the complete HF data.

Generate a training set data  $[\mathbf{x}_{HF}, \mathbf{y}_{HF}] = 75\%[\mathbf{x}_{HF}, \mathbf{y}_{HF}]_{tot}$

Set  $\mathbf{x}_{HF} \leftarrow input$  and  $\mathbf{y}_{HF} \leftarrow output$ ,

While termination criteria **not true**:

TRAIN SVR

$$\text{Calculate } MSE_{SVR} = \frac{1}{N_{LF}} \sum_{i=1}^{N_{LF}} (|y_{LF} - y_{HF}|^2) + \beta_1 \|\phi_{SVR}\|_2$$

Tune SVR parameters

Generate LF dataset  $[\mathbf{x}_{HF}, \mathbf{y}_{LF}]_{tot}$  using HF input  $[\mathbf{x}_{HF}]_{tot}$

Utilize the correlation  $y_H = F(x, y_L)$  to model error between LF and HF outputs  $[y_{LF}]_{tot}$  and  $[y_{HF}]_{tot}$  respectively using a NN

Set  $[\mathbf{x}_{HF}, \mathbf{y}_{LF}]_{tot} \leftarrow input$  and  $\mathbf{y}_{HF} \leftarrow output$ ,

While termination criteria **not true**:

TRAIN NN

$$\text{Calculate } MSE_{NN} = \frac{1}{N_{HF}} \sum_{i=1}^{N_{HF}} (|y_{LF}^* - y_{HF}|^2) + \beta_2 \|\phi_{NN}\|_2$$

Tune NN parameters

---

In the Workflow 1,  $y_{LF}, y_{LF}^*$  represent output from SVR and NN.  $N_{LF}, N_{HF}$  is the number of training data points and total HF data points.  $\beta_1, \beta_2$  represent the regularization weights and  $\phi_{SVR}, \phi_{NN}$  represent the associated parameters with SVR and NN respectively. A schematic for the Workflow 1 is shown in Figure 3.

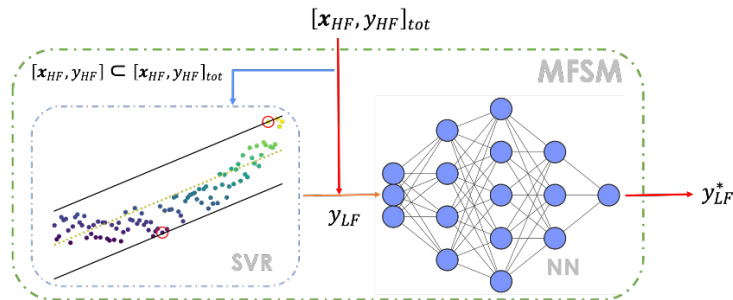


Figure 3: A schematic of the workflow for fitting the MFSMs. Part of the total HF data set (shown in blue) is used to train LF SVR model and generate LF data  $y_{LF}$

### 3. Results and discussion

To understand and visualize the effect of MF and MFSM approach on underestimator construction, we first take a 1-dimensional case study. Let us consider that the function  $f(x) = \sin(x) + \sin(10x/3)$  is available as a black-box function, with  $x \in [0, 9]$  and known global solution at  $x_{opt} = 5.14574$ . Figure 4 shows underestimators constructed using HF, MF and MFSM approaches. We can see among all three approaches, the

optimum solution identified is more accurate in MFSM approach and LF data in case of MFSM approach is more accurate in comparison to MF approach, without increasing the number of HF data collected.

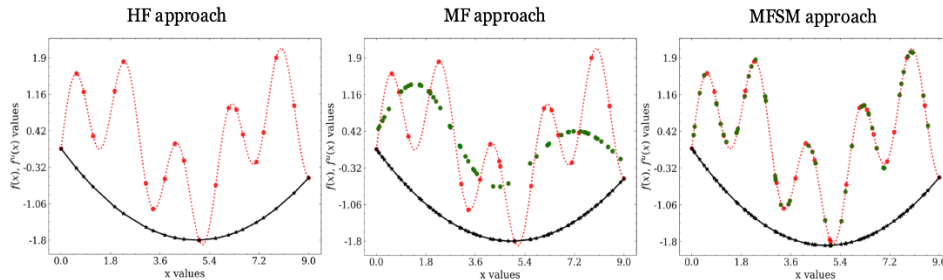
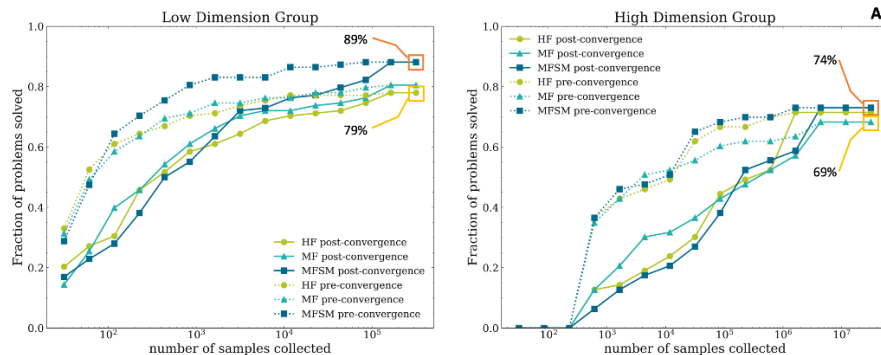


Figure 4: Comparison of constructed underestimators and LF data using HF, MF and MFSM approaches. Black-box function is shown in red dotted line, HF samples in red circles and LF samples in green circles. Underestimator is shown in solid black line.

Next, we test the performance of all three approaches on a large set of continuous box-constrained benchmark problems with known global solutions (BARON 2022). The benchmark problems were divided into two groups based on their dimensionality: lower dimensional group containing 118 problems with 2-3 variables and higher dimensional group containing 63 problems with 4-10 variables. All three approaches were initialized with  $10 \times \text{dimension} + 1$  initial samples and converge when the  $|UB - LB| \leq 0.05$  or  $|UB - LB|/|LB| \leq 0.001$ . For the performance analysis, we use the criterion  $f_{best} \leq \max(f^* + 0.01, (1.01)f^*)$  to allow a tolerance limit towards the optimal solution found.  $f^*$  and  $f_{best}$  represent the known global solution and the solution reported by DDSBB as optimum. We study the fractions of problems solved with sampling and CPU requirements based on pre and post convergence solutions. *Pre-convergence* and *post-convergence* solutions represent the optimum solution found by the algorithm *before* and *after* closing the UB-LB gap respectively. The performance curves for all three approaches are shown in Figure 5. In Figure 5A, we can see that the MFSM approach solves higher fraction of benchmark problems, and the pre-convergence solution quality is better in both low and high dimensions groups. In Figure 5B, we show the CPU requirements for the three approaches. The HF approach does not involve any surrogate fitting, so the CPU time largely corresponds to HF sampling. On the other hand, CPU time in MF and MFSM approaches also include surrogate modeling costs. As expected MFSM approach takes higher CPU time for fitting the complex MFSM structure. Thus, the advantage of this method is expected to be even greater when sampling cost increases.





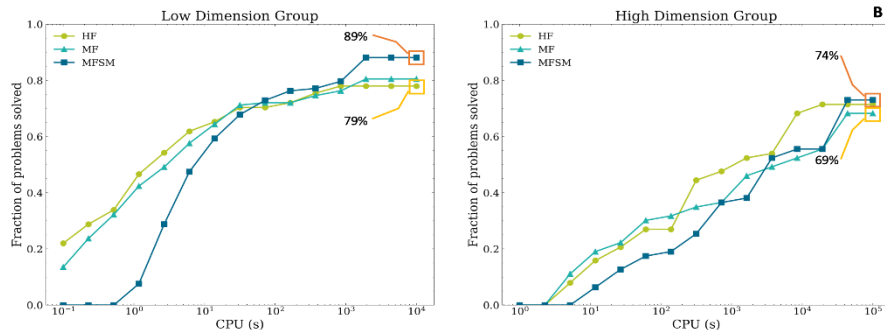


Figure 5: Performance curves of HF, MF, MFSM approaches. A) Fraction of problems solved vs sampling requirement, compared with pre and post convergence solutions reported by the algorithm. B) Fraction of problems solved vs CPU requirement.

#### 4. Conclusions

In this work, we utilized multifidelity surrogate models as composite structures to model the error between HF and LF data to improve the validity of constructed data-driven underestimators embedded within a branch-and-bound framework. Results show that using composite/hybrid multifidelity models for surrogate-based optimization is promising, because it leads to more accurate surrogates with the same sampling cost but requires additional CPU time for training.

#### 5. Acknowledgements

The authors acknowledge support from the National Science Foundation (NSF-1944678)

#### References

- BARON (2022). Bound-constrained programs.
- Bradley, W., J. Kim, Z. Kilwein, L. Blakely, M. Eydenberg, J. Jalvin, C. Laird and F. Boukouvala (2022). "Perspectives on the integration between first-principles and data-driven modeling." *Computers & Chemical Engineering* **166**: 107898.
- Fisher, O. J., N. J. Watson, J. E. Escrig, R. Witt, L. Porcu, D. Bacon, M. Rigley and R. L. Gomes (2020). "Considerations, challenges and opportunities when developing data-driven models for process manufacturing systems." *Computers & Chemical Engineering* **140**: 106881-106881.
- Kim, S. H. and F. Boukouvala (2020). "Surrogate-based optimization for mixed-integer nonlinear problems." *Computers & Chemical Engineering* **140**: 106847-106847.
- Li, Z., Z. Dong, Z. Liang and Z. Ding (2021). "Surrogate-based distributed optimisation for expensive black-box functions." *Automatica* **125**: 109407.
- Meng, X. and G. E. Karniadakis (2020). "A composite neural network that learns from multi-fidelity data: Application to function approximation and inverse PDE problems." *Journal of Computational Physics* **401**: 109020-109020.
- Rios, L. M. and N. V. Sahinidis (2013). "Derivative-free optimization: a review of algorithms and comparison of software implementations." *Journal of Global Optimization* **56**(3): 1247-1293.
- van de Berg, D., T. Savage, P. Petsagkourakis, D. Zhang, N. Shah and E. A. del Rio-Chanona (2022). "Data-driven optimization for process systems engineering applications." *Chemical Engineering Science* **248**: 117135-117135.
- Zhai, J. and F. Boukouvala (2022). "Data-driven spatial branch-and-bound algorithms for box-constrained simulation-based optimization." *Journal of Global Optimization* **82**(1): 21-50.

# Intensified Process for the Production of Biojet Fuel from Mexican Biomass

Rosa Angelica Morales-Gutierrez<sup>a</sup>, Eduardo Sánchez-Ramírez<sup>a</sup>, Brenda Huerta-Rosas<sup>a</sup>, Juan José Quíroz-Ramírez<sup>b</sup>, Gabriel Contreras-Zarazua<sup>b</sup>, Juan Gabriel Segovia-Hernández<sup>a</sup>

<sup>a</sup> *Universidad de Guanajuato, Noria Alta S/N, Guanajuato 36050, México*

<sup>b</sup> *CONACyT – CIATEC A.C. Centro de Innovación Aplicada en Tecnologías Competitivas, Omega 201 Col. Industrial Delta, León 37545, México*

## Abstract

The use of fossil fuels has direct environmental impacts, including the use of jet fuel. An alternative to jet fuel is biojet fuel, which can be produced from biomass. Recently the ATJ (alcohol to jet) process has been promoted as an alternative that may offer better biojet production performance compared to some other process routes. Although there are several proposals for jet fuel production using the ATJ route, no intensified alternative has been reported to increase the performance, and decrease the economic and environmental impact. In this proposal, an intensified ATJ production process is presented, considering the intensification of the pretreatment stage using a saccharification-fermentation reactor, ethanol purification considering a divided wall column, and a distillation column with two reaction zones for biojet fuel production. Once the intensified proposal was evaluated in a multi-objective optimization framework, a saving of 18% in the environmental impact was obtained, and a saving of 12% in the environmental impact.

**Keywords:** biojet production, ATJ process, process intensification.

## 1. Introduction

Despite the COVID-19 pandemic's reduced jet fuel demand, growth in the services provided by the industry is anticipated once more. This poses a difficulty, as does the fuel's intrinsic environmental impact throughout its life cycle, in an economic climate that is still dependent on finite oil. Given this, using biojet fuel as a renewable aviation fuel is a possible alternative. The production of biojet can be done by several routes, for example, from the conversion of oils. There are some other routes such as hydroprocessing of fatty acids and esters (HEFA), catalytic hydro thermolysis (CH) and hydroprocessing to depolymerized cellulosic biojet (HDCJ). These processes are not very feasible despite their low-cost processing. There are other routes to obtain it, for example, starting from alcohol (Alcohol-to-jet, ATJ) that was recently certified in April 2016 by ASTM. The ATJ process consists of 4 stages starting from ethanol, i) dehydration, ii) oligomerization, iii) hydrogenation, and iv) purification. The ATJ route has bioethanol as a raw material, which has lignocellulosic agroindustrial waste as a precursor raw material that does not require additional water or a long time to obtain it. However, one of the disadvantages is that it is a relatively expensive process and is still not sufficiently competitive in the market, compared to the conventional jet fuel production process. The process of obtaining biojet fuel has been previously studied by several authors; however, it is a process with a wide field for improvement.

Process intensification, on the other hand, is an improvement strategy that can involve the reduction of equipment size, the development of several operations in a single piece of equipment, the reduction of process times, etc. An example of an intensified process is the reactive distillation column, where a chemical reaction and a separation process based on liquid-vapor equilibrium can be carried out. Taking into consideration that the process of obtaining biojet fuel has already been previously studied and applied through the ATJ route; however, it is a process with a wide field of improvements where process intensification strategies can be implemented. Thus, the goal of this work was to design an intensified process to produce biojet fuel from Mexican lignocellulosic biomass (corn stover, and cane bagasse) having alcohols as intermediates. The process was designed and modeled considering an intensified process for pretreatment/hydrolysis/fermentation/purification for the biomass-ethanol process.

## 2. Description of the intensified process

The creation of intermediate alcohols is necessary for the whole process of producing biojet fuel from lignocellulosic biomass. The general process for producing alcohols from lignocellulosic agricultural waste includes the following steps: (1) biomass pretreatment (2) enzymatic hydrolysis (3) fermentation of sugars (4) and purification of the alcohols produced. In this manner, Figure 1 shows how the conventional turns into an intensified process in some sections of the ATJ process.

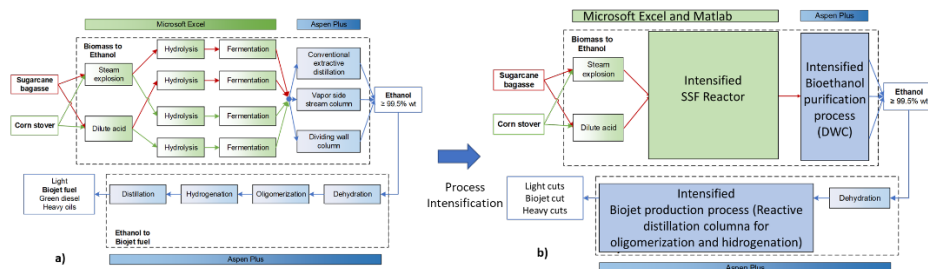


Figure 1. Description of conventional (a) and intensified (b) biojet fuel production.

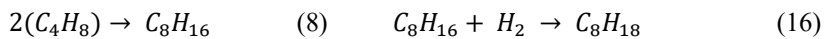
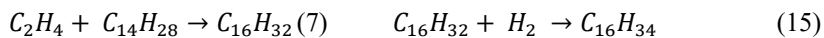
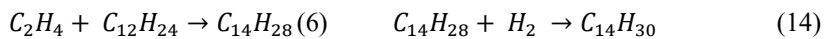
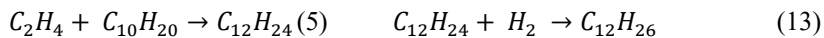
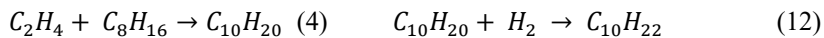
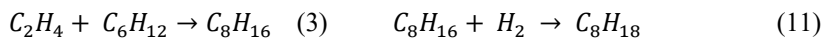
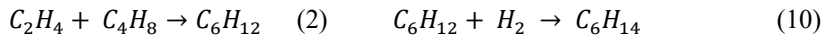
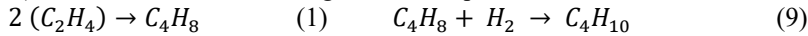
The main difference between the conventional process and the intensified process is the use of the intensified fermentation-saccharification (SSF) reactor instead of separate process steps. Additionally, an intensified process is used for bioethanol separation. Finally, in the final part of the process, the oligomerization, hydrogenation, and purification stages are carried out simultaneously using a double-reactive distillation column.

## 3. Case study

According to SENER (2017), Mexico needs to replace at least 5.5% of its current demand for conventional jet fuel with biojet fuel, or 258 million liters, to develop a viable biojet fuel market. Two pretreatment alternatives, dilute acid, and steam explosion, are considered in this proposal. For the hydrolysis and fermentation section, an intensified saccharification-fermentation (SSF) reactor will be considered. Once an effluent of ethanol, water, etc. has been produced from the reactor, it will be purified in an intensified separation scheme (divided wall column). The alcohol is dehydrated in a dehydration reactor to obtain ethylene. Finally, starting from ethylene, the biojet fuel stream is obtained.

#### 4. Methodology and Process Modeling

The two feedstocks and pretreatments were combined to create the biomass-to-ethanol process, which was created using the superstructure scheme. Kinetic data were gathered from the literature to describe the pretreatment. The modeling of the intensified saccharification-fermentation reactor was carried out in the Matlab computational package considering the kinetics reported by Kadam et al. 2004. The *S. cerevisiae* strain was considered for fermentation. Aspen Plus was used to simulate the ethanol purification zone and the module ethanol-to-biojet fuel using the NRTL and ENRTL thermodynamic models, respectively. A column sequence with a vapor side stream and a column sequence of dividing wall (DWC) was proposed as enhanced methods because the expense of separation represents a well-known area of opportunity. For the dehydration reaction, a performance reactor as reported by Rivas-Interian (2022) was considered. For the replacement of the oligomerization and hydrogenation process by a reactive double distillation column, the approach previously proposed by Heydenrych et al. was considered (Heydenrych et al., 2001). The complexity of the system (alkenes starting from C2) is reflected in the following chemical equations.



The estimated Kinect parameters of power law are shown in Table 1.

Table 1. Estimated kinetic parameters of the power-law model in equations 1-16

Oligomerization-Hydrogenation production					
Reaction Number	k	E (kJ/kmol)	Reaction Number	k	E (kJ/kmol)
1 (C2)	38728.5907	4.623	9 (C2)	6146.3487	5.1584
2 (C4)	74825.7294	14.934	10 (C4)	7150.2255	2.4686
3 (C6)	48935.42	164.8806	11 (C6)	6409.9637	5.5759
4 (C8)	110614.018	16.2112	12 (C8)	2465.1438	2.8607
5 (C10)	83400.8644	17.5098	13 (C10)	6052.2691	7.6072
6 (C12)	116517.955	21.4748	14 (C12)	4027.1747	7.9935
7 (C14)	27140.8737	19.7567	15 (C14)	1723.794	6.935
8 (C16)	3349.1216	158.5374	16 (C16)	746.1161	6.0205

## 5. Process Optimization

An optimization effort is implied by the quest for the design and operation parameters that maximize savings and minimize the environmental effect. The following describes these goals as well as the optimization method.

### 5.1. Objective functions

The total annual cost (TAC) and eco-indicator-99 (EI99) were chosen as the objective functions to evaluate the process' sustainability. Equations 17 and 18 were used to calculate them, respectively:

$$TAC = \frac{\text{Capital cost}}{\text{Recovery time}} + \text{Operation cost} \quad (17) \quad EI99 = \sum_b \sum_d \sum_{k \in K} \delta_d \omega_d \beta_b \alpha_{b,k} \quad (18)$$

Where  $\beta_b$  is the total amount of chemical  $b$  released per unit of reference flow due to direct emissions,  $\alpha_{b,k}$  is the damage caused in category  $k$  per unit of chemical  $b$  released to the environment,  $\omega_d$  is a weighting factor for damage in categories  $d$ , and  $\delta_d$  is the normalization factor for damage of category  $d$ . Guthrie's approach was used to compute the capital cost. The annual production of alcohol was also factored into the ethanol production process to meet Mexico's demand for biojet fuel, which accounts for 5.5% of the country's demand for conventional jet fuel (SENER, 2017). Following is a description of the objective function:

$$\begin{aligned} \text{Min}(TAC, EI99, IR) &= f(X) \\ \text{Subject to } x_m^{\vec{}} &> y_m^{\vec{}} \end{aligned} \quad (19)$$

Being the variables involved in the whole modeling process can serve as degrees of freedom of the process. The Differential Evolution with Tabu List (DETL) stochastic approach was used to optimize the case of research. The best design parameters and the best design were chosen using this methodology. It was carried out on a hybrid platform that combined Matlab, Aspen Plus, and Microsoft Excel. The following criteria were employed: 120 individuals, 1000 generations, 50% of the entire population on the tabu list, 0.0001 tabu radius, and probability of crossover and mutation of 0.9 and 0.3, respectively.

## 6. Results

In the following section, the results obtained will be shown. The results will be shown in three sections: general pretreatment section, ethanol purification, and biojet production.

### 6.1.1. Pretreatment section

As mentioned in the methodology section, two scenarios were analyzed considering two types of feedstock. As shown in Figure 2, the use of an intensified strategy (SSF reactor) considerably reduces costs compared to a conventional process in both pretreatments and with the two feedstocks considered. Figure 2 shows the optimization results of all costs associated with the entire reaction section for both processes (conventional and intensified) when considering various pretreatments. Regarding the ethanol purification process, a direct comparison with the conventional process was not performed, since in the previous work of Rivas-Interian et al. (2022) a substantial saving of the intensified process versus the conventional one was observed. Thus, Figure 3a) shows the Pareto front of the optimization results schemes of all alternatives that accomplish recovery and purity restrictions considering economic and environmental objectives for ethanol purification. On the other hand, Figure 3b) shows similar information for the biojet fuel

section after being fed with ethane to be converted by oligomerization and hydrogenation.

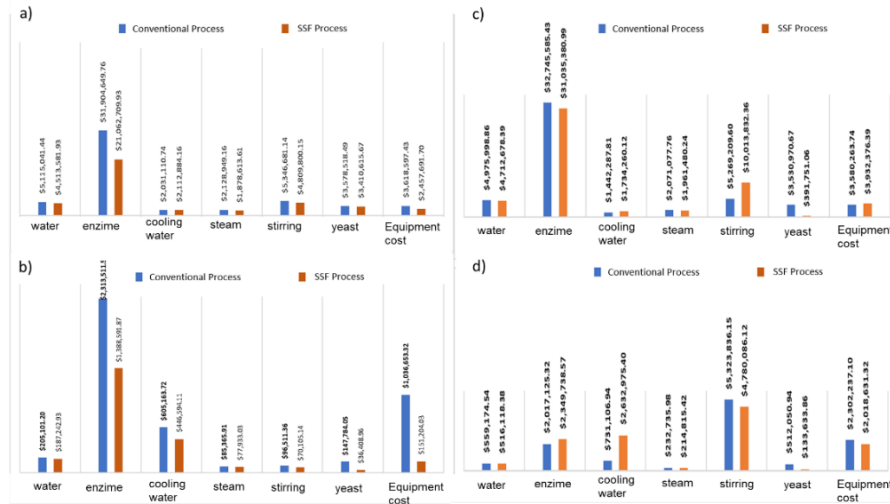


Figure 2. Comparison of pretreatment processes cost considering steam explosion a) sugarcane bagasse, b) corn stover; and dilute acid c) sugarcane bagasse, d) corn stover; and dilute acid.

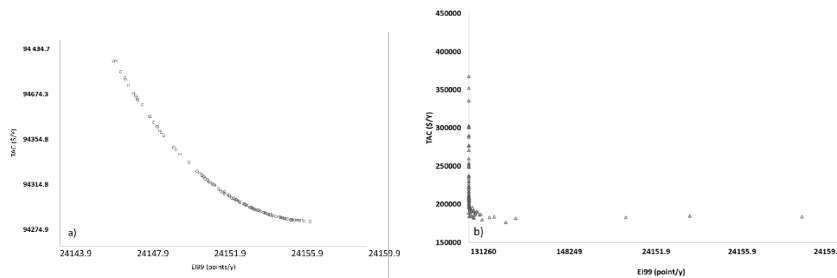


Figure 3. Pareto front for the ethanol purification zone considering a split-wall column (DWC) (a), and for the biojet production zone (b).

Figure 3a) shows that a TAC and EI99, weighted by the number of kilograms produced 1.18USD/kg ethanol and 0.3015 ecopoints/kg ethanol, respectively, were obtained. These values represent a reduction of 10% for the TAC, and close to 40% in the EI99 concerning the conventional design that reports TAC values of 1.29USD/kg ethanol and EI99 is 0.4716 ecopoints/kg ethanol. Finally, Figure 3b) shows the Pareto obtained from the transformation of ethanol to biojet fuel considering the dehydration, oligomerization, and hydrogenation stages. An important aspect to consider for savings in all stages is the reduced use of smaller equipment. In the case of the pretreatment and fermentation section, the use of equipment was reduced, the purification section was carried out in a piece of single process equipment, and in the biojet production zone, the double reactive distillation column performed the work of two reactors and a separation column. These capital cost savings are immediately reflected in the overall cost of the entire process. Table 2 shows the design parameters for the SSF reactor section as well as the double reactive column for biojet production. Data for the ethanol separation section can be found in the work of Rivas-Interian et al. (2022).

Table 2. Design parameter for SSF reactor, and reactive distillation column.

Scheme	Residence time (h)	Volume (m <sup>3</sup> )	Enzyme (FPU/g cellulose)	Yeast (FPU/g cellulose)	Stirring (RPM)	Oligomerization-Hydrogenation Reactive Column	
						Number of stages	Reflux ratio
Sugar Cane/Steam	29.63	20492	9.013	0.9986	150.8353	Feed stage	24 and 30
						Side stream stage	33
Corn/Steam	29.6492	6633	18.75	0.9952	200.81	Reactive stages	12-20, 21-29
						Hold Up (l)	3.5, 48
Sugar cane/acid	32.82	62528	10.008	1.0031	150.03	Distillate flowrate (kg h <sup>-1</sup> )	5422.44
						Condenser duty (kcal h <sup>-1</sup> )	9.56216 x 10 <sup>6</sup>
Corn/acid	25.034	16455	10.7791	1.0059	18086	Reboiler duty (kcal h <sup>-1</sup> )	854153
						Jet Fuel Production (kg/h)	19935.821

The synergistic effect generated by joining the oligomerization and hydrogenation processes in a column with two reactive zones is quite good. During the parametric optimization process, it was possible to obtain design conditions that allowed obtaining a significant amount of product with better operating conditions.

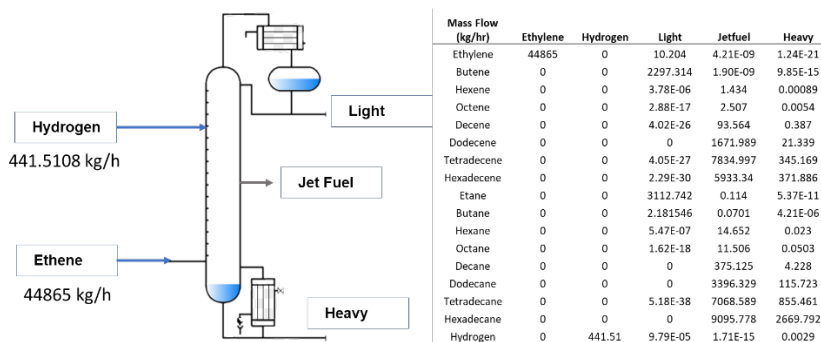


Figure 4. Mass balance for the reactive distillation column.

## Conclusions

In this work, an intensified approach to the production of biojet from biomass was presented. Process intensification strategies were applied in the pretreatment section, ethanol purification, as well as in biojet production area. Once the intensified alternatives were implemented, an improvement was observed in the economic and environmental indicators that were evaluated. In general terms, an improvement of 18% was observed in the economic indicator, and 12% in the environmental indicator when comparing the conventional and the intensified strategy.

## References

- SENER. Mapa de Ruta Tecnológica Bioturbosina, 2017. [https://www.gob.mx/cms/uploads/attachment/file/324219/MRT\\_Bioturbosina\\_Final.pdf](https://www.gob.mx/cms/uploads/attachment/file/324219/MRT_Bioturbosina_Final.pdf) (accessed April 07, 2020). Y. Brown, Year, Article or Chapter Title, etc.
- Heydenrych, M. D., et al. (2001). Oligomerization of ethene in a slurry reactor using a nickel(II)-exchanged silica-alumina catalyst. *Journal of Catalysis*, 197(1), 49–57.
- Kadam, K., et al. (2004). Development and Validation of a Kinetic Model for Enzymatic Saccharification of Lignocellulosic Biomass. *Biotechnology Process*, 20(3), pp.698- 705.
- Rivas-Interianet al. (2022). Economic and Environmental Optimization and Feedstock Planning for the Renewable Jet Production Using an Intensified Process. In *Computer Aided Chemical Engineering* (Vol. 51, pp. 643-648). Elsevier.

# Metrics for Evaluating Machine Learning Models Prediction Accuracy and Uncertainty

Yushi Deng, Mario R. Eden, Selen Cremaschi

*Department of Chemical Engineering, Auburn University, Auburn, AL 36830, USA  
selen-cremaschi@auburn.edu*

## Abstract

Models with stochastic outputs are commonly employed to make decisions for design and operation of chemical processes. The accuracy and precision of the stochastic outputs, i.e., model predictions, should be compared to the observation based on the design and operation requirements. The area metric measures the overall mismatch between prediction and observation but does not provide precision and accuracy information separately. In this paper, we proposed a new metric called uncertainty width that decomposes area metrics into precision and bias components. This new metric enables the consideration of precision and accuracy simultaneously for model evaluation. We applied this metric to evaluate and compare the performance of three models fitted using three liquid entrainment fraction measurement datasets with different flow orientations, using Gaussian Process (GP) and Bayesian Neural Network (BNN). The results reveal that the overall prediction accuracy of GP model (GPM) is better than that of BNN model. The overestimation of the GPM prediction precision causes the high average area metric.

**Keywords:** Metrics, Machine Learning, Area Metrics, Accuracy, Precision, Uncertainty

## 1. Introduction

Design and operations decisions for chemical processes are regularly made considering various uncertainties (Sharifian et al., 2021). Models that provide predictions and their uncertainty over the target domain are essential in the risk assessment of these decisions. Before applying these decisions in chemical engineering analysis and design, the model performance evaluation must be performed. For models with stochastic outputs, the model performance is evaluated by measuring the mismatch between the model output and the true observation. Various metrics exist for this evaluation.

In classical hypothesis testing, a test is constructed under the hypothesis that the observations come from the predicted population. Decisions about rejecting or not rejecting the hypothesis can be made according to whether the test statistic falls within the critical region (Liu et al., 2011). Common hypotheses include the comparison of several moments (e.g., t-test, F-test) and comparison of full statistical distributions (e.g., Kolmogorov-Smirnov test (Haldar and Mahadevan, 2000), Anderson-Darling test (Rebba and Mahadevan, 2008)). The Bayes factor, expressed as the ratio of the relative likelihood of the null hypothesis over the alternative hypothesis, is used as the criteria (Mahadevan and Rebba, 2005). The null hypothesis is not rejected when the Bayes factor is larger than one. The Frequentist Metric evaluates the distance between the estimated mean of the observations and the mean of the predictions (Oberkampf and Trucano, 2002). The confidence bound is generated from the uncertainty of the distance.

The area metric (AM) measures the disagreement area of the cumulative distributions between the prediction and the observation (Ferson and Oberkampf, 2009). The u-pooling



method derived from the area metric is proposed to measure the global dispersion of the distributions of observation and prediction. The measures of the cumulative distribution function at each validation site, called the u-value, are aggregated over the entire input space. The evaluation of the overall disagreement between the observation and model prediction is expressed as the area difference between the u-value distribution and the uniform distribution (Ferson et al., 2008). The model acceptance is determined according to a predefined accuracy requirement for the frequentist and area metrics.

For the area metric, the mismatch between the stochastic observation and prediction can be decomposed into precision and accuracy. The precision refers to the variance of multiple predictions at one validation site. Accuracy measures the degree of closeness between the prediction and the observation (Eisenhart, 1969). Generally, models with high prediction accuracy (i.e., low average prediction bias) and precision (i.e., narrow average uncertainty) are good performers. The area metric measures the sum of accuracy and precision. For predictions with high area metric values, it is difficult to determine whether this is due to low accuracy or poor precision. The accuracy and precision should be quantified separately when the decisions depend on both accuracy and precision.

This study proposes uncertainty width to evaluate the prediction precision within the area metric. By separating the area metric into the prediction bias related to accuracy and the uncertainty width related to precision, the performance between various models can be compared, and a model can be selected according to the acceptance criterion. Large bias may be acceptable with narrow uncertainty width when the model is employed to provide a prediction range. High uncertainty may be tolerated with a low bias if a point prediction is adopted for solving the problem. The proposed metric is applied to evaluate the performance of two machine learning models built using Gaussian Process Modeling (GPM) and Bayesian Neural Network (BNN).

The uncertainty width is defined and discussed in Section 2. Section 3 presents the details of the application of the uncertainty width for model performance evaluation. The results are given and discussed in Section 4. Finally, Section 5 summarizes the conclusions and future directions.

## 2. Methodology

A graphical illustration of the area metric is given in Figure 1. For one sample, the cumulative distribution functions (CDFs) of the model prediction and true observation are plotted as  $f_1(y)$  and  $f_2(y)$ . According to the definition, the area metric is the area between the two curves. For symmetric distributions, the means of the two curves,  $\mu_1$  and  $\mu_2$ , are located at the cumulative probability value equal to 0.5. The bias of the prediction  $\mu_2 - \mu_1$ , defined as the difference between the two means, is shown as the rectangular area between  $\mu_1$  and  $\mu_2$ . Assume the two curves intersect at  $y = \mu_I$ .

The area metric is defined as the integral of the difference in area between the two CDFs and its analytical expression is given in Equation 1.

$$\int_{-\infty}^{\infty} |f_1(y) - f_2(y)| dy = \int_{-\infty}^{\mu_I} [f_2(y) - f_1(y)] dy + \int_{\mu_I}^{\infty} [f_1(y) - f_2(y)] dy \quad (1)$$

The bias between the prediction and observation mean is equal to the integral of the cumulative probability  $\int_{\mu_1}^{\mu_2} 1 dx$ . By shifting the  $f_1(y)$  to the left by  $2(\mu_2 - \mu_1)$  and shifting  $f_2(y)$  to the right by  $2(\mu_2 - \mu_1)$ , the bias can be cut into several pieces. The integrals of the resulting pieces are given in Equation 2.

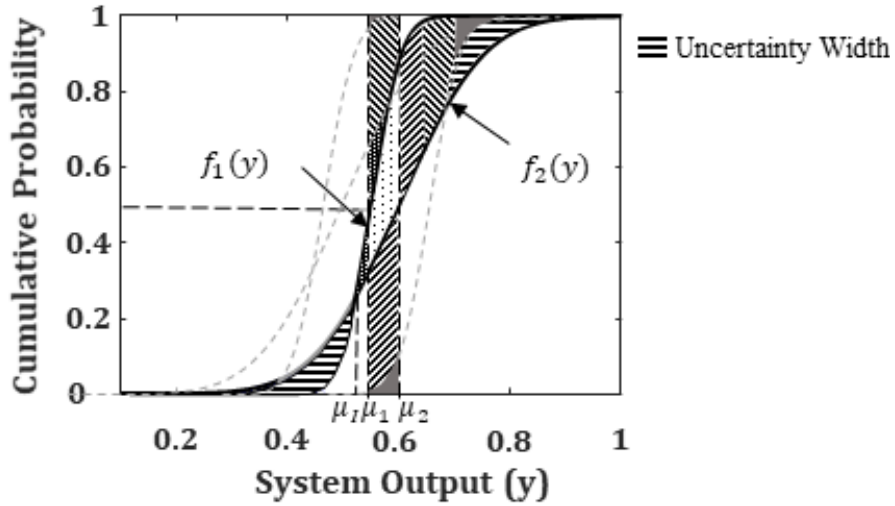


Figure 1 Area Metric Decomposition

$$\begin{aligned}
 \int_{\mu_1}^{\mu_2} 1 dy &= \int_{\mu_1}^{\mu_2} [1 - f_1(y + 2(\mu_2 - \mu_1))] dy \\
 &+ \int_{\mu_1}^{2\mu_1 - \mu_1} [f_2(y + 2(\mu_2 - \mu_1)) - f_1(y + 2(\mu_2 - \mu_1))] dy \\
 &+ \int_{2\mu_1 - \mu_1}^{\mu_2} [f_1(y + 2(\mu_2 - \mu_1)) - f_1(y)] dy \\
 &+ \int_{\mu_1}^{\mu_2} [f_2(y + 2(\mu_2 - \mu_1)) - f_1(y)] dy + \int_{\mu_1}^{\mu_2} [f_1(y) - f_2(y)] dy \\
 &+ \int_{\mu_1}^{\mu_2} [f_2(y) - f_1(y - 2(\mu_2 - \mu_1))] dy \\
 &+ \int_{\mu_1}^{\mu_2} [f_1(y - 2(\mu_2 - \mu_1))] dy
 \end{aligned} \tag{2}$$

The second term in area metric shown in Equation 1 can also be decomposed into integrals given in Equation 3.

$$\begin{aligned}
 \int_{\mu_1}^{\infty} [f_1(y) - f_2(y)] dy &= \int_{\mu_1}^{\mu_1} [f_1(y) - f_2(y)] dy + \int_{\mu_1}^{\mu_2} [f_1(y) - f_2(y)] dy \\
 &+ \int_{\mu_2}^{2\mu_2 - \mu_1} [f_1(y) - f_2(y)] dy + \int_{2\mu_2 - \mu_1}^{2\mu_2 - \mu_1} [f_1(y) - f_2(y)] dy \\
 &+ \int_{2\mu_2 - \mu_1}^{3\mu_2 - 2\mu_1} [f_1(y) - f_1(y - 2(\mu_2 - \mu_1))] dy \\
 &+ \int_{3\mu_2 - 2\mu_1}^{5\mu_2 - 4\mu_1} [f_1(y) - f_1(y - 2(\mu_2 - \mu_1))] dy \\
 &+ \int_{5\mu_2 - 4\mu_1}^{\infty} [f_1(y - 2(\mu_2 - \mu_1)) - f_2(y)] dy \\
 &+ \int_{5\mu_2 - 4\mu_1}^{\infty} [f_1(y) - f_2(y)] dy
 \end{aligned} \tag{3}$$

Because  $f_2(y)$  is symmetric at  $(\mu_2, 0.5)$  and  $f_2(y + 2(\mu_2 - \mu_1))$  is shifted from  $f_2(y)$ , the functions  $f_2(y + 2(\mu_2 - \mu_1))$  and  $f_2(y)$  are symmetric at  $(\mu_1, 0.5)$ . It can also be inferred that the functions  $f_1(y)$  and  $f_1(y - 2(\mu_2 - \mu_1))$  are symmetric at  $(\mu_2, 0.5)$ . As such,  $\int_{\mu_1}^{2\mu_1 - \mu_1} [f_2(y + 2(\mu_2 - \mu_1)) - f_1(y + 2(\mu_2 - \mu_1))] dy = \int_{\mu_1}^{\mu_1} [f_1(y) - f_2(y)] dy$ . Other equal integrals of Equation 2 and Equation 3 are given in Equations 4 to 7.

$$\int_{3\mu_2-2\mu_1}^{5\mu_2-4\mu_1} [f_1(y) - f_1(y - 2(\mu_2 - \mu_1))] dy$$

$$= \int_{\mu_1}^{\mu_2} [1 - f_1(y + 2(\mu_2 - \mu_1))] dy + \int_{\mu_1}^{\mu_2} [f_1(y - 2(\mu_2 - \mu_1))] dy \quad (4)$$

$$\int_{\mu_1}^{2\mu_1-\mu_1} [f_1(y + 2(\mu_2 - \mu_1)) - f_2(y + 2(\mu_2 - \mu_1))] dy$$

$$= \int_{2\mu_2-\mu_1}^{2\mu_2-\mu_1} [f_1(y) - f_2(y)] dy$$

$$+ \int_{2\mu_2-\mu_1}^{3\mu_2-2\mu_1} [f_1(y) - f_1(y - 2(\mu_2 - \mu_1))] dy \quad (5)$$

$$\int_{\mu_1}^{2\mu_1-\mu_1} [f_2(y + 2(\mu_2 - \mu_1)) - f_1(y + 2(\mu_2 - \mu_1))] dy$$

$$= \int_{\mu_1}^{\mu_1} [f_1(y) - f_2(y)] dy \quad (6)$$

$$\int_{\mu_1}^{\mu_2} [f_2(y) - f_1(y - 2(\mu_2 - \mu_1))] dy = \int_{\mu_2}^{2\mu_2-\mu_1} [f_1(y) - f_2(y)] dy \quad (7)$$

For visualization, the equal integral pairs are shaded with the same patterns in Figure 1. The uncertainty width is defined as the difference between the area metric  $\int_{-\infty}^{\infty} |f_1(y) - f_2(y)| dy$  and the prediction bias  $\int_{\mu_1}^{\mu_2} 1 dy$ . If Equations 4 to 7 are inserted into the uncertainty width definition, Equation 8 is obtained to calculate the uncertainty width.

*Uncertainty width*

$$= \int_{-\infty}^{\infty} |f_1(y) - f_2(y)| dy - \int_{\mu_1}^{\mu_2} 1 dy$$

$$= \int_{-\infty}^{\mu_1} [f_2(y) - f_1(y)] dy + \int_{3\mu_2-2\mu_1}^{5\mu_2-4\mu_1} [f_1(y - 2(\mu_2 - \mu_1)) - f_2(y)] dy$$

$$+ \int_{5\mu_2-4\mu_1}^{\infty} [f_1(y) - f_2(y)] dy \quad (8)$$

### 3. Case Study

#### 3.1. Machine Learning Models

Two Bayesian-based machine learning techniques, GPM and BNN, are adopted to build models in this study. GPM is a supervised learning method with a theoretical basis in statistics. A GPM is characterized by its mean function and covariance function or kernel function (Williams and Rasmussen, 2006). A constant mean function and square exponential kernel function with different length scales for each input dimension are utilized. The GP prior is defined in Equation 9:

$$f \sim \mathcal{GP}(m, k)$$

$$m(x) = a, \text{ and } k(x_p, x_q) = \sigma_f^2 \exp\left(-\sum_{h=1}^d \frac{(x_p - x_q)^2}{2l_h^2}\right) + \sigma_n^2 \Delta_{pq} \quad (9)$$

In Equation 9,  $f$  represents the underlying function value at the input  $x$ ,  $m(x)$  is the mean function of the GP, and  $k(x_p, x_q)$  is the covariance function representing the spatial covariance between any two points ( $x_p$  and  $x_q$ ) at the process. In the covariance function,  $d$  is the dimension of input  $x$ ,  $l_h$  represents the characteristic length scale corresponding to the  $h^{\text{th}}$  dimension of input  $x$ ,  $\sigma_n^2$  is the output variance, and the parameter  $\sigma_f^2$  is output-scale amplitude.  $\Delta_{pq}$  is Kronecker delta, which is one if  $p = q$  and zero otherwise.

BNN applies Bayesian statistics to Neural Networks (NNs) (Goan and Fookes, 2020). Instead of treating the model weights as a fixed value, BNN assumes the weights are

random variables. After specifying the prior distribution of weights and likelihood, the Bayes theorem is applied to calculate the posterior distribution. BNN model parameters are optimized using variational learning that minimizes the Kullback-Leibler (KL) divergence to avoid intractable computations (Blundell et al., 2015). The prediction mean and variance are calculated using Monte Carlo sampling.

### 3.2. Model Performance Evaluation using Uncertainty Width

The uncertainty width is applied to evaluate and compare the performance of the selected machine learning (ML) models from both accuracy and precision perspectives. Three datasets containing liquid entrainment fraction measurements and two machine learning techniques, GPM and BNN, are used to build six ML models. The three datasets for vertical, horizontal, and inclined flow orientations are composed of 1,083, 478, and 100 samples with 9 input features representing the experimental operating conditions. For each sample, the liquid entrainment fraction prediction with its uncertainty is calculated and evaluated using the area metric. The area metric is then decomposed into accuracy (absolute bias) and precision (uncertainty width). The overall model performance is evaluated using the average area metric, average absolute bias (mean absolute error), and average uncertainty width of all the samples.

## 4. Results and Discussion

The model performance evaluation results are shown in Figure 2. The average area metric, mean absolute error (MAE), and the average uncertainty width of the models built using experimental datasets for vertical, horizontal, and inclined flow orientations are shown in Figures 2 (a), (b), and (c), respectively. The total height of each bar represents the average area metric. The MAE and the average uncertainty width values are separated and marked using different colors and shades within the area metric. While the average area metric shows the overall performance of the model, models with low MAE have better accuracy, while models with low uncertainty have better precision.

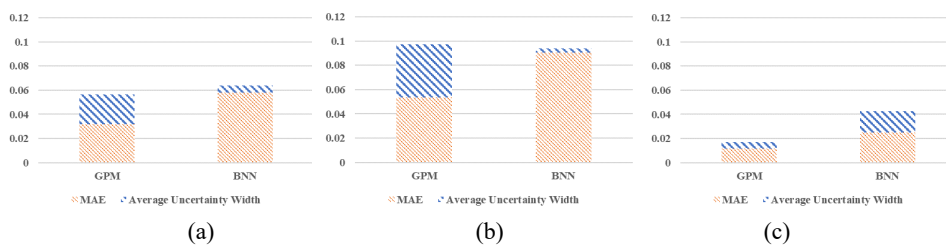


Figure 2 Model Performance of the models built using liquid entrainment fraction measurements for (a) vertical (b) horizontal (c) inclined flow orientations

Considering the overall model performance evaluated using the average area metric, the GPM and BNN fitted using datasets for vertical and horizontal flow orientations have identical performance. This observation suggests that when an entrainment fraction prediction range is used for developing a solution from the model output, the GPM and BNN performances are close. The GPM outperforms the BNN when fitted using the dataset for inclined flow orientation. For the model accuracy, the MAE of GPM, is lower than that of BNN for all three pairs of models. This observation indicates that the accuracy of the GPM is always better than the BNN. The GPM should be selected when entrainment fraction point prediction is employed to make a decision. For models built using the dataset for inclined flow orientation, both MAE and average uncertainty width

of BNN are higher than those of GPM, indicating that the GPM outperforms BNN in both accuracy and precision.

## 5. Conclusions

In this study, a metric called uncertainty width is developed to evaluate the performance of machine learning models with stochastic outputs. The prediction performance at each sample is separated into accuracy and precision by decomposing the area metric into absolute bias and uncertainty width. The average absolute error and average uncertainty bias from all samples are used to simultaneously evaluate the overall model performance, considering accuracy and precision. The new metric is applied to evaluate the performances of six models built using GPM and BNN using three datasets containing liquid entrainment fraction measurements for vertical, horizontal, and inclined flow orientations. The results suggest that GPM performs better than BNN, considering both accuracy and precision using the dataset for the inclined flow orientation. For vertical and horizontal flow orientation datasets, although the overall model performance evaluated using the average area metric are identical, the accuracy of GPM is better than that of BNN. The GPM performs better when an entrainment fraction point prediction is needed for making design and operation decisions using these models. The high area metric of GPM built using vertical and horizontal flow orientation datasets comes from the high uncertainty width, which means that GPM overestimates the prediction uncertainty. Future work will investigate the application of uncertainty width to outputs with skewed distributions.

## References

- Blundell, C., Cornebise, J., Kavukcuoglu, K., Wierstra, D., 2015. Weight uncertainty in neural networks. 32nd Int. Conf. Mach. Learn. ICML 2015 2, 1613–1622.
- Eisenhart, C., 1969. Realistic Evaluation of the Precision and Accuracy. NBS Spec. Publ. 21.
- Ferson, S., Oberkampf, W.L., 2009. Validation of imprecise probability models. *Int. J. Reliab. Saf.* 3, 3–22.
- Ferson, S., Oberkampf, W.L., Ginzburg, L., 2008. Model validation and predictive capability for the thermal challenge problem 197, 2408–2430. <https://doi.org/10.1016/j.cma.2007.07.030>
- Goan, E., Fookes, C., 2020. Bayesian Neural Networks: An Introduction and Survey. *Lect. Notes Math.* 2259, 45–87. [https://doi.org/10.1007/978-3-030-42553-1\\_3](https://doi.org/10.1007/978-3-030-42553-1_3)
- Haldar, A., Mahadevan, S., 2000. Probability, reliability, and statistical methods in engineering design. John Wiley & Sons Incorporated.
- Liu, Y., Chen, W., Arendt, P., Huang, H.Z., 2011. Toward a better understanding of model validation metrics. *J. Mech. Des. Trans. ASME* 133, 1–13. <https://doi.org/10.1115/1.4004223>
- Mahadevan, S., Rebba, R., 2005. Validation of reliability computational models using Bayes networks. *Reliab. Eng. Syst. Saf.* 87, 223–232.
- Oberkampf, W.L., Trucano, T.G., 2002. Verification and validation in computational fluid dynamics. *Prog. Aerosp. Sci.* 38, 209–272.
- Rebba, R., Mahadevan, S., 2008. Computational methods for model reliability assessment. *Reliab. Eng. Syst. Saf.* 93, 1197–1207.
- Sharifian, S., Sotudeh-Gharebagh, R., Zarghami, R., Tanguy, P., Mostoufi, N., 2021. Uncertainty in chemical process systems engineering: a critical review. *Reviews in Chemical Engineering* 37, 687–714. <https://doi.org/doi:10.1515/revce-2018-0067>
- Williams, C.K.I., Rasmussen, C.E., 2006. Gaussian Processes for Machine Learning. MIT press Cambridge, MA.

# Optimization of Process Families for Deployment of Carbon Capture Processes using Machine Learning Surrogates

Georgia Stinchfield<sup>a</sup>, Bashar L. Ammari<sup>a</sup>, Joshua C. Morgan<sup>c,d</sup>, John D. Siirola<sup>b</sup>, Miguel Zamarripa<sup>c,d</sup>, Carl D. Laird<sup>a\*</sup>

<sup>a</sup>*Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA, 15232*

<sup>b</sup>*Sandia National Laboratories, P.O. Box 5800, Albuquerque, NM, 87185*

<sup>c</sup>*National Energy Technology Laboratory (NETL), 626 Cochran Mill Road, Pittsburgh, PA, 15236*

<sup>d</sup>*NETL Support Contractor, 626 Cochran Mill Road, Pittsburgh, PA, 15236*

\**claird@andrew.cmu.edu*

## Abstract

Traditional process design approaches focus on exploiting economies of scale but are inefficient when designing a large number of similar processes for decentralized applications. By optimizing each process individually, these approaches do not allow for manufacturing standardization or make use of economies of numbers. In this work, we design a family of processes (i.e., multiple processes) of a carbon capture facility with various performance requirements. We identify sub-components within the process and create a small set of sub-component designs that can be shared across all the processes in the process family. We formulate this optimization problem as a nonlinear generalized disjunctive program (GDP) and, in previous work, developed two approaches for reformulating and solving this problem: one based on full-discretization of the design space (Zhang et al., 2022) and one that used Machine Learning (ML) surrogates to replace the nonlinear process models (Stinchfield et al., 2022). Using ML surrogates to predict required system costs and performance allows us to reformulate the nonlinearities in the GDP to generate an efficient MILP formulation. In this work, we apply the ML surrogate approach to design a family of carbon capture systems to cover a set of different flue gas flow rates and inlet CO<sub>2</sub> concentrations, where we consider the absorber and stripper as sub-component types.

**Keywords:** Manufacturing, Optimization, Energy Systems, Machine Learning.

## 1. Introduction

Effectively mitigating climate change requires the broad deployment of green energy and industrial decarbonization processes. When designing *multiple* processes, a conventional design approach produces a set of unique unit designs for each process that can be expensive to manufacture. Modularity, a well-explored design approach, provides significant reductions in manufacturing costs and has proven beneficial in particular when designing multiple instances of a process (Baldea et al., 2017). However, modular designs are typically ‘numbered-up’ to achieve capacity requirements and often fail to fully exploit economies of scale. Rather than using traditional or modular design, we propose a hybrid approach that applies concepts from product family design to the design of

process systems to gain the manufacturing benefits of modularity while still obtaining process customization and economies of scale.

In product family design (PFD), a product family is broadly defined as “a set of products that share one or more common ‘element(s)’ yet target a variety of different market segments” (Simpson et al., 2014). This approach is used extensively in automotive industries, for example, where a wide variety of products are available, but designed and assembled using a smaller, commonly designed platform of sub-components. As process vendors tackle the broad deployment of large numbers of processes for industrial decarbonization, they face similar design goals. We are extending the ideas of PFD into this area of process design. In our work, the “products” are the set of processes we must design. Each individual process is referred to as a process variant; each variant is associated with a set of conditions that the process design must meet. The “common ‘element(s)’” are sub-components shared among the process variants. For our case study, we consider the design of a carbon capture system (CCS). Typically, a CCS is designed for a specific point source capture location (i.e., coal-fired power plants, natural gas combined cycle, cement, refineries, etc.) mainly based on the flue gas conditions and flowrate. PFD has the potential to accelerate the deployment of CCS for industrial processes through the simultaneous design of process families that share common sub-components. Our case study considers 12 process variants. Each process variant in the family is designed for a particular flue gas flow rate and CO<sub>2</sub> concentration. We wish to design each of these process variants using a common platform that includes shared sub-component types, in this case, the absorber and the stripper. We require one of each sub-component for each process variant (i.e., 12 absorbers and 12 strippers in total) but hope to meet all the process variant requirements with a small number of sub-component designs (e.g., here we consider a platform with only two absorber designs and two stripper designs from which to design all 12 process variants).

This approach captures the benefits of manufacturing standardization, akin to modularity, since a large number of process variants share the same common sub-component designs. At the same time, this retains economies of scale as the platform of manufactured sub-components is designed simultaneously with the set of process variants in the family. This approach has proven beneficial in other industries; companies such as Nissan, Toyota, and Boeing have reported substantial cost and time savings by using this approach (Simpson et al., 2014). However, despite documentation of significant success in other industrial manufacturing settings, PFD has largely not been applied to chemical process design. Furthermore, reported approaches have largely been heuristic, and there is a need for rigorous optimization approaches.

Our optimization problem simultaneously determines the individual *designs* of the sub-components in the platform (i.e., if we only offer two absorber designs, what should they be) and *which design* each process variant should use (i.e., for each of 12 process variants, which of the two absorber designs should each variant use). We first formulated this problem as a nonlinear GDP. We solved this problem by reformulating it as a MILP formulation based on full discretization of the design space (Zhang et al., 2022). We extended this approach, proposing another MILP reformulation that avoids discretization by utilizing ReLU activated Neural Network (NN) as piecewise linear surrogate models (Stinchfield et al., 2022, Ceccon et al., 2022). This significantly decreased pre-computation requirements for simulations and allowed us to search for an approximation of the continuous design space. In this paper, we further demonstrate this piecewise linear

surrogate approach to optimize a process family of carbon capture systems, using both an ReLU activated NN and a linear model decision tree as surrogates. We design 12 carbon capture systems with different flue gas flow rates and CO<sub>2</sub> concentrations with the absorber and stripper columns as the shared sub-components.

## 2. Optimization Formulation

The following formulation for optimization of a process family design was first proposed by Stinchfield, et al., (2022). A process variant is represented by  $v$ , and is characterized by the design requirements for that variant. The parameter set  $V$  represents all process variants. The boundary conditions of a process variant are parameterized in  $b_v$ . In our case study,  $|V| = 12$  and the boundary conditions for each variant include the flue gas flow rate and CO<sub>2</sub> concentration. The total annualized cost of a process variant is represented by variable  $c_v$ . The objection function includes a weight for each process variant,  $w_v$ , that captures the expected sales or number of installations of each variant. The variable  $p_v$  captures the performance metrics of a particular process variant. In many cases, this will be an indicator function that determines whether a particular design is feasible for the variant. This ensures that an infeasible combination of sub-component designs is not selected for a process variant  $v$ . We use trained piecewise linear surrogate models, represented by the system of equations  $f^c$  and  $f^p$ , to predict the variable cost of the variant,  $c_v$ , and the performance indicator,  $p_v$ . Piecewise linear surrogates are used because they can be represented exactly as MILPs within the optimization formulation.

The set of common sub-component types considered in the platform is given by  $K$ . In our case study,  $K = \{\text{absorber, stripper}\}$ . The set  $J_k$  includes the sub-component designs for a particular sub-component of type  $k$ . The cardinality of  $J_k$  dictates the number of sub-component designs to consider in the platform for sub-component type  $k$  (i.e., if we want two absorber designs, then  $J_k = \{1,2\}$  where  $k = \text{absorber}$ ). The variable  $\hat{d}_{k,j}$  represents the design variable values for design  $j$  of sub-component type  $k$ , while variable  $d_{v,k}$  represents the design variable values of sub-component type  $k$  for process variant  $v$ . With this, if design  $j$  is chosen for sub-component type  $k$  in variant  $v$ , then  $d_{v,k} = \hat{d}_{k,j}$ .

$$\min. \sum_{v \in V} w_v c_v \quad (1a)$$

s.t.

$$c_v = f^c(b_v, d_{v,1}, \dots, d_{v,k}) \quad \forall v \in V \quad (1b)$$

$$p_v = f^p(b_v, d_{v,1}, \dots, d_{v,k}) \quad \forall v \in V \quad (1c)$$

$$\bigvee_{j \in J_k} \begin{bmatrix} Y_{v,k,j} \\ d_{v,k} = \hat{d}_{k,j} \end{bmatrix} \quad \forall v \in V, k \in K \quad (1d)$$

$$\hat{d}_{k,j}^L \leq \hat{d}_{k,j} \leq \hat{d}_{k,j}^U \quad \forall k \in K, j \in J_k \quad (1e)$$

$$p_v^L \leq p_v \leq p_v^U \quad \forall v \in V \quad (1f)$$

$$c_v \geq 0 \quad \forall v \in V \quad (1g)$$

$$Y_{v,k,j} \in \{\text{True, False}\} \quad \forall v \in V, k \in K, j \in J_k \quad (1h)$$



The objective, Eq. (1a), is to minimize the total weighted annualized cost of all process variants,  $v \in V$ . Piecewise linear surrogates predict the cost of each variant  $c_v$ , Eq. (1b), and the performance indicator  $p_v$ , Eq. (1c). Both surrogates are functions of the boundary conditions at a particular variant,  $b_v$ , and the design of each sub-component type  $k \in K$ . The disjunctions in Eq. (1d) select which of the sub-component designs,  $j \in J_k$ , for a sub-component of type  $k$  are selected for process variant  $v$ . The number of disjunctions is equal to  $|K| \times |V|$  because we make this decision for each sub-component  $k$  and each process variant  $v$ . The number of disjuncts corresponds to  $|J_k|$  for a particular sub-component of type  $k$ , because we decide on which sub-component design,  $j \in J_k$ , to select. Eq. (1e) bounds the design variable of sub-component type  $k$ , which also corresponds to the ranges used to train the surrogate. Eq. (1f) captures the performance constraints for each process variant.

### 3. Numerical Case Study

We demonstrate this formulation by designing a process family of 12 aqueous monoethanolamine (MEA) solvent-based carbon capture variants. Each variant requires a design that captures at least 90% of CO<sub>2</sub> entering the facility. In this case study, each process variant  $v$  is described by one of three flue gas flow rates and one of four carbon dioxide concentrations. This led to  $3 \times 4 = 12$  possible combinations, representing the boundary conditions for the 12 process variants. Considering the units in the capture facility flowsheet, we chose the common sub-component types to be the absorber and stripper,  $K = \{\text{absorber, stripper}\}$ , since they contribute heavily to the overall purchase cost and have large effects on the performance of the system, thereby directly affecting operating costs.

To gather surrogate training data, we discretized the design ranges for the absorber and stripper. For the absorber, we considered 10 diameters in the range of 0.3 m – 1.2 m. For the stripper, we considered 10 diameters in the range of 10 in.–55 in. We performed a simulation of the carbon capture system for all possible combinations of the process variant boundary conditions and sub-component sizes. This led to  $12 \times 10 \times 10 = 1,200$  simulations. We used an Aspen Plus model of the aqueous MEA system for the simulations (Morgan et al., 2018). Each simulation was marked as *feasible* if the key performance requirements (e.g., CO<sub>2</sub> recovery) were met and *infeasible* otherwise. This set of data and corresponding *feasible*, *infeasible* labels were used to train a 15-node, single-layer ReLU activated NN to determine the performance indicator,  $p_v$ . The NN was imported directly in the formulation, as shown in Eq. (1c), using OMLT: Optimization and Machine Learning Toolkit (Ceccon et al., 2022). Purchase cost was estimated using the IDAES costing framework (Lee et al., 2021). Combined with estimates for annual operating cost, total annualized cost was used to train a linear decision tree (Ammari et al., 2022) and imported using OMLT, represented by Eq. (1b), to predict  $c_v$  of process variant  $v$ . Note that both surrogates are captured as piecewise linear functions, resulting in an MILP formulation.

For this case study, we considered two absorbers and two strippers for the sub-component designs (i.e.,  $J_k = \{1,2\}$  for each sub-component type). The problem was formulated using Pyomo (Bynum et al., 2021) and solved using Gurobi (Optimization, LLC Gurobi, et al., 2020). Additionally, we used OMLT and the Pyomo extension, Pyomo.GDP (Chen et al., 2022), to capture the disjunctions. The optimal designs for the two absorbers and two strippers are reported in Table (1), and Figure (1) shows which combinations of the

two possible absorber designs and two possible stripper designs were selected for each of the 12 process variants. The optimization determined the sizes for the common sub-components. The larger absorber was used for the upper right corner of the variant process space, while the larger stripper was needed only for the most demanding variant in the upper right-hand corner.

Table 1. Optimal Designs for Common Units

Sub-component $k$	Design 1 $j_1 \in J_k$	Design 2 $j_2 \in J_k$
Absorber	0.3 m. diameter	0.61 m. diameter
Stripper	0.254 m. diameter	1.075 m. diameter

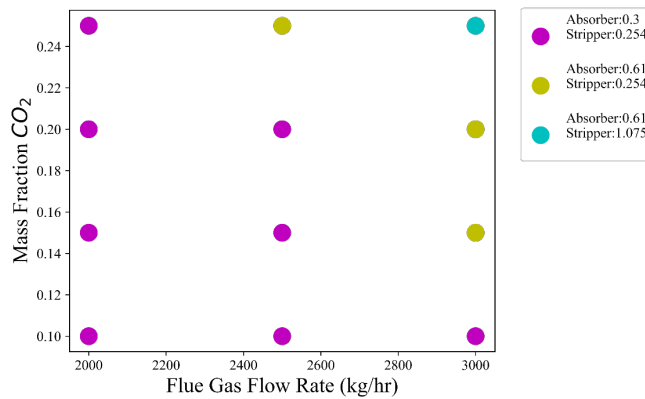


Figure 1. Process Family Design

Since the optimization is performed with surrogate models, we verified this result using the original model to simulate each of the 12 process variant boundary conditions with the corresponding set of assigned absorber and stripper designs. The total annualized verified cost was  $\$1.29 \times 10^6$ , which is an overall surrogate error of approximately 2%.

Table 2. Optimization Problem Specs

Num. of Binary Variables	191
Num. of Continuous Variables	273
Num. of Constraints	720
Total Annualized Cost (Obj.)	$\$1.26 \times 10^6$

The combination of training time for both the ReLU NN surrogate and linear decision tree was approximately 300 seconds, while the optimization formulation itself took approximately 1 second along with the problem structure as described in Table (2).

#### 4. Conclusions and Future Work

We demonstrated a surrogate-based approach for optimally designing a process family of 12 carbon capture variants. We considered a platform of common sub-component designs that considered two absorbers and two strippers, whereas a conventional design approach would have generated 12 unique absorber designs and 12 unique stripper designs. This approach uses surrogates to predict required system costs and performance. This removes

the nonlinearities in the GDP and allows the transformation to an MILP. The machine learning surrogates were easily included in the optimization with OMLT, and the overall formulation solved in under a second, demonstrating tractability in this approach.

For future work, we plan on expanding the sub-component design ranges and incorporating more sub-component types. Furthermore, we plan on incorporating an estimation of savings due to economies of numbers within the optimization formulation, exploring uncertainty in weights, and quantifying the degree of flexibility our designs offer across the design space.

**Disclaimer:** This project was funded by the United States Department of Energy, National Energy Technology Laboratory an agency of the United States Government, in part, through a support contract. Neither the United States Government nor any agency thereof, nor any of their employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

SNL is managed and operated by NTESS under DOE NNSA contract DE-NA0003525

**Acknowledgments:** This effort is part of the U.S. Department of Energy's (DOE) Institute for the Design of Advanced Energy Systems (IDAES) supported by the DOE Office of Fossil Energy and Carbon Management's Simulation-Based Engineering/Crosscutting Research Program.

## References

- Ammari, B., et al., 2022, Machine Learning Surrogates with OMLT and ideas for Improved Design and Analysis of Energy Systems, INFORMs Annual Meeting 2022, Indianapolis, IL.
- Baldea, M., et al., 2017, Modular manufacturing processes: Status, challenges, and opportunities. *AIChE journal* 63(10), 4262-4272
- Bynum, M. L., et al., 2021, Pyomo- Optimization Modeling in Python. *Springer*.
- Chen, Q., et al., 2022, Pyomo gdp: an ecosystem for logic based modeling and optimization development. *Optimization and Engineering* 23(1), 607-642-
- Ceccon, F., et al., 2022, Omlt: Optimization & machine learning toolkit, submitted, preprint on arXiv.
- Lee, A., et al., 2021, The idaes process modeling framework and model library- flexibility for process simulation and optimization. *Journal of Advanced Manufacturing and Processing* 3(3)
- Morgan, J.C., et al., 2018. Development of a rigorous modeling framework for solvent-based CO<sub>2</sub> capture. part 2: steady-state validation and uncertainty quantification with pilot plant data. *Industry & Engineering Chemistry Research* 57, 10464-10481.
- Optimization, LLC Gurobi, et al., 2020, Gurobi optimizer reference manual, URL <http://www.gurobi.com>
- Stinchfield, G., et al., 2022, Optimization-based Approaches for Design of Chemical Process Families Using ReLU Surrogates, submitted to Foundations of Computer Aided Process Operations, 2022.
- Wilberforce, T., et al., 2019, Outlook of carbon capture technology and challenges, *Science of the total environment*, 56-72

Zhang, C., et al., 2022, Optimization-based design of product families with common components. *Computer Aided Chemical Engineering, Proceedings of 14<sup>th</sup> International Symposium on Process Systems Engineering (PSE 2021+)*. Volume 49, pp. 91-96. Elsevier.



# Global Optimal Explainable Models for Biorefining

Jiayang Ren<sup>a</sup>, Kaixun Hua<sup>a</sup>, Heather Trajano<sup>a</sup>, Yankai Cao<sup>a,b</sup>

<sup>a</sup> *Department of Chemical and Biological Engineering, University of British Columbia, 2360 East Mall, Vancouver V6T 1Z3, Canada*

<sup>b</sup> *yankai.cao@ubc.ca*

## Abstract

Hemicelluloses are amorphous polymers of various sugar molecules and have been widely utilized in bioenergy, mining, and textile. Through hydrolysis, hemicellulose is transformed into sugar oligomers and monomers. In this paper, we build a global optimal decision tree (GODT) model on an extensive hemicellulose hydrolysis dataset containing 1955 experimental data points from 71 published papers from 1985 to 2019. The GODT model is trained to predict xylose yield from hardwood hemicellulose hydrolysis in batch reactors. Compared with the heuristic method, our global optimal algorithm can obtain an average absolute improvement of 1.54% in test accuracy. Moreover, we demonstrate that the reasoning procedure of predictions is easy to comprehend by human decision-makers, thus contributing an explainable model for biorefining.

**Keywords:** Biorefining; Global Optimization; Machine learning; Decision Tree

## 1. Introduction

Hemicelluloses are amorphous polymers of various sugar molecules, such as xyloses, mannoses, arabinoses, glucoses, and galactoses (Scheller & Ulvskov, 2010). It is the second most component in lignocellulosic biomass, ranging from 25 to 35 wt% (Isikgor & Becer, 2015). Hemicellulose has been broadly utilized in the bioenergy, mining, textile, cosmetic and pharmaceutical industries (Spiridon & Popa, 2008). Industrial use of hemicellulose often requires that the polymer be hydrolyzed into constituent oligomers and monomers. Scale-up and commercialization of hemicellulose hydrolysis require robust models. However, the traditional kinetic models can only consider limited operating regimes and species (Fearon et al., 2020). Fortunately, extensive studies on kinetic models have contributed abundant data to enable the development of statistical models. In our previous work (Wang et al., 2022), we mined the literature on hardwood hemicellulose hydrolysis and collected a dataset containing 1955 experimental points on batch hemicellulose hydrolysis from 71 published papers dating from 1985 to 2019. Three machine learning models (i.e., ridge regression, support vector regression, and artificial neural networks (ANN)) were developed to predict the xylose yield of hemicellulose hydrolysis utilizing this dataset. This

work also showed the superior performance of ANN in predicting the xylose yield of hemicellulose hydrolysis, compared with kinetic models. However, the black-box nature of ANN makes the interpretability of predictions difficult.

This paper aims to develop an interpretable model based on decision trees to predict the xylose yield of hemicellulose hydrolysis. The decision tree is a powerful machine-learning method for classification and regression tasks. The key advantage of decision tree modeling lies in interpretability, as such models closely resemble human reasoning and are easy to comprehend by human decision-makers (Weiss & Kulikowski, 1991). For example, the operator of a biorefinery plant is more likely to understand and trust a decision tree model over black-box models (e.g., neural network) since the decision tree breaks a complex decision-making process (e.g., the efficiency of the operation) into a collection of more straightforward logical tests (e.g., temperature range).

A fundamental problem is decision tree learning, a task to find the most accurate decision tree model from a dataset. For decades, we have relied on greedy-based heuristics, such as CART (Loh, 2011), ID3 (Quinlan, 1986), and C4.5 (Quinlan, 2014), to obtain suboptimal decision trees. Recent studies based on a comprehensive benchmark have shown that global solutions obtained through mixed-integer optimization (MIO) led to an average absolute improvement in test accuracy of 1 - 5% (Bertsimas & Dunn, 2017). Despite its superior performance, existing MIO-based approaches (Bertsimas & Dunn, 2017; Aghaei et al., 2021; Verwer & Zhang, 2019) rely on all-purpose global optimization solvers and can only address a small dataset (e.g., several hundred samples). Recently, we developed a tailored reduced-space branch and bound algorithm to train optimal decision trees for classification tasks (Hua et al., 2022). The algorithm is based on the observation that decision tree learning can be formulated as a two-stage stochastic programming problem, and thus various approaches proposed in the stochastic programming community can be utilized to exploit the problem structure. This work extended the solvable optimal decision tree problem to 245,000 samples.

In this paper, we apply the global optimal algorithm proposed by Hua et al. to the biorefining dataset and show that the average training and testing accuracy can be improved by 1.70% and 1.54% compared with the heuristic method.

## **2. Problem Description**

### *2.1. Dataset Description and pre-process*

We adopt the hemicellulose hydrolysis dataset from our previous work (Wang et al., 2022). This dataset collected 1955 data points from 71 published works from 1985 to 2019. Each data point has eight numeric variables and two categorical variables. Wang et al. have demonstrated that the two categorical variables (wood and acid species) contribute little in predicting the xylose yield.

Besides, these two categorical variables need one-hot encoding to be utilized in the decision tree model, which will introduce redundant binary variables and make the decision tree modeling unstable. Therefore, we only use eight numeric variables in this paper, as shown in Table 1. Notably, there are some repeating experiments with the same reaction and biomass conditions but resulting in different xylose yields. We unified these data points with the average value of xylose yields and obtained a unique dataset with 1828 data points.

Table 1 Dataset variable list

No.	Variables	Type	Units	Range
1	Total Reaction Time	Numeric	minute	0 - 1155
2	Reactor Temperature	Numeric	Kelvin	313 - 553
3	Liquid Solid Ratio	Numeric	g liquid / g solid	1.53 - 50
4	Initial Proton Concentration	Numeric	mols/L	0-1.22
5	Particle Size	Numeric	mm	0.117 - 50
6	Isothermal Reaction Time	Numeric	minute	0 - 1140
7	Initial Hemicellulosic Xylose in Feedstock	Numeric	weight %	11.932 - 34.9
8	Xylose Yield	Numeric	%	0 - 100

To adapt to the classification decision tree algorithms, we further encoded the variable xylose yield from a numeric variable to a categorical variable with  $K$  categories. Specifically, we created two classification datasets with three and five categories. For the first dataset, the yield is either lower [0%, 30%], medium (30%, 75%), or high (75%, 100%]. For the second one, the labels of yield are very low [0, 20%], low (20%, 40%], medium (40%, 60%], high (60%, 80%], or very high (80%, 100%].

## 2.2. Problem Description

The paper aims to develop a classification decision tree to predict the encoded xylose yield of hemicellulose hydrolysis utilizing variables 1-7 in Table 1. Firstly, we randomly divide the whole dataset into training and testing datasets. Then, let  $\mathcal{S} = \{1, \dots, n_{train}\}$  be the index set of the training dataset,  $\mathcal{S}_{test} = \{1, \dots, n_{test}\}$  be the index set of the testing dataset,  $\mathcal{K} = \{1, \dots, K\}$  be the class set of the encoded xylose yield. For  $s$ th data point, we denote the encoded xylose yield as  $\mathbf{y}_s \in \{0,1\}^K$  and the feature variables as  $\mathbf{x}_s \in \mathbb{R}^P$ . Sequentially, the training dataset can be represented as  $\mathbf{X} = \{\mathbf{x}_s | \mathbf{x}_s \in \mathbb{R}^P, s \in \mathcal{S}\}$  with the label set  $\mathbf{Y} = \{\mathbf{y}_s | \mathbf{y}_s \in \{0,1\}^K, s \in \mathcal{S}\}$ . Then, the problem of this paper can be represented as finding a decision tree  $F: \mathbb{R}^P \rightarrow \{0,1\}^K$  with maximum depth  $D$  to solve the following optimization problem:

$$\min \sum_{s \in \mathcal{S}} E(F(\mathbf{x}_s), \mathbf{y}_s) + \lambda R(F) \quad (1)$$

where  $E(\cdot)$  is the misclassification error between the predicted label of tree  $F(\mathbf{x}_s)$  and the corresponding ground truth  $\mathbf{y}_s$ ,  $\lambda$  is the complexity parameter to balance the accuracy and tree complexity,  $R(\cdot)$  is the complexity function to measure the tree's complexity. After obtaining the decision tree, the labels of testing samples are predicted by  $F(\mathbf{x}_s), s \in \mathcal{S}_{test}$ .



### 3. Global Optimal Decision Tree (GODT)

Given the maximum depth of a decision tree  $D$ , the nodes of the tree are denoted as  $t \in \{1, \dots, T = 2^{D+1} - 1\}$ . We further define branch nodes as  $t \in \mathcal{T}_B = \{1, \dots, \lfloor T/2 \rfloor\}$  and leaf nodes as  $t \in \mathcal{T}_L = \{\lfloor T/2 \rfloor + 1, \dots, T\}$ . Then, the decision tree problem (1) can be reformulated in the following MIO form:

$$\begin{aligned}
 \min_{a,b,c,d,z,L} \quad & \frac{1}{L} \sum_{s \in \mathcal{S}} L_s + \lambda \sum_{t \in \mathcal{T}_B} d_t & (2) \\
 \text{s.t.} \quad & \frac{1}{2} \sum_{k \in \mathcal{K}} (y_{sk} + c_{kt} - 2y_{sk}c_{kt}) - L_s \leq 1 - z_{st}, \forall t \in \mathcal{T}_L \\
 & \sum_{k \in \mathcal{K}} c_{kt} = 1, \forall t \in \mathcal{T}_L \\
 & \sum_{t \in \mathcal{T}_L} z_{st} = 1 \\
 & a_m^T (\mathbf{x}_s + \epsilon - \epsilon_{\min}) + \epsilon_{\min} \leq b_m + (1 + \epsilon_{\max})(1 - z_{st}), \forall m \in A_L(t), t \in \mathcal{T}_L \\
 & a_m^T \mathbf{x}_s \geq b_m - (1 - z_{st}), \forall m \in A_R(t), t \in \mathcal{T}_L \\
 & \sum_{j=1}^P a_{jt} = d_t, \forall t \in \mathcal{T}_B, j \in \{1, \dots, P\} \\
 & 0 \leq b_t \leq d_t, \forall t \in \mathcal{T}_B \\
 & d_t \leq d_{p(t)}, \forall t \in \mathcal{T}_B \\
 & 0 \leq L_s \leq 1 \\
 & a_{jt}, d_t \in \{0,1\}, 0 \leq b_t \leq 1 \forall t \in \mathcal{T}_B, j \in \{1, \dots, P\} \\
 & z_{st}, c_{kt} \in \{0,1\}, \forall t \in \mathcal{T}_L
 \end{aligned}$$

where  $L_s \in [0,1]$  is the loss of sample  $s$ .  $L_s$  equals to 1 if sample  $s$  is misclassified, otherwise 0. Variables  $a, b, c, d$  are the structure variables of the decision tree. Specially,  $a_t = [a_{1t}, \dots, a_{pt}] \in \{0,1\}^P$  represents the split dimension on node  $t$ .  $b_t \in [0,1]$  is the split value on node  $t$ .  $c_{kt} \in \{0,1\}$  is the class indicator at the leaf node  $t \in \mathcal{T}_L$ .  $d_t \in \{0,1\}$  is the split indicator at the decision node  $t \in \mathcal{T}_B$ . A more detailed description of this MIO formulation can be found in Hua et al., 2022.

The fundamental idea to solve this MIO problem is to decompose the problem into a two-stage problem, in which structure variables  $a, b, c, d$  are the first-stage variables. The two-stage problem is shown as follows:

$$\text{First-stage:} \quad f(M_0) = \min_{m \in M_0} \sum_{s \in \mathcal{S}} Q_s(m) \quad (3)$$

$$\text{Second-Stage:} \quad Q_s(m) = \min_{z_s, L_s} \frac{1}{L} L_s + \frac{\lambda}{n} \sum_{t \in \mathcal{T}_B} d_t \quad (4)$$

s.t. Constraints in problem (2)

where  $m = (a, b, c, d)$  is the first-stage variables, and  $M_0 := [m^l, m^u]$  is the initial region of the first-stage variables. Then we can use the branch and bound scheme to solve the resulting two-stage problem. The key advantage here is that

we can prove that branching only on variables denoting the tree structures  $(a, b, c, d)$ , our branch and bound training algorithm can guarantee the convergence to a global optimal solution with properly designed lower and upper bounds. Moreover, the decomposable nature of our lower and upper bound methods allows us to easily parallelize the branch and bound decision tree training algorithm on thousands of CPU cores. The detailed procure of the tailored reduced-space branch and bound algorithm to solve problem (2) can be found in Hua et al., 2022.

## 4. Results and Discussions

### 4.1. Experiment setup

In our study, we randomly split the dataset into a training dataset (75% of the entire dataset) and a testing dataset (25%). We compare the performance of our global optimal decision tree (GODT) with the popular heuristic algorithm CART (Loh, 2011). Different tree depths, including 2 and 3 levels, were considered in our experiments. All experiments were executed using Julia 1.7.0 with 4 hours time limit on a high-performance computing cluster, of which each node contains 40 Intel cores at 2.4 GHz and 202 GB RAM. We executed all the experiments in parallel mode with 40 cores for two-level trees and 80 cores for three-level trees.

To compare the performance of each algorithm, we reported train and test accuracy (%) :=  $\frac{\text{correct classification}}{\text{all classification}}$ . We also reported the optimality gap for GODT as  $Gap (\%) := \frac{\text{upper bound} - \text{lower bound}}{\text{upper bound}} \times 100\%$ . This optimality gap is the worst gap of the current solution from the global optimal solution, which is a unique property of the deterministic global optimization algorithms. Heuristic algorithms such as CART cannot provide this worst gap. Besides, to fairly compare the performance, we repeated the experiments five times with five different random splits and reported means and standard deviations of each index in the form of *mean (standard deviation)*.

### 4.2. Comparison results

Table 2 shows the comparison results of CART and GODT on different datasets with 3 and 5 categories. The method with better training and test accuracy is bolded in this table. Generally, both GODT and CART perform better on the dataset with fewer categories and trees with more levels. For the optimality gap, the GODT method can obtain a small and acceptable optimality gap on two-level trees while maintaining a larger optimality gap on three-level trees. Despite the larger optimality gap on three-level trees, GODT still performs better than CART. As for accuracy, GODT can obtain average absolute improvements of 1.70% and 1.54% in train and test accuracies, respectively.

Table 2 Comparison results of CART and GODT. Each index is in the form of *mean (standard deviation)*.

Yield Category	Tree Depth	Method	Optimality Gap (%)	Train Accuracy (%)	P value	Test Accuracy (%)	P value
3	2	CART	-	64.62 (0.86)	9.80E-04	64.74 (2.62)	6.98E-04
		<b>GODT</b>	<b>0.92 (0.04)</b>	<b>65.73 (0.92)</b>		<b>66.14 (2.72)</b>	
3	3	CART	-	68.46 (1.45)	1.96E-01	67.98 (2.23)	1.96E-01
		<b>GODT</b>	<b>88.86 (0.61)</b>	<b>68.86 (0.98)</b>		<b>68.99 (1.07)</b>	
5	2	CART	-	44.24 (0.30)	2.18E-05	43.78 (1.00)	4.54E-03
		<b>GODT</b>	<b>5.27 (0.79)</b>	<b>48.61 (0.24)</b>		<b>46.51 (0.37)</b>	
5	3	CART	-	49.37 (0.59)	4.88E-02	47.12 (1.45)	1.40E-01
		<b>GODT</b>	<b>79.55 (1.70)</b>	<b>50.30 (0.83)</b>		<b>48.13 (2.91)</b>	

*Remark 1.* GODT focuses on optimizing the training accuracy in problem (2). We have proved that global optimality can be reached on the problem (2). However, there is no guarantee that the test accuracy of GODT can also be globally optimal or consistently exceed CART (considering over-fitting). Nevertheless, extensive numerical results in this paper and previous works have demonstrated that an improvement in the test accuracy can be expected on most datasets for GODT, especially when the number of samples is large (e.g.,  $\geq 1000$ ) and the tree depth is small (e.g., 2-4), since over-fitting is avoided.

*Remark 2.* The p-values of one-tailed paired t-tests show that when tree depth is 2, GODT's train and test accuracy is statistically significantly larger than CART. When tree depth is 3, the results are not statistically significant. That is possible because we have not reached a small optimality gap in these experiments. We may obtain a more significant result if we develop a more efficient global algorithm for the decision tree.

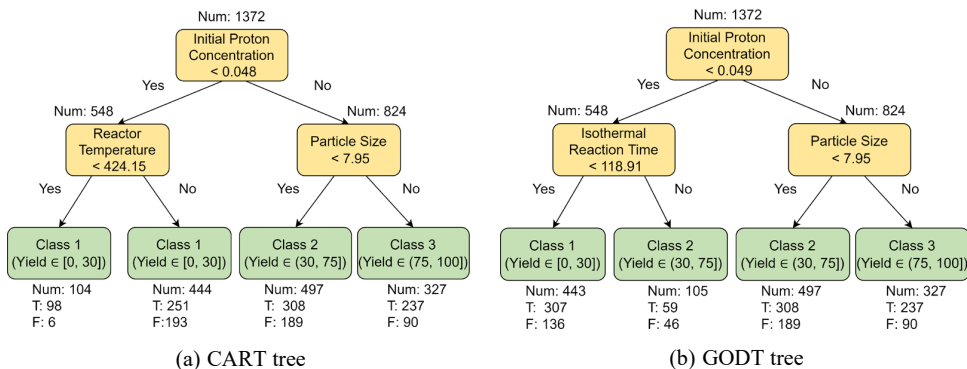


Figure 1 Decision Trees with two levels

### 4.3. Interpretability of models

The reasoning procedures of decision trees are similar to human reasoning and easy to comprehend by human decision-makers. Figure 1 shows the detailed reasoning procedures of two-level trees on the three-category training dataset in one experiment. In this experiment, the CART and GODT trees obtained a training accuracy of 65.16% and 66.40%, respectively. Notably, the optimality gap of this GODT tree is 0.92%. The CART tree selected initial proton

concentration, reactor temperature, and particle size as the reasoning variables, while the GODT tree selected initial proton concentration, isothermal reaction time, and particle size as the reasoning variables. It should be noted that the CART trees have two leaf nodes with the same parent node and the same predicted class. This kind of split is abnormal since this split does not improve accuracy. The reason for this abnormal split is that CART utilizes the Gini index as the objective function at each node to make its split decisions, which differs from the accuracy we want to optimize. This kind of split may reduce the Gini index but not influence the accuracy. As an outcome, the accuracy of the GODT tree is better than the CART tree.

## 5. Conclusion

In this paper, we developed a global optimal explainable decision tree model for hemicellulose hydrolysis. This global optimal decision tree algorithm utilizes the MIO form of the decision tree and solves the problem by decomposing the MIO problem into a two-stage stochastic programming problem. Experiments on a hemicellulose hydrolysis dataset show that the global optimal decision tree model can obtain an average absolute improvement of 1.54% in test accuracy compared to the heuristic decision tree model while providing an acceptable optimality gap. Moreover, as a nature of decision tree models, the proposed models closely resemble human reasoning and are easy to comprehend by human decision-makers. In the future, we aim to develop more efficient global optimal explainable decision tree models to achieve higher prediction accuracies and lower computational costs.

## References

- H. V. Scheller, & P. Ulvskov, 2010, Hemicelluloses, *Annual Review of Plant Biology*, 61, 263–289
- F. H. Isikgor, & C. R. Becer, 2015, Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers, *Polymer Chemistry*, 6(25), 4497-4559.
- I. Spiridon, & V. I. Popa, 2008, Hemicelluloses: major sources, properties and applications, *Monomers, Polymers and Composites from Renewable Resources*, 289-304.
- O. Fearon, V. Nykänen, S. Kuitunen, K. Ruuttunen, R. Alén, V. Alopaeus, & T. Vuorinen, 2020, Detailed modeling of the kraft pulping chemistry: carbohydrate reactions, *AIChE Journal*, 66(8), e16252.
- E. Wang, G. Cai, R. Ballachay, Y. Cao, & H. L. Trajano, 2022, Predicting Xylose Yield in Prehydrolysis of Hardwoods: A Machine Learning Approach, *Frontiers in Chemical Engineering*, 84.
- S. M. Weiss, & C. A. Kulikowski, 1991, *Computer systems that learn: classification and prediction methods from statistics, neural nets, machine learning, and expert systems*, Morgan Kaufmann Publishers Inc..
- W. Y. Loh, 2011, Classification and regression trees, *Wiley interdisciplinary reviews: data mining and knowledge discovery*, 1(1), 14-23.
- J. R. Quinlan, 1986, Induction of decision trees, *Machine learning*, 1(1), 81-106.
- J. R. Quinlan, 2014, *C4. 5: programs for machine learning*, Elsevier.
- D. Bertsimas, & J. Dunn, 2017, Optimal classification trees, *Machine Learning*, 106(7), 1039-1082.

- S. Aghaei, A. Gómez, & P. Vayanos, 2021, Strong optimal classification trees, arXiv preprint arXiv:2103.15965.
- S. Verwer, & Y. Zhang, 2019, Learning optimal classification trees using a binary linear program formulation, In Proceedings of the AAAI conference on artificial intelligence, 33(1), 1625-1632.
- K. Hua, J. Ren, & Yankai Cao, 2022. A Scalable Deterministic Global Optimization Algorithm for Training Optimal Decision Tree. Advances in Neural Information Processing Systems.

# NUMERICAL SIMULATION OF THE L-V EQUILIBRIUM WITHIN A STAGE IN A DISTILLATION COLUMN USING CFD

Perla G. Canchola-López<sup>a</sup>, Ariadna E. Vázquez-Hernández<sup>a</sup>, Jazmín Cortez-González<sup>a</sup>, Rodolfo Murrieta-Dueñas<sup>a</sup>, Roberto Gutiérrez-Guerra<sup>b</sup>, Carlos E. Alvarado-Rodríguez<sup>c</sup>

<sup>a</sup> *Tecnológico Nacional de México/ Irapuato, Ingeniería Química y Bioquímica,; Carretera Irapuato-Silao Km. 12.5, C.P.:36821 Irapuato, Guanajuato, MEXICO;*

<sup>b</sup> *Universidad Tecnológica de Leon, Departamento de Sustentabilidad para el desarrollo. Blvd. Universidad Tecnológica 225, Universidad Tecnológica, San Carlos la Roncha, 37670 León, Gto.*

<sup>c</sup> *Departamento de Ingeniería Química, División de Ciencias Naturales y Exactas, Universidad de Guanajuato, Noria Alta S/N, C.P. 36050, Guanajuato, Guanajuato, México.*

## Abstract

This work presents the numerical simulation in L -V equilibrium stage in a plate distillation column using the SPH method, considering Sieve and Bubble cap plates. To perform the simulation of the equilibrium stage, periodic conditions in temperature were established. The sizing of the column was performed in Aspen One considering an equimolar mixture of Benzene-Toluene and an operating pressure that guarantees that the cooling water temperature of the condenser is 120°F, the thermodynamic model used was Chao-Seader. The following information was obtained from the Aspen simulation: liquid and vapor velocity per stage, viscosity and density of the mixture, operating pressure, and column diameter. Data from Aspen simulation was used for the CFD simulations using DualSPHysics. The results allow it to compare the thermal distribution and the velocity fields between plate design and the time to achieve the thermal equilibrium.

**Keywords:** CFD, Simulation of Distillation, thermal equilibrium.

## 1. Introduction

Traditionally, the numerical simulation of hydrodynamics and heat transfer in distillation column has been performed with Eulerian methods. These methods use mesh for the discretization of the medium, thus generating averages in the transfer zones and interfaces between fluids. On the other hand, Lagrangian methods allow to see in detail the phenomena at the interface and to discretize the continuous medium through points meshless. In this work we propose to perform a numerical simulation of distillation columns using the SPH method (Lagrangian nature).

Distillation is a highly inefficient unitary operation and its energy consumption is high, so some authors have proposed new alternatives to improve efficiency and reduce energy consumption. Yin et al., (2002) studied the liquid retention in a packed column with Pall rings and the distribution of the fluid over the packing's. They find that the distribution of the liquid retention is not uniform due to the design of the distributor.

Fu et al., 2020 investigated the flow and retention of a liquid in a packed column with pall rings finding that the randomness of the packing generates transitions from viscous flow to turbulent regimes in a Reynolds Number range between 6.7 and 40.2. Jain et al., 2021 developed an Eulerian model to simulate multiphase gas-liquid flow in a wide variety of regimes over a manifold, finding that the discharge patterns, the uniformity and symmetry of the flow lines and the pressure drop across the manifold exhibit significant differences from each other. Pichler et al., 2021, studied heat transfer in packed columns by numerical simulations considering conduction and natural convection suggesting a correction as a function of particle diameter that can reduce up to 75% the error in surface temperature. In this work the numerical simulation of the hydrodynamics and heat transfer in one stage inside a distillation column using the Smoothed Particle Hydrodynamics (SPH) method is performed in order to compare velocity fields, retention times and heat transfer in one stage. The SPH method is a mesh-free method with which it is possible to track the interface between the fluids immersed in a flow in a clear and simple way, thus allowing a better analysis of the phenomena occurring in it.

## 2. Methodology

In this work a methodology for the simulation of the hydrodynamics and thermal equilibrium of plate distillation columns is proposed. This methodology consists of 3 stages: rigorous design of the distillation column in Aspen Plus, 3D design of the column - sieve plates and hydrodynamic analysis of the columns using the SPH method. The Aspen Plus simulation results were used as input data for SPH. The 3D design was performed in SolidWorks.



Figure 1. Methodology for the numerical simulation of the hydrodynamics and equilibrium of distillation columns.

### 2.1. Rigorous design of the distillation column in Aspen Plus

The design of the distillation column was carried out considering a binary equimolar Benzene-Toluene mixture, whose equilibrium is modeled with Chao-Seader. A Rad-Frac model was used for column equilibrium. It was proposed at 120°F to ensure cooling water in the condenser. The feed is 100lbmol/hr. According to the rigorous design, 10 stages were obtained to achieve 98% purity and recovery for each component. Table 1 shows the results obtained from the Aspen One® simulation of stages 6 and 7, which were used as input data for the SPH method.

### 2.2. Column and sieve plate sizing

The 3D design of the column and plates was carried out in SolidWorks. The design parameters used were: column diameter, effective and downspout areas, obtained in Aspen One. The distance between plates was proposed to be 0.15m. Stainless steel 316L

was chosen as construction material. Figure 1 shown in detail the 3D design of the sieve and bubble cap plates.

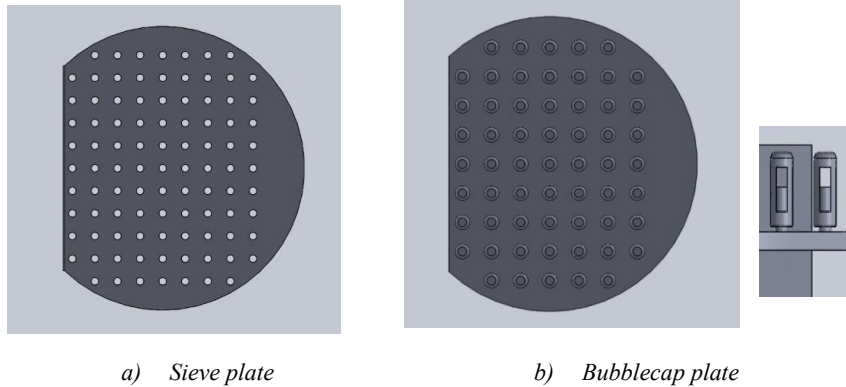


Figure 1. 3D design plates using in distillation columns. a) sieve and b) bubble cap.

Table 1. Results of the rigorous simulation of the distillation column in Aspen Plus.

Parameter	Liquid	Vapor
Density (kg/m <sup>3</sup> )	783.8023	4.0206
Viscosity (Ns/m <sup>2</sup> )	2.5x10 <sup>-4</sup>	9.5x10 <sup>-6</sup>
Temperature (°C)	110.39	105.99
Mass fraction	0.5886	0.5405
Velocity (m/s)	0.0261	0.0214
Coefficient of thermal diffusivity (m <sup>2</sup> /s)	7.9247E-8	
Surface tension	0.018 N/m	0.018 N/m

### 2.3. Hydrodynamics and heat transfer in SPH

SPH is a Lagrangian meshless method with applications in the field of Computational Fluid Dynamics. Originally invented for astrophysics in the 1970s (Monagan J, 1992) it has been applied in many different fields, including fluid dynamics (Alvarado-Rodríguez C.E. et al., 2019) and wastewater treatment (Mokos et al., 2015). The method uses points named particles to represent the continuum and these particles move according to the governing equations in the fluid dynamic. When simulating free-surface flows, no special surface treatment is necessary due to the Lagrangian nature of SPH, making this technique ideal for studying violent free-surface motion. Moreover, the movement of the boundaries can be set easily without the necessity of update the mesh as in the Eulerian methods. The SPH formalism used in the simulations is reported by Dominguez et al. (2021) which is set in the DualSPHysics code, in this work only the continuity equation (Eq. 1), the momentum equation (Eq. 2), the equation of state (Eq. 3) and the energy transport equation (Eq. 4) in the SPH formalism are reported. Mass transport can be added for future work (Chiron et al., 2019)



$$\frac{d\mathbf{v}_a}{dt} = -\sum_b m_b \left( \frac{p_a + p_b}{\rho_a \rho_b} + \sum_b m_b \left( \frac{4v_0 r_{ab} \cdot \nabla_a W_{ab}}{(\rho_a + \rho_b)(r_{ab}^2 + \eta^2)} \right) \mathbf{v}_{ab} + \sum_b m_b \left( \frac{\tau_{ab}^j}{\rho_b^2} + \frac{\tau_{ab}^i}{\rho_a^2} \right) \nabla_a W_{ab} \right) \nabla_a W_{ab} + g, \quad (1)$$

$$\frac{d\rho_a}{dt} = -\rho_a \sum_b \frac{m_b}{\rho_b} (\mathbf{v}_b - \mathbf{v}_a) \cdot \nabla_a W_{ab}, \quad (2)$$

$$P = B \left[ \left( \frac{\rho}{\rho_0} \right)^\gamma - 1 \right], \quad (3)$$

$$\frac{dT_a}{dt} = -\frac{1}{C_p} \sum_b \frac{m_b (k_a - k_b) (r_a - r_b) \cdot \nabla_a W_{ab}}{\rho_a \rho_b (r_{ab}^2 - \eta)} (T_a - T_b), \quad (4)$$

where the subscripts  $a$  and  $b$  are denoted for the mean particle “ $a$ ” and the neighbors particles “ $b$ ”,  $\mathbf{v}$  is the velocity,  $t$  is time,  $m$  is mass,  $P$  is pressure,  $\rho$  is density,  $v_0$  is the kinematic viscosity,  $\tau$  is the stress tensor,  $B = c_0^2 \rho_0 / \gamma$ ,  $c_0$  is an artificial sound speed, and  $\gamma = 7$ ,  $C_p$  is the specific heat of the fluid,  $k$  is the conductivity constant,  $T$  is the temperature and  $W$  is the kernel function defined in the SPH method. This method was implemented for the numerical analysis of the distillation columns under the following considerations: periodic conditions were used to perform the hydrodynamic and equilibrium analysis. The initial properties of the fluids are shown in Table 1. In all cases, no-slip conditions were considered in the interaction between fluid and boundary particles using the dynamic particle method reported by (Crespo et al., 2007a). A different resolution was used in each case due to the specific characteristics for the geometries in order to save computational time, obtained through a convergence analysis for each phase. SPH results have so far only been compared with those obtained from the Aspen One simulation.

### 3. Discussion of results

With the results obtained from the SPH numerical simulations, it is possible to analyze the velocity profile and the temperature profile inside the stage, for each plate design. In addition, it is possible to approximate the required time that each stage needs to reach equilibrium under the initial conditions established.

#### 3.1. Bubble cap plate

The simulation was performed with a total of 1,686,948 fluid particles of which 1,042,345 are liquid and 684,936 are vapor. The particles modeling the plate and column are 535,069. We simulated 30 seconds of real-time. The simulation took a total of 10.5h on the ITESI computer with an Nvidia GPU Geforce RTX 3060. Figure 2 shows the velocity field obtained in the simulation from the top and side perspectives. According to the results obtained, it can be deduced that the vapor velocity below the cap does not register fluid movement, however above it, the profile is almost homogeneous. This can be contrasted with the results of the temperature profile shown in Figure 3, for the same stages analyzed. The stagnation of the fluid shown in Figure 2 generates a retard in the homogeneity of the temperature in that zone. This is due to the fact that the temperature variation is dominated by the effect of diffusion, which means that even after 30 seconds, the final temperature corresponds to the initial value.

### 3.2. Sieve plate

The simulation was performed with a total of 1,433,006 fluid particles of which 962,243 are liquid and 470,763 are vapor. The particles modeling the plate and column are 387,365. 10 seconds of real-time were simulated. The simulation was completed in a total of 2.5h on the same ITESI computer. Figure 4 shows the velocity field obtained in the simulation from the top and side perspectives. As in the bubble cap plate, a higher velocity is obtained in the area where the fluid enters from the downspout, however, the velocity distribution is more homogeneous compared to the bubble cap plate due to the velocity generated by the vapor in the orifices. This homogeneity in the flow also generates homogeneity in the temperature as shown in Figure 5.

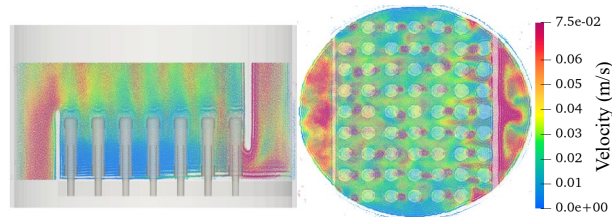


Figure 2. Side and top perspectives of the velocity fields in the stage for the bubble cap plate.

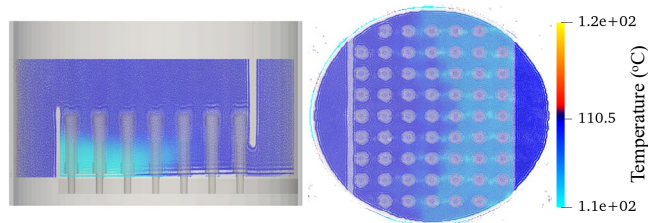


Figure 3. Side and top perspectives of the temperature fields in the stage for the bubble cap plate.

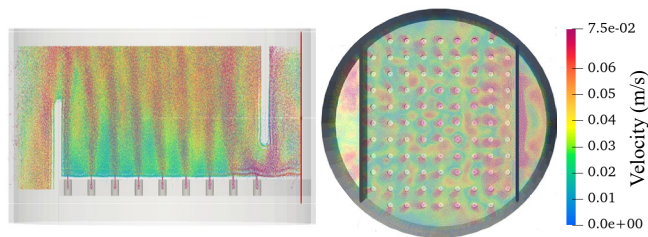


Figure 4. Side and top perspectives of the velocity fields in the stage for the sieve plate.

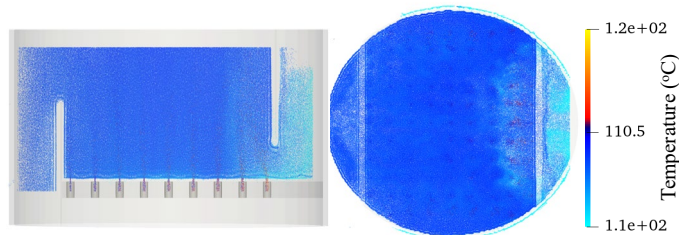


Figure 5. Side and top perspectives of the temperature fields in the stage for the sieve plate.

#### 4. Conclusions

In this work, the simulation of a stage inside a distillation column is presented. Two case studies were developed: bubble cap and sieve plates. The geometry of the column and the plates were made in SolidWorks. Then, the geometries were exported to DualSPHysics to perform the corresponding simulations, based on the parameters obtained with the rigorous simulation in Aspen One. The numerical results indicate that the height of the bubble caps is not adequate since there is no homogeneous temperature distribution in the area near the downspout towards the next stage. A proper configuration in the distribution of the caps or in the height of the caps could benefit a homogeneous distribution of velocity and temperature. In contrast to the bubble cap plate, in the sieve plate system, the velocity distribution is more homogeneous and reaches a constant temperature after about 10 seconds. Based on the methodology proposed in this paper, it is possible to evaluate the hydrodynamic and heat transfer performance of different geometries for packed or plate columns. These analyses will allow obtaining mathematical correlations with which it will be possible to quantify the impact of a design modification on the hydrodynamic performance of the column. This methodology can be extended to packed columns to determine how liquid retention increases or decreases with packing design and distribution.

#### References

- Alvarado-Rodríguez, C. E., Klapp, J., Domínguez, J. M., Uribe-Ramírez, A. R., Ramírez-Minguela, J. J., & Gómez-Gesteira, M. (2019, March). Multiphase Flows Simulation with the Smoothed Particle Hydrodynamics Method. In *International Conference on Supercomputing in Mexico* (pp. 282-301). Springer, Cham.
- Domínguez JM, Fourtakas G, Altomare C, Canelas RB, Tafuni A, García-Feal O, Martínez-Estévez I, Mokos A, Vacondio R, Crespo AJC, Rogers BD, Stansby PK, Gómez-Gesteira M. (2021), DualSPHysics: from fluid dynamics to multiphysics problems. *Computational Particle Mechanics*.
- Fu, Y., Bao, J., Singh, R., Wang, C., & Xu, Z. (2020). Investigation of countercurrent flow profile and liquid holdup in random packed column with local CFD data. *Chemical Engineering Science*, 221, 115693.
- Mokos, A., Benedict, D. R., Stansby, K. P., & Dominguez, J. M. (2015). Multi-phase SPH modelling of violent hydrodynamics on GPUs. *Journal Computer Physics Communications*, 196, 304-31.
- Monagan J J, "Smoothed Particles Hydrodynamics", *Annu. Rev. Astron. Astrphys.*, 30:534-74, 1992.
- Pichler, M., Haddadi, B., Jordan, C., Norouzi, H., & Harasek, M. (2021). Effect of particle contact point treatment on the CFD simulation of the heat transfer in packed beds. *Chemical Engineering Research and Design*, 165, 242-253.
- Jain, E., Sau, M., & Buwa, V. V. (2021). Eulerian simulations of liquid distribution generated by chimney and bubble cap distributors. *Chemical Engineering Journal*, 421, 127799.
- Yin, F., Afacan, A., Nandakumar, K., & Chuang, K. T. (2002). Liquid holdup distribution in packed columns: gamma ray tomography and CFD simulation. *Chemical Engineering and Processing: Process Intensification*, 41(5), 473-483.
- Chiron, S. Marrone, A. Di Mascio, D. Le Touzé, (2018). Coupled SPH-FV method with net vorticity and mass transfer, *Journal of Computational Physics*, Volume 364, Pages 111-136, ISSN 0021-9991.
- Crespo, A.J.C. ,Gómez-Gesteira, M.,Dalrymple,R.A., Boundary conditions generated by dynamic particles in SPH methods, *Computers, Materials & Continua* ,5,173–184, 2007a.

# MONITORING OF THE DRILLING REGION IN OIL WELLS USING A CONVOLUTIONAL NEURAL NETWORK

Caroline Dias Grossi<sup>a,b</sup>, Vinícius P. Barbosa<sup>c</sup>, Rubens Gedraite<sup>c</sup>, Maurício B. de Souza Jr.<sup>a</sup>, Cláudia Mirian Scheid<sup>b</sup>, Luís Américo Calçada<sup>b</sup>, Luiz Augusto da Cruz Meleiro<sup>b</sup>.

<sup>a</sup>*Federal University of Rio de Janeiro, EPQB/UFRJ, Rio de Janeiro/RJ, 21941-598, Brasil*

<sup>b</sup>*Federal Rural University of Rio de Janeiro, IT/DEQ/UFRRJ, Seropédica/RJ, 23897-000, Brasil*

<sup>c</sup>*Federal University of Uberlândia, DEQ/UFU, Uberlândia/MG, 38400-902, Brasil*  
*cdiasgrossi@eq.ufrj.br*

## Abstract

Currently, in the oil well drilling industry, the process is monitored by pressure and temperature sensors in the well and of the drilling fluid. Instabilities can be noticed with the gravel that returns in the mud, demanding the implementation of monitoring systems. This work proposes applying computer vision techniques to monitor vibrating screens and detect particulate material flow. Experiments were conducted on a pilot screen using a suspension of sand and 0.1% xanthan gum, simulating drilling mud. Each experiment varied operating conditions, obtaining 26 videos. The methodology evaluated the response of U-Net neural networks in semantic segmentation with Python. A database of 11,140 images and corresponding templates was built. The evaluation metrics were accuracy, F1-Score, and MeanIoU. The results show high segmentation ability, with accuracy 97%, F1-Score 92%, and MeanIoU 91%. The results show that the U-Net has good segmentation capabilities and provide a promising alternative for non-invasive process monitoring.

**Keywords:** computational vision, convolutional neural networks, oil wells drilling.

## 1. Introduction

Artificial Intelligence is a science that studies the ability of machines to act similarly to humans. For this, a machine learning step is necessary<sup>[1]</sup>. Currently, several algorithms can be found in the literature capable of solving problems in various industry segments. In this context, the choice of the appropriate algorithm will depend on the final objective and nature of the data<sup>[2]</sup>.

Image acquisition of industrial processes is a common practice in facility surveillance. The rise of machine learning models and the increased processing power of computers have allowed the development of models capable of processing increasingly complex data, such as images, which has motivated various industries to explore the information that the so-called computer vision can offer. Such a perspective can bring advantages to the automation of quality control systems, sensing, and safety monitoring, with a high potential of identifying signs of operational accidents before their occurrence<sup>[3]</sup>.

Shallow learning algorithms, in general, need a feature extraction phase before they can make predictions and classifications. The manual process of feature extraction from images is costly and can generate attributes irrelevant to the proposed objective, leading

to highly complex models that do not solve the problem. When these methods are not able to provide feasible results, deep learning techniques are applied<sup>[3][4]</sup>.

One example is Convolutional Neural Networks (CNN's), traditionally consisting of two modules: general learning and specific learning. The general learning module is responsible for receiving the original image and identifying, from it, information common to a diversity of objects from convolution and pooling operations. The response is the so-called attribute vector, which stores the learned characteristics<sup>[4]</sup>. The specific learning module consists of a fully connected layer responsible for solving the problem. It takes as input the attribute vector and provides the class to which the image belongs.

Besides the classification of images as a whole, the classification of each image pixel can be performed, a particular feature of Fully Convolutional Neural Networks (FCN). In such a type of FCN, the specific learning module is replaced by another convolutional layer capable of classifying individual pixels. This type of neural network can be trained without the need for feature extraction steps and exhibits a more accurate response regarding the shape of the objects of interest<sup>[5]</sup>.

The U-net (Figure 1), a particular type of FCN, was developed for biomedical image segmentation. Given the difficulty of creating a sizeable medical image bank, U-Net was initially designed to provide high-accuracy responses regarding segmentation with a low number of images available for training<sup>[6]</sup>.

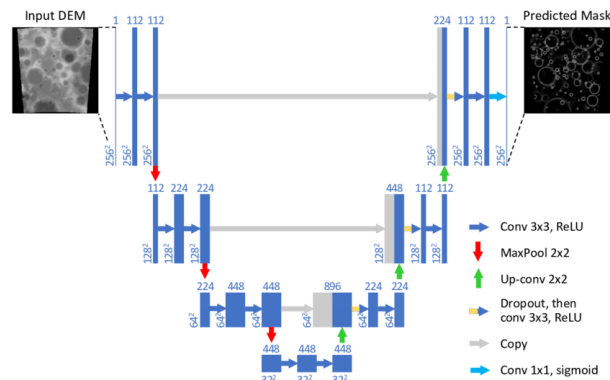


Figure 1. Representation of the U-Net architecture. Source: Silburt et al. 2018.

The U-Net architecture is formed by a sequence of contraction and expansion layers. In the shrinking step occurs dimensionality reduction, generating feature maps that preserve essential information, such as object contours. At the end of the contraction step, the expansion step occurs, which performs the opposite path, increasing the dimensionality of the image from the generated feature maps. One of the main advantages of this network is in the expansion path, which uses the output of the previous unpooling concatenated with the corresponding pooling layer to increase the number of feature maps. This feature improves the high-level feature extraction capability and makes the network learn the content and location of each pixel in the image<sup>[7]</sup>.

Some relevant uses of the U-Net, far from those cases applied to medical imaging, have been published in recent works with adaptations of this type of FCN, as follows: the detection of manufacturing defects in clothing fabrics<sup>[8]</sup>; mapping impervious surfaces from satellite images<sup>[9]</sup>; the identification of falsification in image splices using dual encoder U-Nets<sup>[10]</sup>; quantification of gas leakage in the subsea environment using CFD and fully convolutional U-net networks<sup>[11]</sup>.

The process of drilling oil wells is complex, and with advances in-depth, instabilities are more likely to occur<sup>[12]</sup>. Therefore, it is convenient to monitor all stages of the process, which is already done using sensors installed on the rig. Most of the variables currently monitored are related to the pressure and temperature conditions along the well and the drilling fluid. An interesting approach would be to evaluate the mud that reaches the surface, which can also indicate the beginning of some instabilities, such as the well collapsing and insights from the column trapped.

Thus, this work proposes using computer vision methods to detect the gravel that reaches the surface. The method evaluated the response of U-Net type fully convolutional neural networks in the semantic segmentation of images that simulate the flow in vibrating screens, the first equipment to receive the drilling mud. Another justification for using this type of network is the difficulty in obtaining images of gravel flow on actual screens, in rigs, and the adaptability of U-Net for the scenario of small image banks.

## 2. Methodology

Due to the sake of industrial secrecy related to obtaining images from the operation of the actual sieves, the behavior of this process equipment was investigated from experiments performed in a pilot-scale sieve using a suspension of sand and xanthan gum. A total of 26 videos with varied operating conditions was employed in the present study. The experiments are detailed in [P].

The tests with the proposed U-Net architecture used 11,140 flow images with their respective templates; 20% of these were used to test the FCN and 80% for the training stage. From the training data bank, 10% was used for the validation stage, which used a cross-entropy loss function. Examples of images from the database can be seen in Figure 2. The templates were obtained in a semi-automated way from the construction of image processing algorithms specific to the color, brightness, and contrast conditions present in each video. Conversion to color spaces other than RGB, histogram equalization, and digital filtering for image quality improvement were explored.

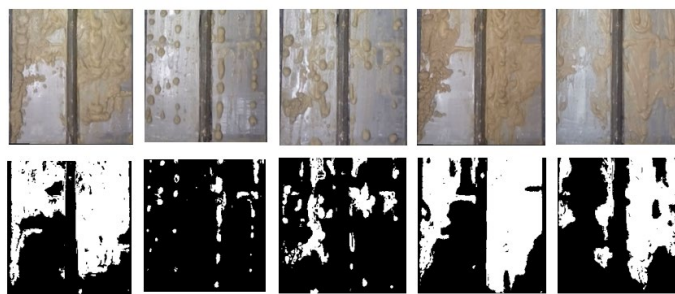


Figure 2. Examples of images from the database.

The U-Net algorithm was developed in Python, using Keras and TensorFlow. The preprocessing step performed the image acquisition in its original form, resizing to size [128, 128] and storing the data in tensors. The procedure was the same for the network input images and the template, with the difference that the input images are in color channels and the templates in binary masks.

Next, the network architecture was built using convolution, pooling, and ReLu activation layers in both the contraction and expansion stages. After input, the data goes through a normalization step that forces the pixels in each channel to assume values between zero

and one. Next, the image passes through each layer of the network, which has a depth equal to 4 levels, until it reaches the final layer, which uses a sigmoidal function to perform the pixel-by-pixel classification.

At the end of the model construction, training parameters were defined, such as early stopping criteria, which helps minimize the occurrence of overfitting and the maximum number of epochs. The model was trained, and the network was saved for later use.

### 3. Results and Discussions

During training, accuracy and loss data were collected for training and validation and used in constructing the graphs in Figure 3. It was noticed that the curves converged quickly to a model with high accuracy and low loss rate, with no indication of overfitting. The final accuracy found for this data set was 98.85% for training and 98.88% for testing.

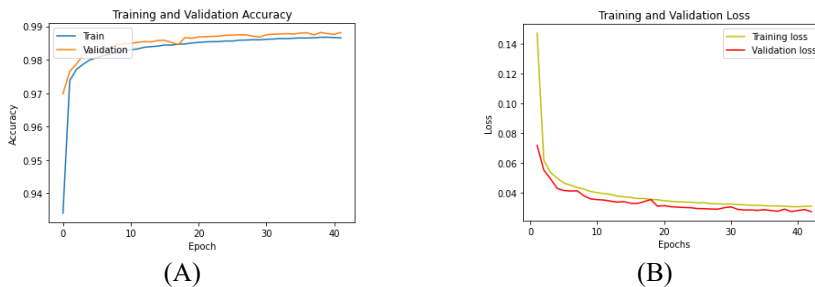


Figure 3. (A) accuracy and (B) loss data for training and validation data.

Other metrics were implemented to evaluate the network response: accuracy, recall, F1-Score, and MeanIoU. Accuracy evaluates the hits in each class considering all positive predictions; recall is the proportion of positives correctly classified as positives. The F1-score is a balancing metric between the response of precision and recall.

The MeanIoU metric was considered the most important for this work since it is the primary metric for semantic segmentation evaluation. It checked the overlap between the prediction and the template, considering the classification value of each pixel and its location. To calculate this metric, the number of pixels that were classified as true positives, which represented the intersection between the template and the prediction, was divided by the sum of all positive pixels in both the template and the prediction, which represented the union between them. The results of the metrics and the confusion matrix can be seen in Figure 4 and Tables 1 and 2.

	TARGET		PREDICT
	SAND	BACKGROUND	
SAND	98.35%	0.97%	
BACKGROUND	1.65%	99.03%	

	TARGET		PREDICT
	SAND	BACKGROUND	
SAND	98.36%	0.95%	
BACKGROUND	1.64%	99.05%	

Figure.4 Results for the (A) training and (B) test image set.

Table 1. Metrics by class for the training and test image set.

CLASS	TRAINING			TEST		
	Recall	Precison	F1-score	Recall	Precision	F1-score
<b>SAND</b>	0.983	0.973	0.978	0.983	0.973	0.978
<b>BACKGROUND</b>	0.994	0.990	0.992	0.994	0.990	0.992

Table 2. Global metrics for the training and test image set.

	TRAINING	TEST
Mean Accuracy	0.998	0.998
Mean IoU	0.973	0.942

The results showed high segmentation ability for both training and test data, so it can be concluded that the chosen parameters for the proposed U-Net were adequate for solving this problem. The adequate segmentation performance measured how many pixels of sand suspension and background were in the image and evaluated the location of each. The prediction of the network by image response was qualitatively verified, and the result is shown in Figure 5.

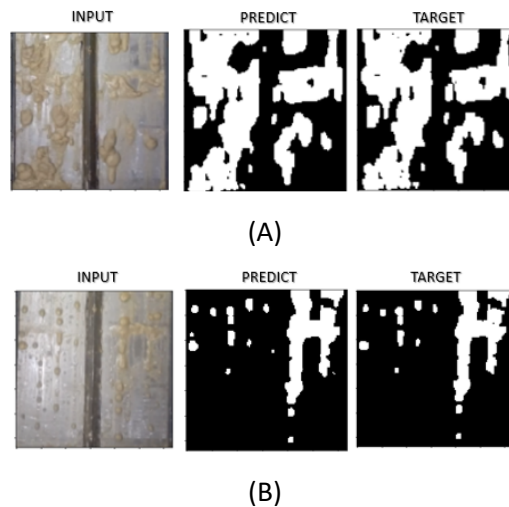


Figure 5. Segmentation examples provided by the trained U-Net for images from the (A) training and (B) test set.

Inputs from the training and test data sets were tested, and the prediction results showed that the U-Net appropriately described the expected template, showing its potential for feature learning in future predictions.

#### 4. Conclusion

It can be concluded that convolutional networks are powerful tools for image processing, presenting great results with minimal pre-processing procedures for pattern corrections. Parameter tuning is an essential step for the best neural network design. In the case of supervised learning, the quality of the templates is essential to ensure the reliability of the predictions. As a proposed sequence of the present work, the authors expect to evaluate the proposed U-Net in analyzing video images from actual industrial sieves operations for estimating useful process variables for monitoring the drilling process.

This study allows the extraction of operational information from the objects of interest. In the case of this application, the calculation of estimates of percentage filling of the cutout analysis and average flow velocity of the agglomerates is under development. These parameters are not currently measured for rig screens but are useful in the online identification of instabilities and well-collapse. Extension to other areas, such as mineral



and food processing, is possible with the estimation of other variables, such as flow rate and morphological properties of particles, at macro or micro scales, provided images are available.

### Acknowledgments

CAPES (Finance code 001), CENPES/PETROBRAS (Grant No. 4600580875), CNPq (Grant No. 311153/2021-6) and FAPERJ (Grant No. E-26/201.148/2022).

### References

1. GONÇALVES, L. B. **Intelligent image classification system of macroscopic rock for the oil and gas industry**. PhD Thesis (In Portuguese). Universidade Federal Fluminense. Niterói – RJ, Brazi. p. 309, 2010.
2. OSAROGIAGBON, A. U., KHAN, F., VENKATESAN, R., GILLARD, P. **Review and analysis of supervised machine learning algorithms for hazardous events in drilling operations**. *Process Safety and Environmental Protection*, Volume 147, 2021, Pages 367-384, ISSN 0957-5820, <https://doi.org/10.1016/j.psep.2020.09.038>.
3. MAGANA-MORA, A., AFFLECK, M., IBRAHIM, M., MAKOWSKI, G., KAPOOR, H., OTALVORA, W. C., JAMEA, M. A., UMAIRIN I. S., ZHAN, G., GOONERATNE, C. P. **Well Control Space Out: A Deep-Learning Approach for the Optimization of Drilling Safety Operations**. *IEEE Access*, vol. 9, pág. 76479-76492, 2021, <https://doi.org/10.1109/ACCESS.2021.3082661>.
4. LECUN, Y., BENGIO, Y. & HINTON, G. **Deep Learning**. *Nature* 521, 436-444 (2015).
5. IJJEH, A. A., ULLAH, S., KUDELA, P. **Full wavefield processing by using FCN for delamination detection**. *Mechanical Systems and Signal Processing*, Volume 153, 2021, 107537, ISSN 0888-3270, <https://doi.org/10.1016/j.ymssp.2020.107537>.
6. JOHN, L. D., ZHANG, C.. **An attention-based U-Net for detecting deforestation within satellite sensor imagery**. *International Journal of Applied Earth Observation and Geoinformation*, Volume 107, 2022, 102685, ISSN 0303-2434, <https://doi.org/10.1016/j.jag.2022.102685>.
7. SILBURT, ARI & ALI-DIB, MOHAMAD & ZHU, CHENCHONG & JACKSON, ALAN & VALENCIA, DIANA & KISSIN, YEYGENI & TAMAYO, DANIEL & MENOUE, KRISTEN. (2018). **Lunar Crater Identification via Deep Learning**. *Icarus*. 317. <https://doi.org/10.1016/j.icarus.2018.06.022>.
8. JING, J., WANG, Z., RÄTSCH, M., ZHANG, H. **Mobile-Unet: An efficient convolutional neural network for fabric defect detection**. *Article Information*. Volume: 92 issue: 1-2, page(s): 30-42. Article first published online: May 29, 2020; Issue published: January 1, 2022. <https://doi.org/10.1177/0040517520928604>
9. MCGLINCHY, B. JOHNSON, B. MULLER, M. JOSEPH AND J. DIAZ. **Application of UNet Fully Convolutional Neural Network to Impervious Surface Segmentation in Urban Environment from High Resolution Satellite Imagery**. *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 2019, pp. 3915-3918, <https://doi.org/10.1109/IGARSS.2019.8900453>.
10. BI, X., LIU, Y., XIAO, B., LI, W., PUN, C. M., WANG, G., GAO, X. **D-Unet: A Dual-encoder U-Net for Image Splicing Forgery Detection and Localization**. *Computer Vision and Pattern Recognition*. <https://doi.org/10.48550/arXiv.2012.01821>
11. CALDAS, G. L. R., BENTO, T. F. B., MOREIRA, R. M., SOUZA, M. B. **Quantifying Subsea Gas Leakages using Machine Learning: a CFD-based study**, Editor(s): Yoshiyuki Yamashita, Manabu Kano, *Computer Aided Chemical Engineering*, Elsevier, Volume 49, 2022, Pages 1345-1350, ISSN 1570-7946, ISBN 9780323851596, <https://doi.org/10.1016/B978-0-323-85159-6.50224-4>.
12. M. BARBOSA, V.P.; MENEZES, A.L.; GEDRAITE, R.; ATAÍDE, C.H. **Vibration screening: A detailed study using image analysis techniques to characterize the bed behavior in solid-liquid separation**. *Minerals Engineering*, Volume 154, 2020, 106383, ISSN 0892-6875

# An inverse optimization approach to decision-focused learning

Rishabh Gupta<sup>a</sup>, Qi Zhang<sup>\*a</sup>

<sup>a</sup>*Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455, USA*

*\* qizh@umn.edu*

## Abstract

Decision-focused learning is an emerging paradigm specifically aimed at improving the data-driven learning of input parameters to optimization models. The main idea is to learn predictive models that result in the best decisions rather than focusing on minimizing the parameter estimation error. Virtually all existing works on decision-focused learning only consider the case where the unknown model parameters merely affect the objective function. In this work, extend the framework to also consider unknown parameters in the constraints, where feasibility becomes a major concern. We address the problem by leveraging recently developed methods in data-driven inverse optimization, specifically applying a penalty-based block coordinate descent algorithm to solve the resulting large-scale bilevel optimization problem. The results from our computational case study demonstrate the effectiveness of the proposed approach and highlight its benefits compared with the conventional predict-then-optimize approach, which treats the prediction and optimization steps separately.

**Keywords:** Decision-focused learning, inverse optimization, constraint learning.

## 1. Introduction

In traditional data-driven optimization, we often follow a two-step predict-then-optimize approach, i.e. we first predict the unknown model parameters from data with external features and then solve the optimization problem with those predicted inputs. Here, the learning step focuses on minimizing the parameter estimation error; however, this does not necessarily lead to the best decisions (evaluated with the true parameter values) in the optimization step. In contrast, decision-focused learning (Wilder et al., 2019), also known as smart predict-then-optimize (Elmachtoub and Grigas, 2022), integrates the two steps to explicitly account for the quality of the optimization solution in the learning of the model parameters (i.e. minimize the decision error).

Existing works on decision-focused learning, many of which are based on deep learning with differentiable optimization layers (Amos and Kolter, 2017), have shown that significantly improved solutions can be achieved compared to the traditional predict-then-optimize approach. However, virtually all of them consider the case where the unknown model parameters only affect the objective function, which simplifies the problem considerably since feasibility is not a concern. Yet in many applications, we also need to use data to predict parameters in the constraints; the treatment of this case is in theory possible but difficult using existing methods. In this work, we address this problem by leveraging methods that we recently developed for data-driven inverse optimization (Gupta and Zhang, 2022a), which provide a natural way of incorporating constraints with unknown parameters.

The goal of inverse optimization is to infer an unknown optimization model from decisions that are assumed to be optimal solutions to that optimization problem (Chan et al., 2021). While early works primarily addressed the deterministic setting in which observations are assumed to be exactly optimal solutions of the optimization model, more recent contributions focus on the case with multiple noisy observations (Aswani et al., 2018; Chan et al., 2019; Gupta and Zhang, 2022b). Decision-focused learning can be viewed as a data-driven inverse optimization problem by treating the predictive model for the input parameters as the unknown part of the overall optimization model.

In the remainder of this paper, we first present the mathematical formulation of the decision-focused learning problem where we explicitly incorporate constraints that ensure feasibility of the optimal solutions obtained from the model with the estimated input parameters. To solve the resulting large-scale bilevel optimization problem, we apply our recently proposed penalty-based block coordinate descent algorithm. In a computational case study, we demonstrate the effectiveness of the proposed approach and highlight its benefits compared with the conventional predict-then-optimize approach.

## 2. Mathematical formulation

We assume that the optimization problem to be solved can be generally formulated in the following compact form:

$$\begin{aligned} & \text{minimize} && f(x, u) \\ & \text{subject to} && g(x, u) \leq 0, \end{aligned} \tag{1}$$

where  $x \in \mathbb{R}^n$  are the decision variables, and the model parameters are denoted by  $u$ . In this work, we assume that problem (1) is convex, with  $f$  and  $g$  being differentiable and convex in  $x$ . The model parameters  $u$  (or a subset of them) change with some external features  $r$  and are not exactly known; hence, they need to be estimated from data. The goal is to construct a predictive model  $u = m(r)$  given a set of  $N$  data points, where each data point  $i$  corresponds to a feature-output pair  $(\bar{r}_i, \bar{u}_i)$ . Given a new  $r$ , problem (1) will then be solved using the predicted values  $u = m(r)$ .

### 2.1. Conventional predict-then-optimize approach

In the conventional two-step process, the prediction of the model parameters is carried out independent from the later optimization. To obtain a predictive model  $m$ , one typically solves an empirical risk minimization problem of the following form:

$$\text{minimize} \quad \frac{1}{N} \sum_{i \in I} \ell(\bar{u}_i, m(\bar{r}_i)), \tag{2}$$

where  $I = \{1, \dots, N\}$  denotes the set of data points, and the loss function  $\ell$  is some measure of the difference between the true output  $\bar{u}_i$  and the prediction  $m(\bar{r}_i)$ . The underlying assumption is that if we minimize the difference between the true and predicted parameter values, this will also lead to optimal solutions to problem (1) that are the closest possible to the solutions we would obtain if we knew the true parameter values.

### 2.2. Decision-focused learning

In decision-focused learning, we integrate the prediction and optimization steps to construct a predictive model that directly takes the quality of the resulting optimization solution into account. We do so by solving the following problem:

$$\text{minimize } \frac{1}{N} \sum_{i \in I} f(\hat{x}_i, \bar{u}_i) \tag{3a}$$

$$\text{subject to } \hat{x}_i \in \arg \min_{\tilde{x}} \{f(\tilde{x}, u): g(\tilde{x}, u) \leq 0, u = m(\bar{r}_i)\} \quad \forall i \in I \tag{3b}$$

$$g(\hat{x}_i, \bar{u}_i) \leq 0 \quad \forall i \in I, \tag{3c}$$

where per constraints (3b),  $\hat{x}_i$  is an optimal solution to problem (1) with  $u = m(\bar{r}_i)$ . The objective is to minimize the true cost averaged over the training set  $I$ , i.e. it considers the cost of  $\hat{x}_i$  evaluated at the true parameter values  $\bar{u}_i$  for each  $i \in I$ . Importantly, constraints (3c) ensure feasibility of each  $\hat{x}_i$  given  $\bar{u}_i$ . These last set of constraints are omitted in virtually all existing works on decision-focused learning since they consider the case in which only the objective function  $f$  depends on  $u$  such that feasibility is not an issue.

### 3. Solution approach

The decision-focused learning problem (3) is a bilevel optimization problem with  $|I|$  convex optimization problems in its lower-level. We reformulate (3) into a single-level problem by replacing the lower-level problems with their KKT conditions. This results in a nonconvex nonlinear optimization problem which generally lacks regularization. To address the convergence difficulties of standard nonlinear solvers on this problem, we consider a penalty reformulation and apply an efficient block coordinate descent (BCD) algorithm. We do not provide more details about our solution algorithm here but refer the reader to Gupta and Zhang (2022a) for more details. We end this section by highlighting the fact that our approach is restricted to the case where problem (1) is a strictly convex problem and satisfies Slater’s condition.

### 4. Case study

In this section, we apply the proposed decision-focused learning approach to a (single-period) production planning problem for a small interconnected process network. This network, as depicted in Figure 1, consists of 5 materials and 3 processes. The goal is to determine the optimal quantities of raw materials to purchase and the amounts of products to manufacture to satisfy a given demands. This problem can be formulated as follows:

$$\text{minimize } z = \sum_{m \in \mathcal{M}} (\sum_{p \in \mathcal{P}} c_p y_p^2 + f_m w_m^2) \tag{4a}$$

$$\text{subject to } q_m^{\min} \leq q_m^0 + \left( \sum_{p \in \hat{\mathcal{P}}_m} \mu_{pm} y_p - \sum_{p \in \bar{\mathcal{P}}_m} \mu_{pm} y_p + w_m - d_m \right) \tag{4b}$$

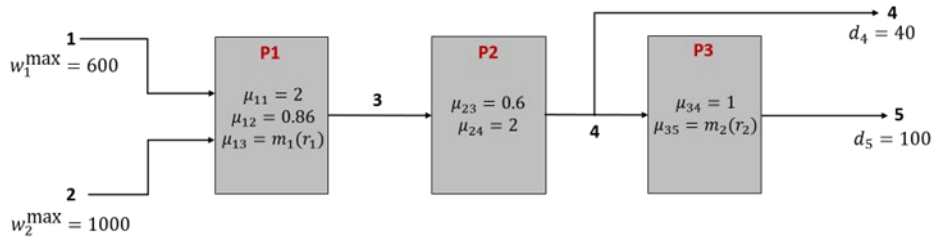
$$\leq q_m^{\max} \quad \forall m \in \mathcal{M}$$

$$0 \leq w_m \leq w_m^{\max} \quad \forall m \in \mathcal{M} \tag{4c}$$

$$0 \leq y_p \leq y_p^{\max} \quad \forall p \in \mathcal{P}, \tag{4d}$$

where  $\mathcal{M}$  and  $\mathcal{P}$  are the sets of materials and processes, respectively. Further, the set  $\hat{\mathcal{P}}_m$  consists of the processes that can produce material  $m$ , and the set  $\bar{\mathcal{P}}_m$  contains the processes that consume  $m$ . The amount of a reference material produced by process  $p$  is denoted by  $y_p$  and we use  $w_m$  to specify the amount of material  $m$  purchased from the market. The conversion factor  $\mu_{pm}$  determines the amount of a material  $m$  produced or consumed by process  $p$  for one unit of the reference material. Constraints (4b) restrict the inventory levels while accounting for product demand represented by  $d_m$ , (4c) limit the

amount of a material that can be acquired from the market, and (4d) set the capacities of processes. The objective is to minimize the total production and material purchasing cost.



**Figure 1** Process network for the production planning problem (4). The minimum and maximum allowed inventory values are 0 and 200, respectively, for all materials. For all processes, the values of  $y_p^{\max}$  is set to 400.

For this case study, we consider a scenario where the conversion factors vary based on some observable external feature  $r$ . For the sake of simplicity, we assume that the change in most of the conversion factors is negligible; only  $\mu_{13}$  and  $\mu_{35}$  deviate significantly enough from their nominal values to affect optimal production decisions. Our goal is to build predictive models for these two uncertain parameters using a data set containing observed  $(r_1, \mu_{13})$  and  $(r_2, \mu_{35})$  values.

4.1. Synthetic data generation

We now describe the process used to generate the training data set for the case study. We start by assigning models (5a) and (5b) to the uncertain parameters. These are the underlying true models which are assumed to be unknown. To obtain the training data, we sample  $|I|$  values of the features  $r_1$  and  $r_2$  from the uniform distributions  $\mathbb{U}(0.1, 0.45)$  and  $\mathbb{U}(-2, 1)$ , respectively. Following that, we evaluate the models for  $\mu_{13}$  and  $\mu_{35}$  at each of the sampled feature values to complete the training data set.

$$\mu_{13}(r_1) = \frac{1}{10} (\sin(20\pi r_1) + 7r_1) + 2 \tag{5a}$$

$$\mu_{35}(r_2) = 2 + \frac{1}{10} ((r_2 - 1) r_2 (r_2 + 2)^2) \tag{5b}$$

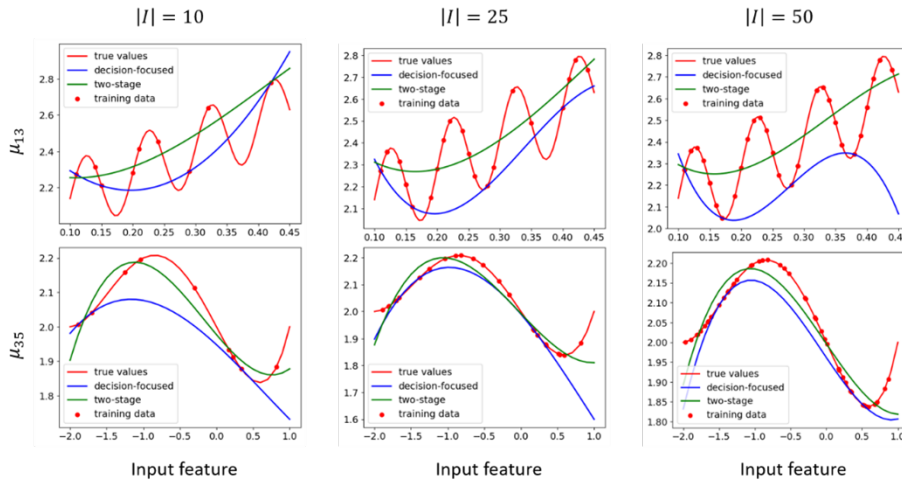
In order to estimate predictive models  $m_1$  and  $m_2$  with the proposed approach, we assume a hypothesis class consisting of cubic polynomials. We solve problem (3) using our BCD algorithm with training data sets of four different sizes: 10, 25, 50, and 100. The quality of the resulting model estimates is evaluated through a test data set of 100 unseen data points, which is generated using the same parameter generation scheme as the training data set. A model is considered good if the produced  $\hat{\mu}$  values result in production decisions that are not only close to the true optimal decisions but also feasible for the true model (i.e., problem (4) with the true  $\mu$  values).

In addition to the decision-focused approach, we also estimate cubic models for  $m_1$  and  $m_2$  using the traditional two-stage predict-then-optimize approach. Specifically, we use least squares regression to fit a cubic polynomial to the observed data. This estimated model is then used to solve problem (4) for the points in the test data set.

4.2. Results and discussion

Here we compare the performance of the models estimated using the decision-focused and two-stage approaches. Figure 2 compares the plots of the true  $\mu$  models with their

estimates obtained using the two approaches. In all cases, we find that the decision-focused approach constructs an underestimator function for the training data points. This happens because in problem (4), an optimal solution will always be such that inventories of all materials are close to or at their minimum values. If the estimated  $\hat{\mu}$  values are such that production gets overestimated, then there is a high probability that the inventory will fall below its permissible value when the process is actually run with a lower conversion value, resulting in an infeasible operation. Therefore, decision-focused learning obtains an underestimate of  $\mu$  to avoid violating the lower bound on the inventory constraint. From Figure 2, we find that as we provide more training data, the proposed approach finds better underestimators. With 50 data points, it is able to find almost perfect underestimators for both  $\mu$  parameters.



**Figure 2** True  $\mu$  models compared with their approximations estimated using the decision-focused and two-stage learning approaches

In contrast, the goal of the two-stage approach is to build the best approximation of the function itself using the provided data. The estimated model tries to closely mimic the behavior of the actual function to the extent that the assumed hypothesis class allows. This difference in approach leads to differences in performance, as seen in Table 1. The "feasible fraction" column indicates the fraction of the points in the test data set for which the estimated  $\mu$  models produced a feasible decision. The data in this column shows that the decision-focused approach significantly outperforms the two-stage approach. Moreover, one can see that for the decision-focused case, the better the estimated function underestimates the true function, the higher the fraction of the feasible points. Here since the two-stage approach does not focus on yielding underestimators, the fraction of feasible decisions is very low.

For the test data points that yield feasible decisions, Table 1 also compares the distance of those decisions from the true optimal solutions. While the decision-focused learning generates feasible solutions with a high degree of confidence, these solutions are slightly more different from the true optimal solutions compared to the two-stage approach. However, as can be seen from the last column in the table, which compares the optimality gap of the decisions generated by the two approaches, the mean optimality gap of the decision-focused approach is still less than 10% (compared to  $\sim 5\%$  in the two-stage case). This suggests that the decision-focused approach produces high quality decisions while

almost guaranteeing their feasibility.

$ I $	feasible fraction		mean $\frac{\ y(\mu) - y(\hat{\mu})\ _1}{\ y(\mu)\ _1}$		mean $\frac{\ w(\mu) - w(\hat{\mu})\ _1}{\ w(\mu)\ _1}$		mean $\frac{ z(\mu) - z(\hat{\mu}) }{z(\mu)}$	
	decision-focused	two-stage	decision-focused	two-stage	decision-focused	two-stage	decision-focused	two-stage
10	0.54	0.21	0.04	0.02	0.09	0.06	0.08	0.04
25	0.7	0.32	0.04	0.02	0.11	0.06	0.08	0.05
50	0.9	0.25	0.03	0.03	0.13	0.07	0.07	0.06
100	0.91	0.27	0.03	0.02	0.12	0.06	0.07	0.04

**Table 1.** A comparison of the performance of the models estimated using the decision-focused and two-stage approaches

## 5. Conclusions

In this work, we extended the decision-focused learning framework to include cases where the unknown parameters are in the constraints. We used an inverse optimization approach in which the problem is formulated as a bilevel program. Our approach allows inclusion of constraints that, with a high degree of confidence, ensure that the estimated model produces decisions that remain feasible for the true model. We illustrated our approach by applying it on a production planning problem with unknown process parameters. Our results show that the models obtained using decision-focused learning produce feasible decisions at a significantly higher rate compared to traditional two-stage learning without substantially sacrificing the optimality of these decisions.

## Acknowledgements

The authors gratefully acknowledge the financial support from the National Science Foundation under Grant #2044077. R.G. acknowledges financial support from a departmental fellowship sponsored by 3M and a Doctoral Dissertation Fellowship from the University of Minnesota.

## References

- Amos, B. and Kolter, J.Z., 2017. Optnet: Differentiable optimization as a layer in neural networks. *Proceedings of the International Conference on Machine Learning*, pp. 136-145.
- Aswani, A., Shen, Z.-J. M., and Siddiq, A., 2018. Inverse optimization with noisy data. *Operations Research*, 66(3), 870–892.
- Chan, T. C., Lee, T., and Terekhov, D., 2019. Inverse optimization: Closed-form solutions, geometry, and goodness of fit. *Management Science*, 65(3), 1115–1135.
- Chan, T.C., Mahmood, R., and Zhu, I.Y., 2021. Inverse optimization: Theory and applications. arXiv:2109.03920.
- Elmachtoub, A.N. and Grigas, P., 2022. Smart “predict, then optimize”. *Management Science*, 68(1), pp. 9-26.
- Gupta, R. and Zhang, Q., 2022a. Efficient learning of decision-making models: A penalty block coordinate descent algorithm for data-driven inverse optimization. arXiv:2210.15393.
- Gupta, R. and Zhang, Q., 2022b. Decomposition and adaptive sampling for data-driven inverse linear optimization. *INFORMS Journal on Computing*.

Wilder, B., Dilkina, B., and Tambe, M., 2019. Melding the data-decisions pipeline: Decision-focused learning for combinatorial optimization. *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 1658-1665.





# Multi-scale modeling and techno-economic analysis of biogas catalytic reforming for hydrogen & power production with CO<sub>2</sub> capture feature

Alessandra-Diana Selejan <sup>a</sup>, Simion Dragan <sup>a\*</sup>, Ana-Maria Cormos <sup>a</sup>, Mihaela Dragan <sup>b</sup>, Calin-Cristian Cormos <sup>a</sup>

<sup>a</sup> Faculty of Chemistry and Chemical Engineering, Babes-Bolyai University, 11 Arany Janos street, RO-400028, Cluj-Napoca, Romania

<sup>b</sup> Faculty of Economics and Business Administration, Babes-Bolyai University, Teodor Mihali street, No. 58-60, RO-400591, Cluj-Napoca, Romania

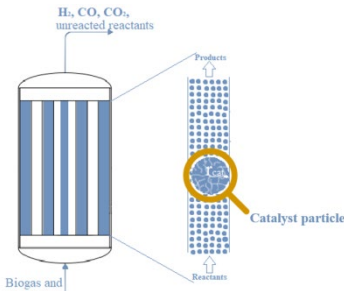
## Abstract

Hydrogen is one of the main pillars needed for a future low-carbon economy. The biogas steam reforming process can produce hydrogen without consuming fossil-based (non-renewable) resources. In this work, a dynamic 2D heterogeneous model (at macroscale and microscale) is developed for catalytic hydrogen production, applying the biogas steam reforming process. The model was implemented in Matlab/Simulink and the simulation results were compared with literature data, obtaining good agreement. The temperature profiles and those for the gas composition were obtained, related to the height of the catalyst layer in the reactor, as well as methane profile in the catalyst particle. A special emphasis was placed on highlighting the influence of a high content of carbon dioxide on the reactor design parameters. The dynamic behavior was also studied for ramp input changes on the feed flow. Based on mass and energy balanced derived from whole process simulation, the detailed techno-economic assessment of decarbonized biogas catalytic reforming plant for flexible hydrogen & power generation was assessed to evaluate the overall advantages.

**Keywords:** Biogas catalytic reforming; Green hydrogen production; Multiscale modeling; Techno-economic assessment; Negative CO<sub>2</sub> emissions.

## 1. Introduction

Hydrogen is seen as the fuel of the future, because it has the potential to be produced and used with near-zero greenhouse gas emissions. There are several industries using hydrogen, from fertilizer production and refining processes to methanol production and synthesis of organic compounds. Biogas got significant attention as a promising renewable energy source and an energy-efficient way of converting various biowastes into energy carriers. Steam methane reforming is the most used process for hydrogen production. By using biogas to produce hydrogen, the whole process can be considered environmentally friendly (IEA 2022, Lepage et al. 2021, Nahar et al. 2017). The steam reforming process, which uses biogas as feedstock for hydrogen production, takes place in a multi-tubular reactor with side furnaces flanking tubes filled with a packed-bed of catalyst (figure 1). The tubes are made of IN-519 alloy, while the catalyst consists of nickel particles on MgAl<sub>2</sub>O<sub>4</sub> support. The reactor operates at a temperature range of 700 to 1200 K and a pressure range of 28 to 37 bar. The inlet gas contains one-part biogas and at least three parts steam (Soliman et al. 1988).



**Figure 1.** The multi-tubular reforming reactor with catalyst

between the catalyst particles, inside of reactor' tubes represent the macroscale of the system. The model of the microscale described the pores inside the catalyst particle. A special emphasis was placed on highlighting the influence of a high content of carbon dioxide on the reactor design parameters. The dynamic behavior was also studied for step and ramp input changes on the feed flow. As other relevant research contributions of present work, the detailed technical and economic assessment of decarbonized biogas catalytic reforming process for flexible hydrogen & power generation was assessed.

## 2. Mathematical model

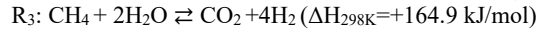
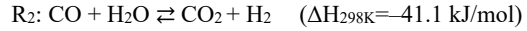
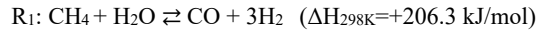
### 2.1. Model development

The multi-scale heterogeneous dynamic model of catalytic hydrogen production by steam reforming of biogas is developed based on following assumptions: along the length of the reactor the mass transfer through the catalyst layer takes place by convective and diffusional mechanism, while for the mass transfer through the catalyst pores, just the radial direction is considered, while it is thought that the catalyst particles have a homogenous porous structure. In the case of the catalyst phase, the mass balance equations are functions of time and catalyst particle radius. Any particle at the  $z$ -position is equivalent to all other particles at the same position. The partial differential equations describing the model are presented in Table 1 (Ghouse and Adams 2013).

**Table 1.** Mass and energy balances of the reforming reactor

Observations	Equation and boundary conditions
Mass balance for the gas phase	$\frac{\partial c_i}{\partial t} = -\frac{F}{A} \frac{\partial c_i}{\partial z} - k_{Ti} (c_i - c_{ci} _{r=R_p}) - r_i \rho_{cat}$
Energy balance for the gas phase	$\frac{\partial T}{\partial t} = -\frac{F}{A} \frac{\partial T}{\partial z} - \frac{\sum_{i=1}^N \Delta H_i k_i (c_i - c_{ci})}{c p_{mix} \rho_g} + \frac{K_T A_T (T_w - T)}{c p_{mix} \rho_g A dz}$ at $z=0$ and $t>0$ $c_i _{z=0} = c_{i,inlet}$ $T_g _{z=0} = T_{inlet}$
Mass balance for the catalyst phase	$\frac{\partial c_{ci}}{\partial t} = \frac{1}{\theta_c} \left\{ 2 D_{eimix} \frac{\partial c_{ci}}{\partial r} + \frac{\partial}{\partial r} \left[ D_{eimix} \frac{\partial c_{ci}}{\partial r} \right] - r_i \rho_{cat} \right\}$ at $t>0$ and any $z$ $\left[ \frac{\partial c_{ci}}{\partial r} \right]_{r=0} = 0$ $\left[ D_{eimix} \frac{\partial c_{ci}}{\partial r} \right]_{r=R_p} = k_{Ti} (c_i - c_{ci} _{r=R_p})$

There are three reactions involved in the endothermic process described above,  $R_1$  and  $R_2$  are parallel reactions, while  $R_3$  follows them (Calistru et al. 1984, Xu and Froment 1989, Ghouse and Adams 2013):



In this work, a dynamic 2D heterogeneous model (at macroscale and microscale) is developed for catalytic hydrogen production, applying the biogas steam reforming process. The spaces

**Table 2.** Process parameters

Inlet flow, $F$ – 23.271 [kmol/h]
Inlet temperature, $T_0$ – 723 [K]
Inlet pressure, $P$ – 30 [bar]
Inlet gas composition (mole fraction):
$x_{CH_4}$ : 0.1808
$x_{H_2O}$ : 0.7981
$x_{CO}$ : 0.0049
$x_{CO_2}$ : 0.0061
$x_{H_2}$ : 0.0098
Catalyst particle diameter, $D_p$ – 0.017 [m]
Catalyst void fraction, $\theta_c$ – 0.519
Tub diameter, $D_t$ – 0.107 [m]
Tub wall thickness, $X_t$ – 0.015 [m]
Packed-bed length, $L_{cat}$ – 9.18 [m]

The developed mathematical model was implemented in Matlab/Simulink, using process parameters presented in Table 2 (Calistru et al. 1984, Soliman et al. 1988, Ghouse and Adams 2013).

Since all 247 tubes present in the reforming reactor are operated and designed identically, the simulation was performed for a single tube (Calistru et al. 1984, Ghouse and Adams 2013). The connection between the two models (macroscale and microscale) have been made with multiple functions, repetitive and nested structures.

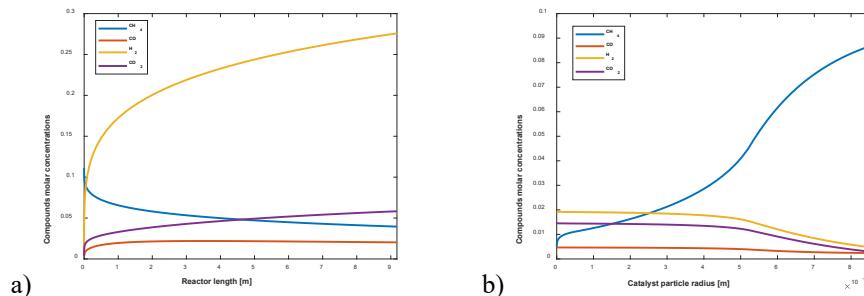
2.2. Results and discussions

**Table 3.** Model validation

Parameter	Model	Literature
CH <sub>4</sub> conversion	69.10%	71.66%
H <sub>2</sub> fraction	69.00%	70.60%

The predicted values of the developed mathematical model were compared with literature data in terms of CH<sub>4</sub> percentage conversion and H<sub>2</sub> mole fraction at exit in dry basis - Table 3 (Calistru et al. 1984, Ghouse and Adams 2013, Soliman et al. 1988).

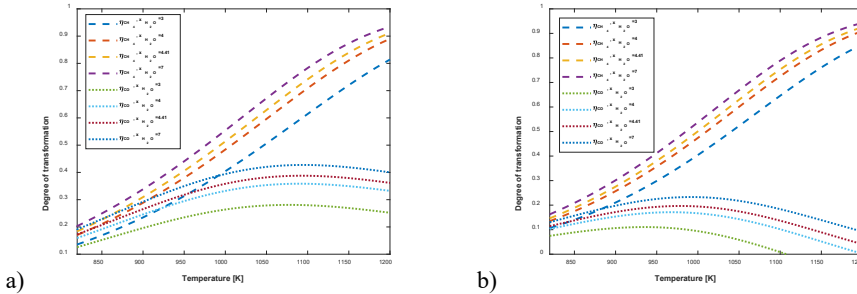
Based on simulations results, the profiles for the gas molar concentrations in the macroscale (reported to the length of the reforming reactor) and in the microscale (reported to the catalyst particle radius) were obtained (Figure 2). It can be observed that the hydrogen is produced manly in the first half of the reforming reactor length. In the case of the microscale, close to the interface, due to the mass transfer, the density of reactants is higher than into the particle center.



**Figure 2.** Variations of chemical species concentrations  
a) macroscale and b) microscale

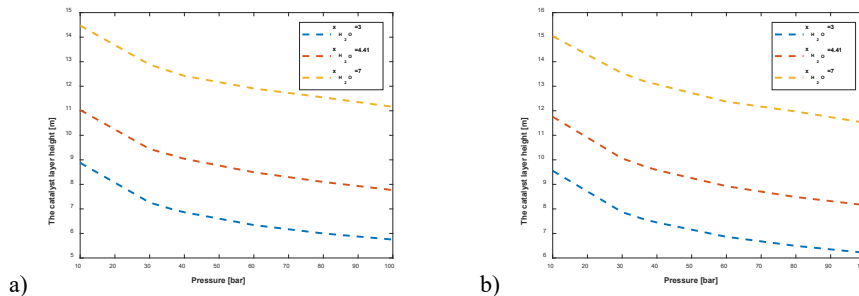
The biogas high content of carbon dioxide could influence the reactor design parameters. Simulations of the model were performed to observe how the degrees of transformation, and hence the conversions of the reactions are changed at a pressure of 30 bar, different  $\dot{x}_{H_2O} = H_2O_{(g)} : CH_4$  and  $\dot{x}_{CO_2} = CO_2 : CH_4$  molar ratios in the reactor feed gas, for a temperature between 820-1200 K. The upper temperature limit was chosen according to the limitations imposed by the material from which the reactor tubes are made. Comparing Figures 3a and 3b for the same molar ratio of  $H_2O_{(g)} : CH_4$ , reaction methane conversion ( $R_1$ ) is favoured by an excess of  $CO_2$ . If in Figure 3a the corresponding yellow line  $\dot{x}_{H_2O}=4.41$  barely reaches the value of 0.90 for  $\eta_{CH_4}$ , in

Figure 3b it exceeds it, reaching a value of 0.92. The carbon monoxide conversion ( $R_2$  reaction), it is strongly based by an excess of  $CO_2$ , (Figure 3a), for  $\dot{x}_{H_2O}=4.41$ ,  $\eta_{CO}=0.35$ , and in Figure 3b) for the same  $\dot{x}_{H_2O}$ ,  $\eta_{CO}<0.1$ .



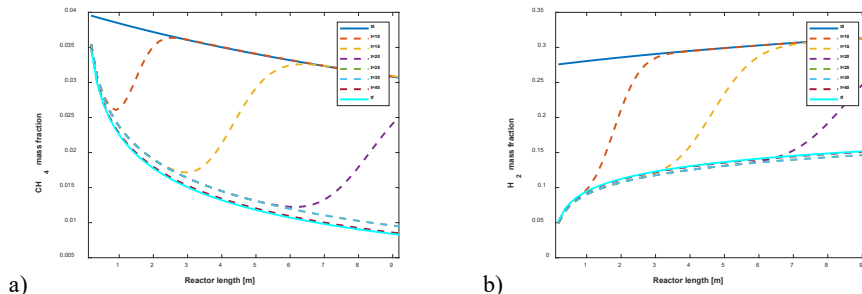
**Figure 3.** Variation of transformation degrees with temperature for a)  $P=30$  [bar]  $\dot{x}_{CO_2}=0.03$  and b)  $P=30$  [bar]  $\dot{x}_{CO_2}=0.75$

How an excess of  $CO_2$  in the reactor inlet flow requires a change in the height of the catalyst packed-bed to achieve the same results is presented in Figure 4. Although the excess  $CO_2$  is very large, the catalyst layer height increases by a maximum of one meter and the desired reaction is favored.



**Figure 4.** Variation of catalyst layer height with pressure at different  $\dot{x}_{H_2O}$  a)  $P=30$  [bar]  $\dot{x}_{CO_2}=0.03$  and b)  $P=30$  [bar]  $\dot{x}_{CO_2}=0.75$

The dynamic behaviour of the process was investigated by implementing a ramp signal disturbance of -20% on the biogas feed flow (Figure 5). The simulations showed that for the ramp signal disturbance, the system needs a long time to reach the new steady state values, as the disturbance propagated with a delay from one end of the reactor to the other. It is also observed that when the biogas feed flow rate is reduced, the amount of methane is reduced and the produced hydrogen quantity increases.



**Figure 5.** Dynamic profiles at -20% biogas feed flow a) Methane and b) Hydrogen

### 3. Techno-economic evaluation

The biogas catalytic reforming process for hydrogen & power co-generation with CO<sub>2</sub> capture feature was evaluated also in term of overall techno-economic performances. As decarbonization method, the pre-combustion CO<sub>2</sub> capture using reactive gas-liquid absorption with Methyl-D-Ethanol-Amine (MDEA) solvent was used. The overall process was modeled and simulated using ChemCAD software, the mass and energy balances being employed for calculation of main performance indicators. As illustrative concept, 100 MW thermal hydrogen output with 65% CO<sub>2</sub> capture rate were considered. For the economic evaluation methodology, the overall International Energy Agency - Greenhouse Gas R&D Programme method (reported in IEAGHG, 2017) was used. In terms of main economic assumptions, the ones defined in Cormos et al. (2022) was used. For instance, the capital costs were estimated using the cost correlation method with the most recent values of Chemical Engineering Plant Cost Index (CEPCI). The levelized costs of hydrogen and power were calculated based on Net Present Value (NPV) method with an internal rate of return of 8%. The CO<sub>2</sub> removal and avoidance costs were calculated using as benchmark a biogas reforming plant without CO<sub>2</sub> capture feature as reported in Cormos et al. 2022.

Table 4 presents the overall techno-economic and environmental performances. As can be noticed, the cumulative energy efficiency is high enough (66.50%) as well the decarbonization rate (65%). The economic parameters (e.g., hydrogen production cost, CO<sub>2</sub> capture costs) look promising in comparison to relevant studies (IEAGHG, 2017) based on natural gas as fuel input. For instance, the utilization of biogas (renewable source) into the decarbonized reforming plant gives overall negative CO<sub>2</sub> emissions.

**Table 4.** Techno-economic performance indicators

Parameter	Unit	Value
Biogas input flowrate	t/h	31.22
Lower heating value of biogas	MJ/kg	17.58
Thermal energy of biogas	MW <sub>th</sub>	152.45
Gross power production	MW <sub>e</sub>	8.00
Ancillary electricity consumption	MW <sub>e</sub>	6.50
Hydrogen thermal energy	MW <sub>th</sub>	100.00
Net power production	MW <sub>e</sub>	1.50
Hydrogen thermal efficiency	%	65.59
Net power efficiency	%	0.98
Cumulative energy efficiency	%	66.57
CO <sub>2</sub> capture rate	%	65.00
Specific CO <sub>2</sub> emissions	kg/MWh	172.50
Capital investment cost	M€	129.95
Specific capital investment cost	€/kW net	1280.00
Operational & maintenance cost	€/MWh	37.95
Levelized cost of hydrogen	€/MWh	61.42
Levelized cost of electricity	€/MWh	60.60
CO <sub>2</sub> removal cost	€/t	24.12
CO <sub>2</sub> avoidance cost	€/t	28.65

### 4. Conclusions

The 2D heterogeneous multiscale dynamic model for the biogas steam reforming process presented in this paper has been validated with data available in the literature. Particular attention was paid to how a higher CO<sub>2</sub> concentration influences the reactor design parameters while favoring the methane conversion (R<sub>1</sub>). The dynamic response of the system was highlighted by its response to a ramp disturbance of the feed flow.

When the feed flow rate decreases, the amount of methane decreases, but the amount of hydrogen increases in the reactor. In addition, a detailed techno-economic evaluation of hydrogen and power co-generation based on biogas reforming with CO<sub>2</sub> capture feature was done to illustrate its key advantages e.g., high energy efficiency, competitive hydrogen production cost, negative CO<sub>2</sub> emissions (due to CO<sub>2</sub> capture capability) etc.

### Acknowledgments

This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI - UEFISCDI, project number PN-III-P4-ID-PCE-2020-0632, within PNCDI III.

### Nomenclatures

$c_i$	concentration of component $i$ in the gas phase [kmol/m <sup>3</sup> ]
$c_{ci}$	concentration of component $i$ in the catalyst phase [kmol/m <sup>3</sup> ]
$k_{Ti}$	mass transfer coefficient [mol/m <sup>2</sup> s]
$A$	section area [m <sup>2</sup> ]
$r_i$	reaction rate of component $i$ [kmol/kg h]
$\rho_{cat}$	catalyst density [kg/m <sup>3</sup> ]
$T$	gas phase temperature [K]
$T_w$	temperature of the tube wall [K]
$\Delta H_i$	enthalpy of reaction [J/mol]
$k_i$	rate coefficient of reaction $i$
$cp_{mix}$	specific heat of the gas mixture [J/mol K]
$\rho_g$	gas mixture density [kg/m <sup>3</sup> ]
$D_{eimix}$	effective diffusivity of component $i$ inside the catalyst pores [cm <sup>2</sup> /s]
$\dot{x}_{H_2O}$	H <sub>2</sub> O : CH <sub>4</sub> moles ratio in the inlet stream
$\dot{x}_{CO_2}$	CO <sub>2</sub> : CH <sub>4</sub> moles ratio in the inlet stream

### References

- C. Calistru, C. Leonte, I. Siminiceanu, C. Hagi, O. Popa, 1984, Mineral Fertilizer Technology Vol. I -Nitrogen Fertilizers, Bucharest Technical Publishing House, Romania
- C.C. Cormos, A.M. Cormos, L. Petrescu, S. Dragan, 2022, Techno-economic assessment of decarbonized biogas catalytic reforming for flexible hydrogen and power production, Appl. Therm. Eng., 207, 118218
- G. Nahar, D. Mote, V. Dupont, 2017, Hydrogen production from reforming of biogas: Review of technological advances and an Indian perspective, Renewable and Sustainable Energy Reviews, 76, 1032-52
- International Energy Agency - Greenhouse Gas R&D Programme (IEAGHG), 2017, Techno-economic evaluation of SMR based standalone (merchant) hydrogen plant with CCS, Report 2017/02, Cheltenham, UK
- J. Xu, G.F. Froment, 1989, Methane Steam Reforming, Methanation and Water-Gas Shift: I. Intrinsic Kinetics, AIChE Journal Jan, 35(1), 88-96
- J.H. Ghose, T.A. Adams II, 2013, A multi-scale dynamic two-dimensional heterogeneous model for catalytic steam methane reforming reactors, International Journal of Hydrogen Energy, 38, 9984-99
- M.A. Soliman, S.S.E.H. El-Nashaie, A.S. Al-Ubaid, A. Adris, 1988, Simulation of steam reformers for methane, Chemical Engineering Science, 43(8), 1801-6
- T. Lepage, M. Kammoun, Q. Schmetz, A. Richel, 2021, Biomass-to-hydrogen: A review of main routes production, processes evaluation and techno-economical assessment, Biomass and Bioenergy, 144, 105920
- The International Energy Agency (IEA), 2022, Hydrogen IEA Paris, <https://www.iea.org/reports/hydrogen>; [accessed 1 December 2022].

# Modeling hierarchical systems via nested generalized disjunctive programming

Hector D. Perez,<sup>a</sup> Ignacio E. Grossmann<sup>a</sup>

<sup>a</sup>*Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh 15207, USA*

## Abstract

Modeling systems with discrete-continuous decisions is traditionally done in algebraic form with mixed-integer programming models, which can be linear or nonlinear in the continuous variables. A more systematic approach to modeling such systems is to use Generalized Disjunctive Programming (GDP), which extends the Disjunctive Programming paradigm proposed by Egon Balas. GDP allows modeling systems from a logic-based level of abstraction that captures the fundamental rules governing such systems via algebraic constraints and logic. Although GDP provides a more general way of modeling systems, it warrants further generalization for systems presenting a hierarchical structure. This work extends the GDP literature to address alternatives for modeling and solving systems with nested (hierarchical) disjunctions. We also provide two theorems on the relaxation tightness of such alternatives, establishing that a explicitly modeling nested disjunctions is superior to the traditional approach for dealing with nested disjunctions.

**Keywords:** generalized disjunctive programming, hierarchical systems, discrete-continuous optimization.

## 1. Introduction

Discrete-continuous optimization is one of the main modeling approaches for addressing design, planning, and scheduling problems in Process Systems Engineering (PSE) (Grossmann, 2012). Raman and Grossmann (1994) present a new modeling paradigm that extends the work done by Balas (1985) on disjunctive programming. This new paradigm, called Generalized Disjunctive Programming (GDP), was further developed by others in the PSE community over the years to account for additional features, such as nonlinearities and nonconvexities, in the problems encountered (Grossmann & Trespalacios, 2013). GDP relies on the intersection of disjunctions of convex sets to model the feasible space. Boolean variables are used as the indicator variables for each convex set, meaning that if True, the constraints in the disjunct are enforced. Logical constraints are also included to describe the relationships between the Boolean indicator variables via propositional logic and constraint programming logic.

The present work extends the GDP theory to allow modeling hierarchical systems, which are commonly encountered in PSE, and more particularly in Enterprise-Wide Optimization (EWO) (Grossmann, 2012), and flowsheet superstructure optimization (Türkay & Grossmann, 1996). Multi-level decisions can be concisely modelled via nested disjunctions. However, traditional GDP does not consider such formulations. Existing GDP literature suggests reformulating nested disjunctions into equivalent single-level disjunctions (Vecchiotti & Grossmann, 2000). Such an approach requires introducing



additional Boolean variables and logical propositions. An alternate approach is used in the work by van den Heever and Grossmann (1999), in which a direct or inside-out reformulation to MI(N)LP is performed. We formalize these two approaches and provide theoretical proofs on the tightness of their continuous relaxations. The model tightness and computational performance of the different approaches are compared. A series of examples are used to show the modeling and computational advantages obtained by explicitly modeling nested disjunctions.

## 2. GDP Modeling Overview

The classical GDP formulation is given below, where  $x$  is the set of continuous variables (bounded between  $x^{LB}$  and  $x^{UB}$ ),  $f(x)$  is the objective function,  $r(x) \leq 0$  is the set of global constraints,  $g_{ij}(x) \leq 0$  is the set of constraints applied when the indicator Boolean  $Y_{ij}$  is *True* for disjunct  $j$  in disjunction  $i$ .  $\Omega(Y)$  defines the set of logic constraints, which are described via propositional logic on a subset of Boolean variables. These constraints describe the interactions between the Boolean variables via clauses that contain with one or more of the following logic operators: AND ( $\wedge$ ), OR ( $\vee$ ), implication ( $\Rightarrow$ ), equivalence ( $\Leftrightarrow$ ), and negation ( $\neg$ ). The set of logic constraints may also include cardinality clauses of the form *choose exactly* (or *at least* or *at most*)  $m$  Boolean variables from a subset of Booleans to be *True* (Yan & Hooker, 1999). Cardinality clauses are indicated with the notation  $(Y_s \ \forall s \in S)_m$ , which means that at least  $m$  of the Boolean variables  $Y_s$  are *True*. An alternative is to use predicate logic to define the following predicates:  $\Xi(m, Y_s \ \forall s \in S)$  for *exactly*  $m$  are *True*,  $\Lambda(m, Y_s \ \forall s \in S)$  for *at least*  $m$  are *True*, and  $\Gamma(m, Y_s \ \forall s \in S)$  for *at most*  $m$  are *True*. GDP models typically include a cardinality clause to enforce that exactly 1 disjunct is chosen in each disjunction, i.e.,  $\Xi(1, Y_{ij} \ \forall j \in J_i) \ \forall i \in I$ . The GDP literature often uses the exclusive OR operator,  $\underline{\vee}$ , to define this constraint. However, such an operator is only correct for proper disjunctions (non-overlapping disjuncts). Thus, to avoid any ambiguity, we use the predicate logic notation,  $\Xi(1, Y)$ , here instead.

$$\begin{array}{ll}
 \min z = f(x) & \text{(GDP)} \\
 \text{s. t.} & r(x) \leq 0 \\
 & \bigvee_{j \in J_i} \left[ \begin{array}{l} Y_{ij} \\ g_{ij}(x) \leq 0 \end{array} \right] \quad \forall i \in I \\
 & \Xi(1, Y_{ij} \ \forall j \in J_i) \quad \forall i \in I \\
 & \Omega(Y) \\
 & x^{LB} \leq x \leq x^{UB} \\
 & x \in \mathbb{R}^n \\
 & Y_{ij} \in \{\text{True}, \text{False}\} \quad \forall i \in I, j \in J_i
 \end{array}$$

One of the main advantages of modeling discrete-continuous problems using GDP is the collection of methods that are available for optimizing such systems (Grossmann & Trespalcios, 2013; Trespalcios & Grossmann, 2016). These include, 1) reformulating to mixed-integer (non)linear models (MI(N)LP) via either Big-M or Hull reformulations, 2) logic-based decomposition methods such as Logic-based Outer Approximation (LOA), 3) disjunctive branch-and-bound, 4) basic steps, and 5) hybrid cutting planes. The reader is referred to the above references for a detailed understanding of each of these solution methods.

### 3. Extended formulation for hierarchies

#### 3.1. Hierarchical GDP

We propose extending the GDP paradigm to include multi-level decisions by means of nested disjunctions. Although the notion of nesting disjunctions to represent hierarchical decisions is not new, the limitations in the traditional GDP notation have made it difficult to exploit the benefits of nesting disjunctions. One of the first references to nested disjunctions is found in the work by Vecchietti and Grossmann (2000), which describes the transformations required to conform to the current GDP notation.

Therefore, from a model development point of view, the use of disjunction nesting is shown to add value. However, its implementation has often required breaking the explicit hierarchical structure. An exception is the work by Van Den Heever and Grossmann (1999), which does not transform the nested GDP into a logically equivalent single-level GDP, but rather suggests performing the hull reformulation on the inner disjunction and then reformulating the outer disjunction.

The proposed extension to the classical GDP notation for hierarchical systems is given below for a 2 level GDP (*2L-GDP*), where the upper-level decisions,  $Y$ , enforce the constraints  $g(x) \leq 0$  and the nested decisions,  $W$ , which have constraints  $h(x) \leq 0$ . Here the cardinality clause of selecting exactly one disjunct from the upper-level decisions,  $Y$ , is expressed explicitly, along with a new set of cardinality rules that enforce selecting exactly one of the lower-level decisions,  $W$ , if and only if the upper-level decision has been selected, and selecting no lower-level decisions when the upper-level decision is not selected. This constraint is expressed as the conjunction of two cardinality rules:  $[Y_{ij} \Rightarrow \exists(1, W_{ijkl} \forall l \in L_{ijk})] \wedge [\neg Y_{ij} \Rightarrow \exists(0, W_{ijkl} \forall l \in L_{ijk})] \forall i \in I, j \in J_i, k \in K_{ij}$ . In the GDP literature, this constraint has been traditionally written as  $Y_{ij} \Leftrightarrow \bigvee_{l \in L_{ijk}} W_{ijkl} \forall i \in I, j \in J_i, k \in K_{ij}$ . However, such a logic proposition is incomplete because it would allow the following to occur:  $Y_{ij} = False$  and  $W_{ijkl} = True$  for more than 1 index  $l \in L_{ijk}$ . As a result, the cardinality rule  $\Gamma(1, W_{ijkl} \forall l \in L_{ijk}) \forall i \in I, j \in J_i, k \in K_{ij}$  should be added to ensure that no more than 1 literal  $W_{ijkl}$  can be *True*. A more compact form would be to use the predicate constraint,  $\exists(t(Y_{ij}), W_{ijkl} \forall l \in L_{ijk})$ , where  $t(\cdot)$  is a unary function that maps a Boolean variable to its binary counterpart (i.e.,  $t(True) = 1$  and  $t(False) = 0$ ). For simplicity, we make a slight abuse of notation, dropping the mapping function  $t(\cdot)$ , and using the expression  $\exists(Y_{ij}, W_{ijkl} \forall l \in L_{ijk})$  instead. Such a model can be generalized to a multi-level hierarchical GDP with  $n$  levels.

$$\begin{aligned}
 \min z &= f(x) && (2L-GDP) \\
 \text{s. t. } & r(x) \leq 0 \\
 & \bigvee_{j \in J_i} \left[ \begin{array}{c} Y_{ij} \\ g_{ij}(x) \leq 0 \\ W_{ijkl} \\ \bigvee_{l \in L_{ijk}} \left[ h_{ijkl}(x) \leq 0 \right] \forall k \in K_{ij} \end{array} \right] && \forall i \in I \\
 & \exists(1, Y_{ij} \forall j \in J_i) && \forall i \in I \\
 & \exists(Y_{ij}, W_{ijkl} \forall l \in L_{ijk}) && \forall i \in I, j \in J_i, k \in K_{ij} \\
 & \Omega(Y, W) \\
 & x^{LB} \leq x \leq x^{UB}
 \end{aligned}$$

$$\begin{array}{ll}
x \in \mathbb{R}^n & \\
Y_{ij} \in \{True, False\} & \forall i \in I, j \in J_i \\
W_{ijkl} \in \{True, False\} & \forall i \in I, j \in J_i, k \in K_{ij}, l \in L_{ijk}
\end{array}$$

### 3.2. Equivalent Single-level GDP

Previous references to GDP with nested disjunctions in literature have proposed transforming the 2L-GDP model into the equivalent single-level GDP (2E-GDP) given below (Vecchiotti & Grossmann, 2000). Here, the nested disjunction is extracted and a dummy or “slack” disjunct is added to preserve feasibility. Thus, if none of the nested disjuncts is selected, the slack disjunct is selected, which contains the entire feasible set for  $x$ . The exclusive cardinality rule on the inner Boolean variables,  $W$ , is also augmented to include the slack Boolean variable,  $W_{ijk0}$ . This slack variable is, however, not included in the linking logic constraint for the upper and lower-level decisions. This ensures that the nested decisions are only selected if their master Boolean is *True*. This method for transforming a nested disjunction can also be applied to the multi-level system *ML-GDP*.

$$\begin{array}{ll}
\min z = f(x) & (2E-GDP) \\
s. t. \quad r(x) \leq 0 & \\
\bigvee_{j \in J_i} \left[ \begin{array}{l} Y_{ij} \\ g_{ij}(x) \leq 0 \end{array} \right] & \forall i \in I \\
\left( \bigvee_{l \in L_{ijk}} \left[ \begin{array}{l} W_{ijkl} \\ h_{ijkl}(x) \leq 0 \end{array} \right] \right) \bigvee \left[ \begin{array}{l} W_{ijk0} \\ x^{LB} \leq x \leq x^{UB} \end{array} \right] & \forall i \in I, j \in J_i, k \in K_{ij} \\
\Xi(1, Y_{ij} \forall j \in J_i) & \forall \square \in \square \\
\Xi(1, W_{ijkl} \forall l \in L_{ijk} \cup \{0\}) & \forall i \in I, j \in J_i, k \in K_{ij} \\
\Xi(Y_{ij}, W_{ijkl} \forall l \in L_{ijk}) & \forall i \in I, j \in J_i, k \in K_{ij} \\
\Omega(Y, W) & \\
x^{LB} \leq x \leq x^{UB} & \\
x \in \mathbb{R}^n & \\
Y_{ij} \in \{True, False\} & \forall i \in I, j \in J_i \\
W_{ijkl} \in \{True, False\} & \forall i \in I, j \in J_i, k \in K_{ij}, l \in L_{ijk}
\end{array}$$

Although the above formulation, allows modeling hierarchical systems in the standard GDP notation, it has two major drawbacks: 1) the explicit hierarchical structure is lost, and 2) although the equivalent single-level GDP model is logically equivalent to the nested GDP model, it requires introducing additional disjuncts and Boolean variables. Introducing “slack” disjuncts and Boolean variables results in models whose continuous relaxations are less tight, as given in the below theorem.

### 3.3. Tightness of Continuous Relaxations

The following two theorems address the advantages of modeling multi-level decisions problems via Hierarchical GDP, rather than the traditional single-level GDP approach. The advantages are shown by discussing the tightness of the continuous relaxations of these two approaches. For proofs of these theorems see Perez and Grossmann (2023).

*Theorem 1.* Let *rML-GDP-HR* denote the continuous relaxation of the mixed-integer program (MIP) obtained from a multi-level hierarchical GDP via the hull reformulation, and let *rME-GDP-HR* denote the continuous relaxation of the MIP obtained from the

equivalent single-level GDP via the hull reformulation. The feasible space of the former is contained within the feasible space of the latter, meaning  $rML-GDP-HR \subseteq rME-GDP-HR$ .

*Theorem 2.* Let  $rML-GDP-BM$  denote the continuous relaxation of the mixed-integer program (MIP) obtained from a multi-level hierarchical GDP via the Big-M reformulation, and let  $rME-GDP-BM$  denote the continuous relaxation of the MIP obtained from the equivalent single-level GDP via the Big-M reformulation. The feasible space of the former is contained within the feasible space of the latter, meaning  $rML-GDP-BM \subseteq rME-GDP-BM$ .

### 3.4. Graphical Example of Model Tightness

Consider the hierarchical (nested) GDP constraint system given in (1), which can be expressed as the equivalent single-level GDP in (2). Each of these models is reformulated to a MIP using the Big-M reformulation, with both a loose  $M$  value and a tight  $M$  value, and the Hull reformulation. Their continuous relaxations are then projected onto the  $x_1, x_2$  plane in *Figure 1*. The projections show that flattening via basic steps is advantageous when the hull reformulation is performed, but not necessarily when the Big-M reformulation is performed with a tight  $M$  value. Preserving the explicit hierarchical relationship via nested GDP reduces the feasible region of the continuous relaxation more than the equivalent single-level GDP approach in both the Big-M and hull reformulation cases. Furthermore, in this example the Big-M reformulation with a tight  $M$  on the nested GDP model produces the same relaxation as the hull reformulation on the equivalent GDP model with only a fraction of the model size. It should also be noted that, in this example, the convex hull of the system is obtained when either the hull reformulation is applied to the nested GDP or to the flattened GDP after two basic steps. The continuous relaxation of either of these two formulations will yield the optimum.

$$\left[ \begin{array}{c} Y_1 \\ 1 \leq x_1 \leq 3 \\ 4 \leq x_2 \leq 6 \end{array} \right] \vee \left[ \begin{array}{c} Y_2 \\ 8 \leq x_1 \leq 9 \\ 1 \leq x_2 \leq 2 \end{array} \right] \quad (1)$$

$$\left[ \begin{array}{c} W_1 \\ 1 \leq x_1 \leq 2 \\ 5 \leq x_2 \leq 6 \end{array} \right] \vee \left[ \begin{array}{c} W_2 \\ 2 \leq x_1 \leq 3 \\ 4 \leq x_2 \leq 5 \end{array} \right]$$

$$\Xi(1, \{Y_1, Y_2\})$$

$$\Xi(Y_1, \{W_1, W_2\})$$

$$\left[ \begin{array}{c} Y_1 \\ 1 \leq x_1 \leq 3 \\ 4 \leq x_2 \leq 6 \end{array} \right] \vee \left[ \begin{array}{c} Y_2 \\ 8 \leq x_1 \leq 9 \\ 1 \leq x_2 \leq 2 \end{array} \right] \quad (2)$$

$$\left[ \begin{array}{c} W_1 \\ 1 \leq x_1 \leq 2 \\ 5 \leq x_2 \leq 6 \end{array} \right] \vee \left[ \begin{array}{c} W_2 \\ 2 \leq x_1 \leq 3 \\ 4 \leq x_2 \leq 5 \end{array} \right] \vee \left[ \begin{array}{c} W_3 \\ 1 \leq x_1 \leq 9 \\ 1 \leq x_2 \leq 6 \end{array} \right]$$

$$\Xi(1, \{Y_1, Y_2\})$$

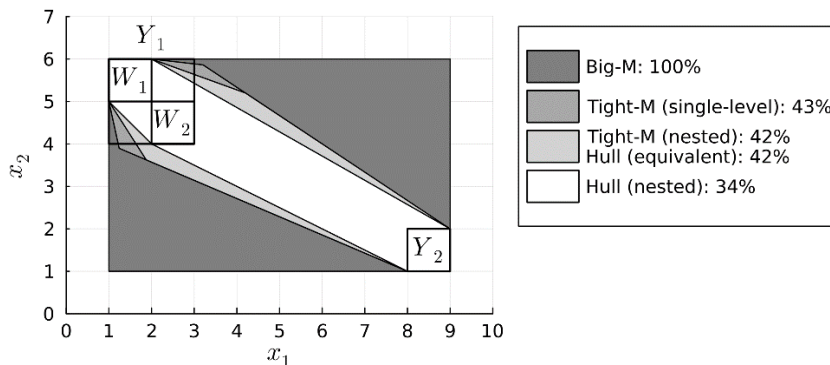
$$\Xi(1, \{W_1, W_2, W_3\})$$

$$\Xi(Y_1, \{W_1, W_2\})$$

## 4. Conclusions

The contribution of this paper has been to extend GDP for modeling hierarchical systems via nested disjunctions. Such an approach results in more intuitive models, but has not

been formalized in the past, as classical GDP does not consider disjunction nesting. The notation and logic constraints for such structures is provided, along with theoretical proofs to the tightness of such models, versus equivalent single-level GDP models. It is shown that mixed-integer programming reformulations of nested GDP models have tighter relaxations than the reformulations of their single-level counterparts in both the hull reformulation, as well as the Big-M reformulation when tight  $M$  values are used.



**Figure 1.** Projections of the continuous relaxations of Example 4.1 for different reformulations (Big-M = Big-M Reformulation, Tight-M = Big-M Reformulation with tight  $M$  values, Hull = Hull Reformulation; equivalent = equivalent single-level GDP, nested = multi-level hierarchical GDP). Projection areas, relative to Big-M are indicated in %.

## Acknowledgement

The authors gratefully acknowledge the financial support from the Center of Advanced Process Decision-making at Carnegie Mellon University.

## References

- E. Balas, 1985, Disjunctive Programming and a Hierarchy of Relaxations for Discrete Optimization Problems, *SIAM J. Alg. Disc. Meth.*, 6(3), 466–486.
- I.E. Grossmann, 2012, Advances in mathematical programming models for enterprise-wide optimization, *Comput. Chem. Eng.*, 47, 2–18.
- I.E. Grossmann, F. Trespalcios, 2013, Systematic modeling of discrete-continuous optimization models through generalized disjunctive programming. *AIChE J.*, 59(9), 3276–3295.
- H.D. Perez, I.E. Grossmann, 2023, An Extension to Generalized Disjunctive Programming for Modeling Hierarchical Systems, submitted for publication.
- R. Raman, I.E. Grossmann, 1994, Modelling and computational techniques for logic based integer programming, *Comput. Chem. Eng.*, 18(7), 563–578.
- F. Trespalcios, I.E. Grossmann, 2016, Cutting Plane Algorithm for Convex Generalized Disjunctive Programs, *INFORMS J. Comput.*, 28(2), 209–222.
- M. Türkay, I.E. Grossmann, 1996, Logic-based MINLP algorithms for the optimal synthesis of process networks, *Comput. Chem. Eng.*, 20(8), 959–978.
- S.A. Van Den Heever, I.E. Grossmann, 1999, Disjunctive multiperiod optimization methods for design and planning of chemical process systems, *Comput. Chem. Eng.*, 23(8), 1075–1095.
- A. Vecchietti, I.E. Grossmann, 2000, Modeling issues and implementation of language for disjunctive programming, *Comput. Chem. Eng.*, 24(9–10), 2143–2155.
- H. Yan, J.N. Hooker, 1999, Tight representation of logical constraints as cardinality rules. *Math. Program.*, 85(2), 363–377.

# Fusion and integrated correction of chemometrics and machine learning models based on data reconciliation

Pál P. Hanzelik,<sup>a,b</sup> Alex Kummer,<sup>b</sup> Ádám Ipkovich,<sup>b</sup> János Abonyi<sup>b</sup>

<sup>a</sup> *Dombóvári út 28, Budapest, Hungary*

<sup>b</sup> *Egyetem Street 10, Veszprém, Hungary*

## Abstract

Soft-sensors can provide real-time estimates of unmeasured process variables to improve process control and performance. A Multi-Input-Multi-Output system is considered, where the outputs need to satisfy a linear constraint (e.g. balance equations), while the output variables are predicted using machine learning (ML) models. The key idea is that data reconciliation (DR) can be used to modify both the measurements and the model predictions to meet the prescribed constraints. In addition, the model prediction errors can be used as weighting factors for data reconciliation.

The paper summarizes the mathematical derivation and application of the approach to predict the composition of various rocks in an industrial environment. The results show that the proposed, integrated ML+DR approach greatly contributes to the models' usability. Furthermore, we obtained smaller model errors in four cases out of the six properties, and the prediction sum also considers the constraints.

**Keywords:** machine learning, algorithm development, data reconciliation, chemometrics, model prediction

## 1. Introduction

This paper proposes a methodology that can improve ML models applied in industrial processes and help optimization based on digitization tools. Moreover, during the Industry 4.0 processes, software sensors' use is becoming more widespread [1]. The motivation of the work is that the use of correlations between process parameters can make the industrial application of machine learning algorithms more efficient and robust. These options include various ML algorithms that can replace traditional quality assurance processes. During the development of ML models of Quality 4.0 techniques, we can estimate only specific parameters with sufficient accuracy in many cases. The key idea is that in many cases the variables to be estimated are not independent of each other, for example they have to satisfy some balance equation, but these constraints are not taken into account when teaching ML models. We recommend a data reconciliation technique to meet the limits, thus increasing the applicability of ML models.

The DR technique has already been used for energy monitoring chemical processes, which considers the balance equations' consistency. In addition to the DR method, principal component analysis (PCA) can also use to forecast operating costs. The advantage of this method is that if we apply the PCA technique, the reliability of the data increases significantly, which increases the accuracy of this estimated cost.[2], furthermore, DR and PCA are also excellent for error detection [3] [4]. Xie et al.

investigated the industrial application of sodium-aluminium solution vaporization. They presented that DR is a key technique in improving the accuracy of measured and predicted data in the case of a chemical process [5]. The DR-based approach results in better data quality and, thus, fewer operational errors [6]. Furthermore, DR is used in the control design of online processes, where the goal is to estimate the valid values of erroneous measurements within limits. Non-linear DR was used in material flow analysis for the systematic and efficient management of material flow and inventory [7] or even for eliminating errors caused by measurement noise in the control of reactor processes [8]. For example, the hierarchical DR provides a high-quality measurement of predicting a natural industrial system's material and heat energy consumption [9]. In addition, DR can be used in 4.0 quality, and there is a large literature on Fourier transform infrared spectroscopy (FTIR) measurement techniques and dielectric spectroscopy techniques for predicting the concentration of different metabolites. DR using FTIR belongs to Dabros et al. according to them, the predictive performance of the primary models was found to be poor, but with DR techniques the problem can be improved by real-time mass and elemental balance [10].

We present a machine learning and data reconciliation-based approach (see Section 2) to predict elemental composition of rocks, where the predictions need to satisfy a linear constraint (a balance equation). Furthermore, since the model predictions are loaded with errors, the standard deviation of the model errors can be used in data reconciliation to modify the predicted values. The accuracy of the ML models with and without using DR are compared (see Section 3).

## 2. Data reconciliation-based integration of machine learning models

This section introduces the methodology of our approach to integrate ML and DR techniques to improve the quality of ML predictions. The goal is to predict several related variables, where the predictions need to satisfy some constraints, e.g. balance equations. In order to achieve the highest accuracy of the ML predictions, and to fulfil the constraints, we recommend to integrate ML and DR techniques.

We consider a Multi-Input-Multi-Output system with  $N$  samples and  $N_m$  measurement variables. The measurement outputs of the  $k$ -th sample can be written as in Equation 1.

$$\mathbf{y}_k = f(\mathbf{x}_k, \boldsymbol{\theta}) + \boldsymbol{\varepsilon}_k \quad (1)$$

where  $\mathbf{y}_k = [\mathbf{y}_k^{(1)}, \dots, \mathbf{y}_k^{(N_m)}]^T$  is the measurement vector of the  $k$ -th sample,  $\mathbf{x}_k$  is the input features of the  $k$ -th sample, and  $\boldsymbol{\varepsilon}_k$  is the measurement error at the  $k$ -th sample. The  $k$ -th sample of the  $i$ -th output variable can be defined as Equation 2.

$$y_k^{(i)} = f_i(\mathbf{x}_k^{(i)}, \boldsymbol{\theta}^{(i)}) + \varepsilon_k^{(i)} \quad (2)$$

where  $y_k^{(i)}$  is the  $i$ -th output variable as the  $k$ th sample,  $f_i$  is the  $i$ -th model,  $\mathbf{x}_k^{(i)}$  stands for the input features for the  $i$ th model at the  $k$ th sample,  $\boldsymbol{\theta}^{(i)}$  is the model parameters of the  $i$ th model and  $\varepsilon_k^{(i)}$  represents the model error for the  $i$ th model at the  $k$ th sample.

The prediction of the model at the  $k$ th sample is represented as  $\hat{\mathbf{y}}_k = [\hat{y}_k^{(1)}, \dots, \hat{y}_k^{(N_m)}]^T$ , and the prediction of the  $i$ th model at the  $k$ th sample ( $\hat{y}_k^{(i)}$ ) can be calculated as presented in Equation 3.

$$\hat{y}_k^{(i)} = f_i(\mathbf{x}_k^{(i)}, \boldsymbol{\theta}^{(i)}) \quad (3)$$

The measurement variables are not independent of each other, because they need to satisfy a linear constraint, while the measurements are loaded with errors ( $e_k$ ), hence this constraint is not fulfilled.

$$g(\mathbf{y}_k) + e_k = 0 \quad (4)$$

where  $g()$  is an equality constraint function.

Data reconciliation calculates minimal correction of the predicted variables to make them verify a set of model constraints. It minimizes the difference between the predicted and reconciled values considering the variance of the variables and ensures that reconciled parameters satisfy some equality and inequality constraints. Similarly, the corrected model predictions for the  $k$ th sample are represented as  $\tilde{\mathbf{y}}_k = [\tilde{y}_k^{(1)}, \dots, \tilde{y}_k^{(N_m)}]^T$ .

$$\min_{\tilde{\mathbf{y}}_k} (\hat{\mathbf{y}}_k - \tilde{\mathbf{y}}_k)^T \mathbf{Q} (\hat{\mathbf{y}}_k - \tilde{\mathbf{y}}_k) \quad (5)$$

$$\text{s.t.} \quad g(\tilde{\mathbf{y}}_k) = 0 \quad (6)$$

where  $\tilde{\mathbf{y}}_k$  is the corrected value by data reconciliation, the  $\mathbf{Q}$  is the inverse covariance matrix of the measurements/predictions error, the  $g(\tilde{\mathbf{y}}_k)$  is a model equality constraint. In the inverse covariance matrix  $\mathbf{Q}$  the standard deviation of the measurement errors are presented, as in Equation 7.

$$\mathbf{Q} = \begin{bmatrix} 1/\sigma_1^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1/\sigma_{N_m}^2 \end{bmatrix} \quad (7)$$

In case of a linear constraint Equation 5 becomes the following:

$$\mathbf{A}\tilde{\mathbf{y}}_k = \mathbf{b} \quad (8)$$

where  $\mathbf{A}$  is the incidence matrix, and  $\mathbf{b}$  stands for constraint values.

The analytical solution of the system of Equations 5 and 8 is the following:

$$\tilde{\mathbf{y}}_k = (\mathbf{I} - \mathbf{Q}\mathbf{A}^T(\mathbf{A}\mathbf{Q}\mathbf{A}^T)^{-1}\mathbf{A})\hat{\mathbf{y}}_k + \mathbf{Q}\mathbf{A}^T(\mathbf{A}\mathbf{Q}\mathbf{A}^T)^{-1}\mathbf{b} \quad (9)$$

where  $\mathbf{I}$  is the identity matrix,  $\mathbf{Q}$  is the inverse covariance matrix,  $\mathbf{A}$  is the incidence matrix and  $\mathbf{b}$  is the constraint values.

In the developed methodology, industrial data collection in a data store provides input data to the ML models. The essence of the developed methodology is that various DR



restrictions are applied after the models are incorporated, which enables more accurate and robust ML models. As a result, we can build more effective multivariate data analysis on our industrial processes. Figure 1 shows in a general diagram how the developed technique can be applied in an industrial environment. The essential part of the methodology is shown with the orange arrows, and the blue dashed line shows the traditional approach in the following figure.

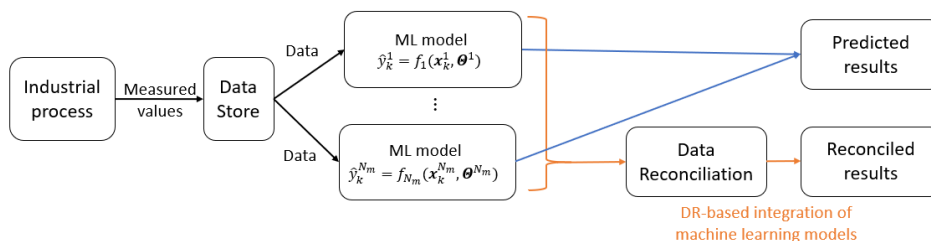


Figure 1: Diagram of the recommended method

### 3. Data reconciliation-based integration of machine learning models

A mineral composition prediction model was built on 618 rock samples, where the target variables were mass percentage of quartz, clay content without kaolinite, kaolinite, plagioclase with K-feldspar, calcite and dolomite. The reference measurements were performed with an X-ray diffraction (XRD) device. PLS models were trained on pretreated spectra from FTIR measurements, and individual models were trained for each of the six target variables. Optimal number of factors for PLS models were tuned based on the results of 10-fold cross-validation. The ML input spectra of samples were scanned (4 scans co-added) from 1998 to 600  $\text{cm}^{-1}$  in attenuated total reflectance (ATR) mode by using a Spectrum 400 spectrometer equipped with a Universal ATR (UATR) accessory, which contains a single reflection diamond/ZnSe composite crystal.

Table 1 summarizes the target values of a chosen sample, the predictions (ML(PLS)) and the reconciled predictions (ML(PLS)+DR).

Table 1. Prediction and reconciled results a chosen sample (weight%)

	Quartz	Clay without kaolinite	Kaolinite	Plagioclase + K-feldspar	Calcite	Dolomite	sum
Target	49	16	1	17	10	7	100
ML (PLS)	49.39	17.95	1.637	14.32	10.51	8.053	101.86
ML(PLS)+DR	49.21	17.55	1.085	14.19	10.36	7.605	100

The reconciled predictions for five of the six output variables were better compared to the non-reconciled model predictions. Table 2 summarizes the results obtained for all 618 rock samples, as well as the standard deviation of the reference measurement. The performance indicators of the models were the cross-validation error ( $RMSE_{cv}$ ).

Table 2. Reference measuring standard deviation, optimal factor number of the PLS,  $RMSE_{cv}$  (%) results of ML and ML + DR models

	Quartz	Clay without kaolinite	Kaolinite	Plagioclase + K-feldspar	Calcite	Dolomite
Factor number	8	5	9	9	5	8
ML(PLS) $RMSE_{cv}$	5.934	6.265	1.467	3.263	4.173	2.414
ML(PLS) + DR $RMSE_{cv}$	5.924	6.232	1.624	3.260	4.172	2.481

The values of the six target variables ranged from 1.47 to 6.27 with the ML solution and from 1.62 to 6.23 with the ML+DR technique. The accuracy of the ML models constructed from FTIR spectra approaches the uncertainty of the reference measurement in two cases (kaolinite, dolomite), and the ML+DR models give better results when the six properties are considered together. In the latter case, the ML+DR models perform worse for kaolinite and dolomite (Table 2). This can be explained by the fact that these models already performed very close to the accuracy of the reference models.

#### 4. Conclusion

The article proposes a methodology that improves the usability of ML models by reconciling the model predictions. The proposed methodology takes into account the prediction error of the ML models, by which the DR technique corrects the prediction values while satisfying certain constraints. Since in many cases the variables to be estimated are not independent of each other, for example, they correspond to balance equations. However, these constraints should be fulfilled by ML models. We present a methodology on how to integrate machine learning and data-reconciliation, and we tested the approach to estimate the elemental composition of rock samples.

With the inclusion of the constraints, the results showed more accurate outputs for four of the six target variables. The results will help geologists by enabling them to obtain mineral composition data of rocks. Furthermore, the presented methodology allows faster and cheaper access to compositional data, either with a mobile device or with devices installed in a soft sensor.

Our future research is to take constraints into account using DR during the training of the ML models, rather than performing the training and data reconciliation serially.

#### References

1. P. P. Hanzelik, A. Kummer, J. Abonyi, 2022, Edge-Computing and Machine-Learning-Based Framework for Software Sensor Development, *Sensors*, 22, 4268.
2. B. Farsang, S. Nemeth, J. Abonyi, 2014, Synergy between data reconciliation and principal component analysis in energy monitoring. *Chemical Engineering*, 39.
3. T. Amand, G. Heyen, B. Kalitventzeff, 2001, Plant monitoring and fault detection: Synergy between data reconciliation and principal component analysis. *Computers & Chemical Engineering*, 25(4-6), 501-507.

4. B. Farsang, S. Nemeth, J. Abonyi, 2014, Life-cycle Modelling for Fault Detection–Extraction of PCA Models from Flowsheeting Simulators. In *Computer Aided Chemical Engineering*, 33, 421-426, Elsevier
5. S. Xie, C. Yang, X. Yuan, X. Wang, Y. Xie, 2018, Layered online data reconciliation strategy with multiple modes for industrial processes. *Control Engineering Practice*, 77, 63-72.
6. X. Jiang, P. Liu, Z. Li, 2013, A data reconciliation based approach to accuracy enhancement of operational data in power plants. *Chemical Engineering Transactions*, 35, 1213-1218.
7. J. H. Taylor, R. del Pilar Moreno, 2013, Nonlinear dynamic data reconciliation: In-depth case study. In 2013 IEEE international conference on control applications (CCA) (pp. 746-753). IEEE.
8. M.J. Leibman, T.F. Edgar, L.S. Lasdon, 1992, Efficient data reconciliation and estimation for dynamic processes using nonlinear programming techniques, *Computers & Chemical Engineering*, 16, 10-11, 963-986
9. S. Xie, H. Wang, J. Peng, X. Liu, X. Yuan, 2020, A hierarchical data reconciliation based on multiple time-delay interval estimation for industrial processes. *ISA transactions*, 105, 198-209.
10. M. Dabros, M. Amrhein, D. Bonvin, I. W. Marison, U. von Stockar, 2009, Data reconciliation of concentration estimates from mid-infrared and dielectric spectral measurements for improved on-line monitoring of bioprocesses. *Biotechnology progress*, 25(2), 578-588.

# Model-based Optimisation of Regional Nutrient Flow and Recovery for Resource and Environmental Sustainability

Purusothmn Nair S Bhasker Nair,<sup>a</sup> Nan-Hua Nadja Yang,<sup>a</sup> Wei Zhang,<sup>a</sup> Kok Siew Ng,<sup>a,b</sup> Aidong Yang<sup>a</sup>

<sup>a</sup> *Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, United Kingdom.*

<sup>b</sup> *Department of Chemical Engineering, College of Engineering, Design and Physical Sciences, Brunel University London, Kingston Lane, Uxbridge UB8 3PH, United Kingdom.*

## Abstract

The unsustainably managed flow of nutrients combined with the rise in the generation of waste calls for urgent actions involving nutrient recovery. This work develops a mathematical programming model intended to optimise regional nutrient recovery from various waste treatment technologies. Food waste, green waste, and wastewater are evaluated in this work with an option of treating these wastes individually or as mixed streams. Five waste treatment technologies are available, each with a specific profile of nutrient recovery. The results of this work demonstrate the attractiveness of anaerobic digestion as well as the benefit of a mixed waste stream for optimum nutrient recovery.

**Keywords:** Mathematical Optimisation, Nutrient Recovery, Circular Economy, Climate Change, Sustainable Resource Management

## 1. Introduction

Nutrients such as nitrogen and phosphorus are vital for food production. Presently, there is a vast usage of inorganic fertilisers derived from ammonia and phosphate (van der Wiel et al., 2020). Phosphorus is solely sourced from phosphate rocks which are finite resources and limited in terms of their geographic availability (Álvarez et al., 2018) while nitrogen in ammonia production represents one of the major industrial energy demands and sources of greenhouse gas emissions (Liu et al., 2020). Furthermore, the existing managed flow of phosphorus and nitrogen from its source (mines) to the ocean or other aquatic environments is not sustainable (Cordell and White, 2015). The current farming and waste management practices create a significant nutrient influx to water streams, resulting in undesirable consequences e.g., eutrophication (Cordell and White, 2015). Aside from that, nutrient depletion due to current practices would induce instability in the food systems, thus driving up the costs of food production. A key solution to these problems lies in innovative waste management. In the United Kingdom, 27 million tonnes of household waste were produced in the year 2018 (DEFRA, 2022). Although close to half of this waste is recycled, much more needs to be done to align with environmental sustainability and thus create a circular economy (DEFRA, 2022). With the rise in population, the total volume of waste generated annually is projected to rise. Nutrients such as nitrogen and phosphorus should be extracted from these waste streams as they may be re-utilised as fertilisers.

In the past, several studies concerning nutrient recovery had been conducted (Mehta et al., 2014; Perera et al., 2019; Sniatala et al., 2023), which were all focused on technological advancement for nutrient recovery. There was a limited focus in terms of modelling and optimising nutrient flows from their sources (water, food, and green waste) to potential sinks e.g., land dispersion, fertilisers etc. Despite some work being done recently for optimising nutrient recovery, much of the focus was associated with transportation costs (Metson et al., 2020) and biogas generation from anaerobic digestion (Feiz et al., 2022). There has been limited emphasis on optimising nutrient recovery from these waste treatment technologies. This work, therefore, seeks to model and optimise the regional nutrient flows and maximise recovery from waste streams into valuable products such as fertilisers. Firstly, the flows of wastewater, food and green wastes are mapped to understand their resource recovery potential in creating a circular economy. Based on the available data, an optimisation model is subsequently built toward determining potential opportunities that may be present for nutrient recovery. Potential technologies such as anaerobic digestion, composting and incineration are considered in the optimisation model. This work makes use of a case study from Leicestershire County (LC) in the UK and is expected to consolidate efforts toward building a circular economy by tapping into the underexplored potentials of nutrients embedded in the waste streams, thus promoting a sustainable practice of nutrient management in a regional setting.

## 2. Problem Statement

Given a set of waste streams i.e., food waste, green waste and wastewater that may be treated by processing unit (PU)  $j$  as either an individual waste (IW) stream  $i$  or mixed waste (MW) stream  $im$ . The treatment of IW  $i$  and MW  $im$  by PU  $j$  results in the recovery of nutrient  $l$ . A superstructural mathematical programming formulation is developed to synthesise a waste treatment network for optimised nutrient recovery.

## 3. Mathematical Programming Formulation

The total flow of IW  $i$  to PU  $j \in J$  ( $x1_{i,j}$ ) and MW  $im \in IM$  ( $x2_{i,im}$ ) should equate to the total waste available ( $TW_i$ ), as shown in Eq.(1).

$$\sum_j x1_{i,j} + \sum_{im} x2_{i,im} = TW_i \quad \forall i \quad (1)$$

Next, the total flow of IW  $i \in I$  and MW  $im \in IM$  ( $x3_{im,j}$ ) to PU  $j$  should not exceed the maximum treatment capacity of PU  $j$  ( $M_j$ ), as shown in Eq.(2).

$$\sum_i x1_{i,j} + \sum_{im} x3_{im,j} \leq M_j \quad \forall j \quad (2)$$

Eq.(3) ensures that the flow of MW  $im$  to PU  $j$  equate the flow summations of IW  $i$  to MW  $im$ .

$$\sum_i x2_{i,im} = \sum_j x3_{im,j} \quad \forall im \quad (3)$$

Following this, the total recovery of nutrient  $l$  from PU  $j$  treating IW  $i$  ( $NT1_{i,j,l}$ ) and MW  $im$  ( $NT2_{im,j,l}$ ) are depicted in Eq.(4) and Eq.(5) respectively.

$$NT1_{i,j,l} = x1_{i,j} \times a1_{i,j,l} \quad \forall i \forall j \forall l \quad (4)$$

$$NT2_{im,j,l} = x3_{im,j} \times a2_{im,j,l} \quad \forall im \forall j \forall l \quad (5)$$

where  $a1_{i,j,l}$  and  $a2_{im,j,l}$  are the recovery of nutrient  $l$  from PU  $j$  treating IW  $i$  and MW  $im$  respectively.

The objective function of this optimisation problem is to maximise the total nutrient recovery, i.e., the sum of nutrients  $l \in L$  from PU  $j \in J$  treating IW  $i \in I$  and MW  $im \in IM$ , as shown in Eq.(6):

$$\max NTT = \sum_i \sum_j \sum_l NT1_{i,j,l} + \sum_{im} \sum_j \sum_l NT2_{im,j,l} \tag{6}$$

The linear programming model is implemented in GAMS using the CPLEX solver.

#### 4. Nutrient Flow Data and Operational Constraints

This work makes use of a case study from LC in the UK. Firstly, the type and mass of waste available within the LC region are presented in Table 1.

Table 1: Type and volume of waste

Waste	Waste Mass, t y <sup>-1</sup>
Food (F)	149,845
Green (G)	61,143
Wastewater sludge (W)	463,696

Based on Table 1, the three types of waste investigated in this work are food waste, green waste, and wastewater sludge. It is estimated that there are 149,845 tonnes of food waste generated annually in LC (WRAP, 2020). Meanwhile, the annual volume of green waste amounted to 61,143 tonnes (WasteDataFlow, 2021). Moving on to the wastewater, it is estimated that the total wastewater generated in LC is 0.7 m<sup>3</sup> d<sup>-1</sup> capita<sup>-1</sup>, covering both domestic households and small businesses (Severn Trent Plc, 2022). Considering that the population in LC is 712,300 (Richardson, 2022), the total annual volume of wastewater in LC is estimated to be 180.9 million m<sup>3</sup>, resulting in 463,696 tonnes of wastewater sludge annually. In this work, five PUs i.e., anaerobic digestion (AD), windrow composting (WC), bin composting (BC), incineration (IN) and the landfill (LF) are considered for treating various waste in LC. Table 2 presents the assumed maximum treatment capacity of each PU.

Table 2: Assumed maximum treatment capacity of PU

Processing Unit	Waste processing capacity, t y <sup>-1</sup>
AD	700,000
WC	300,000
BC	60,000
IN	700,000
LF	700,000

Both IN and LF are well-established technologies and are assumed to treat a maximum of 700,000 tonnes of waste, as observed in Table 2. Meanwhile, BC which is done within households is assumed to have the least waste processing capacity. In this work, AD is assumed to have a significant annual waste treatment capacity of 700,000 tonnes of organic waste due to the potential for nutrient recovery and biogas generation. Next, the recovery of nutrients  $l$  from the PU  $j$  treating IW  $i$  and MW  $im$  is presented in Table 3.

Note that the two nutrients considered in this work i.e., nitrogen (N) and phosphorus (P) may be recovered in solid (NS/PS) and liquid (NL/PL) streams.

Table 3: Nutrient recovery

Waste Stream	Processing Unit	NL, kg t <sup>-1</sup>	NS, kg t <sup>-1</sup>	PL, kg t <sup>-1</sup>	PS, kg t <sup>-1</sup>
Food <sup>a</sup>	AD	2.28	1.94	0.230	0.61
Water <sup>b</sup>	AD	0	1.80	0	1.47
Food + Green <sup>c, d</sup>	AD	2.15	1.82	0.350	0.92
Green <sup>e</sup>	WC	0	3.22	0	0.61
Food + Green <sup>e</sup>	BC	0.05	2.48	0.010	0.56
Food + Green <sup>f, g</sup>	IN	0	0	0	1.20
Food + Green <sup>f, h</sup>	LF	0.15	0	0.005	0

References: a) (Dimambro, 2015) b) (Havukainen et al., 2022) c) (Banks et al., 2018) d) (Cesaro, 2021) e) (Andersen et al., 2010) f) (Jeswani and Azapagic, 2016) g) (Šyc et al., 2020) h) (Kulikowska and Klimiuk, 2008)

Based on Table 3, AD may be used to treat food wastes, wastewater sludge as well as a mixture of food and green wastes. WC may be used for the treatment of green waste. Besides, the mixture of food and green wastes may be treated using BC, IN or LF. Although the ratio of food and green waste may change in a MW stream, this work assumes a fixed ratio of the food and green waste (9:1) in MW streams to all PUs.

## 5. Case Study

This work is solved according to Eq.(6) subject to the constraints in Eq.(1) till Eq.(5). Table 4 presents the optimised volumes of waste flows to the PUs. Note that Table 4 also includes the results of two supplementary runs that maximise the recovery of total N or P.

Table 4: Waste Flow Volume to PU <sub>j</sub>

Waste Stream	Waste flow to PU (max NTT/max total N/max total P)				
	AD	WC	BC	IN	LF
Food, t y <sup>-1</sup>	0/149,845/0	0	0	0	0
Green, t y <sup>-1</sup>	0	44,493/61,143/44,493	0	0	0
Wastewater, t y <sup>-1</sup>	463,696	0	0	0	0
Food + Green, t y <sup>-1</sup>	166,495/0/166,495	0	0	0	0

Based on Table 4, maximising the total nutrient recovery (Eq.(6)) resulted in no food waste being treated via AD as an IW stream. All food waste is combined with approximately 27% of the green waste for treatment via AD. The higher nutrient recovery of the MW stream via AD resulted in it being the primary choice of waste treatment in this optimisation problem (see Table 3). Note that all wastewater sludge is treated via AD for a similar reason. The objective function in this work is maximised at 2,559 tonnes of nutrients, comprising 1,639 tonnes of N and 920 tonnes of P. Moving on, the total recovery of N is maximised at 1,663 tonnes. All wastes are treated via AD and WC as an

IW stream. The greater N recovery from all types of waste being treated as an IW stream led to this solution. The total recovery of P is maximised at 920 tonnes, with similar results observed as to maximising the total nutrient recovery. It is good to note that if incineration is selected, the P recovery from incineration ash would require additional extraction processes before it may be applied to agricultural land or replacing phosphate for fertiliser production (Kalmykova and Karlfeldt Fedje, 2013). The results of this optimisation problem align with the aim of this work in maximising nutrient recovery and exploring the potential offered by these waste streams.

## 6. Conclusions

An optimal waste treatment plan for maximising nutrient recovery is suggested in this work based on mathematical programming models. This work is expected to aid in value-creation among waste streams that would otherwise remain unutilised. The results of this work demonstrate the importance of AD in maximising nutrient recovery. Additionally, combining the food and green waste streams presents an added value for better use of resources. Future work should include the cost analysis of potential by-products i.e., biogas, and digestate as well as emissions that may arise from the waste treatment. Also, the operational and capital costs associated with the operations of waste treatment plants must be considered for a more realistic waste management scenario.

## Acknowledgement

This work is supported by the Natural Environment Research Council (NERC), UK as part of the AGILE initiative (grant number NE/W004976/1). The authors would like to also thank Dr John Ingram and Dr Bhawana Gupta from the Environmental Change Institute at the University of Oxford for the support received throughout this work.

## References

- Álvarez, J., Roca, M., Valderrama, C., Cortina, J.L., 2018. A Phosphorous Flow Analysis in Spain. *Science of The Total Environment* 612, 995–1006. <https://doi.org/10.1016/J.SCITOTENV.2017.08.299>
- Andersen, J.K., Boldrin, A., Christensen, T.H., Scheutz, C., 2010. Greenhouse gas emissions from home composting of organic household waste. *Waste Management* 30, 2475–2482. <https://doi.org/10.1016/J.WASMAN.2010.07.004>
- Banks, C., Heaven, S., Zhang, Y., Baier, U., IEA Bioenergy Programme, 2018. Food waste digestion : Anaerobic digestion of food waste for a circular economy. IEA Bioenergy Task 37 12.
- Cesaro, A., 2021. The valorization of the anaerobic digestate from the organic fractions of municipal solid waste: Challenges and perspectives. *J Environ Manage* 280, 111742. <https://doi.org/10.1016/J.JENVMAN.2020.111742>
- Cordell, D., White, S., 2015. Tracking phosphorus security: indicators of phosphorus vulnerability in the global food system. *Food Secur* 7, 337–350. <https://doi.org/10.1007/S12571-015-0442-0/FIGURES/6>
- DEFRA, 2022. UK statistics on waste [WWW Document]. Department for Environment, Food & Rural Affairs. URL <https://www.gov.uk/government/statistics/uk-waste-data/uk-statistics-on-waste#data-revisions-in-this-update> (accessed 9.27.22).
- Dimambro, M.E., 2015. Novel Uses for Digestates: Protected Horticulture. 20th European Biosolids & Organic Resources Conference & Exhibition.
- Feiz, R., Johansson, M., Lindkvist, E., Moestedt, J., Pålledal, S.N., Ometto, F., 2022. The biogas yield, climate impact, energy balance, nutrient recovery, and resource cost of biogas production from household food waste—A comparison of multiple cases from Sweden. *J Clean Prod* 378, 134536. <https://doi.org/10.1016/J.JCLEPRO.2022.134536>



- Havukainen, J., Saud, A., Astrup, T.F., Peltola, P., Horttanainen, M., 2022. Environmental performance of dewatered sewage sludge digestate utilization based on life cycle assessment. *Waste Management* 137, 210–221. <https://doi.org/10.1016/J.WASMAN.2021.11.005>
- Jeswani, H.K., Azapagic, A., 2016. Assessing the environmental sustainability of energy recovery from municipal solid waste in the UK. *Waste Management* 50, 346–363. <https://doi.org/10.1016/J.WASMAN.2016.02.010>
- Kalmykova, Y., Karlfeldt Fedje, K., 2013. Phosphorus recovery from municipal solid waste incineration fly ash. *Waste Management* 33, 1403–1410. <https://doi.org/10.1016/J.WASMAN.2013.01.040>
- Kulikowska, D., Klimiuk, E., 2008. The effect of landfill age on municipal leachate composition. *Bioresour Technol* 99, 5981–5985. <https://doi.org/10.1016/J.BIORTECH.2007.10.015>
- Liu, X., Elgowainy, A., Wang, M., 2020. Life cycle energy use and greenhouse gas emissions of ammonia production from renewable resources and industrial by-products. *Green Chemistry* 22, 5751–5761. <https://doi.org/10.1039/D0GC02301A>
- Mehta, C.M., Khunjar, W.O., Nguyen, V., Tait, S., Batstone, D.J., 2014. Technologies to Recover Nutrients from Waste Streams: A Critical Review. *Crit Rev Environ Sci Technol* 45, 385–427. <https://doi.org/10.1080/10643389.2013.866621>
- Metson, G.S., Feiz, R., Quttineh, N.H., Tonderski, K., 2020. Optimizing transport to maximize nutrient recycling and green energy recovery. *Resources, Conservation & Recycling: X* 9–10, 100049. <https://doi.org/10.1016/J.RCRX.2021.100049>
- Perera, M.K., Englehardt, J.D., Dvorak, A.C., 2019. Technologies for Recovering Nutrients from Wastewater: A Critical Review. *Environ Eng Sci*. <https://doi.org/10.1089/ees.2018.0436>
- Richardson, H., 2022. Leicestershire population rises by more than 100,000 in last ten years - Leicestershire Live [WWW Document]. LeicestershireLive. URL <https://www.leicesterm Mercury.co.uk/news/local-news/leicestershire-population-rises-more-100000-7290712> (accessed 11.16.22).
- Severn Trent Plc, 2022. Who We Are [WWW Document]. About Us. URL <https://www.severntrent.com/about-us/who-we-are/> (accessed 11.16.22).
- Sniatala, B., Kurniawan, T.A., Sobotka, D., Makinia, J., Othman, M.H.D., 2023. Macro-nutrients recovery from liquid waste as a sustainable resource for production of recovered mineral fertilizer: Uncovering alternative options to sustain global food security cost-effectively. *Science of The Total Environment* 856, 159283. <https://doi.org/10.1016/J.SCITOTENV.2022.159283>
- Šyc, M., Simon, F.G., Hykš, J., Braga, R., Biganzoli, L., Costa, G., Funari, V., Grosso, M., 2020. Metal recovery from incineration bottom ash: State-of-the-art and recent developments. *J Hazard Mater* 393, 122433. <https://doi.org/10.1016/J.JHAZMAT.2020.122433>
- van der Wiel, B.Z., Weijma, J., van Middelaar, C.E., Kleinke, M., Buisman, C.J.N., Wichern, F., 2020. Restoring nutrient circularity: A review of nutrient stock and flow analyses of local agro-food-waste systems. *Resour Conserv Recycl* 160, 104901. <https://doi.org/10.1016/J.RESCONREC.2020.104901>
- WasteDataFlow, 2021. WasteDataFlow [WWW Document]. What is WasteDataFlow? URL <https://www.wastedataflow.org/> (accessed 11.16.22).
- WRAP, 2020. UK progress against Courtauld 2025 targets and UN Sustainable Development Goal 12.3 [WWW Document]. Report. URL <https://wrap.org.uk/resources/report/uk-progress-against-courtauld-2025-targets-and-un-sustainable-development-goal-123> (accessed 11.16.22).

# Improving results of a Continuous Fluidized Bed Process for the Separation of Enantiomers by applying mathematical optimization

Nadiia Huskova<sup>1,3</sup>, Jonathan Gänsch<sup>1</sup>, Michael Mangold<sup>2</sup>, Heike Lorenz<sup>1,3</sup>,  
Andreas Seidel-Morgenstern<sup>1,3</sup>

<sup>1</sup>*Max Planck Institute for Dynamics of Complex Technical Systems,  
Sandtorstraße 1, 39106 Magdeburg*

<sup>2</sup>*University of Applied Sciences Bingen, Berlinstraße 109, 55411 Bingen*

<sup>3</sup>*Otto von Guericke University Magdeburg, Universitätsplatz 2, 39106  
Magdeburg*

## Abstract

The focus of the presented work is on the model-based optimization of a purification process of chemical compounds called enantiomers. As an example, the separation of asparagine monohydrate from the racemic mixture is considered, which has been proven to be effective in experiments. The experimental results indicate the need for a careful selection of operating conditions to provide the desired process performance. To describe the studied phenomena, a validated mathematical model based on population balance equations is used. Numerical optimization finds a set of operating conditions ensuring high product purity, while keeping the product contamination under a defined threshold

**Keywords:** Separation of Enantiomers, Continuous Crystallization, Population Balance Equations, Mathematical Optimization, Genetic Algorithms, Parallel Computations.

## 1. Introduction

More than ten million substances have been discovered in organic chemistry over the past two centuries (Dengale et al., 2016). Organic synthesis is becoming more and more prevalent due to the need to solve fundamental problems, such as determining the relationship between a chemical substance and its reactivity. To prevent and treat diseases, there is a need for affordable and effective drugs. To satisfy this need, chemists and pharmaceutical researchers are joining forces.

To reduce the time of drug production, great efforts are made to study already known compounds that are optical or geometric isomers.

One of the potential candidates are enantiomers, which are pairs of optically active isomers differently orientated in space (McConathy and Owens, 2003). A large interest in producing single enantiomers is confirmed by the life science industries, pharmaceuticals, and agrochemical companies. To supply pure enantiomers different approaches can be performed. The focus of this study is on the racemic approach based on separating mixtures of both enantiomers to obtain a significant amount of the desired single enantiomer characterized by high purity.

## **2. Process description**

Chemical industries have increasingly relied on crystallization as a separation and purification technology, especially for those chiral molecules that crystallize in conglomerates, where the two enantiomers are mechanically mixed in the solid phase crystallizing separately as enantiopure crystals. In this kinetically controlled process, different rates of crystallization of each enantiomer are important.

Implementing a fluidized bed crystallization approach, two tubular crystallizers are connected spatially and are capable of crystallizing both enantiomers of chiral molecules selectively out of racemic solutions. Each of the operated crystallizers consists of the cylindrical top and the conical bottom parts. Liquid solvent without crystals enters the crystallizer at the bottom, and a mixture of solvent and formed crystals quits from the crystallizer at the top. The mechanical milling element located at the bottom ensures the continuity of the process and the supply of seed crystals. The volumetric flow rate characterizes the incoming stream; the removal of crystals is carried out periodically, and the outflow rate is determined by the applied withdrawal strategy. Corresponding scheme of the process is given in Figure 1. Special attention within this work is paid to modeling the crystal's withdrawal in a pulse-wise manner how it is performed in a laboratory.

## **3. Modeling crystallization process in fluidized beds**

The fundamentals of the mathematical model applied are proposed in (Huskova, 2022). The model consists of three main components: population balances for the enantiomers in both crystallizers, distributed in one space and one property coordinate; a total mass balance for describing changes in the liquid phase; an energy balance to account for the non-isothermal nature of the process. The model distinguishes between the two types of enantiomers, as they may

crystallize at different rates The structural change in the crystal shape is not considered. Instead, the physical parameters of the model are sorted into four groups:

- crystallizer design parameters;
- thermodynamic and kinetic properties of a particular chemical compound;
- operational parameters and parameters related to the product withdrawal strategy;
- set of parameters related to the model assumptions that have been adapted within the validation procedure.

A detailed description of the mathematical model and all the related model equations can be found in (Huskova, 2022). The reference numerical solutions are obtained in Matlab by a time integration with the ODE23 integrator applying Runge–Kutta method.

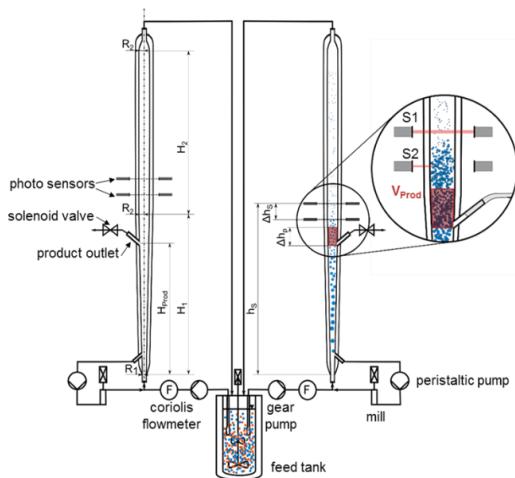


Figure 1. Schematic representation of continuous crystallization in a fluidized bed described in (Gänsch et al., 2021)

#### 4. Model-based optimization

The next step after finding the reference solution and providing a model validation is to determine relevant model parameters to improve the process. The

main goal of the optimization is to improve productivity of the target compound, which in general can be expressed as follows:

$$OF = Pr^L \rightarrow \max ,$$

where  $L$  stands for a target enantiomer.

A sensitivity analysis presented in (Gänsch et al., 2021) indicates that the operational conditions and geometrical parameters of the crystallizer listed in Table 1 have the biggest leverage.

Table 1. List of optimization variables

$v_1$	$H$	Crystallizer height
$v_2$	$H_1$	Height of conical section
$v_3$	$R_1$	Radius of the bottom
$v_4$	$R_2$	Radius of the top
$v_5$	$T_{cryst}$	Crystallizer Temperature
$v_6$	$u_{emptytube}$	Empty tube velocity

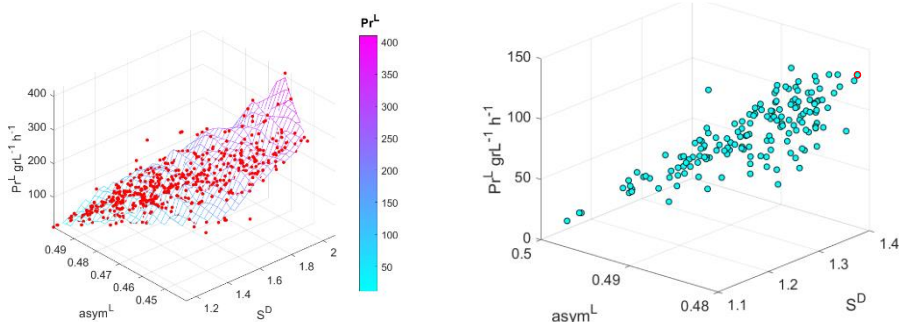
#### *Empirical threshold to nucleation assessment*

However, while analyzing productivity, it should be noted that when the sensor is in a high position, the risk of product contamination significantly increases, as was experimentally proven in (Gänsch et al., 2021). The occurrence of contamination can be explained by critical values of nucleation factors – supersaturation of the counter-enantiomer,  $S^D$ , and asymmetry of the solution,  $asym^L$  (from here on, the target enantiomer is indexed  $L$ , and the counter-enantiomer  $D$ ). From previous work, it is known that a low value of  $asym^L$  as well as a high value of  $S^D$  cause a drop in purity. Based on the experimental study, the following restrictions are proposed to ensure an acceptable product contamination:

$$\begin{cases} S^D \leq 1.4 \\ asym_L \geq 0.48 \end{cases} \quad [\text{Eq.1}]$$

#### **Optimization results**

The simulation results shown on the left-hand side of Figure 1 do not include imposed constraints. It demonstrates that the productivity of the process can reach  $400 \text{ gL}^{-1} \text{ h}^{-1}$ , however with low purity and high contamination of the product. When the nucleation factors from [Eq.1] are considered in the optimization, the achieved productivity is lower – see the right-hand side of Figure 2.



*Figure 2.* Simulation results achieved ignoring (left) following (right) the imposed nucleation factors constraints.

Also in this case, the obtained values significantly exceed the productivity noted in the available literature. Compared to the reference case study reported in (Gänsch et al., 2021), the increase in productivity reaches about 60%. The proposed approach makes it possible to ensure the minimum acceptable level of product purity and focus on increasing productivity by selecting the previously identified optimization variables.

The use of evolutionary algorithms for finding optimal solutions made it possible to avoid getting stuck in a local minimum, which would have happened with a high probability when using gradient-based approaches due to the non-convexity problem.

*Applicability of parallel computing*

Another challenge for the mathematical optimization is related to the resource intensity of the numerical solution. To ensure high accuracy of the numerical solution, the number of differential equations resulting from a discretization of the space and property coordinate exceeds 20 thousand. The elapsed time to calculate the transient to a cyclic steady state for one set of parameters exceeds several hours. The size of the developed model and time constraints lead to the inability to use a single processor to obtain optimal parameters and improve the process. In this regard, an attempt is made to directly solve the problem of reducing the computational requirements without compromising the accuracy of the developed mathematical model through parallelization methods. At each iteration, the formed sets of parameters describing the operating and design conditions were distributed among the processors, which made it possible to significantly speed up the search for the optimal solution.

## 5. Conclusions

The proposed model-based optimization approach is applied to improve the productivity and robustness of the separation of asparagine monohydrate in a fluidized bed reactor. It is demonstrated that the proposed optimal design of the crystallizer setup increases productivity, and additionally, it allows the evaluation of the nucleation factors once the suggested geometry is simulated. These recommendations can be used by experimenters to further achieve the desired purity and productivity of the process.

By shifting the focus away from asparagine monohydrate, the developed mathematical model and optimization procedure can be applied to design the purification of other substances. The study showed that substances for which the proposed fluidized bed approach is applicable should have a certain ratio of the densities of the solid and liquid phases. As an outlook, the approach may also be used to optimize particle characteristics like average particle size.

## Acknowledgments

The authors acknowledge the financial support of Deutsche Forschung Gemeinschaft (DFG) within the Priority Program “Dynamic Simulation of Interconnected Solids Processes (SPP 1679)”. N.H. is grateful for the support by the “Research Center of Dynamic Systems (CDS)” of Otto von Guericke University Magdeburg.

## References

- (Dengale et al., 2016). Dengale, S.J., Grohganz, H., Rades, T., Löbmann, K., 2016. Recent advances in co-amorphous drug formulations. *Adv. Drug Delivery Rev.* 100,116-125.
- (Gänsch et al., 2021). Gänsch, J., Huskova, N., Kerst, K., Temmel, E., Lorenz, H., Mangold, M., Janiga, G., Seidel-Morgenstern, A., 2021. Continuous enantioselective crystallization of chiral compounds in coupled fluidized beds. *Chem. Eng. J.* 422.
- (Gooding, W. B., Pekny, J. F., McCroskey, P. S. (1994). Enumerative Approaches to Parallel Flowshop Scheduling via Problem Transformation. *Comput. Chem. Eng.*, 18, 909.
- (Huskova, 2022). Huskova, N. 2022. Dynamic Modeling and Optimization of a Continuous Fluidized Bed Process for the Separation of Enantiomers by Preferential Crystallization. PhD thesis. Otto von Guericke University, Magdeburg.
- (McConathy and Owens, 2003). McConathy, J., Owens, M., 2003. Stereochemistry in Drug Action. *J. Clin. Psychiatry* 2, 70-73.

# Machine-learning-based optimization of operating conditions of naphtha cracking furnace to maximize plant profit

Chonghyo Joo,<sup>a,b</sup> Hyukwon Kwon,<sup>a,b</sup> Junghwan Kim<sup>a,b</sup>, Hyungtae Cho<sup>a</sup>, and Jaewon Lee<sup>a,\*</sup>

<sup>a</sup>*Green Materials and Processes R&D Group, Korea Institute of Industrial Technology, 55, Jongga-ro, Ulsan 44413, Korea*

<sup>b</sup>*Department of Chemical and Biomolecular Engineering, Yonsei University, 50, Yonsei-ro, Seoul 03722, Korea*

**[j.lee@kitech.re.kr](mailto:j.lee@kitech.re.kr)**

## Abstract

Naphtha cracking is the primary process for propylene (PL) and ethylene (EL) production and depends on the operating conditions, such as the coil outlet temperature and feedstock composition. The product yields, in turn, determine the profit of the naphtha cracking plant. Therefore, the operating conditions should be optimized to maximize plant profit. However, it is challenging to optimize the conditions using conventional simulation methods since a high-fidelity model is hard to develop owing to the nonlinearity and complexity of the cracking process. In this study, we used machine learning to optimize the operating conditions of a naphtha cracking furnace to maximize plant profit. First, a data-driven model was developed to predict the product yields for different operating conditions using a deep neural network (DNN). The model could predict the PL and EL yields with high accuracy ( $R^2 = 0.965$  and  $0.900$ , respectively). Next, a genetic algorithm was used for optimization based on the developed DNN model. Finally, the developed model was used with real-world plant data and product prices for 2020. The plant profit under the optimized operating conditions was 30% higher than that corresponding to the original operating conditions. Thus, the proposed method is suitable for determining the optimal operating conditions of various types of plants in order to maximize profit.

**Keywords:** naphtha cracking furnace, optimization, machine learning, genetic algorithm

## 1. Introduction

Growth in the petrochemical industry has increased the demand for ethylene (EL) and propylene (PL), which are the elementary materials for various petrochemical products. To meet this demand, EL and PL are generally produced in a naphtha cracking center (NCC), which breaks down naphtha into light and valuable hydrocarbons.

The NCC consists of four main units: a cracking furnace and quenching, compression, and distillation units, as shown in Figure 1. In the first unit, that is, the cracking furnace, naphtha is passed through the coils of the furnace as the feedstock, where it receives thermal energy for the cracking reaction from the furnace. As a result, it undergoes cracking into light hydrocarbons. In the second unit, that is, the quenching unit, the light



hydrocarbons are quenched to prevent the formation of coke and tar in the hydrocarbon mixture and recover the waste heat. In the third unit, that is, the compression unit, the light hydrocarbons receive energy for transportation. Moreover, catalyst poisons and impurities, such as  $H_2S$  and  $CO_2$ , are removed in this unit. Lastly, in the distillation unit, the light hydrocarbons are fractionated into different products, such as EL and PL.

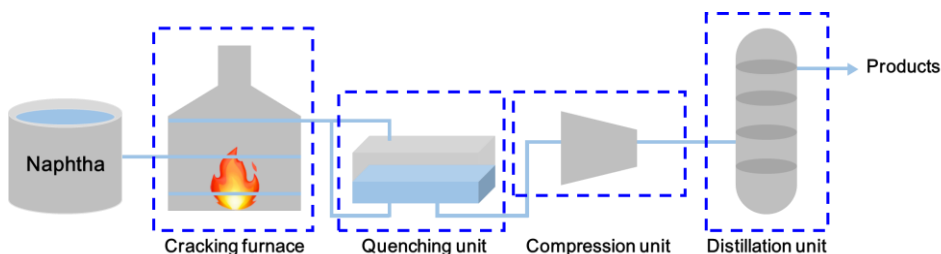


Figure 1. Illustration of NCC for EL and PL production

Of the various units that form the NCC, the cracking furnace has the greatest impact on the product yields, which determine the profit of the naphtha cracking plant because the cracking reaction occurs in the furnace. Figure 2 describes the cracking reaction in the furnace. First, naphtha is preheated before being mixed with diluting steam. After being preheated in the convection section, this mixture is transported to the coil in the radiant section and broken into EL and PL via the cracking reaction. The products are subsequently transported to the quenching unit through the coil outlet.

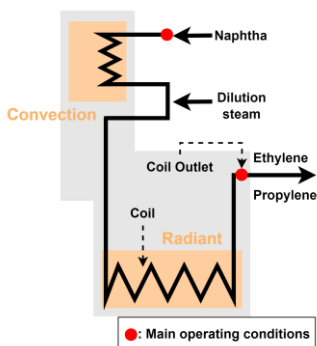


Figure 2. Schematic of cracking furnace showing main operating conditions: naphtha composition and COT

The yields of EL and PL depend on the operating conditions, that is, the naphtha composition and coil outlet temperature (COT). Thus, the operating conditions of the cracking furnace must be optimized to maximize the profit of the naphtha cracking plant. However, it is challenging to optimize the conditions using conventional simulation methods since a high-fidelity model is hard to develop owing to the nonlinearity and complexity of the cracking process.

In this study, we use machine learning to optimize the operating conditions in a naphtha cracking furnace in order to maximize the plant profit. Figure 3 provides an overview of the study. First, 784 actual industrial operating data points were collected from an NCC plant. In this dataset, the naphtha composition, which included 25 components, and the

COT are used as the input variables, while the EL and PL yields are used as the output variables for the developed model. Next, a deep neural network (DNN)-based model is developed using the data. The developed model is evaluated based on the  $R^2$  values before optimization. Finally, the operating conditions are optimized for profit maximization using a genetic algorithm. For model evaluation, the optimal conditions and the corresponding profit are compared with the actual plant data for 2020.

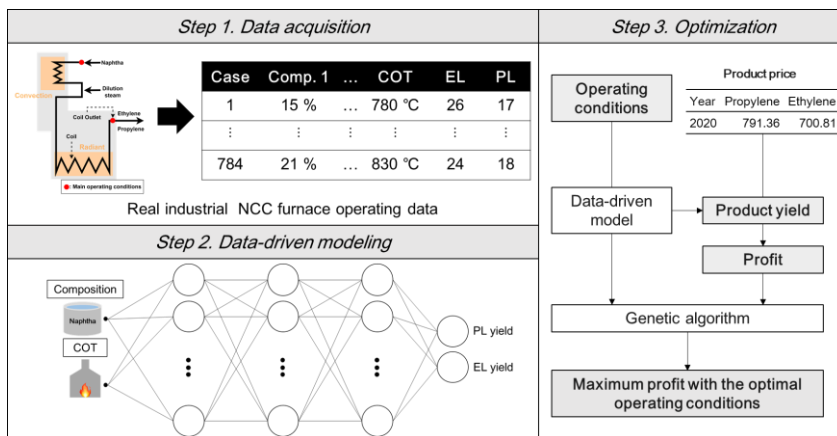


Figure 3. Overview of study

## 2. Method

### 2.1. Deep neural networks (DNNs)

DNNs have been used widely for data-driven modeling. A DNN consists of layers, including nodes and edges, that contain mathematical relationships. During data training, these relationships are updated by backpropagation. After training, the updated relationships are used as the equations for predicting the output variables based on the input variables. Therefore, a significant advantage of DNNs is that they can express the relationships that exist in a system regardless of the nonlinearity and complexity of the system.

Figure 4 shows the structure of the DNN-based model used in this study. The model consists of one input layer, three hidden layers, and one output layer, and these layers include 26 nodes, 100 nodes, and two nodes, respectively. This means that the 26 input variables, namely, the composition of the 25 naphtha components and the COT, were used to calculate the two output variables, namely, the EL and PL yields, while considering 300 functions.

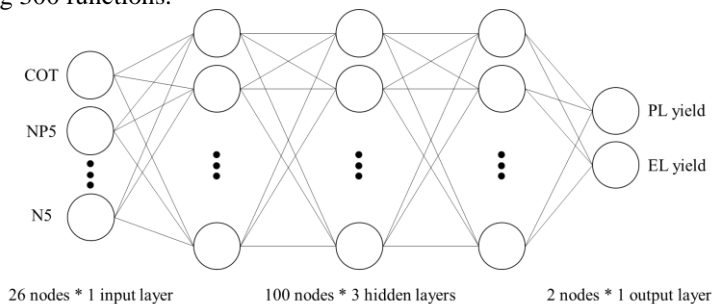


Figure 4. Structure of DNN-based model used in this study

## 2.2. Genetic algorithm

A genetic algorithm (GA) is a heuristic search algorithm inspired by the biological evolution process. GAs repeat three steps for optimization, as shown in Figure 5: selection, crossover, and mutation. In the selection step, a few data points, called parents, are selected randomly and contribute to the population of the next generation. Here, “ $Obj_n$ ” is the result of the objective function of the  $n^{\text{th}}$  data point. In the crossover step, two parents are combined to form the offspring for the next generation by switching some of the input variables. Lastly, the random changes in the input variables are applied to obtain different offspring. By repeating these steps, “ $Obj_n$ ” is enhanced gradually, and the optimal data point is suggested eventually.

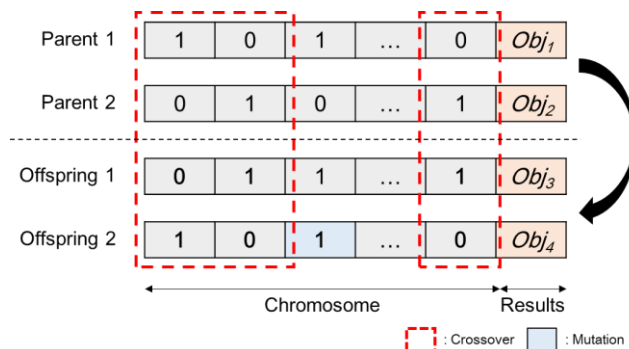


Figure 5. Main steps of GA: selection, crossover, and mutation

This study aimed to optimize the operating conditions of NCC furnaces for profit maximization. Thus, the objective function was defined as shown in Eq. (1), and a GA was employed to solve the optimization problem.

### Objective function

$$\text{maximize Profit} = \max \left( \text{flowrate} \times \frac{\text{yield}_{EL} \times \text{Price}_{EL} + \text{yield}_{PL} \times \text{Price}_{PL}}{\text{yield}_{total}} \right) \quad (1)$$

where  $\text{flowrate}$  is assumed to be 29 ton/h in the steady state, and  $\text{yield}_{EL}$ ,  $\text{yield}_{PL}$ , and  $\text{yield}_{total}$  indicate the yields of EL, PL, and the total product, respectively.  $\text{Price}_{EL}$  and  $\text{Price}_{PL}$  are the product prices in 2020, as shown in Table 1.

Table 1. Prices of primary naphtha cracking products in 2020

Year	Propylene	Ethylene
2020	791.36 \$/ton	700.81 \$/ton

## 3. Results and discussion

### 3.1. Model development

To develop the data-driven model, the 784 industrial data points was randomly split into 60% training, 20% validation, and 20% test datasets. In other words, 470 data points were used for model training, 157 data points are used for model validation, and the rest (157 data points) were used for model testing.

After the model had been developed using the training and validation datasets, its performance was evaluated using the test dataset based on the coefficient of determination ( $R^2$ ), which is a representative evaluation criterion for regression.  $R^2$  was calculated using Eq. (2):

$$R^2 = 1 - \frac{\sum_{i=1}^N (S_{ia} - S_{ip})^2}{\sum_{i=1}^N (S_{ia} - \bar{S}_{ip})^2} \quad (2)$$

where  $N$  indicates the number of data points,  $S_{ia}$  indicates the  $i^{\text{th}}$  actual value, and  $S_{ip}$  indicates the  $i^{\text{th}}$  value predicted by the model.

The results of the model performance evaluation are shown in Figure 6. As can be seen, even though the data point within the red box shows a relatively large error, the DNN-based model predicted the EL and PL yields with high  $R^2$  values (0.965 and 0.900, respectively). Moreover, the figure shows that the more EL the NCC furnace produces, the less the amount of PL produced. This is because it is likely that PL is decomposed into EL owing to the high COT or a lack of heavy components in naphtha. Therefore, the model showed good prediction performance and confirmed that the product yields could be changed by varying the cracking reaction conditions.

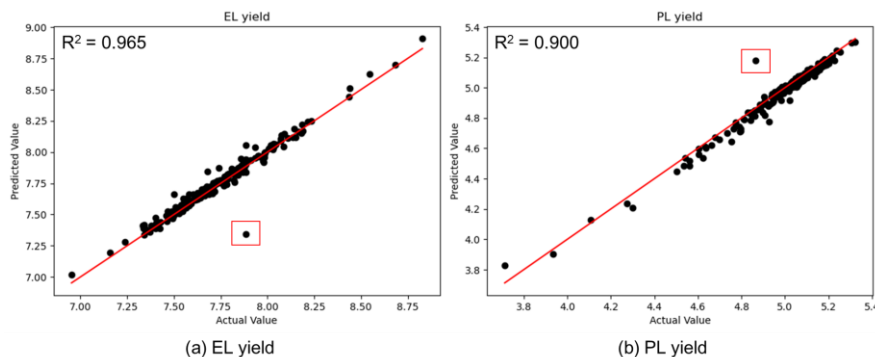


Figure 6. Parity plots and  $R^2$  values for (a) EL yield and (b) PL yield

### 3.2. Profit optimization

Using the DNN-based model and the GA, the operating conditions, that is, the naphtha composition corresponding to 25 components and the COT, were optimized to maximize the profit of the NCC plant. In addition, the base case, that is, the actual operating data for 2020, were compared with the optimization results for validation.

Table 2 shows the operating conditions and the profits for the base and optimal cases; the 25 components are encoded as “Comp.N” for reasons of confidentiality. The optimal case corresponds to a profit of 12,202 \$/h, which is 30% higher than that for the base case, which is 9,382 \$/h. Moreover, for the optimal case, COT is 42 °C higher than that for the base case, and some of the naphtha components are very different from those for the base case. This means that a high COT will increase the profit because it accelerates the cracking reaction. Furthermore, the use of the appropriate naphtha composition also increases the profit. Thus, the blending process can be considered to increase the profit.

It should be noted that the optimization process only considered the product prices. However, other factors, such as the fuel price, energy consumption, and coke formation, should also be considered in the future to obtain a more practical model.

Table 2. Operating conditions and profits for base and optimal cases

Case	COT (°C)	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6	Comp.7	Comp.8
Base	808	3.20	8.09	2.28	0.01	0.00	2.86	0.21	0.00
Optimal	850	2.79	6.70	0.76	2.02	0.55	1.28	0.03	0.00
	Comp.9	Comp.10	Comp.11	Comp.12	Comp.13	Comp.14	Comp.15	Comp.16	Comp.17
	(wt%)								
Base	0.00	0.00	0.13	5.12	22.20	11.97	1.06	0.00	0.00
Optimal	0.00	0.00	1.93	7.84	15.87	9.14	3.87	2.55	1.10
	Comp.18	Comp.19	Comp.20	Comp.21	Comp.22	Comp.23	Comp.24	Comp.25	Profit (\$/h)
	(wt%)								
Base	0.00	0.51	22.34	16.86	3.13	0.01	0.01	0.00	9,382.01
Optimal	0.29	2.32	19.20	11.47	5.56	3.06	0.88	0.81	12,202.90

#### 4. Conclusions

The profit of NCC plants depends on the operating conditions of the cracking furnace, composition of the 25 naphtha components, and COT. Thus, these conditions must be optimized to maximize the profit. However, it is challenging to optimize the conditions because developing a high-fidelity model for optimization is difficult owing to the nonlinearity and complexity of the cracking process. In this study, we used machine learning to optimize the operating conditions in a naphtha cracking furnace to maximize the profit. First, a DNN-based model was developed using 784 industrial operating data points. The model could predict the EL and PL yields with high  $R^2$  values (0.965 and 0.900, respectively). Next, the model was integrated with a GA for optimization. The profit under the optimal operating conditions was 30% higher than that for the base case.

Despite the success in profit maximization, this study has a limitation in that it only considers the product prices for profit calculation. Thus, in the future, the model should be improved by considering the other factors that also affect the profit. The updated model should yield more practical solutions for maximizing the profit.

#### Acknowledgments

This study has been conducted with the support of the Korea Institute of Industrial Technology as “Development and application of carbon-neutral engineering platform based on carbon emission database and prediction model (kitech KM-22-0348)”.

#### References

- E. Joo, K. Lee, M. Lee, S. Park, CRACKER — a PC based simulator for industrial cracking furnaces. *Comput Chem Eng* 2000;24:1523–8. [https://doi.org/10.1016/S0098-1354\(00\)00558-5](https://doi.org/10.1016/S0098-1354(00)00558-5).
- M. Masoumi, M. Shahrokhi, M. Sadrameli, J. Towfighi, Modeling and control of a naphtha thermal cracking pilot plant. *Ind Eng Chem Res* 2006;45:3574–82. <https://doi.org/10.1021/ie050630f>.
- Y. Brown, Year, Article or Chapter Title, etc. *Global ethylene demand & capacity 2015-2022* | Statista, (n.d.). <https://www.statista.com/statistics/1246694/ethylene-demand-capacity-forecast-worldwide/> (accessed November 23, 2022).

# Methanol synthesis and hematite reduction using waste tires

Athi-enkosi Mavukwana,<sup>a</sup> Celestin Baraka Sempuga,<sup>b</sup>

<sup>a</sup>*Department of Civil and Chemical Engineering, College of Science, Engineering, and Technology, University of South Africa (UNISA), c/o Christiaan de Wet & Pioneer Avenue, Florida Campus 1710, Johannesburg, South Africa*

<sup>b</sup>*Institute for the Development of Energy for African Sustainability, College of Science, Engineering, and Technology, University of South Africa (UNISA), c/o Christiaan de Wet & Pioneer Avenue, Florida Campus 1710, Johannesburg, South Africa*

## Abstract

Iron production and Steelmaking are very energy-intensive processes with an estimated global energy intensity of 16.5 GJ/ton and global CO<sub>2</sub> emissions of 1.8 tons per ton of steel produced. Key solutions are required to optimize this industry but maintain its output or production. This work proposes to supplant the coal used in steelmaking with waste tires to reduce operational costs and CO<sub>2</sub> emissions through the polygeneration of products. Process synthesis techniques are used to determine the fundamental thermodynamic limitations of such a process through process performance targeting. It is found that one ton of iron and 430 kg of methanol can be potentially produced from 1.42 tons of haematite and 479 kg of waste tires. However, 63% of the carbon resource will end up as carbon dioxide, equivalent to 985 kg/ton of iron produced. This is lower than the coal route which releases about 1.05 CO<sub>2</sub> tons/ ton of iron.

**Keywords:** Waste tires, Reduction, Methanol, Iron, Thermodynamics.

## 1. Introduction

Global energy needs are constantly increasing and as a result, the natural resources that sustain these energy demands are becoming scarcer and more expensive, whilst human-generated emissions and other solid waste are increasing to critical levels. All companies are now under pressure from societies to decarbonize or reduce emissions from their processes. One such process industry is iron ore reduction and steelmaking. Iron production and Steelmaking are very energy-intensive processes with an estimated global energy intensity of 16.5 GJ/ton [1]. The whole process is dependent on natural resources for both energy needs and products. This results in global CO<sub>2</sub> emissions of 1.8 tons per ton of steel produced [2]. Key solutions are required to optimize this industry but maintain its output or production. The major contributor to carbon emissions from the industry is the use of coal. This work proposes to reduce the coal used in steelmaking with waste tires. Waste tires are derived from fossil fuels and currently at their end-of-life tires have no fundamental use and end up being dumped in landfills creating an unnecessary environmental problem. For example, South Africa produces an estimated 250 000 tonnes of waste tires per annum adding to an existing stockpile excess of 900 000 tonnes spread across 26 national storage depots without a solid plan for reprocessing or recycling [3]. Tires possess a higher volatile content, higher heating value, and low ash content than

coal [4], and can be used to replace or reduce coal in thermochemical conversion processes. This paper evaluates and determines theoretical performance targets for combined methanol synthesis and iron ore reduction using waste tires as a carbon source. The main constituents of tires are carbon and hydrogen and thus the reduction process simply involves reacting the carbon and hydrogen in waste tires with the oxygen in the iron ore to produce iron in its reduced form and carbon dioxide and water. However, in this work, we explore co-production opportunities to improve the efficiency of the process by reducing CO<sub>2</sub> and H<sub>2</sub>O production and thereby improving the material and energy efficiency of the process.

## 2. Methodology

Two thermodynamic properties, i.e., the enthalpy ( $\Delta H$ ) and the Gibbs free energy ( $\Delta G$ ) are used for process targeting. Where,  $\Delta H$  is the amount of energy requirement (input or output) across the process boundary, and  $\Delta G$  is the quality of the energy expressed in terms of the equivalent amount of mechanical work required to effect the change from feed to the product [5]. We use a systematic approach to identify feasible process targets by first defining the attainable region of the process and systematically identifying targets of interest subject to material and energy balances and thermodynamic constraints. The feasibility of the process can be analyzed in terms of their heat and work requirements. The target mass balance for all processes discussed here is met by setting two limits; the process must be adiabatic ( $\Delta H = 0$ ), and the mass balance is based on one mole of a tire. The utilization of raw materials, waste generation, and the thermodynamic efficiency of the process will be assessed using the following three metrics [6].

$$\text{Carbon Efficiency} = \frac{\text{moles of C in the desired product}}{\text{moles of C in the feed}} \quad (1)$$

$$\text{Atom Economy} = \frac{\text{mass of desired product}}{\text{Total mass of feed}} \quad (2)$$

$$E - \text{factor} = \frac{\text{Mass of waste produced}}{\text{Mass of desired product}} \quad (3)$$

Where equation 2 quantifies how much of the carbon in the feed stream ends up in the desired products. Any carbon that is not contained within a product will be lost in undesired by-products. Equation 3 provides useful information on the utilization of all the feed and waste generation. Table 1 shows the ultimate and proximate analysis of waste tires and coal used in this study.

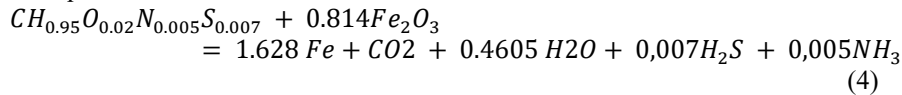
Table 1: Ultimate and proximate analysis [7].

	Ultimate Analysis wt.%		Proximate Analysis wt.%			
	Tire	Coal		Tire	Coal	
C	85.05	67.69	Moisture	1.14	7.76	
H	6.79	4.59	Fixed Carbon	32.28	47.14	
N	0.5	1.13	Volatile Matter	62.24	34.05	
O	1.75	5.48	Ash	4.35	11.05	
S	1.53	2.3	LHV	34.9	30	MJ/kg

### 3. Results and discussion

#### 3.1. Process targeting

Let us consider the case where waste tyres, are used for the direct reduction of iron ore. The material balance is based on one mole of a tyre. For simplicity of theoretical calculation, the S and N in the tyre feed are assumed to be transferred only to  $H_2S$  and  $NH_3$ . Equation 4 shows the mass balance for iron ore reduction.



$$\Delta H = 223.393 \text{ kJ}, \quad \Delta G = 150.680 \text{ kJ}$$

Equation 4 shows that the overall process requires energy ( $\Delta H > 0$  and  $\Delta G > 0$ ) for it to be feasible. A positive  $\Delta G$  across a process indicates the amount of energy equivalent to mechanical work that must be supplied for the process to be feasible, while a positive  $\Delta H$  indicates the total amount of energy to be supplied in the form of heat or work. An additional feed of oxygen must be added to combust some of the tyre feed for heat. The addition of oxygen must meet the target of  $\Delta H = 0$  and one-mole tyre. As shown in figure 1, one mole of the tyre, can reduce 0.542 mols of iron ore and produce 1.084 mols of Fe. The process is feasible with excess work, but all the carbon available in the feed is transferred to  $CO_2$ . Thus, the carbon efficiency of the process is zero. This corresponds to about 730 kg of  $CO_2$  released to the environment per ton of iron produced during this process. When compared to coal, tyres marginally perform better in terms of iron production and carbon dioxide emissions. Figure 2 shows that for one mole of coal, the iron production decreases by 5% and the carbon dioxide emissions increase to 763 kg  $CO_2$  per ton of iron produced. The Atom efficiency as well as the E-factor for the waste tyre process is 54% and 0.87 compared to coal which is 53% and 0.89. This indicates the ability of waste tyres to replace coal. However, this is not a solution for the industry if tyres can only reduce carbon emissions by 5%. An alternative route is required where the carbon emissions per ton of iron are reduced as an incentive to transition to a waste tyre feed process. One way is to consider a co-production of methanol and iron using waste tyres. In this way, not all the carbon feed in the tyre is used for reduction, but an indirect route is utilized where the tyres are converted to syngas, and the syngas is used for the reduction and the excess syngas is used to produce methanol in a downstream process unit.

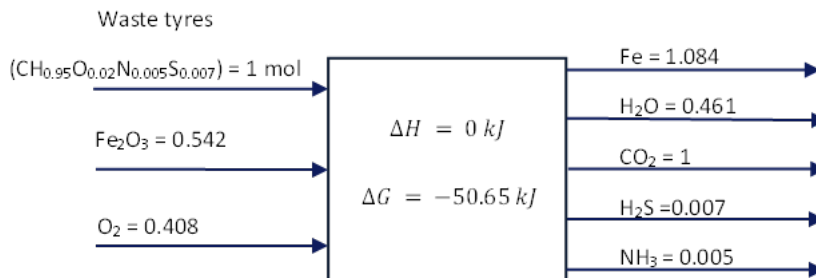


Figure 1: Reduction of iron ore with waste tyres



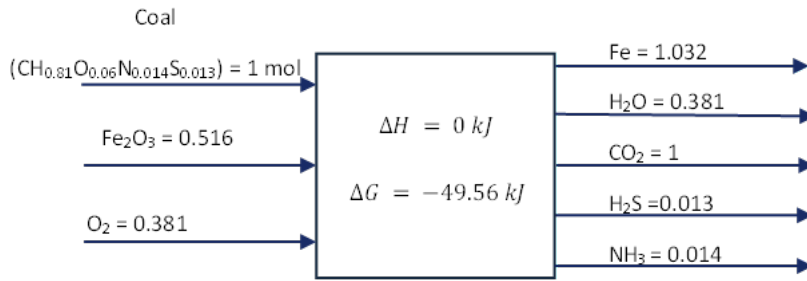
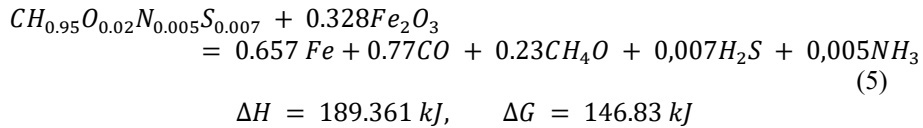
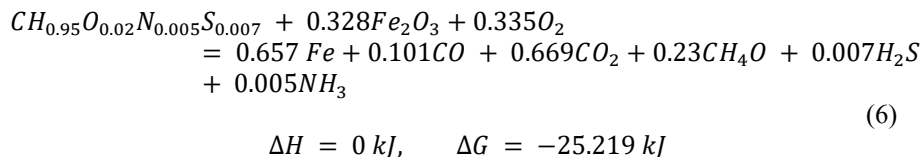


Figure 2: Reduction of iron ore with coal

For 100% raw material conversion efficiency and minimal impact on the environment, the law of conservation of mass implies that the elemental composition of the desired products must match that of the feed. Equation 5 shows the material balance for an integrated methanol synthesis and ironmaking process with 100% raw material conversion efficiency on tire feed. Tires contain a ratio of carbon to hydrogen close to one-to-one, therefore the eventual set of desired products (in combination), regardless of the process, must contain the same ratio of carbon to hydrogen, because any deviation would lead to the formation of unwanted products (i.e.,  $\text{CO}_2$ ), thus reducing the process material efficiency. The overall material balance for this process based on one mole of waste to produce methanol and iron ore from using waste tires for reduction is as follows:

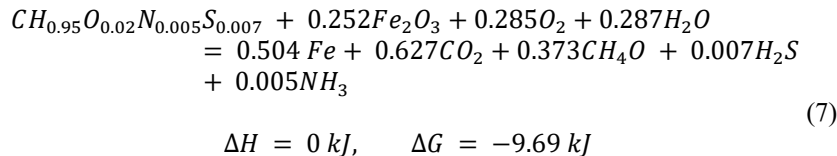


Equation 5 shows that under ideal conditions one mole of tire feed can reduce 0.328 moles of iron ore to produce 0.657 moles of iron and 0.23 moles of methanol. However, in equation 5 we see that excess oxygen from iron ore leaves as CO and that the small ratio of hydrogen to carbon in the feed constrains the full conversion of Carbon to methanol. One can immediately see from the above equation that waste tires can never achieve the 2:1 ratio of hydrogen to carbon monoxide required for liquid fuel production. To achieve this ratio an additional amount of feed containing hydrogen often water is required, however often this additional feed comes in with extra oxygen which must leave the system as carbon dioxide. Equally, the system described in equation 5 requires energy input for it to be feasible as indicated by the positive  $\Delta H$  and  $\Delta G$ . Thus, additional oxygen can be fed to achieve a zero-energy requirement as shown in eq.6.



Therefore, supplying the required energy needed by the process in eq.5 by feeding more oxygen generates 0.669 mols of carbon dioxide thus reducing the carbon efficiency to just 33% and the chemical potential efficiency decreases from 86% to 70%. In the process

described by eq.6, further optimization can be done to convert the excess CO to methanol through the addition of water. The easiest route is to burn an additional number of tires and use water to increase the ratio of hydrogen to carbon. However, this approach also leads to increased CO<sub>2</sub> emissions. Thus, the overall material balance for a hypothetical fully integrated and adiabatic process to produce methanol and iron is shown in eq.7 and figure 3:



Carbon Efficiency = 37.3%  
 Atom Economy = 59%  
 E - factor = 0.7

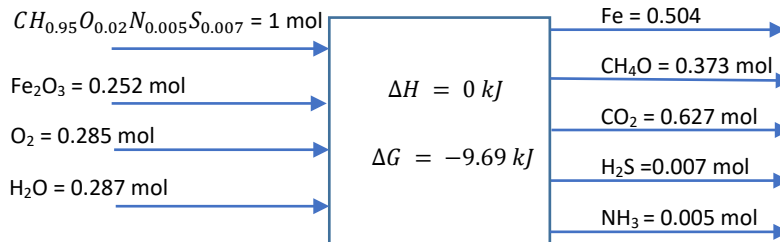


Figure 3: Process diagram for methanol and iron production from waste tires.

Eq. 7 shows that having an additional amount of water increases methanol production by 62%. The carbon efficiency is 37.3%. However, the amount of iron ore reduced decreases thus which decrease elemental iron production by 23%, and by 50% when compared to the direct routes. This, therefore, means that combined iron ore reduction with methanol synthesis with waste tires as carbon feed is achievable. The atom economy is 59% an improvement of about 10% compared to the no-methanol case discussed in figure 1. The E-factor also reduced by 25% suggesting that more feed is transferred to the desired product than left as waste. By equivalence, 1 ton of iron produced consumes 478 kg of waste tires, and 1.42 tons of iron ore and 430 kg of methanol are co-produced. The production of methanol reduces the molar carbon capacity to reduce iron ore as such, the quantity of iron reduces by almost 50% this, therefore, means that per ton of iron, the process produces 980 kg of CO<sub>2</sub>. This is still lower than the coal route which is 1.05 tons per ton of iron.

### 3.2. Effect of temperature

The process described in figure 3 is an ideal process where all units in it operate at Carnot temperatures. The Carnot temperature is the single temperature at which heat can only be used to satisfy the work requirements of a process and such a process will not lose or take additional energy to the environment, in that it will be reversible. However, in real processes reaction kinetics and materials of construction limits the use of Carnot temperature which necessitates a deviation from the ideal. The blast furnace temperature is > 1200 °C,

whereas the water gas shift and methanol synthesis reactor operate at 450 °C and 220 °C respectively. Operating at a temperature other than the Carnot temperature results in the loss of some of the potential work and the energy released from other process units cannot be integrated directly into the blast furnace due to the temperature gradient. To offset the lost work, more tires are burned. However, this reduces iron production by 15% and increases CO<sub>2</sub> emissions by 6%.

### 3.3. Revenue potential

A high-level market-related revenue potential analysis is conducted using the production rates in figure 3. The revenue potential of the process is obtained by a difference in the price of selling products and the cost of feed materials. The analysis excludes the capital costs associated with each pathway. During the time of writing this manuscript, the price of pig iron was \$ 1000 per ton, methanol was \$ 370/ ton, and iron ore was \$ 101.5 per ton. The revenue analysis of the different process pathways suggests the polygeneration of methanol and iron is the most profitable. A revenue of \$1015 per ton of iron can be achieved for the polygeneration route compared to \$855 per ton for the iron-only route. But more capital investment is required for the polygeneration due to the different units required to achieve the required products.

## 4. Conclusion

The fundamental thermodynamic analysis was used to investigate the performance limits of an idealized process of co-production of methanol and iron from a feed of waste tire and iron ore. The results show that co-production with methanol reduces CO<sub>2</sub> emissions by 37.3% (from 1 mole CO<sub>2</sub> to 0.627 CO<sub>2</sub> for the methanol case) and increases revenue by 15.6% (from 855 \$ to 1015\$). The purpose of this study is to present a method that takes energy, the environment, and economics into account, that can be used to select an acceptable process pathway. It is intended to give the process engineer a macroscopic perspective of the process and its possibilities rather than to illustrate the kinetics or the reaction path.

## References

- [1] Gupta, R. C., 2014, *Energy Resources, Its Role and Use in Metallurgical Industries*, Elsevier Ltd.
- [2] Holappa, L., 2020, "A General Vision for Reduction of Energy Steel Industry," *Metals (Basel)*, **10**, p. 1117.
- [3] Jenkin, N., 2022, "Industrial Usage Opportunities for Products Derived from Waste Tyres in South Africa," (April 2022), pp. 1–60.
- [4] Portofino, S., Casu, S., Iovane, P., Russo, A., Martino, M., Donatelli, A., and Galvagno, S., 2011, "Optimizing H<sub>2</sub> Production from Waste Tires via Combined Steam Gasification and Catalytic Reforming," *Energy and Fuels*, **25**(5), pp. 2232–2241.
- [5] Mavukwana, A. E., Fox, J. A., and Sempuga, B. C., 2020, "Waste Tyre to Electricity: Thermodynamics Analysis," *J. Environ. Chem. Eng.*, **8**(4).
- [6] Mavukwana, A. E., Stacey, N., Fox, J. A., and Sempuga, B. C., 2021, "Thermodynamic Comparison of Pyrolysis and Gasification of Waste Tyres," *J. Environ. Chem. Eng.*, **9**(2), p. 105163.
- [7] Martínez, J. D., Puy, N., Murillo, R., García, T., Navarro, M. V., and Mastral, A. M., 2013, "Waste Tyre Pyrolysis - A Review," *Renew. Sustain. Energy Rev.*, **23**, pp. 179–213.

## Unlocking the potentials of integrating direct air capture with HVAC system

Yasser M. ABDULLATIF \*1, 2, Odi Fawwaz Alrebei<sup>2</sup>, Ahmed Sodiq<sup>2</sup>, Tareq AL-ANSARI<sup>1</sup>, and Abdulkarem I. Amhamed\*<sup>2</sup>

<sup>1</sup> College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Education City, Doha, Qatar.

<sup>2</sup> Qatar Environment and Energy Institute (QEERI), Doha, Qatar.

\*Corresponding author's email: [Yabdellatif@hbku.edu.qa](mailto:Yabdellatif@hbku.edu.qa) and [aamhamed@hbku.edu.qa](mailto:aamhamed@hbku.edu.qa)

### Abstract

The increasing concentration of carbon dioxide (CO<sub>2</sub>) in the atmosphere has compelled researchers and policymakers to seek urgent solutions to address the current global climate change challenges. In order to keep the global mean temperature at approximately 1.5°C above the preindustrial era, the world needs increased deployment of negative emission technologies. Among all the negative emissions technologies reported, direct air capture (DAC) is positioned to deliver the needed CO<sub>2</sub> removal in the atmosphere. DAC technology is independent of the emissions origin, and the capture machine can be located close to the storage or utilization sites or in a location where renewable energy is abundant or where the price of energy is low. Notwithstanding these inherent qualities, DAC technology still has few drawbacks that need to be addressed before the technology can be widely deployed. The main findings point to undeniable facts that DAC overall system energy requirement is high, and it is the main bottleneck in DAC commercialization. The proposal of this study is to unlock the potential of integrating DAC with HVAC systems. DAC with HVAC system integration can be proposed to allow for reducing indoor CO<sub>2</sub> levels to improve indoor air quality (IAQ), HVAC energy reduction using higher air recirculation ratios, a more efficient DAC system by capturing from more elevated CO<sub>2</sub> concentrated streams (indoor air), applying cooler adsorption, and adsorption at different relative humidity points. The expected reduction in the net present value of tons of CO<sub>2</sub> after 50 years is estimated to be 3000\$/tonCO<sub>2</sub> compared to the highest in the literature of 1000 \$/tonCO<sub>2</sub> [1]. This reduction in NPV is estimated based on possible energy recovery by quantifying only one benefit which is HVAC energy reduction using higher air recirculation ratios. Therefore, coupling DAC with HVAC can be economically favorable in comparison to only DAC. In the present work, a comparison between efficiencies of the stand-alone DAC unit and the one integrated with Air Handling Unit are quantitatively evaluated based on validated mathematical model. The positions of direct air capture unit within the HVAC are chosen to maximize the possible benefits. The DAC unit efficiencies throughout the hot months in Qatar are improved to be within the range of 21% to 22% compared to the range of 16% to 21% by only positioning the DAC at point 2 in the Air Handling Unit.

**Keywords:** DAC; green sorbents; HVAC; indoor air quality

## 1 Introduction

The high energy requirements of DAC lead to high capital and operating costs, which is a major challenge of the existing technologies [2]. As a possible solution, the DAC and HVAC integration was proposed as it brings benefits for both subsystems. These possible benefits are; reducing indoor CO<sub>2</sub> levels to improve indoor air quality (IAQ), HVAC energy reduction using higher air recirculation ratio and finally, more efficient DAC system by capturing from more elevated CO<sub>2</sub> concentrated streams (indoor air) and applying cooler adsorption. The integration between direct air capture unit and HVAC system is shown in Figure 2.

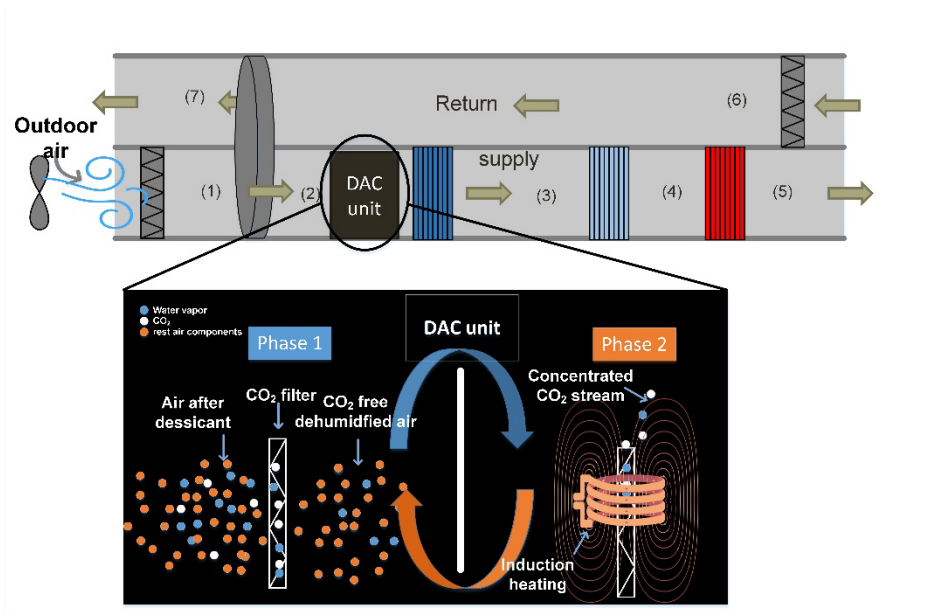


Figure 1: Air handling unit and Direct air capture Integration scheme

The current work includes two subsystems, which are a direct air capture unit and a building air handling unit. Both systems will be thermodynamically simulated for evaluating their performance and energy consumption. These two systems are expected to have better performance when combined, which will be assessed in the present work. The main AHU parameters and setpoints are presented in Table 1. A benchmark direct air capture cycle that uses temperature vacuum swing regeneration will be used in the present study. This cycle was chosen as its model accounts for the role of humidity, it uses commercial adsorbent material, and the model was experimentally validated [3].

Table 1: Air Handling Unit technical data during hot months obtained from Qatar foundation facilities

Parameter	Value
Humidity ratio	Same as point after energy recovery wheel
Relative humidity After precooler (%)	95
Dry Bulb temperature After Cooling coil (°C)	13.2
Relative humidity After Cooling coil (%)	95
Dry Bulb temperature After Reheater (°C)	21.66
Relative humidity After Reheater (%)	56
Air pressure drop through the Cooling coil (Pa)	245
Supply volume flow rate (m <sup>3</sup> /h)	39906
Exhaust volume flow rate (m <sup>3</sup> /h)	20500
Pressure drop through energy wheel (Pa)	239
Sensible energy wheel effectiveness $\varepsilon_s$	0.40
Latent energy wheel effectiveness $\varepsilon_L$	0.32

### 1.1 Methodology/Analysis

The main components of air handling unit such as energy recovery wheel, precooler, cooler, heater and fan was simulated based on the preset points in Table 1 using mass balance equation, first and second laws of thermodynamics, properties such as enthalpy, entropy, and specific exergy are determined for each component. The general energy balance equation for each component is described in [4].

$$\dot{m}_i \times h_i + \dot{Q}_i + \dot{W}_i = \dot{m}_o \times h_o + \dot{Q}_o + \dot{W}_o \quad (1)$$

The sensible and latent energy recovery wheel effectiveness  $\varepsilon_s$ ,  $\varepsilon_L$  were used to calculate temperature and humidity after the energy wheel [5].

$$\varepsilon_s = \frac{\dot{m}Cp}{(\dot{m}Cp)_{max}} \frac{T_{s,o} - T_{s,i}}{T_{s,i} - T_{e,i}} \quad (2)$$

$$\varepsilon_t = \frac{\dot{m}Cp}{(\dot{m}Cp)_{max}} \frac{\omega_{s,o} - \omega_{s,i}}{\omega_{s,i} - \omega_{e,i}} \quad (3)$$

Where  $\dot{m}$  is air mass flow rate in (kg/s),  $h$  is the enthalpy in (J/(kg K)),  $\dot{Q}$  is heat in kW,  $\dot{W}$  is the power in kW,  $Cp$  is the air specific heat in (J/(kg K)),  $T$  is the temperature in (K) and  $\omega$  is the humidity ratio in (g/kg). The subscripts s, e, i and o stand for supply, exhaust, inlet, and outlet respectively.

The adsorbent isotherm was simulated based on mechanistic co-adsorption model, which accounts for the humidity effect on CO<sub>2</sub> adsorption and was validated by experimental work [3]. The energy required for regeneration was calculated from work in reference [6]. The performance of DAC unit in different position within AHU was evaluated. The proposed positions are stand-alone DAC and at state points 2, 4, 5, and 6 which were chosen to minimize adsorption temperature and humidity while maximize achieving other benefits such as better IAQ and less HVAC cooling load. To assess the performance of DAC, the second-law efficiency of TVSA cycle-based DAC system was calculated as follow [7]:

$$\eta_{TVSA} = \frac{W_{min}}{W_{pump} + Q_H \left(1 - \frac{T_0}{T_H}\right) - Q_L \left(1 - \frac{T_0}{T_L}\right)} \quad (4)$$

$\eta_{TVSA}$  is the efficiency of DAC unit uses temperature vacuum swing adsorption,  $Q_H$  is the heat input per cycle,  $Q_L$  is the cooling load required for reaching the cycle adsorption temperature and  $W_{pump}$  is the work required for reaching the vacuum pressure,  $T_0$  is the ambient temperature,  $T_H$  is the desorption temperature and  $T_L$  is the adsorption temperature.

## 2 Results and discussion

The efficiencies of the direct air capture unit based on different relative humidity points and temperatures were shown in Figure 1. The range of temperatures and relative humidity values were chosen to cover the weather data for Qatar and the points within the air handling unit. The efficiencies in Figure 2 below were calculated for temperature vacuum swing adsorption system and a chemisorption material.

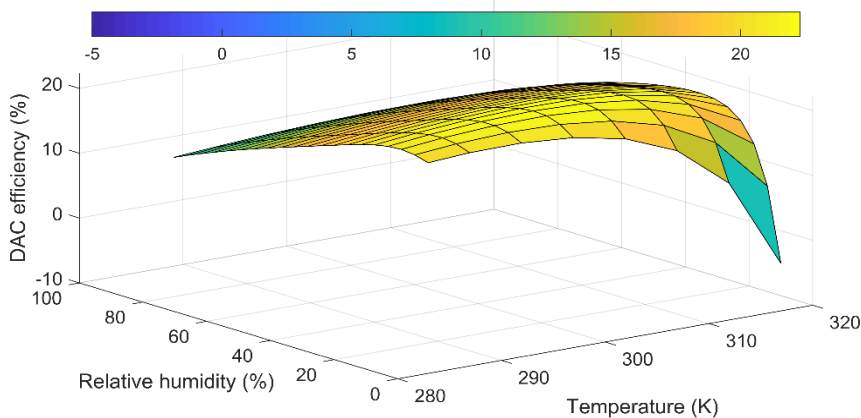


Figure 2: Direct Air Capture efficiency for different temperatures and relative humidity values

Figure 1 shows that the DAC efficiency increases as the adsorption temperature and relative humidity decrease. Based on the mentioned trend, three position of DAC in AHU were chosen and assessed to get a more efficient DAC system.

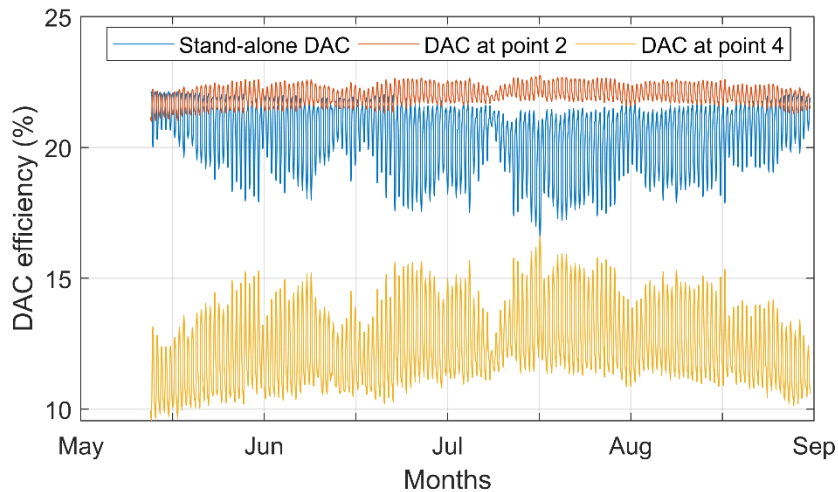


Figure 3: DAC efficiency for different positions in AHU

The efficiencies of three direct air capture units and their variation through the hot months in Qatar were presented in Figure 3. The three positions equally achieve lower indoor CO<sub>2</sub> levels and less HVAC cooling load by allowing air recirculation however, the adsorption conditions such as temperature and relative humidity is different for each position leading to different efficiencies. The calculated efficiency in Figure 3 considers the effect of these adsorption conditions only without considering other benefits. The figure shows that the best position between the three is at point 2. Although the temperature is lowest at point 4, the DAC unit efficiency was low because of the high relative humidity reaching 95%. The efficiency in position 2 was better than the stand-alone DAC unit because in position 2, the temperature and relative humidity is always lower than the ambient temperature during the hot months in Qatar. The analysis shows that the range of efficiencies can be raised from 16%-21% to 21% -22% by only positioning the DAC unit in point 2 in the Air Handling Unit.

### 3 Conclusion

In conclusion, buildings consume almost 40% of the world's energy demands, and about 50 to 82% of buildings' energy is consumed in the HVAC systems. The necessity to provide fresh indoor air, especially during and after COVID-19, and the rising energy demands, challenge researchers to reduce energy losses in the HVAC systems. The DAC and HVAC integration provides a solution for



achieving better indoor air quality, lower HVAC energy consumption and a more efficient DAC system. The present work investigated the effect of positioning the DAC unit within HVAC system and provided a higher efficiency prediction, however, the other benefits need to be quantified and a technoeconomic analysis for the integration should be conducted to compare the integration to the normal DAC system. Moreover, improving the adsorbent material CO<sub>2</sub> uptake and affinity to water could further enhance the performance of the integration.

## References

- [1] L. Baus, and S. Nehr, "Potentials and limitations of direct air capturing in the built environment," *Building and Environment*, vol. 208, pp. 108629, 2022.
- [2] O. S. Board, E. National Academies of Sciences, and Medicine, *Negative emissions technologies and reliable sequestration: A research agenda*: National Academies Press, 2019.
- [3] J. Young, E. García-Diez, S. Garcia, and M. van der Spek, "The impact of binary water–CO<sub>2</sub> isotherm models on the optimal performance of sorbent-based direct air capture processes," *Energy & Environmental Science*, vol. 14, no. 10, pp. 5377-5394, 2021.
- [4] Y. M. Abdullatif, E. C. Okonkwo, Y. Bicer, and T. Al-Ansari, "Thermodynamic analysis of gravity assisted solar-powered reverse osmosis unit for greenhouses situated in a depleted zone," *Case Studies in Thermal Engineering*, vol. 25, pp. 100990, 2021/06/01/, 2021.
- [5] A. S. Al-Ghamdi, "ANALYSIS OF AIR-TO-AIR ROTARY ENERGY WHEELS."
- [6] J. A. Wurzbacher, C. Gebald, N. Piatkowski, and A. Steinfeld, "Concurrent Separation of CO<sub>2</sub> and H<sub>2</sub>O from Air by a Temperature-Vacuum Swing Adsorption/Desorption Cycle," *Environmental Science & Technology*, vol. 46, no. 16, pp. 9191-9198, 2012/08/21, 2012.
- [7] R. Zhao, L. Liu, L. Zhao, S. Deng, S. Li, Y. Zhang, and H. Li, "Thermodynamic exploration of temperature vacuum swing adsorption for direct air capture of carbon dioxide in buildings," *Energy Conversion and Management*, vol. 183, pp. 418-426, 2019.

# Model development of amine scrubbing and CO<sub>2</sub> liquefaction process using waste LNG cold energy

Jonghun Lim<sup>a,b</sup>, Yurim Kim<sup>a,b</sup>, Yup Yoo<sup>a,b</sup>, Hyungtae Cho<sup>a</sup>, Jaewon Lee<sup>a</sup>,  
Jinwoo Park<sup>a,c</sup>, Junghwan Kim<sup>a,b,\*</sup>

<sup>a</sup>*Green Materials and Processes R&D Group, Korea Institute of Industrial Technology, 55, Jonga-ro, Ulsan, 44413, Republic of Korea*

<sup>b</sup>*Department of Chemical and Biomolecular Engineering, Yonsei University, 50, Yonsei-ro, Seoul, 03722, Republic of Korea*

<sup>c</sup>*Department of Chemical Engineering, Kongju National University, 1223-24, Cheonan-daero, Seobuk-gu, Cheonan-si, Chungcheongnam-do, 31080, Republic of Korea*

## Abstract

Amine scrubbing and CO<sub>2</sub> liquefaction processes are generally employed in most power plants for CO<sub>2</sub> capture and storage (CCS) because they are suitable for large-scale plants. However, a significant amount of cold energy is required to cool the lean amine solution in the amine scrubbing system and liquefy compressed CO<sub>2</sub>. Hence, liquefied natural gas (LNG) cold energy, which is generally wasted in seawater during LNG regasification, should be recovered. In this study, we developed an amine scrubbing and CO<sub>2</sub> liquefaction process using waste LNG cold energy. The proposed process model comprises the following three steps: (1) LNG regasification, (2) CO<sub>2</sub> capture and regeneration, and (3) CO<sub>2</sub> liquefaction. First, in LNG regasification, the boil-off gas (BOG) is compressed, and the LNG with compressed BOG is regasified in an LNG vaporizer. To regasify the LNG, seawater is used as a heat source, and the waste cold energy in the chilled seawater is recovered in steps 2 and 3. Second, the CO<sub>2</sub> capture and regeneration process involves the use of an absorber and a regenerator, and the waste cold energy is recovered at the lean amine cooler to reduce the temperature of the lean amine solution. Finally, to liquefy CO<sub>2</sub>, cold energy derived from LNG regasification is recovered to a heat exchanger to cool the compressed CO<sub>2</sub>. Consequently, the total cold energy consumption can be decreased by 31% by recovering waste LNG cold energy. This increases the overall energy efficiency of the amine scrubbing and CO<sub>2</sub> liquefaction processes for CCS using waste LNG cold energy, thus resulting in less fuel use. We believe that the proposed approach will afford significant economic improvements and environmental protection effects.

**Keywords:** Amine scrubbing, CO<sub>2</sub> liquefaction, Liquefied natural gas, CO<sub>2</sub> capture and storage

## 1. Introduction

In many countries, natural gas (NG) emits less air pollutants than other fossil fuels and is considered the energy resource that is developing the most rapidly. NG resources are generally located in specific areas; thus, to ensure the safe long-distance transportation of

NG, it is liquefied by cooling to reduce its volume. However, a significant amount of cold energy is wasted during liquefied natural gas (LNG) regasification; thus, waste LNG cold energy recovery is crucial for improving energy efficiency. Waste LNG cold energy derived from LNG regasification can be used for CO<sub>2</sub> capture and storage (CCS). In general, amine scrubbing and CO<sub>2</sub> liquefaction processes are employed for CCS in LNG power plants. The amine scrubbing system is primarily composed of an absorber and a regenerator, and CO<sub>2</sub> capture depends on the variation in equilibrium constants based on the pKa difference between the amine and CO<sub>2</sub> via temperature control. The CO<sub>2</sub> captured during amine scrubbing is liquefied, which requires a significant amount of cold energy. To control the temperature of the absorber during amine scrubbing and liquefy CO<sub>2</sub>, waste LNG cold energy should be recovered. However, most previous studies have only focused on waste hot energy recovery, which is generated from LNG power plants, to reduce the steam energy consumed by the regenerator in the amine scrubbing system. Dual et al. conducted an optimization study to reduce steam energy consumption by a regenerator reboiler. By recovering waste hot energy from an LNG power plant, the energy consumption of the regenerator reboiler was reduced by 56%. Garlapalli et al. proposed a heat recovery unit that directly recovers waste hot energy in flue gas for CCS. Using this unit, the steam energy consumption of the regenerator reboiler can be reduced by 15% to 31%. Despite the significant amount of relevant studies performed hitherto, the following challenges are yet to be addressed. Although waste LNG cold energy can be recovered during amine scrubbing and CO<sub>2</sub> liquefaction, it was not recovered in previous studies. Cold energy constitutes approximately 50% of the total energy required. To increase the overall energy efficiency, waste LNG cold energy must be recovered. Hence, a process model for amine scrubbing and CO<sub>2</sub> liquefaction using waste LNG cold energy is developed in this study. The objective of this study is to overcome the challenge of conventional studies to increase the overall energy efficiency of LNG power plants for CCS by efficiently recovering waste LNG cold energy, which will ultimately result in less fuel use.

## 2. Methods

This section describes the process model developed for amine scrubbing and CO<sub>2</sub> liquefaction using waste LNG cold energy. Section 2.1 presents an overview of the proposed process, and Section 2.2 presents a detailed description of the process model.

### 2.1. Overview of proposed process

Figure 1 shows a simplified diagram of the developed amine scrubbing and CO<sub>2</sub> liquefaction process model based on waste LNG cold energy. The proposed process comprises three steps: (1) LNG regasification, (2) amine scrubbing, and (3) CO<sub>2</sub> liquefaction. First, in the LNG regasification process, the transported LNG is assumed to be burned with boil-off gas (BOG) in an LNG power plant. Second, the CO<sub>2</sub> in the flue gas emitted at the combustor of an LNG power plant is captured by the amine scrubbing system. In the amine scrubbing system, waste LNG cold energy is recovered at the lean amine cooler. Finally, the captured CO<sub>2</sub> is liquefied during the CO<sub>2</sub> liquefaction process, and waste LNG cold energy is used to reduce the temperature of the captured CO<sub>2</sub>.

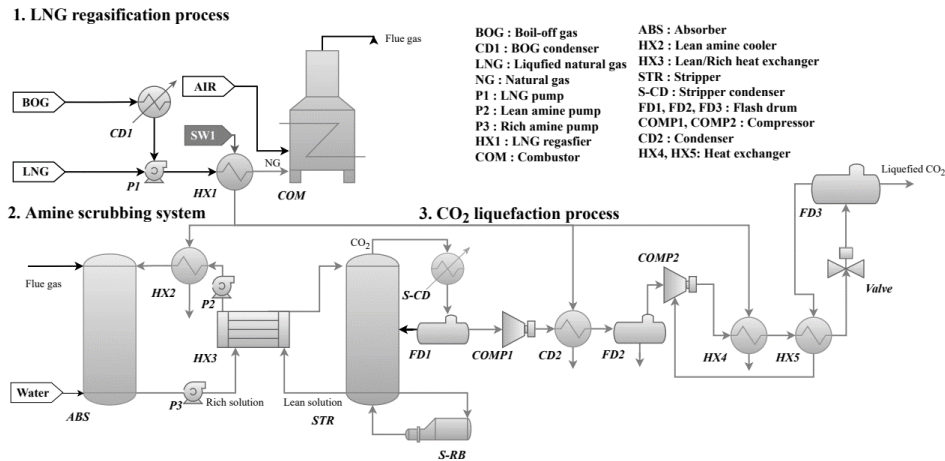


Figure 1. Simplified diagram of proposed process

## 2.2. Process model

Figure 2 shows the proposed process model for amine scrubbing and CO<sub>2</sub> liquefaction using waste LNG cold energy. To design the proposed model, Aspen Plus V10.0, and ELECNRTL thermodynamic models were employed to consider the liquid–liquid equilibria of amine solution with captured CO<sub>2</sub> and vapor–liquid equilibria of amine solution with flue gas. Subsequently, the following assumptions were considered for model development: (1) The developed process was in a steady state; (2) the temperature of the seawater for the LNG regasifier was 15 °C; (3) the LNG composition was CH<sub>4</sub> (97 wt%), C<sub>2</sub>H<sub>6</sub> (1 wt%), C<sub>3</sub>H<sub>8</sub> (1 wt%), and N<sub>2</sub> (1 wt%).

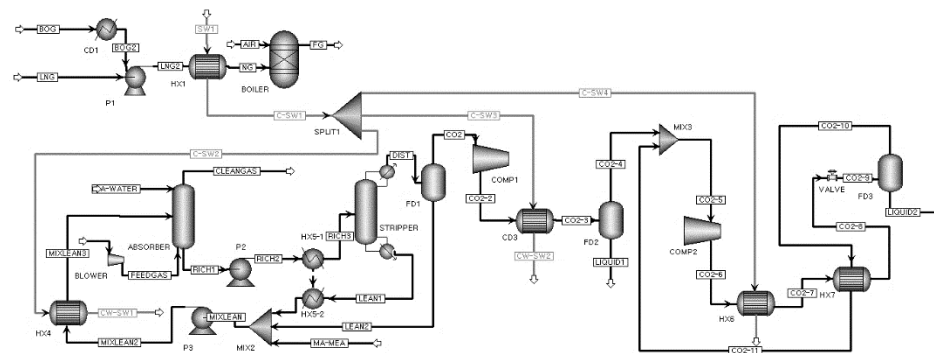


Figure 2. Developed process model of amine scrubbing and CO<sub>2</sub> liquefaction using waste LNG cold energy

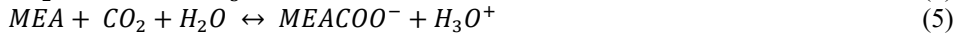
### 2.2.1. LNG regasification

First, to obtain the driving force that can be supplied to an LNG power plant, the transported LNG is input to a high-pressure LNG pump with a BOG. In this pump, the pressure of the LNG is increased from 1 to 60 bar. Subsequently, the compressed LNG is input to an LNG vaporizer, and the temperature of the LNG is increased from -159 °C to

20 °C using seawater as the heat source. During LNG vaporization, a significant amount of LNG cold energy is wasted via dissipation into the sea. Thus, the waste LNG cold energy is recycled to the lean solution cooler (HX2) and each heat exchanger (CD2 and HX4) during LNG liquefaction.

### 2.2.2. Amine scrubbing system

The CO<sub>2</sub> from the LNG power plant is captured by the amine scrubbing system, which uses monoethanolamine (MEA) as a CO<sub>2</sub> absorbent. An amine scrubbing system is generally composed of an absorber and a regenerator. In the absorber, CO<sub>2</sub> is captured using a lean amine solution via an exothermic reaction (Equations 1–4).



The rich amine solution that captures CO<sub>2</sub> is discharged into the lower section of the absorber and enters the regenerator. Subsequently, it is preheated via the lean/rich heat exchanger (HX3) before entering the regenerator, which reduces the heat load on the regenerator reboiler. Finally, the rich amine solution is regenerated at the regenerator by supplying heat energy from the reboiler.

### 2.2.3. LNG liquefaction

Finally, the absorbed CO<sub>2</sub> is liquefied, and the liquefied CO<sub>2</sub> requires a significant amount of cold energy for storage. In this study, a Linde dual pressure system that utilizes a cold gas stream, which is generated by controlling a Joule–Thompson valve to cool CO<sub>2</sub>, is employed for CCS. First, the captured CO<sub>2</sub> is compressed at a low-pressure compressor (COMP1); subsequently, a high-pressure compressor (COMP2) is used to increase the pressure of the CO<sub>2</sub>, which contains CO<sub>2</sub> recycled from the flash drum (FD1). Subsequently, the temperature of the compressed CO<sub>2</sub> is decreased at each heat exchanger (HX4 and HX5) using waste LNG cold energy. Finally, the cooled compressed CO<sub>2</sub> expands through the Joule–Thomson valve (valve).

## 3. Results and discussion

This section presents the simulation results obtained using the proposed amine scrubbing and CO<sub>2</sub> liquefaction process model, which is based on waste LNG cold energy. The process for the base case involving the non-recovery of waste LNG cold energy from LNG regasification is compared with the proposed process. Table 1 shows the simulation results obtained using the proposed and base processes. Based on the results, the energy consumption of the cooler (HX2), condenser (CD2), and each heat exchanger (HX4 and HX5) was completely substituted by waste LNG cold energy. The total cold energy consumption decreased from 261 to 179 MW; thus, the cold energy saving rate was calculated to be 31% based on the recovery of the waste LNG cold energy. Finally, to conserve waste LNG cold energy, the total energy consumed by the proposed process should be 655 MW, which is 11% lower than that consumed by the base process.

Table 1. Simulation results of proposed and base processes

Classification		Base process	Proposed process	Units
LNG regasification process	CD1 (Condenser)	1.5	1.5	kW
	P1(Pump)	13.7	13.7	MW
CO <sub>2</sub> capture process	HX2 (Lean amine cooler)	13.9	0	MW
	P2 (Pump)	0.5	0.5	MW
	P3 (Pump)	3.8	3.8	kW
	Regenerator reboiler	420	420	MW
	Regenerator condenser	179	179	MW
CO <sub>2</sub> liquefaction process	COMP1 (Compressor)	30	30	MW
	COMP2 (Compressor)	12	12	MW
	CD2 (Condenser)	35	0	MW
	HX4,5(Cooler)	33	0	MW
Total energy consumption		741	659	MW

Figure 3 shows the total energy consumed by the proposed and base processes. By operating the condenser (CD1) and pump (P1), the energy consumed by the base and proposed processes was 14 MW each. In the amine scrubbing system, the energy consumed by the base process was 613 MW. Meanwhile, the energy consumed by the proposed process was calculated by 600 MW. Thus, based on the recovery of waste LNG cold energy to the lean amine cooler (HX2), the energy consumed by the proposed process is 2% lower than that consumed by the base process. The energy consumed during CO<sub>2</sub> liquefaction based on the base and proposed processes was 110 and 42 MW, respectively. Based on the recovery of the waste LNG cold energy to the condenser (CD2) and each heat exchanger (HX4 and HX5), the energy consumed during CO<sub>2</sub> liquefaction based on the proposed process was 61% lower than that consumed by the base process. Finally, the total energy consumed by the proposed process was 11% lower than that consumed by the base process, which indicates a high energy efficiency. Therefore, we believe that this study can provide valuable insights into the efficient recovery of waste LNG cold energy for CCS.

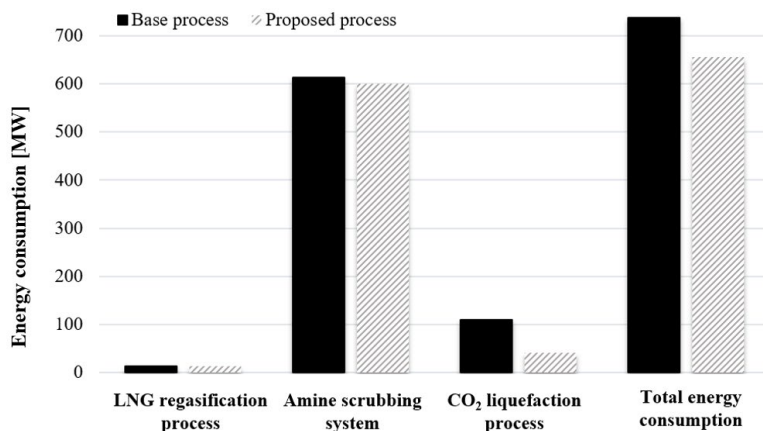


Figure 3. Total energy consumed by proposed and base processes

#### 4. Conclusion

In this study, novel amine scrubbing and CO<sub>2</sub> liquefaction processes using waste LNG cold energy was developed. The main contributions of this study are twofold. First, waste LNG cold energy wasted during LNG regasification was efficiently reused for amine scrubbing and CO<sub>2</sub> liquefaction; thus, a higher overall energy efficiency was obtained. Second, the results showed a higher energy efficiency compared with that afforded by the base process, wherein waste LNG cold energy was not recovered; thus, the proposed processes can result in significant energy and economic improvements. Additionally, the results showed that the energy consumed during CO<sub>2</sub> amine scrubbing and CO<sub>2</sub> liquefaction based on the proposed process was lower than that based on the base process by 2% and 61%, respectively; thus, the total energy consumed by the proposed process was 11% lower than that consumed by the base process. Therefore, we believe that this study can provide valuable insights into the efficient recovery of waste LNG cold energy for CCS.

#### Acknowledgement

This research was supported by the Korean Institute of Industrial Technology within the framework of the following projects: “Development and application of carbon-neutral engineering platform based on carbon emission database and prediction model [grant number KM-22-0348] and Development of complex parameter smart analysis modules for color customering [grant number EH-22-0011]”.

#### References

- J Lim, J Lee, I Moon, H Cho, J Kim, Techno-economic comparison of amine regeneration process with heat-stable amine salt reclaiming units, *Energy Science & Engineering*, 9, 12, 2529-2543.
- J Lim, J Lee, H Cho, J Kim, Model development of amine regeneration process with electro dialysis reclamation unit, *Computer Aided Chemical Engineering*, 50, 579-584.
- Y Kim, J Lim, JY Shim, H Lee, H Cho, J Kim, Optimizing wastewater heat recovery systems in textile dyeing processes using pinch analysis, *Applied Thermal Engineering*, 214, 118880.

# A parametric analysis of the co-gasification of biomass and plastic waste using Aspen Plus

Nomadlozi L. Khumalo<sup>a</sup>, Bilal Patel<sup>a</sup>,

*Department of Chemical Engineering, University of South Africa, Corner of Christiaan de Wet Road and Pioneer Avenue, Florida 1709, South Africa*

## Abstract

The syngas produced from biomass (BM) gasification is typically of low quality and energy content. This inhibits it from being utilized for downstream application for manufacturing of high value products. Plastic waste (PW) is a suitable co-feed since it possesses useful properties, such as high hydrogen content, low oxygen, and low moisture content.

In this study, the effects of blend ratios of (25% PE - 50% PE), and operating condition such as Equivalence Ratio (ER)) on the product gas composition, H<sub>2</sub>/CO ratio and the LHV, were evaluated using an oxygen-steam mixture as a gasifying agent. A non-stoichiometric equilibrium model for the co-gasification of BM and PW was developed in Aspen Plus. The co-gasification of BM and PW has led to high Lower Heating Value (LHV) of 9.6 MJ/Nm<sup>3</sup> and H<sub>2</sub>/CO ratio of 2 was achieved at low ER values.

**Keywords:** co- gasification, blend ratio, equivalence ratio; synergistic effect.

## 1. Introduction

Biomass is an attractive renewable resource, particularly since it is the only renewable source of carbon, that can be used to generate carbon- based fuels and chemicals. Thus, biomass can be utilized to produce a wide variety of products including biofuels, chemicals, bioproducts and energy. There are various technologies to convert biomass, however, among these technologies, gasification is a versatile technology that allows the production of a wide range of products. Gasification process is a thermochemical process that converts carbonaceous material into syngas in the presence of a gasifying agent (s) at high temperatures. The syngas needs to be of good quality and high energy content, so that it can be suitable for downstream applications. However, the syngas produced from biomass gasification has low energy density and low quality, for instance its H<sub>2</sub>/CO ratio needs to be adjusted by adding a feedstock with high hydrogen content to increase in quality.

Plastic waste has been identified as a suitable gasification feedstock It is estimated that global plastic usage will triple by 2060, with almost half ending up in landfills and less than a fifth recycled. (OECD, 2022). Therefore, valorization of waste plastic, via gasification, is seen as a viable option for large-scale implementation (Lopez et al. 2018). Plastic waste can compensate for qualities lacking in biomass such as high oxygen content, low energy density. Plastic waste has high volatile matter and is rich in hydrogen



and does not contain oxygen (Al Amoodi et al. 2013). When biomass and plastic waste are combined the syngas produced is of improved quality and has high energy density, as compared to that of pure biomass (Singh et al. 2022). The blended feedstocks of biomass and plastic waste may enhance hydrogen content in the syngas, resulting in an increased energy density of the product gas (Huber et al. 2006). During the co-gasification technology tar is reduced in the syngas as the plastic percentage increases in the feedstock mixture (Lopez et al. 2015). The quantity of syngas and the composition (e.g.  $H_2/CO$ ) ratio improves (Block et al. 2019). Co-gasification of biomass and plastic waste also assists in overcoming difficulties of seasonal biomass availability. Moreover, co-gasification eliminates operational challenges that are encountered during plastic waste gasification.

This study addresses a gap in literature in terms of determining the effect of using oxygen - steam mixture as a gasifying agent instead of single gasifying agent during the co-gasification of biomass (pine sawdust) and plastic waste (Low Density Polyethylene). Previous studies such as Pinto et al. (2002), Burra et al. (2018) and Ahmed et al. (2011), have considered the co-gasification of biomass and plastic waste, typically, using one gasifying agent such as air, steam, or oxygen, but not mixtures of gasifying agents. The effect of feedstock composition (biomass and plastic waste and their mixtures (25% PE – 50% Biomass)) in the presence of oxygen – steam mixture as gasifying agent on the product gas composition ( $H_2$ , CO), on the  $H_2/CO$  ratio of the syngas, on the Lower Heating Value (LHV) of the syngas, gas yield (GY) and on the cold gas efficiency (CGE). The synergistic interaction between the biomass and plastic waste will be evaluated. This study also aims to determine the effect of the equivalence ratio at a fixed SFR of 0.6 At value of SFR equal to 0.6 is considered optimal for the steam flowrate to enhance the steam reforming reaction and water gas shift reaction and the production of  $H_2$  and CO in the product gas (Arjoloo et al. 2022). Increasing SFR above 0.6, leads to a decrease in process temperature and that affects the syngas quality as well as it reduces the process efficiency.

## 2. Modelling framework

Aspen Plus was used to develop a non- stoichiometric equilibrium model to determine the effect of feedstock composition and operating conditions such as the equivalence ratio on the product gas composition,  $H_2/CO$  ratio of the syngas and the Lower Heating value (LHV) of the syngas. The synergistic interaction between biomass and plastic waste is also evaluated. The ultimate analysis of biomass (sawdust) was determined using an elemental analyser (C; 45.5%, H; 5%, O; 47.1%, N;0.05 % and 2.35% Ash) and the proximate analyses using a thermogravimetric method (18.45%, VM, 79.2% and Ash 2.35%). Data for plastic waste (low-density polyethylene) was sourced from the Al Amoodi et al. (2013).

Various components such as  $H_2$ ,  $H_2O$ , S, N,  $O_2$ ,  $CH_4$ , where specified as conventional, whereas, biomass, ash and plastics were specified as non- conventional species. HCOALGEN and DCOALIGT algorithm were used as enthalpy and density functions for biomass and plastic waste. The Peng-Robinson – Boston – Mathias (PR-BM) thermodynamic model was employed. This property method was selected based on its ability to correlate the low operating pressure with the higher operating temperatures  $>700^\circ C$  that are typically used in the co- gasification process (Ramzan et al. 2011). In addition, the process was assumed to operate at steady state, the gasifier is operated at

atmospheric pressure, tar and carbon formation is not considered. The model was developed based on three-unit blocks; the first unit is the RYield reactor which represents the decomposition of the non-conventional components into conventional constituents. A FORTRAN statement is utilised within Aspen Plus to accomplish this. The ultimate analyses of the feedstocks is used to determine the yield distribution. The product stream from the RYield reactor enters the second unit, modelled by the RGibbs reactor, uses the minimization of Gibbs energy to predict the product distribution from the gasifier. Gasifying agents such as steam and oxygen were introduced to the gasifier (RGibbs). The equivalence ratio (ER) and steam – to – fuel ratio (SFR) are the operating parameters that controls the amount of oxygen and steam that enter the gasifier. The gasification products from RGibbs reactor (mainly syngas) are sent to the third unit (Cyclone) which separates solids from gaseous product. The validity of the model was assessed using the relative error ( $\epsilon_r$ ) method. The output values of the product gas obtained from the experimental work (Yu et al. 2014), were compared with the values obtained from Aspen Plus for the co-gasification of biomass and plastic waste. Based on the relative error, the model and experimental values for the syngas composition are in good agreement and are considered acceptable i.e. (the relative errors for each of the variable were within +/- 15%) (De Andres et al. 2019).

### 3. Results and Discussion

#### 3.1 Effect of the equivalence ratio on H<sub>2</sub> and CO composition in the product gas for various blend ratios

Figure 3.1 (a) shows that an increase in the equivalence ratio (ER) at fixed SFR equal to 0.6 causes the H<sub>2</sub> to increase and reach maximum values and thereafter decrease as ER increases further, for feedstocks that consists of higher plastic percentage such as 50% PE and 100% PE. In Figure 3.1(b) the same pattern is observed for the CO composition for all the blended feedstocks except for biomass feedstock. It is observed that high H<sub>2</sub> and CO composition in the product gas are favoured at low ER values below 0.4. This is because at low ER values the chemical reactions such as the steam reaction, partial oxidation reaction, water – gas shift reaction, and steam reforming reaction and methanation reaction, which are responsible for the production of H<sub>2</sub> and CO are favoured.

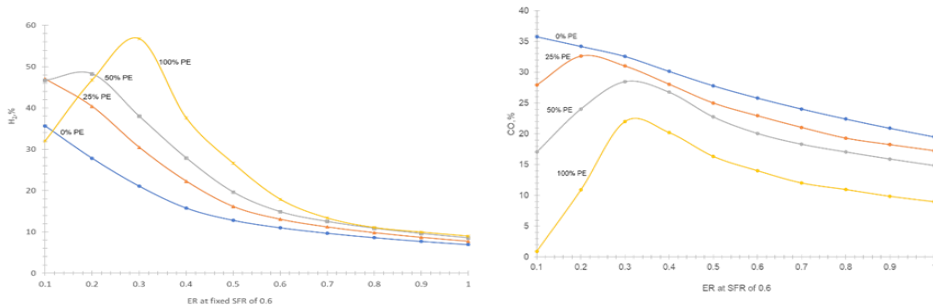


Figure 3.1 Effect of the (a) equivalence ratio (ER) on the (a) H<sub>2</sub> and (b) CO (b) composition in the product gas.

However, at high ER values above 0.4, the combustion reactions are favoured, which are responsible for the production of  $H_2O$  and  $CO_2$ . An increase in the percentage of plastic waste in the feedstock mixture increases the  $H_2$  composition in the product gas and reduces the  $CO$  composition in the product gas. The reason for this is that plastic waste consists of high volatile matter, which enhances the  $H_2$  composition in the syngas but does not consist of oxygen content. Figure 3.1 (a) shows that a  $H_2$  composition of 49% is obtained at ER equal to 0.18 and SFR equal to 0.6, from the blend ratio of (50% PE + 50% Biomass). From the blend ratio of (25% PE + 75% Biomass) a  $H_2$  composition of 46% is attained at ER equal to 0.1 and SFR equal to 0.6. The attainment of a higher  $H_2$  composition in the product gas from the gasification of blended feedstocks indicates that there is a synergistic interaction between the biomass and plastic waste feedstocks. The highest  $CO$  composition is attained from the single feedstock (100% Biomass), which shows that in terms of the  $CO$  composition in the product gas, the co-gasification of biomass and plastic waste does not have the synergistic effect.

### 3.2 Effect of the equivalence ratio on the $H_2/CO$ ratio and the Lower Heating Value (LHV) of the syngas.

Figure 3.2 (a) and (b) exhibits that an increase in the equivalence ratio (ER) from 0.1 – 1 at fixed SFR equal to 0.6 causes the  $H_2/CO$  ratio and the Lower Heating Value (LHV) of the syngas to decrease, for all the feedstock compositions.

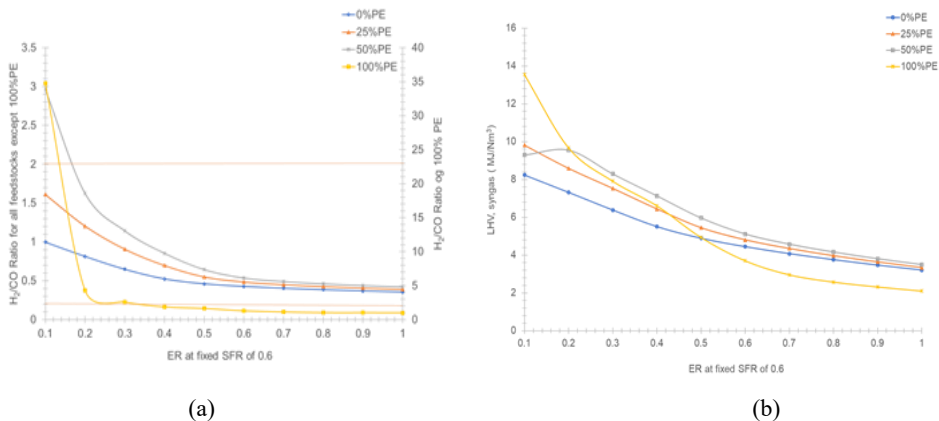


Figure 3.2 Effect of the equivalence ratio (ER) on the (a)  $H_2/CO$  ratio of the syngas and (b) Lower Heating Value (LHV) of the syngas.

This is due to the oxidation reactions that are enhanced by the increase in the amount of oxygen that enters the reactor. In Figure 3.2 (a) an increase in the  $H_2/CO$  ratio at low ER values and a fixed SFR equal to 0.6, can be attributed to the enhancement of the partial oxidation reaction, and the addition of steam which favours the endothermic reactions such as the steam-methane reforming reactions and the water gas shift reaction, which results in more  $H_2$  being produced, and this increases the  $H_2/CO$  ratio of the syngas. Moreover, the enhanced  $H_2/CO$  ratio of the syngas, is also caused by the addition of plastic waste, which contains high volatile matter and is rich in hydrogen. and the promotion of the water gas shift reaction, are the main factors that causes  $H_2$  to increase while  $CO$  decreases, thus leading to an enhanced  $H_2/CO$  ratio of the syngas. Figure 3.2 (b) shows that as ER increases the Lower Heating Value (LHV) of the syngas decreases. The high Lower Heating Value (LHV) of the syngas is favoured at low ER values below

0.4. This pattern can be attributed to the fact that at low ER values, the production of combustible gases is high, thus increasing the LHV of the syngas. In addition, an increase in plastic waste in the feedstock also contributes towards the increase in the LHV of the syngas, since increasing plastic waste results in a high hydrogen and methane composition in the syngas. Figure 3.2 (a) and (b) shows that an LHV of 9.6 MJ/Nm<sup>3</sup> was obtained from a feedstock with the blend ratio of (50% PE + 50% Biomass). The recommended value of 2 for the H<sub>2</sub>/CO ratio of the syngas and the high Lower Heating Value (LHV) of the syngas attained from the blended feedstock in the blend ratio of (50% PE + 50% Biomass) when compared to pure feedstocks indicates that there is a synergistic effect between biomass and plastic waste.

### 3.3 Effect of the Equivalence ratio on the GY and CGE

Figure 3.3 (a) shows that an increase in the equivalence ratio (ER) at fixed SFR of 0.6, causes the gas yield to increase. This is attributed to the linear relationship between the equivalence ratio and the gasifier temperature. An increase in ER, increases the gasifier temperature.

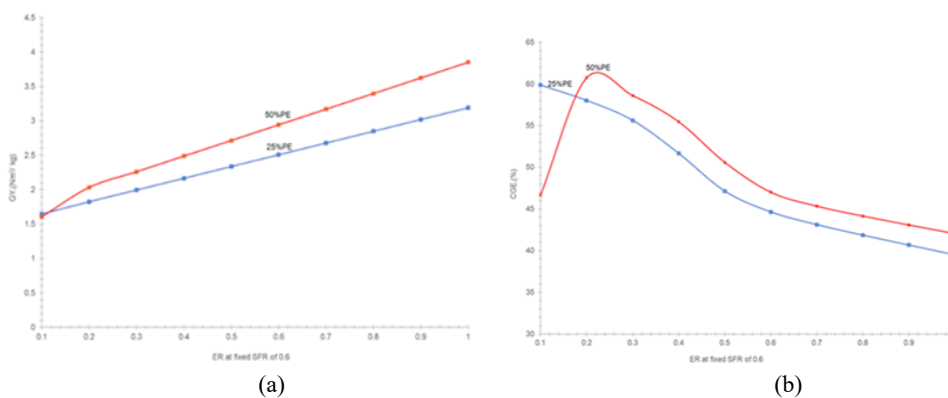


Figure 3.3 Effect of the equivalence ratio (ER) on the (a) gas yield (GY) and (b) cold gas efficiency.

The increase in the gasifier temperature increases the rate of devolatilization reactions, thus enhancing the gas production in the initial pyrolysis stage. Furthermore, the cracking and endothermic char gasification reactions are favoured at high temperatures, these reactions contribute toward the increase of the syngas yield. Figure 3.3 (b) exhibits that at low ER values, the cold gas efficiency is high, however as ER increases above 0.3, the cold gas efficiency decreases. This is because at low ER values the combustible gases such as H<sub>2</sub>, CO and CH<sub>4</sub> are high in the product gas, and that increases the LHV of the syngas, which consequently increases the cold gas efficiency. Whereas, at high ER values, there is a decrease in the combustible gases in the product gas, as these are oxidized to form CO<sub>2</sub> and H<sub>2</sub>O and thus resulting in a lower LHV of the syngas, which leads to a low cold gas efficiency. The LHV and cold gas efficiency have a linear relationship. Figure 3.3 (a) and (b) shows that increasing plastic waste fraction in the feedstock mixture increases both the gas yield and cold gas efficiency. As a result, high gas yield and high cold gas efficiency were obtained from the blend ratio of (50% PE + 50% Biomass) instead of from the blend ratio of (25% PE + 75% Biomass). This is

expected due to the nature of the plastic waste by nature which they contain high volatile matter and no oxygen content. A high cold gas efficiency of 61% is attained at ER equal to 0.24 and SFR equal to 0.6 from the blend ratio of (50% PE + 50% Biomass) (Figure 3.3 (b)).

#### 4. Conclusion

The co-gasification of biomass and plastic waste in the presence of an oxygen – steam mixture as gasifying agent has shown that the increase in polyethylene percentage in the blends, increases the H<sub>2</sub>, the H<sub>2</sub>/CO ratio of the syngas, the Lower Heating Value (LHV) (9.6 MJ/Nm<sup>3</sup>) of the syngas, the gas yield (GY) and the cold gas efficiency (GCE) (61%) of the system. This can be attributed to the low fixed carbon of polyethylene materials and that it has no oxygen. The improved outputs are determined from the blend ratio of (50% PE + 50% Biomass) instead of (25% PE + 75% Biomass) or from the biomass or plastic waste. The high yield of syngas composition, quality and energy content are favored at low ER values below 0.4, at fixed SFR equals to 0.6.

#### References

- I.I Ahmed, N. Nipattummakul, A.K. Gupta, (2011), Characteristics of Syngas from Co-gasification of Polyethylene and Wood – Chips. *Applied Energy*, 88,165 – 174.
- M. Arjorloo, M. Ghordrat, J. Scott, V. Strezov, (2022), Modelling and Statistical Analysis of Plastic Biomass Mixture Co-gasification, *Energy*, 256,124638.
- N. Al Amoodi, P. Kannan, A. Al Shoaibi, C. Shrinivasakannan, (2013), Aspen Plus Simulation of Polyethylene Gasification under Equilibrium Conditions, *Chemical Engineering Communication*, 200, 7 – 9.C. Block, A. Ephraim, E. Weiss-Hortala, D. Pham Minh, A. Nzihou, C. Vandecasteele., (2019), Co- pyrogasification of Plastic and Biomass, a Review, *Waste and Biomass Valorization*, 10, 483 – 509.
- K.G. Burra, A.K..Gupta, (2018), Synergistic Effects in Steam Gasification of Combined Biomass and Plastic Waste Mixtures, *Applied Energy*, 211, 230 – 236.
- J. M. De Andres, M. Vedrenne, M. Brambilla, E. Rodriguez, (2019), Modeling and Model Performance in Fluidized Bed Gasifier using Aspen Plus, *Journal of Air and Waste Management Association*, 6901,23 – 33.
- G. Lopez, A. Erkiaga, M. Amutio, J. Bilbao, M. Olazar, (2015), Effect of Polyethylene Co-feeding in the Steam Gasification of Biomass in a Conical Spouted Bed Reactor, *Fuel*, 393 – 401.
- F. Pinto, C. Franco, R.N. Andre, M. Miranda, I. Gulyurtlu, I. Cabrita, (2002), Co-gasification Study of Biomass mixed with Plastic Wastes. *Fuel*, 81,291 – 297.
- N. Ramzan, A. Ashraf, S. Naveed, A. Malik, (2011), Simulation of Hybrid Biomass Gasification Using Aspen Plus: A Comparative Performance Analysis for Food, Municipal Solid and Poultry Waste, *Biomass and Bioenergy* ,35, 3962 – 3969.
- M. Singh, S.A. Salaudeen, B.H. Gilroyed, A. Dutta, (2022), Simulation of Biomass – Plastic Co-gasification in a Fluidized Bed Reactor using Aspen Plus, *Fuel*, 319,123708.
- The Organization for Economic Co-operation and Development (OECD), Plastic waste, (2022), <https://www.oecd.org/newsroom/global-plastic-waste-set-to-almost-triple-by-2060.htm>.
- G. W. Huber, S. Iborra, A. Corma, (2006), Synthesis of Transportation Fuels from Biomass: Chemistry, Catalysts and Engineering, *Chemical Reviews*, 106,4044 -98.
- H. Yu., X. Yang., L. Jiang, Chen. D, (2014), Experimental Study on Co-gasification Characteristics of Biomass and Plastic Wastes, *Bioresources*, 9, 5615 – 5626.

# Food security in an oligopolistic EWF nexus system: a cooperative vs a non-cooperative case

Sarah Namany<sup>a</sup>, Maryam Haji<sup>a</sup>, Mohammad Alherbawi<sup>a</sup>, Tareq Al-Ansari<sup>a</sup>

*<sup>a</sup>College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar*

*[\\*talansari@hbku.edu.qa](mailto:*talansari@hbku.edu.qa)*

## Abstract

The ever-increasing population growth along with the diverse socio-economic pressures it creates, have led to rising demands for energy, water and food (EWF) resources. This pressure to satisfy the needs for such products has paved the way to competitive behaviors amongst the sectors' main players which might influence the efficient achievement of their objectives. Food security is one critical target that is heavily affected by the dynamics within the EWF nexus sectors. In fact, food production and supply depend primarily on the performance of water and energy systems, which involve diverse stakeholders characterised with both diverging and converging goals. Therefore, there is a need to understand and capture the competitive and collusive behaviors of each nexus sub-system as means to avoid any shortages and enhance the EWF nexus overall efficiency. The purpose of this paper is to optimise the economic and environmental performance of an EWF nexus system supplying a local food basket considering a set of renewable and non-renewable energy technologies and a range of water systems. The model is formulated as a multi-objective integer linear program that minimizes the economic costs and emissions. The model is run under two different scenarios investigating the impact of competition and cooperation on the optimal solution. Competition is modeled using Cournot game and cooperation is formulated as a Cartel, and results of the games are used as constraints restricting the optimisation framework. Results of the study assert that under competition and cooperation the water mix is mostly supplied by GW and TSE with more than 80%, while the energy production is evenly distributed amongst the three technologies. These results are highly influenced by the affordability and environmental emissions and less impacted by the behavior of the market. Findings of this study can be used to quantify the impact stakeholders' behaviors on the achievement of food security targets.

**Keywords:** EWF nexus, food security, game theory, Cournot, Cartel, optimisation

## 1. Introduction

Food security is one of the most challenging targets that many nations are striving to achieve and maintain. Food systems are governed by multiple pressures inflicted by the surrounding environment, counting climate change, political instability and market dynamics (Namany & Al-Ansari, 2021). Heavily relying on the energy and water resources to operate, the food sector's efficient functioning is also subject to the challenges faced by the water and energy sectors. Competition between water and energy industries is one of the critical issues that is often neglected when tackling food security. Competition amongst power generators and water producers can lead to fluctuations in

the quantities produced and supplied to the food system, which can disturb its operations and productivity, hence threatening the food security target. A proper modeling of competitive and cooperative behaviors is essential to reduce the impact of such volatility on the water and energy supplies to the food sector. In this regard, game theoretic techniques have been widely adopted in literature as a tool to capture competition. These models have demonstrated their efficacy in simulating competitive and collusive behaviors and in identifying the best course of action that results in mutually beneficial outcomes for the contending parties in many domains, including resources sectors (Ho et al., 2021). Considering the food sector, Saghaian and Asgari (2013) examined the non-cooperative import system of pistachios to Japan in the context of a duopolistic market made up of two significant exporting nations. The technique examined quantity competition between American and Iranian pistachio suppliers using Cournot and Stackelberg games. Addressing cooperation within the energy system, oil and gas more specifically, Gausmann (2012) proposed a model to analyze the strategic motivations for cooperation among the three major players in the European gas market, namely Qatar, Algeria, and Russia. A game-theoretic model was developed to determine each agent's interest in gas policy. Since each of these nations has distinct geopolitical goals, they will each look for a certain kind of alliance. The goal of this work is to investigate the impact of competition and cooperation in oligopolistic energy and water markets on the performance of a food system in charge of supplying a diversified food basket. Competition and cooperation between the water and energy technologies are modeled using Cournot and Cartel games and the resulting payoffs are used as constraints for an optimisation framework. The latter consists of a multi-objective integer linear program which seeks to minimize both the economic costs and environmental emissions of a configuration of renewable and non-renewable technologies. This paper is constructed such that the following section 2 presents the scenarios formulation and the mathematical representation of the optimisation model. Then, section 3 describes the main findings of the study.

## **2. Data and Methods**

The methodology developed in this work suggests a framework that assesses the influence of resource sectors' behaviors on the economic and environmental performances of an EWF nexus system delivering a food basket for the country of Qatar. Food groups considered in this study consist of perishable products, counting fruits, vegetables, poultry and meat products (Namany et al., 2019). The energy and water sectors in charge of supplying irrigation water, electricity for cultivation and fertilisers production are represented by a set of renewable and non-renewable technologies that are driven with the need of satisfying the food sector's demand for water and energy while maximising their individual profits. Two different scenarios are investigated wherein the set of energy and water technologies are either competing or cooperating to meet the food demands. Competition is modeled using Cournot game while cooperation is described by a Cartel game. Results of the games are incorporated as constraints in a multi-objective optimisation model that determines the optimal water and energy mix that would supply 40% self-sufficiency while minimising the cost and environmental impacts. The following sections describes the two different scenarios along with the optimisation model formulation.

## 2.1. Scenarios formulation

### 2.1.1. Cournot competition

Industries producing the same goods tend to compete as means to maximise their individual economic benefits, either through rising prices or increasing their outputs. In this first scenario investigated, technologies within the water and energy sectors compete between each other in order to maximise the amounts of water or power generated. The competition and the payoff of each competing technology are modeled using the Cournot game. The latter is founded on a set of assumptions entailing that players decide simultaneously and independently on the quantities to be produced and the market price of the goods is set once all players disclose their optimal quantities. In addition, Cournot is only applied when rivals are producing identical products. In this work, the water technologies counting reverse osmosis (RO), multi-stage flash (MSF), treated sewage effluent (TSE) and ground water pumping (GW) compete to produce water, and the power producers including combined-cycle gas turbine (CCGT), photovoltaics (PV) and biomass integrated gasification combined cycle (BIGCC) compete to generate power. The best response function for each water and energy technology is depicted by the following equation (1) where  $q_i$  is the individual Nash equilibrium quantity produced by each player  $i$ .  $c$  is the marginal cost of production of either water or energy, as for  $a$  and  $b$  are the parameters representing the linear inverse demand function described by equation (2), where  $P$  is the price and  $Q$  is the total quantity demanded.

$$q_i^* = \frac{a-c_i}{2b} - \frac{\sum_{i \neq j} q_j}{2} \quad (1)$$

$$P = a - b \times Q \quad (2)$$

### 2.1.2. Cartel cooperation

In the second scenario, each sector's set of technologies decide to collaborate and form cartels. Cartels occur when competitors decide to collude in order to rise prices and maintain their stability in the market. Players act in a monopoly where the optimal payoffs are distributed based on agreed-upon percentages. In this scenario, the water sector technologies and the power plants form two different cartels to maximise profits and restrict quantities. The optimal quantities produced under each scenario can be described by the following equations (3) and (4), wherein  $p_i$  is the pre-determined distribution percentage for each technology, which in this case is based on the current share of the technology in the market. As for  $j$ , it represents the index of the remaining technologies.

$$q_i = p_i \times Q \quad (3)$$

$$Q = \frac{\frac{a-c_i}{2b}}{p_i \frac{\sum p_j}{b}} \quad (4)$$

## 2.2. Optimisation model formulation

Results of the games, which are mainly payoffs of each participating technology are used as constraints for the optimisation model. The contribution of each technology to the mix is bounded by the level of the expected profit that each technology would make. If the quantity produced while participating in the mix is less than the Nash equilibrium quantity  $q_i^*$ , the technology has no incentive to be part of the mix.



The optimisation model is formulated as a multi-objective linear program aiming to minimise greenhouse gases (GHG) and the total economic costs of a configuration of water and energy technologies supplying a 40% of Qatar's food sector. The optimisation model can be described as follows:

*Objective functions:*

$$\text{Minimise: } Env = Q_w \sum_{i=1}^m x_i^w e_i^w + Q_e \sum_{i=1}^n x_i^e e_i^e \quad (5)$$

$$\text{Minimise: } Eco = Q_w \sum_{i=1}^m x_i^w c_i^w + Q_e \sum_{i=1}^n x_i^e c_i^e \quad (6)$$

*Constraints:*

$$Q_w x_i^w \geq q_i^{*w} \quad (7)$$

$$Q_e x_i^e \geq q_i^{*e} \quad (8)$$

$$\sum_{i=1}^m x_i^w = 100\% \quad (9)$$

$$\sum_{i=1}^n x_i^e = 100\% \quad (10)$$

$$x_i^w; x_i^e > 0$$

Such that:

$Env$  is the total environmental emissions in kg of CO<sub>2eq</sub>;

$Eco$  is the total economic cost in \$;

$x_i^w$  and  $x_i^{cc}$  are the decision variables representing the percentage contribution of each water and energy technology, respectively;

$Q_w$  and  $Q_e$  are the total quantities of water and energy required by the food sector in m<sup>3</sup> and kWh, respectively;

$e_i^w$  and  $e_i^e$  represent the unit carbon footprint associated the water production and energy generation in in kg of CO<sub>2eq</sub>/m<sup>3</sup> and kg of CO<sub>2eq</sub>/kWh, respectively (table 1);

$c_i^w$  and  $c_i^e$  are the unit costs associated with water production and energy generation in kg of \$/m<sup>3</sup> and kg of \$/kWh, respectively (table 1);

$m$  is the number of water technologies, counting RO, MSF, TSE and GW;

$n$  is the number of energy technologies, counting CCGT, PV and BIGCC;

**Table 1.** The optimisation model input data

Objective function	CCGT	PV	BIGCC	RO	MSF	GW	TSE
Cost(\$/kWh-\$/m <sup>3</sup> )	0.093	0.215	0.11	0.80	2.43	0.07	0.11
GWP(kg of CO <sub>2eq</sub> )	9.94E+08	9.83E+04	1.31E+08	1.71E+09	2.02E+09	1.90E+08	4.99E+08

### 3. Results and discussion

#### 3.1. Results of the games

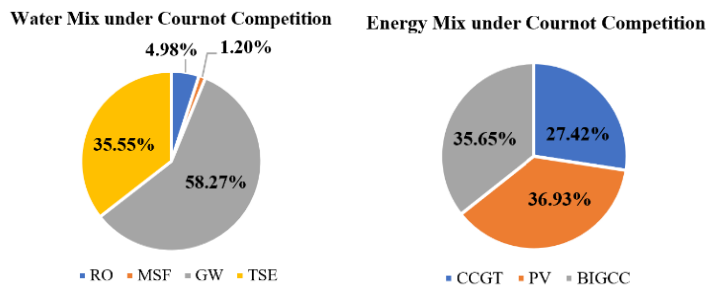
Under Cournot competition, TSE and GW are allowed to produce larger quantities compared to the desalination technologies, due to their relatively small marginal costs. As for Cartel, the largest quantities are granted to GW, RO and MSF, thanks to their important market shares. Considering the energy technologies, the same pattern is noticed for both scenarios. This means that low marginal costs and high market shares are the factors influencing the quantities produced and hence CCGT is generating the largest quantity of energy and GW is producing the highest amount of water in both scenarios (table 2).

**Table 2.** Cournot and Cartel quantities results

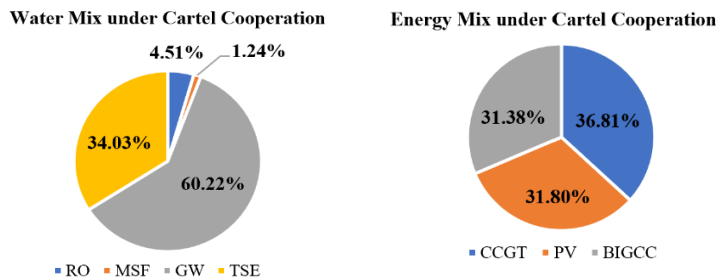
		Cournot	Cartel
<b>Water Technologies (Mm<sup>3</sup>)</b>	<b>RO</b>	42.93	50.01
	<b>MSF</b>	10.68	50.01
	<b>GW</b>	160.68	125.04
	<b>TSE</b>	154.22	25.00
<b>Energy Technologies (MW)</b>	<b>CCGT</b>	2810.00	1751.29
	<b>PV</b>	1450.00	350.25
	<b>BIGCC</b>	2610.00	1401.03

### 3.2. Results of the optimisation model

Results of the multi-objective optimisation model suggest 5 optimal solutions for both Cournot and Cartel cases. Considering Cournot, the average optimal mix costs around \$210M and generates 715,6M kg of CO<sub>2eq</sub>. The optimal water configuration under this scenario is mainly consisting of GW and TSE with a slight contribution of desalination. This can be explained by the affordability of these technologies and their relatively low environmental impact. As for the energy mix, the three power generators are contributing with almost the same share, with CCGT slightly leading thanks to its low price (Figure 1). Considering the Cartel cooperation scenario, the average optimal solution has an economic cost of \$203M and emits around 792M kg of CO<sub>2eq</sub>. The slight increase in environmental impacts and decrease of the economic costs is due to the slightly higher contribution of CCGT which is affordable, yet environmentally damaging (Figure 2). Under both scenarios the participation of the water and energy technologies is almost the same, showing that the competition has almost no impact on the technology selection in this case study. This can also be explained by the fact that the equilibrium quantities generated when playing the games are profitable for all the studied technologies under both scenarios.



**Figure 1.** The water and energy configuration under Cournot competition.



**Figure 2.** The water and energy configuration under Cartel competition.

#### 4. Conclusion

Food security is a multi-faceted target involving the contribution of multiple sectors and entailing numerous risks. Competition between water technologies and energy technologies is one of the risks that can influence the productivity of the food sector due to the uncertain supplies of resources. In this paper, competitive and cooperative behaviors of these two sectors are modeled by means of Cournot non-cooperative game and Cartel cooperative game and the generated payoffs were used to restrict the contribution of some water and energy technologies in an EWF mix as part of an optimisation framework. Results of the study assert that under competition and cooperation the water mix is mostly supplied by GW and TSE with more than 80%, while the energy production is evenly distributed amongst the three technologies. These results are highly influenced by the affordability and environmental emissions and less impacted by the behavior of the market.

#### 5. Acknowledgment

This research is supported by Hamad Bin Khalifa University (HBKU) and the Qatar National Research Fund (QNRF), members of Qatar Foundation, through the National Priorities Research Program (NPRP) grant No NPRP11S-0107-180216. The findings herein reflect the work, and are solely the responsibility, of the authors.

#### 6. References

- Gausmann, J. P. (2012). *A Game-Theoretic Approach to Cooperation in the European Gas Market. Why a Gas Exporters' Cartel Has Not Yet Materialized.*
- Ho, E., Rajagopalan, A., Skvortsov, A., Arulampalam, S., & Piraveenan, M. (2021). *Game Theory in defence applications: a review.* <https://doi.org/10.3390/S22031032>
- Namany, S., & Al-Ansari, T. (2021). Energy, water, food nexus decision-making for sustainable food security. In *Environmental Footprints and Eco-Design of Products and Processes* (pp. 191–216). [https://doi.org/10.1007/978-981-16-0239-9\\_7](https://doi.org/10.1007/978-981-16-0239-9_7)
- Namany, S., Al-Ansari, T., & Govindan, R. (2019). Optimisation of the energy, water, and food nexus for food security scenarios. *Computers & Chemical Engineering*, *129*, 106513. <https://doi.org/10.1016/J.COMPCEMENG.2019.106513>
- Saghaian, S. H., & Asgari, M. (2013). Oligopolistic market structure in the Japanese pistachio import market. *Journal of Agricultural and Food Industrial Organization*, *11*(1), 87–99. <https://doi.org/10.1515/jafio-2013-0005>

# Multiple-output Gaussian Process based Global Sensitivity Analysis for the cost-performance of electric vehicle in the United Kingdom

Min Tao<sup>a</sup>, Jude O. Ejeh<sup>a</sup>, Robert A. Milton<sup>a</sup>, Joan Cordiner<sup>a</sup>, Solomon F. Brown<sup>a\*</sup>

<sup>a</sup>*Department of Chemical & Biological Engineering, The University of Sheffield, S102TN, United Kingdom*

## Abstract

Transportation is one of the major sources of greenhouse gas (GHG) in the United Kingdom (UK), which motivates the government's policies to perform electrification of transportation and stimulate the use of electric vehicle (EV). However, many practical factors including battery costs and charging infrastructure are significantly affecting the wide utilisation of EVs. In this work, a multiple-output Gaussian processes (GPs) based Global Sensitivity Analysis (GSA) method is employed to investigate the relationship between these EV properties and the cost-performance. The variance-based GSA computations provide the detailed understandings of these key vehicle parameters for EV end-users and manufacturers. It is initially found that the total travel distances and its relevant variables available ID and fraction, and battery efficiency are the dominant factors for the cost of performance within all optimal solutions.

**Keywords:** Gaussian Process, greenhouse gas emissions, electric vehicle, global sensitivity analysis

## 1. Introduction

The impact of greenhouse gas (GHG) emissions has led to the growing commitments to net-zero emission by 2050. According to the BEIS (BEIS, 2020), Dominguez, et al, 2020), transportation is a significant source of GHG emission. The UK has therefore chosen to extensively electrify transportation and stimulate the use of electric vehicle (EV) to alleviate the GHG emissions. However, many practical factors, such as battery and ownership costs, and charging infrastructure, still inhibit the desired adoption of EVs. Thus, it is of importance to understand the relationships between the EV properties and their cost and performance. Such relationship can be obtained via Global Sensitivity Analysis (GSA). GSA calculates how much the uncertainty in outputs is attributed to the uncertainty in its individual inputs, or combinations thereof (Saltelli et al, 2008), and is a powerful tool to identify the key input variables to system outputs.

In this work, we adopt a variance-based GSA method using Gaussian Processes (GPs) (Williams & Rasmussen, 2006) to obtain the order of importance of vehicle properties to the cost and performance of EVs. The multi-output GP model is generated through the open-source ROM-COMMA software library (ROM-COMMA, 2022). GPs with automatic relevance determination (ARD) kernels allow for the semi-analytic calculation of Sobol' indices which measures the proportion of output variance attributable to each (combination of) input parameter Sobol' indices (Yearley et al., 2021). The multiple-

output GP-based GSA computations can thus provide an accurate and efficient understanding of key vehicle parameters for EV end-users and manufacturers.

## 2. Gaussian process based global sensitivity analysis

### 2.1. Gaussian process regression

In this work, we utilised GP surrogate models to build explicit relationship between the vehicle parameters  $\mathbf{x}$  and the cost-performance  $\mathbf{y} = \mathbf{f}(\mathbf{x}) + \sigma_e^2 \mathbf{I}_f$ , along with the observe data noises  $\sigma_e^2$  and the unit matrix  $\mathbf{I}_f$ . GP regression is an efficient non-parametric model technique, which also work as one statistic machine learning approach. One key assumption of GP regression is that the random output variables  $\mathbf{f}(\mathbf{x}_i)$  at any location point  $\mathbf{x}_i$  follows the normal distribution with mean value  $\mu(\mathbf{x}_i)$  and variance  $\sigma^2(\mathbf{x}_i)$ , and then prior joint Gaussian distribution  $\mathbf{f}(\mathbf{X})$  for the finite random variables  $\mathbf{X}$ :

$$\begin{aligned} \mathbf{f}(\mathbf{X}) &\sim N(\mu(\mathbf{X}), \sigma^2(\mathbf{X})) \\ \mathbf{y}|\mathbf{f} &\sim N(\mathbf{f}, \sigma_e^2) \end{aligned} \quad (1)$$

where  $\mathbf{f}(\mathbf{X}) = (\mathbf{f}(\mathbf{x}_1), \mathbf{f}(\mathbf{x}_2) \dots \mathbf{f}(\mathbf{x}_n))$ ,  $\mu(\mathbf{X}) = (\mu(\mathbf{x}_1), \mu(\mathbf{x}_2) \dots \mu(\mathbf{x}_n))$  and  $n$  is the number of samples. which are typically replaced by the zero mean function, and the covariance matrix  $\sigma^2(\mathbf{X})$  are expressed as the kernel matrix  $\mathbf{K}(\mathbf{X}, \mathbf{X})$  along with each element  $\mathbf{K}_{i,j} = \mathbf{k}(\mathbf{x}_i, \mathbf{x}_j)$ .

Here, the common radial basis function kernel was used since the automatic relevance determination kernel could provide semi-analytical computations for further variance-based GSA. The GP kernel function formulation indicates another assumption that when the inputs  $\mathbf{x}_i, \mathbf{x}_j$  are close, the corresponding output functions  $\mathbf{f}(\mathbf{x}_i)$  and  $\mathbf{f}(\mathbf{x}_j)$  are close:

$$\mathbf{k}(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 \exp\left(-\frac{1}{2l^2}(\mathbf{x}_i - \mathbf{x}_j)^T(\mathbf{x}_i - \mathbf{x}_j)\right) \quad (2)$$

where the kernel hyperparameter  $\sigma_f^2$  is global scale parameter while  $l = (l_1, l_2 \dots l_d)$  represents the bandwidth parameters,  $d$  is the dimension of inputs  $\mathbf{x}$ .

If the prior samples  $(\mathbf{X}, \mathbf{y})$  are collected, the predictive variable  $\mathbf{f}(\mathbf{X}')$  at the new location  $\mathbf{X}'$  and  $\mathbf{y}$  would satisfy a joint Gaussian distribution:

$$\begin{bmatrix} \mathbf{f}(\mathbf{X}') \\ \mathbf{f}(\mathbf{X}) \end{bmatrix} \sim N\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{K}(\mathbf{X}', \mathbf{X}') & \mathbf{K}(\mathbf{X}', \mathbf{X}) \\ \mathbf{K}(\mathbf{X}, \mathbf{X}') & \mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_e^2 \mathbf{I}_f \end{bmatrix}\right) \quad (3)$$

Then posterior predictive distribution  $\mathbf{f}(\mathbf{X}')|(\mathbf{X}', \mathbf{X}, \mathbf{y})$ :

$$\begin{aligned} \mathbf{f}(\mathbf{X}')|(\mathbf{X}', \mathbf{X}, \mathbf{y}) &\sim N(\mathbf{K}(\mathbf{X}', \mathbf{X})(\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_e^2 \mathbf{I}_f)^{-1} \mathbf{y}, \mathbf{K}(\mathbf{X}', \mathbf{X}') - \\ &\quad \mathbf{K}(\mathbf{X}', \mathbf{X}')(\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_e^2 \mathbf{I}_f)^{-1} \mathbf{K}(\mathbf{X}, \mathbf{X}')) \end{aligned} \quad (4)$$

The hyperparameters  $\sigma_f^2, l$  and  $\sigma_e^2$  can be obtained through maximizing the marginal likelihood on the collected data  $(\mathbf{X}, \mathbf{y})$ .

2.2. Variance-based global sensitivity analysis

With the built explicit GP surrogate model, the variance-based GSA method, also referred to as the Sobol' indices, is then performed on the mean predictive function  $\boldsymbol{\mu}(\mathbf{X}') = \mathbf{K}(\mathbf{X}', \mathbf{X})(\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_e^2 \mathbf{I}_f)^{-1} \mathbf{y}$ . From the probabilistic perspective, the inputs-outputs  $(\mathbf{X}', \boldsymbol{\mu}(\mathbf{X}'))$  of GP mean function are considered as the random variables. The key idea of Sobol indices method is to decompose the variance of the random outputs  $\boldsymbol{\mu}(\mathbf{X}')$ :

$$\text{Var}(\boldsymbol{\mu}) = \sum_{i=1}^d V_i + \sum_{i < j}^d V_{ij} + \dots + V_{12\dots d} \tag{5}$$

where  $V_i = \text{Var}_{\mathbf{X}'_i}(E(Y|\mathbf{X}'_i))$ ,  $V_{ij} = \text{Var}_{\mathbf{X}'_i, \mathbf{X}'_j}(E(Y|\mathbf{X}'_i, \mathbf{X}'_j)) - V_i - V_j$  and so on.

Then the important first-order index  $S_i$ , quantifying the effect of the single input  $\mathbf{X}'_i$ , can be computed as follows:

$$S_i = \frac{V_i}{\text{var}(\boldsymbol{\mu})} \tag{6}$$

Then the cumulative first-order index  $S^j$  quantifying the sum effects of the single inputs  $(\mathbf{X}'_1, \mathbf{X}'_2, \dots, \mathbf{X}'_j)$  on the total variance, can be calculated

$$S^j = \sum_{i=1}^j S_i \tag{7}$$

where  $j$  ( $j = 1, 2, \dots, d$ ) is the ranking position in the input index order  $\alpha = \{1, 2, 3, \dots, d\}$ .

Meanwhile, we computed the closed index  $S_{\ell_j}^C$  to represent the single and interaction effects of the inputs  $(\mathbf{X}'_1, \mathbf{X}'_2, \dots, \mathbf{X}'_j)$ , following the equation:

$$S_{\ell_j}^C = \sum_{\ell_i} S_{\ell_i}, \ell_i \subseteq \ell_j \text{ \& } \ell_i \not\subseteq \emptyset \tag{8}$$

where the  $\ell_j = \{1, 2, \dots, j\} \subseteq \alpha$  is the index set of inputs  $(\mathbf{X}'_1, \mathbf{X}'_2, \dots, \mathbf{X}'_j)$ .

Thus the interaction effects  $S_{\sim j}$  among the first  $j$  inputs  $(\mathbf{X}'_1, \mathbf{X}'_2, \dots, \mathbf{X}'_j)$  could be obtained as

$$S_{\sim j} = S_{\ell_j}^C - S^j \tag{8}$$

In this paper, we focus on the effects of the key parameters on the outputs. Through comparing the cumulative first-order indices  $S^j$  and the closed indices  $S_{\ell_j}^C$  within the order  $\alpha$  by decreasing the first-order index in a GSA test computation, the important input group is identified.

Here, all the Gaussian Process based Global Sensitivity Analysis computations were implemented in our open-sources ROM-COMMA software library (ROM-COMMA, 2022).

### 3. Computational results

#### 3.1. Process introduction

The objective of this work is to seek the key EV parameters for the cost and CO<sub>2</sub> emission. Here, firstly the important EV properties are considered including EV pool, 2WP (two-way power), acceleration, battery capacity, battery efficiency, battery power, range of photovoltaic (PV) output profiles, vehicle capital cost, vehicle maintenance cost, total travel distance, availability fraction, peak PV and PV cost. While the major attributions for the cost and CO<sub>2</sub> emission contain CO<sub>2</sub> emission, electricity import cost, electricity export revenue, total power imported, total power exported, total power of battery charging, total cost and total power of battery discharging. The multiple inputs-outputs relationship will be built using GP surrogate and further be analysed through the models.

#### 3.2. Data preprocessing

Through a bio-objective optimisation procedure for the cost-performance under multiple scenarios, around 2000 data pairs are generated for the multiple cost-performance under the different EV properties. By filtering the dependent inputs-outputs, the inputs of acceleration, battery capacity, battery efficiency, battery power, range of PV output profiles, vehicle capital cost, vehicle maintenance cost, total travel distance) and the outputs of CO<sub>2</sub> emission, electricity import cost, total power imported, total power of battery charging, total power of battery charging, total cost under two different optimisation strategies are chosen to construct the GP models.

#### 3.3. Computational results: GP models

For the GP model training process, a two-fold cross-validation was implemented to ensure the generalization ability of the obtained GP models. Table 1 displays the results of prediction validation. all validation criteria (the roots of mean square errors, R<sup>2</sup> value and outlier) shows that the errors of the surrogate of the selected CO<sub>2</sub> emission, total cost, the total power of battery charging and the total power imported are small enough. All the R<sup>2</sup> values are larger than 0.8, the RMSE errors are lower while the percentage of outliers are less 10%. The computational results verify the accuracy of the generated GP models. Meanwhile, Figure 1 compared the prediction of GP models and the actual observes. In generally, almost all the prediction points from the selected four parameters are close to the original observes, which indicates the efficiency of the generated GP models. Furthermore, small deviations are found for the predicted CO<sub>2</sub> emission and total imported power. That may be caused by the noises inside of the observed data.

Table 1: Prediction validation of the generated GP models for the selected output attributes

Output attributes	RMSE	R <sup>2</sup>	Outlier (%)
CO2 emission (g)	0.41	0.830	0.564
Total cost (GBP)	0.037	0.999	0.410
Total Power, Battery charging (kW)	0.061	0.996	6.410
Total Power imported (kW)	0.36	0.873	0.615

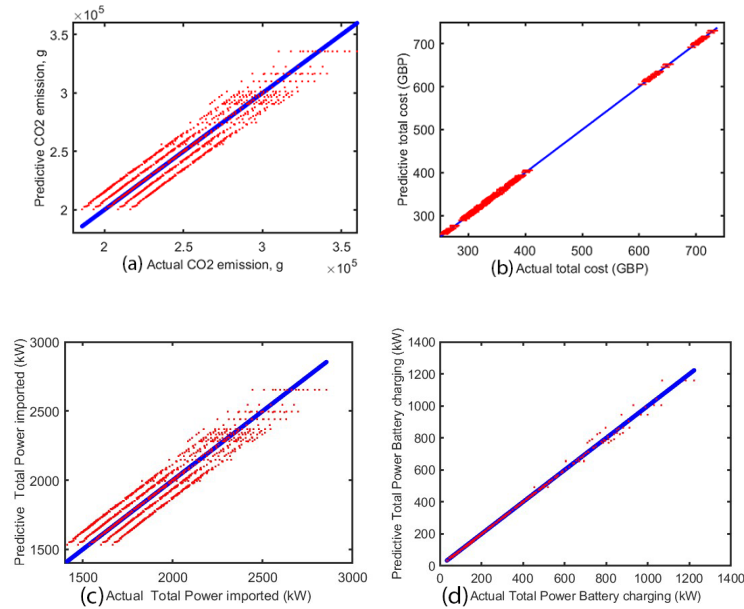


Figure 1: Tests of the GPs prediction (a) CO2 emission, (b) The total cost (c) The total power imported and (d) The total power battery charging

### 3.4. Computational results: GSA computations

With the validated GP models, GSA computations were then performed. Explicit ARD kernels contributes to the semi-analytical computations of GSA calculations, reducing the computational costs. Figure 2 summarises the results of cumulative first-order and closed indices for the four selected parameters (CO2 emission, total cost, total power imported and total power battery charging) on the eight inputs (acceleration, battery capacity, battery efficiency, battery power, range, vehicle capital cost, vehicle maintenance cost and total travel distance). It can be found, the cumulative closed indices of the first two terms all are larger than 94%, very close to 1 for all the selected four outputs. That indicates the first two input parameters total travel distance and battery efficiency are the dominant factors for the outputs. Furthermore, the small difference between the cumulative first-order and closed indices implies weak interactions effects among the input parameters. Thus, the total travel distance and its dependent variables available ID and fraction, together with another input battery efficiency are the key parameters for the cost and CO2 emission with the optimal solutions.



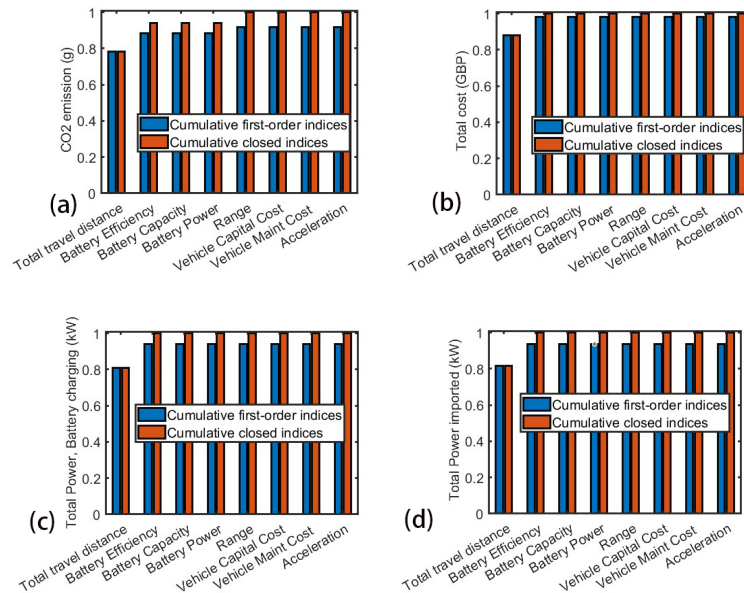


Figure 2: Computational results of the Sobol' indices (a) CO<sub>2</sub> emission, (b) The total cost (c) The total power imported and (d) The total power battery charging

#### 4. Conclusions

In this work, the multiple-output Gaussian Process based global sensitivity analysis computations are implemented between the EV properties and the cost-performance using different optimal solutions. The prediction validations indicate the high accuracy of the generated GP models while the GSA computations identify the dominant factors. It was initially found that the total travel distance and its highly relevant available ID and fraction, and battery efficiency are the key parameters for the cost and CO<sub>2</sub> emission within the optimal solutions. Further conclusions will be reached through more and general data for detailed analysis.

#### References

- A. S. Yeardley, S. Bellinghausen, R. A. Milton, J. D. Litster, S. F. Brown, 2021. Efficient global sensitivity-based model calibration of a high-shear wet granulation process. *Chemical Engineering Science* 238, 116569.
- A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, ... & S. Tarantola, 2008. *Global sensitivity analysis: the primer*. John Wiley & Sons.
- BEIS, 2020, *Electricity Generation Costs*.
- C. K. Williams, C. E. Rasmussen, 2006, *Gaussian processes for machine learning*.
- D. A. Dominguez, A. Dunbar, S. Brown, 2020, The electricity demand of an EV providing power via vehicle-to-home and its potential impact on the grid with different electricity price tariffs, *Energy reports*, 6, pp.132-141
- S. Razavi, A. Jakeman, A. Saltelli, C. Prieur, B. Iooss, E. Borgonovo, E. Plischke, S.L. Piano, T. Iwanaga, W. Becker, S. Tarantola, 2021, The future of sensitivity analysis: an essential discipline for systems modeling and policy support. *Environmental Modelling & Software*, 137, p.104954
- ROM-COMMA, 2022. ROM-COMMA. URL <https://github.com/C-O-M-M-A/rom-commma>

# Extended Multiple-Curve Resolution framework for the calibration of first-principles models

Daniel Casas-Orozco<sup>a</sup>, Jaron Mackey<sup>a</sup>, Ilke Akturk<sup>a</sup>, Gintaras V. Reklaitis<sup>a</sup>, Zoltan K. Nagy<sup>a\*</sup>

<sup>a</sup>*Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN, 47907, USA*

\* *Corresponding author: [zknagy@purdue.edu](mailto:zknagy@purdue.edu)*

## Abstract

In this work, a parameter estimation framework is presented for applications where both spectral and non-spectral information are used for model parameter estimation purposes. We illustrate our methodology by studying a crystallization system, where the liquid phase is monitored by spectroscopic means, while the solid phase is characterized by particle sizing techniques. The obtained results show that our framework is capable of handling experimental data of different types, and that crystallization model parameters and associated uncertainty regions can be successfully estimated. We also found a strong influence of data uncertainty on the accuracy of the estimated parameters for secondary nucleation.

**Keywords:** Multiple-Curve Resolution, parameter estimation, crystallization kinetics, PharmaPy

## 1. Introduction

Construction of digital twins in the pharmaceutical industry is a current trend motivated by the flexibility offered by in-silico tools, allowing exploration of process conditions and configurations early in process design/planning, and by the decision support digital twins can provide for process operation in the form of process control and monitoring. In both cases, mathematical models describing the unit operations involved typically include parameters which must be estimated by comparing model predictions to experimental data collected at carefully selected process conditions.

Some of the experimental data arise in the form of spectral information (UV-Vis, IR, Raman) gathered in real time for monitoring or parameter estimation purposes. In particular, spectra from the manufacturing of active pharmaceutical ingredients (APIs) is abundant in operations such as reaction and crystallization, and also in solid processing and product formulation. The availability of such data calls for frameworks that fully utilize the rich information provided by spectroscopic data in the parameterization of first-principle models. A layer of complexity in systems characterized by spectral data lies in need for the joint treatment of spectral and non-spectral data. For instance, building first-principles crystallization models usually requires information regarding both liquid and solid phases, the former often measured by spectroscopic techniques, and the latter by particle sizing instruments. In these cases, state and sensitivity information has to be

utilized in such a way that general-purpose parameter estimation frameworks are still usable. Although selecting fixed wavelengths as traditionally practiced is also an option, fully utilizing spectroscopic information offers a more robust and direct way of calibrating the models of interest [1], and also facilitates the analysis of multicomponent systems, of special interest in crystallization for impurity analysis.

In this work, we present a framework that uses spectral and non-spectral information from a crystallization system, to estimate kinetic parameter and their uncertainty regions. For this purpose, we expand a multiple-curve resolution/alternating least-squares (MCR-ALS) methodology [1] to allow the use of both liquid-phase and solid-phase batch cooling crystallization data within a parameter estimation framework built into the PharmaPy numerical library [2]. A case study is presented demonstrating the potential of the framework for handling different data sources and types in modeling of pharmaceutical systems. Moreover, the model parametric uncertainty estimated with this framework can be further utilized in probabilistic design space determination and robust optimization.

## 2. Mathematical framework

### 2.1 Crystallization modeling

The crystallization system is modeled by using the standard method of moments:

$$\frac{d\mu_0}{dt} = B \cdot V, \quad \frac{d\mu_n}{dt} = n \cdot \mu_{n-1} G + B L_0^n, \quad n \in \{1, 2, 3\}, \quad (1)$$

with the liquid phase described by component and global material balances:

$$\begin{aligned} \frac{dC_j}{dt} &= -\frac{tr}{V_L} \left( \delta_{j,API} - \frac{C_j}{\rho_L} \right), \quad j \in \{1, \dots, n_{comp}\}, \\ \frac{dV_L}{dt} &= -\frac{tr}{\rho_L}. \end{aligned} \quad (2)$$

Liquid-to-solid mass transfer is given by  $tr = 3k_v \rho_s G \mu_2$ , and  $\delta_{j,API} = 1$  for  $j = API$  and 0 otherwise. Cooling profile  $T(t)$  was modeled as a linear trajectory at constant cooling rate. Primary and secondary nucleation ( $B = B_p + B_s$ ) and growth ( $G$ ) crystallization mechanisms are modeled by power law expressions of the form:

$$B_p = k_p S^p, \quad B_s = k_s S^s (k_v \mu_3), \quad G = k_g S^g, \quad (3)$$

where the crystallization driving force is relative supersaturation  $S = (C_{API} - C_{sol}) / C_{sol}$ , with solubility  $C_{sol}(T, x_{solvent})$ . With the included mechanisms, the model parameter set is  $\boldsymbol{\theta} = \{k_p, p, k_s, s, k_g, g\}$ . An enlarged ODE system was constructed by including the sensitivity system  $\mathbf{S} \in \mathbb{R}^{n_y \times n_\theta}$  of the crystallization model:

$$\frac{d\mathbf{S}}{dt} = \frac{\partial \mathbf{f}}{\partial \mathbf{y}} \mathbf{S} + \frac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}}, \quad \mathbf{S} \equiv \frac{\partial \mathbf{y}}{\partial \boldsymbol{\theta}}, \quad \mathbf{y} = \{\mu_1, \dots, \mu_n, C_1, \dots, C_n, V_L\}, \quad (4)$$

where the vector function  $\mathbf{f}$  in Eq. (4) represents the right hand sides of Eqs. (1) and (2).

Analytical forms for matrices  $\partial \mathbf{f} / \partial \mathbf{y} \in \mathbb{R}^{n_y \times n_y}$  and  $\partial \mathbf{f} / \partial \mathbf{y} \in \mathbb{R}^{n_y \times n_\theta}$  were adapted from [3] and incorporated into PharmaPy.

### 2.2 Parameter estimation framework

The parameter estimation problem was formulated as

$$\min_{\boldsymbol{\theta}} \frac{1}{2} \sum_{e \in E} \sum_{\ell \in L_e} (Y - Y_{data})_{\ell}^T W^{-1} (Y - Y_{data})_{\ell}, \quad (5)$$

$$\text{subject to } Y = \mathbf{h}(\mathbf{y}). \quad (6)$$

with  $\mathbf{y}$  are the states of the dynamic mode (see Eq. (4)). Summations in Eq. (5) are indexed for a set of experiments  $e \in E$ , each experiment measured at  $L_e$  sampling times.  $Y \in \mathbb{R}^{L_e \times n_y}$  collects the monitored states computed from the dynamic model,  $Y_{data} \in \mathbb{R}^{L_e \times n_y}$  collects the corresponding experimental data, and  $W \in \mathbb{R}^{n_y \times n_y}$  is a diagonal matrix of weighting factors. Since our methodology includes spectral (liquid phase) and non-spectral (solid phase) information, states are arranged as  $Y = [Y_{sp}, Y_{nsp}]$ . Following Eq. (6), simulated spectra [1] and (number-based) mean crystal size are given by:

$$Y_{sp} = \mathbf{C}(\mathbf{C}^+ A), \quad Y_{nsp} = \boldsymbol{\mu}_1 / \boldsymbol{\mu}_0, \quad (7)$$

where  $A \in \mathbb{R}^{L_e \times n_{sp}}$  is a matrix of liquid phase absorbance measured at  $n_{sp}$  wavelength values, and  $\mathbf{C}^+$  is the Moore-Penrose inverse of the concentration matrix  $\mathbf{C} = [\mathbf{C}_1, \dots, \mathbf{C}_{n_{comp}}]$ ,  $\mathbf{C} \in \mathbb{R}^{L_e \times n_{comp}}$ ,  $\boldsymbol{\mu}_0, \boldsymbol{\mu}_1 \in \mathbb{R}^{L_e}$ .

Derivative information required to solve Eq. (5) was also decomposed into spectral and non-spectral contributions (Eq. (9)), with related dynamic sensitivities  $\partial \mathbf{C} / \partial \theta_i(t)$  and  $\mathbf{S}_{\mu_0}(t) = \partial \boldsymbol{\mu}_0 / \partial \boldsymbol{\theta}$ ,  $\mathbf{S}_{\mu_1}(t) = \partial \boldsymbol{\mu}_1 / \partial \boldsymbol{\theta}$  retrieved after solving Eq. (4).

$$J = \begin{bmatrix} J_{sp} \\ J_{nsp} \end{bmatrix}, \quad J_{sp} = [J_{sp, \theta_1}, J_{sp, \theta_2}, \dots, J_{sp, \theta_{n_\theta}}]. \quad (8)$$

Each column vector  $J_{sp, \theta_i} \in \mathbb{R}^{L_e \cdot n_{sp}}$  in  $J_{sp}$  is obtained by vertically stacking the columns of the corresponding matrix given by the variable projection shown in Eq. (10) [4]:

$$J_{sp, \theta_i} = \left[ \left( P_C^\perp \frac{\partial \mathbf{C}}{\partial \theta_i} \mathbf{C}^+ \right) + \left( P_C^\perp \frac{\partial \mathbf{C}}{\partial \theta_i} \mathbf{C}^+ \right)^T \right] \cdot A, \quad \forall \theta_i \in \boldsymbol{\theta}. \quad (9)$$

where  $P_C^\perp = (\mathbf{I} - \mathbf{C} \cdot \mathbf{C}^+)$  is an orthogonal projector of the column space of  $\mathbf{C}$  [4]. On the other hand, the non-spectral contribution  $J_{nsp} \in \mathbb{R}^{L_e \times n_\theta}$  is given by:

$$J_{nsp} = \frac{\mathbf{S}_{\mu_1} \circ \boldsymbol{\mu}_0 - \mathbf{S}_{\mu_0} \circ \boldsymbol{\mu}_1}{\boldsymbol{\mu}_0^2} \quad (10)$$

where the  $\circ$  operator represents element-wise column multiplication.

### 2.3 Materials and methods

Table 1. Experimental design

Exp	%seed	Cooling rate (K/s)
1	0	-0.1
2	0	-0.3
3	5	-0.1
4	5	-0.3

Four in-silico API concentration and mean size ( $\mu_1/\mu_0$ ) datasets were generated by simulating the system described by Eqs. (1) and (2) at the experimental conditions shown in Table 1, and using the nominal parameter values shown in Table 2. Time was sampled every 5 seconds until a final temperature  $T_f = 278$  K was reached for all the simulations. Normal noise  $\mathcal{N}(0, 3)$  was added to the resulting concentration and mean size time profiles. Using the simulated concentration  $\mathbf{C}$ , in-silico spectra was calculated by the Beer-Lambert law as  $Y_{sp}^{data} = \mathbf{C} \cdot \mathbf{a}$ , where  $\mathbf{a}$  is a vector of pure component

molar absorptivity, modeled as the summation of gaussian peaks centered at the wavenumber where maximum absorbance for the three-component system (API + THF + heptane) occurs. Only initial and final mean size data were passed to the parameter estimation framework, as in practice these are the typical measurement gathered in crystallization experiments.

Initial temperature was set as the saturation temperature at a given solvent composition. Values of initial API concentration, and molar fraction of the THF + heptane solvent mixture were fixed at  $C_{API,init} = 167$  kg m<sup>-3</sup> (0.7 mol L<sup>-1</sup>) and  $x_{THF} = 0.4$  (30 %v/v), respectively according to previous works [5]. Furthermore, initial seed mass was calculated from the theoretical crystal yield using %seed in Table 1 as  $m_{seed} = (C_{API} - C_{solub}(T_f))V$  (%seed / 100).

For experiments 3 and 4 (seeded experiments), initial crystal size distribution was assumed lognormal,  $f_{init} \sim \ln \mathcal{N}(\mu = 35, \sigma^2 = 150)$ , resulting in  $\mu_1 / \mu_0 = 25 \mu\text{m}$  ( $t = 0$ ), values in the reported ranges for continuous lomustine crystallization [5].

The parameter estimation problem, Eq. (5) was solved using an in-house implementation of the Levenberg-Marquardt algorithm in PharmaPy. A weighing matrix  $W = \text{diag}([W_{sp}^T, 1])$  where  $W_{sp} \in \mathbb{R}^{n_{sp}=113}$  is a vector of 0.1's was used to account for the difference in the scales of the spectra versus crystal size, and was tuned by testing the optimizer performance at different candidate values.

## 2. Results

In order to decrease the chance of terminating at local minima, Eq. (5) was evaluated at 2<sup>6</sup> parameter realizations according to a Sobol sampling scheme, within the parameter bounds shown in Table 2. Then, the parameter realization that lead to the lowest objective function value was used as an initial estimate to perform optimization. After solving the optimization problem, the converged parameter estimates and their corresponding 95% asymptotic confidence intervals were obtained, as shown in Table 2.

The simulated dynamic profiles using the converged parameters are shown in Figure 1, where a close agreement between data and model prediction is observed. It must be emphasized that our optimization framework uses spectra as the simulated state ( $Y_{sp}$  in Eq. (7)), thus API concentration is used in Figure 1 only for illustration purposes. It is also noteworthy that primary nucleation and growth parameter estimates are very close to their corresponding nominal values, with some of their nominals being included in their corresponding confidence intervals. This shows how the methodology is effective in recovering the original parameters used to generate the in-silico data.

Table 2. Parameter information and optimization results

		Lower bound	Upper bound	Nominal	Initial	Converged
Primary nucleation	$\ln(k_p)$	11.513	34.539	19.519	21.543	$19.39 \pm 0.18$
	$p$	0.500	5.000	3.000	1.270	$2.78 \pm 0.16$
Secondary nucleation	$\ln(k_s)$	11.513	34.539	17.613	13.726	$22.02 \pm 0.32$
	$s_l$	0.500	5.000	2.000	2.707	$1.39 \pm 0.25$
Growth	$\ln(k_g)$	-36.148	2.303	-0.693	-0.665	$-0.65 \pm 0.024$
	$g$	0.000	2.000	1.000	1.054	$1.05 \pm 0.026$
<i>WSSE</i>		--	--	--	12041.82	367.46

On the other hand, it can be seen that secondary nucleation parameters significantly deviate from their nominal values, even though the model closely follows the provided data using the converged parameter estimates. As seen in Figure 1 (bottom), the level of noise used for the mean size profile significantly distorts the general trend of the in-silico data, which can have a pronounced effect on the final estimates, especially when only initial and final mean size information is being used. As a way to check the effect of noise, new in-silico datasets were generated, this time setting noise to zero for mean size. After parameter estimation (started at the initial parameter values shown in Table 2) the converged secondary nucleation parameters were closer to their corresponding nominal values ( $\ln(k_s) = 19.84 \pm 1.69$ ,  $s = 1.62 \pm 1.35$ ).

These results indicate how challenging it can be to determine true secondary nucleation parameters, and also call attention to the feasibility of decoupling this mechanism from primary nucleation. In this regard, the order of magnitude of primary and secondary nucleation can lead to one of them being underrepresented for a particular application. For illustration, the smallest ratio between the nucleation mechanisms at the highest supersaturation for the in-silico data occurs in experiment 3 ( $B_p / B_s = 95$ ). This suggests that the effect of secondary nucleation on the system dynamics will be overshadowed by primary nucleation and growth, making sensitivity with respect to secondary nucleation parameters particularly small. This has been proven to be detrimental in the context of optimization, making the numerical system ill-conditioned, which in turn makes optimization problems harder to solve. This clearly indicates the need for more experimental data that allows one to better characterize secondary nucleation, for which different model-based designs of experiment frameworks exist in the literature [6].

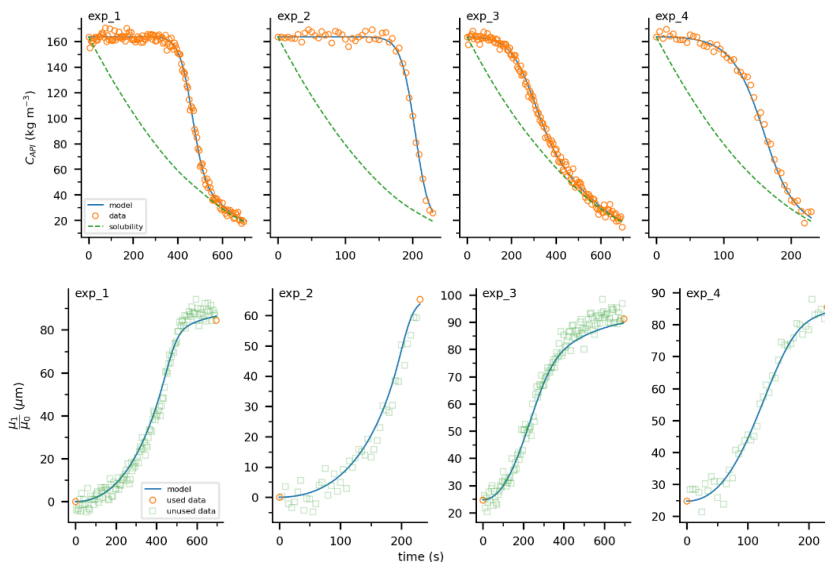


Figure 1. Predicted dynamic concentration and mean size profiles vs. corresponding in-silico data

### 3. Conclusions

In this work, we described a framework that allows the use of spectral and non-spectral information in the context of parameter estimation, and use it for the first time (to the best of our knowledge) in the context of crystallization systems. Two key elements contributed to the successful application of the framework, namely i) scaling the spectral and non-spectral contributions to the objective function and ii) properly arranging the sensitivity sources of the numerical system. This methodology can be applied to other studies in which both spectral and non-spectral information are gathered.

### 4. Acknowledgements

Funding for this work was made possible, in part, by the Food and Drug Administration through grant (U01FD006738). Views expressed herein do not necessarily reflect the official policies of the Department of Health and Human Services; nor does any mention of trade names, commercial practices, or organization imply endorsement by the United States Government.

### References

- [1] G. Puxty, M. Maeder, and K. Hungerbühler, “Tutorial on the fitting of kinetics models to multivariate spectroscopic measurements with non-linear least-squares regression,” *Chemom. Intell. Lab. Syst.*, vol. 81, no. 2, pp. 149–164, 2006.
- [2] D. Casas-Orozco et al., “PharmaPy: An object-oriented tool for the development of hybrid pharmaceutical flowsheets,” *Comput. Chem. Eng.*, vol. 153, p. 107408, Oct. 2021.
- [3] Z. K. Nagy and R. D. Braatz, “Distributional uncertainty analysis using power series and polynomial chaos expansions,” *J. Process Control*, vol. 17, no. 3, pp. 229–240, 2007.
- [4] G. Golub and V. Pereyra, “Separable nonlinear least squares: the variable projection method and its applications,” *Inverse Probl.*, vol. 19, no. 2, pp. R1–R26, Apr. 2003.
- [5] J. Mackey et al., “End-to-end reconfigurable process development for the cancer drug Lomustine,” 2021. In: AIChE Meeting, Boston MA.
- [6] S. P. Asprey and S. Macchietto, “Statistical tools for optimal dynamic model building,” *Comput. Chem. Eng.*, vol. 24, no. 2–7, pp. 1261–1267, 2000.

# Energy analysis of CO<sub>2</sub> capture by flash and distillation

Giorgia De Guido<sup>a</sup>

*<sup>a</sup>GASP – Group on Advanced Separation Processes & GAS Processing, Dipartimento di Chimica, Materiali e Ingegneria Chimica “G. Natta”, Politecnico di Milano, Piazza Leonardo da Vinci, I-20133, Milano, Italy  
[giorgia.deguido@polimi.it](mailto:giorgia.deguido@polimi.it)*

## Abstract

The aim of this work is to compare the performances of CO<sub>2</sub> cryogenic separation from a flue gas stream by means of flash and of distillation. Simulations have been carried out with the process simulator Aspen HYSYS<sup>®</sup> V11 (AspenTech, 2019), selecting the Peng-Robinson Equation of State and a sensitivity analysis has been performed on the CO<sub>2</sub> concentration in the feed gas and on the operating pressure (in the range 1.5-4 MPa) of the unit. Given the better performances of cryogenic distillation, which allows obtaining a CO<sub>2</sub> product at high purity at lower energy expenses, distillation has been considered for CO<sub>2</sub> capture from a flue gas from an oxy-fuel combustion power plant. A process scheme has been developed, properly selecting the operating conditions so to avoid the formation of solid CO<sub>2</sub>, demonstrating the better energy performances (ca. 10 % lower “Net Equivalent Methane” consumption) compared with those of a flash-based process.

**Keywords:** CCUS, cryogenic CO<sub>2</sub> capture, distillation, oxy-fuel combustion

## 1. Introduction

Carbon Capture Utilization and Storage (CCUS) is a key strategy to reach the goal of the Paris Agreement for the path to net-zero or even negative emissions in some scenarios. CCUS involves capturing CO<sub>2</sub> from a gas source and utilizing or storing it, with transportation being an essential step (Mazzocchi et al., 2013). Among the known carbon capture technologies, low-temperature/cryogenic separation processes - recently also considered for a profitable exploitation of low-quality natural gas reserves (Pellegrini et al., 2019) - are particularly suitable for gases with high CO<sub>2</sub> concentrations. Thus, it is one of the best options for treating flue gases from oxy-combustion processes or as a finishing treatment in hybrid systems (De Guido et al., 2021). Two aspects, namely high CO<sub>2</sub> recovery rates and high purity have contributed to the considerable attention cryogenic CO<sub>2</sub> capture technologies are gaining (Font-Palma et al., 2021). This is the motivation for this work, which focuses on distillation-based CO<sub>2</sub> capture not involving dry ice formation by a careful selection of the proper operating conditions. To the author’s knowledge, only few literature works on this topic deal with distillation techniques. After explaining the methods for the analysis (section 2), the advantage of distillation over flash has been demonstrated for further purification of a flue gas downstream of a bulk removal step. Then, a distillation-based process has been considered for CO<sub>2</sub> capture from a flue gas from an oxy-fuel combustion power plant for different CO<sub>2</sub> contents (66.8-85.5 mol%) in it and different recovery rates.



## 2. Methods

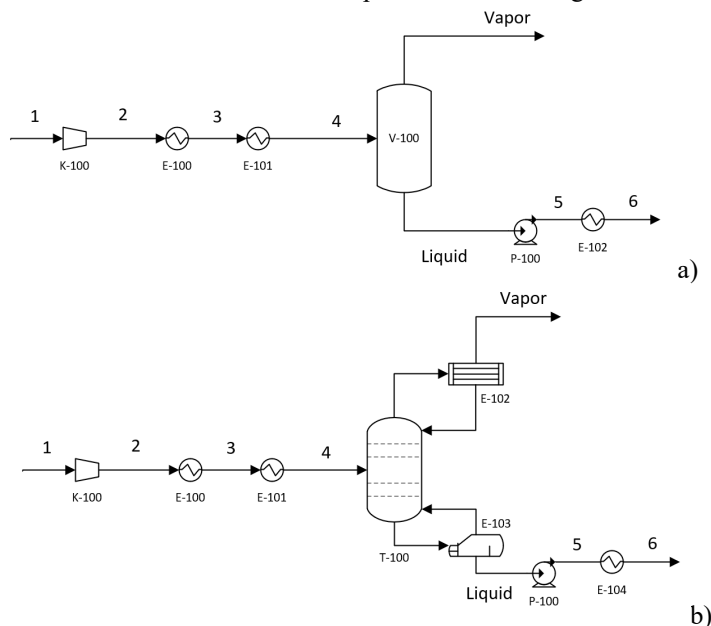
Process simulations have been carried out with Aspen HYSYS® V11 (AspenTech, 2019), selecting the Peng-Robinson Equation of State to properly represent the thermodynamic behavior of the CO<sub>2</sub>-N<sub>2</sub> system, which is representative of a flue gas mixture. In order to identify and avoid the operating conditions causing solid CO<sub>2</sub> formation, the “CO<sub>2</sub> Freeze Out” utility (De Guido et al., 2019; Spatolisano and Pellegrini, 2021a, b) was used.

The comparison among the different process schemes analyzed in this work has been carried out on the basis of an energy analysis. Since two types of energy consumption are involved in a typical cryogenic CO<sub>2</sub> capture process, both thermal and electrical, to determine the overall energy consumption the different contributions have been expressed on the same basis adopting the “Net Equivalent Methane” (NEM) method (De Guido et al., 2018). In selecting the most suitable flue gas for cryogenic CO<sub>2</sub> separation, previous findings have been taken into account. For example, van Benthum et al. (van Benthum et al., 2012) reported that a minimum concentration of 40% of CO<sub>2</sub> in the feed gas is necessary to achieve recovery higher than 70%. Therefore, according to typical flue gas compositions (Berstad et al., 2013), suitable flue gases for CO<sub>2</sub> capture at low temperature are those from oxy-fuel combustion.

## 3. Process Description and Analysis

### 3.1. Cryogenic CO<sub>2</sub> capture by flash or distillation

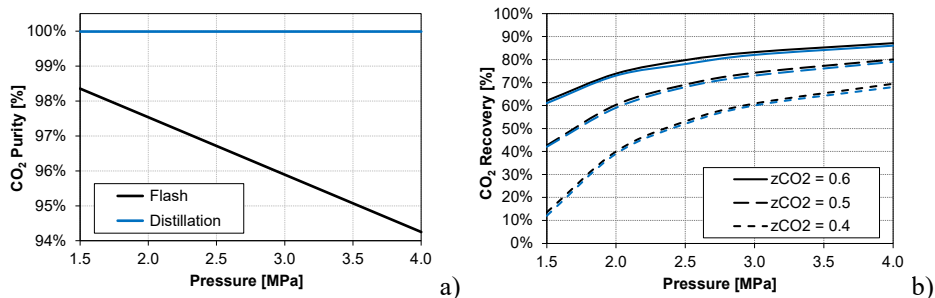
Cryogenic separation of CO<sub>2</sub> from a flue gas mixture by means of flash or distillation downstream of a bulk removal step is illustrated in Figure 1.



**Figure 1.** CO<sub>2</sub> separation from a CO<sub>2</sub>-N<sub>2</sub> mixture by: a) flash expansion; b) distillation.

To this purpose, a flue gas mixture consisting of CO<sub>2</sub> and nitrogen with a CO<sub>2</sub> content in the range 40-60 mol%, available at 0.15 MPa, has been compressed to different pressures ranging from 1.5 to 4.0 MPa. In the case of separation carried out by flash expansion, such stream has been cooled down to a suitable operating temperature selected as the lowest allowing the liquid-vapor separation without formation of solid CO<sub>2</sub>, which

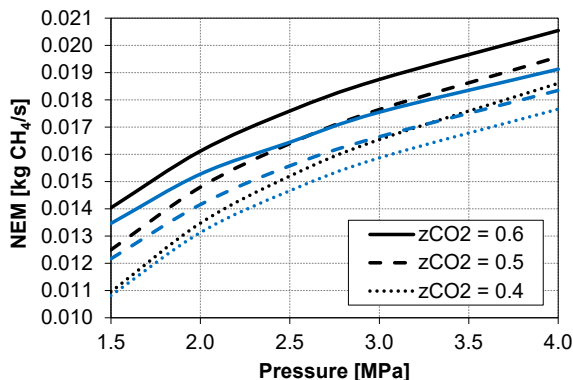
maximizes CO<sub>2</sub> recovery. Such temperature was determined using the “CO<sub>2</sub> Freeze Out” utility considering a 2 K margin from the freezing temperature. Two heat exchangers are shown in Figure 1 for cooling with cooling water (CW) first (in E-100) – not accounted for in the energy analysis – and, then, in E-101 with a suitable refrigeration cycle (De Guido et al., 2015). In the case of separation by distillation, the flue gas stream has been compressed to the operating pressure of the tower (in the range 1.5-4.0 MPa), and cooled down to the corresponding dew-point temperature. The saturated vapor stream enters the distillation column, equipped with a full reflux condenser, for which the number of trays, the liquid CO<sub>2</sub> purity (99.99 mol%) and the liquid CO<sub>2</sub> recovery (determined as the highest not forming solid CO<sub>2</sub> in any tray) have been specified. The number of trays has been determined using the short-cut method implemented in the process simulator Aspen HYSYS<sup>®</sup> V11 (AspenTech, 2019), knowing CO<sub>2</sub> product purity and recovery and assigning an actual reflux ratio 30 % higher than the minimum one. In both cases, the liquid CO<sub>2</sub>-rich product is pumped to 6 MPa and heated to 293.15 K (Hasse et al., 2013). As shown in Figure 2a, increasing the operating pressure of the flash drum, the CO<sub>2</sub> product purity is lower and it is not significantly affected by feed gas composition, since the flash operating temperature is about the same (215.86 K at 1.5 MPa and 214.26 K at 4.0 MPa) for all considered pressures. On the contrary, when separation is performed by distillation, the CO<sub>2</sub> product purity remains constant and equal to the specified value. Figure 2b shows that the CO<sub>2</sub> product recovery is higher as the operating pressure increases at constant feed gas composition, and it is lower for lower CO<sub>2</sub> content in the feed gas at constant pressure, with a little difference (within 1.7 %) between separation by flash and by distillation.



**Figure 2.** Liquid CO<sub>2</sub> purity (a) and recovery (b) as a function of the operating pressure of the flash unit / distillation column (1.5-4.0 MPa) and for different feed gas compositions. Comparison between flash (black) and distillation (blue).

Separation by distillation allows obtaining a high-purity CO<sub>2</sub> product with no penalty on energy consumptions, as shown in Figure 3. Indeed, the energy requirement in terms of NEM of the distillation column (related to flue gas compression, refrigeration both at the inlet and at the condenser of the column, and pumping of the CO<sub>2</sub> product) is lower in all the simulations, and the difference with the energy requirement of the flash (related to flue gas compression, refrigeration at the inlet of the drum, and pumping of the CO<sub>2</sub> product) is higher for higher operating pressures, which correspond to higher CO<sub>2</sub> recoveries (Figure 2b). The lower energy requirement of the distillation process could be explained by taking into account the distribution of the cooling duty in the two cases. In flash separation, the whole cooling duty is due to cooling down the feed to the flash operating temperature (ca. 215 K), while in the distillation process the cooling duty is partly due to cooling down the feed to its dew-point temperature, that is higher than flash

temperature, and only a portion of the cooling duty is due to refrigeration in the column condenser, resulting in lower equivalent methane consumption.



**Figure 3.** Net Equivalent Methane (NEM) comparison between flash (black) and distillation (blue).

### 3.2. Distillation-based CO<sub>2</sub> capture process

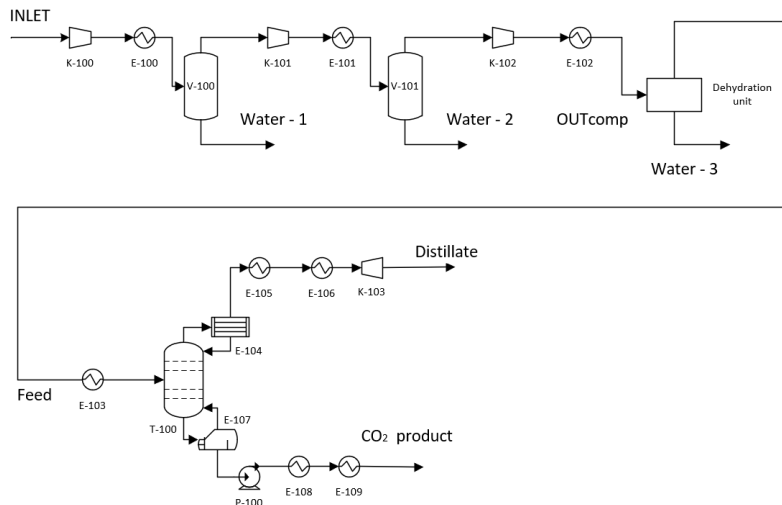
Following the results reported in the previous section, a distillation-based CO<sub>2</sub> capture process is proposed (both in a base and in an improved configuration) selecting an oxy-fuel combustion flue gas. The specific composition considered has been taken from Amann et al. (Amann et al., 2009) and is reported in Table 1.

**Table 1.** Composition (mol%) of the flue gas (Amann et al., 2009) entering the distillation-based CO<sub>2</sub> capture process.

CO <sub>2</sub>	H <sub>2</sub> O	N <sub>2</sub>	O <sub>2</sub>	Ar
66.8	3.1	18.9	5.5	5.5
74.1	3.0	11.0	5.7	6.0
85.5	3.0	0.4	5.7	5.4

The process scheme is illustrated in Figure 4 and differs from the one proposed by Amann et al. (Amann et al., 2009) for the replacement of the flash units by a distillation column for the reasons reported in the previous section. The feed gas is compressed in a three-stage process to 3.5 MPa, with intermediate cooling to 303.15 K using CW. The pressure of each stage has been determined as explained in a previous work (Pellegrini et al., 2015). After cooling in the first two stages, water is separated as liquid in two drums (V-100 and V-101), whereas the vapor stream is fed to the next compression stage. The stream exiting the compression train enters a dehydration unit (not accounted for in the energy analysis that follows, for comparison purposes with the flash-based process (Amann et al., 2009)), where the reduction of the water content to 20 ppm is accomplished. The dehydrated stream is, then, cooled to its dew-point temperature and is fed to the distillation column, equipped with a full reflux condenser, where an external refrigerant is used. The reboiler employs CW as heating medium because the heating is below ambient temperature. For all the investigated case studies (i.e., different CO<sub>2</sub> contents in the flue gas and different recoveries), the number of theoretical trays turned out to be 4 from a short-cut calculation and has been specified together with the CO<sub>2</sub> purity (99 mol%) and recovery set to the same values considered in the flash-based process (Amann et al., 2009) for comparison purposes. The distillate stream, exiting the top of the distillation column, is heated (E-105) to 293.15 K using CW, then it is heated

(E-106) to 423.15 K with steam and is expanded to atmospheric pressure in a turbine. The stream exiting the bottom of the column is pumped (P-100) to 15 MPa, heated (E-108) to 293.15 K with CW and finally heated (E-109) to 293.15 K.



**Figure 4.** Scheme of the distillation-based CO<sub>2</sub> capture process without heat integration.

The process scheme in Figure 4 (in the following, referred to as “distillation-based (A)”) can be improved from an energy point of view by using the distillate product for pre-cooling of the feed stream to the distillation column, so reducing the cooling duty supplied by an external refrigerant for bringing the feed to its dew-point temperature. Such an improved configuration is referred to as “distillation-based (B)” in the following. The results of the energy analysis are summarized in Table 2, where the relative difference has been computed with respect to the flash-based process presented in the literature (Amann et al., 2009).

**Table 2.** Results of the energy analysis (“Net Equivalent Methane” method).

CO <sub>2</sub> content [mol%]	CO <sub>2</sub> recovery [%]	NEM [kgCH <sub>4</sub> /s]			Relative difference [%]	
		flash-based (Amann et al., 2009)	distillation-based (A)	distillation-based (B)	distillation-based (A)	distillation-based (B)
66.8	75	0.6945	0.7257	0.6366	+ 4	- 8
74.1	85	0.6873	0.7138	0.6603	+ 4	- 4
85.5	75	0.5527	0.5416	0.4930	- 2	- 11
85.5	90	0.6364	0.6274	0.6025	- 1	- 5

The “distillation-based (A)” process is more energy demanding than the flash-based separation for low CO<sub>2</sub> contents in the flue gas, whereas the configuration with heat integration (“distillation-based (B)”) is less energy demanding in all cases, with the highest difference at low carbon dioxide recovery due to a decrease of column condenser and cooling (E-103) duties. An important advantage of the proposed process, compared with the flash-based one, is that a liquid CO<sub>2</sub> product ready for pumping is obtained, which is suitable for further utilization (e.g., for Enhanced Oil Recovery). Moreover, it allows specifying even a stricter specification (by carefully selecting the proper operating

conditions) on the CO<sub>2</sub> product to meet the quality requirements, in particular with respect to the residual O<sub>2</sub> content, for food-grade CO<sub>2</sub>, which will be the subject of a future work together with a techno-economic analysis.

#### 4. Conclusions

The comparison between cryogenic CO<sub>2</sub> capture by flash and distillation, not addressed in previous literature works, has shown that a CO<sub>2</sub> product at higher purity can be obtained in the latter case with lower energy consumptions, especially at higher operating pressures, although with a slightly lower (within 1.7 %) recovery. Therefore, a distillation-based process has been proposed for separating CO<sub>2</sub> out of flue gas streams typical of oxy-fuel combustion power plants. An energy comparison with the flash-based analogous process has demonstrated the better performances of the distillation-based process for several CO<sub>2</sub> concentrations (66.8-85.5 mol%) in the flue gas to be treated and different recovery rates. An important advantage of the proposed process, compared with the flash-based one, is that a high purity CO<sub>2</sub> product is obtained, which is suitable for further utilization, including in the food and beverage industry if a stricter specification, in particular with respect to the residual O<sub>2</sub> content, is assigned.

#### References

- J.-M. Amann, M. Kanniche, C. Bouallou, 2009, Natural gas combined cycle power plant modified into an O<sub>2</sub>/CO<sub>2</sub> cycle for CO<sub>2</sub> capture, *Energy Conversion and Management*, 50, 510-521.
- AspenTech, 2019, Aspen HYSYS®, Burlington (MA), United States.
- D. Berstad, R. Anantharaman, P. Nekså, 2013, Low-temperature CO<sub>2</sub> capture technologies—applications and potential, *International Journal of Refrigeration*, 36, 1403-1416.
- G. De Guido, M.R. Fogli, L.A. Pellegrini, 2018, Effect of Heavy Hydrocarbons on CO<sub>2</sub> Removal from Natural Gas by Low-Temperature Distillation, *Industrial & Engineering Chemistry Research*, 57, 7245-7256.
- G. De Guido, M. Gilardi, L.A. Pellegrini, 2021, Novel technologies for low-quality natural gas purification, *Computer Aided Chemical Engineering*, 50, 241-246.
- G. De Guido, S. Langè, L.A. Pellegrini, 2015, Refrigeration cycles in low-temperature distillation processes for the purification of natural gas, *Journal of Natural Gas Science and Engineering*, 27, 887-900.
- G. De Guido, F. Messinetti, E. Spatolisano, 2019, Cryogenic Nitrogen Rejection Schemes: Analysis of Their Tolerance to CO<sub>2</sub>, *Industrial & Engineering Chemistry Research*, 58, 17475-17488.
- D. Hasse, S. Kulkarni, E. Sanders, E. Corson, J.-P. Tranier, 2013, CO<sub>2</sub> capture by sub-ambient membrane operation, *Energy Procedia*, 37, 993-1003.
- M. Mazzocchi, G. De Guido, B. Bosio, E. Arato, L.A. Pellegrini, 2013, CO<sub>2</sub>-mixture properties for pipeline transportation in the CCS process, *Chemical Engineering Transactions*, 32, 1861-1866.
- L.A. Pellegrini, G. De Guido, V. Valentina, 2019, Energy and exergy analysis of acid gas removal processes in the LNG production chain, *Journal of Natural Gas Science and Engineering*, 61, 303-319.
- L.A. Pellegrini, S. Langè, M. Baccanelli, G. De Guido, 2015, Techno-Economic Analysis of LNG Production Using Cryogenic vs Conventional Techniques for Natural Gas Purification, *Offshore Mediterranean Conference and Exhibition*, Ravenna, Italy, 25-27 March 2015.
- E. Spatolisano, L.A. Pellegrini, 2021a, CO<sub>2</sub>-Tolerant Cryogenic Nitrogen Rejection Schemes: Analysis of Their Performances, *Industrial & Engineering Chemistry Research*, 60, 4420-4429.
- E. Spatolisano, L.A. Pellegrini, 2021b, Solid-Liquid-Vapor Equilibrium Prediction for Typical Helium-Bearing Natural Gas Mixtures, *Journal of Chemical & Engineering Data*, 66, 4122-4131.
- R. van Benthum, H. van Kemenade, J. Brouwers, M. Golombok, 2012, Condensed rotational separation of CO<sub>2</sub>, *Applied Energy*, 93, 457-465.

# Feature Embedding of Molecular Dynamics-Based Descriptors for Modeling Electrochemical Separation Processes

H. K. Gallage Dona,<sup>a, 1</sup> T. Olayiwola,<sup>b, 1</sup> L. A. Briceno-Mena,<sup>b</sup> C. G. Arges,<sup>c</sup> R. Kumar,<sup>a</sup> J. A. Romagnoli,<sup>b, \*</sup>

<sup>a</sup>*Department of Chemistry, Louisiana State University, Baton Rouge, Louisiana 70803, United States.*

<sup>b</sup>*Cain Department of Chemical Engineering, Louisiana State University, Baton Rouge, Louisiana 70803, United States.*

<sup>c</sup>*Department of Chemical Engineering, The Pennsylvania State University, University Park, PA 16802, United States.*

\*corresponding author: [jose@lsu.edu](mailto:jose@lsu.edu)

<sup>1</sup>Equal contribution

## Abstract

Machine learning is increasingly being used as a modeling technique for the analysis of complex systems even when limited knowledge is available. Here, we demonstrate the application of machine learning models to predict macroscopic properties of electrochemical polymers from molecular attributes. The computational framework proposes a novel approach to estimate molecular dynamic simulations (MD) attributes and experimental activity coefficients. Results showed that the data augmentation and embedding strategy effectively produce unique representations of each polymer. The findings from this study allow for the estimation of activity coefficients of novel polymers without the need for new time-consuming MD simulation runs. This data processing technique could then be used map to material properties that influence for electrochemical ionic separation units.

**Keywords:** Feature embedding, machine learning modeling, molecular dynamics simulations, electrochemical separations.

## 1. Introduction

It is estimated that the world's population will clock 9 billion by 2050; these fast-growing populations lead to an increase in demand for energy and water supply. For example, over 4 billion people lives in areas with severe scarcity of water (Mekonnen and Hoekstra, 2016). Also, in the energy industry, there is strong demand for greater energy supply and to reduce the carbon footprints of these processes. Both energy and water industries require the selective separation of millions of liquid components. These industries have a high demand for effective separation procedures to obtain clean water, gases, or solvents, which today make up a sizable share (10–15%) of global energy consumption (Sholl and Lively, 2016). Thus, there is a strong interest on the design and optimization of high efficiency separation processes. In ionic separation using electrodialysis and membrane

capactive deionization, the selection of the ion-exchange membrane (IEM) is very important as it can affect the energy efficiency during deionization (Palakkal et al., 2018).

Computational models that can bridge molecular descriptors to material properties and performance in electrochemical separation platforms are vital to advancing the field (Briceno-Mena et al., 2021). To understand the ionic transport between the solution and polymer, there exists a need to model the relationship defining this transport mechanism. Our primary macroscopic property of interest are the ion activity coefficients in polyelectrolytes. Activity coefficients influence the partitioning coefficient of an ion between the polymeric ion-exchange membrane and the liquid solution, and it can also affect ionic conductivity. However, accurate ion activity coefficient models based upon the physics and chemistry of the ion-exchange membranes is lacking. In fact, most activity coefficient models for ions in liquids rely upon empirical and semi-empirical approaches (e.g., Debye-Hückel or Pitzer models). These empirical and semi-empirical models often need large data sets for estimating the empirical constants in these models. Because there is a lack of data for ion activity in various ion-exchange membrane chemistries and structures, attaining effective and accurate activity coefficient models for these soft materials are elusive. Machine learning represents a unique tool to streamline model development for attaining accurate activity coefficients of polymeric ion-exchange membranes.

The Proposed ML network will bridge molecular scale attributes from molecular dynamics (MD) simulations to the ionic activity coefficients in polymeric ion-exchange membrane materials. Machine learning methods such as Support Vector Regression (SVR) and Random Forest (RFR) were used to connect the polymer fingerprint (based on chemical structure) and molecular level attributes (based on MD descriptors) to the macroscopic attribute of polymer electrolytes (i.e., activity coefficient). This work developed a ML framework that involves predicting the MD descriptors of ionomers. Ionomer molecular fingerprints, salt concentration and one-hot encoded categorical features of the salt ions were used. The main goal of this work is to predict ion activity coefficient in polymer electrolytes from the predicted MD descriptors, thereby serving as a tool for future users when there exists no MD data and experimental activity coefficient for a chosen polymer and salt ions.

## 2. Methods

More than ten original experimental reports containing ion activity coefficient values in ion-exchange membranes and thin films were collected from literature. We restricted our search to publications with detailed information about the structure of the copolymer, ion exchange capacity, water uptake and experimental ion activity coefficient. Polymer structures and salt ions represented by their SMILES representations and categorical variables, respectively were converted into numeric arrays using *RDKit* and the OneHotEncoder method in *Scikit-Learn*, respectively as shown in Figure 1.

With explicit counterions and salt ions, we simulated polyelectrolyte systems using molecular dynamics simulations. To represent polyelectrolytes, a system with 17 polymer chains, water and salt ions was prepared with PACKMOL package in conformance with experimental conditions. We adopted the OPLSAA or GAFF2 forcefield parameters and

TIP3P model for polymer and water respectively. The MD simulations were implemented in LAMMPS package. The simulations were then analyzed to compute the solvation properties of these polymer electrolyte systems including radial distribution function, coordination numbers and diffusion coefficients using final production run of 40ns long in the isothermal (fixed temperature at 300K using a Nose Hoover thermostat, constant number of particles and volume) ensemble.

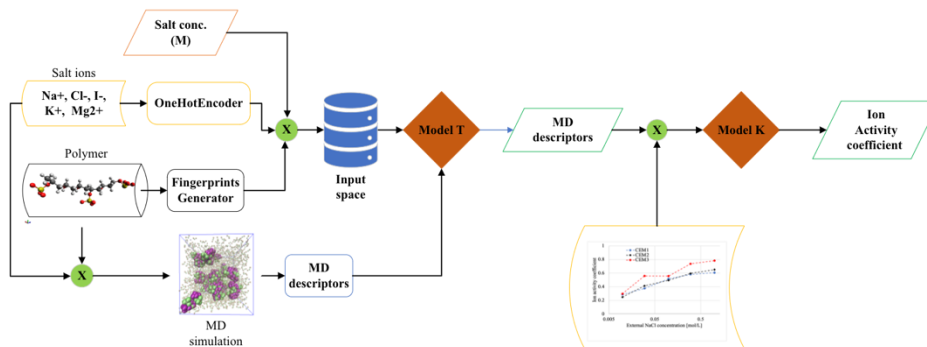


Figure 1: Workflow for modeling the prediction of MD descriptors and ion activity coefficient.

In this contribution, two regression models (Random Forest regression, RFR and Support Vector Regression, SVR) are compared to estimate the MD descriptors and ion activity coefficient. Figure 1 shows the framework for the model development. First, an ML model (*Model T*) was created to predict MD descriptors based on an input vector consisting of arrays of polymer fingerprints, one-hot encoded salt ions, salt concentration and number of water molecules per tethered ions. Secondly, another regression model (*Model K*) was created as shown in Figure 1 to predict the experimental ion activity coefficient. The input vectors consist of the MD descriptors (predictable from *Model T*) and the input matrix of model T. To train these models, the available dataset was split into training and test data. The model hyperparameters are tuned using the NSGA-II algorithm as implemented by *Pymoo*. The final model performance was quantified using coefficient of determination ( $R^2$ ), mean square error (MSE) and mean absolute error (MAE).

### 3. Results and Discussions

#### 3.1. Polymer space

Our studied dataset consists of 8 block copolymers and 3 random copolymers linked to experimental ion activity coefficients. To assess the similarity between the available copolymer, we computed the Tanimoto similarity ( $T_c$ ) which is the most popular and widely accepted similarity metrics for polymer chemistry. The results of the analysis are depicted in Figure 2. Two polymers are considered identical when the  $T_c$  equals 1 and completely different when the  $T_c$  equals 0. The similarity matrix showed that our selected polymers are very different from one another. Only the RCE and BCE forms of a polymer have high similarity values ( $>0.7$ ). This is expected because these polymer forms have



the same chemistry except with different arrangement of repeat units. Thus, a model generated from these polymers and will have a good generalization to other polymers with similar constituents because the sampled polymers exhibit great variation.

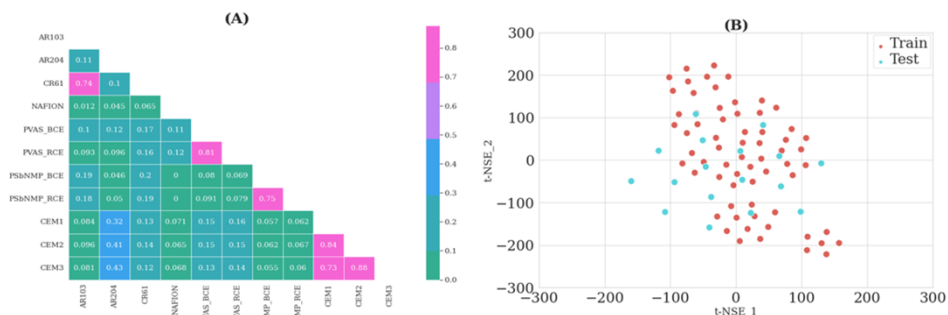


Figure 2: (a) Visualization of the similarity among the studied polymers and (b) sample space of the training and test input data used in machine learning.

During ML model development, greater attention must be given to the sample data. Specifically, by using similar train and test data, the impact of data discrepancies can be ameliorated and thus, gain a better understanding of the model's properties. To assess the conformational space of the training and test data, we computed the t-SNE plot, and the result is shown in Figure 2b. The analysis showed that the proportion of the sample space of the test data aligns strongly with that of the training set. Also, the t-SNE profile showed that the input matrix is well scattered and thus, an accurate ML model based on this input matrix will exhibit the better generalization to unseen data.

### 3.2. Prediction of MD descriptors

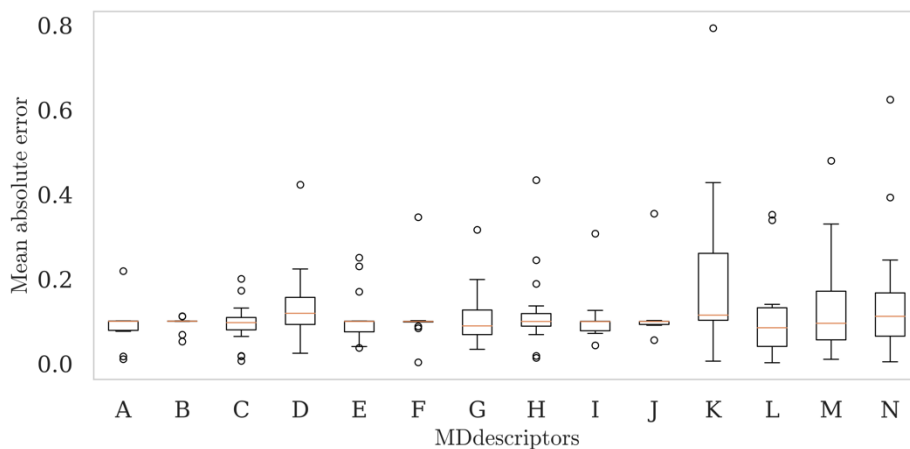


Figure 3. Boxplot showing the performance of the SVR model. Table 1 shows the description of the legends (A-N).

The SVR model was developed for a multi-output system. Here, the input features were transformed using principal component analysis to correct the effect of dimensionality reduction. Also, the target parameters (i.e., MD descriptors) were converted into a range (-1, 1). Afterwards, the transformed input and target data were imported to a SVR model and then trained based on the training set only. Figure 3 shows the statistical measure of the best model. From the perspective of the ML algorithm, the predicted MD descriptors exhibit close agreement with the target MD descriptors. Furthermore, the ML model based on the MF fingerprints shows a better predictive performance than the RDKit descriptor (not shown here); thus, a better representation of the studied polymer systems.

Table 1. The descriptions of the MD descriptions shown by the boxplot in Figure 3. (RDF refers to radial distribution function).

A	first minima of RDF between counterion and oxygen of water.	H	coordination number (at first minima) from RDF between charge bearing group (in polymer) and oxygen (in water).
B	peak position of RDF between counterion and oxygen of water	I	first minima of RDF between charge bearing group (in polymer) and counterion (in salt).
C	peak height of RDF between counterion and oxygen of water	J	peak position of RDF between charge bearing group (in polymer) and counterion (in salt).
D	coordination number (at first minima) of RDF between counterion and oxygen of water.	K	peak height of RDF between charge bearing group (in polymer) and counterion (in salt)
E	first minima of RDF between charge bearing group (in polymer) and oxygen of water.	L	coordination number (at first minima) from RDF between charge bearing group (in polymer) and counterion (in salt).
F	peak position of RDF between charge bearing group (in polymer) and oxygen of water	M	diffusion coefficient of water in polymer electrolyte per diffusion coefficient of pure water.
G	peak height of RDF between charge bearing group (in polymer) and oxygen of water	N	diffusion coefficient of counterion in polymer electrolyte per diffusion coefficient of pure water

### 3.3. Activity Coefficients Prediction

The model K was validated against existing datasets from the literature, giving confidence for generating additional ML models with both RFR and SVR. The test set is used to assess the model's performance on previously unexplored data. Prediction accuracy of Both RFR and SVR regression models were compared for predicting activity coefficients in figure 4a. Figure 4b illustrates the correlation between the model K predictions of the activity coefficients using RFR. An RFR architecture with max depth ( $D = 82.95$ ) with minimum sample split ( $W = 2$ ) was found to have the best hyperparameters for the model K implementation. With the optimized neural network, MSE values were 0.002 for fivefold cross-validation model training with RFR. This model might further be improved by expanding the scope of data set.

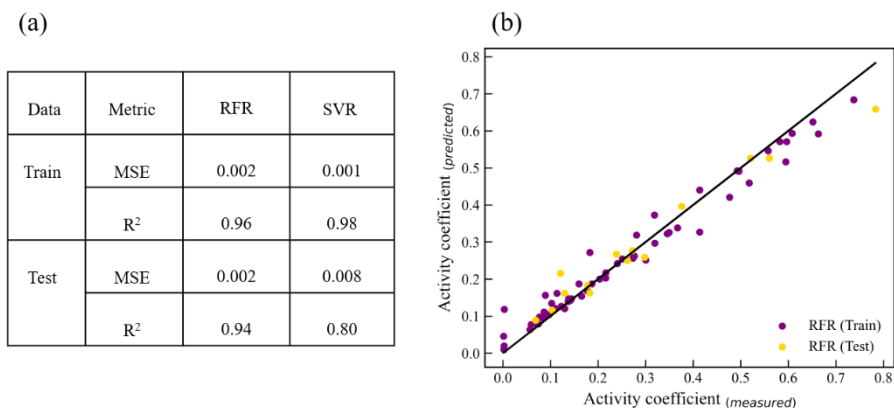


Figure 4. (a) Comparison of different model  $K$  for predictions of activity coefficients using RFR and SVR (b) Parity plot of predicted and experimental activity coefficients using the RFR.

#### 4. Conclusions

Machine learning identified physically meaningful patterns between molecular structure, molecular dynamics data and experimentally determined ionic activity coefficients in polymeric ion-exchange membranes and thin films. The outcomes of this framework can inform the rationale design of the materials that enable energy efficient and potentially selective ionic separations.

#### Acknowledgments

This material is based upon work supported by the U.S. Department of Energy, Office of Science, under the Office of Basic Energy Sciences Separation Science program under Award No. DE- SC0022304. We gratefully acknowledge the computer time allotted by the high-performance computing center at LSU and the Louisiana Optical Network Initiative.

#### References

- Briceno-Mena, L.A., Venugopalan, G., Romagnoli, J.A., Arges, C.G., 2021. Machine learning for guiding high-temperature PEM fuel cells with greater power density. *Patterns* 2, 100187.
- Mekonnen, M.M., Hoekstra, A.Y., 2016. Four billion people facing severe water scarcity. *Science Advances* 2, e1500323. <https://doi.org/10.1126/sciadv.1500323>
- Palakkal, V.M., Rubio, J.E., Lin, Y.J., Arges, C.G., 2018. Low-Resistant Ion-Exchange Membranes for Energy Efficient Membrane Capacitive Deionization. *ACS Sustainable Chem. Eng.* 6, 13778–13786. <https://doi.org/10.1021/acssuschemeng.8b01797>
- Sholl, D.S., Lively, R.P., 2016. Seven chemical separations to change the world. *Nature* 532, 435–437. <https://doi.org/10.1038/532435a>

# An NLP-based framework for extracting the catalysts involved in Hydrogen production from scientific literature

Avan Kumar,<sup>a</sup> Hariprasad Kodamana,<sup>a,b,\*</sup>

<sup>a</sup>*Department of Chemical Engineering, Indian Institute of Technology Delhi, New Delhi-110016, India*

<sup>b</sup>*Yardi School of Artificial Intelligence, Indian Institute of Technology Delhi, New Delhi-110016, India*

*corresponding author: - \*[kodamana@chemical.iitd.ac.in](mailto:kodamana@chemical.iitd.ac.in)*

## Abstract

Hydrogen production as a renewable fuel source is an active area for the industry and researchers. NLP has been growing for the last few years as the need for extracting non-structural text data is accelerating with time. Here, in this work, we focused on the peer review articles on “Hydrogen production from alcohol” as a keyword for search in an Elsevier database. The abstracts are retrieved with the help of the “Elsevier API” key. We mainly focus on the article's abstract because it contains crucial points. The collated text data set combines two broad classes, i.e., relevant abstracts and non-relevant abstracts. The relevant abstract class stands for hydrogen production from alcohol-based abstracts. Other left-over abstracts will belong to the non-relevant class. The relevant abstracts are followed by the task of “Question and Answering” (Q&A), where we have selected a question, e.g. (a) What is the catalyst material used? Transfer learning implemented on the emerging model “BERT” will be utilized for these tasks. The finetuned BERT model is renamed H<sub>2</sub>-BERT. The H<sub>2</sub>-BERT performs well, with an accuracy of 0.980 for the train set and 0.974 for the test dataset for the classification task. For the task of Q&A, the data is prepared, fed to the “BERT” model and retrained the last layer weights. This downstream task gave outstanding results with 1.000 accuracy for the train set and an accuracy of 0.823 for the test dataset for accurately predicting the answer to their question based on the given context. This model will help readers extract the utilized catalyst name from the given text and make the process less time-consuming.

**Keywords:** Hydrogen production, NLP, H<sub>2</sub>-BERT, Classification, Q&A

## 1. Introduction

Nowadays, the requirement for energy is increasing with time, and the depletion of non-renewable energy resources is also happening [1]. Hydrogen is an excellent energy carrier and an environmental-friendly energy source for the future. Recently, many researchers have focused on the production methodology search of H<sub>2</sub> to fulfil the energy demand [2]. A quick search on google scholar will give us millions of patents and research articles. Accessing all of those sources of knowledge and digging deep will be challenging and tedious. It requires expertise in this field and a considerable time to get essential information, such as the catalyst used. All articles hide some crucial parameters that may help the industries to grow. A quick recommendation system can be built for the above reference to extract all the crucial information from available articles to ease access for academicians and industry practitioners. Machine learning and Natural language process (NLP) fields are gaining high popularity to help extract non-structural text data into structural information.

The BERT is a standard model in the NLP field, which stands for "Bi-directional encoder representation of transformers developed by google in 2018. The BERT model is well-trained on Wikipedia, Books, Journals corpus, etc. The downstream tasks performed by the BERT model are Classification, Name entity recognition (NER), and Q&A. It has two classes based on the size of layers of the encoder (i) BERT<sub>large</sub> (total parameters= 340 M, encoder layers =24) and (ii) BERT<sub>base</sub> (total parameters = 110 M, encoder layers =12). The performance of BERT encourages many researchers, so many versions of BERT have been launched. The most famous version of BERT is "Sci-BERT". It adopts the architecture of BERT and is trained from scratch with the vocab of computer science (85%) and biomedical (15%) literature. Other finetuned "BERT" are BioBERT, clinical-BERT, m-BERT, and FinBERT [3].

Implementing transfer learning on the BERT model has helped learn the hidden information from a new corpus of data. In this context, Kumar et al. have reported Classification followed by the with an accuracy of 0.997(test dataset) for Classification, and NER also has a good 0.890 (test dataset) accuracy for a finetuned BERT model called "Ex-SciBERT" [4]. Huang et al. created six battery-related BERT models with downstream tasks, Classification and Q&A [5]. Qu et al. have built a ConQA model based on the BERT model; it helps to get insights into historical data [6]. Some recent projects introduced an attention-based deep learning approach for NER tasks for material science literature and disclosed a synthesis procedure for all-solid-state batteries using a 243-article literature corpus.

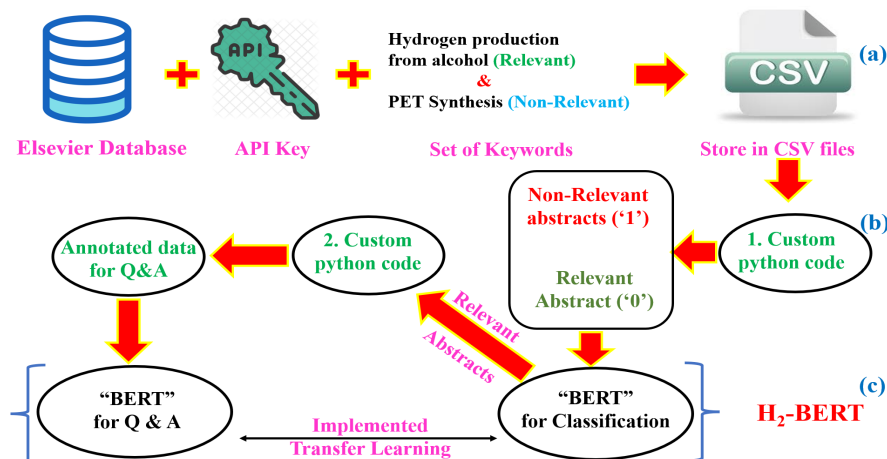


Figure: - (1) A Schematic diagram of the steps followed in this study. Where (a) signifies the steps for text extract, (b) stands for data annotation, and (c) shows transfer learning implementation on the BERT model.

In this study, we propose to develop a finetuned BERT mode for Q&A tasks related to the text that contains information on "Hydrogen production from alcohol" ". First, we will classify relevant class abstracts (means related to hydrogen production information) from the group of abstracts. The relevant abstracts will be processed, and a specific question will be asked. For clear visualization, see Figure 1, where each step of the proposed framework is showcased. The structure of the proposed study is as follows: Section 2 for data extraction and annotation, discussed the proposed model in Section 3, Section 4 for Result & Discussion and at last, a conclusive remark on the proposed study.

## 2. Data collection and annotation

We have utilized the application programming interface (API) API key to access the complete database of Elsevier. The keywords for the search are “Hydrogen production from alcohol” and a random keyword like "PET synthesis". The total number of accessible abstracts is 11397 using mentioned keywords.

The preliminary steps for cleaning all available abstracts are as follows; delete duplicate entries, some unwanted symbols, etc. These steps are implemented with the help of the "NLTK" library. For the task of Classification, relevant abstracts are marked as '0', and the rest of the other abstracts are labelled as '1'. Afterwards, only relevant abstracts are further processed for configuring the data for the task of Q&A. For both annotations of the dataset, two highly dedicated python scripts are written. The Q&A dataset has a specific question: (i) What is the catalyst material used?

The 'Chemdataextractor' library extracts the chemical entities from given relevant class abstracts, followed by a catalyst filter. It means a set of most probable catalyst elements. It helped to dig out only feasible chemical entities as catalyst material. In Figure 2, all reliable catalyst entities are shown with the help of a word cloud plot. After completion of data annotations, the datasets are fed to the BERT model for respective tasks.



Figure: - (2) All extracted catalyst entities are shown in the word cloud.

## 3. The proposed model: H<sub>2</sub>-BERT

The Classification and Q&A tasks are performed using the BERT architecture and only updating the last layer weight. The BERT architecture that we have chosen has 12 encoder layers, with each encoder's input layer could have a maximum of 512 tokens. The attention layer is the significant difference that makes the "BERT" model more robust for text mining than other sequence-to-sequence models. The BERT model can quickly produce sentiment and contextual outputs.

### 3.1. Transfer Learning Framework

Transfer learning is a method to reuse the pre-trained weights of the model by skipping the high computational power and time. In our work, we utilize transfer learning to finetune the pre-trained BERT model's last layer weights on our annotated datasets. The training is carried out on a compute cluster with a 32GB GPU in the Python environment using the “PyTorch” framework. The hyperparameters during finetuning are learning rate, number of epochs, and sequence length. We optimized the following hyperparameters for the best performance: epoch= 3, maximum sequence length = 256, learning rate = 0.000003, hidden activation =” gelu”, hidden dropout = 0.1, hidden size = 768, maximum

position embedding 512, number attention heads = 12, vocabulary size = 28996, and loss function = "Cross-entropy loss". Similarly, for the Q&A task, we have optimized hyperparameters: epochs= 10, learning rate= 0.00004, maximum sequence length = 512, optimizer= AdamW, number of best sizes =3, and loss function is cross entropy. We have implemented the early stopping in both cases to avoid overfitting the model.

## 4. Results and Discussion

### 4.1. Statistical analysis of available relevant abstracts

We have obtained a total number of relevant abstracts equal to 5864 related to "Hydrogen production from alcohol" from an Elsevier database search. We have done some statistical analysis of the abstracts represented in the box plot. The box plot presents the dataset maximum, third quartile, median, first quartile, and minimum. The upper and lower bound of the median have covered the 50%-50% of the dataset, as indicated in Figure 3. The number of sentences in abstracts and the average length of each word in the abstract is shown in Figure 3(a); approximate mean values are 7 and 5, respectively. Similarly, the number of words and a number of stop words are represented in Figure 3(b), with mean values around 200 and 143, respectively.

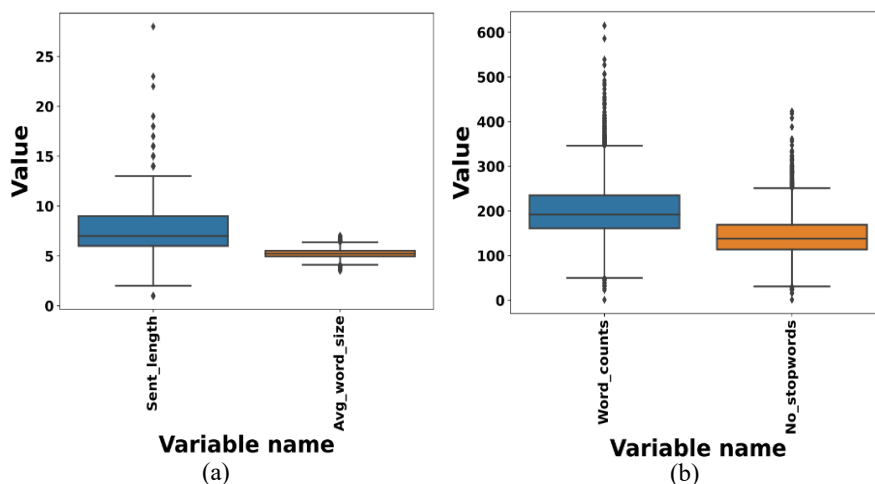


Figure: - (3) The box plots with statistical analysis of some key parameters are shown, (a) Number of sentences and Average length of words present, (b) Number of words and Count of stop words present in each relevant abstract.

### 4.2. Classification and Q&A using $H_2$ -BERT

An annotated dataset for Classification has two classes: relevant and non-relevant. The dataset is well balanced, approximating 50%-50% of both classes. The data set is divided into the training dataset of 70% test dataset of 30%. While training, cross-validation is 30% of the training dataset to make the model robust. The learning curve with train loss and validation loss with values 0.0694 and 0.0785, respectively. The loss function is cross-entropy.

Moreover, the confusion matrix for the test dataset is shown in Figure 4. The final accuracy, precision and recall, and F1-score value of the model for the Train and test dataset are tabulated in Table 1. The accuracy of the Train set is 0.980, and the test has an accuracy value of 0.974.

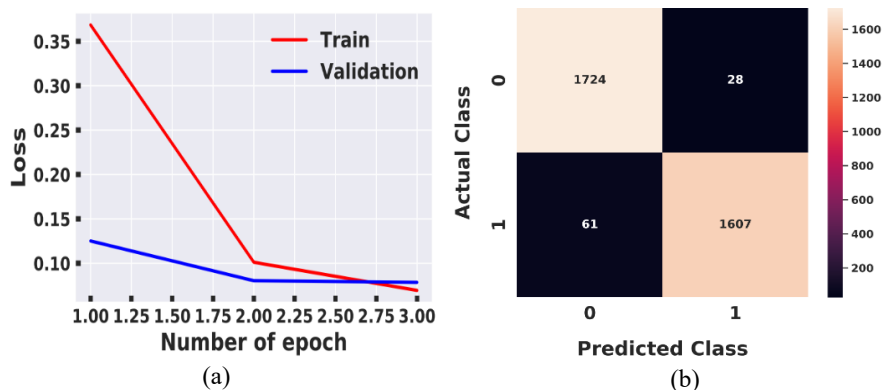


Figure: - (4) The performance evaluation of the classification task by the H<sub>2</sub>-BERT model (a) The variation of loss per epoch (b) Confusion matrix for the test dataset.

Table: - (1) The performance metric values for the task of Classification of the H<sub>2</sub>-BERT.

Type of data	Name of label	Precision	Recall	F1-score	Counts
Train	Relevant (0)	0.98	0.99	0.98	4112
	Non-relevant (1)	0.99	0.97	0.98	3865
Accuracy	0.980				
Test	Relevant (0)	0.97	0.98	0.97	1752
	Non-relevant (1)	0.98	0.96	0.97	1668
Accuracy	0.974				

Similarly, the Q&A task is performed on the architecture of "BERT" and finetuned weights of the last layer only. We have tabulated the performance of this model in Table 2, where Correct, Similar, and Incorrect predictions and total data points are mentioned. The correct answer means that the prediction and true answer are an exact match, and a similar answer means some portion of the prediction and true answer matched; the rest predictions are entirely wrong. If we consider [correct + similar] answers as right out of total data points, the calculated accuracy for Train and test are 1.000 and 0.823, respectively. At last, we took some random abstracts and checked the class of abstract, whether relevant or non-relevant abstract. Subsequently, we took only relevant classes for further processing. A question (What is the catalyst material used?) was asked, the given abstract as context. The predictions are tabulated in Table 3. These two tasks are performed sequentially.

Table: - (2) Performance of the H<sub>2</sub>-BERT for the task of Q&A.

Type of data	Correct	Similar	Incorrect	Total
Train	1734	5	0	1739
Accuracy	1.000			
Test	243	64	66	373
Accuracy	0.823			

Table: - (3) The Q&A prediction on some random abstracts, **bold**, *italic*, and underlined, used for highlighting catalyst entities are present in the abstract, which are predicted correctly, similarly and incorrectly, respectively.

Abstract	Predicted-Catalyst
.... over a series of <b>y-Al<sub>2</sub>O<sub>3</sub></b> supported metal NPs catalyst....	y-Al <sub>2</sub> O <sub>3</sub>



.... For bimetallic <b>nickel-indium</b> alloys, catalysts supported....	nickel-indium
<b>Ni</b> catalysts were prepared by wet impregnation of three....	Ni
.... Recoverable thioporphyrzine catalyst ( <i>CoPz(S-Bu)8/SiO2@Fe3O4</i> ) was prepared...	CoPz(S-Bu)8/SiO2
...Species on the surface of the <u>CuxNi3-xAlOy</u> catalyst can be...	hydrotalcite

## 5. Conclusion

In this proposed work, we have implemented NLP tools for information extraction in the form of catalyst entities from scientific literature data. The text data are accessed with the help of the "Elsevier API key" from the complete database of Elsevier. The extracted abstracts are a mixture of relevant and non-relevant classes. The abstracts are annotated with python scripts for the classification task and Q&A task. These annotated datasets are utilized to retrain the weights of the last layer for task classification and Q&A of "BERT" architecture. The model's performance for Classification and the Train and test accuracy are 0.980 and 0.974, respectively.

Similarly, the accuracy of the finetuned model for Q&A for the Train is 1.000, and the test has 0.823, precisely answering the question. As a result, human effort is expected to be reduced in extracting the names of the catalyst entities for identifying catalysts from scientific text and making the process less time-consuming.

## 6. Acknowledgement

The authors acknowledge the Indian Institute of Technology, Delhi (IIT Delhi) for providing computational resources and a place to carry out this work.

## 7. References

1. Miller, H. A., Bellini, M., Vizza, F., Hasenöhrl, C., & Tilley, R. D. (2016). Carbon supported Au–Pd core–shell nanoparticles for hydrogen production by alcohol electroreforming. *Catalysis Science & Technology*, 6(18), 6870-6878.
2. Kumar, A., Pant, K. K., Upadhyayula, S., & Kodamana, H. (2022). Multiobjective Bayesian Optimization Framework for the Synthesis of Methanol from Syngas Using Interpretable Gaussian Process Models. *ACS omega*.
3. Kuniyoshi, F., Makino, K., Ozawa, J., & Miwa, M. (2020). Annotating and extracting synthesis process of all-solid-state batteries from scientific literature. *arXiv preprint arXiv:2002.07339*.
4. Kumar, A., Ganesh, S., Gupta, D., & Kodamana, H. (2022). A text mining framework for screening catalysts and critical process parameters from scientific literature-a study on Hydrogen production from alcohol. *Chemical Engineering Research and Design*.
5. Huang, S., & Cole, J. M. (2022). BatteryBERT: A Pretrained Language Model for Battery Database Enhancement. *Journal of Chemical Information and Modeling*.
6. Qu, C., Yang, L., Qiu, M., Croft, W. B., Zhang, Y., & Iyyer, M. (2019, July). BERT with history answer embedding for conversational question answering.

# Simulation of low-temperature district heat networks from mine water energy

Thomas Cowley,<sup>a</sup> Timothy Hutty,<sup>a</sup> Solomon Brown<sup>a</sup>

<sup>a</sup> *Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom, s.f.brown@sheffield.ac.uk*

## Abstract

Mine water sourced district heat networks (DHN) are a pathway for reducing emissions of the heating sector. Here, an agent-based model for network dynamics is developed and showcased through a case study in the South Yorkshire Mayoral Combined Authority (SYMCA), consisting of 3437 domestic buildings and 8 non-domestic buildings used as anchor loads. The results show that the mine water DHN achieved a levelised heat (LCOH) cost of 12.6p/kWh compared to the boiler-only scenario of 6.8p/kWh. Emissions in the region could be reduced by 83.6% and can compete with boilers with a natural gas price of 10.8 p/kWh, highlighting the potential of mine water energy for decarbonising the heating sector.

**Keywords:** energy, heat decarbonisation, district heat networks, agent-based modelling

## 1. Introduction

To meet Net-Zero by 2050, the heating sector in the UK needs innovation and development. Currently, 85% of residential buildings utilise a natural-gas grid connection composed of standalone boilers (BEIS, 2022). The UK's Sixth Carbon Budget restricts new gas grid connections by 2025 and prescribes the phase-out of boilers by 2035 (Committee on Climate Change, 2020). This allows district heat networks (DHN) to penetrate the heating sector. Approximately 25% of UK homes are built above abandoned mine workings which are flooded with water between 10-20°C (The Coal Authority, 2021). With this source of low-grade heat, large-scale heat pumps can achieve a higher coefficient of performance (COP) due to the warmer energy source and better thermal efficiency due to the proximity of the source to heat demand.

Mine water networks are susceptible to location-specific parameters, and models should include features such as GIS to improve accuracy and spatial interactions between agents. A review (J. Keirstead et al, 2012) of 219 publications found that only 44% of models incorporated spatial features, and 58% used a yearly temporal scale or greater. Additionally, only a few studies incorporate fully dynamic models, most focusing on techno-economic studies, singular node simulation, and steady-state models (J. Wang et al, 2016). The work of E. Guelpa et al (E. Guelpa, 2020) analyses the relevance of implementing network dynamics such as thermal losses, time delay propagation and thermal transience. The results found that these phenomena had a significant impact when changes to the operating mode (i.e., during the night, start-up, shut-down), demand peaks and large variations in volumetric flow rate. Similarly, work by J. Allegrini (J. Allegrini, 2015) confirms that network dynamics contributed to a 20% discrepancy between the simulated thermal peak request and the measured data. Capturing these dynamics will

vastly improve the accuracy of an extensive network simulation with variable flow at a short time resolution.

It will be of interest to project planners to gauge if a heat network project utilising mine water can reduce emissions and provide competitive heat prices to consumers (BEIS, 2018). Therefore, this work aims to model the network dynamics using an agent-based approach to decrease computational intensity, allow flexibility of location-based parameters, and incorporate network dynamics in an extensive network. Undertaking techno-economic analysis while capturing these features will help understand the trade-off between the cost and emissions of a region with mainly standalone gas boilers. Hence, the objectives are to:

- Develop an agent-based model to capture the network dynamics of DHNs
- Showcase the model through a techno-economic case study and determine at what price of natural gas the proposed heat network becomes a viable investment.

## 2. Methodology

This work aims to simulate the extraction of mine water energy from unused coal mine workings to supply domestic and non-domestic heating demand. The mine water heat is upgraded via large-scale heat pumps at a geothermal plant located above the mine workings, intersected by a borehole, and distributed around a network of pipes. A visual representation of the model is shown in Figure 1. Simulation of a DHN is an inherently complex system; therefore, to reduce the computational intensity of the model, only the phenomena that have a significant impact on the accuracy of the network dynamics are included. Thus, the following assumptions have been made:

- Water properties are constant parameters.
- Water is considered an incompressible fluid.
- Flow within pipeline agents is one-dimensional.
- Heat loss to the surroundings is one-dimensional.
- There is no degradation of piping material or insulation over the simulation time
- Axial conduction in each pipe is considered insignificant.
- The pressure drops in the pipeline agents have negligible viscous heating effects.
- Perfect mixing and adiabatic operation when flow combines at junctions.
- Non-constant terms in the momentum equation are neglected
- Pressure loss and efficiency losses within the HIU are considered insignificant

The model uses node agents, which consist of geothermal plant centres (GCs), demand centres (DCs), junctions (J) and pipes with start and end points at nodes and are classed as pipeline agents (P). An arbitrary number of pipeline agents can connect to a single node. Each node agent has several parameters, which are pre-determined and describe its location and whether there is a physical connection to the network. Each pipeline agent is tied to the corresponding start and end node agents. The length of each pipeline agent is calculated by geolocating the distance between the start and end node agents. All primary agents (GC, DC, J, P) interact with each other within a GIS space. Water packet agents are classed as secondary agents and are instantaneously generated in the case of volumetric flow or temperature changes in the network. The water packet agents allow heat propagation from a source at a finite speed, so that time delays over large distances are captured. The pipeline agents form the distribution network that follows the network's topology, while junction agents allow the splitting of flow and integration of loops. Generation centres and junctions have pumping stations that maintain the network's

pressure, and the pressure block calculates its pumping requirements. The demand model block uses a linear regression model at multiple time intervals and assigns buildings to five archetypes: non-domestic (housing), leisure, education, commercial and residential. The thermal sub-model calculates the radial heat loss from the pipe to the soil and accounts for transient heat loss when the water in the pipes is stationary, and there is no flow. The mass flow model uses linear programming to maximise the flow through the distribution network.

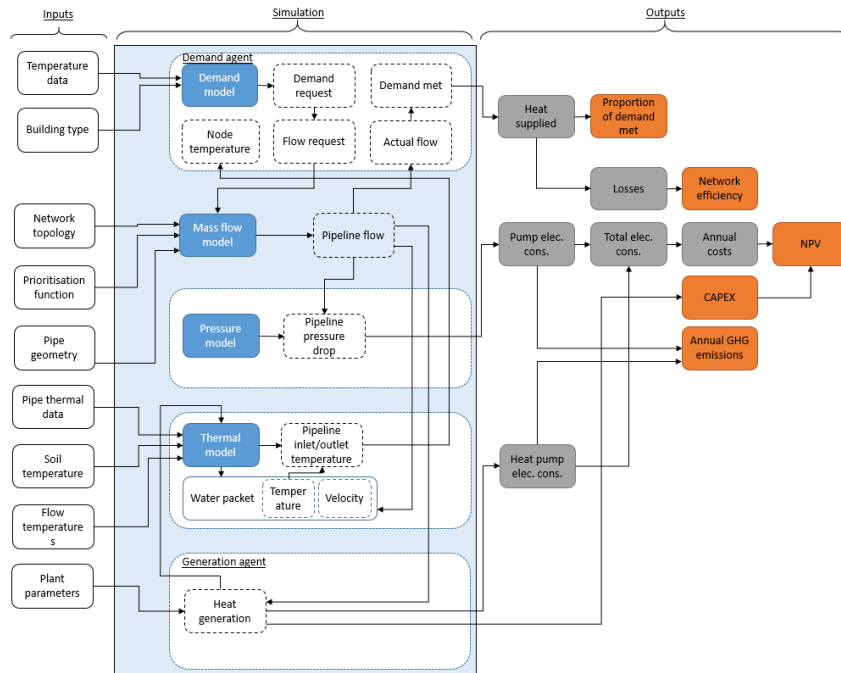


Figure 1 - Flowchart of heat network dynamic model

### 2.1. Case study

The case study location has many flooded mine workings; this provides an ambient and stable body of renewable low-grade heat. A borehole was chosen from the Coal Authority BGS database (British Geological Survey, 2022) nearest the heat demand and intersected the flooded mine workings at a temperature of 18 °C and a depth of 332 m. 3445 buildings were allocated to the five building archetypes (Commercial: 2, Education: 4, Leisure: 1, Residential: 1), with 3437 domestic buildings aggregated into 13 different demand agents. Eight non-domestic buildings are used as anchor loads to ensure consistent demand on the network. Analysis of the costs incurred by negating emissions from standalone boilers will inform the region of the use of mine water energy to decarbonise its heating sector; therefore, the case study is run with the same demand profiles generated from the demand model block for both the mine water heat network and boilers-only scenarios, while assuming an 80% efficiency for each individual boiler. Table 1 shows the parameters for the case study. A COP of 4.9 (A. David et al, 2017) has been selected as it is aligned with values from the literature that aggregate the supply and return temperatures of 58 different European heat networks. Table 2 shows the cost parameters used for the techno-economic analysis. Temperature data from 2021 (Met Office, 2021) is used as input into the demand

sub-model. The network length is generated by the model and is ca. 12 km. Pipe dimensions were calculated pre-simulation using insulation class 1 pipes (Fintherm, 2019). The maximum pipe velocity was 5 m/s to prevent rapid degradation of piping material and to adhere to noise, safety and vibration regulation.

Table 1 – Network parameters

	Value
Pump efficiency (%)	80
Motor efficiency (%)	80
Network pressure (kPa)	1,500
Supply temperature (°C)	68
Return temperature (°C)	18
Soil temperature (°C)	10
Water conductivity (W/mK)	0.598
Water diffusivity (mm <sup>2</sup> /s)	0.16
Lifetime (years)	30
Carbon factor natural gas (t/MWh)	0.210
Average carbon factor electricity (t/MWh)	0.136

Table 2 – Cost parameters

	Value
Electricity price (p/kWh)	13
Plant cost (£k/MW)	666.8
Gas price (p/kWh)	4.94
Operation costs (£/kWh)	35.9
Demand Centre costs (£/kWh)	919
Substation cost (£/kWh)	16
Ancillary cost (£/kWh)	68
Pipework cost (£/kWh)	516
Boiler CAPEX and OPEX cost (£/kWh)	2.28
Inflation rate (%)	2.7
Discount rate (%)	3.5

### 3. Results & Discussion

The simulation was run for one year (8760 hours) for the mine water scenario and the boiler-only scenario using temperature data from 2021. Tables 3 and 4 show the KPIs (Key performance indicators) and costs for both scenarios. The mine water scenario showed a high efficiency of 96.3% by the heat network and met 99.98% of the region's demand. This highlights the model's success in distributing the heat effectively to consumers while minimising the heat lost to the surrounding soil compared to the generated heat. 0.02% of the demand was not supplied to the consumers due to the water velocity constraints. These consumers tended to be at the end of the network.

Figure 2 shows the total demand on the network across the year and the variations over each day. As the network consists of mainly non-domestic buildings, this has a more significant effect on the demand profile of the network. The demand on the network during winter is greater than during summer, so the buildings are maintained at a comfortable temperature. Demand peaks between 6:30 and 10 am as more hot water is needed for showers before people go to work, and building heating is turned on. There are also peaks between 16:00 and 23:00, as the heating is turned back on as people get back from work. This is less prominent in the summer months, but there is still, on average, some smaller peaks in the morning. During the night, the network cools, and this is due to heat being lost when the water in the network is stagnant; this is the effect of transience in the network. Positive demand is seen on the network throughout the year due to the anchor loads provided by the non-domestic building types.

BEIS reported that the average household energy consumption of a typical domestic dwelling was 12.3 MWh/year (A. O'Mahoney, 2020); this corroborated the demand model results with an average of 12.2 MWh/year over the 3437 domestic buildings. The mine water network scenario emitted 1807 tCO<sub>2</sub>e, whereas the boiler-only scenario 11013 tCO<sub>2</sub>e, thus achieving achieved 83.59% reduction in emissions compared

to boiler-only scenario. Over a 30-year lifetime, 276 million tCO<sub>2e</sub> could be saved by 2050. Mine water LCOH (Levelised cost of heat) obtained was 12.4 p/kWh, which is slightly more economical than work by O. Gudmundsson et al (O. Gudmundsson et al, 2013), which found that geothermal DHN achieved 14.0 p/kWh for outer city projects. The boiler-only scenario obtained an LCOH of 6.8 p/kWh, significantly lower than the literature value of 11p/kWh. The disparity may have been due to lower efficiencies in actual gas boiler installation or underestimated cost parameters. The NPV (Net present value) of the boiler-only scenario is -£90.5 million, whereas the heat network NPV requires 46.3% more investment totalling -£132.4 million. The disparity in NPV is primarily due to the already extensive gas infrastructure in the UK and the price of natural gas. The mine water network scenario can abate emissions at an average cost of 151.7£/tCO<sub>2e</sub> over its lifetime.

Table 3 – KPIs from the case study

	Mine water	Boiler-only
Efficiency (%)	96.3	80
Heat demand (MWh/year)	41,964	-
Heat supplied (MWh/year)	41,955	-
Heat generation (MWh/year)	43,594	52444
Demand supplied (%)	99.98	100
Pumping power (MWh)	529.55	-
Heat pump output (MW)	18.75	-
Emissions (tCO <sub>2</sub> )	1,807	11,013
Emission savings (%)	83.59	0

Table 4 – Costs over a 30-year lifetime

	Mine water	Boiler-only
Total capital cost (£)	60,880,000	24,421,000*
Total operating cost (£)	43,077,000	-
Total energy cost (£)	28,418,000	66,054,000
Total cost over lifetime (£)	132,379,000	90,475,000
LCOH (p/kWh)	12.38	6.77

\*Both capital and operational cost

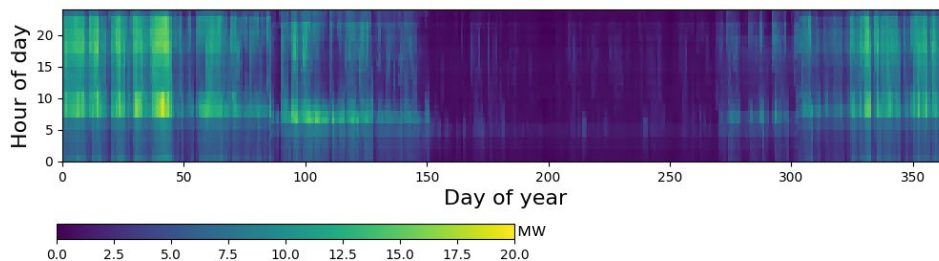


Figure 2 – Average heat demand of nodes in the heat network over the year (and over the day)

The wholesale gas price has seen a significant increase in the final quarter of 2021 due to the high demand for natural gas, and even more so in 2022 with the start of the war in Ukraine. 2020 saw the wholesale gas price at a minimum of 4.39 p/kWh, and a maximum of 27.30 p/kWh (Trading Economics, 2021); therefore, DHN would provide a more stable price of heat and improved energy security for customers by providing protection from global fluctuating gas prices. If the gas price trends upward, the mine water scenario becomes more economically attractive. The sensitivity analysis in Figure 3 shows that mine water can provide competitive prices at a natural gas price of 10.8p/kWh.

Figure 4 shows the percentage change in the LCOH when the total heat demand, capital cost and operational costs are modified by plus or minus 50%. The network is more sensitive to demand changes than operational costs, as it increases by roughly 18%

when the demand on the network is reduced by half. In contrast, it decreases by roughly 5% when demand is increased by half. Moreover, decreasing the capital and operational costs by 50% leads to a decrease in the LCOH of 16% and 24%, respectively. The energy costs of the Levelized cost method are assumed to be constant over the technologies' lifetimes. They are not the prices to sell energy. Hence, it does not reflect energy prices' short-term or long-term volatilities.

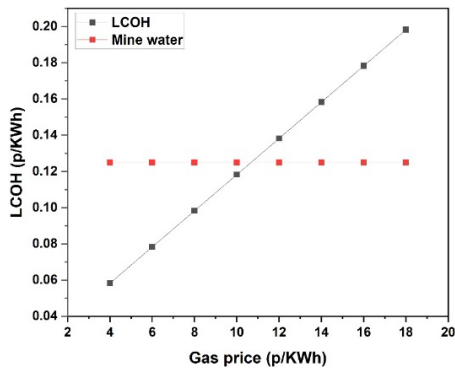


Figure 3 – LCOH as a function of gas price

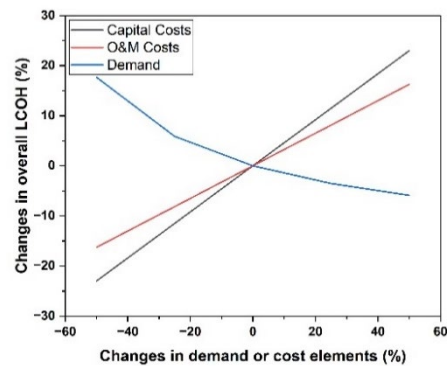


Figure 4 – LCOH as a function of changes to costs and demand

#### 4. Conclusion

Utilising renewable energy sources with DHN is a possible way of decarbonising the UK's heating sector. Mine water energy paired with large-scale heat pumps is proven throughout European projects. Current research lacks accurate modelling of network dynamics. The model successfully captured the network dynamics and expected network demand profiles and obtained an efficiency of 96.3%. The case study was found to reduce the region's carbon emissions by 83.59% and provide an average LCOH of 12.38 p/kWh over the project's lifetime. Sensitivity analysis shows that the project became economically feasible at a natural gas price of 10.8p/kWh, highlighting a promising pathway to reduce emissions.

#### 5. References

- A. David, et al, 2017, "Heat Roadmap Europe: Large-Scale Electric Heat Pumps in District Heating Systems", *Energies*, 10, 4, 578.
- A. O'Mahoney, 2020, "Review of Typical Domestic Consumption Values 2021," The Office of Gas and Electricity Markets.
- BEIS, 2018, "Heat Networks: Ensuring sustained investment and protecting consumers," GOV.UK.
- BEIS, 2022, "Decarbonising heat in homes," GOV.UK.
- British Geological Survey, "Coal Authority Borehole Log Data," GOV.UK.
- Committee on Climate Change, 2020, "The Sixth Carbon Budget,".
- E. Guelpa, 2020, "impact of network modelling in the analysis of district heating systems," *energy*, vol. 213, p. 118292.
- FINTHERM, 2019, "Pre-insulated pipes and accessories catalogue,".
- J. Allegrini, K. Orehounig, G. Mavromatidis, F. Ruesch, V. Dorer, a. R and Evins, 2015, "A review of modelling approaches and tools for the simulation of district-scale energy systems," *Renewable and Sustainable Energy Reviews*, pp. 1391-1404.
- J. Keirstead, M. Jennings and A. Sivakumar, 2012, "A review of urban energy system models: Approaches, challenges and opportunities," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 3847-3866.
- J. Wang, Z. Zhou and J. Zhao, 2016, "A method for the steady-state thermal simulation of district heating systems and model parameters calibration," *Energy Conversion and Management*, vol. 120, pp. 294-305.
- Met Office, 2021, "MIDAS Open: UK daily temperature data, v202107", NERC EDS Centre for Environmental Data Analysis.
- O. Gudmundsson et al, 2013, "Cost analysis of district heating compared to competing technologies", *WIT Transactions on Ecology and the Environment*, 176 107-118.
- The Coal Authority, 2021, "Mine Water Heat," GOV.UK.
- Trading Economics, 2021, "UK Natural Gas,".

# **In-silico investigation of microalgae culture performance in lab and pilot scale: The trade-off between reactor size and productivity.**

Christos Chatzidoukas\*, Vasileios Parisis

*Department of Chemical Engineering, Aristotle University of Thessaloniki*

*(AUTH), P.O. Box: 472, 54124, Thessaloniki, Greece*

[chatzido@auth.gr](mailto:chatzido@auth.gr)

## **Abstract**

The biomass productivity of photo-autotrophic microalgae is known to be influenced by a plethora of environmental factors, with light being the most notable one. Given the critical role of light in the complex metabolic network of microalgae, conditionally directing the carbon flow either towards biomass growth or lipid accumulation, the empirical operational design of a photobioreactor system can only be suboptimal. This task becomes even more mazy in larger scales, where the culture growth results in more intensive light attenuation. The present work analyses the importance of a structured, distributed mathematical model for the design of optimal lighting policies of a photobioreactor, and investigates the successful shift of microalgae cultures from a lab to a pilot scale unit. The key contribution of this study is to help realize the performance bounds of a photobioreactor depending on its geometry and size and to showcase how to deal with the lighting policy when scaling-up of a microalgae bioreactor is desired.

**Keywords:** Lab/pilot scale photobioreactor; Structured macroscopic/kinetic model; Dynamic optimization; Light path; Photobioreactor design.

## **1. Introduction**

Despite the ever increasing academic and industrial attention and research on microalgae based biorefineries over the last decades, the commercialization of microalgae products is hindered by the low productivity of photobioreactors and open culture systems. Particularly in autotrophic cultures, light is the most notable limiting growth factor. Its availability inside the culture volume is significantly limited by the high optical density of microalgae suspensions. Since the measurement of light intensity inside the culture often requires expensive equipment, the empirical deduction of lighting policies that maximize biomass production is unfeasible. Model-based optimization has the potential to facilitate the above task, supporting the sustainable use of microalgae in nutritional and biofuel applications, and fill in the gaps in our understanding of the impact of the lighting on microalgae growth.

In this study, a multi-physics mathematical model is developed and validated, with the view to systematically obtain the optimal lighting policies that maximize biomass growth inside stirred photobioreactors. Three scenarios are devised that reflect the impact of scale up on process performance, which is assessed by proper process KPIs.

## **2. Model development**

In a photoautotrophic culture, microalgae fix CO<sub>2</sub> into organic carbon photosynthetically, and subsequently mobilize it, in the presence of nitrogen and phosphorus, to synthesize



the cellular machinery needed for cell division and maintenance. Furthermore, microalgae store their organic carbon in the form of carbohydrates and/or lipids, especially when their environment becomes limiting, i.e., nutrient exhaustion. A distributed/structured dynamical model is developed to describe the dynamic operation of typical stirred photobioreactors (PBRs) employed for microalgae cultures. The model has been built up upon the well-established Droop model (Droop, 1983), incorporating biological phenomena like photoinhibition and photo acclimation, via chlorophyll-a dynamics. Carbon allocation and product synthesis (i.e., carbohydrates, proteins, and lipids) are dependent on the internal state of the cell as well as the nutrient availability in the culture medium. The one-dimensional Cornet model (Cornet et al., 1992a, 1992b) is adopted and modified in order to describe the light gradients that develop inside the photobioreactor along with the evolution of microalgal biomass and pigment content. Tables 1 and 2 summarize the dynamic mass balances and the specific rates of the model.

**Table 1:** Dynamic mass balances of the model.

$$\begin{aligned} \frac{d(V_l[O_{2,l}])}{dt} &= Q_{l,fd}[O_{2,l,fd}] - Q_{l,out}[O_{2,l}] + V_lOTR + V_lY_{CO_2}\langle\mu\rangle X - V_lY_{CO_2}RX \\ \frac{d(V_l[CO_{2,l}])}{dt} &= Q_{l,fd}[CO_{2,l,fd}] - Q_{l,out}[CO_{2,l}] + V_lCTR + V_lY_{CO_2}RX - V_lY_{CO_2}\langle\mu\rangle X \\ \frac{d(V_l[O_{2,g}])}{dt} &= Q_{g,fd}[O_{2,g,fd}] - Q_g[O_{2,g}] - V_lOTR \\ \frac{d(V_g[CO_{2,g}])}{dt} &= Q_{g,fd}[CO_{2,g,fd}] - Q_g[CO_{2,g}] - V_lCTR \\ \frac{d(V_l)}{dt} &= -\frac{d(V_g)}{dt} = Q_{l,fd} - Q_{l,out} \\ \frac{d(V_l[N_l])}{dt} &= Q_{l,fd}[N_{l,fd}] - Q_{l,out}[N_l] - V_l\langle\rho_N\rangle X \\ \frac{d(V_l[P_l])}{dt} &= Q_{l,fd}[P_{l,fd}] - Q_{l,out}[P_l] - V_l\rho_P X \\ \frac{d(V_lXq_N)}{dt} &= \langle\rho_N\rangle \\ \frac{d(V_lXq_P)}{dt} &= \rho_P \\ \frac{d(V_lX)}{dt} &= Q_{l,fd}X_{in} - Q_{l,out}X + V_l(\langle\mu\rangle - R)X \\ \frac{d(V_l\theta X)}{dt} &= V_l\frac{K_{acc1}\theta}{\theta_{ad}}(\theta_{ad} - \theta), \frac{\theta_{ad}(I)}{q_N} = \varphi_{max}\frac{k_I^*}{k_I^* + (I)} \\ \frac{d(V_lC)}{dt} &= Q_{l,fd}C_{fd} - Q_{l,out}C + V_l\langle\mu\rangle X - V_lr_{C,Pr}X - V_lr_{C,L}C - V_lRX \\ \frac{d(V_lL)}{dt} &= Q_{l,fd}L_{fd} - Q_{l,out}L + V_lr_{C,L}C + V_lD_pPr \\ \frac{d(V_lPr)}{dt} &= Q_{l,fd}Pr_{fd} - Q_{l,out}Pr + V_lr_{C,Pr}X - V_lD_pPr \end{aligned}$$

where:  $[O_{2,i}]$   $[CO_{2,i}]$  oxygen and carbon dioxide concentrations (gas & liquid phase);  $[N_l]$ ,  $[P_l]$  nitrogen and phosphorus concentrations;  $q_N$ ,  $q_P$ ,  $\theta$  nitrogen and phosphorus and chlorophyll-a cellular quotas;  $X$ : biomass;  $C$ : carbohydrates;  $L$ : lipids;  $Pr$ : proteins;  $R$ : respiration rate;  $OTR$ ,  $CTR$ :  $O_2$  and  $CO_2$  transfer rates;  $Y_{CO_2}$ :  $CO_2$  to biomass yield coefficient;

### 3. Model validation and biomass optimization.

Experimental data from the work of Psachoulia and Chatzidoukas, (2021) were utilised in the scope of parameter tuning. The datasets contain dynamic information about the main macroscopic process variables (nutrients, biomass, intracellular products) in *Stichococcus sp.* photoautotrophic cultures in a lab stirred photobioreactor. The modified Cornet sub-model was validated offline using the experimental data from Ma et al., (2022).

**Table 2:** Specific growth/uptake/production/consumption rates of the model.

$$\rho_P = \rho_{P,max} \left( \frac{[P_l]}{K_{S,P} + [P_l]} \right) \left( \frac{q_{P,max} - q_P}{q_{P,max} - q_{P,min}} \right)$$

$$\rho_N(r) = \rho_{N,max} \left( \frac{[N_l]}{K_{S,N} + [N_l]} \right) \left( \frac{q_{N,max} - q_N}{q_{N,max} - q_{N,min}} \right) \left( \frac{q_P - q_{P,min}}{q_{P,max} - q_{P,min}} \right) \left( \eta + (1 - \eta) \frac{l}{l + \epsilon_l} \right)$$

$$\mu(r) = \mu_{max} \frac{l}{l + K_{S,l} + \frac{l^2}{K_{il}}} \frac{S_{CO_2,l}}{K_{S,C} + S_{CO_2,l}} \left( 1 - \frac{[O_2,l]}{[O_2,l,max]} \right) \left( 1 - \frac{q_{N,min}}{q_N} \right) \left( 1 - \frac{q_{P,min}}{q_P} \right)$$

$$r_{C,Pr} = \alpha \langle \rho_N \rangle$$

$$r_{C,L} = \mu_{c,s} + \mu_{c,m} = \mu_{c,m,max} \frac{q_N}{q_{N,max}} + \mu_{c,s,max} \left( 1 - \frac{q_N}{q_{N,max}} \right)$$

$$D_p = \gamma \mu_{c,s}$$

$$I(r) = I_0 \frac{4\alpha_1}{(1 + \alpha_1)^2 e^{a_2 r} - (1 - \alpha_1)^2 e^{-a_2 r}}; \alpha_1 = \sqrt{\left( \frac{E_a' k_a \theta}{E_a' k_a \theta + E_s} \right)}; \alpha_2 = (E_a' k_a \theta + E_s) \alpha_1 X$$

$$\langle z_i \rangle = \frac{1}{V_l} \int_{V_l} z_i(x) dx, \text{ where } z_i: \rho_N(r), \text{ or } \mu(r), \text{ or } I(r)$$

where:  $\rho_P$ : specific phosphorus uptake rate;  $\rho_N(r), \mu(r)$ : specific nitrogen uptake rate & specific growth rate at depth r;  $r_{C,Pr}, r_{C,L}$  specific carbon allocation rate to proteins & lipids;  $D_p$  specific protein degradation rate to lipids;  $I_0, I(r)$  Incident light intensity & light intensity at depth r.

Subsequently, the validated model was used to guide the optimization of a batch photobioreactor, in the context of biomass production, on three different spatial scales. The incident light intensity on the photobioreactor lateral surface,  $I_0$ , is the sole control variable of the problem, and is modelled as a piecewise-constant function using ten 1-day time intervals. The three different scenarios are presented below, whereas the characteristic dimensions of the three photobioreactors are detailed in Table 3:

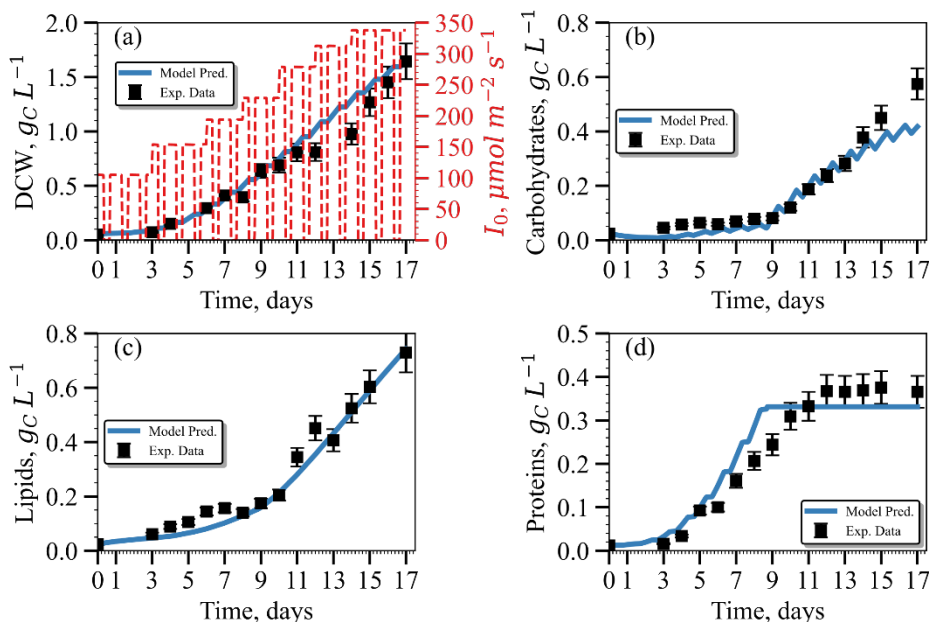
**Table 3:** Characteristic dimensions and optimization strategies of the three PBRs.

PBR ID	Diameter, T (m)	Height, H (m)	Working volume, $V_l$ (L)
1	0.16	0.16	3.2
2	0.32	0.32	25
3	0.16	9.9	200
<b>Scenario 1:</b> Maximization of the final biomass concentration in the lab PBR.			
Task: $\max_{I_0} \mathbf{Obj}_1$	$Obj_1 = X_{PBR1}$		Operational constraint $50 \leq I_0 \leq 2,000$ $\mu mol m^{-2} s^{-1}$
w.r.t model equations of <b>Tables 1 and 2</b>			
<b>Scenario 2:</b> Optimal tracking of the biomass profile achieved in the lab scale by the semi-pilot PBR.			
Task: $\min_{I_0} \mathbf{Obj}_2$	$Obj_2 = \int_0^{t_f} \left( \frac{X_{PBR1} - X_{PBR2}}{X_{PBR1}} \right)^2 dt$		Operational constraint $50 \leq I_0 \leq 3,000$ $\mu mol m^{-2} s^{-1}$
w.r.t model equations of <b>Tables 1 and 2</b>			
<b>Scenario 3:</b> Optimal tracking of the biomass profile achieved in the lab scale by the pilot PBR.			
Task: $\min_{I_0} \mathbf{Obj}_3$	$Obj_3 = \int_0^{t_f} \left( \frac{X_{PBR1} - X_{PBR3}}{X_{PBR1}} \right)^2 dt$		Operational constraint $50 \leq I_0 \leq 3,000$ $\mu mol m^{-2} s^{-1}$
w.r.t model equations of <b>Tables 1 and 2</b>			

An energy efficiency factor, adopted to compare the performance of the three PBRs, is defined as the produced biomass per unit of lighting power input:  $e = \frac{x}{\alpha \cdot I_0 \cdot A_{lat}} \cdot \left(\frac{gL^{-1}}{W}\right)$ , where  $A_{lat}$  is the illuminated surface area of the PBR and  $\alpha = 0.219$  is a conversion factor between  $\mu\text{mol m}^{-2} \text{s}^{-1}$  and  $W \text{ m}^{-2}$ .

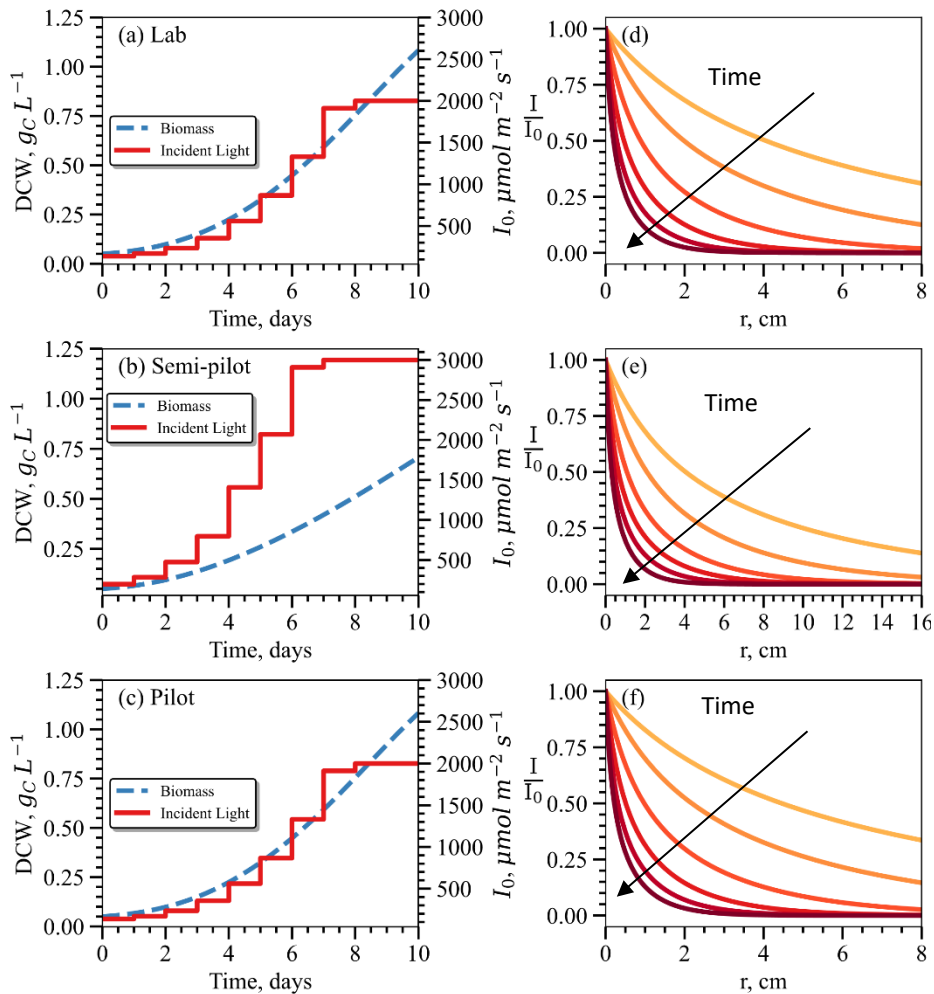
#### 4. Results and discussion.

The model predictions are plotted against the experimental data in Figure 1. The dynamic trajectory evolution of the main state variables is represented by the model, thereby establishing a sound basis for further exploration of the optimal lighting policies that maximize production on different scales.



**Figure 1:** Comparison of the model predicted time trajectories to the experimental data: a) biomass dry cell weight concentration; b) carbohydrates; c) lipids and d) proteins.

Figure 2 depicts the optimal lighting policy for a batch mode cultivation along with the corresponding biomass evolution trajectory, on the three spatial scales. The increasing culture density promotes progressively steeper light gradients inside the culture volume. This results in constantly escalating lighting demand until the point of saturation. In Scenario 1, the overall increased energy input to the system, compared to the validation experiment, results in approximately a 35% increase in volumetric productivity, demonstrating the value of model-based optimization. In Scenario 2, the lighting profile that will steer the semi-pilot PBR on the respective biomass growth trajectory attained in the lab PBR is sought. Despite the increase in light provision in this setup serious loss in biomass productivity is observed. This is entirely attributed to the large proportion of “dark volume” in the semi-pilot PBR as one can deduce inspecting the spatial profile of light distribution reducing sharply over time (Figure 2 d-f). On the contrary, in Scenario 3,



**Figure 2:** Light optimization in three spatial scales. (a-c): biomass trajectory and incident light profile; (d-f): Light gradients inside the PBRs during a batch cultivation.

the biomass evolution in the PBR with the largest height/diameter ratio is identical to its lab scale counterpart. Though it may seem that tall and skinny geometries can pave the way to a successful scale-up of microalgae cultures, one should address this matter from a techno-economic viewpoint as well. Indeed, Figure 3 demonstrates that the photobioreactor size is practically correlated with a reduced lighting efficiency, and this is consistent across the culture time horizon. This point is in agreement with the findings of Nikolaou et al., (2016). The value of ‘ $e$ ’ decreases by about one order of magnitude during scale up from the lab scale to the semi-pilot scale (increase in reactor radius from 8 cm to 16 cm), and furthermore decreases by about two orders of magnitude on the pilot scale, compared to the lab scale. This comparison indicates that PBRs with the same specific illuminated area per unit volume can attain the same productivity with identical light flux, regardless of their volume. However, in our viewpoint, the absolute light power (W) and not the luminous flux ( $W m^{-2}$ ) spent is a more appropriate criterion to gauge PBR efficiency on a techno-economic basis.

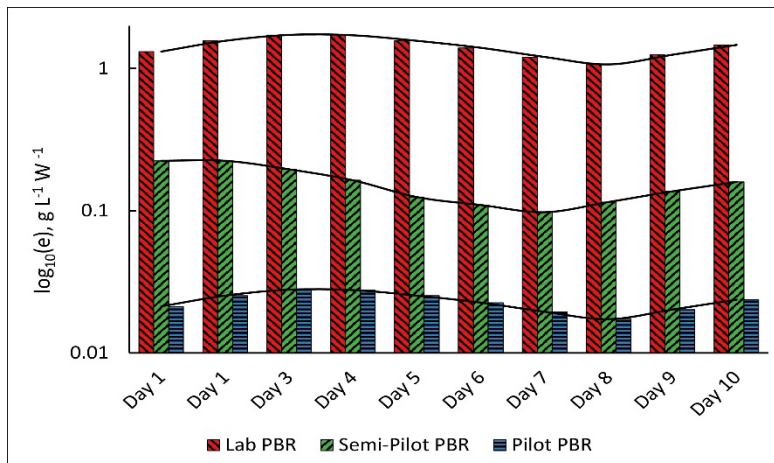


Figure 3: Dynamic evolution of the energy efficiency factor on the three different scales.

## 5. Conclusions.

In this work the notion of light limitation on photobioreactor is confirmed and quantified, and an attempt to explore strategies that maximize biomass production in photobioreactors on three different scales has been made. Results suggest that the industrial cultivation of microalgae on “traditional” fermenter like photobioreactors seems unrealistic. Furthermore, it is considered that is of utmost importance that photobioreactor design, process optimization and scale-up should be coupled activities, due to the dominant effect of the characteristic length of light distribution on the performance of these systems. We are confident that the present modeling framework can serve at the epicenter of a multi-physics platform that can provide valuable insight and guide the successful commercialization of microalgae based biorefineries.

## References

- Cornet, J.F., Dussap, C.G., Cluzel, P., Dubertret, G., 1992a. A structured model for simulation of cultures of the cyanobacterium *Spirulina platensis* in photobioreactors: II. Identification of kinetic parameters under light and mineral limitations. *Biotechnol. Bioeng.* 40, 826–834.
- Cornet, J.F., Dussap, C.G., Dubertret, G., 1992b. A structured model for simulation of cultures of the cyanobacterium *Spirulina platensis* in photobioreactors: I. Coupling between light transfer and growth kinetics. *Biotechnol. Bioeng.* 40, 817–825.
- Droop, M.R., 1983. 25 Years of Algal Growth Kinetics A Personal View. *Bot. Mar.* 26.
- Ma, S., Zeng, W., Huang, Y., Zhu, Xianqing, Xia, A., Zhu, Xun, Liao, Q., 2022. Revealing the synergistic effects of cells, pigments, and light spectra on light transfer during microalgae growth: A comprehensive light attenuation model. *Bioresour. Technol.* 348, 126777.
- Nikolaou, A., Booth, P., Gordon, F., Yang, J., Matar, O., Chachuat, B., 2016. Multi-Physics Modeling of Light-Limited Microalgae Growth in Raceway Ponds. *IFAC-Pap.* 49, 324–329.
- Psachoulia, P., Chatzidoukas, C., 2021. Illumination Policies for *Stichococcus* sp. Cultures in an Optimally Operating Lab-Scale PBR toward the Directed Photosynthetic Production of Desired Products. *Sustainability* 13, 2489.

# Process modelling of the hydrothermal liquefaction of oil-palm waste for biocrude and hydrochar production

Muhammad Shahbaz<sup>a\*</sup>, Mohammad Alherbawi<sup>a</sup> Prakash Parthasarathy<sup>a</sup>,  
Gordon McKay<sup>a</sup>, Tareq Al-Ansari<sup>a\*</sup>

<sup>a</sup> *College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar*

<sup>b</sup> *Centre of Bio-Fuel and Biochemical Research, Department of Chemical Engineering, Universiti Teknologi PETRONAS, 32610 Bandar Seri Iskandar, Perak, Malaysia.*  
Email: [talansri@hbku.edu.qa](mailto:talansri@hbku.edu.qa), [mshahbaz@hbku.edu.qa](mailto:mshahbaz@hbku.edu.qa)

## Abstract

Energy production from biomass has promising prospects in the domain of clean and sustainable sources. Hydrothermal Liquefaction (HTL) is an emerging process to convert biomass into energy products and is less energy intensive than other thermal conversion processes including gasification and pyrolysis, as HTL does not require pre-drying of biomass. In this study, a process simulation is developed using Aspen Plus<sup>®</sup> V10 to quantify the yield of liquefaction products including bio-crude, hydrochar and gas from various palm oil wastes; such as palm kernel shell (PKS), empty fruit bunches (EFB), palm oil fronds (POF) and their blends. The use of PKS as feedstock for HTL resulted in the highest biocrude yield of ~441 kg/tonne, whereas the EFB was linked to the lowest biocrude yield of ~293 kg/tonne. Meanwhile, the highest yield of hydrochar was achieved at ~193 kg/tonne for EFB. The gas production was also higher in the case of EFB, at about 317 kg/tonne. The blend of all three wastes yielded ~356 kg/tonne of biocrude, 102 kg/tonne of hydrochar and 387 kg/tonne of gas. In addition, the produced gas composition is dominated by CO<sub>2</sub> followed by CH<sub>4</sub>. The blended feed yielded the lowest CO<sub>2</sub> composition, which indicates the importance of feedstock blending in CO<sub>2</sub> mitigation. POF is found to be a more economical feedstock which has shown the lowest cost of 0.331USD/kg and 0.045 USD/kg. The generated biocrude replicates several petroleum properties, granting it a high potential for conventional refining into clean transportation fuels. This study provides insights and technical knowledge on HTL technology and its possible upgradation from lab and pilot into commercial scales.

**Keywords:** Liquefaction; Oil-palm waste; Aspen Plus; Biocrude; Hydrochar.

## 1. Introduction

The existence of modern life standard is in danger due to the increasing energy demand and its associated environmental impact due to the high dependency on fossil fuels. The domino effect of these issues is not only extended to energy deficiency, and the world existence inquest is due to climate change, which caught the world's attention at the research and policy level to explore new and alternative renewable and sustainable energy

sources (Ali et al., 2022). In all alternative energy resources, biomass emerges as an attractive and viable source due to its wide availability and possible conversion into all forms of energy, materials, and chemicals that are currently obtained from fossil fuels (AlNouss et al., 2022).

Biochemical and thermochemical routes can convert biomass into energy derivatives and chemicals. Biomass can be converted into gaseous, liquid, and solid fuels and value-added chemicals through thermochemical and biochemical conversion processes (Chan et al., 2019; Chan et al., 2015). The thermochemical conversion processes including gasification and incineration are used for gaseous fuel, heat, and electricity production. In contrast, pyrolysis and liquefaction are utilised for liquid fuel and char generation. All the above processes are relatively energy-extensive, with liquefaction being the least-energy consuming. Liquid fuel is vital in energy products due to its high energy density, ease of storage and transportation (Chan et al., 2015). Liquefaction of biomass is a relatively new process compared to the other process. Still, it appears to be very effective for liquid fuel production due to its low energy requirement, no drying and pre-treatment process (Chan et al., 2015). In the liquefaction process biomass is treated at a lower temperature than 400 °C and high pressure of 4 - 20 MPa using an appropriate solvent such as; water, organic solvent, and their mixture (Chan et al., 2018).

Palm oil waste (POW) is Southeast Asia's major crop, supplying 79% of the world's palm oil demand (Umar et al., 2021). POW is primarily used in the industry for electricity and heating purposes. Chan et al. (2019) reviewed the utilization of waste in Malaysia for energy production via thermochemical conversion pathway and noticed that most of POW were investigated for gasification and pyrolysis processes for gaseous fuel and bio-oil production. Umar et al. (2021) recently reviewed the POW utilisation for energy production and found that the gasification and pyrolysis processes are well investigated at pilot and commercial scales for gaseous fuel generation. On the other hand, liquefaction was mainly investigated at the lab and pilot scale levels (Chan et al., 2019). The first objective of this study is to develop a flowsheet model of liquefaction process using Aspen Plus® to investigate the process parameters on the yield of solid crude, hydro crude, and gas for each type of POW and their blend. HTL is used to produce for liquid production and gas is obtained as a by product or waste. While the second objective is to conduct a techno-economic analysis using Aspen Plus Economic Analyzer to evaluate the production cost of each product for all types of selected feedstocks. This study may provide insights for researchers, as well as policymakers to make decisions on the selection of appropriate wastes for suitable products generation.

## 2. Methodology

### 2.1 Description of process modelled

The high tech, advanced and reliable simulation modelling software are in research and commercial use for many polymer, chemicals, oil and gas, power, and biofuel sectors. (Shahbaz et al., 2021). Aspen plus is one of the most important simulation modelling software that can perform high-level simulations for solid, fluids, and gases by providing suitable fluid packages and properties. Hydrothermal liquefaction process is developed using Aspen Plus® V10 to yield the biocrude and hydrochar as illustrated in the process flow diagram presented in Figure 1. To perform a steady state smooth simulation process, some important assumptions were considered, including no temperature and pressure gradient and carbon is graphite carbon (Alherbawi et al., 2022). The feedstock including

EFB, PKS, POF, and their blend are introduced into the reactor as biomass slurry where non-conventional components are converted into conventional elements using Fortran routine based on the composition presented in Table. 1 as presented in previous study (Parthasarathy et al., 2022). Fortran routine is also employed to calculate the quantity of water to make a 20% solid concentrated slurry. The pressure of the slurry is raised up to 100 bar and heated up to 350 °C before being fed into the HTL reactor for the hydrothermal liquefaction process. The maximum biocrude and minimum hydrochar yields are restricted using a Fortran code calculation based on previous prediction models (Alherbawi et al., 2021).

Table 1. Composition of palm waste (PKS, EFB, POF and blends).

	PKS	EFB	POF	Blend
<b>Proximate analysis</b>				
Moisture	9.70	5.30	5.30	6.77
FC	14.42	20.61	17.74	17.59
VM	80.57	76.17	74.97	77.24
Ash	5.00	3.22	7.29	5.17
<b>Ultimate analysis</b>				
C	46.44	43.70	45.03	45.06
H	5.43	4.74	4.92	5.03
N	0.96	0.76	0.59	0.77
Cl	-	-	-	-
S	-	-	-	-
O	42.17	47.59	42.17	43.98

The HTL reactor is set to operate at Gibbs free energy minimization approach in which all possible products and biocrude components have been defined. Hydrochar is separated from the product stream in a hydro cyclone unit as depicted in flow diagram. The solid free product stream is directed to the three-stage flash where biocrude (liquid) and gases are separated along with aqueous phase which contain other dissolved organics.

### 2.2. Cost analysis

A techno economic analysis is performed using the Aspen Plus Economic Analyzer and based on the correlation and values presented in previous study (Parthasarathy et al., 2022). The levelized cost of biocrude and hydrochar per kg feed of PKS, EFB, POF, and their blend is calculated based on the levelized cost of energy concept.

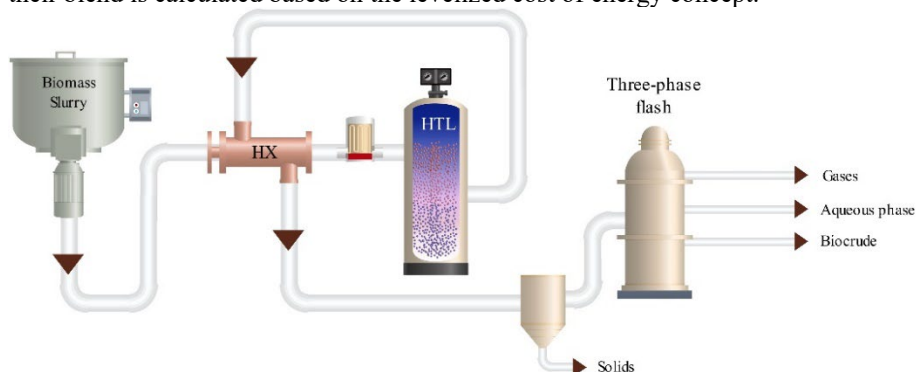


Fig.1 Process flow diagram of hydrothermal liquefaction of (PKS, EFB.POF. and Blend).



### 3. Results and discussion

#### 3.1 Product yield distribution and gas composition

The product yield of biocrude, hydrochar and gas per tonne of all feedstocks (PKS, EFB, POF and Blend) are presented in Figure 2. It can be noticed that biocrude is the dominant product for all cases except EFB. The higher yield of biocrude is due to the higher degradation rate for the cracking of the matrix structure of lignocellulosic biomass at higher temperature of 500 °C (Chan et al., 2015). PKS has shown the highest biocrude yield of 441.56 kg/tonne. POF and the blend yielded biocrude at 359.69 and 356.02 kg/tonne, respectively. Chen et al. (2018) have also observed the increase of bio oil yield with the increase in temperature for the liquefaction of PKS. A similar trend is also reported for bio-oil production from the liquefaction of giant fennel with the increase in temperature (Aysu & Küçük, 2013). Hydrochar production is found to be the highest in the case of EFB at about 192.95 kg/tonne. The higher hydrochar production for EFB is due to higher fixed carbon content of 20% as presented in Table 1. POF was associated to hydrochar production of 159.66 kg/tonne, which is the second highest amongst all feedstocks. PKS and the blend of all palm waste have a significant lower yield of 100.73 and 101.25 kg/tonne, respectively. Although liquefaction is not known for hydrochar production, it is obtained as a by-product. In this study, the yield of hydrochar is still significant, contributing to the process's commercial viability. Gas production is an important by-product that can be obtained significantly, but its usage is limited due to the domination of CO<sub>2</sub> in gas yield. The gas production for EFB, PKS, and POF are 316.48, 236.14, and 259.85 kg/tonne, respectively. The blend of feedstocks generated highest quantity of gas at 387.48 kg/tonne, which is the highest within the products array.

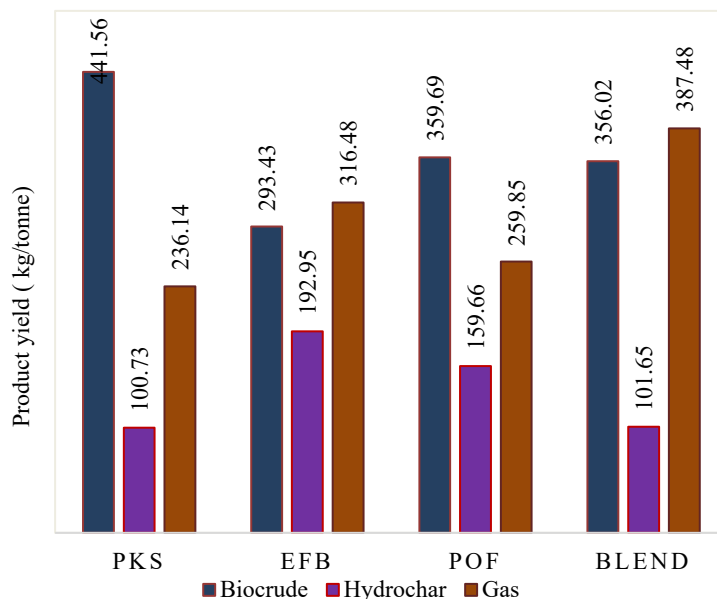


Figure 2. Product yield at temperature of 500 °C for PKS, EFB, POF and Blend.

Figure 3 demonstrates the composition of the syngas attained from liquefaction of PKS, EFB, POF and their blend. It is very clear from Figure 3 that CO<sub>2</sub> is the dominant gas in the array of gasses present in syngas. The CO<sub>2</sub> content is more than 90% in syngas for all feedstocks, being the highest in PKS case at about 98.14 %. However, the palm waste blend has yielded the lowest CO<sub>2</sub> composition at 90.14%. CH<sub>4</sub> content is the second highest component which is in the range of 1 to 10% in which the blend has also shown the highest syngas composition of CH<sub>4</sub> at 9.73%. The content of H<sub>2</sub> and CO is very low (less than 0.1%) which makes it insignificant. Due to this reason, the composition of syngas and their yield for liquefaction is very rarely reported in the literature, as such gas composition is incomparable to gasification but may be compared with hydro-pyrolysis.

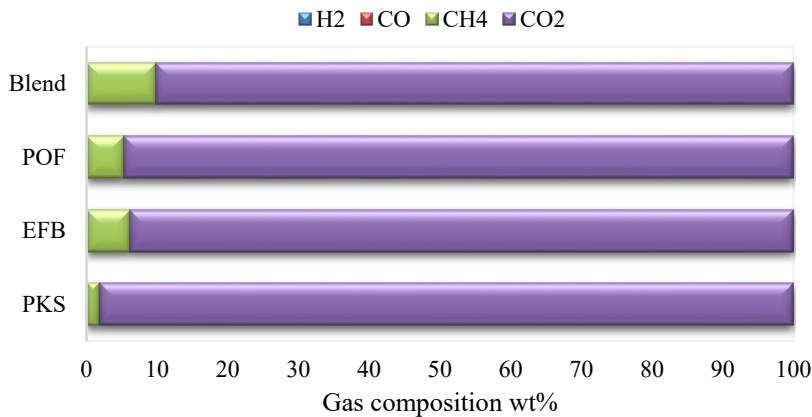


Figure 3. Gas composition at a temperature of 500 °C for PKS, EFB, POF and Blend.

### 3.2 Cost analysis

The cost of biocrude and hydrochar production per kg of PKS, EFB, POF and their blend is presented in Figure 4. The highest cost of biocrude and hydrochar are found for EFB which are at 0.402 USD/kg and 0.203 USD/kg, repeatedly. Whereas the second highest cost is noticed for palm oil waste blend. For POF, the cost of biocrude and hydrochar production is estimated at 0.391 USD/kg and 0.045 USD/kg, respectively. In the case of PKS, the cost of production for biocrude is achieved at 0.345 USD/kg, whereas hydrochar is obtained as a complementary product. It can be concluded that the POF can be a good source in terms of cost for biocrude and hydrochar production, which is linked to lower cost of feedstock.

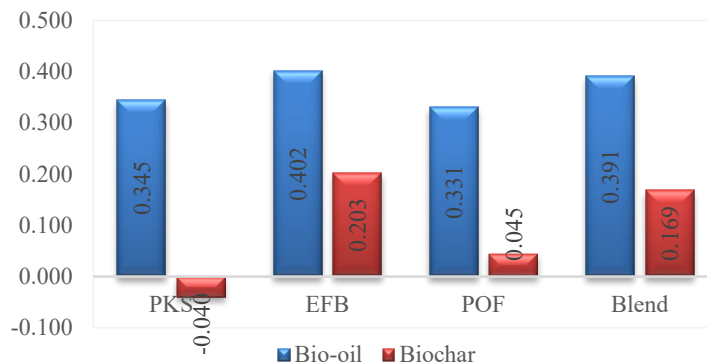


Figure 4. Cost of Biocrude and Hydrochar at 500 °C using PKS, EFB, POF and Blend.

#### 4. Conclusion:

A process simulation model of Hydrothermal Liquefaction (HTL) of PKS, EFB, POF and their blend is instituted for biocrude, hydrochar, and gas production using Aspen Plus® V10. The highest production yield of biocrude is achieved at 441.56kg/tonne using PKS, with hydrochar and gas yields at 100.73 kg/tonne and 236.14 kg/tonne, respectively. Whereas the lowest yield of biocrude is attained at 293.43 kg/tonne using EFB. The blend of all three palm wastes is found to be a good option for enhanced composition and yield of gas. Where the blended feedstock yielded 387.48 kg/tonne of gas with the lowest CO<sub>2</sub> composition. While it yielded biocrude and hydrochar at 356 kg/tonne and 101.65 kg/tonne, respectively. The gas crude composition is dominated by the CO<sub>2</sub> Content which is within the range of 90 to 98% for all feedstocks. The economic analysis indicated that the lowest biocrude cost is attained using POF at 0.331 USD/kg, while PKS was associated to the highest biocrude cost at 0.345 USD/kg.

**Acknowledgment:** The authors thank Hamad Bin Khalifa University (HBKU) for providing financial and technical support.

#### References

- M. Alherbawi, T. Al-Ansari, H.R. Mackey, G. McKay, 2021, A technoeconomic assessment of an on-site biocrude production from sewage sludge in Qatar's wastewater treatment plants. *Computer Aided Chemical Engineering*, 50, 1929-1935.
- M. Alherbawi, P. Parthasarathy, G. McKay, H.R. Mackey, T. Al-Ansari, 2022, Investigation of optimal blending of livestock manures to produce biocrude via hydrothermal liquefaction. In *Computer Aided Chemical Engineering* (Vol. 51, pp. 1243-1248): Elsevier.
- A.M. Ali, M. Inayat, A.A. Zahrani, K. Shahzad, M. Shahbaz, S.A. Sulaiman, H. Sadig, 2022, Process optimization and economic evaluation of air gasification of Saudi Arabian date palm fronds for H<sub>2</sub>-rich syngas using response surface methodology. *Fuel*, 316, 123359.
- A. AlNouss, M. Shahbaz, G. McKay, T. Al-Ansari, 2022, Bio-methanol production from palm wastes steam gasification with application of CaO for CO<sub>2</sub> capture: techno-economic-environmental analysis. *Journal of Cleaner Production*, 341, 130849.
- T. Aysu, M.M. Küçük, 2013, Liquefaction of giant fennel (*Ferula orientalis* L.) in supercritical organic solvents: Effects of liquefaction parameters on product yields and character. *The Journal of Supercritical Fluids*, 83, 104-123.
- Y. Chan, K.W. Cheah, B. How, A. Loy, M. Shahbaz, H. Singh, A. Shuhaili, S. Yusup, W. Ghani, J. Rambli, 2019, An overview of biomass thermochemical conversion technologies in Malaysia. *Science of The Total Environment*, 680, 105-123.

- Y. Chan, A. Quitain, S. Yusup, Y. Uemura, M. Sasaki, T. Kida, 2018, Optimization of hydrothermal liquefaction of palm kernel shell and consideration of supercritical carbon dioxide mediation effect. *The Journal of Supercritical Fluids*, 133, 640-646.
- Y. Chan, S. Yusup, A. Quitain, R. Tan, M. Sasaki, H. Lam, Y. Uemura, 2015, Effect of process parameters on hydrothermal liquefaction of oil palm biomass for bio-oil production and its life cycle assessment. *Energy Conversion and Management*, 104, 180-188.
- P. Parthasarathy, M. Alherbawi, M. Shahbaz, H.R. Mackey, G. McKay, T. Al-Ansari, 2022, Conversion of oil palm waste into value-added products through pyrolysis: a sensitivity and techno-economic investigation. *Biomass Conversion and Biorefinery*.
- M. Shahbaz, A. AlNouss, S. Yusup, G. McKay, T. Ansari, 2021, Techno-economic evaluation of sorption enhanced steam gasification of PKS system for syngas using CaO for CO<sub>2</sub> capture. In M. Türkyay & R. Gani (Eds.), *Computer Aided Chemical Engineering* (Vol. 50, pp. 129-134): Elsevier.
- H. Umar, S. Sulaiman, M. Meor Said, A. Gungor, M. Shahbaz, M. Inayat, R. Ahmad, 2021, Assessing the implementation levels of oil palm waste conversion methods in Malaysia and the challenges of commercialisation: Towards sustainable energy production. *Biomass and Bioenergy*, 151, 106179.



# Heuristic approaches to solving multistage stochastic programs with type II endogenous uncertainty

Yasuhiro Shoji,<sup>a</sup> Selen Cremaschi<sup>\*</sup>,<sup>a</sup>

<sup>a</sup>*Department of Chem. Eng., Auburn University, Auburn, AL 36849, United States  
szc0113@auburn.edu*

## Abstract

Many optimization problems in the process industry, such as production planning and scheduling, contain endogenous (decision-dependent) uncertainties. The endogenous uncertainties can be represented as scenarios based on the realized outcomes; hence, by nature, multistage stochastic programming (MSSP) is one approach to modeling and solving these problems. MSSP problems quickly grow and become computationally intractable as the problem size increases. We solve four MSSP problems with Type II endogenous uncertainty using absolute expected value solution (AEEV) and generalized Knapsack-problem-based decomposition algorithm (GKDA) to address the computational tractability issues. We bound the solutions using a modified Lagrangian relaxation (mLR). The results reveal that AEEV provides the smallest optimality gap.

**Keywords:** multistage stochastic program, endogenous uncertainty, heuristic approaches

## 1. Introduction

MSSP is a mathematical decision-making approach used in the process industry, such as for process synthesis (Tarhan and Grossmann, 2008), to solve optimization problems with the sequential realization of uncertainties. Uncertainty in optimization problems is classified as exogenous or endogenous, which can be further classified into Type I: decisions affect the probability distribution, and Type II: decisions affect the realization time (Goel and Grossmann, 2006). For MSSPs with type II endogenous uncertainty, implicitly enforcing non-anticipativity (NA) in the model formulation is not possible. Hence, NA constraints (NACs) are added to the model to prevent decisions from anticipating unrealized future outcomes. However, the number of NACs grows exponentially as the number of scenarios increases, resulting in computational intractability (Apap and Grossmann, 2017). Various heuristic, approximation, and decomposition methods exist for solving MSSPs with endogenous uncertainty, e.g., sample average approximation algorithm (Solak et al., 2010), improved Lagrangian decomposition framework (Gupta and Grossmann, 2014), and rolling-horizon heuristic approach (Colvin and Maravelias, 2009).

This paper applies two heuristic primal bounding approaches, AEEV (Zeng and Cremaschi, 2019) and GKDA (Zeng et al., 2018), to solve MSSP problems with Type II endogenous uncertainty and bound the solutions using mLR (Zeng and Cremaschi, 2020). It has been suggested that AEEV and GKDA yield tight primal bounds quickly, and the goal is to investigate their performance for a wider range of MSSP problems.

### 2. A general MSSP formulation under endogenous uncertainties

A general MSSP formulation under endogenous uncertainties is given in Eqs. 1-6.

$$OV = \min \sum_s p_s \sum_{i,t} G_{i,t,s}(V_{i,t}, \theta_i^s, b_{i,t}^s, \gamma_t^s) \tag{1}$$

$$g_{i,t,s}(b_{i,t}^s, \gamma_t^s, \theta_i^s) \leq 0 \quad \forall i \in I, t \in T, s \in S \tag{2}$$

$$h_{i,t,s}(b_{i,t}^s, \gamma_t^s, \theta_i^s) = 0 \quad \forall i \in I, t \in T, s \in S \tag{3}$$

$$b_{i,1}^s = b_{i,1}^{s'} \quad \forall i \in I, \forall (s, s') \in S \tag{4}$$

$$\begin{bmatrix} Z_{i,t}^{s,s'} \\ b_{i,t}^s = b_{i,t}^{s'} \end{bmatrix} \vee \begin{bmatrix} -Z_{i,t}^{s,s'} \end{bmatrix} \quad \forall (i, s, s') \in S_E, \forall t \in T, t > 1 \tag{5}$$

$$\gamma_t^s \in \mathbb{R}, b_{i,t}^s, Z_{i,t}^{s,s'} \in \{0,1\} \quad \forall (i, s, s') \in S_E, \forall t \in T, \forall i \in I \tag{6}$$

The objective function (Eq. (1)) calculates the optimum OV. The model has two decision variables,  $b_{i,t}^s$  and  $\gamma_t^s$ , with uncertain resources  $i \in I = \{1,2, \dots, I\}$ , planning horizon  $t \in T = \{1,2, \dots, T\}$ , and scenarios  $s \in S = \{1,2, \dots, S\}$ . The binary decision variable  $b_{i,t}^s$  is enforced to be identical by either initial (Eq. (4)) or conditional NACs (Eq. (5)) until scenario pair  $(s, s')$  becomes distinguishable. The decision variable  $\gamma_t^s$  is determined by scenario-specific constraints (Eqs. (2)-(3)).  $V_{i,t}$  and  $\theta_i^s$  represent deterministic and endogenous uncertain parameters. The function  $G_{i,t,s}(V_{i,t}, \theta_i^s, b_{i,t}^s, \gamma_t^s)$  calculates the contribution of a specific scenario  $s$  with probability  $p_s$  to the objective. The conditional NACs (Eq. (5)) are enforced depending on the values of the Boolean variable  $Z_{i,t}^{s,s'}$ , which is equal to one if scenario pair  $(s, s')$  is indistinguishable, and zero otherwise.

### 3. Solution approaches

#### 3.1. Absolute Expected Value Solution Approach (AEEV)

AEEV (Zeng and Cremaschi, 2019) is a primal bounding approach that converts an MSSP problem into two sub-problems, deterministic expected value sub-problems (*DEVSPs*) and recourse deterministic expected value sub-problems (*DEVSPs<sup>recourse</sup>*), by removing all NACs (Eq. (4), (5)) and indices  $s$  from the MSSP formulation. *DEVSPs* and *DEVSPs<sup>recourse</sup>* determine here-and-now decisions and corresponding recourse actions. AEEV starts by constructing a *DEVSP* at the initial time, as shown in Figure 1.

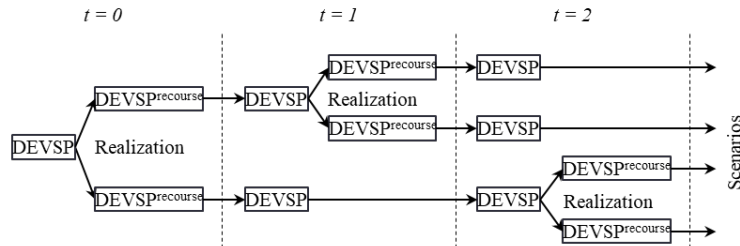


Figure 1 The schematic diagram of AEEV

The values of endogenous uncertain parameters are set to their expected values over all scenarios because there are no realizations before solving the initial *DEVSP*. AEEV stores the here-and-now decisions obtained from the *DEVSP* solutions and determines which scenarios can be differentiated based on the realized outcomes. If realizations occur and

the MSSP problem has decision variables associated with recourse actions,  $DEVSP_s^{\text{recourse}}$  is generated and solved to obtain these decision variables. If there is no realization, AEEV skips constructing  $DEVSP_s^{\text{recourse}}$ . AEEV continues to construct and solve  $DEVSP_s$  and  $DEVSP_s^{\text{recourse}}$  for the subsequent time periods with stored decisions set to solutions of prior  $DEVSP_s$  and  $DEVSP_s^{\text{recourse}}$  until the end of the time horizon. The values of endogenous uncertain parameters are expected-valued over possible scenarios at each  $DEVSP$  and  $DEVSP^{\text{recourse}}$ . At termination, AEEV provides a feasible solution and an upper bound for the MSSP model.

### 3.2. Generalized Knapsack-Problem based Decomposition Algorithm (GKDA)

GKDA (Zeng et al., 2018) is a similar primal bounding approach. It translates the MSSP problem into two types of knapsack sub-problems (KSPs) by decomposing scenario and time indices, i.e., it removes all NACs (Eq. (4), (5)) and indices  $s$  and  $t$  from the MSSP formulation. The remaining index  $i$  is translated into an item set of KSPs by enumerating all eligible decisions. Scenario-specific constraints (Eq. (2), (3)), except sequencing constraints, are transformed into KSP weight constraints. Sequencing constraints and bounds are used to determine the allowable items in a KSP from the eligible items. Item values are generally objective function coefficients, i.e.,  $V_{i,t}$  in Eq (1) translates into  $V_e$  as item values in KSPs, where  $e$  notates item  $e \in E$ . Under certain conditions, this translation does not accurately approximate the expected contribution of the associated decision to the MSSP objective function. Zeng et al. (2018) defined two conditions with approaches to better approximate item values. Condition 1 applies to problems in which gains/revenues are only realized at certain stages or are only associated with certain decision variables, and the item values are generated using the maximum potential gain approach (MPGA). Condition 2 applies to problems where an item represents a capital investment decision variable in MSSP, and the item value is generated using annual capital charges. Similar to AEEV, GKDA continues to construct and solve KSPs along the planning horizon with stored decisions of prior KSPs. At termination, GKDA provides a feasible solution and an upper bound for the MSSP problem.

### 3.3. Modified Lagrangian Relaxation (mLR)

Lagrangian relaxation (LR) is a common approach to removing complicating constraints and adding them to the objective function as penalty terms (Lin et al., 2011). For MSSP problems, NACs (Eq. (4), (5)) are complicating constraints. A standard LR dualizes these NACs to the MSSP objective function, as shown in Eq. (7).

$$\begin{aligned} OV_{LR} = \min \sum_s p_s \sum_{i,t} G_{i,t,s}(V_{i,t}, \theta_i^s, b_{i,t}^s, \gamma_t^s) + \sum_{i,s,s'} \lambda 1_{i,s,s'} (b_{i,1}^s - b_{i,1}^{s'}) \\ + \sum_{i,t,s,s'} \lambda 2_{i,t,s,s'} (1 - Z_{i,t}^{s,s'} - b_{i,t}^s + b_{i,t}^{s'}) \\ + \sum_{i,t,s,s'} \lambda 3_{i,t,s,s'} (1 - Z_{i,t}^{s,s'} + b_{i,t}^s - b_{i,t}^{s'}) \end{aligned} \quad (7)$$

In Eq. (7),  $\sum_{i,s,s'} \lambda 1_{i,s,s'} (b_{i,1}^s - b_{i,1}^{s'})$  corresponds to initial NACs (Eq. (4)), and  $\sum_{i,t,s,s'} \lambda 2_{i,t,s,s'} (1 - Z_{i,t}^{s,s'} - b_{i,t}^s + b_{i,t}^{s'})$  and  $\sum_{i,t,s,s'} \lambda 3_{i,t,s,s'} (1 - Z_{i,t}^{s,s'} + b_{i,t}^s - b_{i,t}^{s'})$  correspond to conditional NACs (Eq. (5)).  $\lambda 1_{i,s,s'}$ ,  $\lambda 2_{i,t,s,s'}$ , and  $\lambda 3_{i,t,s,s'}$ , called multipliers, are updated over the iterations to obtain tighter bounds. LR provides a lower bound. For  $\sum_{i,t,s,s'} \lambda 2_{i,t,s,s'} (1 - Z_{i,t}^{s,s'} - b_{i,t}^s + b_{i,t}^{s'})$  and  $\sum_{i,t,s,s'} \lambda 3_{i,t,s,s'} (1 -$



$Z_{i,t}^{s,s'} + b_{i,t}^s - b_{i,t}^{s'}$ ) in Eq. (7), NACs are always satisfied when scenario pairs  $(s, s')$  are distinguishable, i.e.,  $Z_{i,t}^{s,s'} = 0$ , thereby leading to looser bounds. To tighten it, Zeng and Cremaschi (2020) proposed mLR by transforming Eq. (7) into Eq. (8) where the conditional NACs are dualized only when  $Z_{i,t}^{s,s'} = 1$ .

$$OV_{LR} = \min \sum_s p_s \sum_{i,t} G_{i,t,s}(V_{i,t}, \theta_i^s, b_{i,t}^s, \gamma_i^s) + \sum_{i,s,s'} \lambda_{i,s,s'} (b_{i,1}^s - b_{i,1}^{s'}) + \sum_{i,t,s,s'} \lambda_{i,t,s,s'} Z_{i,t}^{s,s'} (b_{i,t}^s - b_{i,t}^{s'}) \quad (8)$$

Multiplier updating schemes are critical for obtaining tighter bounds. The sub-gradient optimization method (Eqs. (9) and (10)) is a commonly used technique (Fisher, 1985).

$$\lambda_n^{k+1} = \max\{0, \lambda_n^k - \gamma^k(Ax_n^k - b)\} \quad (9)$$

$$\gamma^k = scale \frac{Z(\lambda_n^k) - Z^*}{\sum_n (Ax_n^k - b)^2} \quad (10)$$

The total  $|N|$  number of multipliers at  $k + 1^{th}$  iteration,  $\lambda_n^{k+1}$  ( $n \in N$ ), is calculated with Eq. (9) (Fisher, 1981), where  $Ax_n^k - b$  is the dualized constraints transferred to the objective function, where  $x_n^k$  is solved variables, and  $A$  and  $b$  are parameters of dualized constraints. The step size at iteration  $k$ ,  $\gamma^k$ , is calculated by Eq. (10) based on the gap between the best-known primal  $Z^*$  and dual bound  $Z(\lambda_n^k)$ . *scale* in Eq. (10) is an adjustable parameter in  $0 < scale < 2$ , which is often set via trial and error. To achieve tighter bounds efficiently, Zeng and Cremaschi (2020) proposed updating *scale* at the  $k^{th}$  iteration using Eq. (11) if the ratio of dualized constraints value  $D$  to the dualized objective function value  $OV_{LR}$  is greater than or equal to 0.1. The multipliers are updated using Eq. (12), where  $N^v$  is the total number of violated dualized constraints.

$$\text{If } \frac{D}{OV_{LR}} \geq 0.1 \text{ Then } scale^k = \frac{scale^{k-1}}{1.618} \quad (11)$$

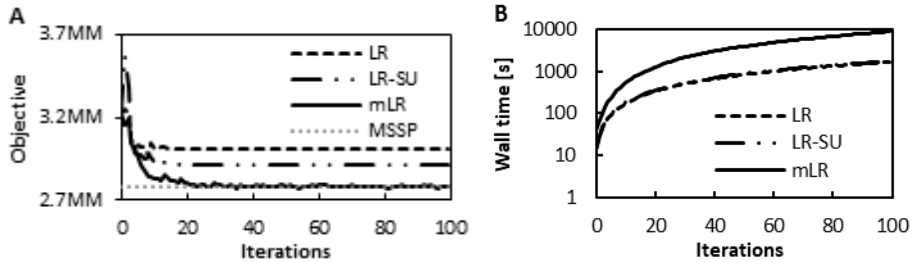
$$\gamma^k = scale^k \frac{Z(\lambda_n^k) - Z^*}{N^v} \quad (12)$$

#### 4. Case study descriptions, results, and discussion

We applied the solution approaches to four MSSP problems with Type 2 endogenous uncertainty. All problems are Mixed Integer Linear Programming (MILP) problems. Clinical Trial Planning (CTP) (Colvin and Maravelias, 2009) problem determines the timing of the clinical trials, maximizing expected profit. Artificial Lift Infrastructure Planning (ALIP) (Zeng and Cremaschi 2017) problem determines installed equipment and installation timing for exploiting shale gas, maximizing expected profit. Vehicle Routing (VR) (Hooshmand and MirHassani, 2016) problem determines customer visit order, minimizing expected cost. Open Pit Mining (OPM) (Boland et al., 2008) problem determines aggregates to be excavated and processed and their timing, maximizing expected profit. We modeled all instances in Pyomo and solved them with CPLEX 20.10 using 48 processors on a node of Auburn University Easley Cluster.

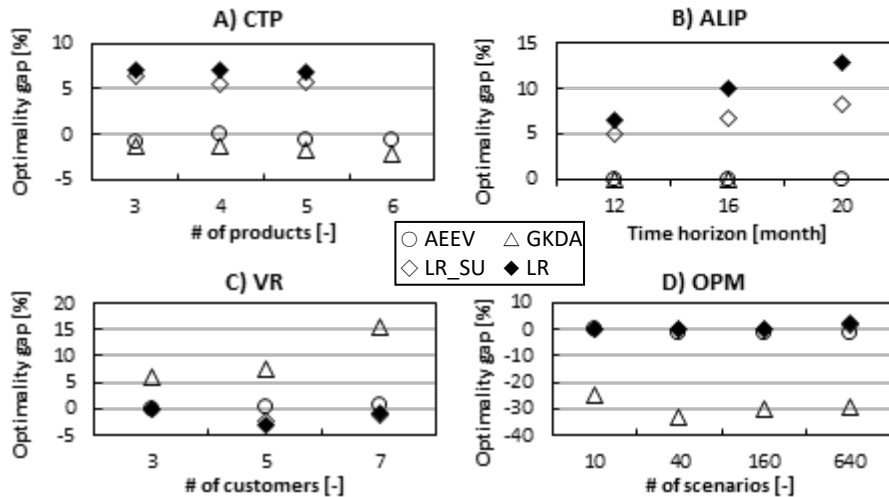
Figure 2 represents the result of mLR for an instance of ALIP with a 12-month planning horizon. mLR required significantly longer times for other instances, yielding a few to 40 iterations within 48 hours. Figure 2 A shows that mLR yields a tighter bound than LR and

LR with the new scale updating scheme only (LR-SU). However, mLR takes longer than LR and LR-SU, as can be seen in Figure 2 B. The objective function of mLR is nonlinear and cannot be decomposed if linearized due to scenario pairs  $(s, s')$  in term  $\sum_{i,t,s,s'} \lambda_{i,t,s,s'} Z_{i,t}^{s,s'} (b_{i,t}^s - b_{i,t}^{s'})$  of Eq. (8). We only applied LR-SU to other instances.



**Figure 2** A) Objective value of LR, LR-SU, mLR, and MSSP and B) Computational time of LR, LR-SU, and mLR for ALIP with a planning horizon of 12 months.

Figures 3 and 4 represent the optimality gap and computational times. LR and LR-SU results are for iteration 50. The optimality gap is defined from MSSP optimum solution. AEEV provided the best feasible solutions, and LR-SU provided tighter bounds than LR. GKDA yielded better feasible solutions for CTP and ALIP than VR and OPM because it employed Conditions 1 and 2 for item value estimation for CTP and ALIP, respectively. These results indicate the importance of item value estimation for GKDA, which is the fastest (Figure 4). LR-SU takes almost the same time as LR but yields tighter bounds. Computational time increased milder for all approaches than for MSSP.



**Figure 3** Optimality gap of solutions yielded by approaches from the MSSP solution

### 5. Conclusion and Future Directions

We applied two heuristics, AEEV, GKDA, and two bounding approaches, mLR and LR-SU, to four MSSP problems with Type II endogenous uncertainty. The results revealed AEEV provides the tightest feasible solutions. The bound yielded with mLR was tighter than LR; however, mLR required a long computational time. Computational time increased milder for all approaches than for MSSP. We plan to apply the approaches to other MSSP problems and study avenues for improving them.

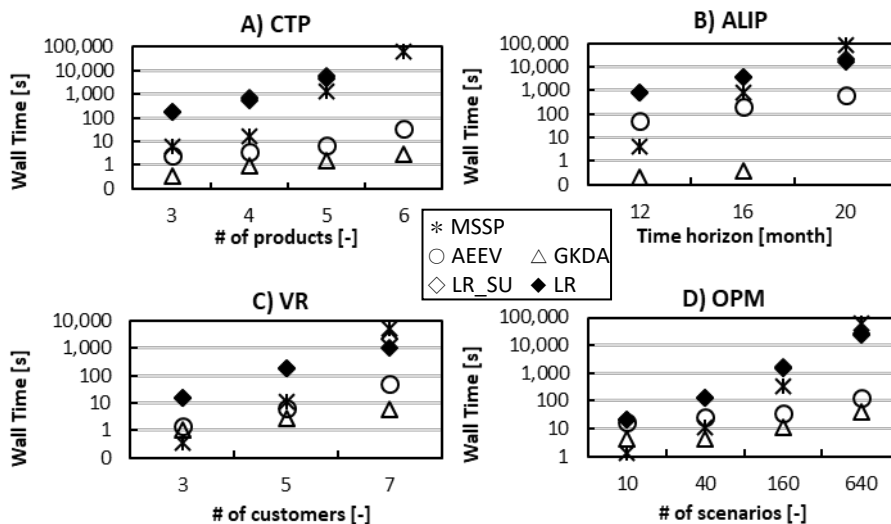


Figure 4 Computational time of MSSP and solution approaches

## References

- Apap, R.M. and Grossmann, I.E. (2017), Models and computational strategies for multistage stochastic programming under endogenous and exogenous uncertainties. *Comput. Chem.Eng.*, 103; pp.233-274
- Boland, N., Dumitrescu, I., & Froyland, G. (2008), A multistage stochastic programming approach to open pit mine production scheduling with uncertain geology. *Optimization online*, 1-33.
- Colvin, M., & Maravelias, C. T. (2009), Scheduling of testing tasks and resource planning in new product development using stochastic programming. *Comput. & Chem. Eng.*, 33(5), 964-976.
- Fisher M. L. (1981), The Lagrangian Relaxation Method for Solving Integer Programming Problems. *Management Science.*, 27(1)
- Fisher M. L. (1985), An applications oriented guide to Lagrangian relaxation. *Interfaces* 15, 10-21
- Hooshmand Khaligh, F., & MirHassani, S. A. (2016), A mathematical model for vehicle routing problem under endogenous uncertainty. *Int. J. Prod. Res.*, 54(2), 579-590.
- Goel, V., & Grossmann, I. E. (2006), A class of stochastic programs with decision dependent. uncertainty *Math. Program.*, 108(2-3), 355-394.
- Gupta, V., & Grossmann, I. E. (2014), Multistage stochastic programming approach for offshore oilfield infrastructure planning under production sharing agreements and endogenous uncertainties. *Journal of Petroleum Science and Engineering*, 124, 180-197.
- Lin, Z., Liu, R., & Su, Z. (2011), Linearized Alternating Direction Method with Adaptive Penalty for Low-Rank Representation. In *Advances in neural information processing systems* (pp. 612-620).
- Solak, S., Clarke, J. P. B., Johnson, E. L., & Barnes, E. R. (2010), Optimization of R&D project portfolios under endogenous uncertainty. *European Journal of Operational Research*, 207(1), 420-433.
- Tarhan, B. and I.E. Grossmann. (2008), A multistage stochastic programming approach with strategies for uncertainty reduction in the synthesis of process networks with uncertain yields. *Comput. & Chem. Eng.*, 32(4-5), 766-788.
- Zeng, Z., Christian, B., & Cremaschi, S. (2018), A generalized knapsack-problem based decomposition heuristic for solving multistage stochastic programs with endogenous and/or exogenous uncertainties. *Ind. Eng. Chem. Res.*, 57(28), 9185-9199.
- Zeng, Z., & Cremaschi, S. (2017), Artificial lift infrastructure planning for shale gas producing horizontal wells. *Proceedings of the FOCAPO/CPC*, Tuscan, AZ, USA, 8-12.
- Zeng, Z., & Cremaschi, S. (2019), A general primal bounding framework for large-scale multistage stochastic programs under endogenous uncertainties. *Chem. Eng. Res. Des.*: 141, 464-480.
- Zeng, Z., & Cremaschi, S. (2020), A new lagrangean relaxation approach for multistage stochastic programs under endogenous uncertainties. *30th ESCAPE*, 47.

# Hybrid modeling of the catalytic CO<sub>2</sub> methanation using process data and process knowledge

Luisa Peterson<sup>a</sup>, Jens Bremer<sup>b</sup>, Kai Sundmacher<sup>a,c</sup>

<sup>a</sup>Max Planck Institute for Dynamics of Complex Technical Systems, Process Systems Engineering, Sandorstraße 1, Magdeburg, 39106, Germany

<sup>b</sup>Clausthal University of Technology, Institute of Chemical and Electrochemical Process Engineering, Leibnizstraße 17, Clausthal-Zellerfeld, 38678, Germany

<sup>c</sup>Otto von Guericke University Magdeburg, Chair for Process Systems Engineering, Universitätsplatz 2, Magdeburg, 39106, Germany

## Abstract

Heterogeneously catalyzed CO<sub>2</sub> methanation is an essential process in the power-to-methane concept. To overcome the limitations of both mechanistic and data-driven models, this study presents three modeling approaches for predicting the CO<sub>2</sub> conversion using different techniques for integrating pseudo-experimental process data into process models. One approach is a data-driven model that uses multilayer perceptron regression (MLPR), a type of feedforward artificial neural network, to represent the data. Another approach is a hybrid model that integrates MLPR to estimate the reaction rate within a mechanistic model. The third approach is another hybrid model that uses MLPR to correct for a simplified mechanistic model that does not fully capture the physical behavior of the system. All models are evaluated in terms of accuracy, complexity, and reliability, and the results show that they all predict the conversion of CO<sub>2</sub> with only minor deviations. Hybrid models have the advantage of combining the strengths of mechanistic and data-driven models and, with basic process knowledge, can make accurate predictions beyond the training data. However, they also require both process knowledge and large amounts of process data.

**Keywords:** Power-to-X, Reactor Modeling, Hybrid Modeling, Machine Learning

## 1. Introduction

The intermittent nature of solar and wind power underscores the need for effective energy storage technologies. Power-to-methane (PtM) is a process that uses electrolysis to convert energy into hydrogen, which is then chemically bonded with carbon to form synthetic methane. This CO<sub>2</sub>-neutral fuel can be used in households and industries through existing distribution infrastructure (Götz, et al., 2016). However, the low efficiency, complexity, and high cost of the PtM process pose challenges to its widespread adoption (Er-rbib & Bouallou, 2013). Optimization of electrolyzer, reactor, and downstream process operation and design can address these challenges. Accurate process models are required to improve the performance of PtM systems. Traditionally, process models are established based on physical and chemical principles, but building first-principle models requires extensive data from specific laboratory experiments and is time-consuming (Zimmermann, et al., 2020). In addition, quantifying the kinetics that depend on elementary reactions and intermediates is challenging using traditional methods (Zahedi, et al., 2005). Data-driven regression models provide an alternative to mechanistic models by identifying relationships between the dependent variable and one or more independent inputs (Kahrs & Marquardt, 2007). They can be generated quickly without prior process knowledge, but

require large amounts of data and may have limited physical interpretation. Hybrid models, which combine mechanistic models with data-driven models, can describe unknown phenomena by integrating data-driven approaches into mass and energy balances or by compensating for discrepancies between the model and the plant (Von Stosch, et al., 2014). Multilevel perceptron regression (MLPR), a type of feedforward artificial neural network, represents all data-driven relationships. In MLPR, data flows from the input layer through the hidden layers to the output layer, where the hidden layers weights the inputs and pass them through a nonlinear activation function to the output. The MLPR is trained using backpropagation by describing the target with input representations. With the advancement of data-driven models, the number of parameters, runtime, and computational resources have increased significantly. Thus, it is imperative to consider metrics such as accuracy, complexity, and reliability in addition to the accuracy of the model (Menghani, 2021). In this study, the models are evaluated based on these metrics.

## 2. Methodology

The modeling approach involves generating pseudo-experimental process data using a mechanistic CSTR model for the CO<sub>2</sub> methanation and predicting this data using three models: one data-driven model and two hybrid models.

### 2.1. Data Generation via CSTR Model

The mass and energy balances for modeling a CSTR are based on the work of Bremer and Sundmacher (Bremer & Sundmacher, 2021). The balances were constructed under the assumptions that (1) the CO<sub>2</sub> methanation is a single reaction, that (2) it occurs in a single phase, and that (3) the pressure  $p$ , the gas heat capacity  $c_p$ , the heat of reaction  $\Delta_R \bar{H}$  and the effectiveness  $\eta$  for limiting catalyst mass transport are constant. Using the conversion of CO<sub>2</sub> ( $X_{CO_2}$ ), the residence time  $\tau$ , the void fraction  $\epsilon$ , the inlet concentration of CO<sub>2</sub>  $c_{CO_2}$ , and the mass flow fraction of the each component  $\omega_i$ , the steady state mass balance of the CSTR is described as

$$0 = X_{CO_2} - \frac{\tau R_{\text{meth}}(T, p, \omega_a)}{\epsilon c_{CO_2, \text{in}}} \quad (1)$$

The intrinsic reaction rate,  $R_{\text{meth}}$ , is a LHHW-type equation proposed by Koschany et al. (Koschany, et al., 2016).  $R_{\text{meth}}$  depends on the reaction temperature  $T$  and the partial pressures  $p_i$  of the components.

$$R_{\text{meth}}^{\text{LHHW}} = kp_{CO_2}^{0.5} p_{H_2}^{0.5} \left( 1 - \frac{p_{CH_4} p_{H_2O}^2}{K_{eq} p_{CO_2} p_{H_2}^4} \right) / \left( 1 + K_{OH} + \frac{p_{H_2O}}{p_{H_2}^{0.5}} + K_{H_2} p_{H_2}^{0.5} + K_{\text{mix}} p_{CO_2}^{0.5} \right) \quad (2)$$

The state variables for the energy balance of the CSTR are  $T$  and  $X_{CO_2}$ , along with the Stanton number  $St$ , the adiabatic temperature rise  $\Delta T_{\text{ad}}$ , and the operation temperature  $T_{\text{op}}$ . The energy balance is solved in steady state.

$$0 = X_{CO_2} - \frac{(1 + St)}{\Delta T_{\text{ad}}} (T - T_{\text{op}}) \quad (3)$$

The steady-state mass and energy balances are solved simultaneously for  $X_{CO_2}$  and  $T_{\text{op}}$ . To obtain the training data, 1000 different operating conditions are selecting using statistical Latin Hypercube Sampling. The sample space includes pressure ( $1 \text{ bar} < p < 10 \text{ bar}$ ), inlet temperature ( $400 \text{ K} < T_{\text{in}} < 800 \text{ K}$ ), residence time ( $0.04 \text{ s} < \tau < 0.4 \text{ s}$ ),

and the ratio of products in the feed ( $0 < x_i < 0.5$ ). The test set is created by selecting 100 evenly distributed data points across the operating temperature range.

## 2.2. Training Strategy

The pipeline used in the study combines data processing and model training. The process data are first scaled using a min-max scaler, ensuring that all representations and target variable of the MLPR were between 0 and 1. In the next step, the NAdam optimization algorithm is used to minimize the mean squared error (MSE) of the CO<sub>2</sub> predictions by training a multi-layer perceptron regression (MLPR) model using the PyTorch library. The SELU activation function is used to compute the output of the hidden layer. The MLPR hyperparameters are determined using a Bayesian optimizer that used Gaussian processes, including the number of hidden layers (1 or 2), the number of neurons per hidden layer (between 5 and 30), the batch size for the optimizer (8, 16, 32, or 64), and the learning rate (1e-2 or 1e-3). The set of hyperparameters that produced the lowest average MSE on the test set when trained with five different sets of initial weights for the MLPR parameters is selected.

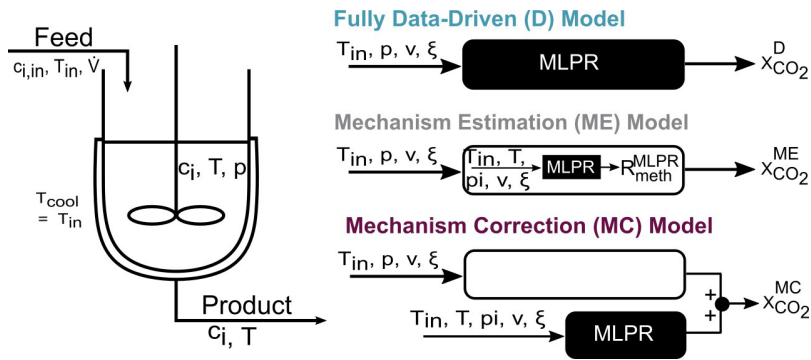


Figure 1: Illustrated schematic depicts CSTR (left) and hybrid modeling approach (right), with white boxes representing first-principle models and black boxes representing MLPR models.

In total, three different models with a data-driven component are set up. Figure 1 shows the models schematically. The fully data-driven (D) model represents mass and energy balances completely with a MLPR.  $X_{CO_2}^D$  is a function of the input representations (*repr*) namely,  $p$ ,  $T_{in}$ ,  $\tau$ ,  $\xi$  as well as the MLPR weights ( $\phi$ ) optimized during MLPR training, and can be expressed as:

$$X_{CO_2}^D = MLPR(repr, \phi). \quad (4)$$

In the Mechanism Estimation (ME) model, the mechanistic part employs the mass and energy balances, while  $R_{meth}$  is represented by an MLPR. The input representations are  $p_i$ ,  $T_{in}$ ,  $T$ ,  $\tau$ , and  $x_i$ . For the model,  $X_{CO_2}^{ME}$  is calculated from the mechanistic arguments (*args*), mechanistic parameters ( $\theta$ ), and  $R_{meth}^{MLPR} = MLPR(repr, \phi)$

$$X_{CO_2}^{ME} = f(args, \theta, MLPR(repr, \phi)). \quad (5)$$

Minimizing the MSE of the  $X_{CO_2}$  predictions determines  $R_{meth}$ . The third model, the Mechanism Correction (MC) model, utilizes a correction term to compensate for a mechanistic model that does not fully capture system behavior. The simplified mechanistic

model calculates  $X_{\text{CO}_2}^{\text{PL}}$  using a mechanistic model with a fitted power law  $R_{\text{meth}}^{\text{PL}}$  as the reaction rate. MLPR serves as the correction term  $\kappa = \text{MLPR}(\text{repr}, \phi)$ , learning the difference between  $X_{\text{CO}_2}^{\text{PL}}$  and  $X_{\text{CO}_2}$  using the representations  $p_i$ ,  $T_{\text{in}}$ ,  $T$ ,  $\tau$ , and  $x_i$ .  $X_{\text{CO}_2}^{\text{MC}}$  is calculated by summing  $X_{\text{CO}_2}^{\text{PL}}$  and  $\kappa$ :

$$X_{\text{CO}_2}^{\text{MC}} = f(\text{args}, \theta, R_{\text{meth}}^{\text{PL}}) + \text{MLPR}(\text{repr}, \phi) \quad \text{with} \quad R_{\text{meth}}^{\text{PL}} = k p_{\text{H}_2}^{n_{\text{H}_2}} p_{\text{CO}_2}^{n_{\text{CO}_2}} \left( 1 - \frac{p_{\text{CH}_4} p_{\text{H}_2\text{O}}^2}{K_{\text{eq}} p_{\text{CO}_2} p_{\text{H}_2}^4} \right). \quad (6)$$

### 2.3. Model Evaluation

The accuracy of the regression is determined by the MSE. The MSE depends on the random initialization of the MLPR parameter weights, so the model is trained 101 times with different initial weights to determine the distribution of the MSE. To ensure reproducibility, a seed is assigned to each set of initial weights, and the model with the mean MSE value among the test models is being further evaluated. Model complexity is defined by the number of parameters in the MLPR and the average CPU time required to fit the training set and predict the test set in 101 runs on a single node of a computing cluster. In addition, the performance of the model under varying amounts of test data is being examined. Fitted models are used to predict  $X_{\text{CO}_2}$  from unknown inlet temperature distributions ( $800 \text{ K} < T < 900 \text{ K}$ ) to assess extrapolation capability. Both training and test data are subjected to 5 % noise to test robustness, and the effect of noise on the predicted  $X_{\text{CO}_2}$  data is analyzed.

## 3. Results and Discussion

Three models are trained on pseudo-experimental process data as described in Section 2.2. Table 1 shows the hyperparameters, average MSEs, and average fitting times for the test and training set of all models. The left part of Figure 2 exhibits the distribution of the MSE over 101 model runs with different MLPR initial values. The results indicate that the MC model achieves the lowest average MSE on the test set, followed by the D model and the ME model. The trained models are applied to pseudo-experimental process data collected at various inlet temperatures with all other variables held constant. Figure 3 demonstrates the prediction of  $X_{\text{CO}_2}$  versus the operating temperature for each model. Despite its second-best performance in the training and test sets, the D model fails to correctly predict the pseudo-experimental temperature curves at small temperatures and residence times, due to underrepresentation of certain operating temperatures in the training set. The D model appears to overfit rather than interpolate the underrepresented data. In contrast, the ME model and MC model accurately describe the temperature curves.

The complexity of the models is determined by the MLPR architecture and the average training and prediction time (refer to Table 1). The D model trains faster. The time required to fit the data correlates with the architecture of the MLPs. Faster fitting times can be achieved by using fewer nodes per layer, larger batches, and lower learning rates. Although the ME model has the same architecture as the other models, it requires more time to train because the reaction rate is calculated using MLPR on  $X_{\text{CO}_2}$  as the target, resulting in more computational steps. The optimization of the models in this work is focused solely on accuracy. Reducing complexity by finding a Pareto front between accuracy and complexity is possible, but not the focus of this study.

Table 1: Selected hyperparameters, average MSEs, and average run times of training and test data for each model.

Model	Hyperparameters			Training		Test			
	Nodes per layer	Learn rate	Batch size	$\overline{\text{MSE}}$	Time in s	$\overline{\text{MSE}}$	Time in s	MSE ex-trap.	MSE 5% noise
D	4-25-25-1	1e-3	16	5.32e-5	61.8	4.62e-5	0.0051	4.39e-3	6.30e-4
ME	8-30-30-1	1e-3	8	6.83e-5	254.1	5.08e-5	0.0066	6.87e-3	5.43e-4
MC	8-25-25-1	1e-3	8	4.13e-5	117.7	2.86e-5	0.0071	8.39e-3	0.186

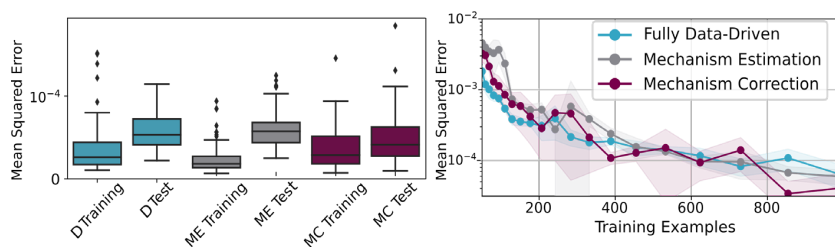


Figure 2: Left: Distribution of MSEs for models with fixed hyperparameters. Right: Model's loss function on the test data against different amounts of training data.

When extrapolating to a temperature range not covered by the training data (see Table 1), the D model performs significantly worse than when predicting unknown data within the trained temperature range. Fully data-driven models perform well only within the trained parameter range. The ME model and the MC models perform reasonably well in predicting the extrapolated data. The mechanistic framework (ME model) or the mechanistic basis (MC model) seems to improve not only the interpretability of the models but also their extrapolation ability. The MLPR part of the ME model can also be used for dynamic simulations. Since  $R$  is independent of time,  $R_{\text{meth}}^{\text{MLPR}}$  can be solved in a steady state. The predictive ability of the D model and the ME model decreases only slightly when noise is added to the training and test data. The MC model is unable to handle noisy data. In the  $\kappa$ -term of the MC model, the structural error due to the simplified model and the error due to the random noise are superimposed, resulting in an incorrect prediction of the MC model.

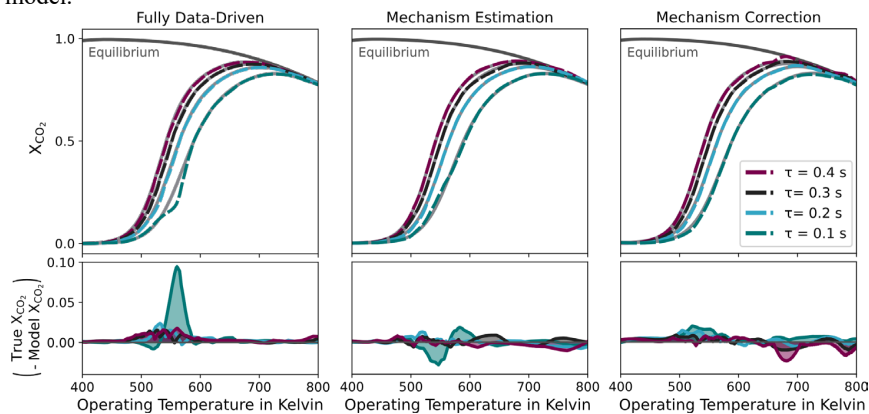


Figure 3: Model prediction of  $X_{\text{CO}_2}$  over the operating temperature. Model predictions are represented by dashed lines.

#### 4. Conclusion

Hybrid models are a promising approach to integrate data-driven and mechanistic models. This study proposes three modeling architectures for  $\text{CO}_2$ -methanation in a CSTR. The models are trained and tested using pseudo-experimental process data and predict  $X_{\text{CO}_2}$



with relatively small deviations. However, obtaining sufficient and high-quality process data is a challenge for all three models. The first model (D model) is straightforward to implement, but its performance is limited to the trained parameter range. The second model (ME model) uses a mechanism estimation approach and integrates a multilevel perceptron regression (MLPR) into the mass and energy balances. This approach provides a better representation of the reaction rate, including extrapolation data and noisy data, without requiring kinetic experiments. The third model (MC model) employs a mechanism correction approach, where a data-driven correction term is added to a mechanistic model. The MC model demonstrates good performance in both the trained and untrained parameter space but struggles with handling noisy data. In conclusion, the proposed hybrid models offer a way to improve the accuracy and reliability of process models for CO<sub>2</sub>-methanation

### Acknowledgment

This contribution is funded by the Bundesministerium für Bildung und Forschung (BMBF) and Project Management Jülich (PtJ) under grant 03HY302R.

### References

- Bremer, J. & Sundmacher, K., 2021. Novel Multiplicity and Stability Criteria for Non-Isothermal Fixed-Bed Reactors. *Frontiers in Energy Research*, Volume 8.
- Er-rbib, H. & Bouallou, C., 2013. Modelling and simulation of methanation catalytic reactor for renewable electricity. *Chemical Engineering Transactions*, pp. 541-546.
- Götz, M. et al., 2016. Renewable Power-to-Gas: A technological and economic review. *Renewable energy*, Volume 85, pp. 1371-1390.
- Kahrs, O. & Marquardt, W., 2007. The validity domain of hybrid models and its application in process optimization. *Chemical Engineering and Processing: Process Intensification*, Volume 46(11), pp. 1054-1066.
- Koschany, F., Schlereth, D. & Hinrichsen, O., 2016. On the kinetics of the methanation of carbon dioxide on coprecipitated NiAl (o) x.. *Applied Catalysis B: Environmental*, Volume 181, pp. 504-516.
- Menghani, G., 2021. Efficient deep learning: A survey on making deep learning models smaller, faster, and better. *arXiv preprint, arXiv:2106.08962*.
- Von Stosch, M., Oliveira, R., Peres, J. & de Azevedo, S. F., 2014. Hybrid semi-parametric modeling in process systems engineering: Past, present and future. *Computers & Chemical Engineering*, Volume 60, pp. 86-101.
- Zahedi, G. et al., 2005. Hybrid artificial neural network-first principle model formulation for the unsteady state simulation and analysis of a packed bed reactor for CO<sub>2</sub> hydrogenation to methanol. *Chemical Engineering Journal*, Volume 115, pp. 113-120.
- Zimmermann, R. T., Bremer, J. & Sundmacher, K., 2020. Optimal catalyst particle design for flexible fixed-bed CO<sub>2</sub> methanation reactors. *Chemical Engineering Journal*, Volume 387.

## Dynamic Optimization of Active Pharmaceutical Ingredient (Semi-)Batch Crystallization using Population Balance Modelling

Gustavo L. Quilló<sup>a,b</sup>, Jan F.M. Van Impe<sup>a\*</sup>, Alain Collas<sup>b</sup>, Christos Xiouras<sup>b\*</sup>, Satyajeet S. Bhonsale<sup>a</sup>

<sup>a</sup>*BioTeC+, Chemical and Biochemical Process Technology and Control, KU Leuven Campus Ghent, Gebroeders De Smetstraat 1, 9000 Ghent, Belgium*

<sup>b</sup>*Chemical & Pharmaceutical Development and Supply, Janssen R&D, Turnhoutseweg 30, 2340 Beerse, Belgium*

[cxiouras@its.jnj.com](mailto:cxiouras@its.jnj.com), [jan.vanimpe@kuleuven.be](mailto:jan.vanimpe@kuleuven.be)

### Abstract

Current industrial (semi-)batch crystallization processes follow recipe-based temperature and antisolvent dosing trajectories obtained by extensive exploratory experimentation. This work optimizes the trajectories based on a 1-D Population Balance Equation coupled with the mass balance, (secondary) nucleation and crystal growth kinetics. The system of differential equations is solved by a High-Resolution Finite Volume Method. The method is demonstrated in a combined cooling-antisolvent crystallization that maximizes the number-average crystal size while satisfying process constraints (e.g., minimum yield). The dynamic optimization problem is solved via single shooting, where the control trajectory is discretized as piecewise linear, and the global solution is determined by multi-start single-objective optimization.

**Keywords:** crystallization, dynamic optimization, thermodynamics, population balance modelling, high-resolution finite volume method.

### 1. Introduction

The Population Balance Equation (PBE) presents a viable mathematical framework for the description of many chemical processes, such as crystallization, deposition, granulation, flocculation, milling, drying, mixing, and reactive multiphase systems such as polymerization and bioreactors (Randolph and Larson, 1971). In all these processes, it is vital to track the time evolution of a distributed property, e.g., the particle size distribution (PSD) in crystallization applications. The formulation for batch processes in the absence of breakage and agglomeration results in a hyperbolic partial differential equation for the crystal number density  $f$  [#/(kg-neat solvent.m)] in the form of eq. (1),

$$\frac{\partial f(L, t)}{\partial t} + \frac{\partial [G(L, t)f(L, t)]}{\partial L} + \frac{f(L, t)}{M} \frac{\partial M}{\partial t} = 0 \quad (1)$$

with boundary conditions  $f(L_0, t) = B(L_0, t)/G(L_0, t)$ ,  $f(\infty, t) = 0$  and initial condition  $f(L, 0) = f_{\text{seed}}(L)$ . This semi-mechanistic approach is then used to engineer processes to satisfy a set of critical quality attributes by numerically solving a system of equations.

### 2. Problem formulation and computational methods

The PBE is supplied with several auxiliary functions. First, the mole-fraction, activity dependent supersaturation (driving force) for crystal nucleation and growth is given by

eq. (2), where the denominator refers to the saturated state and  $\gamma$  and  $x$  are the activity coefficient and molar fraction of the API, respectively. The procedure to estimate the activity at supersaturated state  $\gamma_{\text{API}}$  is available elsewhere (Quilló et al., 2021).

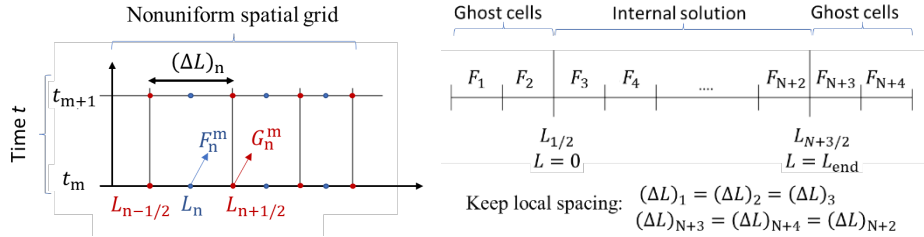
$$\ln(\sigma_{\text{API,MFAD}}) = \ln\left(\frac{\gamma_{\text{API}}x_{\text{API}}}{\gamma_{\text{API}}^{\text{sat}}x_{\text{API}}^{\text{sat}}}\right) \quad (2)$$

The kinetic expressions for the overall linear growth rate  $\bar{G}$  and secondary nucleation are taken as empirical power law functions in this work as in eq. (3).

$$\bar{G} = k_g [\ln(\sigma_{\text{API}})]^g \quad B = k_b [\ln(\sigma_{\text{API}})]^b \quad \sigma_{\text{API}} \geq 1 \quad (3)$$

### 2.1. High-resolution Finite Volume Scheme (HR-FVM)

The HR-FVM approach allows tracking the PSD, although at additional computational effort (LeVeque, 2002). The PBE is discretized in a static, nonuniform, non-smooth, structured spatial grid which results in eq. (4), where  $N$  is the number of cells (see **Figure 1**). The growth rate  $G_n$  can be size-dependent and is evaluated at the upper boundary of cell  $n$ , where it can assume positive (growth,  $G_n^+$ ) or negative (dissolution,  $G_n^-$ ) values. The formulation allows thermal cycling if the dissolution kinetics is known by applying absorbing boundary conditions on both sides of the domain if  $\sigma_{\text{API}} < 1$ , which imply in complete dissolution of any clusters of the lowest size range.



**Figure 1.** Nonuniform grid in the characteristic size coordinate extended by ghost cells for the implementation of boundary conditions, along with indexation. The internal solution is extended by two ghost cells on each side because the flux limiter implies in a stencil of five cells.

$$\begin{aligned} \hat{F}_n^{m+1} = & F_n^m - \frac{\Delta t^m}{\kappa_{n-1/2} \Delta \xi} (G_n^+ F_n^m - G_{n-1}^+ F_{n-1}^m + G_{n+1}^- F_{n+1}^m - G_n^- F_n^m) \\ & - \left[ \frac{\Delta t^m G_n^+}{2\kappa_{n-1/2} \Delta \xi} \left( \frac{\kappa_{n-1}}{\kappa_{n-1/2}} - \frac{\Delta t^m G_n^+}{\kappa_{n-1/2} \Delta \xi} \right) (F_{n+1}^m - F_n^m) \phi_n \right. \\ & - \frac{\Delta t^m G_{n-1}^+}{2\kappa_{n-1/2} \Delta \xi} \left( \frac{\kappa_{n-2}}{\kappa_{n-3/2}} - \frac{\Delta t^m G_{n-1}^+}{\kappa_{n-1/2} \Delta \xi} \right) (F_n^m - F_{n-1}^m) \phi_{n-1} \left. \right] \\ & + \left[ \frac{\Delta t^m G_n^-}{2\kappa_{n-1/2} \Delta \xi} \left( \frac{\kappa_{n-1}}{\kappa_{n-1/2}} + \frac{\Delta t^m G_n^-}{\kappa_{n-1/2} \Delta \xi} \right) (F_{n+1}^m - F_n^m) \phi_n \right. \\ & - \frac{\Delta t^m G_{n-1}^-}{2\kappa_{n-1/2} \Delta \xi} \left( \frac{\kappa_{n-2}}{\kappa_{n-3/2}} + \frac{\Delta t^m G_{n-1}^-}{\kappa_{n-1/2} \Delta \xi} \right) (F_n^m - F_{n-1}^m) \phi_{n-1} \left. \right] \quad (4) \end{aligned}$$

The intermediate solution  $\hat{F}_n^{m+1}$  is corrected by the fractional step method in eq. (5), which allows sequential solution of the homogenous PBE and source terms (Bosetti and Mazzotti, 2020; LeVeque, 2002). The coupling with the mass balance of the solute in eq. (6) and the solvent in eq. (7) occurs by the supersaturation used to compute nucleation and growth. Moreover, the solvent mass balance leads to dilution and changes the API

solubility by altering the solution composition through the dosing stream  $Q_{in}$  [kg-neat solvent] with mass composition  $w_{in}$ . Further, the mixture volume  $V$  is obtained by eq. (8).

$$F_n^{m+1} = \hat{F}_n^{m+1}(M^m/M^{m+1}) \quad (5)$$

$$c_{API}^{m+1} = c_{API}^m - \rho_c k_v \left( 3 \sum_{n=1}^N [G_n^m F_n^m L_n^2 (\Delta L)_n] + B^m L_0^3 \right) \Delta t^m - \frac{c_{API}^m}{M^m} (M^{m+1} - M^m) \quad (6)$$

$$M_i^{m+1} = M_i^m + w_{in,i} Q_{in} \Delta t^m \quad (7)$$

$$V^m = \frac{\sum M_i^m}{\rho_{mix}} = \frac{M^m}{1/\sum(w_i/\rho_i)} \quad (8)$$

The left- and right-going propagating waves are retrieved separately by eq. (9).

$$G_n^- = \min(G_n, 0), \quad G_n^+ = \max(G_n, 0) \quad (9)$$

The nonuniformity in the spatial grid is tackled by the cell capacity  $\kappa$ , which is the ratio of the true cell size  $(\Delta L)_n$  and a computational step  $\Delta\xi$  as in eq. (10). The non-smoothness is addressed by the mean of adjacent cell capacities as in eq. (11), where  $L_{end}$  and  $L_0$  are the right and left boundaries of the discretized spatial domain, respectively.

$$\kappa_n = (\Delta L)_n / \Delta\xi, \text{ where } \Delta\xi = (L_{end} - L_0) / N \quad (10)$$

$$\kappa_{n-1/2} = (1/2)(\kappa_{n-1} + \kappa_n) \quad (11)$$

The flux limiter function  $\phi$  in eq. (12) depends on the smoothness  $\theta$  defined in eq. (13), where  $\varepsilon$  is a small number. Many possibilities of high-resolution flux limiters are available in the literature (LeVeque, 2002) (e.g. minmod, superbee, MC, Koren, van Leer), all of which are algebraic total variation diminishing (TDV) but imply in mildly different capabilities to mitigate numerical problems (e.g., diffusion, dispersion) in the obtained solution. This work uses the superbee flux limiter, as it displays minimal solution smearing while retaining near 2<sup>nd</sup>-order accuracy.

$$\phi_{n,Superbee}(\theta_n) = \max(0, \min((1 + \theta_n)/2, 2, 2\theta_n)) \quad (12)$$

$$\theta_n = \frac{F_n - F_{n-1} + \varepsilon}{F_{n+1} - F_n + \varepsilon} \quad (13)$$

The Dirichlet boundary conditions are applied by ghost cells (see **Figure 1**). The left side of the domain is limited by nuclei size  $L_0$ , and the values assigned to the ghost cells are calculated by linear extrapolation as in eq. (14) to eq. (16), i.e., by the upwind method.

$$F_{L=L_0} = [2/(\Delta L)_1](F_3 - B(t)/G(L_0, t)) \quad (14)$$

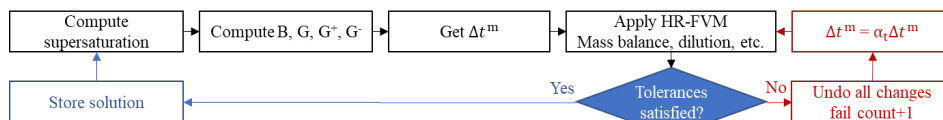
$$F_1 = -2(\Delta L)_1 F_{L=L_0} + F_3 \quad (15)$$

$$F_2 = -(\Delta L)_1 F_{L=L_0} + F_3 \quad (16)$$

The ghost cells on the right-side of the domain embody an absorbing boundary condition, where the  $L$  coordinate is artificially truncated at a sufficiently high value  $L_{end}$  to avoid particle loss to the boundary over the process time frame. The boundary condition is implemented via zero-order extrapolation, as in eq. (17).

$$F_{N+3} = F_{N+4} = F_{N+2} \quad (17)$$

The time coordinate is discretized by a forward Euler scheme, where the time step  $\Delta t^m$  is selected based on the most stringent criterion among the Courant condition for numerical stability  $\min((\Delta L)_n/G_n) \leq 1$ , the acceleration limit based on the previous time step, the minimum and maximum user-specified integration steps and the query time points. After the application of the vectorized HR-FVM scheme, the mass balance for the solute API and the dilution effect are updated using the fractional-step method. Next, the applied time step is approved or rejected based on relative tolerance checks in reference to the previous step supersaturation, solvent mass, and concentration and negative  $F$  due to small oscillations. The rejection cycle is applied up to three times before the time step is forcefully accepted, and each cycle reduces the applied time step ( $\alpha_t = 1/3$ ) prior to reapplication of HR-FVM. The summary of the algorithm is provided in **Figure 2**.



**Figure 2.** Implementation of the adaptive time scheme for the HR-FVM with error control for the custom-built partial differential equation integrator.

### 3. Dynamic optimization of cooling-antisolvent seeded (semi-)batch crystallization

The model is used to find optimal temperature and/or antisolvent trajectories that minimize (or maximize) a given goal under constraints that reflect equipment operation limits. In this case, the objective is to maximize the number average crystal size, defined by the ratio of the first to the zeroth moment of the PSD at final batch time  $t_f$  given the system parameters  $\theta$ . The nonlinear programming problem (NLP) is stated in eq. (18).

$$\begin{aligned} & \max \quad f(t_f, \theta) \\ & \text{subject to:} \quad T(k), M_{in}(k) \\ & \quad T_{\min}(k) \leq T(k) \leq T_{\max}(k) \\ & \quad R_{\min}(k) \leq dT/dt(k) \leq R_{\max}(k) \\ & \quad Q_{\min}(k) \leq dM_{in}/dt(k) \leq Q_{\max}(k) \\ & \quad 0.9 \max(\text{Yield}_{\text{theoretical}}) \leq \text{Yield}_{\text{process}}(t_f) \\ & \quad V(t_f) \leq V_{\max} \end{aligned} \quad (18)$$

The trajectories are obtained by optimizing the piecewise linear control variables at specified nodes  $T(k), M_{in}(k)$  located in regular time intervals and solving the system by HR-FVM. In this case study, the control variables were discretized in 16 nodes after preliminary numerical analysis. The optimal trajectory is determined by derivative-free, trust-region-based, multi-start optimization using the COBYLA algorithm (Powell, 1994). The method is implemented in MATLAB® 2022b, although the optimization and objective function evaluations are compiled in FORTRAN and in C++ by MEX files, respectively, which speed up calculations by at least threefold. The data for the case study with fixed batch time is given in **Table 1**.

**Table 1.** Summary of parameters for the case study.

$k_b = 7.92 \times 10^{14} \text{ \#/(kg.s)}$	$b = 6.22$	$k_g = 2.38 \times 10^{-5} \text{ m/s}$	$g = 2.37$
$\rho_c = 1300 \text{ kg/m}^3$	$k_v = 0.1$	$c_{API,t=0} = 0.33 \text{ kg/kg}$	$M_{t=0} = 0.288 \text{ kg}$
$V_{\max} = 400 \text{ mL}$	$t_{\text{batch}} = 16 \text{ h}$	$T_{\min} = 33 \text{ }^\circ\text{C}$	$T_{\max} = 60 \text{ }^\circ\text{C}$
$R_{\min} = -1 \text{ }^\circ\text{C/min}$	$R_{\max} = 0 \text{ }^\circ\text{C/min}$	$L_{\text{end}} = 1000 \text{ }\mu\text{m}$	$N = 4000$

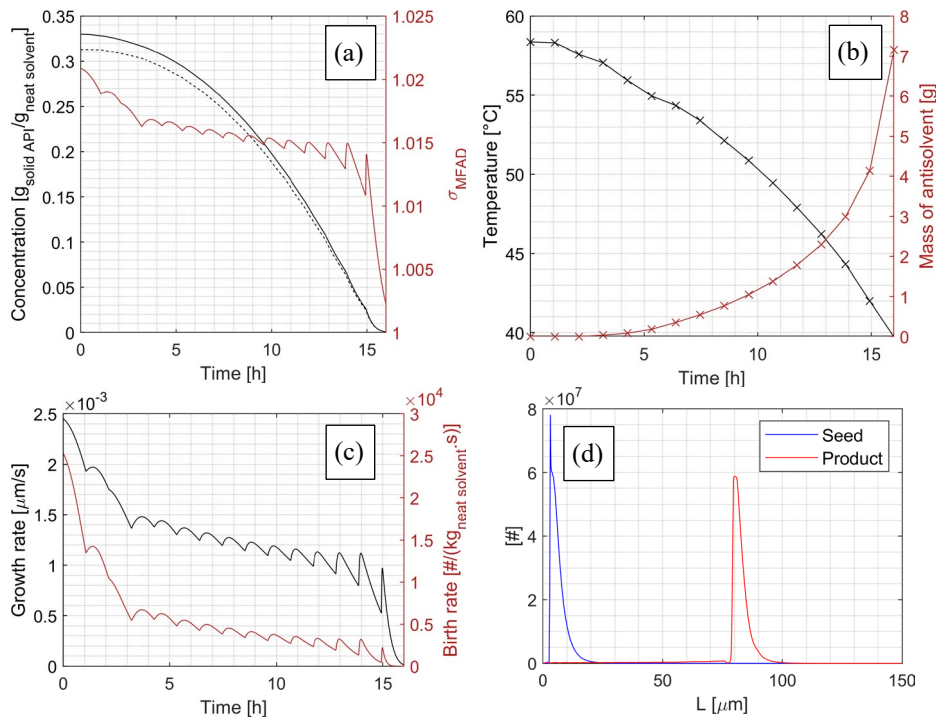
The API solubility is calculated by the Van't Hoff Jouyban-Acree model as in eq. (19).

$$\ln(x_{\text{API}}^{\text{sat}}) = \sum_{p \in \{1,2,3\}} \left( \psi_{0,p} w_p + \psi_{1,p} \frac{w_p}{T} \right) + \sum_{p,q \in \{1,2,3\}; p \neq q} \frac{w_p w_q}{T} \left[ \psi_{2,p,q} + \psi_{3,p,q} (w_p - w_q) + \psi_{4,p,q} (w_p - w_q)^2 \right] \quad (19)$$

Table 2 Van't Hoff Jouyban Acree solubility parameters. Temperature is in degrees Kelvin.

$\psi_{0,1} = -8.77$	$\psi_{1,1} = 9.43 \times 10^6$	$\psi_{2,1,2} = 7.80 \times 10^6$	$\psi_{3,1,2} = -3.40 \times 10^7$	$\psi_{4,1,2} = -3.47 \times 10^7$
$\psi_{0,2} = 17.2$	$\psi_{1,2} = -1.39 \times 10^7$	$\psi_{2,1,3} = -1.55 \times 10^7$	$\psi_{3,1,3} = -6.93 \times 10^6$	$\psi_{4,1,3} = -8.03 \times 10^5$
$\psi_{0,3} = 3.36$	$\psi_{1,3} = -3.06 \times 10^3$	$\psi_{2,2,3} = 2.07 \times 10^7$	$\psi_{3,2,3} = 7.94 \times 10^6$	$\psi_{4,2,3} = 1.15 \times 10^6$

The initial ternary solvent mass composition is  $w_1 = 0.503$ ,  $w_2 = 0.313$  and  $w_3 = 0.184$  and pure antisolvent is dosed ( $w_{\text{in},2} = 1$ ). The theoretical maximum yield is  $\approx 100\%$  due to the strong antisolvent effect. The optimization results are summarized in **Figure 3**.



**Figure 3.** (a) Solution for the mass balance for HR-FVM. Solid line and dashed black line represent the crystallizer concentration and the solubility, respectively, while the red line depicts the supersaturation. (b) Optimal piecewise linear temperature trajectory. (c) Time evolution of the growth and birth rates. (d) Seed and product number density.

The HR-FVM method using the custom-built integrator was validated by comparison with standard ODE integrators in MATLAB (i.e., ode45, ode23s) solving the equivalent NLP by the Standard Method of Moments (SMOM) with adequate solution juxtaposition.

The problem identifies that the initial supersaturation is too high and favors nucleation excessively, so the optimal trajectory starts with an isothermal hold and no antisolvent addition at the beginning. Then, cooling starts before dosing since it is a gentler form of managing the supersaturation, followed by the combined effect of both controls, which are required to satisfy the imposed yield constraint. Furthermore, the optimal profile is consisted of increasingly larger cooling and dosing rates which cause supersaturation to be nearly constant at a level that suppresses nucleation for most of the simulation while growth is consuming the supersaturation. The long but small tail in the product PSD is caused by nucleation, which constantly generates fines, whereas the seeds PSD simply propagate in the size coordinate at the speed determined by the growth rate due to the absence of dissolution and particle death events (breakage, agglomeration). The peak artifact present in the seeds PSD is smeared during the solution, but the shape of the initial PSD is mostly well-retained. The piecewise linear discretization of the controls provides significantly better results than the piecewise constant counterparts (not shown), which implies in supersaturation spikes close to the nodes due to the sudden change in solubility.

#### 4. Conclusions

Piecewise linear optimal temperature and antisolvent addition trajectories were successfully obtained from a custom-built integrator using a versatile and easily extensible implementation of the HR-FVM method to solve a PBE for a (semi-)batch crystallizer. The method at current allows grid tuning, size-dependent growth, and thermal cycles, and can be expanded to cope with more complicated source terms from breakage and agglomeration events by the fractional-step method.

The custom-built integrator for HR-FVM presented comparable accuracy to standard MATLAB integrators using SMOM while delivering more information, however, it needed 2-100 times more optimization time depending on the selected spatial grid resolution. Therefore, there are opportunities to tune the discretization (e.g., via Runge-Kutta, Crank-Nicholson for the time, adaptive spatial grid) to improve the calculations speed. The obtained control trajectory profiles led to crystals that are on average  $\approx 12.4$  times larger than the seeds and low quantity of fines, as nucleation was suppressed.

#### 5. Acknowledgements

This work was supported by VLAIO (Agentschap Innoveren & Ondernemen) and Janssen Pharmaceutica [Baekeland grant number HBC.2020.2214].

#### References

- L. Bosetti., M. Mazzotti, 2020. Population Balance Modeling of Growth and Secondary Nucleation by Attrition and Ripening. *Cryst. Growth Des.* 20, 307–319.
- R.J. LeVeque, 2002. *Finite Volume Methods for Hyperbolic Problems*, 1st ed. Cambridge University Press, Cambridge, United Kingdom.
- M.J.D Powell., 1994. A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation, in: *Advances in Optimization and Numerical Analysis*. Springer Netherlands, Dordrecht, 51–67.
- G.L. Quilló, S. Bhonsale, B. Gielen, J.F. Van Impe., A. Collas, C. Xiouras, 2021. Crystal Growth Kinetics of an Industrial Active Pharmaceutical Ingredient: Implications of Different Representations of Supersaturation and Simultaneous Growth Mechanisms. *Cryst. Growth Des.* 21, 5403–5420.
- A.D. Randolph, M.A. Larson, 1971. *Theory of Particulate Processes: Analysis and Techniques of Continuous Crystallization*, 1st ed. Academic Press, Inc., New York, NY.

# A Systematic Framework for Iterative Model-Based Experimental Design of Batch and Continuous Crystallization Systems

Hemalatha Kilari, Yash Barhate, Yung-Shun Kang, Zoltan K. Nagy\*

*Davidson School of Chemical Engineering, Purdue University, West Lafayette, Indiana 47907, United States. \*znagy@purdue.edu*

## Abstract

A generalized and systematic framework is proposed for rapid model development of batch and continuous crystallization processes. The proposed approach uses iterative model-based experimental design (IMED or MBDoE), whereby the experiments are guided by an imprecise-model to produce information rich data that will enhance its prediction performance for further improvement of model structure, parameters, robustness and accuracy. This enables to generate a precise model that can be used for *in silico* design of experiments in broader range of process parameters as well as for efficient digital design. This paper elucidates various steps for model discrimination, selection as well as precise parameter estimation with a combination of statistical tests and optimal experimental design approaches to arrive at an accurate model. The application benefits of the proposed strategy are demonstrated through case studies for batch and continuous crystallization systems of commercial active pharmaceutical ingredients.

**Keywords:** Optimal experimental design; Model-based experimental design, Design of Experiments; Batch crystallization; Continuous crystallization.

## 1. Introduction

Crystallization is a complex solid-liquid-separation technique that is widely applied in the manufacturing of solid products. High fidelity models are necessary for process improvements to achieve customized product specifications and properties. The usual industrial practice for modelling a crystallization process involves a traditional Quality-by-design (QbD) approach where a predetermined set of crystallization experiments are designed either by prior-knowledge/intuition or traditional statistical design of experiments (DoE) that considers the effect of critical process parameters (CPPs) such as seed, supersaturation, temperature, on the critical quality attributes (CQAs) such as mean crystal size, crystal size distribution (CSD), shape and purity. When large number of CPPs are considered, this approach results in a large set of experiments that is resource intensive. This calls for a more systematic procedure for rapid model development.

Application of model-based experimental design approach for rapid model development has been demonstrated to crystallization processes in earlier studies (Chung et al., 2000; Chen et al. 2004; Pal et al., 2020). It is evident from their studies that such approach is capable of predicting information rich experiments and reduce number of experiments to obtain accurate model with precise estimates of parameters. However, the focus in those studies was on either a particular case study involving batch or a particular type of crystallization process concentrating on either model discrimination or improving



parameter precision only and are sometimes case specific. Also, with increased adoption of continuous manufacturing in the industries, it is necessary to extend this procedure to continuous crystallization systems.

The objective of this work is to propose a generalized framework that gives guidelines to follow a model building procedure for crystallization processes starting with model screening from a model structure database and conduct an iterative sequence of statistical approaches to discriminate between the models along with optimal experimental design criterion to obtain a best model with precise parameters. The experimental design in crystallization processes is posed as an optimal control problem with in a framework that enables to choose cooling recipe, seed recipe, antisolvent flow profile, experiment time, residence time, number of crystallizers, number of sampling points etc. with ease of allowing constraints on inputs and outputs. In the proposed framework, methods are presented accounting to parameter uncertainty to improve model robustness. This work presents the first generalized comprehensive study that is applied to both batch and continuous crystallization processes.

## 2. Framework Development Methodology

### 2.1. Crystallization Model

Population balance equation (PBE) governing the CSD assuming nucleation with constant nucleation size, growth and agglomeration mechanisms can be represented as follows (Szilagyi et al., 2020; Szilagyi et al., 2022).

$$\frac{\partial[(Vn(L,t))]}{\partial t} + \frac{\partial(Gn(L,t)V)}{\partial L} = V[B\delta(L-L_n) + B_{agg} - D_{agg} + B_{bre} - D_{bre}] + F_f n_f(L,t) - Fn(L,t) \quad (1)$$

$$\frac{d(Vc)}{dt} = -3Vk_v \rho_c \left[ (B_p + B_s)L_n^3 + \int_0^\infty GL^2 n(L)dL \right] + F_f c_f - Fc \quad (2)$$

$$B_p = k_p(1 + \gamma_p A)\sigma^p \exp\left(-\frac{k_p}{\ln^2(\sigma+1)}\right) \exp\left(\frac{-E_p}{RT}\right); B_s = k_s(1 + \gamma_s A)\sigma^s \exp\left(-\frac{k_s}{\ln^2(\sigma+1)}\right) \exp\left(\frac{-E_s}{RT}\right); G = k_g(1 + \gamma_g A)\sigma^g (\alpha_g + \beta_g L^{\gamma_g}) \exp\left(\frac{-E_g}{RT}\right) \quad (3)$$

From Eq (1),  $B_{agg}$  and  $B_{bre}$  are birth functions that describe production of new crystals of size  $L$  by agglomeration and breakage.  $D_{agg}$  and  $D_{bre}$  represent Death functions for the rate of consumption of crystals of size  $L$  by agglomeration and breakage. Eq (2) represents mass balance equation for crystallizing compound that accounts for solid-liquid mass transfer generated by nucleation and growth with feed and outflow streams in a crystallizer. Generalized expressions for primary & secondary nucleation and growth are shown in Eq (3). As a first step to modeling crystallization processes, experimental data is obtained for a preliminary set of crystallization experiments by monitoring the process variables throughout the experiment using in situ or inline process analytical technology (PAT) tools and offline characterization methods. An optimization problem for parameter estimation is formulated (Eq (4)) with the objective of minimizing the weighted sum of squared error ( $WMSE$ ) between experimental and model predicted values of the process parameters.

$$\min_{\theta} WMSE = \frac{1}{N_E N_S} \sum_{h=1}^{N_v} \sum_{i=1}^{N_E} \sum_{j=1}^{N_S} w_h (y_{h,i,j}^{\text{exp}} - y_{h,i,j}^{\text{mod}})^2 \quad (4)$$

Where  $N_v$ ,  $N_E$  and  $N_S$  are number of process variables, experiments and samples or data points respectively;  $w_h$  are the weights chosen to normalize the value of the objectives,  $y$  represents the process variables such as concentration, online crystal counts /sec during the operation time, final time mean size or  $D_{10}$ ,  $D_{50}$  and  $D_{90}$  values, or the entire CSD, depending on measurement and modelling techniques used.

*2.2. Optimal Experimental design for Model Identification, Discrimination and Parameter precision*

Selecting a best model from a crystallization model library consisting of combinations of various expressions for different mechanisms is a challenging task. Initial kinetic parameter estimation for some of the chosen models may give an incorrect model that still can closely represent the experimental data. It is of primary importance to statistically assess the reliability and adequacy of the model and if necessary, of the new parameter estimates. This can be done using standard checks (i.e. on residuals distributions) and statistical tests. Statistical goodness-of-fit tests such as chi-square test, F-test, student-t test, etc. can be used for assessing model identifiability and parameter estimability. The problem of over fitting and underfitting a model can be tackled well by using Akaike information criterion (*AIC*) which estimates prediction error and thereby provides the relative quality of statistical models for a given set of data. When all these statistical tests fail to screen one model as the best model, optimal experimental design criterion for model discrimination (MD) can be used, which determines experimental conditions for additional experiments to be conducted in order to maximize the divergence between models. There have been many such approaches reported so far and a modified Hunter-Reiner criterion for many rival models similar to the one reported by Buzzi-Ferraris-Forzatti (1992) has been used in our work. When a best model is selected, it is subjected to parameter refinement (PR) to reduce the parameter uncertainties and improve the predictive capability of the model. Various steps involved in the IMED procedure are illustrated in the Fig.1.

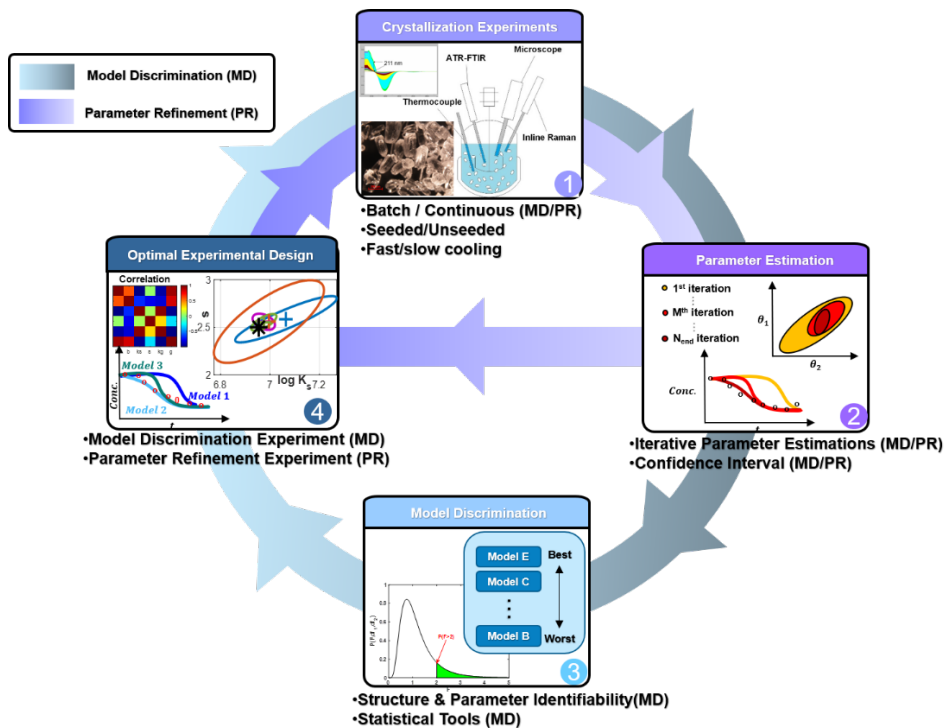


Fig.1. Iterative model-based experimental design procedure for crystallization processes

*D*-optimal design is a popular design procedure widely used for improving parameter precision that uses Fisher Information Matrix (FIM) as a metric to assess experimental information of the present experiments and prescribe a future experiment that provides with more information rich data. FIM is constructed from sensitivity matrix of process states and model parameters (Yang et al., 2022). The new experiment is obtained by maximizing the determinant of the *FIM* (which is an indirect approach of minimizing confidence region of parameter estimates). The generic mathematical representation of this *D*-optimal approach is shown below for a batch cooling crystallization and for a continuous MSMMPR crystallization system.

<i>Batch Cooling Crystallization</i>	<i>Continuous MSMMPR Crystallization</i>
$\max_{N_E, CR, Ms} [\det(FIM)]$ <p>Subject to:</p> $c_1 : Ms_{\min} \leq Ms \leq Ms_{\max}$ $c_2 : CR_{\min} \leq CR \leq CR_{\max}$ $c_3 : CR \leq 0$	$\max_{N_c, \tau_i, T_i} [\det(FIM)]$ <p>Subject to:</p> $c_1 : \tau_{\min} \leq \tau_i \leq \tau_{\max}$ $c_2 : T_{N_c} \leq T_i \leq T_f$ $c_3 : T_i < T_{i-1} \quad (i=1 \dots N_c)$

(5)

Where *CR* is the cooling rate for a linear temperature profile, *Ms* is the seed loading,  $N_E$  is the number of experiments necessary for another iteration for the batch case and  $N_c$  is number of crystallizers in a cascade of MSMMPRs for the continuous case with  $\tau_i$  being the residence time and  $T_i$ , the temperature of  $i^{th}$  Crystallizer. *FIM* is the fisher information matrix and is constructed from the sensitivity matrix with sensitivity coefficients of process variables at various sampling instants with respect to each parameter. This step yields new experimental conditions to be implemented for data collection and for parameter refinement. This procedure is iterated until a predefined model accuracy is obtained with precise parameter estimates.

### 3. Application of the Proposed Framework: Results

The benefits of the framework are tested first by application to a batch cooling crystallization case study with a known model. *In silico* experiments are generated for a combination of high/low seed loading and fast/slow linear cooling experiments. This set of four seeded experiments are used to obtain initial parameter estimates for six plausible model candidates (Table 1) that are subjected to model discrimination. It is observed that all these models closely fit the concentration and final mean size data. The AIC criteria correctly chooses model 6 as the best model. However, WMSE values are very close for other models too. Hence, it is decided to allow these models for an iteration of discriminatory experimental design (MD-OED). F-test (Sobek and Werle, 2020) is efficient in screening 3 models out of 6 for MD-OED procedure. With one iteration of the MD-OED, an unseeded experiment is obtained and this experiment is carried out *in silico* to collect experimental data for the next iteration of parameter estimation procedure. This experiment is added to the initial parameter set and a new parameter estimation is carried out for all 3 models. The unseeded experiment seems to create a better divergence between the 3 models and helps to choose the best model amongst them. Model 6 is chosen as the best model (true model) according to AIC and also F-test rejects other models choosing model 6 as the best one. Model 6 is now chosen for iterative optimal experimental design procedure for improving precision of parameters.

Table 1: Model discrimination and Selection

Model	Models	N <sub>0</sub>	WMSE *10 <sup>3</sup>	AIC	F-value (F <sub>Critical</sub> =2.21)	Decision
<b>Initial Parameter Estimation with a preliminary set of 4 seeded experiments</b>						
1	SN+SIG	3	1.11	-15.07	7.8	Reject
2	SN+SIG <sup>T</sup>	4	1.68	-14.65	11.8	Reject
3	SN+SDG	4	0.22	-16.69	1.5	Accept
4	PN+SN+SIG	5	0.27	-16.48	1.9	Accept
5	PN+SN+ SIG <sup>T</sup>	6	0.76	-15.44	5.3	Reject
6	PN+SN+SDG *True model	6	0.14	-17.12	1.0	Accept
<b>Re-Parameter estimation after including MD-OED experiment</b>						
3	SN+SDG	4	3420.0	-1.88	1.89E+07	Reject
4	PN+SN+SIG	5	20410.0	-1.10	1.12E+08	Reject
6	PN+SN+SDG *True model	6	1.82E-04	-9.15	1.0	Accept
#PN: Primary nucleation; SN: Secondary Nucleation; SIG: Size Independent Growth, SIG <sup>T</sup> : SIG with Temperature term; SDG: Size dependent growth						

The IMED approach for parameter precision or parameter refinement can be carried out using either a sequential or a simultaneous approach. In the simultaneous approach a complete set of experiments with either a fixed number (e.g. in our case  $n_{exp} = 4$ ) or using the number of experiments also as a decision variable can be used to obtain the new set of experimental conditions for all experiments within a single optimization that could be implemented to collect new experimental data for parameter refinement with new parameter estimation. In the sequential approach only one new experiment is obtained at each iteration that can be implemented and the results are added to previous experimental data to carry out the new parameter estimation.

Sample results are shown for the sequential approach where one experiment is obtained with each iteration and new parameters are estimated and analyzed for parameter uncertainty. The confidence region (Fig. 2) plotted for a pair of initial parameter estimates ( $\theta_1 = \log K_p$  &  $\theta_2 = p$ ; primary nucleation parameters shown in Eq(3)) is very broad. With each iteration of the optimal experimental design, the parameters are improved as shown by shrinkage of the confidence ellipse and moving closer to the true parameter values after each iteration. Fig 2(a) shows comparison of confidence ellipse with initial parameter estimation and a model discriminatory experiment. Adding the discriminatory experiment to the parameter estimation set improved the model as it can be clearly observed that the parameter values shifted closer to the true parameter region.

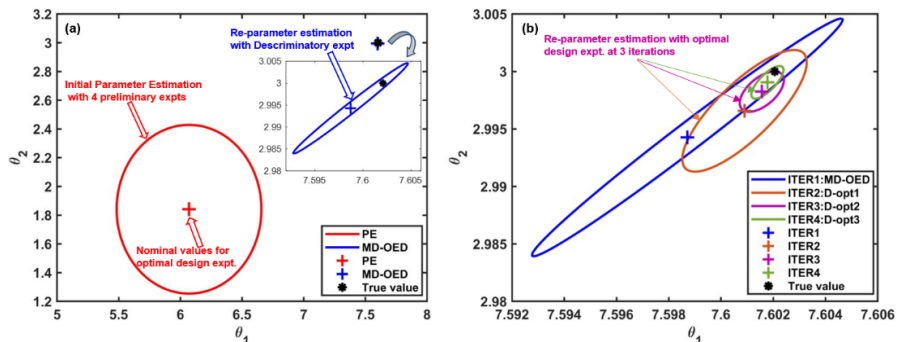


Fig.2. Confidence Ellipse plots (a) Initial Parameter Estimation (PE) & Model Discrimination expt. (MD-OED) (b) Optimally designed experiments using D-design Criterion (D-opt1, D-opt2, D-opt3) at each iteration (ITER)

Fig 2(b) shows the improvement of confidence region and corresponding parameters with each iteration of re-parameter estimation including respective D-optimal experiment. The IMED procedure is continued until no further improvement in the confidence region is obtained. In this case, 3 iterations of IMED brought the parameters closer to the true value of the parameters. Precise parameters with very narrow confidence intervals are obtained at the end of 3 iterations and found to be very close to the true parameters of the model considered. Thus, this batch case study demonstrates powerful predictive capabilities of the proposed framework. Similarly, another case study for continuous crystallization of Diphenhydramine hydrochloride is carried out and it is observed to be effective in arriving at an accurate model with precise parameters and very small confidence intervals.

#### 4. Conclusions

A generalized systematic framework is proposed for batch and continuous crystallization model development using an iterative model-based experimental design framework. The framework combines model discrimination and parameter refinement steps and is able to identify the best model structure that constitutes relevant crystallization mechanisms and drive the model with imprecise initial parameter estimates towards an accurate model with precise parameters. Future work is focused on expanding this knowledge to application scenarios that include other types of crystallization processes and multi-dimensional models to cater to a wider range of applications of crystallization modeling.

#### Acknowledgements

Funding for this work was made possible, in part, by the US Food and Drug Administration (FDA) through grant number 75F40121C00106. Views expressed herein do not necessarily reflect the official policies of the Department of Health and Human Services; nor does any mention of trade names, commercial practices, or organization imply endorsement by the United States Government.

#### References

- S. H. Chung, D. L., Ma., R. D., Braatz, 2000, Optimal model-based experimental design in batch crystallization. *Chemometrics and Intelligent Laboratory Systems*, 50(1), 83-90.
- B.H. Chen, S., Bermingham, A.H., Neumann, H.J. Kramer, S.P., Asprey, 2004, On the design of optimally informative experiments for dynamic crystallization process modeling, *Industrial & engineering chemistry research*, 43(16), 4889-4902.
- K. Pal, B. Szilagyi, Burcham, C.L., D.J. Jarmer, Z.K. Nagy, 2021, Iterative model-based experimental design for spherical agglomeration processes. *AIChE Journal*, 67(5), e17178.
- B. Szilagyi, A. Majumder, Z. K. Nagy, 2020. Fundamentals of Population Balance Based Crystallization Process Modeling, In *The Handbook of Continuous Crystallization* (pp. 51-101).
- B. Szilagyi, W. L. Wu, A. Eren, J. Mackey, S. Kshirsagar, E. Szilagyi, Z. K. Nagy, 2022, Cross-Pharma Collaboration for the Development of a Simulation Tool for the Model-Based Digital Design of Pharmaceutical Crystallization Processes (CrySiV). *Crystal Growth & Design*, 21(11), 6448-6464.
- G. Buzzi-Ferraris, P. Forzatti, 1990, An improved version of a sequential design criterion for discriminating among rival response models. *Chem. Eng. Sci.* 45 (2), 477-481.
- L. Yang, Y. Zhang, P. Liu, C. Wang, Y. Qu, J. Cheng, C. Yang, 2022, Kinetics and population balance modeling of antisolvent crystallization of polymorphic indomethacin. *Chemical Engineering Journal*, 428, 132591.
- S. Sobek, S. Werle, 2020, Kinetic modelling of waste wood devolatilization during pyrolysis based on thermogravimetric data and solar pyrolysis reactor performance. *Fuel*, 261, 116459.

# Evaluation of Factors Affecting Novel Technology Development using Mathematical Programming

Pooja Zen Santhamoorthy<sup>a</sup>, Selen Cremaschi<sup>a,\*</sup>

<sup>a</sup>*Department of Chemical Engineering, Auburn University, AL 36849, USA*  
*scz0113@auburn.edu*

## Abstract

The successful development and timely deployment of novel technologies to compete with existing ones depend on the availability of necessary resources and associated costs. This paper introduces a multiperiod optimization model that integrates processing plants, technologies, and resources under a single network to study the impact of novel technology development on the process network design to meet the annual product demands over a planning horizon. A case study of developing post-combustion adsorption for carbon capture (CC) while meeting the annual capture targets reveals that adsorption should become commercially available within 5 years to compete with the commercially available absorption technology. The total CC cost if only absorption were available is \$13.9 billion, 6% of which could be used towards the research and development (R&D) of the adsorption technology while still making it an economically viable option.

**Keywords:** technology development; multiperiod optimization; CO<sub>2</sub> capture planning.

## 1. Introduction

Technology development requires resources for R&D, pilot plants, and demonstration projects. The availability of such resources plays a major role in the successful development and timely implementation of new technologies to compete with existing and other novel technologies (Anderson Jr and Joglekar, 2005). For example, a high cost for a necessary resource can make a novel technology economically unviable, or prolonged development time can make it uncompetitive against technologies developing at a faster pace (Davila and Wouters, 2004). Hence, it is important to study the impacts of resource availability, the associated costs, and the development time on novel technology development along with the operation of the existing technologies. This work develops a multiperiod optimization model for production planning and resource allocation for developing novel technologies and operating implemented technologies to meet the annual product demands over time.

## 2. Problem Statement

The goal is to develop an optimization-based framework to study the development of novel technologies and deployment and operation of commercial technologies by determining strategic and tactical decisions: (i) allocation of resources for technology development, (ii) selection of processes and commercial technologies to produce products, (iii) determination of production rate based on the annual product demands, and (iv) addition of expansions to the process facilities, if necessary, over time. For given yearly product demands, the objective is to minimize the total cost of technology development, implementation, and operation.

The superstructure in Figure 1 illustrates the overarching network. It consists of a set of processes ( $i \in I$ ), technologies ( $j \in J$ ), resources ( $r \in R$ ), development stages called Technology Stages ( $s \in S$ ), expansions ( $n \in N$ ), and time periods ( $t \in T$ ). The technologies  $T_j$  are in different Technology Stages ( $TS_{j1}, TS_{j2}, \dots, TS_{js}$ ). They mature to subsequent stages if the necessary resources  $R_{rt}$  are allocated during development. The technically developed and commercially available technologies at  $TS_{js}$  are used by processes  $P_i$  in the facilities  $C_{ij}$  to produce products  $X_{it}$  over the planning period  $t$ . Expansion facility  $E_{nij}$  using technology  $j$  is added for process  $i$  when demand for product  $X_{it}$  increases at time  $t$ .

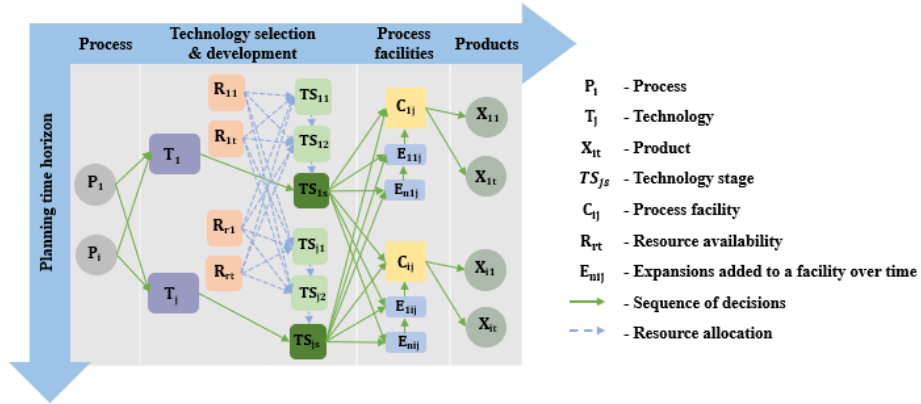


Figure 1. Superstructure of technology development and deployment to produce products.

### 3. Mathematical Model Formulation

#### 3.1. Objective

The objective is to minimize the total cost of producing a set of products over the planning horizon while meeting annual demands. The cost consists of annualized investment and annual operating costs, and R&D and demonstration project costs as given in Eq. (1).

$$\text{MinTC} = \sum_{i,j,t,n} (IC_{i,j,t,n} + OC_{i,j,t,n}) + \sum_{i,j,t} DC_{i,j,t} \quad \forall i \in I, j \in J, t \in T, n \in N \quad (1)$$

The sets  $i \in I, j \in J, t \in T$ , and  $n \in N$  represent the processes, technologies, planning periods, and capacity expansions. For process  $i$ ,  $n=1$  represents the initial facility capacity, and  $n>1$  represents the expansions. The variable  $TC$  is the total cost of investment and operation of implemented technologies and the cost of developing novel technologies. The variables  $IC_{i,j,t,n}$  and  $OC_{i,j,t,n}$  are the investment and operating costs of facility/expansion  $n$  of process  $i$  using technology  $j$  at time  $t$ . The variable  $DC_{i,j,t}$  is the technical development cost of technology  $j$  for process  $i$  at time  $t$ .

#### 3.2. Binary Variables

A binary variable  $yc_{i,j}$  is defined in Eq. (2) to represent the compatibility of technology  $j$  with process  $i$ . Technology Stages are represented by set  $s \in S$ , and a binary variable  $x_{i,j,t,s}$  is introduced in Eq. (3) to represent if technology  $j$  has been developed to at least Technology Stage  $s$  by time  $t$  for process  $i$ . If a technology is not compatible with process  $i$ , it remains in the initial stage of development ( $x_{i,j,t,1}=1 \quad \forall t \in T$  and  $x_{i,j,t,s}=0 \quad \forall t \in T, s \in \{2,3,\dots,S\}$ ). Another binary variable  $x'_{i,j,t,s}$  is introduced in Eq. (4) to denote if the resources necessary for development to stage  $s$  are available at time  $t$ . Once the technology  $j$  is commercially developed for process  $i$  ( $x_{i,j,t,s}=1$ ), a binary variable,  $y_{i,j,t,n}$ , in Eq. (5), represents the technology implementation for facility/expansion  $n$  of process  $i$  at time  $t$ .

$$yc_{i,j} = \begin{cases} 1, \text{if tech } j \text{ is compatible with process } i \\ 0, \text{otherwise} \end{cases} \quad (2)$$

$$x_{i,j,t,s} = \begin{cases} 1, \text{if tech } j \text{ is developed to at least stage } s \text{ for process } i \text{ by time } t \\ 0, \text{otherwise} \end{cases} \quad (3)$$

$$x_{i,j,t,s} = \begin{cases} 1, \text{if resources necessary for tech } j \text{ to develop to stage } s + 1 \\ \text{for process } i \text{ is available by time } t \\ 0, \text{otherwise} \end{cases} \quad (4)$$

$$y_{i,j,t,n} = \begin{cases} 1, \text{if technology } j \text{ is used in facility } n \text{ of process } i \text{ during time } t \\ 0, \text{otherwise} \end{cases} \quad (5)$$

### 3.3. Parameters

The initial Technology Stage of technology  $j$  for process  $i$  is denoted by the parameter  $ITS_{i,j}$ . The resources required for developing technologies are represented by set  $r \in R$ . The amount of resource  $r$  necessary for technology  $j$  to be developed to Technology Stage  $s$  for process  $i$  is denoted by parameter  $Q_{r,i,j,s}$ . These resources should be available for  $R_{i,j,s}$  periods for successful progression to stage  $s$ . The amount of resource  $r$  available at time  $t$  is denoted by parameter  $QT_{r,t}$ . The cost of resources at each period  $t$  is denoted  $QC_{r,t}$ .

### 3.4. Constraints and Other Variables

The Technology Stage of technology  $j$  for process  $i$  during time  $t$  is represented by the variable  $TS_{i,j,t}$ . It is computed from the binary variable  $x_{i,j,t,s}$ . Another variable called Technology Integration Level ( $TIL_{i,j,t}$ ) is introduced to incorporate a commercially developed technology  $j$  into the network for process  $i$ . It is a function of the technology's compatibility with the process ( $yc_{i,j}$ ) and its development stage for the process ( $TS_{i,j,t}$ ). If the final development stage is  $S$ , variable  $TS_{i,j,t}$  takes a value from  $1$  to  $S$ . Variable  $TIL_{i,j,t}$  is expressed in terms of  $TS_{i,j,t}$  and takes a value from  $1$  to  $S$  if the technology is compatible with the process ( $yc_{i,j}=1$ ), else  $TIL_{i,j,t}=0$  when  $yc_{i,j}=0$ . A disjunctive model is used to implement these conditions in Eq. (6). For technology  $j$  that is commercially developed for process  $i$  by time  $t$ ,  $TS_{i,j,t}=S$  and  $TIL_{i,j,t}=S$ , while for an incompatible technology for process  $i$ ,  $TS_{i,j,t}=1$  and  $TIL_{i,j,t}=0$ . Equation (7) ensures that only the technologies commercially developed for the process are implemented in process facility/expansion. The amount of each resource allocated for developing technology  $j$  for source  $i$  to stage  $s+1$  during time  $t$  is given by variable  $q_{r,i,j,t,s}$ . The cost parameter for the resources is  $QC_{r,t}$ . The development cost for technology  $j$  for process  $i$  at time  $t$  is computed by Eq. (8).

$$\left[ \begin{matrix} yc_{i,j} = 1 \\ TIL_{i,j,t} = TS_{i,j,t} = \sum_{s \in S} x_{i,j,t,s} \end{matrix} \right] \vee \left[ \begin{matrix} yc_{i,j} = 0 \\ TIL_{i,j,t} = 0 \end{matrix} \right] \quad \forall i \in I, j \in J, t \in T \quad (6)$$

$$TIL_{i,j,t} / S \geq y_{i,j,t,n} \quad \forall i \in I, j \in J, t \in T, n \in N \quad (7)$$

$$DC_{i,j,t} = \sum_{r \in R, s \in S} q_{r,i,j,t,s} QC_{r,t} \quad \forall i \in I, j \in J, t \in T, r \in R, s \in S \quad (8)$$

Equation (9) ensures that developing technology  $j$  is at least at stage  $s-1$  before it progresses to the subsequent stage  $s$ . Equation (10) computes the Technology Stage of technology  $j$  for process  $i$  at the first planning period using the initial Technology Stage parameter. The resource allocation for Technology Stage progression is represented with the disjunctive model in Eq. (11). The binary variable  $x'_{i,j,t,s}$  takes a value of  $1$  when all necessary resources for Technology Stage progression of technology  $j$  to stage  $s$  for process  $i$  are available at time  $t$ . Equation (12) enables the progression of a technology to



stage  $s$  at time  $t$  if the necessary resources are available throughout  $R_{i,j,s}$  by time  $t-1$ . A technology can progress to only one subsequent development stage for a process per period. Equation (13) ensures this condition by allowing resource allocation for only one Technology Stage progression of technology  $j$  for process  $i$  at time  $t$ . Resource allocation at every period ( $q_{r,i,j,t,s}$ ) is limited by resource availability at that period ( $QT_{r,t}$ ) in Eq. (14).

$$x_{i,j,t,s} \leq x_{i,j,t,s-1} \quad \forall i \in I, j \in J, t \in T, s \in S \quad (9)$$

$$TS_{i,j,1} = \sum_{s \in S} x_{i,j,1,s} = ITS_{i,j} \quad \forall i \in I, j \in J \quad (10)$$

$$\left[ \begin{array}{l} x'_{i,j,t,s} = 1 \\ q_{r,i,j,t,s} = Q_{r,i,j,s} \end{array} \right] V \left[ \begin{array}{l} x'_{i,j,t,s} = 0 \\ q_{r,i,j,t,s} = 0 \end{array} \right] \quad \forall r \in R, i \in I, j \in J, t \in T, s \in S \quad (11)$$

$$\sum_0^{t-1} x'_{i,j,t,s} / R_{i,j,s} \geq x_{i,j,t,s} \quad \forall i \in I, j \in J, t \in T, s \in S \quad (12)$$

$$\sum_{s \in S} x'_{i,j,t,s} \leq 1 \quad \forall i \in I, j \in J, t \in T \quad (13)$$

$$\sum_{i \in I, j \in J, s \in S} q_{r,i,j,t,s} \leq QT_{r,t} \quad \forall r \in R, t \in T \quad (14)$$

Technology  $j$  commercially developed for process  $i$  by time  $t$  ( $TIL_{i,j,t}=S$ ) can be implemented in its facility at time  $t$ . If it is implemented in facility  $n$ ,  $y_{i,j,t,n}=1$ . The facility capacity is denoted by variable  $c_{i,j,t,n}$ , and feed flow rate is denoted by variable  $f_{i,j,t,n}$ . The investment cost ( $CIC_{i,j,t,n}$ ) is a function of  $c_{i,j,t,n}$ , while the operating cost ( $COC_{i,j,t,n}$ ) is a function of  $f_{i,j,t,n}$ . If the technology is not implemented in the facility,  $y_{i,j,t,n}=0$ , and variables  $c_{i,j,t,n}$  and  $f_{i,j,t,n}$  become zero. These conditions are implemented with disjunctive model Eq. (15). Equations (16) and (17) ensure process  $i$  uses at most one technology in a processing facility, and the same technology is used in the subsequent periods.

$$\left[ \begin{array}{l} y_{i,j,t,n} = 1 \\ CIC_{i,j,t,n} = f(X_i, c_{i,j,t,n}) \\ COC_{i,j,t,n} = f(X_i, f_{i,j,t,n}) \end{array} \right] V \left[ \begin{array}{l} y_{i,j,t,n} = 0 \\ CIC_{i,j,t,n} = 0 \\ COC_{i,j,t,n} = 0 \end{array} \right] \quad \forall i \in I, j \in J, t \in T, n \in N \quad (15)$$

$$\sum_{j \in J} y_{i,j,t,n} \leq 1 \quad \forall i \in I, t \in T, n \in N \quad (16)$$

$$y_{i,j,t,n} \geq y_{i,j,t-1,n} \quad \forall i \in I, j \in J, t \in T, n \in N \quad (17)$$

#### 4. Results and Discussion

A case study of planning CC from 10 different sources over a planning period of 10 years (discretized into 10 equal time periods) to meet annual capture targets that increase biannually by 10% (from 30% to 70%) is considered. Post-combustion capture using absorption using aqueous monoethanolamine (ABS-MEA) is assumed to be commercially available, and pressure swing adsorption using methyl viologen exchanged zeolite Y (PSA-MVY) is under development. Previous studies have shown that PSA-MVY is more cost-effective than ABS-MEA for sources with CO<sub>2</sub> composition over 5% (Hasan et al. 2014, Santhamoorthy et al. 2023). The total CC cost should decrease if PSA-MVY is commercially developed during the planning period. The degree of the decrease would depend on when PSA-MVY becomes available and the emission sources using it. We demonstrate the model's capabilities by studying these factors.

Three Technology Stages, lab-scale R&D, pilot-scale testing, and commercialization, are considered. The pilot-scale plants are assumed to capture 0.1 million tons of CO<sub>2</sub> per

annum. The research funds and time for successful lab-scale R&D and pilot-scale testing are not specified as parameters in the model. Instead, it is assumed that each Technology Stage can take from one to ten years for successful testing. The total development time is the total time taken for testing these stages. The maximum research fund available is computed as the difference between total costs with and without the technology.

The emission sources data and the cost models and carbon composition limitations for the capture technologies from Hasan et al. (2014) are used to construct model parameters. The costs of pilot-scale plants are estimated by escalating the cost by a factor of 1.3 to account for the overhead costs associated with testing new technology. The nonlinear cost models for CC yielded a mixed-integer nonlinear programming model (MINLP). The solution approach developed by Santhamoorthy et al. (2023) is used to solve the problem to an optimality gap of 5%. It is a two-step approach that solves a relaxed MILP model to obtain a lower bound and an initialized MINLP model to obtain an upper bound. The steps are employed iteratively with an increasing number of segments in the relaxation step to generate tight bounds by updating the lower and upper bounds.

The optimization models are formulated in Python V3.8.6 using PYOMO V6.4.1. The MILP models are solved using CPLEX V20.10, and the initialized MINLP models are solved using DICOPT V2 through GAMS V24.8.5, all on an Intel Xeon Gold 6248R 3 GHz processor with 48 cores and utilizing a maximum of 50 GB RAM.

Different development times are assumed for PSA-MVY, and the model is solved to generate the results summarized in Figure 2. In Figure 2(a), the marker star shows the cases where there is a commercial development of PSA-MVY during the planning period. There is successful development for all cases with at most 8 years of total development time. The total CC planning cost increases as the total development time increases. Figures 2(b) and (c) show the number of sources using only PSA-MVY and both ABS-MEA and PSA-MVY for capture. More sources use only PSA-MVY for shorter total development times, leading to a lower total CC cost. As the development time increases, ABS-MEA is implemented for capture to meet the annual capture targets, while PSA-MVY is developed and implemented in an expansion facility over time. The number of sources using PSA-MVY is high when the total development time is at most 6 years. The number of cases in which different emission sources use either only PSA-MVY or PSA-MVY in the expansion capacity is shown in Figure 2(d). When the development time is longer for PSA-MVY, sources 1-4, having high emission rates, use it in the expansion capacity, while sources 5-7 still use only PSA-MVY for capture. Sources 8-10 do not use PSA-MVY since they have a very low CO<sub>2</sub> composition. The total cost of CC planning for 10 years using only ABS-MEA technology is \$13.9 billion. The total cost when PSA-MVY is developed in 2, 5, or 9 years is \$12.0, \$13.1, and \$13.7 billion, respectively. Thus, the funds that can be used towards initial R&D and lab-testing of PSA-MVY decrease as its development time prolongs.

## **5. Conclusion and Future Direction**

This paper developed a multi-period optimization model that assists the planning decisions related to technology development, implementation, and operation. A case study of CC demonstrated the capability of the model to minimize the total cost of capture planning while developing a novel technology for implementation in the network. The funds that can be allocated towards R&D of the new technology were determined based on its anticipated development time. The results revealed that some sources used only the new technology for capture despite its long development time, while others used it only

in the expansions. This work will be extended to study the effect of uncertainty in the successful progression of the technologies to subsequent development stages.

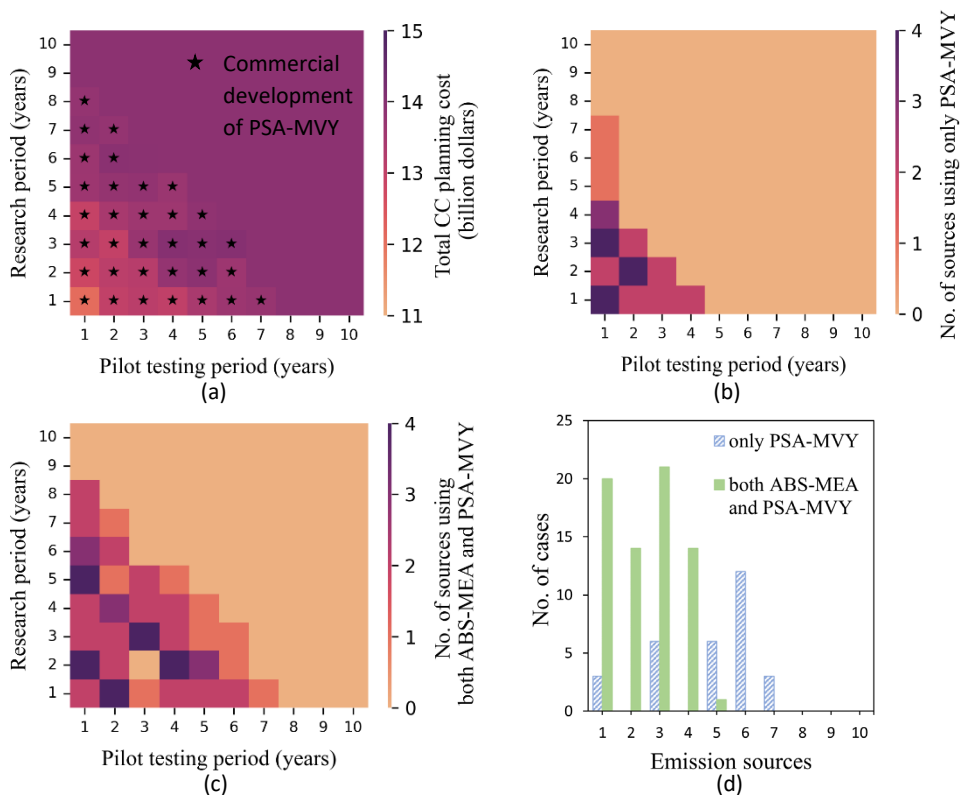


Figure 2. (a) Total cost of CC, (b) No. of sources using only PSA-MVY, and (c) No. of sources using both ABS-MEA and PSA-MVY for different cases of different development periods. (d) No. of cases in which sources use either only PSA-MVY or both ABS-MEA and PSA-MVY.

## Acknowledgment

The authors acknowledge the financial support from Chevron Technical Center.

## References

- Anderson Jr, E. G., & Joglekar, N. R. (2005). A hierarchical product development planning framework. *Production and Operations Management*, 14(3), 344-361.
- Davila, A., & Wouters, M. (2004). Designing cost-competitive technology products through cost management. *Accounting horizons*, 18(1), 13-26.
- Fahmi, I., Nuchitprasittichai, A., & Cremaschi, S. (2014). A new representation for modeling biomass to commodity chemicals development for chemical process industry. *Computers & chemical engineering*, 61, 77-89.
- Hasan, M. F., Boukouvala, F., First, E. L., & Floudas, C. A. (2014). Nationwide, regional, and statewide CO<sub>2</sub> capture, utilization, and sequestration supply chain network optimization. *Industrial & Engineering Chemistry Research*, 53(18), 7489-7506.
- Santhamoorthy, P., Subramani, H. J., & Cremaschi, S. (2023). A multiperiod model for portfolio optimization of carbon capture and utilization. *FOCAPO/CPC*.

# An Encrypted MPC Framework for Security to Cyber-Attacks

Atharva V. Suryavanshi<sup>a</sup>, Aisha Alnajdi<sup>b</sup>, Mohammed S. Alhajeri<sup>a</sup>,  
Fahim Abdullah<sup>a</sup>, Panagiotis D. Christofides<sup>a,b,\*</sup>

<sup>a</sup> *Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA, 90095-1592, USA*

<sup>b</sup> *Department of Electrical and Computer Engineering, University of California, Los Angeles, CA 90095-1592, USA*  
*pdc@seas.ucla.edu*

## Abstract

This work focuses on the design and development of a secure and private communication network using semi-homomorphic encryption to ensure cyber-security of model predictive control (MPC) systems. Specifically, Paillier encryption algorithm is used for encrypting-decrypting signals in the communication links. This implementation involves quantization of signals, which may result in significant errors. Thus, the MPC system is designed to ensure a certain degree of robustness with respect to quantization errors. The Encrypted Lyapunov-based MPC scheme is tuned to ensure that the calculations can be done with the available computational resources. Finally, the proposed design is implemented on a chemical process example.

**Keywords:** Model predictive control (MPC), cyber-security, encrypted control, semi-homomorphic encryption, quantization

## 1. Introduction

Cybersecurity of networked and distributed control systems has become crucial in recent years due to their increased vulnerability to targeted cyber-attacks. The integration of cloud-computing technologies with physical control systems has significantly increased process performance and decreased computational time, thus allowing real-time optimization. However, these advantages do not come without the increased threats of targeted cyber-attacks. Insecure communication channels are vulnerable to data-interception, eavesdropping on sensitive information, and data-manipulation. These cyber-attacks compromise the integrity, stability, and profitability of control systems. 77 cyber-security related incidents, on critical infrastructure, were reported from 2000 to 2019. Out of these reported incidents, a vast number of attacks were carried out on oil and energy sector industries (Parker et al., 2023). Clearly, cyber-attacks on these systems are extremely dangerous and hence it is of crucial importance to design and develop cyber-secure architectures for networked control systems.

In this work, we design and develop an Encrypted Lyapunov-based Model Predictive Control (LMPC) framework for nonlinear systems in which secure communication channels are established between sensor-controller and controller-actuator links. Specifically, Paillier cryptosystem is used to encrypt-decrypt sensitive state and input data in the network links. The designed controller is robust to the quantization errors arising as a consequence of encryption. The computational cost associated with different

quantization parameters is analyzed. Finally, the proposed framework is implemented on a multi-input, multi-output chemical process example.

## 2. Preliminaries

### 2.1. Class of Systems

In this work, we consider continuous-time nonlinear systems, described by the following nonlinear ordinary differential equation,

$$\dot{x}(t) = F(x(t), u(t)) = f(x) + g(x)u \quad (1)$$

where  $u^T = [u_1, u_2, u_3, \dots, u_m] \in \mathbb{R}^m$  is the manipulated input vector denoting  $m$  inputs and,  $x^T = [x_1, x_2, x_3, \dots, x_n] \in \mathbb{R}^n$  is the  $n$ -dimensional state vector. The manipulated inputs are constrained such that  $u \in U$ , where  $U := \{u \in U \mid u_{\min,i} \leq u_i \leq u_{\max,i}, \forall i = 1, 2, \dots, m\}$ , with  $u_{\min,i}$  and  $u_{\max,i}$  being the physical constraints on the  $i^{\text{th}}$  manipulated input.  $f(\cdot)$  and  $g(\cdot)$  are sufficiently smooth vector and matrix functions, respectively. The origin is a steady state of the nonlinear system and, without loss of generality, it is assumed that  $f(0) = 0$ .

### 2.2. Stabilizability Assumptions

We assume that there exists a feedback control law  $u = \Phi(x)$  that can render the steady-state of the nonlinear system of Eq. (1) exponentially stable in the sense that there exists a continuously differentiable control Lyapunov function  $V(x)$  such that the following inequalities hold for all  $x \in D$ , where  $D$  is an open neighborhood around the origin (Wu et al., 2019a):

$$c_1 |x|^2 \leq V(x) \leq c_2 |x|^2, \quad \frac{\partial V(x)}{\partial x} F(x, \Phi(x)) \leq -c_3 |x|^2, \quad \left| \frac{\partial V(x)}{\partial x} \right| \leq c_4 |x| \quad (2)$$

where  $c_1, c_2, c_3, c_4 > 0$ . The stabilizing controller can be constructed using the universal Sontag control law (Lin & Sontag, 1991). The closed-loop stability region for the nonlinear system of Eq. (1) is characterized as a level set of the Lyapunov function  $V(x)$ , and this closed-loop stability region is denoted as  $\Omega_\rho$ , where  $\Omega_\rho := \{x \in D \mid V(x) \leq \rho\}$ ,  $\rho > 0$ .

### 2.3. Paillier Cryptosystem and Quantization

In this paper, Paillier cryptosystem (Paillier, 1999) is used for encryption and decryption operations of state and input signals. Paillier cryptosystem is a semi-homomorphic encryption scheme in the sense that addition operations can be carried out in the encrypted space. The first step in the Paillier cryptosystem is to generate a public key  $(M, g)$  which is used for encrypting plaintext messages. The next step is to generate a private key  $(\lambda, u)$  that is used to decrypt ciphertexts. The detailed key generation steps can be found in Paillier (1999). The encryption-decryption operations in the Paillier cryptosystem are carried out on a subset of  $\mathbb{Z}_+$ . This subset is denoted as  $\mathbb{Z}_M$  and contains positive integers less than  $M$ . Thus, it is important to map the real-valued state and input data to the set  $\mathbb{Z}_M$  (Darup et al., 2017). Here, we consider signed-fixed point numbers in base 2, where  $l_1$  is the number of total bits and  $d$  is the number of fractional bits. In the first step, the function  $g_{l_1,d}(\cdot): \mathbb{R} \rightarrow \mathbb{Q}_{l_1,d}$ ,  $g_{l_1,d}(a) := \arg \min_{q \in \mathbb{Q}_{l_1,d}} |a - q|$  is used to map the real-valued data to the set  $\mathbb{Q}_{l_1,d}$ , where the elements of the set  $\mathbb{Q}_{l_1,d}$  are rational numbers from

$-2^{l_1-d-1}$  to  $2^{l_1-d-1} - 2^{-d}$  and are separated from each other by a resolution of  $2^{-d}$ . Next, we map the quantized signals to a subset of  $\mathbb{Z}_M$  using the bijective mapping  $f_{l_2,d}(\cdot) : \mathbb{Q}_{l_1,d} \rightarrow \mathbb{Z}_{2^{l_2}}$ ,  $f_{l_2,d}(q) := 2^d q \bmod 2^{l_2}$ . Once the quantized signals have been mapped to the set  $\mathbb{Z}_{2^{l_2}}$ , our data is ready for encryption. It is important to note that the encrypted data will be decrypted at the controller and at the actuator and, in that case, the resulting data will be in the set  $\mathbb{Z}_{2^{l_2}}$ . Hence, it is important to map the decrypted integer messages to the set of quantized signals,  $\mathbb{Q}_{l_1,d}$ . The inverse mapping  $f_{l_2,d}^{-1}$ , is defined as,

$$f_{l_2,d}^{-1}(\cdot) : \mathbb{Z}_{2^{l_2}} \rightarrow \mathbb{Q}_{l_1,d}, f_{l_2,d}^{-1}(m) := \frac{1}{2^d} \begin{cases} m - 2^{l_2} & \text{if } m \geq 2^{l_2-1} \\ m & \text{otherwise} \end{cases}.$$

### 3. Encrypted MPC Design

In this section, we propose an encrypted model predictive controller (MPC) scheme. In the closed-loop encrypted MPC design of Figure 1, state measurements  $x(t)$  are encrypted at the sensor, and the encrypted data is then sent to the MPC. At the controller, before performing the nonlinear optimization calculations, the encrypted state data is decrypted to obtain the quantized states  $\hat{x}(t)$ . These quantized states are provided as measurement-feedback to the MPC optimization problem. The optimal control inputs  $u(t)$  are calculated by solving an optimal control problem and are encrypted before being transmitted to the actuator. At the actuator, the encrypted control inputs data is decrypted to obtain the quantized control actions,  $\hat{u}(t)$ . These quantized control actions are applied on the nonlinear system.

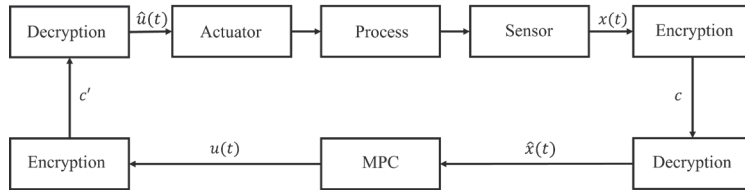


Figure 1: Schematic of nonlinear system under closed-loop encrypted MPC

In the closed-loop design of Figure 1, secure communication is established in the sensor-controller and controller-actuator links. However, two types of quantization errors arise. There is a loss of information in the sensor-controller link due to quantization of states, and in the controller-actuator link due to quantization of inputs. Based on the quantization mappings and the structure of the set  $\mathbb{Q}_{l_1,d}$ , these quantization errors are bounded such that:

$$|x(t) - \hat{x}(t)| \leq \eta_1 2^{-d} \quad \text{and} \quad |u(t) - \hat{u}(t)| \leq \eta_2 2^{-d} \quad (3)$$

where  $\eta_1, \eta_2 > 0$  and  $d$  is the number of fractional bits (quantization parameter). Thus, the larger the magnitude of  $d$ , the smaller are the quantization errors in the design.

#### 4. Encrypted Lyapunov-based Model Predictive Control

In this section, we present the mathematical formulation of an encrypted LMPC for the nonlinear system of Eq. (1), with secure controller-actuator and sensor-controller communication channels. The proposed formulation is as follows:

$$J = \min_{u \in S(\Delta)} \int_{t_k}^{t_k + N} L(\tilde{x}(t), u(t)) dt \quad (4a)$$

$$\text{s.t. } \dot{\tilde{x}}(t) = F(\tilde{x}(t), u(t)) = f(\tilde{x}) + g(\tilde{x})u \quad (4b)$$

$$u(t) \in U, \forall t \in [t_k, t_{k+N}) \quad (4c)$$

$$\tilde{x}(t_k) = \hat{x}(t_k) \quad (4d)$$

$$\dot{V}(\tilde{x}(t_k), u) \leq \dot{V}(\hat{x}(t_k), \phi(\hat{x}(t_k))), \text{ if } \hat{x}(t_k) \in \Omega_{\rho} \setminus \Omega_{\rho_{\min}} \quad (4e)$$

$$V(\tilde{x}(t)) \leq \rho_{\min}, \forall t \in [t_k, t_{k+N}), \text{ if } \tilde{x}(t_k) \in \Omega_{\rho_{\min}} \quad (4f)$$

where  $\tilde{x}(t)$  is the state trajectory predicted by the process model,  $U$  is the set of physical constraints on the control inputs, and  $\Omega_{\rho_{\min}}$  is the level-set of Lyapunov function  $V$  in which the system states are ultimately bounded. The sampling period is denoted by  $\Delta$ , and  $S(\Delta)$  is the set of piecewise constant functions with a period  $\Delta$ . The optimal control input sequence,  $u^*(t | t_k)$ , is calculated by solving the MPC optimal control problem, and the first control action from the sequence is applied on the process. In the encrypted LMPC formulation, the constraint of Eq. (4b) defines the process model, Eq. (4c) defines the physical constraints on the control inputs, and Eq. (4d) is the state measurement feedback. The constraint of Eq. (4e) drives the closed-loop system state to a small neighborhood around the origin, and the constraint of Eq. (4f) ensures that the system states remain ultimately bounded in a small neighborhood,  $\Omega_{\rho_{\min}}$  around the origin. Thus, starting with an initial state  $x_0 \in \Omega_{\rho}$ , the proposed Encrypted LMPC design based on the stabilizing control law  $u = \Phi(x)$  ensures that the states of the nonlinear system of Eq. (1) remain in the stability region  $\Omega_{\rho}$  and are ultimately driven to a small neighborhood around the origin  $\Omega_{\rho_{\min}}$ .

#### 5. Chemical Process Example

In this section, we implement the proposed Encrypted LMPC design to a chemical process, specifically the system from Wu et al. (2019b). In this process, a second-order, irreversible, exothermic reaction occurs in a non-isothermal CSTR. The reaction converts a reactant,  $A$ , to the product,  $B$ , via the reaction  $A \rightarrow B$ . A heating jacket is provided to the CSTR that supplies/removes heat at a rate  $Q$ . Using mass and energy balances, the system of ODEs that describe the nonlinear process can be written as

$$\frac{dC_A}{dt} = \frac{F}{V} (C_{A0} - C_A) - k_0 e^{\frac{-E}{RT}} C_A^2, \quad \frac{dT}{dt} = \frac{F}{V} (T_0 - T) + \frac{(-\Delta H)}{\rho_L C_p} k_0 e^{\frac{-E}{RT}} C_A^2 - \frac{Q}{\rho_L C_p V} \quad (5)$$

where  $C_A$  is the reactant concentration,  $T$  is the temperature inside the reactor,  $C_{A0}$  is the inlet feed concentration, and  $Q$  is the heat added/removed from the reactor. The control objective is to maintain the CSTR at the unstable steady state  $(C_{As}, T_s) = (1.95 \text{ kmol/m}^3, 402 \text{ K})$ . This is achieved by manipulating the control inputs

$\Delta C_{A0}$  and  $\Delta Q$ . The state and input vectors of the closed-loop system are defined in terms of their deviation from steady state values:  $x^T = [C_A - C_{As}, T - T_s]$  and  $u^T = [C_{A0} - C_{A0s}, Q - Q_s]$ . Based on the smallest and largest values of the permissible inputs and possible state values (in the operating region), we get the optimal value of  $l_1 - d$  to be 18. For closed-loop simulations under encrypted LMPC, we vary  $d$  from 1 to 7 in increments of 2, and  $l_1$  changes accordingly. For  $l_2$ , it is important to have  $l_2 > l_1$ , which makes it possible to map quantized numbers from the set  $\mathbb{Q}_{l_1, d}$  to the set  $\mathbb{Z}_M$ . Hence, we select the value  $l_2 = 29$ . The prediction horizon,  $N$ , for MPC is chosen to be 2. The implementation of the encrypted MPC is done in a sample-and-hold fashion with a receding horizon. A first-principles model for the nonlinear CSTR system of Eq. (5) is used in the MPC and the nonlinear optimization package IPOPT (Wächter & Biegler, 2006) is used to solve the MPC optimization problem. The Explicit Euler method is used to simulate the nonlinear system of Eq. (1) with an integration time step of  $h_c = 10^{-4}$  hr. The sampling period is  $\Delta = 10^{-2}$  hr. The closed-loop encrypted MPC simulation results are presented in Figures 2 to 4.

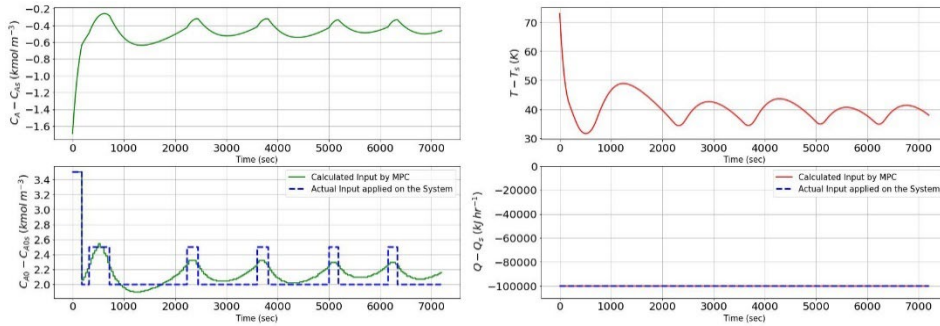


Figure 2: State and input profiles of closed-loop simulations under encrypted LMPC, with  $d = 1$ .

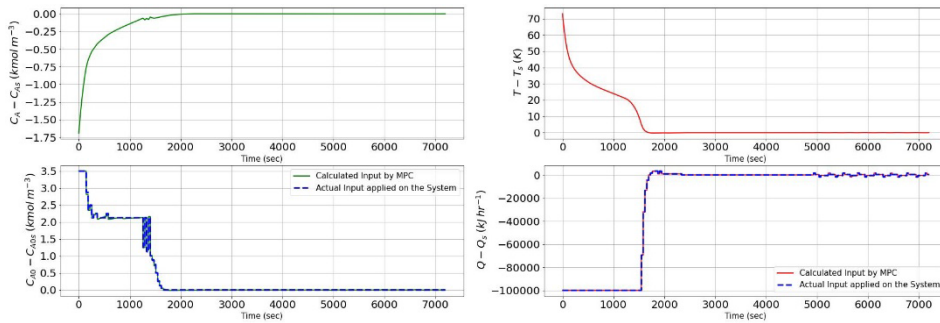


Figure 3: State and input profiles of closed-loop simulations under encrypted LMPC, with  $d = 3$ .

From the results of closed-loop encrypted MPC simulations, it can be observed that, for some small value of the quantization parameter  $d$ , the encrypted MPC is not able to stabilize the nonlinear system around a small neighborhood around the origin. Instead, we observe the oscillations of states around a point other than the unstable steady state.



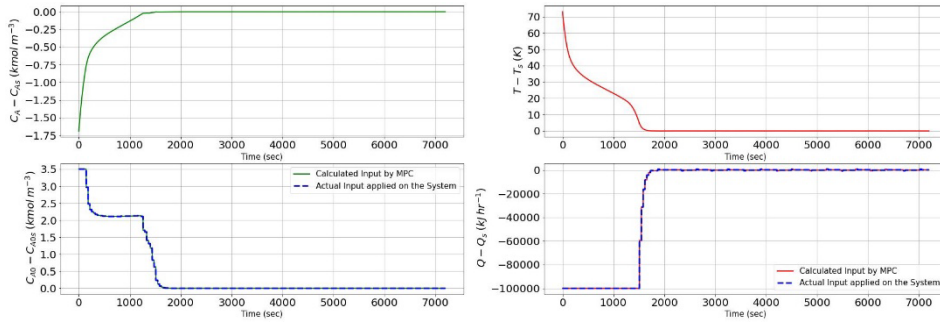


Figure 4: State and input profiles of closed-loop simulations under encrypted LMPC, with  $d = 7$ .

Thus, for systems being operated at an unstable equilibrium, it is possible that the encrypted MPC cannot practically stabilize the system for  $d \leq d_{\text{critical}}$ , and it is important to identify this critical value of the quantization parameter. For the nonlinear system of Eq. (1), we have  $d_{\text{critical}} = 1$ . Additionally, for  $d > d_{\text{critical}}$  as the value of the quantization parameter  $d$  increases, we see improvement in MPC performance in the sense that we achieve faster convergence of states to a small neighborhood,  $\Omega_{\rho_{\text{min}}}$ , around the origin. The MPC performance improved because the quantization error significantly decreases with the increase in  $d$ . However, this improvement in performance comes at a computational cost. The computational cost increases significantly with the increase in  $d$  primarily for two reasons. Firstly, with an increase in  $d$ , the number of elements in the set  $\mathcal{Q}_{l,d}$  increase and, hence, the computational cost required to construct such a set increases significantly. Secondly, with a large number of elements in the set  $\mathcal{Q}_{l,d}$ , the number of search operations required to map real valued signals from  $\mathbb{R}$  to  $\mathcal{Q}_{l,d}$  also increase significantly. Thus, if the computational resources are limited, the MPC performance that can be achieved is also limited but one must ensure that the chosen value of quantization parameter  $d$  is larger than the critical value,  $d_{\text{critical}}$ .

## References

- S. Parker, Z. Wu, and P. D. Christofides, 2023, "Cybersecurity in Process Control, Operations, and Supply Chain," *Proceedings of Foundations of Computer Aided Process Operations/Chemical Process Control*, San Antonio, Texas.
- Z. Wu, A. Tran, D. Rincon, and P. D. Christofides, 2019a, "Machine learning-based predictive control of nonlinear processes. part I: Theory," *AIChE Journal*, vol. 65, e16729.
- Z. Wu, A. Tran, D. Rincon, and P. D. Christofides, 2019b, "Machine learning-based predictive control of nonlinear processes. part II: Computational implementation," *AIChE Journal*, vol. 65, e16734.
- Y. Lin and E. D. Sontag, 1991, "A universal formula for stabilization with bounded controls," *Systems & Control Letters*, vol. 16, 6, pp. 393–397.
- P. Paillier, 1999, "Public-key cryptosystems based on composite degree residuosity classes," in *Proceedings of the International conference on the theory and applications of cryptographic techniques*. Berlin, Heidelberg: Springer, pp. 223–238.
- M. S. Darup, A. Redder, I. Shames, F. Farokhi, and D. Quevedo, 2017, "Towards encrypted MPC for linear constrained systems," *IEEE Control Systems Letters*, vol. 2, 2, pp. 195–200.
- A. Wächter, L. T. Biegler, 2006, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical Programming*, vol. 106, pp. 25–57.

# Machine learning-based product concentration estimation, real-time optimization, and multivariable control of an experimental electrochemical reactor

Berkay Çıtmacı<sup>a</sup>, Junwei Luo<sup>a</sup>, Joon Baek Jang<sup>a</sup>, Carlos Morales-Guio<sup>a</sup>, and Panagiotis D. Christofides<sup>a,b</sup>

<sup>a</sup>*Dept. of Chemical and Biomolecular Engineering, University of California, Los Angeles, Los Angeles 90095, USA.*

<sup>b</sup>*Dept. of Electrical and Computer Engineering, University of California, Los Angeles, Los Angeles 90095, USA; [pd@seas.ucla.edu](mailto:pd@seas.ucla.edu).*

## **Abstract**

Electrochemical reduction of CO<sub>2</sub> gas is a novel carbon capture technology that can potentially not only relieve the global climate crisis caused by increasing CO<sub>2</sub> gas emissions but can also restore electricity generated from clean energy sources in the form of carbon-based fuels. However, due to the complexity of the electrochemical reactions, there has been a limited effort to gain process control intuition for electrochemical reactors and the process development has been restricted to lab scale. This work aims to develop a multivariable control scheme for an experimental electrochemical reactor at UCLA by integrating artificial neural network modeling, real-time optimization, and process controller design. UCLA's experimental reactor is utilized to carry out the experiments required to construct a database to address the modeling, optimization and control tasks.

**Keywords:** Electrochemical CO<sub>2</sub> reduction; Multivariable Control; Experimental data modeling; Real-time optimization; Neural Networks.

## **1. Introduction and Reactor Description**

The percentage of renewable energy in the electricity grid has been increasing as a part of decarbonization efforts against the detrimental effects of global warming. With the current advancing awareness of sustainability, electrochemical methods have emerged as an attractive alternative to transform CO<sub>2</sub> gas into organic chemicals and synthetic fuels (De Luna et al. (2019)). However, scaling up the process to the industrial scale is limited by the lack of an effective control system and a solid understanding of the underlying physico-chemical phenomena of the process. To this end, a machine learning method is adopted to construct data-driven models mapping the critical input-output variables of the reactor and providing predictions to establish a feedback control scheme. As a continuation of the results from the previous study (Çıtmacı et al. (2022)) that focused on single-input single-output control, this work aims to develop an operational model of an electrochemical reactor with a nanoporous copper electrode using experimental data and perform multivariate control.

A gastight rotating cylinder electrode (RCE) cell reactor was used to carry out electrocatalytic experiments in this work. This electrochemical reactor consists of six major components: two electrode (working and counter electrode) chambers separated by an anion-exchange membrane, a potentiostat, a mass flow controller (MFC), a modulated speed rotator (MSR), and a temperature control block. During the operation, gas is

bubbled through 0.2 M  $KHCO_3$  electrolyte in both electrode chambers at a fixed flowrate of 20 sccm controlled by MFC. Nanoporous catalyst structure was synthesized directly on the copper cylinder electrode through the redox cycling in 4 mM KCl, and this roughened copper electrode was used as the cathode (working electrode) to electrochemically reduce  $CO_2$  into 17 different products. There are 5 gas-phase products ( $H_2$ ,  $CO$ ,  $CH_4$ ,  $C_2H_4$  and  $C_2H_6$ ) including hydrogen from the competing hydrogen evolution reaction and 12 liquid-phase products, which are produced on the electrode surface through multiple proton-electron transfer processes. Hydrodynamics and convective mass transport can be controlled systematically with the MSR, to and from the electrode for all species including reactant, intermediate, and final products. Under well-defined hydrodynamics, now the potential applied to the working electrode can be adjusted by the potentiostat, and this potential against the reference electrode and the electrical current are continuously measured and recorded. Therefore, by controlling the electrode rotation speed and the applied potential, the reaction kinetics and mass transfer effects of this process can be decoupled and studied, respectively. Lastly, the reactor is gastight such that online quantification of gas products can be performed using a gas chromatograph (GC). Liquid products, on the other hand, accumulate in the electrolyte solution and are analyzed using nuclear magnetic resonance (NMR) spectroscopy after each experiment.

## 2. Weighted-FNN Steady-State Modeling and Real-Time Optimization for Feasible Setpoints

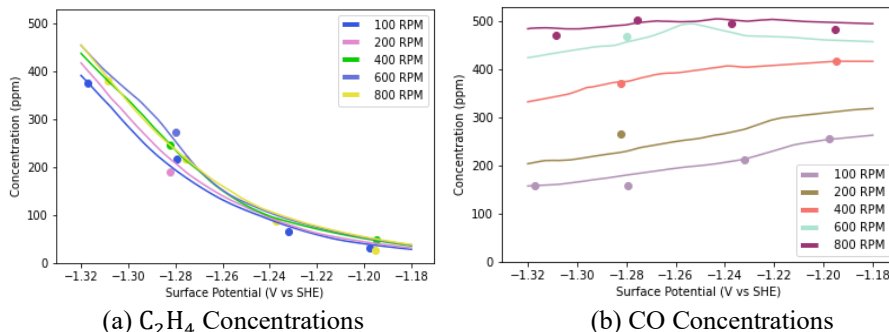
### 2.1. ML Model Trained with Statistical Weight

With the development of data and computation technologies, artificial neural networks (ANN), that have the ability to perform universal approximation for any linear/nonlinear relations (Csáji et al., (2001)), have become a popular method to address chemical and process control challenges. A feed-forward neural network (FNN) is trained with a database generated from open-loop and input step-change experiments to predict steady state concentrations of 17 products based on the applied potential and electrode rotation speed. The FNN model has one hidden layer that contains 64 neurons and a ReLu activation function. The Softplus function is selected as the activation function of the output layer to ensure non-negative prediction. Additionally, the Softplus function predicts the output with a smoother curve which aligns the physical expectation better than other activation functions with similar property such as ReLu. The mean squared error (MSE) function (Eq. 1) is utilized as the cost function to train the FNN model as:

$$L(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1), \quad L(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n \frac{1}{v_i^2} (y_i - \hat{y}_i)^2 \quad (2)$$

Although the trained FNN model can give accurate predictions (i.e., low MSE), it treats every data point in the dataset equally which increases the probability of overfitting high variance data points due to the stochastic nature of the electrochemical reaction and experimental uncertainties. To account for data variance, a weighted MSE function (Eq. 2) is adopted during the training process, where  $v_i$  is the coefficient of variance for the  $i^{th}$  data point. Specifically, data points are grouped by similar operating conditions such that averages and standard deviations under certain input conditions can be calculated and included in the weighted loss function. Therefore, the weighted-FNN is granted more tolerance of prediction error if the corresponding data point has a higher standard deviation. The weighted-FNN predictions for four gas products are represented in Fig. 1,

which displays the ability of FNN to not only provide accurate prediction based on given input conditions, but to also capture the relationship between inputs and outputs.



**Figure 1:** Weighted-FNN prediction of (a)  $C_2H_4$  and (b) CO concentrations under different electrode surface potential and rotation speed conditions.

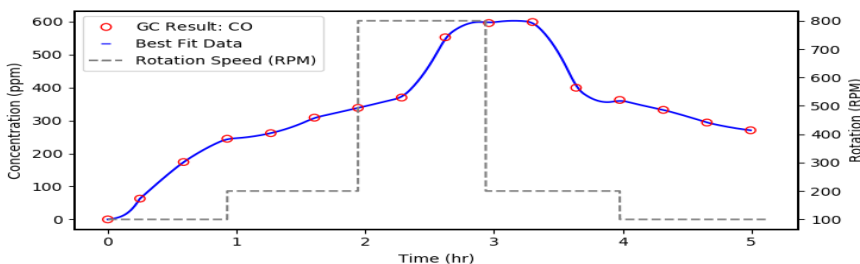
### 2.2. Real-Time Optimization

An optimization problem can be established based on the neural network model prediction to locate the optimum operational setpoint. Specifically, by assuming that electricity consumption is the only operating cost of this process, the optimum setpoint is where the economic profit of the operation is maximized. In this work, the concentrations of  $C_2H_4$  and CO are chosen to be the controlled variables (outputs), due to their strong correlations with surface potential and electrode rotation speed (inputs), respectively. A third-party software, IPOPT is utilized to continuously solve for the optimum setpoint based on the neural network prediction and varying electricity price.

## 3. Recurrent Neural Network Dynamic Modeling of $C_2H_4$ and CO Concentrations

### 3.1. Data Fitting for Experimental Dynamics Approximation

The experimental concentration data from the GC is discrete and deficient to train a neural network model due to the limitation of separation time of chromatography. To tackle this, polynomial and sigmoid best-fits were used to enhance the training data by approximating a probable experimental trajectory between GC results. This significantly increases the number of data points and enables the training of a neural network model. The polynomial and sigmoid best-fits interpolation is demonstrated in Fig. 2, where the original dataset contains only 16 data points collected from the GC readings. After data interpolation, the available concentration estimations increase to 18000.

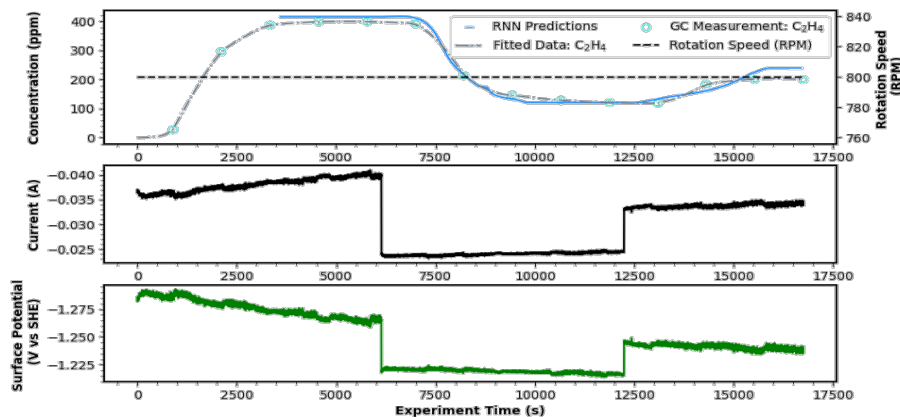


**Figure 2:** Example of GC data interpolation for a probable experimental trajectory from an input (rotation speed) step change experiment under constant applied potential.

### 3.2. Machine Learning Modeling of Process Dynamics

Recurrent neural network (RNN) is a popular structure to model the dynamic behaviors of chemical processes for its proven ability to fit time-series data effectively. Long-Short-Term-Memory (LSTM) models, designed to restore the exploding/vanishing gradients in basic RNN models, have a better performance in predicting time-series data and thus are used in our work as the alternative of explicit operation model in a feedback control loop (Wu et al. (2019)). Specifically, two LSTM models are trained using the enhanced dataset to predict the production rates of  $C_2H_4$  and CO respectively with three input features (i.e., surface potential, rotation speed, and current). The input data sequence contains 36 timestamps to represent a window of one-hour historical data. The time step size is chosen to be 100 seconds to reduce the computational cost. In addition, the model is trained to predict the next 800 seconds of the production rates, with 100 seconds time step. In this way, LSTM models can learn the delay response and process dead time as it is trained with a sequence of future data.

The LSTM model contains one hidden recurrent layer with 200 neurons. Simple grid search is performed to tune the hyper-parameters which demonstrates that the prediction accuracy decreases with decreasing the number of neurons and starts overfitting the data with more neurons or hidden layers. Regularization methods including recurrent dropout and  $L_2$  kernel regularization were performed to prevent overfitting further. A sigmoid function is used as the activation function of the output layer to ensure the model prediction is constrained by the maximum and minimum values in the training set. Finally, one experiment from constant input parameters and two step change experiments were used to test the model. The testing performance is shown in Fig. 3 where the predictions demonstrate a good overlap with the probable experimental trajectories.



**Figure 3:** LSTM predictions for  $C_2H_4$  production rates based on plotted input parameters, where RNN predictions are represented in continuous blue and probable experimental trajectory in dashed gray line.

## 4. Controller Tuning and Multivariable Control Architecture

### 4.1. Process Model Parameter Extraction Using Experimental Data

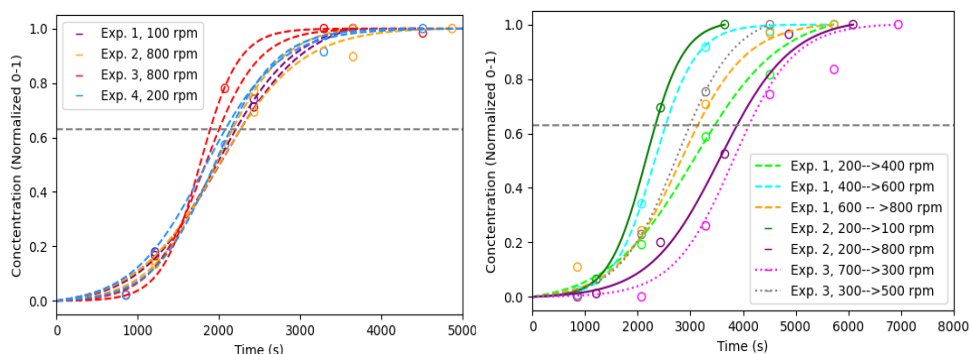
Two first-order plus time delay (FOPTD) models are used to describe the reaction processes, and the corresponding process parameters are extracted from experimental data. Specifically, various open-loop experiments were conducted by applying single-

variable step changes to either applied potential or rotation speed. Subsequently, the response in  $C_2H_4$  and CO concentrations, captured by the GC, are enhanced with a sigmoid function to configure a continuous trajectory of the real-time response as shown in Fig. 4. The responses are normalized for an easy comparison between various step changes. With step change responses, the dead time ( $\theta$ ) is extracted by drawing a tangent line through the inflection point, and time required to reach 63% of the final steady state is defined to be the time constant ( $\tau$ ). The mathematical expression of FOPTD model is demonstrated in Eq. 3, where  $K$  denotes the steady state gain.

$$G(s) = \frac{Ke^{-\theta}}{\tau s + 1} \quad (3)$$

However, a single FOPTD model is not enough to model the entire operation range due to the nonlinearity of the electrochemical  $CO_2$  reduction. To account for that, various FOPTD models are developed to describe different segment of the overall operational range (Fig. 5). We should check general correlations for the dead time and  $\tau$  for different step changes of rotation and potential. If there is a general correlation for the time parameters, it is possible to extract steady state gains from the weighted-FNN model. Fig. 4 is used to observe dynamic responses of  $C_2H_4$  under various rotation speed and potential changes.

The first plot clearly shows that the concentration change has similar dynamic behavior under constant rotation speed. Therefore, we can assume the same dead time and  $\tau$  for the tuning process. The second plot shows that similar dead times and  $\tau$  values can be attained for the rotation changes less than 200 RPM. These values are used in FOPTD model. A similar procedure is applied to develop the transfer function of CO concentration.



(a) Potential change under constant rotation speed (b) Rotation speed change under constant potential  
**Figure 4.** Sigmoid functions fitted to dynamic  $C_2H_4$  concentration change data, normalized between 0 and 1, to observe dead time and  $\tau$  values for FOPTD models

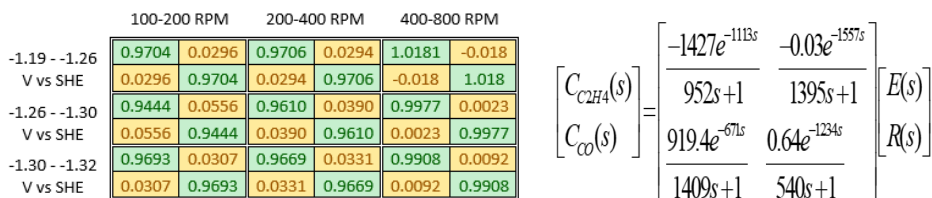
#### 4.2. Relative Gain Array

There are two options to implement the two-input-two-output control for this reactor, which are using 2 PI controllers or a model predictive controller (MPC). The control strategy is selected based on the individual manipulated variable-controlled variable correlations. Specifically, if there is no significant cross-coupling between the input-output pairs, it is feasible to implement the Multi-Input-Multi-Output (MIMO) control with 2 PI controllers. In the opposite case, an MPC should be considered to drive the system to the specific combination of parameters for the targeted set-points. A relative gain array (RGA) is useful to find if there are direct correlations between a specific input and output. The RGA only needs the steady state gains to be used. However, due to the

nonlinearities, we divide the operation range into 9 regions with respect to the input ranges as follows:

- Potential ranges = (-1.19, -1.26 V), (-1.26, -1.30 V), (-1.30, -1.32 V);
- Rotation ranges = (100-200 RPM), (200-400 RPM), (400-800 RPM).

These ranges are anticipated to satisfactorily linearize the process. The steady state gains are calculated from the weighted-FNN model. As a result, we found that rotation speed has a high correlation to CO concentration, whereas the surface potential is strongly correlated to  $C_2H_4$  concentration, which is shown in Fig. 5 that the diagonals of RGA are close to 1.



**Figure 5.** Relative gain array analysis for 100-200 RPM & -1.19 to -1.26 V vs SHE. (4) Multivariable control array example of FOPTD models

Therefore, two PI controllers are sufficient to control this process, in which one of the controllers will manipulate the applied potential to control  $C_2H_4$  production while the other controller will manipulate the rotation speed to control CO production rate. Furthermore, the arrays of FOPTD transfer functions will be used to determine the initial values of the PI tuning parameters using the Cohen-Coon tuning method.

## 5. Conclusion

This study carried out data-based modeling, set-point optimization and preliminary feedback control design using the open-loop experiments conducted using UCLA's experimental electrochemical reactor. Various data fitting methodologies were used to increase the size of the dataset for dynamic ML model development and extract feedback control parameters. Using the calculations presented in this paper, the closed-loop experiments will soon be conducted to drive the process between optimized set-points for  $C_2H_4$  and CO production rates.

## References

- [1] De Luna, P., Hahn, C., Higgins, D., Jaffer, S.A., Jaramillo, T.F., Sargent, E.H., 2019. What would it take for renewably powered electrosynthesis to displace petrochemical processes? *Science* 364, 3506.
- [2] Jeng, E., Qi, Z., Kashi, A.R., Hunegnaw, S., Huo, Z., Miller, J.S., BayuAji, L.B., Ko, B.H., Shin, H., Ma, S., Kuhl, K.P., Jiao, F., Biener, J., 2022. Scalable gas diffusion electrode fabrication for electrochemical  $CO_2$  reduction using physical vapor deposition methods. *ACS Appl. Mater. Interfaces* 14, 7731–7740
- [3] Çitmacı, B., J. Luo, J. B. Jang, V. Canuso, D. Richard, Y. M. Ren, C. G. Morales-Guio, and P. D. Christofides (2022). Machine learning-based ethylene concentration estimation, real-time optimization and feedback control of an experimental electrochemical reactor. *Chemical Engineering Research and Design* 185, 87–107.
- [4] Csáji, Balázs Csanád. Approximation with artificial neural networks. *Faculty of Sciences, Eötvös Loránd University, Hungary* 24, no. 48 (2001): 7.
- [5] Wu, Z., Tran, A., Rincon, D., Christofides, P.D. (2019). Machine learning-based predictive control of nonlinear processes Part I: Theory. *AIChE Journal* 65, e1672

# Symbolic regression based interpretable data-driven soft-sensor for process quality control

Harry Kay<sup>a</sup>, Sam Kay<sup>a</sup>, Max Mowbray<sup>a</sup>, Amanda Lane<sup>b</sup>, Cesar Mendoza<sup>b</sup>,  
Philip Martin<sup>a</sup>, Dongda Zhang<sup>a,\*</sup>

<sup>a</sup>*Department of Chemical Engineering, University of Manchester, Oxford Road, Manchester, M1 3AL, UK.*

<sup>b</sup>*Unilever Research Port Sunlight, Quarry Rd East, Bebington, CH63 3JW, UK.*

\*: [dongda.zhang@manchester.ac.uk](mailto:dongda.zhang@manchester.ac.uk);

## Abstract

Accurately predicting key performance indicators for product quality control is a research priority within the process industry. Physical models capable of predicting real-time process indicators of complex formulations are traditionally difficult to construct or retrofit, especially when the process is operated under dynamic process conditions and incorporates complex mixing phenomena. A solution to this employs the use of machine learning models to develop a soft-sensor using the information rich process data available, however this method lacks interpretability as it is difficult to extract physical information from the machine learning models, and the models are less user-friendly to implement within an existing factory advance process control system. To fill this gap, in this study we propose the use of symbolic regression, an interpretable machine learning technique, which aims to search the space over a subset of pre-defined mathematical functions to make accurate predictions of product quality. The high dimensional historical industrial data is first screened, and partial least squares is used to extract critical information regarding important operating regions and process parameters. Statistical features of critical process parameters are then employed within different symbolic regression based frameworks to obtain suitable models for batch quality prediction. Moreover, Monte Carlo sampling is adopted to estimate model uncertainty and evaluate reliability of the constructed interpretable models. To evaluate the model's performance, predictions were made for batches operated over different seasons and formulations. In addition, these models are benchmarked against a recently developed heteroscedastic noise neural network based data-driven soft-sensor. It is found that symbolic regression provided expressions with high accuracy, robustness, and responsiveness to changing process data suggesting that information about the true underlying correlation between the process data and a key performance indicator has been extracted. This research highlights the great potential of interpretable machine learning techniques for future industrial application.

**Keywords:** Machine learning, data analytics, quality control, symbolic regression, uncertainty estimation.

## 1. Introduction

Historically, mathematical models have been formulated to express relationships describing physical phenomena in order to evaluate functions at both current and future states. The predictive models were derived from first principles using conservation laws, experimental observations, and physical knowledge. However, many industrial processes contain too complex of interactions and phenomena to be described effectively via such



methods. Within Industry 4.0, access to information rich data has become more common due to the large quantity of sensor data being recorded from the industrial processes. The paradigm of machine learning can be exploited with the intention to analyse complex process data and formulate equations to characterise nonlinear and dynamic systems.

Recent works have explored the use of parametric and non-parametric data-driven models as a soft-sensor for process predictions (Kay et al., 2022; Mowbray et al., 2022). These models have proven to mitigate false confidence, have high accuracy, make uncertainty estimations representative of the data, and generalise well. However, a problem associated with such black box machine learning models is the lack of interpretability. To resolve the issue, symbolic regression can be integrated with real process data to generate statistical models representative of the industrial process. In essence, symbolic regression attempts to infer both a model and its parameters such that, for a given set of input and output data  $\{(x_i, y_i)\}_{i=1}^n$  for  $\{(x, y) \in \mathbb{R}\}$ , a function  $f$  with parameters  $\theta$  can be obtained where  $y \approx f(x|\theta)$ . A problem associated with this method is the that the search space increases exponentially with an increasing domain of operators. Knowledge of the system can be integrated to reduce this search space and reduce the dimensionality of the problem. A further complex challenge is the compromise between complexity and accuracy. Parsimonious expressions are traditionally preferred as they hold more relevance to providing understanding of the system at hand. In addition to this, more complex expressions are less robust to process changes and perform worse at extrapolation as they tend to overfit to the training data. Other works have explored the development of symbolic regression models in conjunction with neural networks to overcome the limitations of the combinatorial nature of genetic (Kim et al., 2020; Zhang et al., 2022). However, little attempts have been made to explore the use of symbolic regression in the context of soft sensing and monitoring for batch process.

## 2. Problem statement

This study focuses on the development of a statistical model based soft sensor and attempts to explore the feasibility of using symbolic regression to extract symbolic equations from a non-Newtonian fluid in a batch mixing process. Three datasets are available with the following labels: Alpha, Beta, and Gamma. Alpha and Beta were recorded from the same process line at different seasons of the year, and the Gamma data corresponds to a different formulation obtained from a similar process. We can assume process data of the form  $D = \{X, y\}$  with the predictor matrix,  $X$  being configured with  $N$  batches,  $T$  timesteps, and  $J$  sensors for  $X \in \mathbb{R}^{N \times T \times J}$ . The subsequent response matrix contains a single process measurement per batch, expressed as  $y \in \mathbb{R}^{N \times n_p}$  where  $n_p$  represents the number of process qualities measured. To ensure generalisability within the models, the data must show uniformity, thus all sensors must be shared, so  $J = 28$  for all datasets. To further ensure uniformity, an indicator depicting the beginning of a new phase of operation is used for alignment of batches. Dataset Alpha is characterised by  $N = 25$ ,  $J = 28$  process sensors, and  $T = 8000$  timesteps. Dataset Beta uses the same formulation and process equipment as Alpha in a different season with dimensionality of  $N = 14$  batches,  $J = 28$  sensors, and  $T = 3600$  timesteps. The Gamma data is comprised of  $N = 9$  batches,  $J = 28$  sensors, and  $T = 3700$  timesteps. Although, two different product formulation are tested, the target product quality of both is identical.

Typical problems associated with high dimensional datasets (Min, 2005) can be solved using dimensionality reduction, in which the data is reduced to remove redundant features and multi-collinearity, whilst still retaining sufficient information for quality prediction. We use the same approach as previous work (Hicks et al., 2021) where partial least squares (PLS) is employed, identifying 8 sensors, and 300 timesteps for accurately inferring batch quality predictions. Due to the poor capability of genetic algorithms for handling large datasets, the data is further reduced using timewise unfolding as mentioned in (Hicks et al., 2021) and taking either the mean and standard deviations or the principal components (PCA) of each individual sensor within the critical time region. This results in a 2-rank matrix of dimensions  $\mathbf{X} \in \mathbb{R}^{N \times (16 \times 28)}$ . This approach is adopted to preserve the ability of the symbolic regression models to generalise between the different datasets.

### 3. Methodology

First, we will introduce the concept of genetic algorithms and their function in symbolic regression. The principle of genetic algorithms follows Darwin’s theory of survival of the fittest and through this, populations of symbolic equations can be bred (Koza, 1994) with the objective of fitting a predictor matrix to the response matrix. Within this case study, the individuals of the population are combinations of operators, functions, and numerical constants where the operator domain is restricted to  $(+, -, *, \div, \sqrt{\quad}, \wedge, \exp, \ln)$ . An example of an individual is shown in tree diagram form in figure 1.

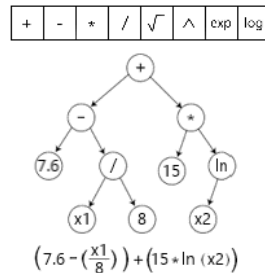


Figure 1: Tree diagram representation of an individual and its operator domain

Different metrics can be used to assess the generated model’s performance including its accuracy and complexity. In this study, the accuracy is defined as the average percentage error (1) and the loss function as the mean squared error (2) such that, given a functional representation,  $f$ , with input data  $\mathbf{x}$ , parameters  $\theta$ , and predictions  $\mathbf{y}$ :

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{(y_i - f(x_i|\theta))}{y_i} \cdot 100 \text{ and } MSE = \frac{1}{N} \sum_{i=1}^N (y_i - f(x_i|\theta))^2. \quad 1$$

The complexity is defined as having a linear dependency to the sum of operators, constants and process variables used within the model structure. This dependency can be altered by assigning unique weightings to specific operators or numerical constants, however, for this study, a value of 1 was assigned to all operators and numerical constants. Let us now discuss the procedure to assess the uncertainty of the statistical models generated via symbolic regression. In essence, we will exploit the fact that, under certain conditions, the covariance matrix is equal to the inverse of the information matrix. The negative log likelihood is employed as the objective function for the optimisation of model parameters and if we consider a gaussian random vector,  $\mathbf{x}$  with mean  $\mathbf{x}^*$  and covariance matrix  $E_\theta$ , the objective function is (Hirschfeld et al., 2020):

$$obj(x) = \operatorname{argmin} \left( \frac{N}{2} (\log 2\pi\sigma^2) + \frac{1}{2} \frac{(\mathbf{x} - \mathbf{x}^*)^T (\mathbf{x} - \mathbf{x}^*)}{\sigma^2} \right) \quad 2$$

Given that the second derivative of the negative log likelihood objective function is independent of  $\mathbf{x}$ , the  $i$ th and  $j$ th components of the hessian matrix  $H$  can be obtained without knowing  $\mathbf{x}^*$ . The parameter variance can be extracted from the diagonal entries of the covariance matrix and used to estimate the model uncertainty. A gaussian distribution of parameters with mean and standard deviation  $N(\theta, \sigma_\theta^2)$  can be sampled from via Monte Carlo sampling, and the statistical models queried to produce a normal distribution of batch quality predictions,  $N(f(x_i|\theta), \sigma_f^2(x))$ . The requirements of the industrial collaborator specify that the  $MAPE < 15\%$  for soft sensor prediction and the measurement uncertainty of the equipment was quantified as being approximately 10% per standard deviation. All statistical models were optimised through using gradient based optimisation techniques in Pyomo, namely IPOPT, allowing for efficient exploration and exploitation of the parameter space such that high-quality solutions were ensured.

#### 4. Results and Discussion

The data is split into two sets, training and testing, where the training set comprises of alpha and beta, and the model's prediction performance is tested using the gamma data. To reduce the dimensionality of the feature space and thus improve the quality of the solutions, gradient boosting was applied. Using this, feature importance charts were generated for selection of process variables to use for training within the symbolic regression models as shown in Figure 2.

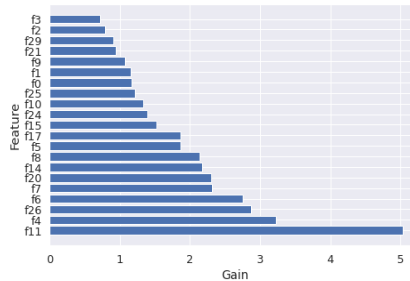


Figure 2: Feature importance chart generated using gradient boosting methods

The extracted symbolic equations that exhibit low complexity and high accuracy are presented in Table 1.

Table 1: MAPE and complexity values of generated symbolic expressions. Models 1-3 are constructed using the data's mean ( $X_{10}$ ) and std ( $X_{11}$ ). Model 4 is constructed using principal components ( $X_{10}$ )

Model	Symbolic expression	MAPE (Training) / %	MAPE (Testing) / %	Complexity
1	$\operatorname{Aln}( BX_{10} + C )$	9.23	13.89	6
2	$\operatorname{Aln}( \frac{BX_{10} + C}{DX_{11}^2 + E} )$	8.53	12.56	12
3	$Ae^{-BX_{11}} \operatorname{Aln}( \frac{CX_{10} + D}{EX_{11}^2 + F} )$	8.44	13.03	15

4	$Ax_{10} + B$	10.87	10.89	4
---	---------------	-------	-------	---

Models 2 and 4 were selected for further analysis due to enhanced performance within their respective groups. The predictions from models 2 and 4 on the training and validation sets are shown below in Figures 2a and 2b after being normalised.

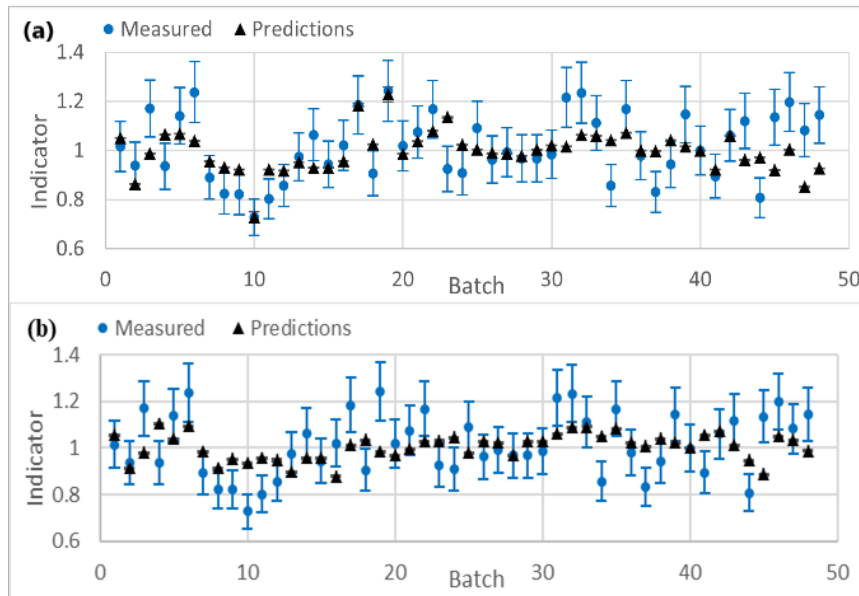


Figure 3: Soft sensor predictions against measured data for the alpha, beta and gamma for models 2 (a) and 4 (b). The error bars represent one standard deviation.

In Figure 2, the alpha, beta and gamma data are shown as batches 1-25, 26-39, and 40-49 respectively. The soft-sensors for both models 2 and 4 are shown to be capable of predicting the training and validation batches with high accuracy with training MAPE's of 8.53% and 10.87%, and validation MAPE's of 12.56% and 10.89% respectively. Notably, the uncertainty estimations of the generated models are far lower than the measured uncertainty so does not provide an accurate representation of the data uncertainty. This is predominantly due to the simple structure used to approximate the required performance indicator. All the statistical model predictions are within three standard deviations of the measured indicator so lie within an acceptable range of values. Model 2 provides a higher training accuracy than its validation error, indicating that some overfitting to the training data has occurred, however the model seems to be more responsive to process variation within the data as the predictions cover a larger domain. Model 4 provides almost identical errors for training and validation, indicating that enough information has been captured from the training set to prevent underfitting and make informed inferences of product quality without overfitting to batch specific trends.

Overall, both models have shown good capacity for predicting on data derived from different process lines, showing that the symbolic regression technique has successfully learned correlations between the process data and batch quality.

A further conclusion to draw is that it is possible to represent the process accurately using minimal features and process sensors. Both models required the use of one shared key process variable, with model 4 only using one principal component ( $X_{10}$ ) from the sensor and model 2 using the mean ( $X_{10}$ ) and standard deviation ( $X_{11}$ ). This is indicative of the sensors individual importance to inferring batch quality and for process monitoring. This conclusion corroborates the results obtained from the generation of importance charts using gradient boosting where the same feature was identified as being highly influential.

#### 4. Conclusion

In conclusion, symbolic regression proved to be a useful tool for developing accurate, interpretable statistical models for predicting process indicators. In addition, the expressions generated were simple and could easily be integrated into an industrial process control system to assist operators with process monitoring and control presenting a substantial benefit in comparison to traditional black box machine learning techniques which are difficult to extract knowledge from. Through the use of information rich statistical features, derived from industrial data, and data driven models, it was shown possible to identify high quality solutions at low computational cost. Upon validation of the optimised statistical expressions, using the training and testing data, the soft-sensor was shown to exhibit good predictive performance and generalisation capabilities. It was also seen to be able to capture relatively large process variation within the industrial data and filter out noise. This study has presented the potential of interpretable machine learning models for use in batch quality prediction in industrial processes.

#### References:

- Hicks, A., Johnston, M., Mowbray, M., Barton, M., Lane, A., Mendoza, C., Martin, P., & Zhang, D. (2021). A two-step multivariate statistical learning approach for batch process soft sensing. *Digital Chemical Engineering, 1*, 100003.
- Hirschfeld, L., Swanson, K., Yang, K., Barzilay, R., & Coley, C. W. (2020). Uncertainty Quantification Using Neural Networks for Molecular Property Prediction. *J. Chem. Inf. Model, 2020*, 3770–3780.
- Kay, S., Kay, H., Mowbray, M., Lane, A., Mendoza, C., Martin, P., & Zhang, D. (2022). Integrating Autoencoder and Heteroscedastic Noise Neural Networks for the Batch Process Soft-Sensor Design. *Industrial and Engineering Chemistry Research, 61*(36), 13559–13569.
- Kim, S., Lu, P., Mukherjee, S., Gilbert, M., Jing, L., Ceperi'c, V. C., & Soljači'c, M. S. (2020). *Integration of Neural Network-Based Symbolic Regression in Deep Learning for Scientific Discovery*.
- Koza, J. R. (1994). Genetic programming as a means for programming computers by natural selection. *Statistics and Computing, 4*, 87–112.
- Min, R. (2005). *A Non-linear Dimensionality Reduction Method for Improving Nearest Neighbour Classification*.
- Mowbray, M., Kay, H., Kay, S., Caetano, P. C., Hicks, A., Mendoza, C., Lane, A., Martin, P., & Zhang, D. (2022). Probabilistic machine learning based soft-sensors for product quality prediction in batch processes. *Chemometrics and Intelligent Laboratory Systems, 228*.
- Zhang, M., Kim, S., Lu, P. Y., & Soljači'c, M. S. (2022). *Deep Learning and Symbolic Regression for Discovering Parametric Equations*.

# Oligomerization Using Reactive Distillation, Design and Model Predictive Control

Gabriel Contreras-Zarazúa<sup>1, a, b</sup> Juan José Quiroz Ramírez,<sup>a</sup> Eduardo Sánchez-Ramírez,<sup>b</sup> Esteban Abelardo Hernández-Vargas,<sup>c</sup> Juan Gabriel Segovia-Hernández<sup>b</sup>

<sup>a</sup>*Country CONACyT- CIATEC, Center for Applied Innovation in Competitive Technologies, Leon Guanajuato*

<sup>b</sup>*Department of Chemical Engineering University of Guanajuato, Noria Alta S/N, 36000, Guanajuato, Gto., Mexico.*

<sup>c</sup>*Department of Mathematics and Statistical Science University of Idaho, 875 Perimeter Drive, MS 1103 Moscow*

## Abstract

Biojet-fuel is a promising alternative to reduce emissions in the aviation industry. The alcohol to jet (ATJ) is considered one of the most attractive alternatives. However, the economic feasibility is an important drawback of this process. With this in mind, the process intensification offers an interesting alternative to improve the ATJ process. Nevertheless, in many cases, the intensification could lead to complex equipment that are difficult to operate. This work proposes a novel reactive catalytic distillation column to improve the oligomerization stage of ATJ process. This intensified process was designed using a sensitivity analysis. In order to determine the feasibility of operating this intensified column, a control study using PI controller and model predictive control was performed. The results showed that a reactive distillation column to intensify the oligomerization zone is feasible. This equipment only consumes 200kW. Finally, the control studies indicated that it is possible to operate the column using a traditional feedback control and model predictive control techniques with good performance.

**Keywords:** Biojet-fuel, reactive distillation, model predictive control, process intensification.

## 1. Introduction

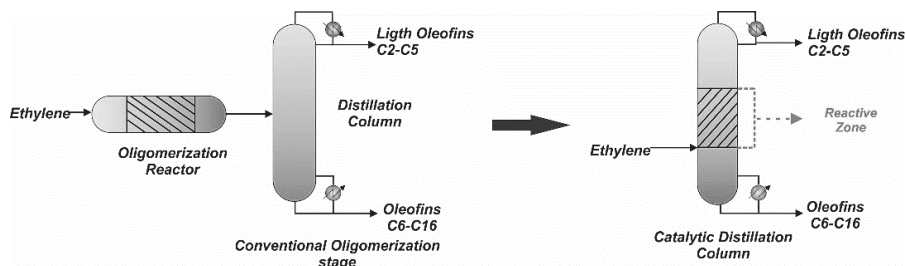
Energy has an utmost important role in our society. Energy is used to produce almost all the necessary commodities such as food, polymers, fertilizers. This energy employed to produce these commodities has several forms such as electricity, natural gas, liquid fuels, among others. In this sense, one sector that is essential for our society and which is characterized by an intensive use of energy is the transport sector. This sector consumes around 21% of the total energy demand annually, and its energy consumption is expected to increase up to 130% by 2050 which also represents an increase in CO<sub>2</sub> emissions. Among all the transport means, the aviation sector is expected to grow the fastest, doubling its energy consumptions and CO<sub>2</sub> emission in the next 20 years. For this reason, the aviation industry has proposed to reduce its CO<sub>2</sub> emissions in order to achieve a more sustainable aviation industry. The use of alternative biofuels for replacing oil-derived fuels is considered by the International Air Transport Association (IATA) as the most interesting alternative to reduce emissions (Rivas-Interian et al., 2022). These biofuels are commonly called biojet-fuels and can be produced from biomass. Nowadays, there

are several routes to produce biojet-fuel such as Fischer-Tropsch (FT-SPK), Hydroprocessed Esters and Fatty Acids (HEFA-SKP), Fisher-Tropsch with Aromatics (FT-SPK/A) and Alcohol to Jet (ATJ-SPK). In this sense, the ATJ process is considered one of the most attractive, since the alcohols are produced in abundance, offering logistical flexibility. The ATJ process consist of three reaction steps: alcohol dehydration, olefin oligomerization and hydrogenation. In this regard, oligomerization is a common stage in many jet fuel production routes, such as ATJ or aqueous phase processing. This stage is utmost importance, because it allows to achieve hydrocarbons with the exact number of carbons required by the aviation industry. This stage has two steps: a reaction stage in which ethylene is converted into larger hydrocarbons and a separation stage, in which the heavy hydrocarbons are separated from the lighter ones. This process stage is characterized by an intense use of energy, both in the reaction stage and in the separation. Therefore, one way to reduce energy consumption and increase the profitability of the ATJ process is through process intensification using reactive distillation.

Nowadays, the reactive distillation has not been explored as alternative to intensify the oligomerization zone. With this mind, this work proposes the design of a novel reactive distillation column for the oligomerization stage of ATJ process. The design parameters and operating conditions of the column were selected to increase the hydrocarbon production required for aviation industry; these hydrocarbons are in the range of C8 to C14. Since this equipment integrates the reaction and separation in a single unit and many reactions are performed inside it, a control study was carried out to determine the operation and control feasibility of this column. The control study consists of two parts: a study using PI controllers, in order to analyze the column control with common controllers in the industry. The second part is a model predictive control study in order to compare its performance with traditional feedback control using PI controllers. The integral absolute error (IAE) was selected as metric to quantify and compare the closed-loop performance for both PI and MPC controllers.

## 2. Design of reactive distillation equipment

The ATJ process considered for this study uses ethanol as raw material. This ethanol is dehydrated and converted to ethylene during the dehydration stage. In this work, it is only considered the oligomerization stage thus the ethylene is the raw material considered to this stage. Additionally, it is assumed that only linear olefins are produced during the oligomerization, this assumption is realistic because the catalysts reported in previous literature produce linear olefins, also it is a common practice to incorporate an isomerization stage after oligomerization stage to generate branched hydrocarbons. On the other hand, in this work, a catalytic distillation column is used to intensify the oligomerization stage, it consists of a distillation column with an intermediate reactive zone where the catalyst is placed and the whole set of reactions take place. The separation stages are located below and above of this reaction zone. Figure 1 shows a scheme of the reactive distillation process. The reactions that occur in the reaction zone and their respective kinetic parameters were taken from Sánchez-Ramírez et al., 2022. It is important to highlight, that the kinetic model corresponds to a set of elementary reactions of second order. A ethylene feed flowrate of 2100 kg/hr was chosen according to the information provided by Rivas-Interian et al., 2022.

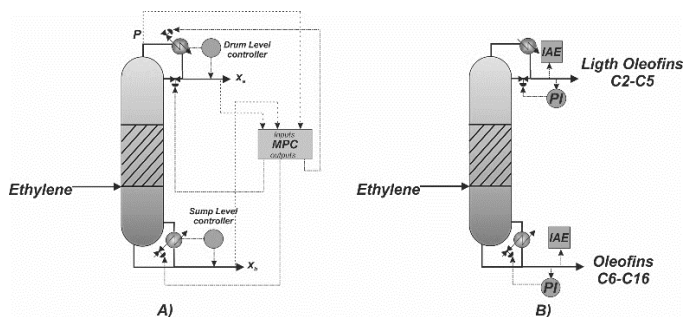


**Figure 1.** Intensification of oligomerization stage by catalytic reactive distillation.

The distillation column was designed and simulated in Aspen Plus using the RADFRAC module. The selection of thermodynamic model for the simulation was performed according to Carlson's algorithm (Carlson, 1996). Consequently, the Soave-Redlich-Kwong (SRK) was used as thermodynamic model to simulate the system, owing to the mixture contains linear hydrocarbons and the pressure required for the oligomerization is around 10–40bar. The design of the reactive distillation was performed using a sensitivity analysis in order to achieve a suitable hydrocarbon distribution and maintain low energy consumption. It is important to highlight, that according to Li et al., (2016), a suitable hydrocarbon distribution corresponds to a hydrocarbon mixture in the range of C10 to C16, being the C12 the most abundant compounds in the blend with the aim of ensuring the suitable physiochemical properties. Finally, in order to determine the improvements of intensification, the total annual costs of both the conventional and intensified oligomerization process were calculated using the Guthrie's method and data reported by Rivas-Interian et al., 2022.

### 3. Control of the reactive distillation column.

Reactive distillation columns are significantly more complex process due to integration of reaction and separation stages. The complexity of performing the separation and reaction processes in a single equipment could result in a process that is difficult to control and operate. Therefore, in order to evaluate the feasibility of operating this intensified process, a control study was performed. This control study considers the implementation of model predictive control (MPC) and compares this control strategy with traditional feedback control strategies using PI controllers. The control schemes using both MPC and PI controllers are shown in the Figure 2.



**Figure 2.** A) Control scheme using MPC B) Control scheme using PI controllers.

The integral absolute error (IAE) was chosen as metric to quantify and compare the closed-loop performance of the PI and MPC controllers. The tuning of controllers were



carried out using the minimization of IAE method, to achieve this, disturbances of -1% were performed on the manipulate variables of distillation column. Mathematically, the IAE can be defined as follows:

$$\text{Integral of Absolute Error (IAE)} = \int_0^{\infty} |y(t) - y_{sp}| dt \quad (1)$$

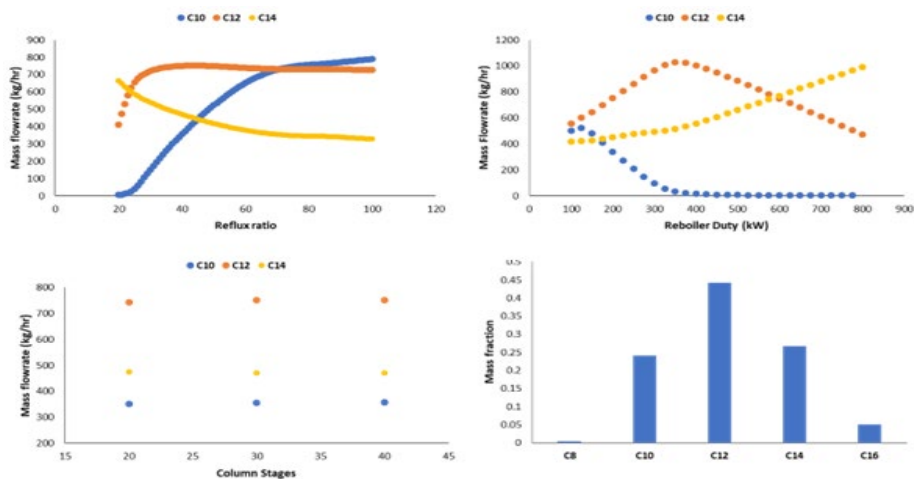
Where  $y_{sp}$  is the value of the set point for *the* control variables, whereas  $y(t)$  is the value of control variable at time  $t$ . It is important to mention, that the MPC controller uses a state space (SS) model as predictor model, this state space model was obtained by subspace identification. The subspace identification is a common method used to obtain a dynamic model when the process is very complex (Li et al., 2021). The rigorous simulation of Aspen Dynamics was used as plant model. The MPC control was implemented in a hybrid platform, which links Aspen Dynamics and the Simulink software. On the other hand, the feedback control using PI controllers was fully implemented in Aspen Dynamics. Disturbances of -1% on manipulate variables were done in order to compare the performance of control systems and to tune the controllers.

#### 4. Results

This section shows the design and control results of the catalytic column. As aforementioned, the design of reactive distillation column was performed by a sensitivity analysis in order to obtain a suitable olefin distribution. In this case, designs for 20, 30, and 40 stages were studied for the column. The results of the sensitivity analysis for the reflux ratio and reboiler duty as representative cases are shown in Figure 3. Please note, in Figure 3 as large values of reflux ratio have a positive effect on the production of C10 and C12, while more reflux ratio has a negative effect on C14. The C12 reaches a maximum concentration in reflux ratios close to 40, while for C10 and C14 the maximum concentrations are reached at higher and lower reflux ratios, respectively. Based on the results, it is concluded that a reflux ratio of 43 is the most adequate, because for this value, there is a suitable distribution of hydrocarbons, being C12 the most abundant and C10 and C14 olefins has a similar production. In addition, it was determined that a reboiler duty of 200kW is suitable since at this point the hydrocarbon distribution is very similar to that reported by Li et al., (2016). The sensitivity analysis was repeated for 30 and 40 stages finding that a greater number of stages does not benefit the production of more C10, C12 and C14. The Hydrocarbon distribution found with the best design is reported in Figure 3., also, some important design parameters are reported in Table 1. In the case of total annual costs (TAC), the results indicated that a conventional process has TAC \$582,204.90 \$/y. On the other hand, the intensified has a TAC of \$511,128.62\$/yr representing savings on costs around 13% in contrast to conventional process.

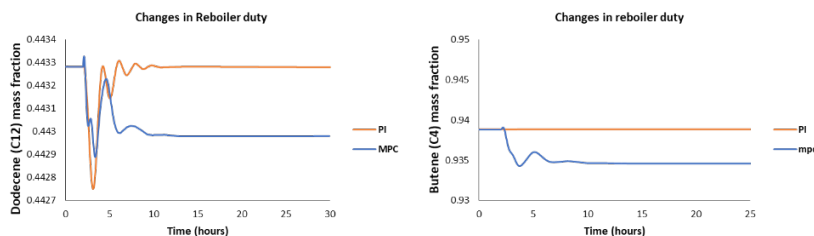
**Table 1.** Representative design parameters for the reactive distillation column.

Design Parameter	Value
Number of stages	20
Reboiler duty (kW)	200
Reflux ratio	43
Feed stage	7



**Figure 3.** Design results of reactive distillation.

With respect to control results, the composition of the most abundant compound in domes and bottoms were used as control variables. In this case, the butane (C4) and dodecane (C12) are de most abundant compounds in the top and bottom of the column, respectively. The composition of C4 was controlled with reflux ratio, while the composition of C12 is controlled using the reboiler duty of bottoms. Once identified the control variables, the tuning of controllers was performed. As aforementioned, the minimization of IAE method was used for tuning the controllers, in this case the parameters that minimize the IAE for the PI controllers at the top of column are 40 min for integral time and 20 for gain whereas the parameters for the PI controller in the bottom of column are 11.443 for gain and 70.99min integral time. In the case of MPC control, it is necessary only one controller to control the composition at the top and bottoms of the column. In this case, the parameters that minimize the IAE are a sample time of 3 min, control horizon of 10 intervals and prediction horizon of 50 intervals. Once the suitable parameters were determined, the control performance for both PI and MPC controllers are compared using their respective IAE values. The results of disturbing the reboiler duty as representative case are shown in Figure 4 whereas the IAE value is shown in Table 2



**Figure 4.** Disturbances of -1% in reboiler duty.

**Table 2.** IAE values for disturbances on reboiler duty

Control loop	C4 control loop (PI)	C4 control loop (MPC)	C12 control loop (PI)	C12 control loop (MPC)
Reboiler duty disturbance	2.86E-05	0.1091	0.0010875	0.00794

Based on the results, there is an appreciable difference between the two types of controllers. In this case, the butene composition can be controller better using a PI controller, having the PI controller up to 30% better performance than the MPC controller. However, for the dodecene control loop, the performance is very similar between both controllers. Note that the MCP control has the capability of controlling the process, however it do not fully fits to the setpoint, being the difference small due to the tolerance with which the objective function is solved in predictive control. However, it is concluded that this difference is not significant and both controllers are adequate to control the process.

## 5. Conclusions

In this work, a reactive distillation equipment was designed to intensify the oligomerization stage in the ATJ process. The intensified reactive distillation column has savings on total annuals costs of 13% in contrast to the conventional process. On the other hand, the control studies indicated that it is possible to operate this catalytic column using traditional control strategies, such as feedback control and other advanced control techniques, such as model predictive control. Finally, it is concluded that both controllers have similar performance, despite of this please note, that the PI controllers have a slight better performance due to they returned to exactly set-point value.

## References

- Rivas-Interian, Eduardo Sanchez-Ramirez, Quiroz-Ramírez, J.J., Segovia-Hernandez., J.G., 2022. Feedstock Planning and Optimization of a Sustainable Distributed Configuration Biorefinery for Biojet Fuel Production via ATJ Process. *Biofuels, Bioprod. Biorefining*.
- Sánchez-Ramírez, E., Huerta-Rosas, B., Quiroz-Ramírez, J.J., Suárez-Toriello, V.A., Contreras-Zarazua, G., Segovia-Hernández, J.G., 2022. Optimization-based framework for modeling and kinetic parameter estimation. *Chem. Eng. Res. Des.* 186, 647–660.
- Carlson, E.C., 1996. Succeeding at Simulation. *Chem. Eng. Prog.* 35–46.
- Li, J., Yang, G., Yoneyama, Y., Vitidsant, T., Tsubaki, N., 2016. Jet fuel synthesis via Fischer-Tropsch synthesis with varied 1-olefins as additives using Co/ZrO<sub>2</sub>-SiO<sub>2</sub> bimodal catalyst. *Fuel* 171, 159–166. <https://doi.org/10.1016/j.fuel.2015.12.062>

# Fault Diagnosis with GAT and PageRank: From Local Attention to Global Attention

Deyang Wu,<sup>a</sup> Jinsong Zhao,<sup>a,b\*</sup>

<sup>a</sup>State Key Laboratory of Chemical Engineering, Department of Chemical Engineering, Tsinghua University, Beijing 100084, China

<sup>b</sup>Beijing Key Laboratory of Industrial Big Data System and Application, Tsinghua University, Beijing 100084, China  
[jinsongzhao@tsinghua.edu.cn](mailto:jinsongzhao@tsinghua.edu.cn)

## Abstract

Fault detection and diagnosis (FDD) is important to ensure process safety in chemical plants. In the past decades, different FDD methods have been proposed. And data-driven methods stood out due to their high diagnosis accuracy rate. When deep learning was introduced to this field, the diagnosis accuracy became higher and higher, but the models became more and more difficult to understand. The lack of explainability and interpretability limited practical application of deep learning-based FDD methods. In this paper, we applied graph attention neural network (GAT) for FDD and achieved high diagnosis accuracy. Attention mechanism in GAT can help understand the inference process of the network. Combining the attention mechanism with the PageRank algorithm, a visualization graph can be generated to highlight the variables that are most relevant to a certain type of fault. This can be used to explain the fault diagnosis results and help us understand deep learning-based fault diagnosis better.

**Keywords:** chemical fault diagnosis, explainability, graph attention network, page rank.

## 1. Introduction

Fault detection and diagnosis (FDD) is quite critical to safe operations of chemical processes to identify abnormal events that can hardly be controlled by distributed control system. In decades, many researchers have been proposed different models of real-time FDD. These models can be categories into quantitative model-based, qualitative model-based and process history-based models (Venkatasubramanian et al., 2003). Process history-based models are further classified into qualitative and quantitative models. The latter is also commonly termed as data-driven models. Data-driven models based on deep learning have drawn much attention of researchers these years, such as DBN-based models (Zhang and Zhao, 2017), CNN-based models (Wu and Zhao, 2018), RNN-based models (Zhang et al., 2020), autoencoder-based models (Bi and Zhao, 2021) and GCN-based models (Wu and Zhao, 2021). These models achieved high diagnosis accuracy, but the lack of explainability made them hard to understand by humans thus limited their practical application.

In this paper, graph attention neural network (GAT) was applied to FDD to achieved high diagnosis accuracy. Attention mechanism built in GAT gave out local attention weights to help understand the model itself. Then a classic algorithm PageRank was applied to extend the local attention to global attention. The global attention can be used to generate

a visualization graph to highlight the variables that are most relevant to a certain type of fault. This was a great improvement of explainability of deep learning-based FDD model.

## 2. Graph Attention Network and PageRank

### 2.1. Graph Attention Network

Graph neural networks (GNNs) were proposed for processing graph-structured data, which are common in the real world. GNNs can be classified into recursive graph neural networks (RecGNNs) and convolutional graph neural networks (ConvGNNs), and the latter includes spectral-based and spatial-based ones. Because of many similarities with CNNs, spatial-based ConvGNNs have gained much attention from researchers, and the representative network structure is graph convolutional network (GCN) (Kipf and Welling, 2017). Spatial-based ConvGNNs define convolutional layers for information passing and aggregation operations. They propagate information through the connections, and aggregate the information of neighbor nodes and corresponding edges to a certain node to update its hidden embedding. Then graph readout technologies are used to obtain a node-level or graph-level representation of the information in the graph.

Graph attention network (GAT) (Veličković et al., 2018) introduced masked self-attentional layers to address the shortcomings of GCN and its variants. In convolutional operations, GAT allows for assigning different weights to different nodes within a neighborhood with efficient computation, as shown in Figure 1. The attention mechanism greatly improved the performance of the model, while bringing some model interpretability. The same advantages can be obtained by applying GAT to chemical process diagnosis.

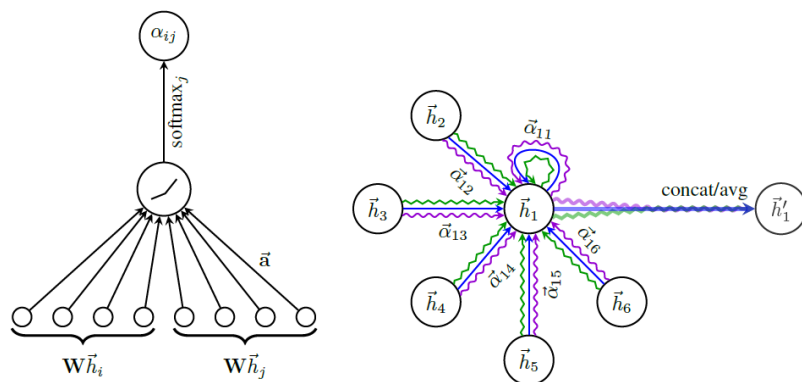


Figure 1: Left: The attention mechanism in GAT. Right: Multi-head attention by node 1 on its neighborhood. (Veličković et al., 2018)

### 2.2. PageRank

PageRank is the founding algorithm of Google's search engine and is one of the most famous algorithms used to rank web pages in order of importance. It's named after Larry Page, one of the founders of Google Inc. Essentially, the algorithm analyzes the importance of a web page by the number and quality of its hyperlinks.

In terms of quantity, the more hyperlinks from other pages point to a web page, the more important the page is. In terms of quality, the more important hyperlinks point to a web page, the more important the page is. This is very similar to the concept of citation networks.

In addition to web ranking, PageRank has a wide range of other application scenarios. When combining with the attention mechanism in GAT, it will be possible to analyze the importance of the nodes in a graph.

### 3. Fault Diagnosis with GAT

#### 3.1. Fault Diagnosis of Chemical Processes

When applying deep learning-based models, fault diagnosis of a chemical process is generally regarded as a classification problem. Given an observation  $\mathbf{X}_t \in \mathbb{R}^{v \times w}$  at time  $t$ , the current operating state  $y \in \mathbb{N}$  of the process should be identified out of a set of operating states consisting of normal state and different types of faults. The observation  $\mathbf{X}_t$  is a matrix with the time window and process variable dimensions. The operating state or label  $y$  is an integer between 0 and  $N_c$  (the number of different fault states).  $y = 0$  means the process is under normal operating state, and  $y = c$  ( $1 \leq c \leq N_c$ ) means the fault  $c$  has occurred. With enough pairs of observations and labels, a dataset  $\{\mathbf{X}_t^{(i)}, y^{(i)}\}_{i=1}^N$  is obtained for training and testing a deep learning-based fault diagnosis model.

#### 3.2. Integrated with Process Topology

The experiments were conducted on the Tennessee Eastman (TE) process (Downs and Vogel 1993). It's a simulation process model modified based on an actual industrial process of Eastman Chemical Company in Tennessee, USA. TE process mainly consists of five unit operations and defines 20 different process disturbances. In the simulation program of TE process, 52 variables can be observed, and 20 different disturbances can be inserted at any time to make the process operate in a fault state.

GAT performs convolutional operations on a graph. Naturally, the topology between process variables can be considered as a graph in a chemical process. The graph structure used in this paper was consistent with PTCN in (Wu and Zhao, 2021). When training the model with process data, GAT performs graph convolutional operations based on the adjacency matrix corresponding to the graph.

#### 3.3. Data Preparation and GAT Model Training

To obtain enough data for training the GAT model, the simulation program of TE process ran for 3000 h in normal state. Under every different fault, the simulation program ran for 20 h after the fault inserted, and 10 sets of parallel simulations were carried out. All the data have been normalized with the mean and standard deviation calculated from data in the normal state. The normalized data were then cut to slices with the time widow of 1 h, and the data are sampled every 3 min. Data from 8 sets of parallel simulations were used for training and the rest 2 sets were used for testing.

The network structure is *GATConv(20)-GATConv(20)-GATConv(20)-FC(300, dropout=0.5)-FC(21)*. It means that the network includes 3 GAT convolutional layers with node embedding of 20 dimensions. This is followed by a fully connected layer with 300 neurons (dropout technique is used with a 0.5 ratio) and a fully connected layer with 21 neurons. It was trained for 50 epochs with Adam optimizer, the mini-batch size was set to 128 and the learning rate was set to 0.001. The trained GAT model finally got an average classification rate of 0.9399 over all the 21 classes.

The confusion matrix of the diagnosis results is shown in Figure 2. Table 1 gives FDR (fault diagnosis rate) and ACR (accurate classification rate) comparison between different types of deep learning models. It shows that the GAT model has the same excellent

diagnosis performance as PTCN. If fault 9 and 15 are not considered, the ACR is as high as 0.9731. Although the GAT model has only a very weak advantage in diagnosis performance compared to PTCN, the attention weights in the network give us an entry for an interpretability study.

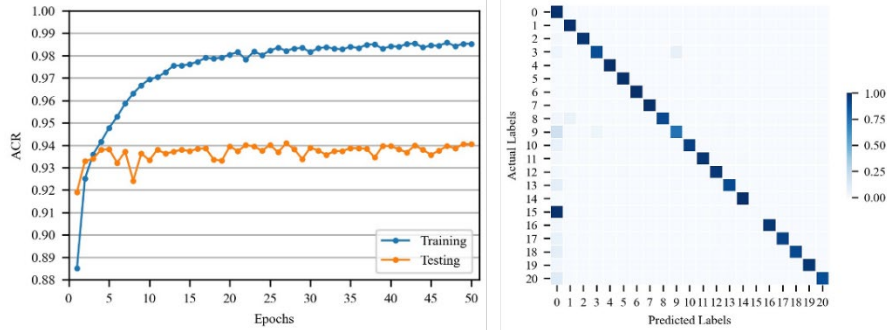


Figure 2: Left: Changes of ACR when training. Right: Confusion matrix of diagnosis results.

Table 1: FDR and ACR comparison of fault diagnosis experiments results on testing dataset

	<b>DBN</b>	<b>DCNN</b>	<b>BiGRU</b>	<b>PTCN</b>	<b>GAT</b>
<b>Faults</b>	(Zhang and Zhao, 2017)	(Wu and Zhao, 2018)	(Zhang et al., 2020)	(Wu and Zhao, 2021)	(ours)
Normal	-	0.978	0.969	0.9924	0.9927
Fault 1	1	0.986	0.986	0.9931	0.9944
Fault 2	0.99	0.985	0.972	0.9819	0.9817
Fault 3	0.95	0.917	0.935	0.8804	0.8872
Fault 4	0.98	0.976	0.974	0.9956	0.9925
Fault 5	0.86	0.915	0.998	0.9786	0.9785
Fault 6	1	0.975	1	1	0.9984
Fault 7	1	0.999	1	1	1
Fault 8	0.78	0.922	0.753	0.9160	0.9136
Fault 9	0.57	0.584	0.807	0.6601	0.6814
Fault 10	0.98	0.964	1	0.9276	0.9288
Fault 11	0.87	0.984	0.965	0.9798	0.9807
Fault 12	0.85	0.956	0.961	0.9704	0.9664

Fault 13	0.88	0.957	0.953	0.8969	0.8971
Fault 14	0.87	0.987	0.996	0.9964	0.9947
Fault 15	0	0.28	0.541	0.0035	0.0029
Fault 16	0	0.442	0.788	0.9685	0.9682
Fault 17	1	0.945	0.97	0.9254	0.9261
Fault 18	0.98	0.939	0.923	0.9049	0.9049
Fault 19	0.93	0.986	0.926	0.9650	0.9682
Fault 20	0.93	0.933	0.981	0.8825	0.8828
ACR	0.821	0.882	0.927	0.9392	0.9399
ACR w/o					
Fault 9 & 15	0.889	0.934	0.952	0.9729	0.9731

#### 4. Extend from Local to Global Attention with PageRank

After the GAT model was trained well, the attention weights in it were also fixed. These weights can be used to investigate exactly which nodes or variables the network is focusing on when making inferences. Visualization of attention weights in GAT is shown in Figure 3. Every node in the attention weight graph represents a variable in TE process. The nodes are connected by edges to represent the physical connection between variables, i.e. the topology of the process. The magnitude of the attention weight is expressed with the thickness of the edges. Note that this attention mechanism only makes sense for arbitrary nodes and their neighbor nodes. For each node, the attention weights of all its incoming edges are normalized. This means that it's a local attention mechanism in GAT.

The attention weight graph also follows the same quantitative and qualitative assumptions as PageRank. When applying PageRank to the graph, it can be iterated until convergence and finally give the importance ranking of the nodes in the whole graph, as shown in Figure 3. The size of the nodes is used to indicate the importance of the variables. This importance ranking is for the whole graph and is therefore a kind of global attention. This helps us to analyze which variable is more important for fault diagnosis. In Table 2, the most important process variables for diagnosing each fault are listed and compared with the results of the manual process analysis. For some faults, affected variables can be located with the importance ranking. For those faults that are hard to identify (such as fault 9 and 15), all variables only change slightly after the disturbance is inserted. Thus, it can hardly localize some certain variables to determine the type of fault. This result indicates that the GAT model, when performing fault diagnosis, does not necessarily focus directly on the root cause of the fault, but rather identifies those sensitive variables that are strongly affected by the root cause.



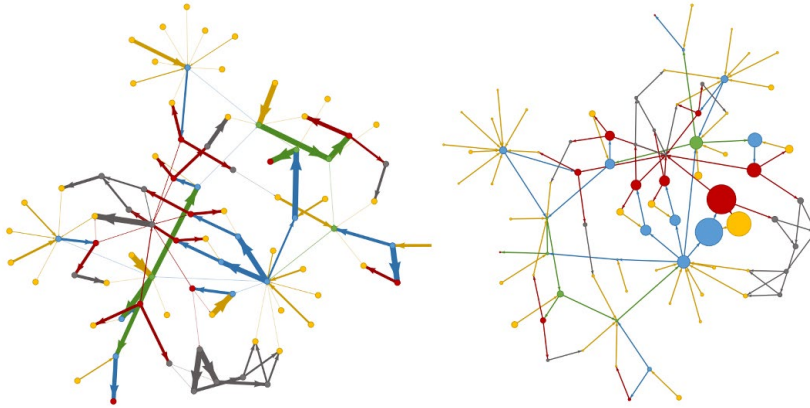


Figure 3: Left: Visualization of attention weights in GAT. Right: Visualization of global attention after applying PageRank to attention weight graph.

Table 2: Comparison between variables affected by disturbances firstly and variables the GAT model focused on when diagnosing faults.  $T, P, L, Q, x, u$  mean temperature, pressure, level, flow rate, composition, control signal respectively. Subscript means unit operations or streams. Superscript means characteristics or components.

Faults	Variables affected firstly	Variables GAT focused on
Fault 1	$x_{feed}^A, x_{feed}^C, Q_{s1}, Q_{s4}, u_{s1}^Q, u_{s4}^Q$	$Q_{s1}, Q_{s4}, u_{s1}^Q, u_{s4}^Q$
Fault 2	$x_{feed}^A, x_{feed}^B, x_{feed}^C, Q_{s1}, Q_{s4}, u_{s1}^Q, u_{s4}^Q$	$x_{feed}^A, x_{feed}^C, Q_{s1}, Q_{s4}, u_{s1}^Q, u_{s4}^Q$
Fault 3	$T_{reac}, P_{reac}, L_{reac}, T_{reac}^{cwr}, u_{reac}^{cwr}$	$x_{feed}^A, x_{feed}^C, Q_{s1}, Q_{s4}, u_{s1}^Q, u_{s4}^Q$
Fault 4	$T_{reac}^{cwr}, u_{reac}^{cwr}, T_{reac}, P_{reac}, L_{reac}$	$T_{reac}^{cwr}, u_{reac}^{cwr}, T_{reac}, x_{feed}^A, Q_{s1}, u_{s1}^Q$
Fault 5	$T_{cond}^{cwr}, u_{cond}^{cwr}, T_{sep}, P_{sep}, L_{sep}$	$u_{cond}^{cwr}, Q_{s4}, u_{s4}^Q$
Fault 6	$Q_{s1}, u_{s1}^Q, Q_{feed}, x_{feed}^A$	$T_{reac}^{cwr}, u_{s1}^Q$
Fault 7	$Q_{s4}, u_{s4}^Q, T_{strip}, P_{strip}, L_{strip}$	$Q_{s4}, u_{s4}^Q, T_{reac}^{cwr}$
Fault 8	$x_{feed}^A, x_{feed}^B, x_{feed}^C, Q_{s1}, Q_{s4}, u_{s1}^Q, u_{s4}^Q$	$Q_{s1}, Q_{s4}, u_{s1}^Q, u_{s4}^Q$
Fault 9	$T_{reac}, P_{reac}, L_{reac}, T_{reac}^{cwr}, u_{reac}^{cwr}$	$Q_{s1}, Q_{s4}, u_{s1}^Q, u_{s4}^Q$
Fault 10	$T_{strip}, P_{strip}, L_{strip}, u_{strip}^{steam}, Q_{strip}^{steam}$	$T_{strip}, Q_{s1}, Q_{s4}$
Fault 11	$T_{reac}^{cwr}, u_{reac}^{cwr}, T_{reac}, P_{reac}, L_{reac}$	$Q_{s1}, u_{s1}^Q$
Fault 12	$T_{cond}^{cwr}, u_{cond}^{cwr}, T_{sep}, P_{sep}, L_{sep}$	$T_{reac}, T_{sep}$

Fault 13	$T_{\text{reac}}, P_{\text{reac}}, L_{\text{reac}}, T_{\text{reac}}^{\text{cwr}}, u_{\text{reac}}^{\text{cwr}}$	$T_{\text{reac}}^{\text{cwr}}, L_{\text{reac}}$
Fault 14	$T_{\text{reac}}^{\text{cwr}}, u_{\text{reac}}^{\text{cwr}}, T_{\text{reac}}, P_{\text{reac}}, L_{\text{reac}}$	$T_{\text{reac}}^{\text{cwr}}, Q_{s1}, u_{s1}^Q$
Fault 15	$T_{\text{cond}}^{\text{cwr}}, u_{\text{cond}}^{\text{cwr}}, T_{\text{sep}}, P_{\text{sep}}, L_{\text{sep}}$	$Q_{s1}, Q_{s4}, u_{s1}^Q, u_{s4}^Q$

## 5. Conclusion

In this paper, GAT was applied to fault diagnosis and achieved high diagnosis accuracy. After the model was trained well, PageRank was used to extend the local attention of GAT to global attention and generated the importance ranking of variables for each type of fault. This ranking helps us to understand the diagnosis process of the GAT model more deeply.

The lack of interpretability and explainability of the deep learning models limits their practical application. Future research on fault diagnosis should pay more attention to inventing more interpretable models, or better integrating first principles into data-driven models. More explainability research on deep learning models should also be conducted.

## References

- Bi, X., Zhao, J., 2021. A novel orthogonal self-attentive variational autoencoder method for interpretable chemical process fault detection and identification. *Process Safety and Environmental Protection*.
- Kipf, T.N., Welling, M., 2017. Semi-Supervised Classification with Graph Convolutional Networks. arXiv:1609.02907 [cs, stat].
- Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., Bengio, Y., 2018. Graph Attention Networks. arXiv:1710.10903 [cs, stat].
- Venkatasubramanian, V., Rengaswamy, R., Yin, K., Ka, S.N., 2003. A review of process fault detection and diagnosis Part I: Quantitative model-based methods. *Computers and Chemical Engineering* 19.
- Wu, D., Zhao, J., 2021. Process topology convolutional network model for chemical process fault diagnosis. *Process Safety and Environmental Protection* 150, 93–109.
- Wu, H., Zhao, J., 2018. Deep convolutional neural network model based chemical process fault diagnosis. *Computers & Chemical Engineering* 115, 185–197.
- Zhang, S., Bi, K., Qiu, T., 2020. Bidirectional Recurrent Neural Network-Based Chemical Process Fault Diagnosis. *Industrial & Engineering Chemistry Research* 59, 824–834.
- Zhang, Z., Zhao, J., 2017. A deep belief network based fault diagnosis model for complex chemical processes. *Computers & Chemical Engineering* 107, 395–407.



# Model predictive control of compact combined cycles in offshore power plants integrating a wind farm

Kang Qiu<sup>a</sup>, Leif E. Andersson<sup>a\*</sup>, Cristina Zotică<sup>a</sup>, Adriana Reyes-Lúa<sup>a</sup>, Rubén M. Montañés<sup>a</sup>, Adriaen Verheyleweghen<sup>b</sup>, Valentin Chabaud<sup>a</sup>, Til Kristian Vrana<sup>a</sup>

<sup>a</sup> SINTEF Energy Research, Sem Sælands vei 11, 7034 Trondheim, Norway

<sup>b</sup> Cybernetica AS, Leirfossvegen 27, 7038 Trondheim, Norway

\* Corresponding author: [leif.andersson@sintef.no](mailto:leif.andersson@sintef.no)

## Abstract

Combined cycle gas turbine plants (CCGTs) fulfill an important role in emission reduction of offshore power systems as the bottoming cycle (BC) produces additional power from exhaust heat of the gas turbines (GTs). With increasing integration of wind turbines, CCGTs offshore must be flexible and provide variation management to the offshore energy system across multiple time scales. This work proposes a model predictive controller (MPC) sending setpoints to the CCGT to satisfy demand in the offshore power system under fluctuating wind power. A high-speed surrogate model suitable for optimizing in an MPC is identified. A linear MPC using a quadratic cost function with process constraints is formulated. The model-based control structure is then validated in simulation for satisfying a constant power demand under disturbances introduced by fluctuating wind power.

**Keywords:** Model predictive control; steam bottoming cycle; surrogate modelling; feedforward disturbance rejection; PI-control; wind power

## 1. Introduction

Currently, natural gas fueled gas turbines (GTs) are the main source of power in offshore oil and gas installations. National and international CO<sub>2</sub> reduction targets are incentivizing offshore operators to develop emission reduction strategies such as increasing energy efficiency and low emission power generation solutions like offshore wind power. It is likely that more than one technology will be implemented (Voldsund et al., 2023), resulting in a hybrid integrated offshore power generation system (Riboldi et al., 2020). In this work, we study a system with a combined cycle (CCGT) and power from wind. In CCGTs, the bottoming cycle (BC) uses the exhaust heat from GTs to produce additional power in a steam turbine (ST) and increase thermal energy efficiency (Nord and Bolland, 2013). Due to weight and space limitations, BCs on offshore installations are designed to be low weight and compact, which affects their dynamic response (Montañés et al., 2021). As simple cycle GTs, CCGTs offshore must provide variation management to the offshore energy system across multiple time scales, stabilizing the power generation system. Decentralized control strategies for compact BCs and CCGT based on PI- and feedforward controllers were studied by Nord and Montañés (2018), and model-based nonlinear feedforward in combination with PI-controllers was developed by Zotică et al. (2022).

The requirements for disturbance rejection of CCGTs are stringent when integrating non-dispatchable wind turbines to the power generation system offshore, due to the increased variability of the net load to be covered by the CCGT. A model predictive controller (MPC) is proposed to control the setpoints to the GTs and provide setpoints to the lower PID control layer of the BC by exploiting information, e.g., of the required system power demand, wind profile forecasts, and a model of the disturbance. This enables the CCGT power output of the integrated system to minimize generation and consumption mismatches, thereby stabilizing the electrical frequency of the power system. Furthermore, the MPC framework allows for additional control objectives, such as minimizing CO<sub>2</sub> emissions and inherently handle steam temperature and pressure constraints.

## 2. System description and model

Figure 1 depicts the system considered in this work, consisting of an integrated wind-thermal electricity generation system. We assume that the thermal system should compensate for variations in wind power generation. The thermal system, a CCGT, consists of two GTs and a BC with two once-through steam generators (OTSG) connected to a common steam turbine (ST) and condenser system. A wind farm produces non-dispatchable power and the CCGT stabilizes the integrated power output. We simulate the power from the wind farm and CCGT using high-fidelity models and develop a surrogate model of the CCGT for the MPC.

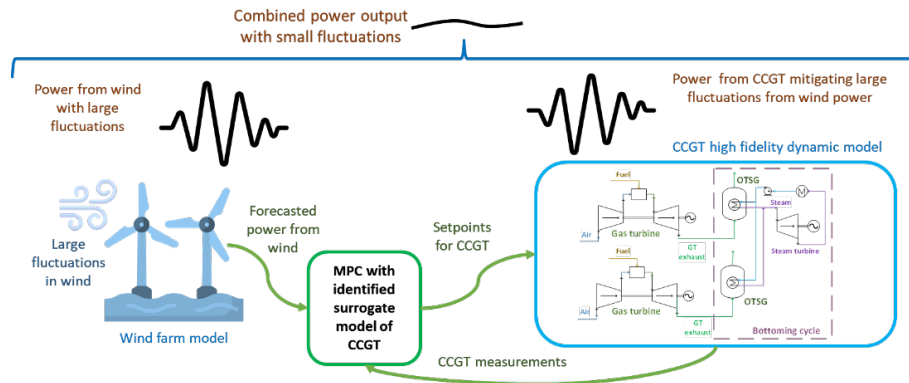


Figure 1. Conceptual description of the system.

### 2.1. Power from wind farm

The considered wind power plant consists of three wind turbines with 8 MW power rating each. The specifications are inspired by the Siemens Gamesa SG 8.0-167 DD wind turbine. The control of the wind turbines is taken as state-of-the-art grid-following maximum-power-point-tracking control, where the turbine actions are almost purely driven by the incoming wind, from which as much energy as possible is harvested. The wind turbines do not actively attempt to stabilize the electric power system. The task of stable control of the electric system is purely left to the gas turbines. This principle reflects the control of commercially available wind turbines, and the typical variation management strategy in wind-thermal systems (Burton, 2011).

The turbine layout is taken as a straight line with a spacing of ten diameters between turbines (standard, conservative spacing). A wind speed of 10 m/s with direction perpendicular to the layout (no wake) and with standard turbulence level is chosen. The

incoming wind is then obtained from a farm-scale synthetic turbulence generator providing the rotor-averaged (power-equivalent) wind speed with consistent low-frequency fluctuations and correlation between turbines (Chabaud, 2023), for a consistent representation of wind power variability.

### *2.2. High fidelity dynamic model of the combined cycle*

SINTEF Energy Research developed a high-fidelity dynamic Modelica model of the CCGT (Montañés et al., 2021; Zotică et al., 2022), using the commercial modeling and simulation environment Dymola. We utilize Modelica models from the Thermal Power Library 1.21 and adapt them to build up the combined cycle power plant models. In this work we model the thermo-hydraulics of the steam cycle of the CCGT with high-fidelity principles and apply a simplified quasi-static model for the GT based on data provided by Siemens Energy through the LowEmission Consortium. The bottoming cycle is based on the design presented in Zotică et al. (2022). For further details on the underlying models for turbine islands and steam generators, refer the work by Montañés et al. (2017) and Zotică et al. (2022). OTSG models are explained in the work by Montañés et al. (2021).

### *2.3. MPC surrogate model of the combined cycle*

In this work, the surrogate model used for the predictive controller is a linear, autoregressive model of the CCGT predicting the steam turbine power output  $P_s$ , and CO<sub>2</sub> emissions from the operation of the GTs. The parameters of the surrogate model are determined with a stability-constrained linear regression performed on data gathered from simulation of the high-fidelity dynamic model of the CCGT.

The GT power  $x_{GT}$  and GT CO<sub>2</sub> emissions  $x_{CO_2}$  are modelled dependent on the current GT load,  $u_{GT}$ ,

$$x_{GT} = f(u_{GT}) \quad (1)$$

$$x_{CO_2} = f(u_{GT}). \quad (2)$$

The dynamics of the BC are highly nonlinear. However, the lower PID control layer results in more linearized dynamics (see section 3.1) and enables the use of a linear prediction model for the MPC. In the BC, the OTSG consists of relatively large volumes and mass of metal walls that accumulate heat over time. Hence, the BC exhibits high thermal inertia at the interface of the gas turbine exhaust and steam generation in the OTSG (Montañés et al., 2021). Therefore, the surrogate model necessitates a higher order,  $n$ , for modelling the predicted ST power in the next timestep  $k + 1$ . The ST power is predicted from the temperature setpoints  $u_t$  and pressure setpoint  $u_p$  of the BC

$$x_{ST,k+1} = f(x_{BC,k-n}, u_{GT,k-n}, u_{t,k-n}, u_{p,k-n}), \quad n = 0, \dots, 5 \quad (3)$$

## **3. Control structure**

The control structure proposed in this paper consists of a lower PID control layer of the BC and GT and an upper MPC control layer to provide setpoints to the lower layer.

### *3.1. PID control layer*

The lower PID control layer keeps the superheated steam temperature at the outlet of the two OTSGs and superheated steam pressure at a setpoint given by the upper supervisory control layer (MPC). The power output of the ST is highly dependent on the enthalpy at the inlet of the steam turbine. The BC heat input is mostly given by the gas turbine operation specified as GT load. The live steam pressure is controlled by manipulating the valve upstream the ST using a pure I-controller. The temperature is controlled by

manipulating the feedwater flowrate, using a combination of model-based nonlinear feedforward for disturbance rejection and a feedback PI-controller to reject unmeasured disturbances and account for plant model mismatch. The nonlinear feedforward is derived from a simple steady-state energy balance over the OTSG. It receives measurements from the hot gas side flowrate and temperatures which are the main disturbances for the BC. The I and PI controllers are tuned from step responses using the SIMC tuning rules (Skogestad, 2003).

### 3.2. MPC control layer

The MPC provides GT setpoints  $u_{GT}$ , superheated steam temperature setpoints  $u_t$ , and the superheated steam pressure setpoint  $u_p$  to the lower PID control layer. A linear MPC using a quadratic cost function with input constraints, and the surrogate model described in 2.3 is formulated

$$\begin{aligned} \min_u \quad & \omega_d J_d + \omega_{CO_2} J_{CO_2} + \omega_a J_a \\ \text{s. t.} \quad & x_{k+1} = f(x_{k-n}, u_{k-n}), \quad n = 0, \dots, 5, \quad k = 0, \dots, N-1 \\ & g(u_k) \leq 0, \quad k = 0, \dots, N-1 \\ & x_0 = \hat{x}_0. \end{aligned} \quad (4)$$

The objective of the MPC consists of, in order priority, a power tracking cost  $J_d$  to satisfy a power demand, a CO<sub>2</sub> emission penalty  $J_{CO_2}$ , and an actuation cost  $J_a$ . The power tracking cost is calculated from the difference between the total predicted power output, consisting of the power output from the GTs  $x_{GT}$  and ST  $x_{ST}$ , forecasted wind power  $p_w$ , and the power demand  $p_d$

$$J_d = \sum_{k=0}^N (p_{d,k} - p_{w,k} - x_{GT,k} - x_{ST,k})^2. \quad (5)$$

The CO<sub>2</sub> emission penalty is a quadratic cost on the predicted CO<sub>2</sub> emissions

$$J_{CO_2} = \sum_{k=0}^N x_{CO_2,k}^2. \quad (6)$$

Lastly, the actuation cost penalizes the change in the GT setpoints  $u_{GT}$ , BC temperature setpoints  $u_t$ , and BC pressure setpoint  $u_p$  with a quadratic cost

$$J_a = \sum_{k=0}^{N-1} (u_{GT,k+1} - u_{GT,k})^2 + (u_{t,k+1} - u_{t,k})^2 + (u_{p,k+1} - u_{p,k})^2. \quad (7)$$

To simplify the tuning of the weights  $\omega$ , both the state and input variables are normalized to ensure comparability of the subobjectives  $J_{CO_2}$  and  $J_a$ . As  $J_d$  is the most important subobjective, the  $\omega_d$  is set first. Afterwards  $\omega_{CO_2}$  and  $\omega_a$  are tuned by trial and error.

## 4. Results and discussion

The proposed MPC is tested in a simulation environment introduced in 4.1. The results (see 4.2) illustrate the disturbance rejection the MPC provides given fluctuating wind power and ST power output.

### 4.1. Simulation environment

The MPC is simulated in loop with a Gaussian process model of the CCGT regressed from data gathered from the high-fidelity model. The sampling time of 5 s and prediction horizon of 5 min were chosen to account for the vastly different time dynamics of the GTs and the BC, which exhibited a closed loop settling time of  $10^0$ - $10^1$  s and  $10^2$ - $10^3$  s, respectively (Montañés et al., 2021). The simulation environment is assumed to provide full-state feedback of the ST power, GT power, and CO<sub>2</sub> emissions of the GTs. Realistic wind profiles as described in Section 2.1, assumed to be forecasted exactly, are utilized.

The simulations are formulated with CasADi (Andersson et al., 2019) and solved with IPOPT (Wächter and Biegler, 2006), where each MPC iteration is solved within 0.05 s.

#### 4.2. Results

Figure 2(a) shows the cumulative power output of the wind turbines and the CCGT. The proposed MPC can control the CCGT to provide variation management under fluctuating wind power to satisfy a constant demand of 70 MW. Figure 2(b), showing the relative power of the wind turbines, GTs, and ST scaled to the total power demand indicates that fluctuations in wind power are mainly compensated by the power output of the GTs due to their fast dynamics.

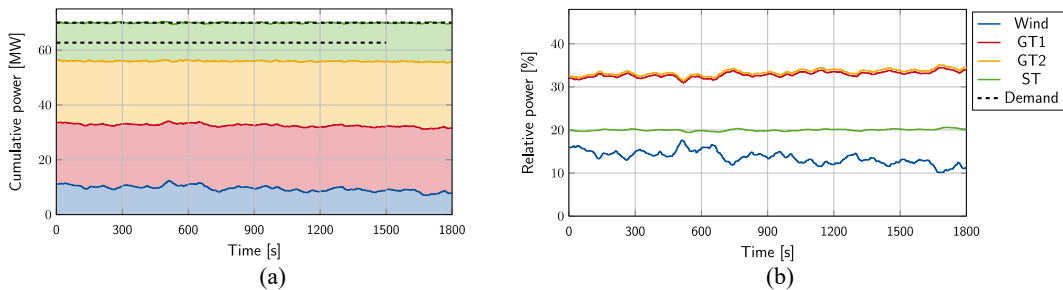


Figure 2. (a) Cumulative power output of the combined wind-thermal system. (b) Relative power output of CCGT and wind scaled to total demand.

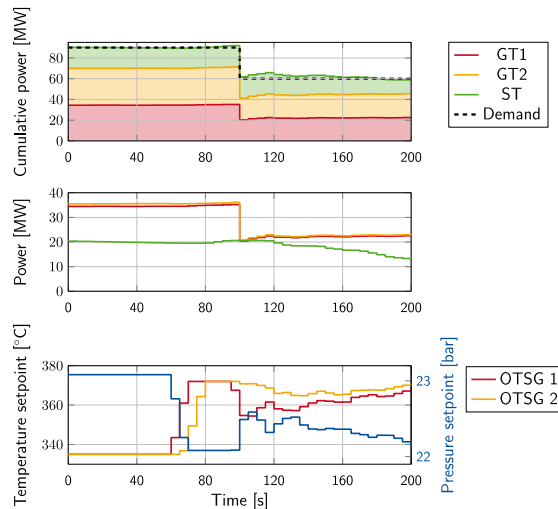


Figure 3. MPC step response showing cumulative power, GT load inputs, and steam temperature and pressure setpoints chosen by the MPC.

Figure 3 illustrates the working principle of the MPC when the power demand of the CCGT changes, e.g., due to a change in wind power or overall wind-thermal system demand. Overall, the CCGT responds sufficiently fast to the step change in power demand. The GT loads are adjusted at time of demand change, as GTs have a fast dynamic response. Furthermore, the MPC accounts for the slower dynamic response of the BC by regulating the gas turbine loads at  $t = 100$  s lower than the steady state gas turbine load at  $t \geq 200$  s. The dynamic response of the BC is improved by MPC as it regulates the setpoints to the lower PID control layer at  $t = 60$  s before the demand step change at  $t = 100$  s to account for the slower BC dynamics



## 5. Conclusion

This work proposes an MPC for CCGT control with variation management, and CO<sub>2</sub> emission reduction objectives. To formulate the MPC, a linear surrogate model is identified. As the CCGT process and the high-fidelity Modelica model are highly nonlinear, it is challenging to represent the dynamics correctly in a linear model.

Simulations show that the MPC helps to achieve variation management of the integrated wind-thermal system by controlling the CCGT under fluctuating, realistic wind profiles. Further work should address uncertain wind profile forecasts and investigate different objectives to the MPC.

## Acknowledgements

This publication has been produced with support from the LowEmission Research Centre ([www.lowemission.no](http://www.lowemission.no)), performed under the Norwegian research program PETROSENTER. The authors acknowledge the industry partners in LowEmission for their contributions and the Research Council of Norway (296207). We acknowledge SIEMENS Energy AG for providing the gas turbine model used as reference.

## References

- Andersson, J.A.E., Gillis, J., Horn, G., Rawlings, J.B., Diehl, M., 2019. CasADi: a software framework for nonlinear optimization and optimal control. *Math. Prog. Comp.* 11, 1–36.
- Burton, T. (Ed.), 2011. *Wind energy handbook*, 2nd ed. ed. Wiley, Chichester, West Sussex.
- Chabaud, V., 2023. Synthetic turbulence modelling for offshore wind farm engineering models using coherence aggregation (preprint). Preprints.
- Montañés, R.M., Skaugen, G., Hagen, B., Rohde, D., 2021. Compact Steam Bottoming Cycles: Minimum Weight Design Optimization and Transient Response of Once-Through Steam Generators. *Front. Energy Res.* 9, 687248.
- Nord, L.O., Bolland, O., 2013. Design and off-design simulations of combined cycles for offshore oil and gas installations. *Applied Thermal Engineering* 54, 85–91.
- Riboldi, L., Alves, E.F., Pilarczyk, M., Tedeschi, E., Nord, L.O., 2020. Optimal Design of a Hybrid Energy System for the Supply of Clean and Stable Energy to Offshore Installations. *Front. Energy Res.* 8, 607284.
- Skogestad, S., 2003. Simple analytic rules for model reduction and PID controller tuning. *Journal of Process Control* 13, 291–309.
- Voldsund, M., Reyes-Lúa, A., Fu, C., Ditaranto, M., Nekså, P., Mazzetti, M.J., Brekke, O., Bindingsbø, A.U., Grainger, D., Pettersen, J., 2023. Low carbon power generation for offshore oil and gas production. *Energy Conversion and Management: X* 17, 100347.
- Wächter, A., Biegler, L.T., 2006. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.* 106, 25–57.
- Zotică, C., Montañés, R.M., Reyes-Lúa, A., Skogestad, S., 2022. Control of steam bottoming cycles using nonlinear input and output transformations for feedforward disturbance rejection. *IFAC-PapersOnLine* 55, 969–974.

# Application of Bayesian Optimization in HME Batch Concentration Process

Chong Liu<sup>a</sup>, Cheng Ji<sup>a</sup>, Chengyu Han<sup>a</sup>, Chenxi Gu<sup>b</sup>, Jindong Dai<sup>a</sup>, Wei Sun<sup>a,\*</sup>,  
Jingde Wang<sup>a,\*</sup>

<sup>a</sup>*College of Chemical Engineering, Beijing University of Chemical Technology, North Third Ring Road 15, Chaoyang District, Beijing, 100029, China*

<sup>b</sup>*Shijiazhuang Lonzeal Pharmaceutical Co., Ltd., No. 16 West Ring Road, Shenze County Industrial Park, Shijiazhuang, Hebei Province, 052560, China*

\*Corresponding Author's E-mail: [jingdewang@mail.buct.edu.cn](mailto:jingdewang@mail.buct.edu.cn)

\*Corresponding Author's E-mail: [sunwei@mail.buct.edu.cn](mailto:sunwei@mail.buct.edu.cn)

## Abstract

Pharmaceutical intermediates are chemical products that play an important role in the synthesis of drugs. Different from the general chemical production process, the production process of pharmaceutical intermediates is characterized by flexible small scale, batch mode and multifunction. Therefore, optimal operation strategies are required at different stages, usually implemented by adjusting operation parameters, such as temperature and pressure along the batch for particular units and instruments. However, it is challenging to obtain the best combination of multiple parameter trajectories. In this paper, the concentration of HME, a pharmaceutical intermediate, is studied as an industrial case. A dynamic model is established to simulate the whole concentration process of HME. To promptly achieve optimal temperature and pressure in HME batch concentration, a Bayesian optimization algorithm based on Gaussian process regression model is used with the goal of minimizing the total cost. It can be seen from the results that Bayesian optimization is very competitive in dynamic simulation.

**Keywords:** batch concentration, Bayesian optimization, pharmaceutical intermediates, dynamic process simulation

## 1. Introduction

Lamivudine is an antiviral drug of nucleoside reverse transcriptase inhibitor, which can selectively inhibit the replication of hepatitis B virus. At present, the chemical synthesis method using L-menthol as the starting material is often adopted in industry (Xue et al., 2010). (2S,5R)-5-Hydroxy- [1,3] -oxathiolane-2-carboxylic acid methyl ester (HME) is an important chemical synthesis intermediate of Lamivudine, and the concentrate process of HME has a significant impact on the synthesis of Lamivudine. The synthesis route is shown in Figure 1 (where I is Lamivudine and III is HME.). Based on industrial projects, this paper focuses on the concentration of HME from its solvent mixture system and optimization of operating parameters, in order to minimize the total production cost.

The dynamic optimization of batch processes has been concerned by many researchers, and many solutions have been proposed. Yang et al. applied iterative particle swarm

optimization (PSO) algorithm to the comprehensive optimization problem of batch processes, in which objective function was set as the sum of product concentration, reciprocal of reaction time and reciprocal of energy loss, and the optimal solution was obtained by searching for optimal trajectory (Yang et al., 2017). Amini et al. proposed a new method to optimize the control of batch cooling crystallizers by genetic algorithm, and obtained the minimum and maximum values of the objective function (Amini et al., 2016). Mujtaba et al. solved the dynamic optimization of the batch reactor with a simple model through neural network, and obtained the optimal operation strategy in terms of reactor temperature (Mujtaba et al., 2006). However, the above algorithms are easy to fall into local optimum or less interpretability. Compared with the traditional optimization methods, the surrogate model of Bayesian optimization has stronger fitting ability and better interpretability, and therefore it is applied for the dynamic optimization of an HME batch concentration process in this work (Shahriari et al., 2016).

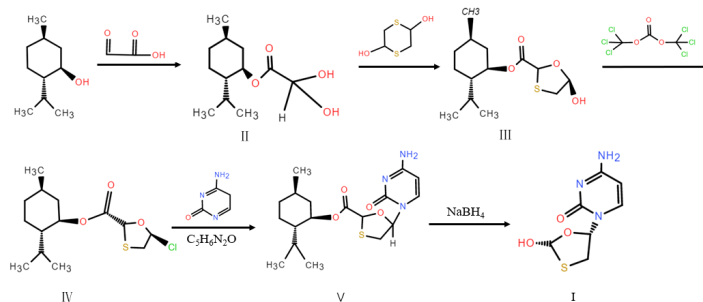


Figure 1 Synthesis route of Lamivudine

## 2. Process Modeling

HME concentration system is mainly composed of toluene, acetic acid, HME and water. The mass and mass fraction of each component are shown in Table 1.

Table 1 Mass and mass fraction of components

Component	Component quality/kg	Mass fraction
Toluene	2480	0.6727
Acetic acid	26.7	0.0072
HME	1080	0.2930
Water	100	0.0271

The dynamic concentration model is established as shown in the Figure 2. The whole concentration system consists of a jacketed distillation tank, two coolers, and two collection tanks. Instruments and controllers are equipped in necessary pipelines and equipment. Details on the control strategies are shown in Table 2.

Table 2 Details on the control strategies

Letter	Name	Control variable	Manipulate variable
A	TICT01	Temperature of T01	Heat medium flow
B	PICT01	Pressure of T01	Gas outlet flow of T01

C	TICE01	Outlet temperature of E01	Refrigerant 1 flow
D	TICE02	Outlet temperature of E02	Refrigerant 2 flow

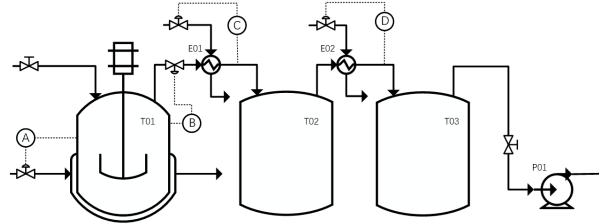


Figure 2 Concentration model

### 3. Bayesian optimization

#### 3.1. Parameter space and Objective function

##### 3.1.1. Manipulated variable selection

In the dynamic batch simulation process, the final production cost of the concentration process is highly correlated with the temperature and pressure control. First of all, pressure has a significant impact on the bubble dew point of the system. The lower the pressure, the lower the bubble point temperature, and the easier it is for concentration. However, to ensure the lower pressure, it is inevitable to increase the load of the vacuum pump. Therefore, it is necessary to balance the pressure in the distillation tank. Secondly, due to the intermittent feature of the concentration process, different temperature control schemes are selected in industry as time changes. Therefore, difference between the set value of temperature controller and bubble point of the system in the concentration process  $\Delta T$  is selected as one of optimization variables, so that temperature control value changes dynamically with the concentration time, making temperature control scheme more suitable for industrial practice. Similarly, difference between dew point at inlet of primary cooler and the set value of temperature controller  $\Delta t_1$ , and difference between dew point at inlet of secondary cooler and the set value of temperature controller  $\Delta t_2$ , are selected as optimization variables.

##### 3.1.2. Objective function

The total production cost is defined as objective function, and it is mainly divided into two parts, one is the real operation cost  $A$ , and the other is cost  $B$  of lost solvent toluene.  $A$  mainly includes labor cost, vacuum pump cost and utility cost.  $B$  is the cost of lost solvent toluene pumped by vacuum pump, including the cost of solvent toluene and the cost of handling the extracted toluene in order to meet the requirements of environmental protection. The calculation of the cost is listed below.

$$c_1 = a \cdot t \cdot n \quad (1)$$

$$c_2 = \left( \frac{d}{10 \cdot 8000} + \frac{f \cdot (101.325 - P) / 101.325 \cdot e}{\eta} \right) \cdot t \quad (2)$$

$$c_3 = F_1 \cdot U_1 \quad (3)$$

$$c_4 = F_2 \cdot U_2 \quad (4)$$

$$c_5 = F_3 \cdot U_3 \quad (5)$$

$$A = c_1 + c_2 + c_3 + c_4 + c_5 \quad (6)$$

$$B = F_4 \cdot (U_4 + U_5) \quad (7)$$

$$\text{Objective}(P, \Delta T, \Delta t_1, \Delta t_2) = A + B \quad (8)$$

where,  $c_1, c_2, c_3, c_4, c_5$ : Cost of labor, vacuum pump, refrigerant 1, refrigerant 2, heat medium, ¥,  $a$ : Hourly salary, ¥,  $t$ : Concentration time,  $h$ ,  $n$ : Number of workers,  $d$ : Fixed cost of vacuum pump, ¥,  $f$ : Rated power of vacuum pump,  $kW$ ,  $\eta$ : Efficiency of vacuum pump,  $e$ : Electricity price, ¥/kw·h,  $F_1, F_2, F_3, F_4$ : Total flow of refrigerant 1, refrigerant 2, heat medium, solvent toluene, kg,  $U_1, U_2, U_3, U_4, U_5$ : Unit price of refrigerant 1, refrigerant 2, heat medium, solvent toluene, solvent toluene treatment, ¥/kg.

### 3.2. Algorithm introduction

In essence, the Bayesian optimization framework uses the probabilistic surrogate model to fit the real objective function, and actively selects the most ‘potential’ evaluation points for evaluation according to the fitting results to avoid unnecessary sampling. Therefore, Bayesian optimization is also called active optimization (Jones, 1998). Overall, Bayesian optimization framework can effectively use process dynamic information to improve the search efficiency. And it mainly includes two parts: probabilistic surrogate model and acquisition function.

#### 3.2.1. Gaussian process regression

Gaussian process regression model aims to quantitatively estimate the probability distribution of the predicted value and obtain the confidence interval by establishing a nonparametric probability model based on the kernel function. It is considered as an extension of multivariate Gaussian function on infinite dimensional set variables. For the existing dataset,  $D = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ ,  $\mathbf{x}_i \in R^d$  is the input variable,  $d$  is the dimension of the input variable,  $y_i \in R$  is the target variable. When there are  $n$  experimental points in the data set, the prior knowledge of the Gaussian process regression algorithm assumes that these  $n$  points obey the multivariate Gaussian distribution  $N$  in space shown in Equation (9).

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \sim N \left( \begin{bmatrix} \mu(\mathbf{x}_1) \\ \vdots \\ \mu(\mathbf{x}_n) \end{bmatrix}, \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & \cdots & k(\mathbf{x}_1, \mathbf{x}_n) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_n, \mathbf{x}_1) & \cdots & k(\mathbf{x}_n, \mathbf{x}_n) \end{bmatrix} \right) \quad (9)$$

where  $\mu$  is the mean value;  $k(\cdot, \cdot)$  is a kernel function used to measure the correlation between two input variables. The kernel function selected in this paper is Matern, which is a stationary kernel and a generalization of the radial basis function (RBF). There is an additional parameter  $\nu$ , which controls the smoothness of the resulting function. In this work,  $\nu = 5/2$ , and specific expression is as follows.

$$k(x_i, x_j) = \left( 1 + \frac{\sqrt{5}}{l} d(x_i, x_j) + \frac{5}{3l} d(x_i, x_j)^2 \right) \exp \left( -\frac{\sqrt{5}}{l} d(x_i, x_j) \right) \quad (10)$$

where  $d(\cdot, \cdot)$  is the Euclidean distance.

### 3.2.2. Confidence Boundary Strategy

Srinivas et al. proposed a confidence boundary strategy for Gaussian processes, GP-UCB (Srinivas et al., 2010). When solving the minimum value of the objective function, the acquisition function of the UCB strategy is:

$$\alpha_t(x; D_{1:t}) = \mu_t(x) + \sqrt{\beta_t} \sigma_t(x) \tag{11}$$

where, parameter  $\beta_t$  balances expectation and variance.

Then Bayesian optimization algorithm is embedded into the batch concentration dynamic model, logic diagram is shown in Figure 3. The left side of the block diagram shows the process of Bayesian optimization algorithm using simulation data to recommend the next point to be evaluated. The batch concentration dynamic model on the right side is responsible for running under the working conditions of the parameter combination recommended by Bayesian optimization algorithm to obtain new cost simulation data, and then adding the simulation data to the dynamic database to iterate until convergence. In order to achieve data interaction between Python and C++, a set of communication master control for data transmission between both parties has been developed, which can greatly improve the speed and reliability of data transmission, and remove the barriers between the algorithm and model operations.

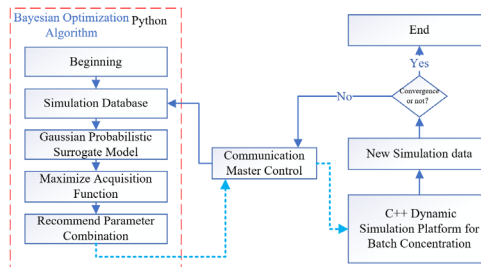


Figure 3 Logic diagram of Bayesian optimization for batch concentration

## 4. Simulation results and discussion

### 4.1. Raw simulation data set

Table 3 Raw Simulation Data

	<i>P</i>	$\Delta T$	$\Delta t_1$	$\Delta t_2$	Total cost		<i>P</i>	$\Delta T$	$\Delta t_1$	$\Delta t_2$	Total cost
1	5	0	10	20	99999	6	15	20	10	20	5736.6
2	5	10	10	20	7004.58	7	25	0	10	20	99999
3	5	20	10	20	7232.51	8	25	10	10	20	12764.32
4	15	0	10	20	99999	9	25	20	10	20	12764.57
5	15	10	10	20	4531.47						

Table 4 Iterative data

Iterations	<i>P</i>	$\Delta T$	$\Delta t_1$	$\Delta t_2$	Total cost	Iterations	<i>P</i>	$\Delta T$	$\Delta t_1$	$\Delta t_2$	Total cost
44	17.38	8.22	11.98	3.96	4443.72	285	5.04	8.39	12.68	4.25	4483.76
71	18.99	9.19	9.59	2.40	4409	320	16.55	2.02	6.65	5.97	4457.54
144	15.38	14.79	9.35	5.62	4407.43	350	24.77	6.19	8.04	7.64	4439.91

162	17.83	3.63	11.21	1.33	4442.43	428	19.05	18.22	11.34	2.16	4462.79
234	13.84	16.88	13.64	6.09	4433.9	467	15.88	14.20	10.14	5.24	4413.62

In order to obtain the initial data set of Bayesian optimization, 9 groups of different parameter combinations are selected, and target values of different combinations are calculated. As shown in Table 3, the selection of initial parameter combinations is kept in discrete state as far as possible, which is conducive to the rapid convergence of subsequent Bayesian optimization iterations.

#### 4.2. Results and discussion

Based on the initial data set, 500 iterations of Bayesian optimization calculations have been carried out for HME batch concentration process. Table 4 shows the optimal target values for every 50 iterations. It can be observed from the Table 4 that after 500 iterations, a lower target value is found, and the total cost of batch concentration is reduced to 4407.43 from 4531.47, the lowest in the initial dataset.

In terms of parameter selection, it can be seen from the iteration selection histogram of four parameters in Figure 4 that, with the support of simulation database and comprehensive consideration of the uncertainty of prediction in the Bayesian optimization iteration process, once a lower value is found, the search can be frequently continued near the point, which greatly improves the search efficiency. In terms of iteration results, after 144 iterations, it can be found that the best parameter combination is  $P=15.38\text{kpa}$ ,  $\Delta T=14.79\text{ }^\circ\text{C}$ ,  $\Delta t_1=9.35\text{ }^\circ\text{C}$ ,  $\Delta t_2=5.62\text{ }^\circ\text{C}$ , the minimum value of the objective function is 4407.43. As shown in Figure 5, the target value close to the lowest value has been obtained after 71 iterations. However, in the subsequent iteration process, there is no lower target value. The reason could be that as the times of iterations increases, the acquisition of the known value of the function increases, and the surrogate model is prone to over fitting. Therefore, the number of iterations can be appropriately reduced to avoid over fitting.

## 5. Conclusion

In this paper, Bayesian optimization algorithm based on Gaussian regression process and UCB is applied to optimize the batch concentration process of HME and obtain the optimal combination of control parameters, so as to minimize the total cost of batch concentration. After 144 iterations, the best control scheme and the lowest total cost value 4407.43 are obtained. It can be seen that Bayesian optimization algorithm has strong competitiveness in dynamic optimization, especially in the screening of control parameters. The surrogate model of Bayesian optimization also has strong fitting ability and good interpretability with the support of simulation database and comprehensive consideration of the uncertainty of prediction in iterations. In the subsequent work, the impact of different probabilistic surrogate models and acquisition functions on the optimization results can be compared.

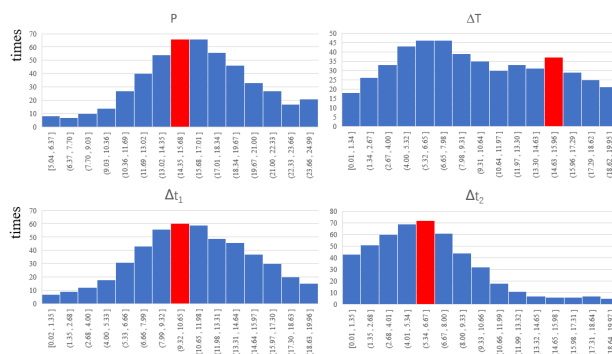


Figure 4 Iterations of each parameter

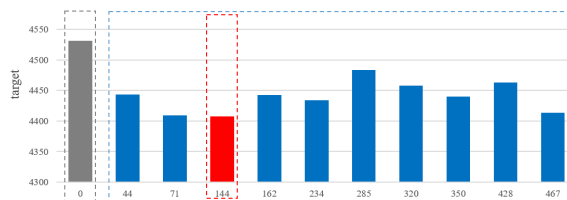


Figure 5 Optimal target values for every 50 iterations

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant numbers 2187081029).

## References

- F. J. Xue, Y. Z. Yang, X. M. Liu, 2010, Synthesis of Lamivudine, *Fine Chemicals*, 27, 6, 589-592.
- L. Yang, H. P. Pan, E. B. Zhang, 2017, Comprehensive Optimization of Batch Process based on Particle Swarm Optimization Algorithm, 29th Chinese Control And Decision Conference (CCDC), pp. 4504-4508.
- Y. S. Amini, M. B. Gerdroodbary, 2016, Optimal Control of Batch Cooling Crystallizers by using Genetic Algorithm, *Case Studies in Thermal Engineering*, 8, 8, 300-310.
- I. M. Mujtaba, N. Aziz, M. A. Hussain, 2006, Neural Network Based Modelling and Control in Batch Reactor, *Chemical Engineering Research and Design*, 84, 8, 635-644.
- B. Shahriari, K. Swersky, Z. Wang, 2016, Taking the human out of the loop: A review of Bayesiaoptimization, *Proc. of the IEEE*, 104, 1, 148-175.
- D. R. Jones, M. Chonlau, W. J. Welch, 1998, Efficient global optimization of expensive black-box functions, *Journal of Global Optimization*, 13, 4, 455-492.
- N. Srinivas, A. Krause, S. M. Kakade, 2010, Gaussian process optimization in the bandit setting: No regret and experimental design, *International Conference on Machine Learning*.





# Symbolic regression-based method for developing a physics-informed surrogate model for a manufacturing process

Utsav Awasthi,<sup>a</sup> and George M. Bollas,<sup>a</sup>

<sup>a</sup>*Department of Chemical and Biomolecular Engineering, Pratt and Whitney Institute of Advanced Systems Engineering, University of Connecticut, 159 Discovery Dr, Storrs, CT, 06269, USA.*

## Abstract

Hybrid models are used increasingly for the simulation of complex physical processes, because they combine the knowledge of the physics of a process and the data obtained from process measurements. This fusion of domain expert knowledge with data leads to hybrid models that are more accurate, can extrapolate, and obey laws of physics. Surrogate modeling can be cast as a hybrid modeling approach that builds reduced-order models for complex physical phenomena with simple model structure and low computational complexity. This study presents the application of hybrid modeling developed through a simple and accurate recursive symbolic regression algorithm for the complex and ill-understood phenomenon of tool wear in precision machining.

**Keywords:** Surrogate model, Symbolic Regression, Physics-explained AI.

## 1. Introduction

With the advent of Industry 4.0, manufacturing industries have undergone the digitalization of manufacturing processes. This digitalization has enabled the collection of better and large sets of sensor data to build data-driven models. Some popular methodologies for building data-driven models are neural networks and support vector machines (Han et al. 2021, 2022) . Though these models are widely used for process health monitoring in the manufacturing industries, they tend to be limited due to data idiosyncrasies. Hybrid models overcome this shortcoming. Hybrid models infuse domain knowledge of the physics of the process in the data-driven methodologies (Yang et al. 2022). Sansana et al. (2021) classified hybrid models based on the structure of the models broadly into three categories, serial structure, parallel structure, and surrogate models. A serial structure combines a hybrid model where the mechanistic part of the model either takes care of the known process behavior, and the data-driven part is used to capture the behavior, a feature of the process that cannot be directly estimated from the mechanistic model. A surrogate model is an approximate process model, and it is a simple representation of a complex phenomenon. These attributes of the surrogate models are helpful in building models for processes/phenomena that are difficult to model directly using data from the sensors or in cases where the physics of the process/phenomenon is too complex to model. In this study, a surrogate model is built to capture the progression of a fault in a manufacturing process. As is often the case in physical processes, the progression of a process fault is dynamic and depends on its state at any given time. Therefore, incorporating the fault state to predict its future values is important in modeling the process and fault accurately. This is accomplished by building dynamic or

recursive models. This work illustrates the development of a recursive symbolic regression model.

## 2. Surrogate modeling approaches

Surrogate models are approximate models of complex physical phenomena that can be used to represent an actual physical process with low complexity. These models require less computational effort and are used in process modeling and optimization. Over the years, several mathematical software have been developed to build surrogate models. Some of the commercially available software are Eureqa (Schmidt and Lipson 2009) (now DataRobot), AI Feynman: A physics-inspired method for symbolic regression (Udrescu and Tegmark 2020), Automated Learning of Algebraic Models for optimization ALAMO (Cozad, Sahinidis, and Miller 2014, 2015; Wilson and Sahinidis 2017), and the Genetic Programming Toolbox for Identification of Physical Systems (GPTIPS) (Searson, Leahy, and Willis 2010; Searson 2015). These software toolboxes exploit symbolic regression to develop surrogate models of response variables for selected input variables. Symbolic regression gives an algebraic expression of input variables. This approach is different from traditional regression techniques because the model structure and the regression coefficients are both evolved in symbolic regression. This helps in finding unique mathematical expressions that would best represent the data. Here, we choose GPTIPS to illustrate a methodology for recursive symbolic regression.

### 2.1. GPTIPS (*Genetic programming toolbox for identification of physical systems*)

GPTIPS is a MATLAB toolbox based on genetic programming (Searson 2015). Genetic programming is a biologically inspired machine learning method that uses the principles of evolution to perform a task or operation. To perform a task/operation, GP generates a random population of tree structures and then performs mutation, selection, and crossover over the best-performing tree to build a new population. This process of GP is repeated until a tree structure is obtained that fulfills the objective or, in the case of symbolic regression, gives a model of desired complexity and accuracy. GPTIPS utilizes the aforementioned features to build empirical mathematical models using data. GPTIPS generates models using multigene genetic programming with a linear combination of nonlinear transformation of input variables. GPTIPS has a variety of tournament selection methods, such as regular tournament selection, Pareto tournament selection, and lexicographic tournament selection. To build models for symbolic regression, GPTIPS has a large set of basis functions, +, -, \*, \, mult3(product of three terms), add3(sum of three terms), tanh, cos, sin, exp, ln (natural logarithm), x2, abs, x3, Ö, exp(-x), if-then-else, >, <, exp(x2), threshold and step functions, for model building. GPTIPS provides options to select a user-defined fitness criterion for model building (the default fitness metric is the root mean squared error, RMSE). GPTIPS provides an extensive list of hyperparameters, such as the number of generations, number of populations, maximum tree depth, maximum mutation depth, tournament size, crossover, mutation, elitism, and direct reproduction, that can be tuned to improve the model fitness. These hyperparameters decide on the complexity, and the model fitness is adjusted to get a model that does not overfit the data.

## 3. Application of surrogate modeling in manufacturing

### 3.1. *Precision machining and tool wear*

Precision machining is performed using Computer Numerical Control (CNC) machines (Awasthi and Bollas 2020). CNC machines use a circular tool with high-grade metal

inserts, rotating at high speed and removing material from the workpiece as metal chips. This process results in high shear stress, friction, and high temperature at the tool interface and workpiece interface (Awasthi et al. 2022). These aggressive conditions lead to the wear of the insert surface. This gradual wear of the insert surface is called tool wear. Analysis of data and understanding of the physics of the tool wear mechanism have identified that tool wear is a function of the operating conditions and the state of the tool. Using this insight, a recursive surrogate model of the tool wear as a function of the tool wear state was developed using GPTIPS. In addition to the tool wear state, the cutting conditions impact tool wear and were considered for model building. Hence, in this study, we utilize the recursive symbolic regression methodology to build a model for tool wear.

Table 1. Machine setting for experiments performed on the HAAS machine

Machine	# of parts	Diameter (mm)	Teeth	Width of cut (mm)	Depth of cut (mm)	Feed rate mm/min	Spindle speed (RPM)
HAAS	#1-#4	20	2	10.16	2.54	710.184	2330
HAAS	#5-#7	20	2	10.16	2.54	970.483	3184

### 3.2. Machining data collection

Machining data were collected by performing experiments in a HAAS Mini Mill machining center. The HAAS mini mill is a 3-axis compact machining center. To generate data, milling was performed on an AISI 4340 cylindrical steel block of 177.8 mm diameter, and 20 HRC hardness using a tool with two Kennametal inserts of grade KC725M and a lead angle of 90°. Each of these steel blocks was called a “Part.” A spiral tool path was selected to generate tool wear data and to make a cylindrical boss height of 10.16 mm and a diameter of 76.2 mm. One complete spiral tool path was called a “Run,” and while following the path, the tool went around the workpiece five times to reduce the diameter from 177.8 mm to 76.2 mm. The width of cut for this spiral path was 10.16 mm, and the depth of cut was 2.54 mm. At the end of the spiral tool path, when a boss of 76.2 mm diameter was obtained, the machining was stopped, and the flank wear measurement was taken. This whole process was repeated four times on a part to form a boss height of 10.16 mm and a diameter of 76.2 mm. This process was performed on the HAAS machine, and tool wear data were collected. A total of seven parts were machined, parts #1-#4 were machined for the machine settings of feed rate 710.184 mm/min and spindle speed 2330 RPM, and parts #5-#7 were machined for a feed rate of 970.483 mm/min and a spindle speed of 3184 RPM. Table 1 shows the machine settings for the experiments. In the case of the HAAS machine, the flank wear measurements were estimated using the values of the tool radius change. The tool diameter was measured after each run, and for the last run of part#7, the flank wear measurements were taken using a Keyence VHX-500, with a VH-Z100 lens and OP-72402 light ring microscope. To determine the flank wear, the tool radius change was calculated from the tool diameter measurements. In milling, the tool radius change follows the same trend as that of the flank wear. Therefore, by using the tool radius change values and the flank wear value from the last run, the flank wear values were estimated because both were assumed to have the same profile. Table 1 shows the values of the machine settings of spindle speed, width of cut, depth of cut, and feed rate, which are used to build a surrogate model for tool wear using GPTIPS.

### 3.3 Methodology for recursive symbolic regression

As shown in Fig. 1, the data of inputs and response variable were collected by performing experiments. Then GPTIPS was used to develop a recursive model of the input response

variable, which contains the machine settings and an initial guess of the state of wear. This guess was updated iteratively, and a new recursive model was generated for the updated value of the state. To update the response variable in the input, the symbolic regression model obtained from GPTIPS was used to calculate the actual response variable values using the machine settings and the predictions of the response variables recursively for all the time steps. Fig. 1 shows the methodology for building a recursive symbolic regression model for the case of machining. Machine settings were the inputs, and the tool wear was the response variable. This method was applied to build a recursive model for tool wear for a milling operation. In the figure, “ $i$ ” and “ $j$ ” are the counters, where the maximum value of  $j$  is  $j_{max}$  and the maximum value of  $i$  is  $N$  (number of datapoints).  $\hat{W}$  is the predicted and  $W$  the experimental tool wear, respectively.

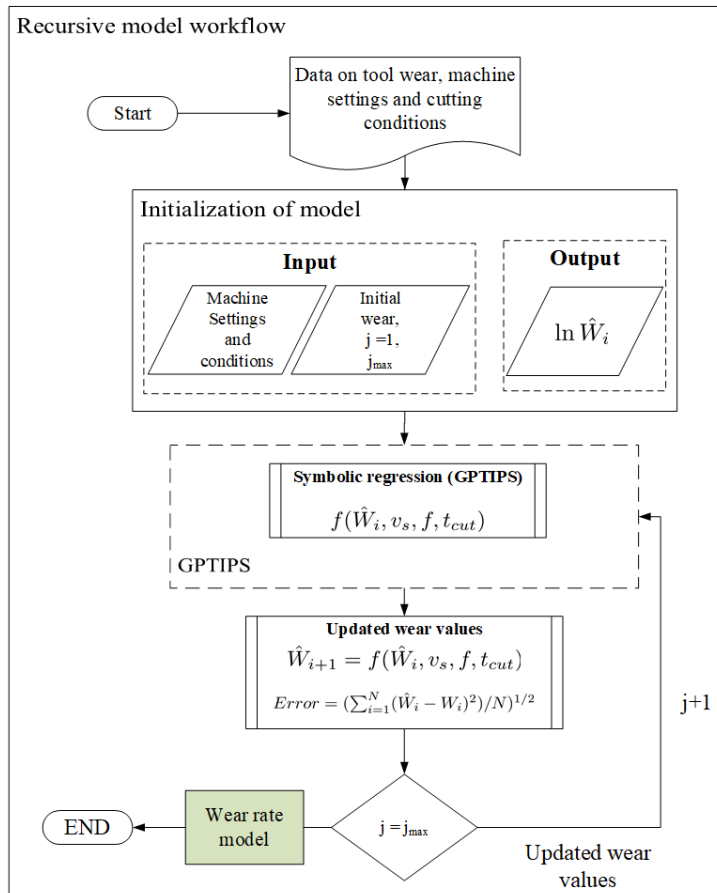


Figure 1: Workflow of the recursive model building methodology.

### 3.4 Surrogate model for tool wear

The recursive surrogate model for tool wear was developed using the methodology shown in Fig. 1. The inputs for model building were the spindle speed ( $v_s$ ), feed rate ( $f$ ), cutting time ( $t_{cut}$ ) and the natural logarithm of the tool wear ( $\ln \hat{W}_i$ ) at the current time step. The output of the model was the natural logarithm of the current wear ( $\ln \hat{W}_{i+1}$ ). The logarithm of the output was taken to reduce the variability of the data and assist in model fitting. Spindle speed and feed rate were the two input settings that changed during machining

and hence are included in model building. The complete experimental dataset was used for model building. Table 2 shows the basis function.  $\exp$ ,  $\ln$ ,  $x^3$ ,  $+$ ,  $*$ ,  $\setminus$ , used for model building. Default values were selected for crossover, elitism, mutation, and direct reproduction per-centages. The population size, number of generations, max. tree depth, and max. mutation depth were the parameters that were tuned to obtain good model fit. The number of iterations,  $j_{max}$ , were 10. After tuning the hyperparameters the best model obtained are shown in Table 2. The number of generations was 100, population size was 100, maximum tree depth was 3 and the maximum mutation depth was 2.

Table 2: GPTIPS toolbox settings for model building.

Basis functions: $\exp$ , $\ln$ , $x^3$ , $+$ , $*$ , $\setminus$
Population size : 100
Number of generations : 150
Tournament size : 10
Max. tree depth : 3
Max. mutate depth : 2
Crossover (%) : 84 (default)
Mutation (%) : 14 (default)
Elitism (%) : 15 (default)
Direct reproduction (%) : 2 (default)

$$\ln \widehat{W}_{i+1} = 108.438 \ln(t_{cut} + 3.0142) - 10.051 t_{cut} - 10.051 \ln t_{cut} - 1.768 \ln t_{cut}^2 - 4.856 t_{cut}^3 \ln \widehat{W}_i - 140.614 \dots (1)$$

Eq. (1) shows the expression of the recursive tool wear model obtained. Fig. 2 shows the plot of the actual tool wear values obtained from the experiment performed on the HAAS machine and the predictions obtained from the tool wear model. The model had an  $R^2$  value of 0.989, and as seen in the figure, the recursive tool wear model was able to capture the progression of tool wear in the machining process. The model shown in Eq. (1) is a function of the tool wear at the previous time step and cutting time. The tool wear increases with the increase in the cutting time and the term  $t_{cut}^3 \ln \widehat{W}_i$ , adds nonlinearity to the cutting profile to capture the effect of the state of the cutting tool and the cutting conditions. GPTIPS added terms  $t_{cut}$ ,  $\ln t_{cut}$ , and  $\ln t_{cut}^2$  to the term to fit the data. Overall, the model gives excellent agreement between predicted and measured tool wear for this process.

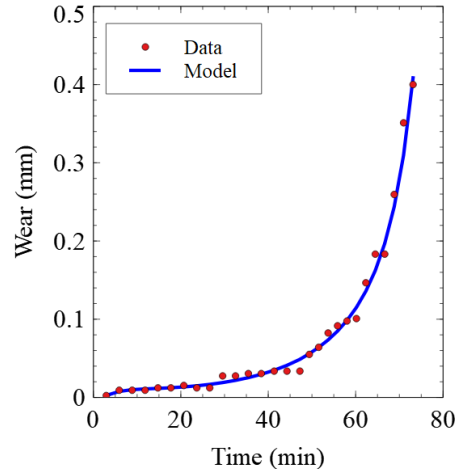


Figure 2: Tool wear predictions of the surrogate model for the wear data of the experiments performed on HAAS.

#### 4. Conclusion

This study demonstrated the use of symbolic regression to build a recursive surrogate model using an open-source genetic programming-based toolbox GPTIPS to perform symbolic regression. This approach helped to build models in the form of simple, transparent expressions, which are interpretable due to their algebraic form. This approach was applied to a manufacturing process to build a model for tool wear. A recursive tool wear model was developed, as a function of the tool wear state at any point in the manufacturing process. The model gave excellent predictions of the tool wear with a very simple and easy to interpret function.

## Acknowledgement

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Advanced Manufacturing Office Award Number DE-EE0007613. We also gratefully acknowledge the Air Force Research Laboratory, Materials and Manufacturing Directorate (AFRL/RXMS) for support via Contract No. FA8650-20-C-5206. Disclaimer: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## References

1. Searson, D.P., Leahy, D.E. & Willis, M.J., GPTIPS: an open source genetic programming toolbox for multigene symbolic regression, Proceedings of the International MultiConference of Engineers and Computer Scientists 2010 (IMECS 2010), Hong Kong, 17-19 March, 2010.
2. Searson, D.P., GPTIPS 2: an open-source software platform for symbolic data mining, Chapter 22 in Handbook of Genetic Programming Applications, A.H. Gandomi et al., (Eds.), Springer, New York, NY, 2015.
3. Awasthi, U., Wang, Z., Mannan, N., Pattipati, K. R., & Bollas, G. M. (2022). Physics-based modeling and information-theoretic sensor and settings selection for tool wear detection in precision machining. *Journal of Manufacturing Processes*, 81(December 2021), 127–140.
4. Awasthi, U., and Bollas, G. M. (2020). "Sensor Network Design for Smart Manufacturing – Application on Precision Machining." *IFAC PapersOnLine* 53 (2): 11440–45.
5. Han, S., Mannan, N., Stein, D. C., Pattipati, K. R., and Bollas, G. M. (2021). "Classification and Regression Models of Audio and Vibration Signals for Machine State Monitoring in Precision Machining Systems." *Journal of Manufacturing Systems*: 45–53.
6. Han, S., Yang, Q., Pattipati, K. R., and Bollas, G. M. (2022). "Sensor Selection and Tool Wear Prediction with Data-Driven Models for Precision Machining." *Journal of Advanced Manufacturing and Processing* 4 (4): e10143.
7. Yang, Q., Pattipati, K. R., Awasthi, U., and Bollas, G. M. (2022). "Hybrid Data-Driven and Model-Informed Online Tool Wear Detection in Milling Machines." *Journal of Manufacturing Systems* 63 (December 2021): 329–43.
8. Cozad, A., Sahinidis, N. V., and Miller, D. C. (2015). "A Combined First-Principles and Data-Driven Approach to Model Building." *Computers and Chemical Engineering* 73: 116–27. <https://doi.org/10.1016/j.compchemeng.2014.11.010>.
9. Cozad, A., Sahinidis, N. V., and Miller, D. C. (2014). "Learning Surrogate Models for Simulation-Based Optimization." *AIChE Journal* 60 (6): 2211–27.
10. Sansana, J., Joswiak, M. N., Castillo, I., Wang, Z., Rendall, R., Chiang, L. H., and Reis, M. S. (2021). "Recent Trends on Hybrid Modeling for Industry 4.0." *Computers and Chemical Engineering* 151: 107365.
11. Schmidt, M., and Lipson, H. (2009). "Distilling Free-Form Natural Laws from Experimental Data." *Science* 324 (5923): 81–85. <https://doi.org/10.1126/science.1165893>.
12. Udrescu, S. M., and Tegmark, M. (2020). "AI Feynman: A Physics-Inspired Method for Symbolic Regression." *Science Advances* 6 (16). <https://doi.org/10.1126/sciadv.aay2631>.
13. Wilson, Z., and Sahinidis, N. (2017). "ALAMO: Machine Learning from Data and First Principles." *Process Development Symposium 2018: Applying New Technologies in Process Development*, 38–48.

# Demand Response in Microgrids with Attention-Based Deep Reinforcement Learning

Jiahan Xie, Akshay Ajagekar, Fengqi You

*Cornell University, Ithaca, NY 14853, USA*

## Abstract

In this work, we propose a novel multi-agent deep reinforcement learning (MA-DRL) based approach to facilitate efficient load shaping through automated demand response in a microgrid. Achieving real-time autonomous demand response with energy management systems in buildings can be challenging due to factors like uncertain system parameters, the dynamic market price, and the complex coupled operational constraints. To develop a scalable approach for automated demand response in a microgrid, it is necessary to allow for coordination between buildings in the grid to smoothen the overall demand curve. We present a MA-DRL agent that utilizes an actor-critic algorithm incorporating a shared attention mechanism to enable an effective and scalable real-time cooperation to prevent the non-coordinated peak shifting actions in the complicated building system. A computational case study conducted with nine residential buildings revealed the attention-based MA-DRL agent's ability to execute decentralized cooperative policies without knowledge of building systems models or electricity price dynamics. Viability of the proposed control approach was also demonstrated by a reduction in net electricity consumption of over 6% accompanied by reduction in carbon emissions as compared to both conventional and state-of-the-art reinforcement learning approaches for automated demand response.

**Keywords:** Deep reinforcement learning, multi-agent, demand response, microgrids.

## 1. Introduction

As the integration of renewable energy sources into microgrids increases, demand response presents an opportunity to reduce the strain on the electric grid by reducing electricity usage in turn lowering greenhouse gas emissions (Pistikopoulos et al., 2021). By promoting the responsiveness of the buildings, demand response can lower the peak demand and the electricity consumption and thus improve the grid's stability in the long term (Yang et al., 2022). Factors like the variability of renewable energy sources, uncertainty associated with weather and load demand, as well as fluctuation of electricity prices further complicate the demand response problem (Silva et al., 2022). Demand response can be framed as an optimization problem to optimize electricity consumption or associated costs subject to coupled spatial and temporal constraints (Sun et al., 2021). Several math programming-based approaches have been previously applied in demand response for optimal scheduling of resources in a microgrid (Jordehi, 2019; Qiu, 2021). However, as the size of these problems increases with the number of buildings and the energy storage devices (Chen, 2022), metaheuristic solution techniques have also been proposed to obtain approximate solutions at the expense of reduced computational times (Shewale et al., 2020). Optimal control techniques like model predictive control have also been successfully incorporated for demand response (Qureshi et al., 2014). Such model-based approaches rely on first-principle models that



are often simplified and may fail in a practical setting owing to the non-ideal behavior of individual energy components (Chu and You, 2015). As a result, it is necessary to develop a comprehensive solution approach that is capable of handling various sources of uncertainties present in a microgrid while overcoming the need for an accurate first-principle model.

Reinforcement learning (RL) is a promising candidate for demand response problems owing to its model-free, and adaptive learning ability (Vázquez-Canteli and Nagy, 2019). Deep reinforcement learning (DRL) utilizes deep neural networks as nonlinear function approximators to overcome computational intractability caused by the increase in problem size. Many DRL-based techniques have been proposed for demand response and energy management in smart microgrids. DRL techniques assume no knowledge of the system dynamics and have been shown to optimize peak load in residential buildings (Mocanu et al., 2018) and minimize energy costs (Yu et al., 2020) by inducing coordination between the HVAC system and energy storage devices. Multi-agent variants of DRL techniques have also been explored to reduce operating costs and peak load through distributed scheduling of energy storage systems (Yang et al., 2019). Multi-agent deep reinforcement learning (MA-DRL) techniques can be directly applied to demand response problems (Vazquez-Canteli et al., 2020), however, they may add dynamicity to the microgrid operation and yield sub-optimal policies. Another challenge lies in ensuring coordination and information exchange between the agents in a multi-agent setting to obtain joint optimal policies. To overcome these challenges, we propose a novel multi-agent approach for demand response in grid-responsive buildings that employ an actor-critic based deep reinforcement learning technique to produce efficient energy management policies resulting in reduced overall electrical demand. Utilizing the attention mechanism in DRL (Iqbal and Sha, 2019) allows for centralized training with a shared critic network while ensuring decentralized execution governed by the multi-agent setting. We also perform several computational experiments to demonstrate that buildings in a microgrid can better learn how to coordinate to decrease the net electricity consumption accompanied by a reduction in carbon emissions as compared to both conventional and state-of-the-art DRL approaches for automated demand response.

## 2. Attention-based MA-DRL for demand response

### 2.1. Problem formulation

Multi-agent Markov decision processes (MMDPs), the generalization of the Markov decision process to the multi-agent setting, is applied to formulate the demand response problem in a grid of interconnected buildings. An MMDP is a tuple  $\langle S, A_1, \dots, A_N, T, R_1, \dots, R_N \rangle$ , where  $S$  is a set of states,  $A_1, \dots, A_N$  is a collection of action sets for  $N$  agents,  $T: S \times A_1 \times \dots \times A_N \rightarrow P(S)$  is the state transition function, which defines the probability distribution over possible next states given the current state and actions for each agent, and  $R_i: S \times A_1 \times \dots \times A_N \rightarrow \mathbb{R}$  is the reward function for each agent that is a function of global state and actions of all agents. As real-world settings often feature incomplete observations, each agent  $i$  receives its private observation  $o_i \in O_i$ , which contains partial information from the global state,  $s \in S$ .

For each building  $i$ , the agent receives observation  $o_{i,t}$  at time  $t$ . The observations contain the time variable accurate to an hour, district's temperature and humidity along with their associated predictions, district's solar radiation along with its prediction, buildings' indoor temperature and humidity, buildings' non-shiftable load, buildings' storage device state of charge (SoC). Each agent has up to three continuous actions: cooling device storage, DHW device storage, and electrical device storage. Since our goal is to minimize the net electricity consumption and the carbon emissions in the district, the reward function is related to the carbon intensity in the district at time  $t$ ,  $CI_t$ , net electricity consumption of building  $i$  at time  $t$ ,  $E_{i,t}$ , and the total net electricity consumption in the whole district,  $E_t = \sum_{i=1}^N E_{i,t}$ . Empirically, we found the following reward function is helpful to reduce the electricity consumption and carbon emissions and to perform more aggressive load flattening:  $500 \min(0, E_t) \times CI_t - 0.01 \text{sign}(E_{i,t}) \times |E_{i,t}| \times E_t^3$ .

## 2.2. Multi-agent attention-based control (MAAC)

In reinforcement learning, methods that learn approximations to both policy and value functions are called actor-critic methods, where 'actor' refers to the learned policy, and 'critic' refers to the learned value function (Li, 2017). The bias introduced through bootstrapping the expected returns with temporal-difference (TD) methods is often beneficial because it reduces variance introduced by the policy gradient estimator. Specifically, the policy gradient method aims to learn a parameterized policy that can select actions without consulting a value function. The gradient estimate is expressed as follows, where  $\theta$  parameterizes the policy network  $\pi$ .

$$\nabla_{\theta} J(\pi_{\theta}) = \nabla_{\theta} \log(\pi_{\theta}(a_t | s_t)) \sum_{t'=t}^{\infty} \gamma^{t'-t} r_{t'}(s_{t'}, a_{t'}) \quad (1)$$

For the learned value function, the action-value function is typically used,  $Q_{\psi}(s_t, a_t) = \mathbb{E}[\sum_{t'=t}^{\infty} \gamma^{t'-t} r_{t'}(s_{t'}, a_{t'})]$ .  $Q_{\psi}(s_t, a_t)$  is learned through TD learning by minimizing the mean squared Bellman error based on the past system transitions sampled from the experience replay buffer  $\mathcal{D}$ :

$$\mathcal{L}_Q(\psi) = \mathbb{E}_{(s,a,r,s') \sim \mathcal{D}} [(Q_{\psi}(s,a) - y)^2] \quad (2)$$

where  $y = r(s,a) + \gamma \mathbb{E}_{a' \sim \pi(s')} [Q_{\bar{\psi}}(s', a')]$

To encourage exploration and avoid converging to non-optimal policies, the maximum entropy framework is adopted in the soft-actor-critic algorithm [11]:

$$\nabla_{\theta} J(\pi_{\theta}) = \mathbb{E}_{s \sim \mathcal{D}, a \sim \pi} [\nabla_{\theta} \log(\pi_{\theta}(a | s)) (-\alpha \log(\pi_{\theta}(a | s)) + Q_{\psi}(s,a) - b(s))] \quad (3)$$

where  $b(s)$  is a state-dependent baseline. The loss function for TD learning of the value function is also revised accordingly:

$$y = r(s,a) + \gamma \mathbb{E}_{a' \sim \pi(s')} [Q_{\bar{\psi}}(s', a') - \alpha \log(\pi_{\bar{\theta}}(a' | s'))] \quad (4)$$

The attention mechanism is used for each agent to query other agents for information about their observations and actions and incorporates that information into calculating its action-value function  $Q_i^\psi(o, a)$  for agent  $i$ . Precisely, the action-value for agent  $i$  is calculated as follows:

$$Q_i^\psi = f_i(g_i(o_i, a_i), x_i) \quad (5)$$

where  $f_i$  is a two-layer multi-layer perceptron (MLP),  $g_i$  is a one-layer MLP embedding function,  $x_i$  denotes the total contribution from other agents,  $h$  is an MLP with  $V$  as the value matrix.

$$x_i = \sum_{j \neq i} \alpha_j v_j = \sum_{j \neq i} \alpha h(V g_j(o_j, a_j)) \quad (6)$$

The attention weight  $\alpha_j$  compares the embedding  $e_j$  with  $e_i = g_i(o_i, a_i)$ , using the query-key system (Vaswani et al., 2017) and apply softmax over the similarity value between these two embeddings,

$$\alpha_j = \frac{\exp((W_k e_j)^T W_q e_i)}{\sum_{j=1}^N \exp((W_k e_j)^T W_q e_i)} \quad (7)$$

where  $W_q$  transforms  $e_i$  into a "query" and  $W_k$  transforms  $e_j$  into a "key". Because of the sharing of the set of parameters  $W_q$ ,  $W_k$ , and  $V$ , all critics are updated to minimize a joint regression loss function:

$$\mathcal{L}_Q(\psi) = \sum_{i=1}^N \mathbb{E}_{(o, a, r, o') \sim \mathcal{D}} [(Q_i^\psi(o, a) - y_i)^2] \quad (8)$$

where  $y_i = r_i + \gamma \mathbb{E}_{a' \sim \pi_\sigma(o')} [Q_i^{\bar{\psi}}(o', a') - \alpha \log(\pi_{\theta_i}(a_i | o_i))]$

Likewise, the individual policies are updated by gradient ascent on the following gradient:

$$\begin{aligned} \nabla_{\theta_i} J(\pi_{\theta_i}) = & \mathbb{E}_{o \sim \mathcal{D}, a \sim \pi} [\nabla_{\theta_i} \log(\pi_{\theta_i}(a_i | o_i)) (-\alpha \log(\pi_{\theta_i}(a_i | o_i)) \\ & + Q_i^\psi(o, a) - b(o, a_{\setminus i}))] \end{aligned}$$

where  $\setminus i$  is denoted as the set of all agents except  $i$ , and  $b(o, a_{\setminus i})$  is the multi-agent baseline. As a controller, each building will store transition tuples  $(o_t, a_t, o_{t+1}, r_{t+1})$  into its corresponding replay buffer  $\mathcal{D}$ . When the replay buffer stores enough experience and the system reaches the update time  $T_{update}$ , we sample a mini-batch  $\mathcal{B}$  and use the batch to update the critics and actors networks. The critics will take in all policy networks and perform enhanced updates through the attention mechanism. During the execution phase, each agent will download the latest updated policy networks and take actions only according to their local observations. This decentralized execution has manifested better computational efficiency.

### 3. Case study: Computational results

We compare our results with three state-of-the-art RL controllers and one Rule-based controller (RBC). The three RL controllers include Deep Deterministic Policy (DDPG) controllers, Soft-Actor-Critic (SAC) controllers, and the Multi-Agent RL with Iterative Sequential Action Selection (MARLISA) controllers [10]. The DDPG controllers and the SAC controllers do not have explicit information sharing and coordination, while the MARLISA controllers have information sharing of buildings' own future electricity consumption following a leader-follower schema. The RBCs take uniform and greedy actions, they charge 9.1% of their storage devices' maximum capacities every hour from 10pm to 8am, and release 8% of their maximum capacities every hour from 9am to 9pm. We conduct simulations with the energy demand for each building in New Orleans using a pre-recorded dataset of nine buildings. Each building consists of an air-to-water heat pump, as well as an electric heater that supplies them with DHW. The primary grid provides electricity for these devices, other electrical equipment, and appliances. Meanwhile, buildings can generate electricity using photovoltaic generation to offset some of their electricity consumption. We run simulations on the same weather data for 10 epochs to test the performance and the convergence of different RL controllers. All the RL controllers perform random action exploration for the first 250 days during the first epoch. For the SAC, MARLISA, and MAAC controllers, they perform stochastic policies for 300 more days by sampling from the policies distribution with entropy regularization. For the DDPG controllers, they have uncorrelated, mean-zero Gaussian noise for 300 more days of exploration. After 550 days into the simulation, the SAC, MARLISA, and MAAC controllers remove stochasticity and exploit their policies by taking the mean action. The DDPG controllers' actions noises scale to 0 for taking deterministic actions. The Adam optimizer is employed to update all neural networks.

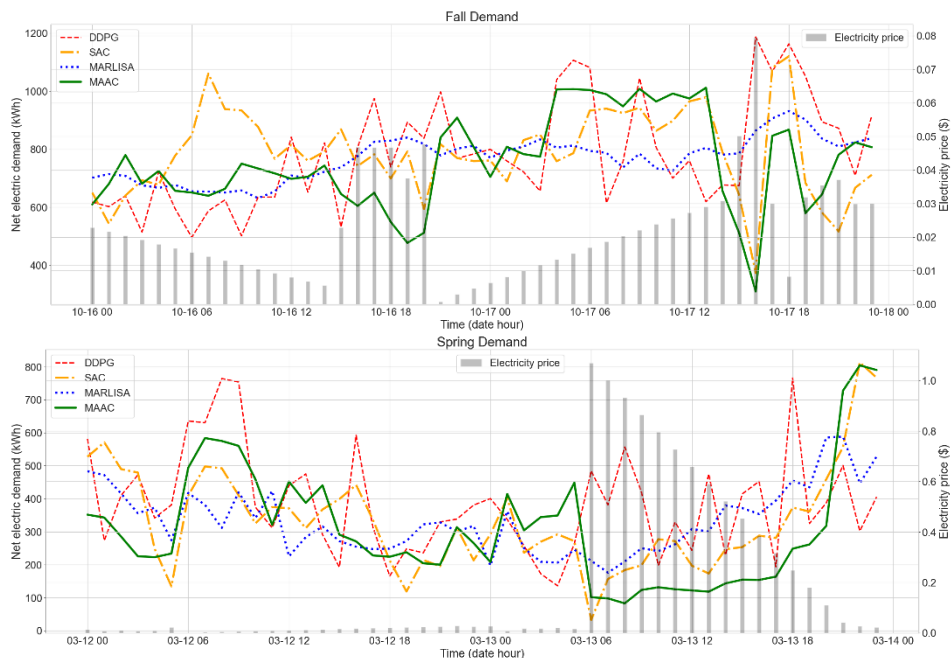


Figure 1. Net electricity consumption of the buildings in the fall and spring seasons over a period of 48 hours along with the electricity pricing.

In Figure 1, we visualize the electricity demand under different RL control algorithms for a fixed period of 48 hours during the fall and spring seasons. With the proposed controller, the electricity demand starts decreasing concurrently with the increase in price, and the demand rises back when the price is lower, during which the agent can store more energy for the preparation of the next peak price period. As the simulation results have demonstrated, all the RL controllers react appropriately to price fluctuations and reduce electric demand except the DDPG method. Since DDPG relies on all kinds of tuning and tends to overestimate the Q-values, it is common to see that the controllers exploit problematic Q-functions and have learned poor policy. Furthermore, during this high-demand period, the MAAC controllers demonstrate their strong ability compared to two other RL algorithms that are also using the soft-actor-critic framework. As the demand curves have shown, the attention-based MA-DRL controllers ensure stability and less fluctuation of the demand curve. We can conclude that the attention-based MA-DRL controllers perform better in minimizing electricity consumption while performing load shaping in the meantime. Their superior abilities to reduce energy consumption and flatten the energy demand curve are essential for the reliability of the microgrids.

#### 4. Conclusion

In this paper, we introduced an attention-based MA-DRL demand response algorithm that can achieve better cooperation between the interconnected buildings to reduce electricity consumption and provide effective demand flattening in a decentralized way in the absence of any prior knowledge of the system model. Simulation results demonstrated the convergence and the efficacy of our proposed attention-based MA-DRL controllers. They achieve a significant reduction in net electricity consumption accompanied by a reduction in carbon emissions compared to both conventional and state-of-the-art RL approaches for automated demand response.

#### References

- W. Chen, F. You, 2022, Sustainable building climate control with renewable energy sources using nonlinear model predictive control. *Renewable & Sustainable Energy Reviews*, 168, 112830.
- Y. Chu, F. You, 2015, Model-based integration of control and operations: Overview, challenges, advances, and opportunities. *Computers & Chemical Engineering*, 83, 2-20.
- J. Gong, F. You, 2015, Sustainable design and synthesis of energy systems. *Current Opinion in Chemical Engineering*, 10, 77-86.
- S. Iqbal and F. Sha, 2019. Actor-attention-critic for multi-agent reinforcement learning. *International conference on machine learning*, PMLR.
- A. R. Jordehi, 2019, Optimisation of demand response in electric power systems, a review, *Renewable and Sustainable Energy Reviews* 103: 308-319.
- Y. Li, 2017, Deep reinforcement learning: An overview, arXiv preprint arXiv:1701.07274.

- E. Mocanu, D. C. Mocanu, P. H. Nguyen, A. Liotta, M. E. Webber, M. Gibescu and J. G. Sloatweg, 2018, On-line building energy optimization using deep reinforcement learning, IEEE Transactions on Smart Grid 10(4): 3698-3708.
- E.N. Pistikopoulos, A. Barbosa-Povoa, J.H. Lee, et al., 2021, Process systems engineering – The generation next? Computers & Chemical Engineering, 147, 107252.
- H. Qiu, W. Gu, F. You, 2021, Bilayer Distributed Optimization for Robust Microgrid Dispatch With Coupled Individual-Collective Profits. IEEE Transactions on Sustainable Energy, 12, 1525-1538.
- H. Qiu, F. You, 2020, Decentralized-distributed robust electric power scheduling for multi-microgrid systems. Applied Energy, 269, 115146.
- F. A. Qureshi, T. T. Gorecki and C. N. Jones, 2014, Model Predictive Control for Market-Based Demand Response Participation, IFAC Proceedings Volumes 47(3): 11153-11158.
- C. Shang, F. You, 2019, Data Analytics and Machine Learning for Smart Process Manufacturing: Recent Advances and Perspectives in the Big Data Era. Engineering, 5, 1010-1016.
- A. Shewale, A. Mokhade, N. Funde and N. D. Bokde (2020) An Overview of Demand Response in Smart Grid and Optimization Techniques for Efficient Residential Appliance Scheduling Problem. Energies 13, DOI: 10.3390/en13164266
- C. Silva, P. Faria, Z. Vale and J. M. Corchado, 2022, Demand response performance and uncertainty: A systematic literature review, Energy Strategy Reviews 41: 100857.
- L. Sun, F. You, 2021, Machine Learning and Data-Driven Techniques for the Control of Smart Power Generation Systems: An Uncertainty Handling Perspective. Engineering, 7, 1239-1247.
- A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, Ł. Kaiser and I. Polosukhin, 2017, Attention is all you need, Advances in Neural Information Processing Systems 30.
- J. R. Vázquez-Canteli, G. Henze and Z. Nagy, 2020. MARLISA: Multi-agent reinforcement learning with iterative sequential action selection for load shaping of grid-interactive connected buildings. Proceedings of the 7th ACM international conference.
- J. R. Vázquez-Canteli and Z. Nagy, 2019, Reinforcement learning for demand response: A review of algorithms and modeling techniques, Applied Energy 235: 1072-1089.
- S. Yang, H. Oliver Gao, F. You, 2022, Model predictive control for Demand- and Market-Responsive building energy management by leveraging active latent heat storage. Applied Energy, 327, 120054.
- Y. Yang, J. Hao, Y. Zheng and C. Yu, 2019. Large-Scale Home Energy Management Using Entropy-Based Collective Multiagent Deep Reinforcement Learning Framework. IJCAI.
- L. Yu, W. Xie, D. Xie, Y. Zou, Z. Sun, Y. Zhang and T. Jiang, 2020, Deep Reinforcement Learning for Smart Home Energy Management, IEEE Internet of Things Journal 7(4): 2751-2762.



# Multi-Kernel Canonical Variate Analysis with Bayesian Optimized Kernel Designs for Nonlinear System Identification

Jan Vincent Madayag,<sup>a</sup> Karl Ezra Pilario,<sup>a,\*</sup>

<sup>a</sup>*Process Systems Engineering Laboratory, University of the Philippines, Diliman, Quezon City, 1101, Philippines*

*\*Corresponding author, email: [kspilario@up.edu.ph](mailto:kspilario@up.edu.ph)*

## Abstract

Kernel Canonical Variate Analysis (KCVA) is a widely used method for nonlinear system identification. Recently, it was shown that combining the Gaussian kernel with a linear kernel can improve the generalization ability of KCVA models. However, there are no known results for multi-kernel CVA as of yet, that is, when more than two kernels are combined to capture the nonlinearities from the process data. In this work, a multi-kernel designs involving two to four kernel functions are investigated in KCVA. Bayesian optimization (BayesOpt) is also proposed to tune kernel parameters by fitting a surrogate Gaussian process on the loss surface, and then searching the next best kernel mixture efficiently via the Expected Improvement policy. We demonstrate that BayesOpt is more data-efficient than other stochastic search methods in finding kernel designs. Results were verified using a series of empirical trials in an benchmark evaporator case study.

**Keywords:** system identification, multiple kernel learning, canonical variate analysis, Newll-Lee evaporator, Gaussian process regression.

## 1. Introduction

Most aspects of process control and operations inherently require a model of the physical process that accurately describe the process dynamics. In order to build these models, two main approaches exist, namely, first-principles modelling or system identification. Due to the highly complex nature of physical systems, first-principles modelling can become too tedious to perform. Hence, system identification (SysID) techniques are more preferred, since they only require an assumed general model structure to be fitted to the real process input-output data.

Many SysID methods already exist, but among these, the subspace methods are favorable since they only require matrix algebra to estimate the model from data, making them fast and non-iterative, whereas prediction error methods require an iterative minimization of a cost function to find an accurate model (Qin, 2006). For processes operating at widely varying operating conditions, nonlinear model structures must be assumed, since nonlinear dynamic behaviors are expected to appear in the data statistics. One of the recently developed nonlinear subspace methods is the feature-relevant mixed kernel canonical variate analysis (FR-MKCVA) (Pilario *et al.*, 2020), which proposed a new kernel design for kernel canonical variate analysis (KCVA) based subspace identification. The proposed kernel is a mixture of a global kernel, given by the linear kernel, and a local



kernel, given by the Gaussian radial basis function (RBF) with automatic relevance determination. Since the kernel is a convex combination of global and local types, it has been demonstrated to improve the generalization ability of the fitted model for the Newell-Lee evaporator case study (Al-Seyab & Cao, 2008). Even with a 25% increase in the input amplitude, the FR-MKCVA model provided superior fitness than linear CVA and recurrent neural nets. The KCVA framework has also been applied to fault detection (Samuel & Cao, 2015) due to its ability to handle nonlinear dynamic processes.

In this work, the FR-MKCVA model is extended by analyzing the impact of other kernel designs such as when the sigmoid and Laplacian kernels are included. We investigate kernel mixtures consisting of two to four kernel functions being combined in a convex manner. Hence, we propose a novel multi-kernel CVA algorithm. Note, however, that the search for kernel parameters become increasingly difficult with larger kernel mixtures. Hence, we also propose in this work a more robust and efficient way to tune kernel parameters via the use of Bayesian Optimization (BayesOpt). We compared the computational efficiency and accuracy of the subspace models when tuned by BayesOpt to those tuned by alternative stochastic algorithms commonly used in literature, such as particle swarm optimization (PSO), genetic algorithm (GA), simulated annealing (SA), and random search (RS). Using the Newell-Lee evaporator case study, the results show that BayesOpt is more efficient, hence, more recommended than other search methods, and that the addition of new kernel mixtures can help improve the accuracy of the subspace models as well.

This paper is organized as follows. Section 2 covers preliminary background on FR-MKCVA. Section 3 presents the BayesOpt method used in this work, as well as the kernel designs being investigated. Section 4 gives the results and discussion for each case study. Finally, Section 5 contains a summary of this work.

## 2. Feature-Relevant Mixed Kernel Canonical Variate Analysis

The FR-MKCVA method is an improved subspace identification method from the original canonical variate analysis (CVA) proposed by Larimore (1990). The idea in FR-MKCVA is to identify a nonlinear dynamic model with the state-space structure:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{f}(\mathbf{u}_k, \mathbf{y}_{k-1}) + \mathbf{w}_k \quad (1)$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{f}(\mathbf{u}_k, \mathbf{y}_{k-1}) + \mathbf{v}_k \quad (2)$$

where  $\mathbf{x}_k$ ,  $\mathbf{y}_k$ ,  $\mathbf{u}_k$  are the states, outputs, and inputs, respectively,  $\mathbf{w}_k$ ,  $\mathbf{v}_k$  are the process noise and measurement noise vectors, respectively,  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ ,  $\mathbf{D}$  are the state-space matrices, and  $\mathbf{f}(\mathbf{u}_k, \mathbf{y}_{k-1})$  is some nonlinear mapping of the input-output data where the mapping is defined by a set of kernel functions (Pilario *et al.*, 2020). To identify the model, the input-output data are first projected onto the kernel subspace using kernel principal components analysis (KPCA). KPCA employs a kernel function  $K(\mathbf{x}, \mathbf{x}')$  to act as a dot product in some nonlinear space. Through KPCA, a nonlinear set of features can be extracted from the data as follows:

$$\mathbf{f}(\mathbf{u}_k, \mathbf{y}_{k-1}) = [\mathbf{t}_k] = \left[ \sum_{j=1}^N \alpha_j^T K(\mathbf{z}_k, \mathbf{z}_j') \right]_{i=1, \dots, l} \quad (3)$$

where  $\mathbf{z}_k = [\mathbf{u}_k \mathbf{y}_k]^T$  is the input-output data set at the  $k$ th sampling instant,  $N$  is the number of training samples,  $\boldsymbol{\alpha}$  are the KPCA loadings, and  $l$  is the number of principal components. The  $l$  is chosen according to the 99% cumulative percent variance rule. The kernel design in Eq. (3) that was proposed in Pilario *et al.* (2020) is given in Table 1. After performing KPCA, the CVA algorithm is now employed. First, the features from Eq. (3) are organized into past and future Hankel matrices,  $\mathbf{Y}_p$  and  $\mathbf{Y}_f$ . Their covariances are given by  $\boldsymbol{\Sigma}_{pp} = \text{cov}(\mathbf{Y}_p, \mathbf{Y}_p)$ ,  $\boldsymbol{\Sigma}_{ff} = \text{cov}(\mathbf{Y}_f, \mathbf{Y}_f)$ , and  $\boldsymbol{\Sigma}_{fp} = \text{cov}(\mathbf{Y}_f, \mathbf{Y}_p)$ . CVA proceeds by taking the singular value decomposition (SVD) of the Hankel matrix, that is  $\boldsymbol{\Sigma}_{ff}^{-1/2} \boldsymbol{\Sigma}_{fp} \boldsymbol{\Sigma}_{pp}^{-1/2} = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^T$ . The CVA states are now estimated from the data as  $\mathbf{X} = [\mathbf{x}_k] = \mathbf{V}_n^T \boldsymbol{\Sigma}_{pp}^{-1/2} \mathbf{Y}_p$  where  $\mathbf{V}_n$  represents the first  $n$  singular vectors from the matrix  $\mathbf{V}$ , signifying that only the subspace of the first  $n$  states are taken as the principal subspace. Finally the state-space matrices are computed from  $\mathbf{X} = [\mathbf{x}_k]$ ,  $\mathbf{U} = [\mathbf{u}_k]$ , and  $\mathbf{Y} = [\mathbf{y}_k]$ , as follows:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \text{cov} \left[ \begin{pmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_k \end{pmatrix}, \begin{pmatrix} \mathbf{x}_k \\ \mathbf{t}_k \end{pmatrix} \right] \cdot \left\{ \text{cov} \left[ \begin{pmatrix} \mathbf{x}_k \\ \mathbf{t}_k \end{pmatrix}, \begin{pmatrix} \mathbf{x}_k \\ \mathbf{t}_k \end{pmatrix} \right] \right\}^{-1} \quad (4)$$

where  $\mathbf{t}_k$  are the KPCA-extracted nonlinear features from Eq. (3).

### 3. Investigating Multi-Kernel Canonical Variate Analysis

#### 3.1. Kernel Designs

The design of the kernel function in Eq. (3) has a large impact in the performance of the identified model. In this work, various kernel designs are tested, shown in Table 1. The first kernel, namely Linear + Isotropic Gaussian RBF, only has 2 kernel parameters to tune, which are the mixture weight  $\omega$  and the kernel width  $\gamma$ . Since the hyper-parameter search space is only 2-D, this mixed kernel was used to compare the performance of BayesOpt against other search methods: Particle swarm optimization, genetic algorithm, simulated annealing, and random search.

The main metric for comparison is the  $R^2$  fitness of the model to the validation data:

$$R^2 = 100\% \times \left( 1 - \frac{\|\mathbf{y}_m - \mathbf{y}\|^2}{\|\mathbf{y}_m - \text{mean}(\mathbf{y}_m)\|^2} \right) \quad (5)$$

where  $\mathbf{y}$  are the predicted outputs from the model and  $\mathbf{y}_m$  are the reference measured outputs. The closer the  $R^2$  value to 100%, the more accurate the model. Aside from the  $R^2$ , the optimizers are also evaluated by their computational time given that all of them are provided the same budget of 100 maximum objective function evaluations.

#### 3.2. Bayesian Optimization

Bayesian optimization is a fast and efficient optimizer for black-box functions that are computationally expensive to sample. It has found many current applications such as in recommender systems, reinforcement learning, and hyper-parameter tuning in machine learning models (Shahriari *et al.*, 2016). For these reasons, BayesOpt is proposed in this work for tuning kernel parameters in multi-kernel CVA.

In BayesOpt, the adopted objective function is the fitness of the multi-kernel CVA model to the validation data. The goal is to find the kernel parameters  $\boldsymbol{\theta}$  that maximize the  $R^2$

value on a held-out validation data set. A surrogate Gaussian process (GP) model is chosen to estimate the mean  $m(\boldsymbol{\theta})$  and variance  $var(\boldsymbol{\theta})$  of the validation  $R^2$  as follows:

$$m(\boldsymbol{\theta}) = \mathbf{K}_*^T [\mathbf{K} + \sigma^2 \mathbf{I}]^{-1} \mathbf{y} \quad (6)$$

$$var(\boldsymbol{\theta}) = k(\boldsymbol{\theta}_*, \boldsymbol{\theta}_*) - \mathbf{K}_*^T [\mathbf{K} + \sigma^2 \mathbf{I}]^{-1} \mathbf{K}_* \quad (7)$$

where  $\mathbf{K}_* = k(\boldsymbol{\theta}_*, \boldsymbol{\theta})$  is the kernel matrix,  $k(\boldsymbol{\theta}, \boldsymbol{\theta}')$  denotes the kernel function in the GP,  $\sigma^2$  is the assumed noise variance and  $\mathbf{y}$  contains the validation  $R^2$  values. In this work, the chosen GP kernel is the *squared exponential* function, which is equivalent to the Gaussian RBF. Next, the Expected Improvement (EI) acquisition function is used to find the next best set of kernel parameter values  $\boldsymbol{\theta}_{next}$  to sample:

$$\boldsymbol{\theta}_{next} = \operatorname{argmax} EI(\boldsymbol{\theta}) \quad (8)$$

$$EI(\boldsymbol{\theta}) = \Delta(\boldsymbol{\theta}) \Phi \left( \frac{\Delta(\boldsymbol{\theta})}{\sqrt{var(\boldsymbol{\theta})}} \right) + \sqrt{var(\boldsymbol{\theta})} \phi \left( \frac{\Delta(\boldsymbol{\theta})}{\sqrt{var(\boldsymbol{\theta})}} \right) \quad (9)$$

where  $\Delta(\boldsymbol{\theta}) = m(\boldsymbol{\theta}) - y_i^* - \xi$ , the  $\xi$  is an exploration/exploitation parameter, the  $y_i^*$  is the maximum  $R^2$  observed so far at the  $i$ th iteration, and the  $\Phi$  and  $\phi$  are the cumulative density function and probability density function of the standard normal distribution, respectively. In this work, the max number of function evaluations is set to 100 and the  $\xi$  is made to decrease linearly from 20 to 0. The EI policy is to first explore the search space by favoring regions with high uncertainty, and then later on exploit on the regions where previous maxima were encountered. This balance between exploration and exploitation makes BayesOpt a data-efficient search algorithm (Shahriari *et al.*, 2016).

Table 1: Kernel functions investigated in this work.

Name and Function	Optimizer
Linear + Isotropic Gaussian RBF $K(\mathbf{z}, \mathbf{z}') = \omega(\mathbf{z}^T \mathbf{z}' + 1) + (1 - \omega) \exp\left(\frac{-\ \mathbf{z} - \mathbf{z}'\ ^2}{\gamma}\right)$	RS, GA, SA, PSO, BayesOpt
Anisotropic Gaussian RBF only $K(\mathbf{z}, \mathbf{z}') = \exp(-(\mathbf{z} - \mathbf{z}')^T \boldsymbol{\Gamma} (\mathbf{z} - \mathbf{z}'))$ $\boldsymbol{\Gamma} = \operatorname{diag}(1/\gamma_1, 1/\gamma_2, \dots)$	BayesOpt
Linear + Anisotropic Gaussian RBF (Pilario <i>et al.</i> , 2020) $K(\mathbf{z}, \mathbf{z}') = \omega(\mathbf{z}^T \mathbf{z}' + 1) + (1 - \omega) \exp(-(\mathbf{z} - \mathbf{z}')^T \boldsymbol{\Gamma} (\mathbf{z} - \mathbf{z}'))$ $\boldsymbol{\Gamma} = \operatorname{diag}(1/\gamma_1, 1/\gamma_2, \dots)$	BayesOpt
Linear + Isotropic Gaussian RBF + Sigmoid + Laplacian $K(\mathbf{z}, \mathbf{z}') = \omega_1(\mathbf{z}^T \mathbf{z}' + 1) + \omega_2 \exp\left(\frac{-\ \mathbf{z} - \mathbf{z}'\ ^2}{\gamma_1}\right)$ $+ \omega_3 \tanh(a\mathbf{z}^T \mathbf{z}' + b) + \omega_4 \exp\left(\frac{-\ \mathbf{z} - \mathbf{z}'\ }{\gamma_2}\right)$ $\omega_1 + \omega_2 + \omega_3 + \omega_4 = 1$	BayesOpt

### 4. Case Study

The full description of the Newell-Lee evaporator system can be found in Pilario *et al.* (2020). It consists of multiple subsystems having 3 inputs and 3 outputs overall. The system was simulated in MATLAB Simulink and it is available online through the following link: <https://www.mathworks.com/matlabcentral/fileexchange/68641-newell-lee-evaporator-system-for-system-identification>.

Since the process has six variables in total, the corresponding search space dimensions for kernel parameter tuning are 2-D, 6-D, 7-D, and 8-D for each kernel mixture proposed in Table 1, respectively. The number of states for this case study is fixed at  $n = 4$ , based on results in Pilario *et al.* (2020).

#### 4.1. Results on Bayesian Optimization

Fig. 1(a) compares various hyper-parameter optimizers in terms of computation time and maximum found, given the same budget of 100 max objective function evaluations. The box plots in this figure summarize the result of 10 trials in each optimizer. BayesOpt-2D is clearly more efficient than the other 2-D optimizers due to having one of the lowest computation times, yet highest maximum found. PSO was found to be the slowest, but it also found consistently high max values. Random Search, although fast, struggles to find the maximum. SA has a better performance than GA, but both are still less efficient than BayesOpt. To further illustrate the efficiency of BayesOpt-2D, Fig. 1(b) and (c) show, respectively, the EI surface and the surrogate GP surface laden with the sampled points. As shown, BayesOpt-2D concentrated more on the promising area near the maximum rather than needlessly exploring across the entire surface. Fig. 1(d) shows the *exact* objective function surface, which, upon comparing with Fig. 1(c), indicates that the surrogate GP can closely estimate the 2-D function given only the sampled points. The consistency of the result of BayesOpt-2D across 10 trials can also be verified in Fig. 1(e), since its converged solutions are more consistent than those from other optimizers.

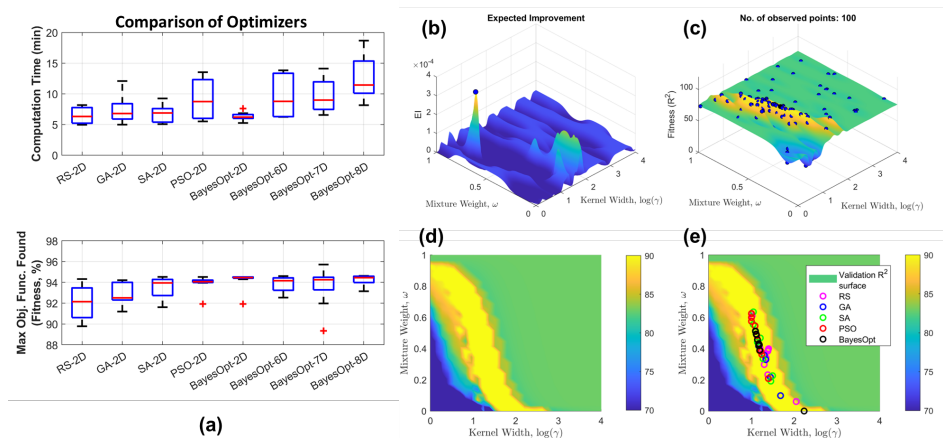


Fig. 1. Optimization results for the evaporator case study: (a) Comparison of computation time and max objective function values obtained from the 2-D, 6-D, 7-D, and 8-D hyper-parameter search repeated for 10 trials each; (b) Sample EI surface at the last iteration of a BayesOpt 2-D run; (c) A surrogate validation  $R^2$  surface showing the 100 points where the algorithm sampled the objective function; (d) The exact validation  $R^2$  surface, top view; (e) Same as (d) but showing the final converged solutions of the various optimizers.

#### 4.2. Results on the Testing of Multi-Kernel CVA models

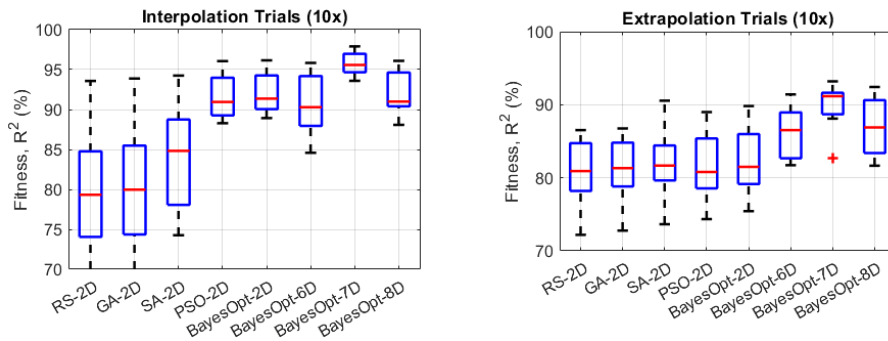


Fig. 2. Testing phase results in the evaporator case study. Each box plot summarizes 10 trials.

In Fig. 2, the validated multi-kernel CVA models are subjected to 10 interpolation data sets and 10 extrapolation data sets (with 25% increase in input amplitude) for testing. These data sets differ in the random input sequences and the random seed for noise generation. Results show that the best model is still the FR-MKCVA, but now Bayesian optimized (BayesOpt-7D) rather than subjected to random search as in Pilario *et al.* (2020). Also, the multi-kernel CVA with 4 kernels (BayesOpt-8D) was seen to be overfitting, since its performance is worse than FR-MKCVA. This indicates that 4 kernel functions may be too complex to validate with such a limited validation data set.

### 5. Conclusion

In this work, various multi-kernel designs were investigated for KCVA based nonlinear system identification. In addition, Bayesian optimization was demonstrated to be an efficient hyper-parameter search method for tuning these kernel designs. Using the benchmark Newell-Lee evaporator case study, the Bayesian optimized multi-kernel consisting of the linear and anisotropic Gaussian RBF was found to provide the best accuracy, as proven by both interpolation and extrapolation test data sets. However, our results also show that more complex multi-kernel designs are more difficult to tune as they become more prone to overfitting. In the future, this issue can be addressed by improving the search policy within BayesOpt for tuning the kernel parameters. Future work also includes the application of multi-kernel CVA to real-world case studies.

### References

- R. Al-Seyab and Y. Cao, 2008, Differential recurrent neural network based predictive control, *Comput. Chem. Eng.*, vol. 32, no. 7, pp. 1533-1545.
- W. E. Larimore, 1990. Canonical variate analysis in identification, filtering, and adaptive control. *Proceedings of the IEEE Conference on Decision and Control* 2, 596–604.
- K.E.S. Pilario, Y. Cao, and M. Shafiee, 2020, A Kernel Design Approach to Improve Kernel Subspace Identification, *IEEE Trans. On Industrial Electronics*, vol. 68, no. 7, pp. 6171-6180.
- S.J. Qin, 2006, An Overview of Subspace Identification, *Computers and Chemical Engineering*, vol. 30 (2006), pp. 1502-1513.
- R. T. Samuel and Y. Cao, 2015, Kernel canonical variate analysis for nonlinear dynamic process monitoring, *IFAC-PapersOnLine*, vol. 28, no. 8, pp. 605-610.
- B. Shahriari, K. Swersky, Z. Wang *et al.*, 2016, Taking the Human Out of the Loop: A Review of Bayesian Optimization, *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148-175.

# Modular Development of Condition Monitoring Systems for the Tennessee Eastman Process

Rexonni B. Lagare,<sup>a</sup> Marcial Gonzalez,<sup>b</sup> Zoltan K. Nagy,<sup>a</sup> Gintaras V. Reklaitis<sup>a</sup>

<sup>a</sup>*Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN 47907, USA*

<sup>b</sup>*School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA*

## Abstract

This paper presents a condition monitoring system development framework that incorporates process knowledge to enhance the performance of machine learning models. Essentially, the framework uses information about the process to create a representation of the process condition, which can be broken down into modules using concepts borrowed from probabilistic graphical modeling. These modules represent simpler problems for fault detection and diagnosis, which allows traditional machine learning (ML) models to perform better without the need for a larger set of training data. Using the Tennessee Eastman Process (TEP) as a case study, the framework was shown to improve detection and diagnosis of all fault types in the TEP fault library under relevant metrics for evaluating condition monitoring systems.

**Keywords:** Tennessee Eastman Process, Machine Learning, Condition Monitoring, Fault Detection and Diagnosis, Condition-based Maintenance

## 1. Introduction

Continuous manufacturing processes often require a condition-based maintenance approach in order to fully realize its benefits. This requires the effective implementation of a condition monitoring (CM) system that can holistically oversee the condition of a process. (Schenkendorf, 2016) It is often challenging to obtain a mechanistic model of the condition of a process, so most of the work in literature focused on data-driven methods such as machine learning (ML). (Yin et al., 2012) The workflow for this development can be visualized in Figure 1, where data from the equipment and sensors of a continuous tableting line are used to train a ML model that can detect and diagnose faults. Given the right dataset, this workflow can be very effective in creating high performing fault detection and diagnosis algorithms. The problem is real scenarios rarely have the right dataset available. Hence, it is often the case that traditional ML models developed in this manner underperform, and the course of action is to acquire a better dataset, which might be prohibitively expensive, and/or to use a better ML model.

Another course of action that one could take is to build on the existing knowledge about the process to develop a mechanistic model. However, level of process knowledge that is required to do this could be even more expensive than acquiring a better dataset for training a machine learning model. Hence, any available process knowledge is often neglected because it is not enough for mechanistic modeling, and modeling efforts would be directly to a purely data-driven approach. An innovative solution would be to find a way to apply this knowledge in enhancing the performance of the ML model development workflow in Figure 1. Such a framework was developed and found to be effective for a continuous pharmaceutical tableting pilot plant. (Lagare et al., 2022a)

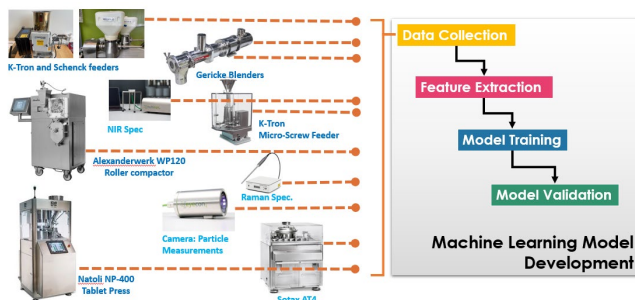


Figure 1. Machine Learning Model Development Workflow Applied to the Continuous Tableting Pilot Plant at Purdue University

It is now interesting to question the limitations of the framework, especially in its applicability towards larger processes with a larger fault library. For this purpose, the Tennessee Eastman Process (TEP) will be considered, (Downs and Vogel, 1993) which is a very different process than a continuous tableting line. First, the material streams are no longer solids, but either liquids or gases. The TEP also has more chemical components and chemical transformations are involved. There are more unit operations in the TEP and to add to its complexity, includes a recycle stream, which the pharmaceutical solids processing system does not have. Finally, the fault library of the TEP is much larger, with 22 conditions that need to be determined from 53 input variables taken from sensors and equipment across the process. (Chiang et al., 2000)

As expected, the condition monitoring system for the TEP, developed using workflow in Figure 1, showed poor performance using metrics relevant to condition monitoring. This study shows that the condition monitoring system development framework, which will be discussed in Section 2.1, was effective in improving the performance of the machine learning models in monitoring the condition of the process.

## 2. Methods

### 2.1. Condition Monitoring System Development Framework

The proposed framework adds several steps to the ML model development workflow in Figure 1, which is now depicted as a node (white) in Figure 2. The two additional steps, i.e., representation and modularization, is mainly responsible for improving the ensuing ML model development step and will be the main subject of this paper.



Figure 2. Proposed Condition Monitoring System Development Workflow

The final step of the proposed framework—integration—is critical in holistically interpreting the predictions of the modules. The result is a more robust condition monitoring system that can still function amid sensor maintenance repairs, reducing the need for product diversions and/or process shutdowns. (Lagare et al., 2022b)

#### 2.1.1. Process Representation

Process representation is the first step in the CM system development framework. This step is responsible for incorporating available process knowledge into the workflow, and its key components are shown in Figure 3.

Based on these components, the minimum process knowledge requirement to be able to perform process representation is a process flow diagram (PFD). The PFD shows the material transformations that taking place, the unit operations responsible for the material transformations, the locations of the measured and manipulated variables, and the locations of the faults in the fault library. Based on the PFD of the TEP (Chiang et al., 2000), the process condition may be represented as in Figure 4.



Figure 3. Process Representation Workflow

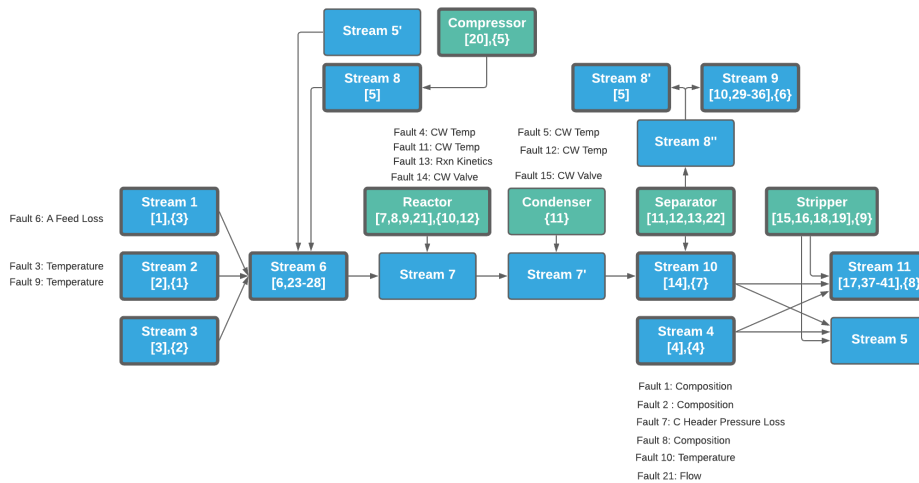


Figure 4. Process Condition Representation of the TEP depicting material and unit operation condition as blue and green nodes respectively. (The numbers in straight brackets are the designated numbers of the measured variables, and the numbers in curly brackets are the designated numbers of the manipulated variables)

### 2.1.2. Process Modularization

Once the process representation is available, it is now possible to modularize the process—i.e., break it down into smaller sections. Although there would be countless ways to do this, there are logical limitations to this combinatorics problem. First, a module must contain at least one fault. Without a fault, there is no condition monitoring job. It is thus interesting to see in Figure 4 that the faults are not dispersed throughout the process but are concentrated on certain nodes. With this first constraint, one can see that the condition representation of the TEP can be broken down into five modules, where the nodes that contain the fault can be considered the central node of the module.

Nodes directly adjacent to the central node may be included in the module, especially if the central node does not have input variables for ML model development. If these secondary nodes, do not have any measured variables, then other nodes that are directly adjacent to the secondary nodes may be added to the module. This process may be repeated until the added node has a measured variable or if there are no more nodes to add. This methodology is represented in Figure 5, and is consistent with the concept of d-separation in probabilistic graphical modeling, where observations in one probabilistic variable/node removes the probabilistic relationship between the parent and the child of



the observed node.(Bishop and Nasrabadi, 2006) Hence, if a measured variable is involved in an added node, adding a node that is conditionally dependent on that measured variable becomes unnecessary since it gives you no further information about the central node that involves the faults for the module.

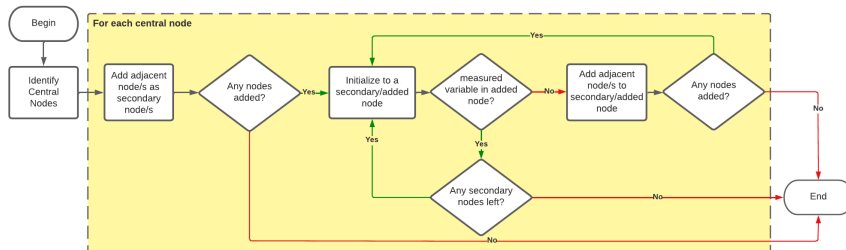


Figure 5. Methodology for modularization of process condition representation

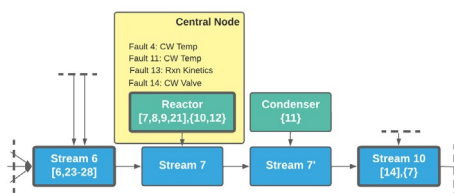


Figure 6. Reactor Condition Module

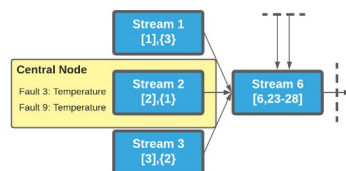


Figure 7. Stream 2 Condition Module

Another noteworthy implication of the workflow in Figure 5 is that only the central nodes are mutually exclusive among the modules, and the measured/manipulated variables in the secondary nodes may be shared. An example of a module based on the reactor condition as the central node is shown in Figure 6 and a module based on the feed stream condition is shown in Figure 7. Notice that both modules have “Stream 6” condition node as a secondary node. The dashed lines indicate the conditional independence of the central node to the rest of the condition nodes in the process condition representation.

### 2.1.3. Machine Learning Model Development Workflow

After the modularization step, ML classifiers are developed using the Model Builder feature by the ML.NET application. Model Builder can take the labelled input variables from each condition module to explore different traditional machine algorithms (i.e., not neural networks) and recommend the best one based appropriate classification metrics.(Microsoft, 2022)

### 2.2. Performance Metrics for Condition Monitoring

In this study, condition monitoring performance will be determined using five metrics: fault detection rate, false alarm rate, accuracy, normal condition certainty index (NCCI), and the overall prediction certainty index (OPCI). Fault detection rate, false alarm rate, and accuracy are standard metrics used for evaluating classifiers(Fawcett, 2006) and fault detection algorithms for the TEP.(Yin et al., 2012)

However, the two certainty indices are new metrics that is a unique contribution of this paper. These indices assume a threshold and work with classifiers (machine learning classification algorithms) that produces a probability for each possible condition in a module. For such classifiers, the condition with the highest assigned probability is considered the prediction condition and its assigned probability can be considered its prediction certainty. If the OPCI for a classifier is 0.95 for a threshold of 0.90, this means

95 out of 100 predictions made had a prediction probability/certainty higher than 0.90. The NCCI works similarly, but only considers the normal condition predictions. These certainty indices reflect the confidence of the predictions by the machine learning algorithm, which the operator can use to evaluate classifier performance. Interestingly, the NCCI has implications for novel fault detection capabilities. If the NCCI is close to 1.0, then future classification predictions that are lower than the threshold could be labelled as novel faults that would require further action by a human operator. Formulas for the OPCI and NCCI are shown below.

$$OPCI = \frac{\sum \text{Predictions} | \text{Prediction Probability} > \text{Threshold}}{\sum \text{Predictions}} \quad \text{Equation 1}$$

$$NCCI = \frac{\sum \text{Normal Condition Predictions} | \text{Prediction Probability} > \text{Threshold}}{\sum \text{Normal Condition Predictions}} \quad \text{Equation 2}$$

### 3. Results and Discussion

To evaluate the effectiveness of the framework, a base case is considered where a machine learning model is developed to use all 53 of the input variables in the TEP for predicting all 22 possible conditions (i.e., normal condition plus 21 fault types). With the framework applied, the ML development task is much simpler after modularization; for the case of the reactor module in Figure 6, the classifier only needs to classify 5 conditions (i.e., normal condition plus 4 fault types) from 16 input variables.

To determine the impact of the d-separation concept in the modularization workflow, a modified case of the framework was considered so that for each module, all the TEP variables would be used as input to determine the faults that are local to the central node of the module. For the reactor module in Figure 6, this would be using all 53 input variables to classify the 5 conditions considered in the module. This results in three cases that can be considered when evaluating the CM performance of each module. The details of these three cases when evaluating the reactor module is summarized in Table 1.

Table 1. Reactor Module Comparison Cases

	Base Case	Modular Case	Modified Modular Case
<b>No. of Faults</b>	21	4	4
<b>No. of Input Variables</b>	53	4	53

Table 2. Performance Summary for Reactor Module Faults

	Base Case	Modular Case	Modified Modular Case
<b>Fault Detection Rate (%)</b>	99-100	88-99	80-100
<b>False Alarm Rate (%)</b>	19-67	0-1	0-2
<b>Accuracy (%)</b>	33-99	89-100	94-99
<b>NCCI</b>	0.57-0.70	0.97-1.00	0.98-1.00
<b>OPCI</b>	0.33-0.89	0.95-1.00	0.92-1.00

For the faults local to the reactor module, the base case performance can be summarized in Table 2. While the fault detection rate is perfect, the other metrics, particularly the high rates of false alarms, make it unusable for condition monitoring. On the contrary, the modular cases show better performance across all metrics, with negligible false alarm rates and high certainty indices which indicate novel fault detection capabilities. For brevity, the performance summaries of the other modules would not be explicitly shown in this paper, albeit the aforementioned trends still hold for other modules. Overall, there seem to be no significant difference between the two modular cases, except for some

faults like fault 5 which saw a lower false alarm rate by 14% compared to the modified modular case.

#### 4. Conclusions

The condition monitoring system development framework that was originally developed for a continuous pharmaceutical tableting line proved to be effective for the Tennessee Eastman Process. The process representation workflow allowed the incorporation of process knowledge into machine learning model development. The ensuing modularization of the process representation simplified the machine learning model development task, which resulted in high performance classifiers that have novel fault detection potential.

For the impact of applying d-separation in the modularization workflow, it was not significant across all faults, although it yielded major improvements for some. Ultimately, this case study validates the applicability of the development framework across two different kinds of continuous manufacturing processes and suggests that it could be effective for other kinds of continuous processes.

#### 5. Acknowledgements

The authors acknowledge Dr. Ziyang Sheriff for sharing valuable insights and knowledge about the Tennessee Eastman Process, which proved valuable in shaping the direction of this research.

This work was supported by the NSF under grant #2140452.

#### References

- Bishop, C.M., Nasrabadi, N.M., 2006. Pattern recognition and machine learning. Springer.
- Chiang, L.H., Russell, E.L., Braatz, R.D., 2000. Fault detection and diagnosis in industrial systems. Springer Science & Business Media.
- Downs, J.J., Vogel, E.F., 1993. A plant-wide industrial process control problem. *Comput Chem Eng* 17, 245–255.
- Fawcett, T., 2006. An introduction to ROC analysis. *Pattern Recognit Lett* 27, 861–874.
- Lagare, R.B., Nagy, Z., Reklaitis, G. v., 2022a. Applying process knowledge for more powerful data-driven condition monitoring systems. Under preparation.
- Lagare, R.B., Sheriff, M.Z., Gonzalez, M., Nagy, Z., Reklaitis, G. v., 2022b. A Comprehensive Framework for the Modular Development of Condition Monitoring Systems for a Continuous Dry Granulation Line. *Computer Aided Chemical Engineering* 49, 1543–1548.
- Microsoft, 2022. What is ML.NET and how does it work? [WWW Document]. URL [https://learn.microsoft.com/en-us/dotnet/machine-learning/how-does-mldotnet-work?WT.mc\\_id=dotnet-35129-website](https://learn.microsoft.com/en-us/dotnet/machine-learning/how-does-mldotnet-work?WT.mc_id=dotnet-35129-website) (accessed 12.1.22).
- Schenkendorf, R., 2016. Supporting the shift towards continuous pharmaceutical manufacturing by condition monitoring. Conference on Control and Fault-Tolerant Systems, SysTol 2016-November, 593–598.
- Yin, S., Ding, S.X., Haghani, A., Hao, H., Zhang, P., 2012. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. *J Process Control* 22, 1567–1581.

# Data-driven predictive model for irrigation management in greenhouses under CO<sub>2</sub> enrichment and high solar radiation

Ikhlas Ghiat, Rajesh Govindan, Tareq Al-Ansari\*

*College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar*

*\*Corresponding author: [talansari@hbku.edu.qa](mailto:talansari@hbku.edu.qa)*

## Abstract

Machine learning models have emerged as a viable method to predict plant water requirements with the ability to form non-linear correlations between plant response variations and microclimate conditions. An artificial neural network is used in this work to predict irrigation water requirements in a greenhouse located in a hyper-arid region with high solar radiation and encompassing an HVAC cooling system and CO<sub>2</sub> enrichment. The prediction model is developed from direct gas exchange measurements of transpiration and takes as input parameters microclimate data including greenhouse solar radiation, temperature, humidity and CO<sub>2</sub> concentration along with hyperspectral imaging-based vegetation indices. Results demonstrate the high performance of the data-driven artificial neural network (ANN) model with microclimate and vegetation index (VI) features. The ANN model resulted in a comparatively higher performance than the FAO56 Penman Monteith empirical model and linear regression with an R<sup>2</sup> of 91.2%, RMSE and MAE of 0.0648 mm/h 0.0528 mm/h respectively. The proposed data-driven predictive model will enable the determination of irrigation water supply in greenhouses under CO<sub>2</sub> enrichment and with varying microclimatic conditions, and support the implementation of precision irrigation.

**Keywords:** Artificial neural network, Data driven model, Vegetation index, Precision irrigation, CO<sub>2</sub> enrichment.

## 1. Introduction

Agriculture in hot and arid regions faces many challenges, primarily due to unfavorable weather conditions, inarable lands, and water scarcity. The continuous development of the agricultural system and the expansion of food production plays an important role in achieving food security (Al-Ansari et al., 2018; Karanisa et al., 2021). Hence, agricultural greenhouse systems have emerged as a viable solution to help secure food by controlling the growing medium microclimate, reducing uncertainties related to the external environment and optimising resource utilisation (Ghiat et al., 2021). Optimum irrigation scheduling is particularly necessary in regions where freshwater resources are limited and require large amounts of energy for desalination and treatment. Agricultural greenhouses in hot and arid climates still face low water use efficiencies due to the inapplicable irrigation schedules applied to these systems, and the various and complex physiological and environmental interactions, which have not been rigorously studied. The current state

of irrigation scheduling in hot and arid regions is based on historical climatic data and expert observations which still poses limitations (Ali, 2010). Intelligent irrigation management such as precision irrigation has emerged as a promising method to improve the operations of agricultural greenhouses and achieve optimum water use efficiencies by ensuring the accurate amount of water inputs at the right time and space (Abioye et al., 2022). Current trends in agriculture are focused on the shift towards data driven models rather than heuristic physical models for system control and management, as they offer better predictions (Mahmood et al., 2021).

The most common methods of irrigation management are based on empirical or mechanistic evapotranspiration models that calculate crop water requirements, with the mostly used one as the Penman Monteith combination model which includes different climatic inputs in the prediction of evapotranspiration mainly solar radiation, relative humidity, temperature, and wind speed (Allen et al., 1998). Crop transpiration is a physiological measure that depends on different microclimatic factors in the greenhouse which interact nonlinearly and dynamically with each other (Stanghellini, 1987).

Machine learning methods have been adopted to estimate evapotranspiration of different vegetation crops and in different growing conditions. Models such as single layer artificial neural networks (ANN) and support vector machines (SVM) have been used because they can provide non-linear correlations between variables and eliminate the parametrisation complexities of internal factors in physical models (Fan et al., 2021). Many studies compared between different machine learning algorithms to predict plant responses. For example, Liu et al. (2009) compared artificial neural networks (ANN) and multiple linear regression (MLR) for the prediction of sap flow of pear trees based on meteorological data and soil water content. Findings from this study indicate that the ANN model had a better performance in predicting sap flow than MLR due to its superiority in predicting non-linear relationships.

This work entails the adoption of artificial neural networks (ANN) for the prediction of transpiration rates in greenhouses under CO<sub>2</sub> enrichment using three different datasets: microclimate data, gas exchange data, and hyperspectral imaging data. The ANN model is compared with the FAO 56 Penman Monteith model and linear regression using statistical performance indicators.

## **2. Greenhouse data**

The study was conducted in a semi-closed, Venlo shaped, and glass-based greenhouse located in the State of Qatar. The microclimate of the greenhouse encompasses CO<sub>2</sub> enrichment (600-800 ppm) and temperature and humidity are controlled by a heating, ventilation, and air conditioning (HVAC) system. The greenhouse is characterised by a hydroponic culture with cucumber as crops. To describe the fast-changing dynamics of plant responses, three different types of data were collected; 1) microclimate data, 2) transpiration rates; 3) hyperspectral-based vegetation indices. Microclimate data including greenhouse temperature, relative humidity and CO<sub>2</sub> concentration were collected using an aspirator box which consists of different sensors, and solar radiation using a pyranometer sensor outside the greenhouse. A gas exchange measurement system (CIRAS-3, PP systems) is used to collect transpiration measurements at the leaf level in

the greenhouse. The gas exchange system is based on infrared gas analysers (IRGA) that trace gases through their absorption in the infrared region. Hyperspectral images were collected using a HSC-2 SENOP camera that takes snapshots in the visible and near-infrared region (400-1000 nm).

### 3. Methodology

A data-based model for transpiration is developed based on microclimate parameters including solar radiation, temperature, relative humidity, and CO<sub>2</sub> concentration along with vegetation indices inside the greenhouse. Several vegetation indices (VIs) belonging to different vegetation categories including broadband greenness, narrowband greenness, light use efficiency, dry or senescent carbon, and canopy water content, were calculated for greenhouse HIS images.

The predictive model is constructed using an artificial neural network (ANN) model which accounts for nonlinearity between parameters. The ANN model is a multi-layer connected network. The input layer in the ANN model receives the input signals and passes them through the hidden layer neurons which in turn transfer them to the output layer to generate the predicted output (Figure 1). Weights are initially assigned between the input layer and the first hidden layer for each input and are passed through the activation function along with the added bias in each neuron. Weights are then updated after acquiring the first predicted output to match the actual value by minimising the loss function through an optimizer.

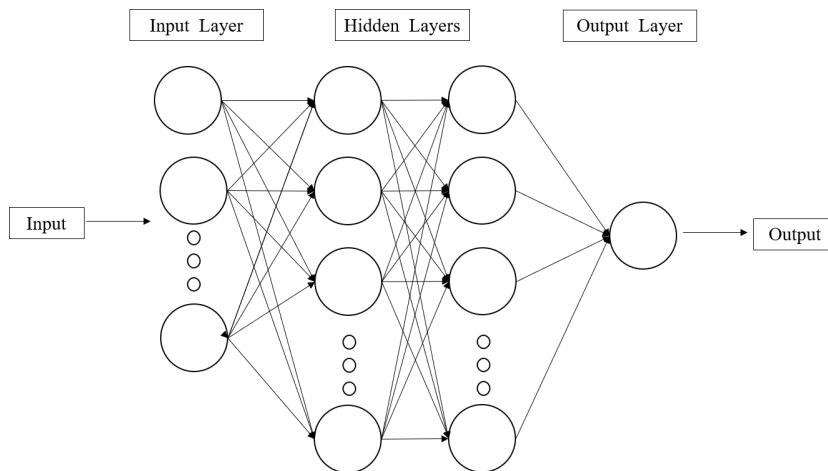


Figure 1: ANN model architecture.

The ANN model is implemented for the prediction of transpiration in the greenhouse. Two activation functions are adopted; the Rectified Linear Unit (Relu) for the input layer (Equation 1) and the hyperbolic tangent (tanh) for the hidden layers (Equation 2). Using different activation functions in the different layers can bring different non-linearities to the problem which can help better solve it. A loss function is used to measure how close the predicted transpiration value is to the actual value. The loss function used is the mean squared error. The ANN model updates the weights through the use of the Adam optimiser.

$$f(x) = \frac{2}{1+e^{-2x}} - 1 \quad [1]$$

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad [2]$$

The optimal number of hidden layers and nodes were experimentally investigated through trial and error. The testing and training sets were chosen as 30% and 70% of the total dataset respectively. Moreover, a hyperparameter optimisation was conducted to select the optimal combination of number of batch size (5, 10, 15, and 20) and Epochs number (5, 10, 50, and 100). Python 3.7 was used in the implementation of the ANN model.

The ANN model is compared against the FAO56 Penman Monteith (Equation 3) empirical model and a linear regression model (Allen et al., 1998).

$$ET_c = k_c \frac{(0.408(R_n - G) + \gamma \frac{900}{T_{mean} + 273} u VPD)}{A + \gamma(1 + 0.34u)} \quad [3]$$

The comparison between the models is assessed using four statistical indicators: the coefficient of determination ( $R^2$ ), the root mean square error (RMSE), and the mean absolute error (MAE) (Equations 4-6). The higher the values of  $R^2$  (closer to 1) and the lower the values of RMSE and MAE indicate good prediction results.

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad [4]$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad [5]$$

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad [6]$$

#### 4. Results and Discussion

In this work, an artificial neural network (ANN) model was implemented for the prediction of transpiration. The number of hidden layers and neurons were decided based on an experimental trial and error method. Two hidden layers with 3 neurons each were decided having only microclimate as input features. However, with the inclusion of vegetation indices as inputs, the dimensionality of the input layer increased, and the optimal number of hidden layers was found as 2 with 6 neurons. The next step entailed the optimisation of the hyperparameters batch size and Epoch number. The hyperparameter optimisation on the first dataset led to a batch size of 5 and an Epochs of 100 as the optimal parameters having the lowest RMSE value.

The ANN model reported good forecasting accuracy as compared to both the FAO56 Penman Monteith (PM) empirical model and the linear regression as shown in Figure 2. The linear regression model reported an  $R^2$  of 0.680, and an RMSE and MAE values of 0.114 mm/h and 0.0893 mm/h respectively (Table 1). The data-driven model, although it is a simple linear regression demonstrated a better performance than the empirical PM model. With the inclusion of vegetation indices, the performance of the linear regression model improved to an  $R^2$  of 0.723, an RMSE of 0.106 mm/h and MAE of 0.082 mm/h.

The ANN model with microclimate data reported values of R<sup>2</sup> of 0.747, RMSE of 0.109 mm/h and MAE of 0.0868 mm/h. In the case where vegetation indices were included as input features, the ANN model was able to fit 91.2% of the data with low tabulated errors of 0.0648 mm/h for RMSE and 0.0528 for MAE (Table 1). The inclusion of vegetation indices increased the ability of the ANN model to predict transpiration variability by 22% (Figure 3). Moreover, the comparative performance of the ANN model including VIs with the FAO56 Penman-Monteith model reveals a higher R<sup>2</sup> by 44%, and lower RMSE and MAE by 68% and 66% respectively.

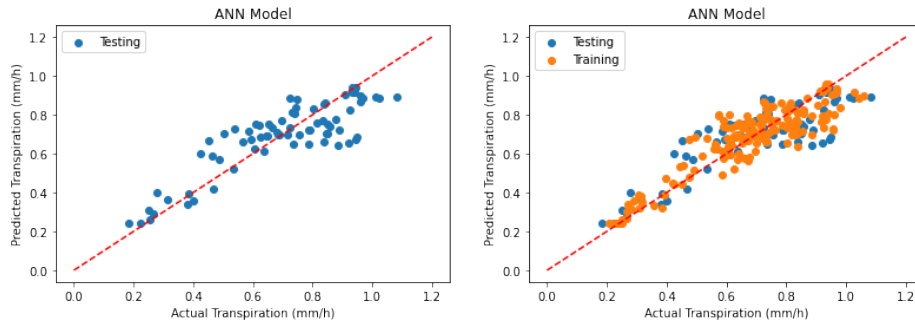


Figure 2: Actual Vs predicted transpiration with ANN model

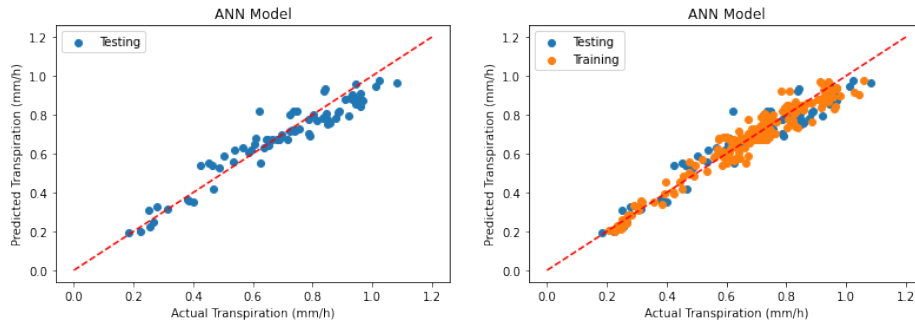


Figure 3: Actual Vs predicted transpiration with ANN model with VIs

Table 1: Model performance on testing set.

Model	RMSE (mm/h)	MAE (mm/h)	R <sup>2</sup>
Linear regression model	0.114	0.0893	0.680
Linear regression with VIs	0.106	0.0820	0.723
ANN model	0.109	0.0868	0.747
ANN model with VIs	0.0648	0.0528	0.912



## Conclusion

Food insecurity is one of the most pressing global challenges. Hot and arid regions are faced to additional environmental limitations for local food production. These regions suffer from climate dependent food production systems that entail large utilisation of natural resources. Irrigation water requirements is one of the most critical inputs in agricultural food systems and needs to be closely assessed, especially for water scarce countries. This work entails the prediction of irrigation water requirements in terms of plant transpiration for a greenhouse with CO<sub>2</sub> enrichment located in a hot and arid climate. A data-driven model is developed from microclimate, physiological, and hyperspectral data using artificial neural networks to predict transpiration in the greenhouse. Results demonstrate the low performance of the FAO56 Penman Monteith empirical model in predicting transpiration for agricultural greenhouses with CO<sub>2</sub> enrichment under high solar radiations. The data-driven artificial neural network (ANN) model with microclimate and vegetation index (VI) features presented a comparatively higher performance than the empirical ET model by increasing the model's ability to explain data variability by 44%, and reducing errors corresponding to RMSE and MAE by 68% and 66% respectively.

## References

- E.A. Abioye, O. Hensel, T.J. Esau, O. Elijah, M.S.Z. Abidin, A.S. Ayobami, O. Yerima, & A. Nasirahmadi, 2022, Precision Irrigation Management Using Machine Learning and Digital Farming Solutions, *AgriEngineering*, 4, 1, 70–103. <https://doi.org/10.3390/agriengineering4010006>
- T. Al-Ansari, R. Govindan, A. Korre, Z. Nie, & N. Shah, 2018, An energy, water and food nexus approach aiming to enhance food production systems through CO<sub>2</sub> fertilization. *Computer Aided Chemical Engineering*, 1487–1492. <https://doi.org/10.1016/B978-0-444-64235-6.50259-X>
- M.H. Ali, 2010, Crop Water Requirement and Irrigation Scheduling. *Fundamentals of Irrigation and On-farm Water Management*, 1, 399–452, Springer New York. [https://doi.org/10.1007/978-1-4419-6335-2\\_9](https://doi.org/10.1007/978-1-4419-6335-2_9)
- R. Allen, L. S. Peirera, D. Raes, & M. Smith, 1998, Crop evapotranspiration - Guidelines for computing crop water requirements, FAO - Food and Agriculture Organization of the United Nations.
- J. Fan, J. Zheng, L. Wu, & F. Zhang, 2021, Estimation of daily maize transpiration using support vector machines, extreme gradient boosting, artificial and deep neural networks models, *Agricultural Water Management*, 245, 106547. <https://doi.org/10.1016/j.agwat.2020.106547>
- I. Ghiat, H. R. Mackey, & T. Al-Ansari, 2021, A Review of Evapotranspiration Measurement Models, Techniques and Methods for Open and Closed Agricultural Field Applications, *Water*, 13, 18, 2523. <https://doi.org/10.3390/w13182523>
- T. Karanisa, A. Amato, R. Richer, S. Abdul Majid, C. Skelhorn, & S. Sayadi, S, 2021, Agricultural Production in Qatar's Hot Arid Climate, *Sustainability*, 13, 7, 4059. <https://doi.org/10.3390/su13074059>
- X. Liu, S. Kang, & F. Li, 2009, Simulation of artificial neural network model for trunk sap flow of *Pyrus pyrifolia* and its comparison with multiple-linear regression, *Agricultural Water Management*, 96, 6, 939–945. <https://doi.org/10.1016/j.agwat.2009.01.003>
- F. Mahmood, R. Govindan, & T. Al-Ansari, 2021, Predicting Microclimate of a Closed Greenhouse Using Support Vector Machine Regression, *Computer Aided Chemical Engineering*, 1229–1234. <https://doi.org/10.1016/B978-0-323-88506-5.50189-3>
- C. Stanghellini, 1987, Transpiration of Greenhouse Crops: an aid to climate management. Wageningen University.

# Greenhouse temperature regulation in the presence of uncertainties using data-driven robust model predictive control

Farhat Mahmood, Rajesh Govindan, Amine Bermak, David Yang, \*Tareq Al-Ansari

*College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar*

*\*Corresponding author: [talansari@hbku.edu.qa](mailto:talansari@hbku.edu.qa)*

## Abstract

Closed environment agriculture is gaining popularity due to the exponentially increasing demand for food, a growing shortage of water, urbanization, reduction in arable land, and degrading soil quality. Greenhouses are closed environment agriculture that provides a viable solution to these problems by maintaining favorable growing conditions. However, maintaining optimum conditions inside a greenhouse is a resource-intensive process, especially in hot and arid climates. Microclimate management, especially the temperature, requires a model-based systematic approach for consistent performance. Therefore, model predictive control is an effective method of managing greenhouse temperature; however, it requires a detailed system model. Moreover, the model predictive control strategy considers the perfect knowledge of the system while not accounting for uncertainties and disturbances. This leads to sub-optimal temperatures inside the greenhouse in the presence of uncertainties. Therefore, this study proposes a robust model predictive control framework to regulate the greenhouse temperature in the presence of external disturbances/uncertainties. An artificial neural network represents the nonlinear and dynamic greenhouse system. The model inputs are solar irradiance, ambient temperature, fan speed, and HVAC control, while the greenhouse temperature is the output. Results illustrate that the robust model predictive control algorithm has superior temperature control compared to a basic model predictive control and the existing greenhouse climate management system with a root mean squared error of 0.25 °C.

**Keywords:** Robust model predictive control, Energy Assessment, Greenhouse control, Robust optimization, Artificial neural network.

## 1 Introduction

Rapidly increasing global population and climate change are adversely affecting the energy, water, and food security of different regions worldwide. Agriculture, which consumes 70% of the total water, is one of the most significantly affected sectors due to irregular weather patterns, high temperatures, droughts, etc. Furthermore, producing crops through open-field agriculture in arid regions is difficult due to the scarcity of freshwater and the harsh external environment. Therefore, efficient and protected crop production environments are required to increase crop yield and minimize resource consumption.

In this context, greenhouses provide a favorable option by protecting the crop from the external environment. Greenhouses regulate the internal climate leading to increased production and reduction of utilized resources by 10-20% and 25-35%, respectively [1].

Greenhouses also provide the benefit of year-round crop production as the microclimate is maintained in the favorable range irrespective of the external environment. The psychrometric conditions of the greenhouse's air significantly affect the cultivated crop's development and yield. Therefore, maintaining optimum conditions, especially the temperature inside the greenhouse, is critical for crop production.

Greenhouses in arid regions require cooling for almost 9-10 months of the year due to high solar irradiance and ambient temperature. In summer, evaporative cooling becomes ineffective; therefore, some greenhouses use active mechanical heat, ventilation, and air conditioning (HVAC) technologies. HVAC technologies perform better than conventional approaches such as evaporative cooling, fogging, etc. however they have a higher capital cost and are more complex to operate.

Different automatic control approaches have been adopted to operate HVAC technologies to improve their performance. One of the most effective control approaches is model predictive control (MPC) which predicts and determines the optimum control. MPC has been successfully implemented in many fields, such as agriculture, aviation, chemical plants, etc.

Mahmood et al. [2] adopted an MPC approach based on a multi-layer perceptron model for a semi-closed greenhouse. Results demonstrated that the MPC performed better than the existing greenhouse climate management system for winter and summer. Similarly, Jung et al. [3] developed an MPC strategy based on an output feedback artificial neural network (ANN) to manage the greenhouse temperature. Results illustrated that the MPC had a root mean squared error (RMSE) value of 2.45 °C while the conventional control strategy had a value of 3.01 °C. However, the performance of MPC drops in the presence of disturbances and uncertainties in the system, leading to the sub-optimal temperature inside the greenhouse. Robust MPC (RMPC) ensures consistent and effective performance in system disturbances. Chen et al. [4] developed a particle swarm optimization-based RMPC strategy for greenhouse temperature control. A nonlinear affine model represents the greenhouse temperature. Results illustrate that the proposed strategy has a superior temperature control performance in the presence of uncertainties. Similarly, González et al. [5] proposed an MPC strategy using a state space model to control the greenhouse temperature. The proposed control strategy reduced the operational cost of the greenhouse; however, it was only applicable to a linear system. Few studies have utilized a data-driven method for greenhouse temperature control; therefore, this study proposes an RMPC framework using an artificial neural network as the system model.

## 2 Methodology

The following methodology is adopted in the present study.

### 2.1 Case study: semi-closed greenhouse

The research was conducted in a semi-closed greenhouse with an active mechanical HVAC system to control the greenhouse microclimate. The greenhouse temperature and relative humidity are measured by a sensor box installed inside the greenhouse. The external climate, such as ambient temperature, solar irradiance, and external relative humidity, is measured by different sensors installed on the outside. The measured data is stored in data loggers and, based on the external and internal climate signals, is sent to the greenhouse actuators. The speed of the ventilation fan and the temperature of the chiller are used to control the microclimate of the greenhouse. The fan speed and chiller temperature are control variables, while the solar irradiance and ambient temperature are the disturbances. The values of the measured variables are given in Table 1.

Table 1: Model parameters measured at the greenhouse site.

Parameter	Value	Unit
Solar irradiance	0 – 1051	Wm <sup>-2</sup>
Ambient temperature	8.1 – 46.5	°C
Fan speed	25 – 100	%
Chiller temperature	17 – 22	°C
Greenhouse temperature	12.1 – 26.8	°C

The greenhouse operates in a closed cycle system where moisture is extracted from the air through a humidifier and then heated to a suitable temperature using an electric heater before returning it to the plantation area. The greenhouse schematic is shown in Figure 1.

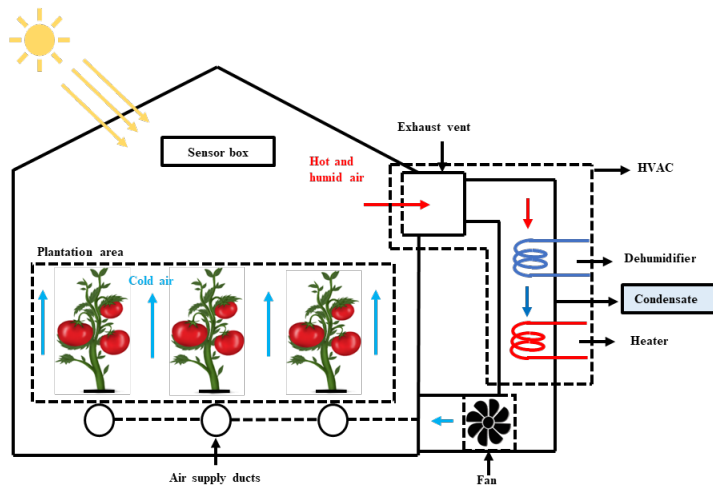


Figure 1: Greenhouse operational schematic. The hot and humid air from the greenhouse is processed through a dehumidification system to remove the moisture. The treated air is then subjected to heating to reach a suitable temperature before being circulated back into the plantation area.

## 2.2 Greenhouse model

An ANN is utilized to predict the dynamic greenhouse temperature. The ANN has solar fan speed, irradiance, ambient temperature, and chiller temperature as the inputs, while the greenhouse temperature is the output. The greenhouse temperature depends on the previous and current values of the control variables ( $u$ ) and system disturbances ( $d$ ). The output is given as follows:

$$\hat{y}(t + 1) = f(u(t), u(t - 1) \dots u(t - n_t), d(t - 1), \dots d(t - n_d)) \quad (1)$$

The structure of the ANN is illustrated in Figure 2.

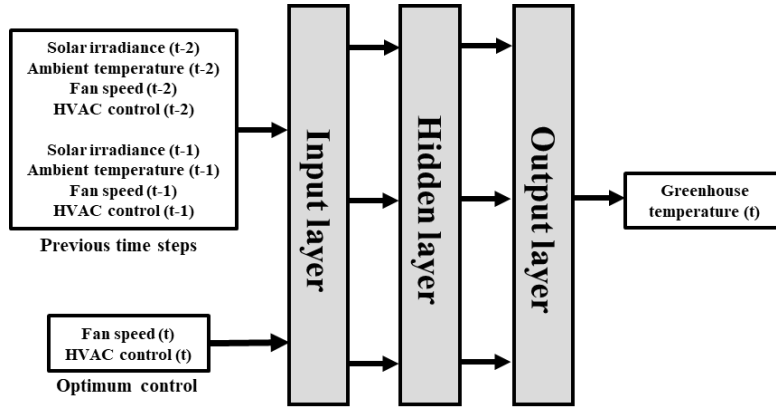


Figure 2: ANN structure

### 2.3 Robust model predictive control

MPC control strategy does not account for the possible uncertainties existing in the system, which may lead to sub-optimal temperatures inside the greenhouse. MPC strategy can be converted into an RMPC by considering the uncertainties in the system. In the present study, a minimax objective function is utilized, which minimizes the worst-case cost and is given by:

$$\underset{u}{\text{minimize}} \quad \underset{k=1,2,\dots,K}{\text{maximize}} \quad \left\{ \sum_{i=1}^P |r(t+i) - \hat{y}(t+1)|^2 \right\} \quad (2)$$

Where  $r$  is the reference trajectory,  $\hat{y}$  is the system output, and  $k=1, 2, \dots, K$  are the different realizations of the system in the presence of uncertainty to which robust optimization is applied to find the control setting.

$$u_{1min} \leq u_1(t+i) \leq u_{1max} \quad p = 0, \dots, N_u = 1, \quad (3)$$

$$u_{2min} \leq u_2(t+i) \leq u_{2max} \quad p = 0, \dots, N_u = 1, \quad (4)$$

The ventilation fan ( $u_1$ ) operates between 25-100% while the chiller temperature ( $u_2$ ) operates between 17-22 °C. The particle swarm optimization (PSO) algorithm determines the optimum control setting at each time step.

## 3 Results and discussions

The proposed control strategy was implemented in Python 3.7. The proposed RMPC strategy is compared with the greenhouse climate management system (GCMS) and a basic MPC strategy.

### 3.1 ANN prediction performance

The accuracy of the system model is critical in the performance of the control strategy. The model is evaluated for 17-18<sup>th</sup> June using 21312 data samples, having ten inputs and the greenhouse temperature as the output. The data set is split into 70% training and 30% testing data sets. The performance of ANN for two days is illustrated in Figure 3. The ANN slightly underestimates the greenhouse temperature during the peak hours while

showing very high accuracy at night. The ANN has an RMSE value of 0.39 °C and  $R^2$  value of 0.98.

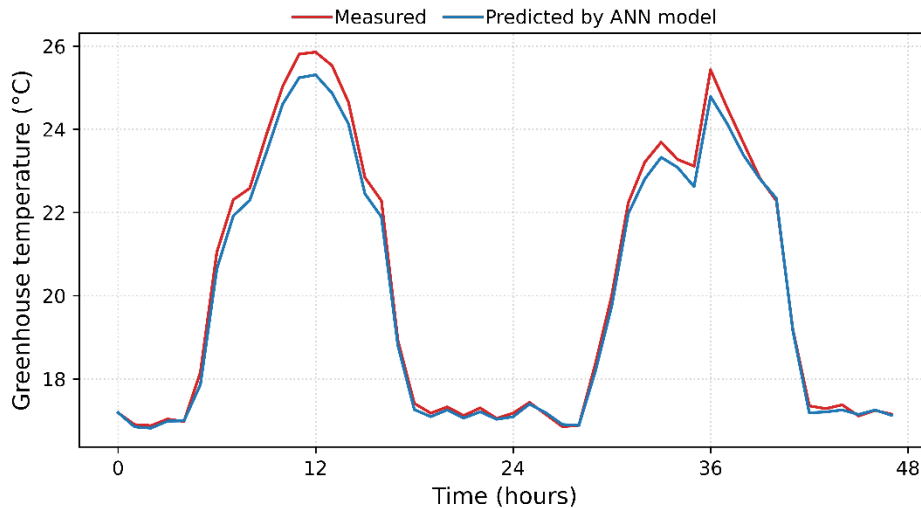


Figure 3: ANN temperature prediction comparison

### 3.2 Robust model predictive control

The greenhouse is in a desert area, and due to higher temperatures, sandstorms are frequent. The solar irradiance reaching the greenhouse fluctuates during the sandstorm due to the dust particles present in the air. The varying solar irradiance affects the greenhouse temperature as the existing climate management does not account for the changing cooling load. The fluctuation leads to sub-optimal temperature during the day, and inefficient use of resources as the amount of cooling required varies. The GCMS also fails to maintain the desired temperature at night as it operates on predetermined control settings. Therefore, an RMPC strategy is proposed to improve the greenhouse's performance in the presence of uncertainty and disturbances. To analyse the performance of the proposed control strategy, additive noise is added to the SI to replicate the sandstorm fluctuations. Figure 4 illustrates the temperature control comparison between the existing greenhouse climate management system, MPC, and RMPC strategy for a two-period (30-31<sup>st</sup> March). RMPC leads to better temperature control performance as compared to GCMS and MPC. RMPC, MPC, and GCSM have RMSE values of 0.25, 0.285, and 2.32 °C.

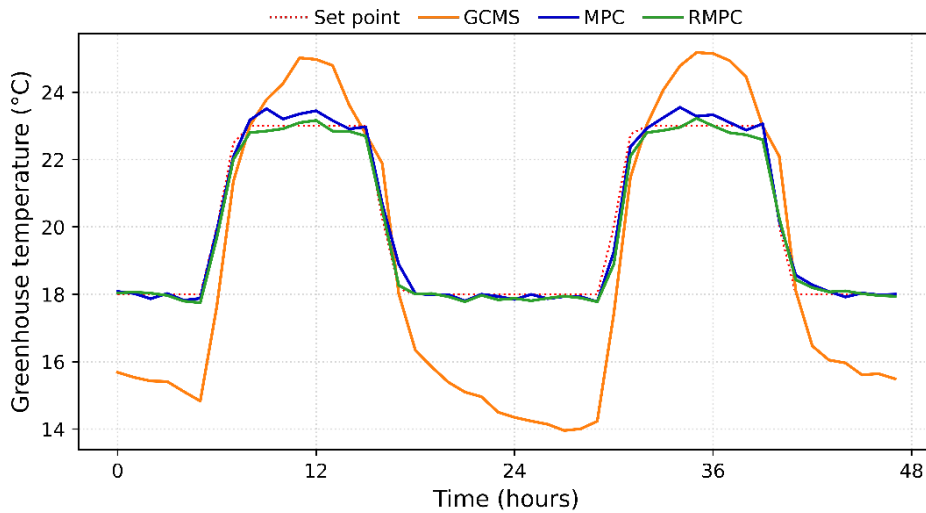


Figure 4: Temperature control performance comparison

#### 4 Conclusion

This study proposes a data-driven robust model predictive control strategy for controlling the greenhouse temperature in the presence of uncertainties. An artificial neural network is used to predict the nonlinear greenhouse temperature. Results illustrate that the minimax-based robust model predictive control has superior performance compared to the greenhouse climate management system and basic model predictive control strategy in the presence of uncertainties due to the external environment. In future studies, the performance of the control strategies in terms of energy consumption will be analysed along with temperature control.

#### Acknowledgment

The authors acknowledge the support provided by Hamad bin Khalifa University, Education, City Doha, Qatar, a member of the Qatar foundation. The research is funded by Qatar National Research Fund (MME01-0922-190049).

#### References

- [1] Mahmood F, Al-Ansari T. Design and analysis of a renewable energy driven greenhouse integrated with a solar still for arid climates. *Energy Convers Manag* 2022;258:115512. doi:10.1016/j.enconman.2022.115512.
- [2] Mahmood F, Govindan R, Bermak A, Yang D, Khadra C, Al-Ansari T. Energy utilization assessment of a semi-closed greenhouse using data-driven model predictive control. *J Clean Prod* 2021;324. doi:10.1016/j.jclepro.2021.129172.
- [3] Jung DH, Kim HJ, Kim JY, Lee TS, Park SH. Model predictive control via output feedback neural network for improved multi-window greenhouse ventilation control. *Sensors (Switzerland)* 2020;20. doi:10.3390/s20061756.
- [4] Chen L, Du S, He Y, Liang M, Xu D. Robust model predictive control for greenhouse temperature based on particle swarm optimization. *Inf Process Agric* 2018;5:329–38. doi:10.1016/j.inpa.2018.04.003.
- [5] González R, Rodríguez F, Guzmán JL, Berenguel M. Robust constrained economic receding horizon control applied to the two time-scale dynamics problem of a greenhouse. *Optim Control Appl Methods* 2014;35:435–53. doi:10.1002/oca.2080.

# A model predictive control framework for the production of functional, safe and sustainable nanomaterials

Argyri Kardamaki<sup>a</sup>, Athanassios Nikolakopoulos<sup>a</sup>, Mihalis Kavousanakis<sup>a</sup>,  
Philip Doganis<sup>a</sup>, Haralambos Sarimveis<sup>a</sup>

<sup>a</sup>*9 Heroon Polytechniou Street, School of Chemical Engineering, National Technical University, Athens 15780, Greece*

[hsarimv@central.ntua.gr](mailto:hsarimv@central.ntua.gr)

## Abstract

In recent years, the production of nanomaterials (NMs) has gained a considerable amount of attention and it has emerged as a flourishing field of research. The increasing interest in nanotechnology and the wide range of innovative industrial applications that NMs offer have highlighted the need to further study NM formation and develop advanced control schemes that can address multiple industrial objectives. This paper presents a new control framework for the production of functional, safe and sustainable silver NMs (Ag NMs) which combines dynamic first-principles modeling, the Model Predictive Control (MPC) architecture and the ability to integrate sensors that provide real-time toxicity data regarding the safety of the produced NMs.

**Keywords:** Nanomaterials, safety, control

## 1. Introduction

The NM case study presented in this paper focuses on the production of Ag NMs through a series of five plug flow (PFR) reactors. The inlet streams of the 1<sup>st</sup> reactor are solutions of the reducing and stabilizing agents, sodium citrate and tannic acid, and the precursor, silver nitrate. All subsequent PFRs are only fed with the precursor solution. The six flows of the above inlet streams are used as the manipulated variables of the process whereas the corresponding concentrations are regarded as measured disturbances that can occur during production. The main objective of this study is to produce Ag NMs of specific diameters that comply with the production and safety specifications. To that end, the mean diameter of the Ag NMs in the product stream of the last PFR is set to be the controlled variable.

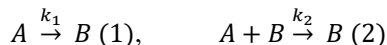
## 2. Methods

### 2.1. First principles modeling

The development of a dynamic process model is essential for the design of the control strategy. A detailed dynamic model of the process was developed based on the discretisation of the PFRs corresponding to the nucleation and growth stages of the Ag NMs production system and the formulation of the respective mass balances. The kinetic



expressions were estimated using the available steady state experimental data. It was assumed that the formation of Ag NMs follows the Finke-Watzky two Step Mechanism (Thanh et al, 2014), which is a process of nucleation and growth where both steps occur simultaneously. The first is a slow continuous nucleation (Eq. 1), and the second is the fast autocatalytic surface growth (Eq. 2) which is not diffusion controlled (Sandoe et al, 2019). The first-principles model describing the system consists of the mass balances of the nuclei ( $A$ ) and the NMs ( $B$ ), while the ordinary differential equation describing the diameter growth depends on the NMs concentration and the autocatalytic surface growth rate.



The above system results in sigmoidal kinetic curves with the FW two-step kinetic model of slow nucleation and fast autocatalytic growth (Watzky and Finke, 2016), with respective, pseudo-elementary nucleation average rate constant,  $k_1$  and surface growth average rate constant,  $k_2$ . Experiments show that with increasing concentration of sodium citrate and/or tannic acid a decrease is observed in both the particle size and UV-vis peak position. Also, the higher concentration of metal atoms within the growing nanoparticle an increase in the particle size was seen (growth favored over nucleation). For the PFR of volume  $V$ , the mass balances for  $A$  and  $B$  and the diameter  $D$  changes are described by the Partial Differential Equations (PDEs) (3), (4) and (5), where  $v_0$  is the volumetric flow rate of the sum of the solutions flowing in the reactor:

$$\frac{\partial C_A(t)}{\partial t} = -v_0 \frac{\partial C_A(t)}{\partial V} + -k_1 C_A(t) - k_2 C_A(t) C_B(t) C_{TA}(t)^{n_{TA}} C_{SC}(t)^{n_{SC}} \quad (3)$$

$$\frac{\partial C_B(t)}{\partial t} = -v_0 \frac{\partial C_B(t)}{\partial V} + k_1 C_A(t) + k_2 C_A(t) C_B(t) C_{TA}(t)^{n_{TA}} C_{SC}(t)^{n_{SC}} \quad (4)$$

$$\frac{\partial D(t)}{\partial t} = -v_0 \frac{\partial D(t)}{\partial V} + k_3 C_A(t) C_B(t) C_{TA}(t)^{n_{TA}} C_{SC}(t)^{n_{SC}} \quad (5)$$

where  $C_A(t)$ ,  $C_B(t)$ ,  $C_{TA}(t)$ ,  $C_{SC}(t)$  are the concentrations of nuclei, Ag NM, tannic acid and sodium citrate respectively. The constants  $k_1$ ,  $k_2$  and  $k_3$  are given by the Arrhenius equation, where the parameters  $k_{o_1}$ ,  $k_{o_2}$  and  $k_{o_3}$  are the pre-exponential factors and  $E_1$ ,  $E_2$  and  $E_3$  are the activation energies. The values of the parameters are estimated using steady state experimental data from the process and the Thompson Sampling Efficient Multi-Objective (TSEMO) optimization procedure (Bradford et al, 2018). Two objectives, namely the error between the predicted NM diameter values and the experimental NM values  $f_1$  and the standard deviation of the wavelength of the peak UV-Vis absorbance  $f_2$ , are placed on the pareto front of the bi-objective optimization problem. From the pareto front, the point that is used for defining the parameter values is  $x_2 = \min_x(f_2)$ , where  $x_1 = \min_x(f_1)$ .

The system of Eqs. (3), (4) and (5) is solved using finite differences and the Method of Lines. Each PFR is discretized along the horizontal axis into  $N$  intervals; elementary Continuous Stirred Tank Reactors, where the solutions are considered homogeneous. Thus, the initial system of PDEs (3 – 5) where  $v_0$  is the volumetric flow rate of the solutions in the reactors is reformulated as a system of Ordinary Differential Equations (ODEs):

$$\frac{dC_{A,i}(t)}{dt} = -v_0 \frac{C_{A,i}(t) - C_{A,i-1}(t)}{\Delta V} - k_1 C_{A,i}(t) - k_2 C_{A,i}(t) C_{B,i}(t) C_{TA,i}(t)^{n_{TA}} C_{SC,i}(t)^{n_{SC}} \quad (6)$$

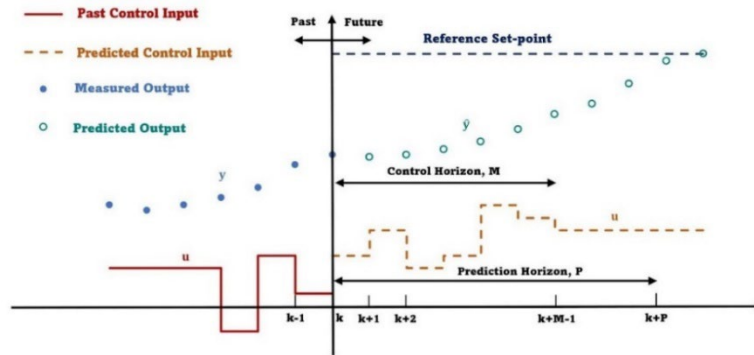
$$\frac{dC_{B,i}(t)}{dt} = -v_0 \frac{C_{B,i}(t) - C_{B,i-1}(t)}{\Delta V} + k_1 C_{A,i}(t) + k_2 C_{A,i}(t) C_{B,i}(t) C_{TA,i}(t)^{n_{TA}} C_{SC,i}(t)^{n_{SC}} \quad (7)$$

$$\frac{dD(t)}{dt} = -v_0 \frac{D_i(t) - D_{i-1}(t)}{\Delta V} + k_3 C_{A,i}(t) C_{B,i}(t) C_{TA,i}(t)^{n_{TA}} C_{SC,i}(t)^{n_{SC}} \quad (8)$$

where  $C_{A,i}$ ,  $C_{B,i}$ ,  $C_{TA,i}$ ,  $C_{SC,i}$  and  $D_i$  are the concentrations and the particle diameter at the  $i^{\text{th}}$  interval.

### 2.2. Model Predictive Control Framework

Model Predictive Control (MPC) is a control strategy that solves an online optimization algorithm at each sampling instant, to compute the optimal control action that drives the predicted output to the reference. MPC stands out among other control strategies because it can handle with ease multi-input multi-output systems and can incorporate explicitly multiple constraints on the state and input variables. The MPC configuration requires a dynamic model of the plant to make predictions about the system's future behavior over a finite time horizon (Camacho et al., 2007). The rolling-horizon concept of MPC is presented graphically in Fig. 1.



**Figure 1:** The rolling horizon concept of MPC

#### 2.2.1. Dynamic Matrix Control

This study employs Dynamic Matrix Control (DMC) as the control algorithm, which is a popular MPC variant. The values of the optimized future manipulated variables are derived by minimizing the quadratic objective function featuring the output prediction errors and the incremental changes of the manipulated inputs. The optimization problem is solved and reformulated at each time instant while the controller only implements the first control action, that has been computed.

In DMC, discrete time step-response models of the plant are integrated in the control algorithm to provide the predictions for the system's future behavior. The discrete-time step response of the system that describes the model output,  $y(t)$ , is given by Eq. (9) where  $g_i$  is the step response coefficient at each time instant  $i$  and  $\Delta u$  is the change of the

manipulated variable  $u$  between two consecutive discrete time instances (Tatjewski, 2007; Camacho et al., 2007; Mulholland, 2016):

$$y(t) = \sum_{i=1}^{\infty} g_i \Delta u(t-i) \quad (9)$$

In systems that reach a steady state within a finite horizon,  $l_{pin}$ , the prediction of the output at each time instant  $k$  can be described by Eq. (10):

$$\hat{y}(t+k|t) = \sum_{i=1}^k g_i \Delta u(t) + \sum_{i=1}^k g_i' \Delta d(t) + \sum_{i=1}^{l_{pin}} (g_{k+i} - g_i) \Delta u(t-i) + \sum_{i=1}^{l_{pin}} (g_{k+i}' - g_i') \Delta d(t-i) + y_m(t) - y_{ss} \quad (10)$$

where  $y_m(t)$  is the current state measurement,  $y_{ss}$  is the steady state,  $g_i'$  is the step response coefficient for the measured disturbance prediction model and  $\Delta d$  is the change of the measured disturbance  $d$  between two consecutive discrete time instances. The measured disturbances are incorporated into the prediction model as additional input variables that cannot be manipulated.

DMC optimizes the future manipulated variables by minimizing the quadratic optimizing function presented in Eq. (11):

$$\min_{\Delta u(k), \Delta u(k+1), \dots, \Delta u(k+N_c)} \left\{ \sum_{k=1}^{N_p} \|\hat{y}(t+k|t) - y_{sp}\|_Q^2 + \sum_{k=0}^{N_c} \|\Delta u(t+k|t)\|_R^2 \right\} \quad (11)$$

where  $\hat{y}(t)$  is the model prediction described by Eq. (10),  $y_{sp}$  is the set-point value,  $N_p$  and  $N_c$  are lengths of the prediction and control horizons and  $Q, R$  are positive definite matrices weighting the output deviations from the set-points and the incremental changes of the manipulated variables, which are defined as follows:

$$\Delta u(t+k|t) = u(t+k) - u(t+k-1) \quad (12)$$

The manipulated variables are constrained within upper and lower bounds with the following inequalities:

$$u_{min} \leq u(t+k) \leq u_{max} \quad (13)$$

The DMC algorithm was developed in MATLAB/Simulink using the external YALMIP library (Löfberg, 2004) for modelling the optimization problem. The GUROBI optimizer (Gurobi Optimization, LLC., 2022) was used to solve the problem. The step response coefficients,  $g_i$  and  $g_i'$ , were generated by introducing a step change of +10% on the initial steady state value of each input variable and by descriptising the dynamic responses of the controlled variables.

### 2.3. Configuration of control simulation platform

The first-principles model described in 2.1 plays the role of the process plant in the integrated MATLAB/Simulink platform. Fig. 2 presents the control loop including the NM production process plant and the DMC controller. The platform can be used both as an open loop simulation to observe the dynamic behavior of the system and as a closed loop simulator, after connecting the DMC controller.

### 3. Results

The tuning parameters of the DMC controller were selected by implementing random sets of disturbances and different setpoint changes on the plant and observing and evaluating the responses of the controlled outputs. Fig. 3 presents an indicative disturbance rejection case study that demonstrates the successful design of the controller, if the system starts from a steady state corresponding to the production of 50 nm Ag NMs. Fig.3 presents the disturbances on the solution concentrations (bottom right), the optimal sequences of manipulated variables (bottom left) and the dynamic response of the mean diameter (top).

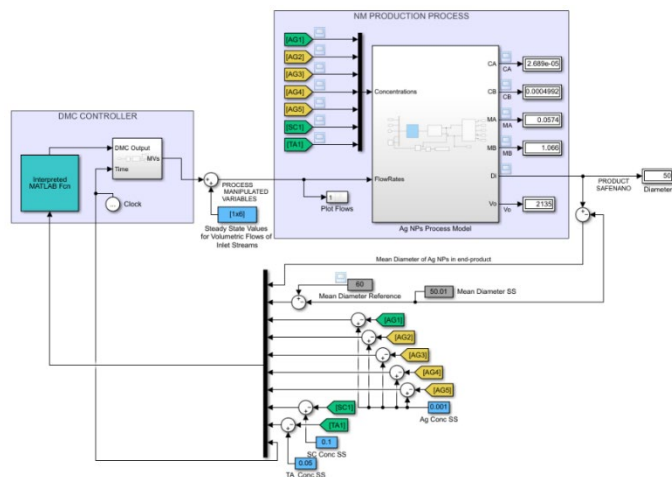


Figure 2: Integrated components of closed loop control platform

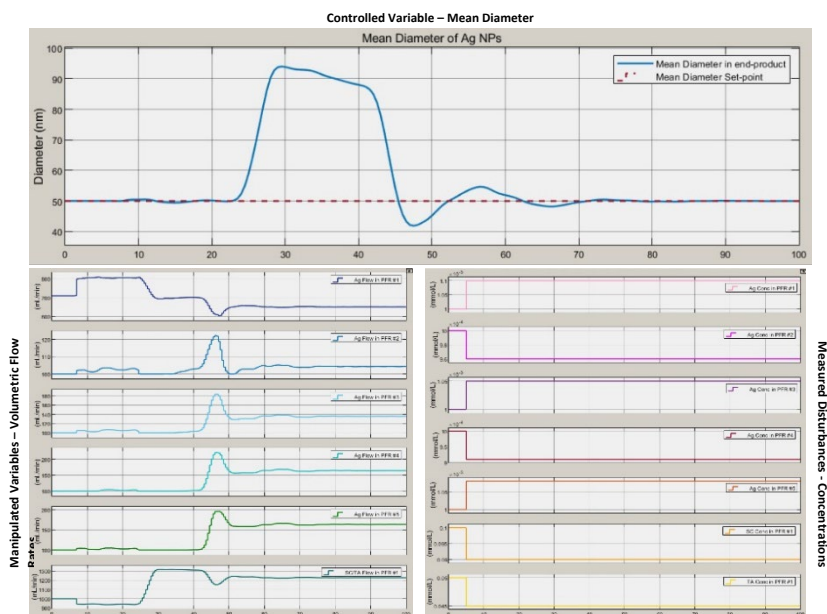


Figure 3: Dynamic response of the DMC control framework at random step changes

After five minutes of simulation time, the system is subject to a series of random step changes in all seven concentrations of the inlet streams (+10%, -4%, +5%, -9%, +8%, -10%, -10% accordingly). These disturbances tend to shift the controlled variable away from its initial state but, as Fig.3 clearly illustrates, the controller succeeds in driving it back to the desired set-point with zero offset by properly manipulating the flow rates of the inlet streams.

#### 4. Discussion

The results demonstrate the effectiveness of the DMC controller in driving the system towards desired set-points while effectively rejecting disturbances that affect closed-loop performance. The proposed approach not only offers the standard benefits of MPC formulations, such as considering constraints on the system variables, but also allows for the incorporation of additional controlled variables relevant to safety specifications for produced NMs. The method is now extended to include feedback from a UV-Vis sensor for real-time information on NM size distribution and a biomembrane sensor for efficient toxicity screening.

**Acknowledgements:** This work has been supported by the SABYDOMA project which has received funding from the European Union's HORIZON 2020 Research & Innovation Programme under grant agreement no. 862296.

#### References

- E. Bradford, A.M. Schweidtmann, A. Lapkin, 2018, Efficient multiobjective optimization employing Gaussian processes, spectral sampling and a genetic algorithm. *Journal of Global Optimization*, 71, 2, p. 407–438.
- E.F. Camacho, C. Bordons, 2007, *Model Predictive Control*. Springer London
- R.J. Field, R.M. Noyes, 1977, Oscillations in chemical systems. 18. Mechanisms of chemical oscillators: Conceptual bases. *Accounts of Chemical Research*, 10,6, p. 214–221.
- Gurobi Optimization, LLC, 2022, *Gurobi Optimizer Reference Manual*.
- J. Löfberg, 2004, YALMIP: A Toolbox for Modeling and Optimization in MATLAB. In *Proceedings of the CACSD Conference*.
- M. Mulholland, 2016, *Applied Control: Essential Methods*. John Wiley & Sons
- R.M. Noyes, R.J. Field, 1977, Oscillations in chemical systems. 19. Mechanisms of chemical oscillators: Experimental examples. *Accounts of Chemical Research*, 10, 8, p. 273–280.
- H.E. Sandoe, M.A. Watzky, M. A., S.A. Diaz, 2019, Experimental probes of silver metal nanoparticle formation kinetics: Comparing indirect versus more direct methods. *International Journal of Chemical Kinetics*, 51,11, p. 861–871.
- P. Tatjewski, 2007, *Advanced control of industrial processes*. Springer London.
- N.T.K. Thanh, N. Maclean, S. Mahiddine, 2014, Mechanisms of nucleation and growth of nanoparticles in solution. *Chemical Reviews*, 114,15, p. 7610–7630.
- M.A. Watzky, R.G. Finke, 1997, Nanocluster size-control and “magic number” investigations. Experimental tests of the “living-metal polymer” concept and of mechanism-based size-control predictions leading to the syntheses of iridium(0) nanoclusters centering about four sequential magic numbers. *Chemistry of Materials*, 9,12, p.3083–3095.

# Multi-period optimisation of oligopolies with contracts: A cooperative approach to customer fairness

Asimina Marousi<sup>a</sup>, Jose M. Pinto<sup>b</sup>, Lazaros G. Papageorgiou<sup>a</sup>, Vassilis M. Charitopoulos<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, Sargent Centre for Process Systems Engineering, University College London, Torrington Place, London WC1E 7JE, UK*

<sup>b</sup>*Linde Digital Americas, 10 Riverview Drive, Danbury CT 06810, United States*  
*v.charitopoulos@ucl.ac.uk*

## Abstract

In an ever-increasing economic instability, cooperation among members of the same market can be proven an efficient way to build resilience. A growing interest has been observed among the process systems engineering community for integrating process design with game theory towards decentralised decision making. In this work, we aim to find the optimal supply chain design in oligopoly case studies by formulating a multi-period cooperative game among manufacturers and customers. Different fairness approaches are examined to maximise the firms' profits and customers' savings by determining the optimal customer allocation and sequencing of different contracts. Since the Nash fairness scheme results in a nonlinear objective, SOS2 approximation is employed resulting in an MILP class. The examined case studies are derived from a duopoly and an oligopoly in the liquid gas market. Computation results suggest that a cooperation can be beneficial for the payoffs of the firms and customers.

**Keywords:** Game theory, Fairness schemes, Supply chain optimisation, Customer allocation

## 1. Introduction

Fairness considerations within supply chain optimisation problems have increasingly attracted the researcher's attention from the PSE community. The interest in this direction stems from the fact that conventional centralised optimisation approaches fail to explicitly consider individual and potentially conflicting objectives of the different parties involved. Game theory is one approach that has been employed by researchers in the field of process design to evaluate both centralised and decentralised decision making. The case of decentralised decision making can be modelled as a cooperative game where all stakeholders' payoffs are considered in the objective function. Different fairness principles can be applied to allocate the payoff to the players of the game, the two main categories are egalitarian and utilitarian schemes. Egalitarian fairness schemes such as Rawlsian Welfare and Schmeidler's nucleolus, aim to equally split the payoff among the players of the game. In contrast, in utilitarian schemes i.e., Nash bargaining and Shapley value, the payoff of a player depends on the contribution of the player to the coalition's payoff.

An early sample of fairness consideration in PSE can be traced in the work of Gjerdrum et al. (2001), where a Nash bargaining approach was employed to solve the problem of fair profit allocation in a multi-echelon supply chain. A common implementation of game

theory can be found in supply chain applications where stakeholders have conflicting objectives (Barbosa-Povoa and Pinto (2020)). To this end, Liu and Papageorgiou (2018) employed a lexicographic min-max approach for the fair profit distribution within a multi-echelon supply chain with firms holding different bargaining powers in the market. Finding the optimum supply chain design of chemical industries under a carbon trading policy framework was examined by Salcedo-Diaz et al. (2021). Different coalition formations were evaluated among the manufacturing firms aiming to maximise their profit while at the same time comply with CO<sub>2</sub> emission caps. Salcedo-Diaz and co-workers suggested that a cooperative framework is beneficial both from an economic and environmental perspective in the examined application.

The impact of the choice of different fairness criteria for the profit allocation of a three-echelon closed loop supply chain was evaluated by Zheng et al. (2019) using the Shapley value, Schmeidler's nucleolus and equal satisfaction mechanism. Charitopoulos et al. (2020) considered a cooperative game to evaluate the fair profit allocation between manufacturing firms in oligopolies using Nash bargaining and a Naïve fairness scheme. The difference between an egalitarian and utilitarian approaches in the formation of an eco-industrial park for fair cost allocation was evaluated by Cruz-Avilés et al. (2021); the examined schemes where Nash bargaining, Social Welfare and Rawlsian Welfare.

## 2. Problem Statement

The present work aims to build on the work of Charitopoulos et al. (2020) by considering a multi-period framework in which customers can be serviced by the different firms and different contracts (Calfa and Grossmann (2015)). In addition, while in Charitopoulos et al. (2020), only the firms were considered as players of the game, here we introduce the customers as additional players aiming to maximise their savings. The proposed methodology can be implemented in different oligopolistic markets of chemical industries, for that reason the model formulation introduced in the following section will be of a generic form.

### 2.1. Model formulation outline

The key elements of the proposed model follow the outline of the model proposed by Charitopoulos et al. (2020) and entail: a) contract modeling, b) material & inventory balances, c) power consumption constraints, d) inter-firm swapping agreements, e) plant production short-cut model, f) service cost. The modeler needs to decide on the sets: i) of the planning horizon of  $P$  time periods, ii) of firms  $F$ , producing  $I$  products at known plant capacities, iii) of existing and new customers  $C$  with deterministic demand and delivery costs, iv) of distinct type of contracts  $K$ , along with the deterministic electricity costs and power consumption limits. The aim of the study is to find the optimal customer allocation and sequencing of contracts, production planning, inventory, swaps, outsourcing and energy consumption levels in order to fairly maximise the firms' profit and customers' savings.

### 2.2. Fair objective function

In the present study two fairness schemes are evaluated: the Naïve and Nash bargaining. Let the parameter  $\pi_f^{sq}$  be the profit of firm  $f$  at the status quo market, i.e., by fixing the customer allocation to a firm and assume no new customer acquisition, and  $\pi_f$  the profit variable of firm  $f$  in a cooperative framework. Both  $\pi_f^{sq}$  and  $\pi_f$  are calculated for the entire duration of the examined time horizon. For the customers' savings consideration, let the parameter  $CC_c^{max}$  be the maximum cost from purchases of a customer  $c$ , and the variable  $CC_c$  be the purchase cost of a customer, both are calculated for the duration of the examined time horizon.

### 2.2.1. Naïve scheme

The objective function for the fair customer allocation with the Naïve fairness scheme is formulated in Eq. (1). It can be observed that in this formulation no consideration of the status quo of the market are considered nor of the maximum cost to be inflicted on the customers.

$$\Omega = \sum_f \pi_f - \sum_c CC_c \quad (1)$$

### 2.2.2. Nash bargaining scheme

The Nash bargaining scheme can be perceived as maximising the geometric mean of the selected payoffs and thus results in a nonlinear objective function (Eq. (4)).

$$\Phi = \prod_f (\pi_f - \pi_f^{sq}) * \prod_c (CC_c^{max} - CC_c) \quad (2)$$

A common reformulation to the Nash fairness objective is separable programming and logarithmic transformation (Liu and Papageorgiou (2018); Charitopoulos et al. (2020); Cruz-Avilés et al. (2021)). Accordingly, it can be observed that Eq. (2) is separable in  $f$  and  $c$ , thus a logarithmic approximation can be applied to the nonlinear terms respectively.

$$\Psi_1 = \ln \Phi_1 = \sum_f \ln(\pi_f - \pi_f^{sq}) \quad (3)$$

$$\Psi_2 = \ln \Phi_2 = \sum_c \ln(CC_c^{max} - CC_c) \quad (4)$$

The reformulated objective function can be derived as the summation of Eq. (3), (4) as follows:

$$\Psi = \Psi_1 + \Psi_2 = \sum_f \ln(\pi_f - \pi_f^{sq}) + \sum_c \ln(CC_c^{max} - CC_c) \quad (5)$$

The given optimisation problem with Eq. (5) as an objective function is of MINLP class, where the nonlinear terms stem from the objective only. In order to transform the problem in a tractable MILP class, a piecewise linear approximation is performed respectively in Eq. (3), (4) using SOS2 variables (Liu and Papageorgiou (2018); Charitopoulos et al. (2020)). It is noteworthy that SOS2 approximations were initially proposed to retrieve global optimum solutions of problems with a single nonlinear function in an otherwise linear programming problem (Beale and Forrest (1976)), which makes it suitable for the examined model. Let  $\tilde{\Psi}_1$  be the linear approximation of  $\Psi_1$ , over a number of grid points  $n$ , parameter  $\tilde{\pi}_{fn}$  the profit of firm  $f$  at grid point  $n$  and  $\lambda_{fn}$  the auxiliary SOS2 variable. Hence, the approximation of Eq. (3) is derived in Eq. (6) with the complementary constraints for the SOS2 variable in Eq. (7), (8).

$$\tilde{\Psi}_1 = \sum_f \sum_n \ln(\tilde{\pi}_{fn} - \pi_f^{sq}) \lambda_{fn} \quad (6)$$

$$\sum_f \lambda_{fn} = 1 \quad \forall f \quad (7)$$

$$\pi_f = \sum_n \tilde{\pi}_{fn} \lambda_{fn} \quad \forall f \quad (8)$$

Analogously the approximation of the customer cost can be reformulated in Eq. (4) with based auxiliary SOS2 variable  $\mu_{cn}$  and the corresponding complementary constraints over the same number of grid points  $n$ . The final linearised objective for the fair customer allocation is formulated in Eq. (9).

$$\tilde{\Psi} = \tilde{\Psi}_1 + \tilde{\Psi}_2 = \sum_f \sum_n \ln(\tilde{\pi}_{fn} - \pi_f^{sq}) \lambda_{fn} + \sum_c \sum_n \ln(CC_c^{max} - \tilde{CC}_{cn}) \mu_{cn} \quad (9)$$

## 3. Case studies

For the purpose of this paper two case studies from an industrial liquid market will be examined. The first case study is a duopoly with 2 firms and 97 customers. At the status quo, Firm A serves 44 customers and Firm B 37, while there are 16 free customers that allow a market share growth. In the second case study, an oligopoly with 3 firms and 81



customers, of whom 13 are new. At the status quo, Firm A holds 21 customers, Firm B 17 and Firm C 30. The selected time horizon is one year discretised into 12 monthly time intervals. Three different contract formulations are evaluated i.e., Open, Formula and Firm, with different duration and pricing as proposed by Marousi et al. (2023). The computational experiments were carried in an Intel®Core™i9-10900K CPU @ 3.70GHZ machine using GAMS-IDE with CPLEX v.38.2 with 12 threads. The optimality gap was set to 1%. The number of grid points for both case studies was selected as  $n=1000$ . The original MINLP problem was addressed with the use of global solver BARON, however no convergence was achieved in reasonable computational time. The aspects that are of interest in this study is the impact of the different fairness schemes on the profit allocation between the firms and the savings of the customers and at the same time how does contract selection facilitate fairness allocation.

### 3.1. Duopoly

Initially it can be observed from Table 1, for the Naïve objective formulation (Eq. (1)), that Firm B has a significant profit increase while Firm A has an only a marginal improvement. This results in the reversal of the market share structure compared to the status quo. In contrast, using the Nash bargaining scheme (Eq. (9)), results in a notable profit increase for both firms, 17% for Firm A and 36% for Firm B. Since Nash scheme follows a utilitarian approach the profit allocation is distributed based on the player's contribution to the total payoff; thus, the resulting market structure closer approximates that of status quo.

Table 1 Profit increase compared to status and market share for different fairness schemes in duopoly market

	<b>%Profit change A</b>	<b>%Profit change B</b>	<b>Market share A</b>	<b>Market share B</b>
<b>SQ</b>	-	-	0.63	0.37
<b>Naïve</b>	+1.8	+80.9	0.49	0.51
<b>Nash</b>	+16.9	+36.4	0.59	0.41

To have a broader picture of the impact of the fairness schemes in the proposed model, it is necessary to evaluate the customer savings. Figure 1 displays the customer savings of the customers with the highest purchasing costs in the status quo, for different schemes. Results in Figure 1 suggest that Nash bargaining scheme leads to increased cost savings ranging from 5 to 8 % for the 5 top customers, in terms of purchasing cost, while the Naïve scheme results in lower customer savings compared to the Nash scheme.

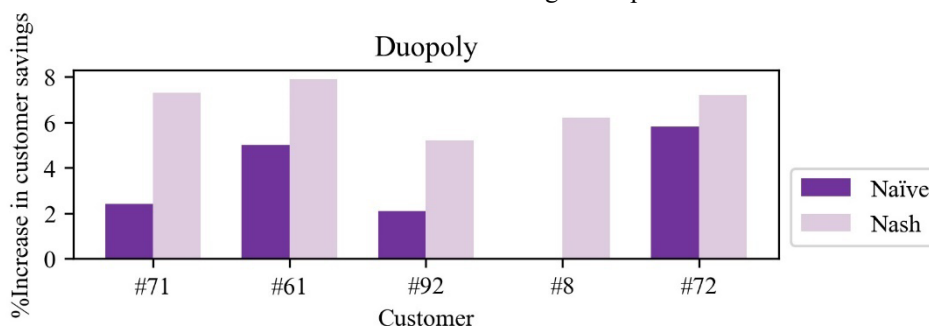


Figure 1 Customer savings compared to status quo for customers with higher cost in duopoly for different fairness schemes

From the Gantt charts of customers #71 and #61 in Figure 2, the savings of the customers can be related to the contract choice. For both customers, status quo allocation results in the use of Formula contract, which inflicts high purchasing cost, in contrast in the Nash fairness scheme the Firm contract prevails in both customers. In the Naïve scheme a combination of all three contracts is used for customer #71 while only Formula and Firm for #61. Consequently, the lower customer savings in the Naïve scheme, compared to the Nash, can be attributed to the fact that Formula contract is selected.

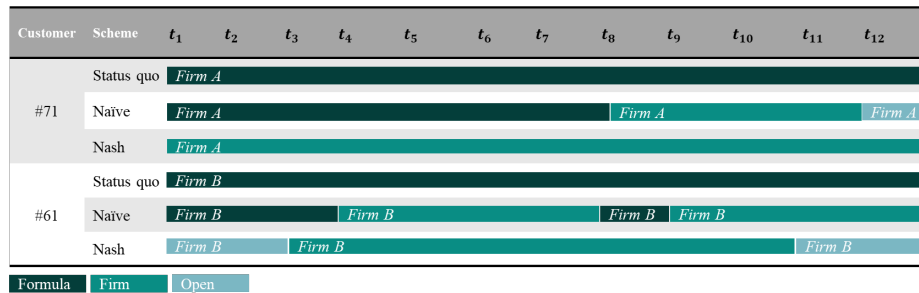


Figure 2 Duopoly Gantt chart for customer allocation in firms and contracts under different fairness schemes

### 3.2. Oligopoly

In the oligopoly case study, an interesting aspect is that there is an imbalance in the market share of the status quo, as suggested by Table 3. Firm C holds more than half of the market share in terms of total profit, while Firm A has a slight advantage over Firm B. The Naïve scheme results in a more balanced market allocation, which is achieved by increasing by 100% the profit of Firm B while maintaining the profit of Firm C at the status quo levels. Although the Nash scheme allows for only a marginal increase for the oligopoly firms, the market share remains stable compared to the status quo.

Table 2 Profit increase compared to status and market share for different fairness schemes in oligopoly market

	%Profit change A	%Profit change B	%Profit change C	Market share A	Market share B	Market share C
<b>SQ</b>	-	-	-	0.25	0.20	0.55
<b>Naïve</b>	+25.0	+104.9	+0.2	0.24	0.32	0.44
<b>Nash</b>	+4.8	+2.8	+6.3	0.25	0.21	0.54

Concerning the customer savings displayed in Figure 3, Nash scheme outperforms the Naïve scheme for up to 10 times for examined customers. On average the Nash scheme results in savings of 22%, while Naïve only to 4%, suggesting that when there is a greater disproportion in the market share of the firms the Nash scheme can result in higher payoffs for the customers.

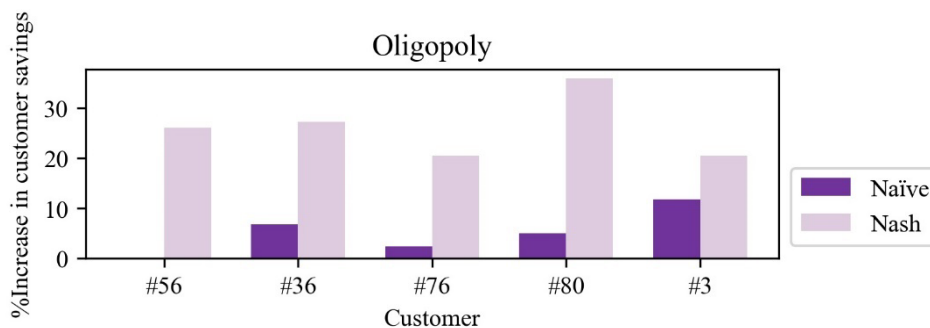


Figure 3 Customer savings compared to status quo for customers with higher cost in oligopoly for different fairness schemes

#### 4. Conclusions

Game theory has been widely used to facilitate supply chain optimisation problems under cooperation. However, the allocation of the coalition's payoff is not always straightforward and different fairness schemes can be evaluated. This paper examines the effect of different fairness schemes on the maximisation of the payoff of manufacturers and customers in cooperative oligopolies, by deciding the optimum contract and customer allocation. For the examined case studies of liquid gas market, Nash fairness resulted in higher customer savings while maintaining the market structure of the status quo.

#### 5. Acknowledgments

The authors gratefully acknowledge financial support from EPSRC grants EP/V051008/1, EP/T022930/1, EP/W003317/1, EP/V050168/1.

#### References

- A. P. Barbosa-Povoa, J. M. Pinto, 2020. Process supply chains: Perspectives from academia and industry. *Comp. Chem. Eng.* 132, 106606.
- E. M. L. Beale, and J. J. H. Forrest, 1976. Global optimization using special ordered sets. *Math. Program.* 10, 52–69.
- B. A. Calfa, I. E. Grossmann, 2015. Optimal procurement contract selection with price ptimization under uncertainty for process networks. *Comp. Chem. Eng.* 82, 330–343.
- V. M. Charitopoulos, V. M. Dua, J. M. Pinto, and L. G. Papageorgiou, 2020. A game-theoretic optimisation approach to fair customer allocation in oligopolies. *Opt. Eng.* 21, 1459–1486.
- D. J. Cruz-Avilés, A. del Carmen Munguía-López, J. M. Ponce-Ortega, 2021. Optimal design of water networks in eco-industrial parks incorporating a fairness approach. *Ind. Eng. Chem. Res.* 275, 8844–8860.
- J. Gjerdrum, N. Shah, L. G. Papageorgiou, 2001 Transfer prices for multienterprise supply chain optimization. *Ind Eng Chem Res* 40:1650–1660
- S. Liu, L. G. Papageorgiou, 2018. Fair profit distribution in multi-echelon supply chains via transfer prices. *Omega* 80, 77–94.
- A. Marousi , J. M. Pinto, L. G. Papageorgiou, V. M. Charitopoulos, 2023. Multi-period fair customer allocation in oligopolies. *FOCAPO/CPC* 2023.
- R. Salcedo-Diaz, J. R. Ruiz-Femenia, A. Amat-Bernabeu, J. A. Caballero, 2021. A cooperative game strategy for designing sustainable supply chains under the emissions trading system. *J. Clean. Prod.* 285, 124845.
- X. X. Zheng, Z. Liu, K. W. Li, J. Huang, J. Chen, 2019. Cooperative game approaches to coordinating a three-echelon closed-loop supply chain with fairness concerns. *Int. J. Prod. Econ.* 212, 92–110.

# Artificial Neural Network-Based Real-Time PID Controller Tuning

Tate Bestwick,<sup>a</sup> Kyle V. Camarda,<sup>a</sup>

<sup>a</sup>*University of Kansas, Department of Chemical and Petroleum Engineering, 1530 West 15<sup>th</sup> Street, Lawrence, Kansas 66045, United States*

## Abstract

Proportional-integral-derivative (PID) control is well established and is still highly prevalent in chemical industries. Many methods for tuning PID controllers to keep a process running at steady-state are also well established; however, they must be retuned if the setpoint must be altered or the system dynamics change at any point in time. The purpose of this work is to use an artificial neural network (ANN) to tune a PID controller on a third order chemical system. It was found that the ANN generated parameters that resulted in smaller overshoot and less settling time than either the Ziegler-Nichols tuning method or the Tyreus-Luyben tuning method in one hundred percent of randomly generated systems.

**Keywords:** Machine Learning, Process Control, Renewable Energy, Neural Network

## 1. Introduction

The focus of this paper is the use of artificial neural networks (ANNs) to predict proportional-integral-derivative (PID) controller tuning parameters for a second-order chemical process which has a third-order closed-loop control scheme. The performance of the PID parameters determined by the ANN are compared with the performance of those obtained from the Ziegler-Nichols tuning method as well as those obtained from the Tyreus-Luyben tuning method.

PID control has been utilized and developed for at least 100 years [1]. In the last century, numerous tuning methods for PID controllers have been developed. The most famous of these methods is likely the Ziegler-Nichols method [2]. However, since 1942, numerous other methods have been introduced, including the integral of the time-weighted absolute error (ITAE) method, Fruehauf's and Chien's internal model control (IMC) method, and the Tyreus-Luyben method [3-5]. All of these methods use system dynamics to obtain heuristical controller parameters which can either be used directly or used as a baseline for finding finer-tuned parameters.

All of the aforementioned tuning methods are primarily for determining values for maintaining steady-state for a single setpoint; in order to transition to a different setpoint, the process must be repeated to obtain new values. Other control methods have been examined in more recent years which allow for more complex and wholistic process control or for automation of PID control tuning. Model predictive control (MPC), fuzzy logic PID control, and active disturbance rejection control (ADRC) are all topics of recent research. [6-8]. However, conservative industries still prefer traditional PID control, particularly if the process is of a more hazardous nature [9]. This has led to the development of using artificial intelligence (AI) to tune PID control parameters. AI-tuned

controllers have been demonstrated to be an effective means of process control for a chemical system [10].

Artificial Neural Networks are a branch of machine learning with considerable predictive power. The input nodes to the network send a signal to the first hidden layer of nodes. This signal is propagated through the first layer to a second layer, and so on until reaching the output layer. Each connection between two nodes is weighted by some function. Finally, each node has a threshold of input signals required to elicit an output greater than zero; this threshold is known as the bias. The number of hidden layers, the number of nodes in each hidden layer, and the weights and biases of each connection are tuned to best apply to the problem of interest. [11-13]

## 2. Physical System and Mathematical Model

### 2.1. Physical System Description

The system examined is two stirred tanks in series. The first tank has two inlets; an initial stream at some concentration and flowrate, and a correction stream which is

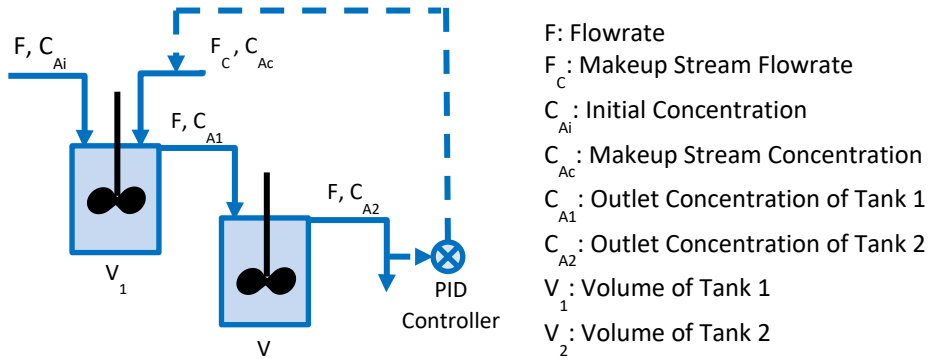


Figure 1: Two stirred-tanks in series. Solid lines represent component streams, dashed lines represent measurements or signals.

highly concentrated and at a much lower flowrate than the previously mentioned initial stream. The two are mixed together in the first tank. The outlet of this first tank is fed into the second stirred tank. The outlet of the second tank is measured, and this signal is sent to a PID controller which is used to adjust the flowrate of the correction stream. The system is depicted in Figure 1.

The following assumptions apply:

- 1) Both tanks are well-mixed at all times.
- 2) The flowrate of the correction stream is orders of magnitude smaller than the flowrate of the initial stream. Thus, the following equation applies:
 
$$F + F_c \approx F \quad 1$$
- 3) The valve on the correction stream has first order dynamics. This causes the overall control scheme to have third order dynamics.

### 2.2. Mathematical Model

The mass balance over each tank is written in equations 2 and 3 with the second assumption applied.

$$\frac{V_1}{F} \frac{d}{dt} C_{A1} + C_{A1} = \tau_1 \frac{d}{dt} C_{A1} + C_{A1} = C_{Ai} + \frac{C_{Ac}}{F} F_c \quad 2$$

$$\frac{V_2}{F} \frac{d}{dt} C_{A2} + C_{A2} = \tau_2 \frac{d}{dt} C_{A2} + C_{A2} = C_{A1} \quad 3$$

Combining these two equations yields a second order equation in which the outlet of tank 2 (measured variable) is obtained from the inlet (disturbance variable) and correction stream (manipulated variable) specifications. Solving for the concentration of the outlet of tank 2 in terms of deviation variables yields equation 4, which is written in the Laplace domain. Deviation variables are denoted as prime variables. Equation 4b relates provides a simplified notation for the inlet variable dynamics and the disturbance variable dynamics.

$$C'_{A2}(s) = \frac{1}{(\tau_1 s + 1)(\tau_2 s + 1)} C'_{Ai}(s) + \frac{C_{Ac}/F}{(\tau_1 s + 1)(\tau_2 s + 1)} F'_c(s) \quad 4a$$

$$C'_{A2}(s) = G_d(s)C'_{Ai}(s) + G_m(s)F'_c(s) \quad 4b$$

The valve on the correction stream was assumed to have dynamics which are modeled in the Laplace domain in equation 5.

$$\frac{0.01}{0.1s + 1} = G_v(s) \quad 5$$

The sensor which measures the concentration of the outlet of the second tank and converting the setpoint from engineering units to electrical signal was assumed to have negligible dynamics (equation 6). Finally, equation 7 models the PID controller.

$$\frac{1}{1} = G_s(s) = G_{sp}(s) \quad 6$$

$$\frac{K_D s^2 + K_p s + K_I}{s} = G_c(s) \quad 7$$

Equations 4-7 allow us to obtain transfer functions which describe the outlet variable response to a change in the setpoint (equation 8), the outlet variable response to the change in the disturbance variable (equation 9), the manipulated variable response to a change in the setpoint (equation 10), and the manipulated variable response to a change in the disturbance variable (equation 11).

$$\frac{G_m(s)G_v(s)G_c(s)G_{sp}(s)}{1 + G_m(s)G_v(s)G_c(s)G_s(s)} = H_{os} \quad 8$$

$$\frac{G_d}{1 + G_m(s)G_v(s)G_c(s)G_s(s)} = H_{od} \quad 9$$

$$\frac{G_v(s)G_c(s)G_{sp}(s)}{1 + G_v(s)G_c(s)G_s(s)G_m(s)} = H_{ms} \quad 10$$

$$\frac{-G_v(s)G_c(s)G_s(s)G_d(s)}{1 + G_v(s)G_c(s)G_s(s)G_m(s)} = H_{md} \quad 11$$

### 3. Methodology

The system described above was modeled in Python using the Python Controls Module [14]. Once modeled, 100,000 variations of the system above were generated. The space-time for tank 1 and the space-time for tank 2 were varied from 1.5 to 11.5 minutes at 0.2-minute increments, and the ratio of the concentration of the correction stream over the inlet flowrate ( $C_{Ac}/F$ ) varied from 20 to 1400 kg/min at an 35 kg/min increments. The Tyreus-Luyben method was used to generate baseline PID tuning parameters which could aptly respond to a setpoint step change for each system. These parameters were then further adjusted to ensure the overshoot for a setpoint step change did not exceed ten percent and the settling time for the system did not exceed the sum of the space-time for tank 1 and the space-time for tank 2. The dataset containing the further adjusted tuning

parameters was randomly split such that eighty percent of it served as the training dataset and twenty percent of it served as the testing data for the ANN which was used to predict the PID controller parameters.

The ANN was generated and trained using Keras's Python interface [15]. The neural network was structured to have three inputs, which can be easily calculated from variables which are directly measurable;  $\tau_1$ ,  $\tau_2$ , and  $\frac{C_{Ac}}{F}$ . The ANN predicts three outputs;  $K_p$ ,  $\tau_I$ , and  $\tau_D$ .

$$\frac{K_p}{\tau_I} = K_i \quad 12$$

$$K_p \tau_D = K_D \quad 13$$

The architecture and the hyperparameters related to training the ANN were varied to find a network which was both accurate and reproducible despite the randomness associated with the training method.

Once the ANN was trained, one hundred variations of the system described in section 2 were randomly generated with the constraint that  $\tau_1$ ,  $\tau_2$ , and  $\frac{C_{Ac}}{F}$  were always within the range set for the generation of the training dataset. The trained ANN was used to predict  $K_p$ ,  $\tau_I$ , and  $\tau_D$  for each system. The response to a step change in the setpoint was measured for each system using the PID controller parameters obtained by the ANN. The response from the system using PID controller parameters generated by the ANN were compared to the responses from the system using PID controller parameters obtained by the Ziegler-Nichols tuning method, and also compared with those obtained by the Tyreus-Luyben tuning method.

#### 4. Results

The ANN generated was able to be trained within a few minutes to predict the tuning parameters of the testing set with an overall  $R^2$  score of 0.919. See Table 1 (below) for ANN performance predicting each controller parameter. The PID parameters predicted by the ANN always resulted in smaller percent overshoot and a shorter settling time than parameters obtained by the Ziegler-Nichols method or those obtained by the Tyreus-Luyben method. The percent overshoot from the system using the Ziegler-Nichols

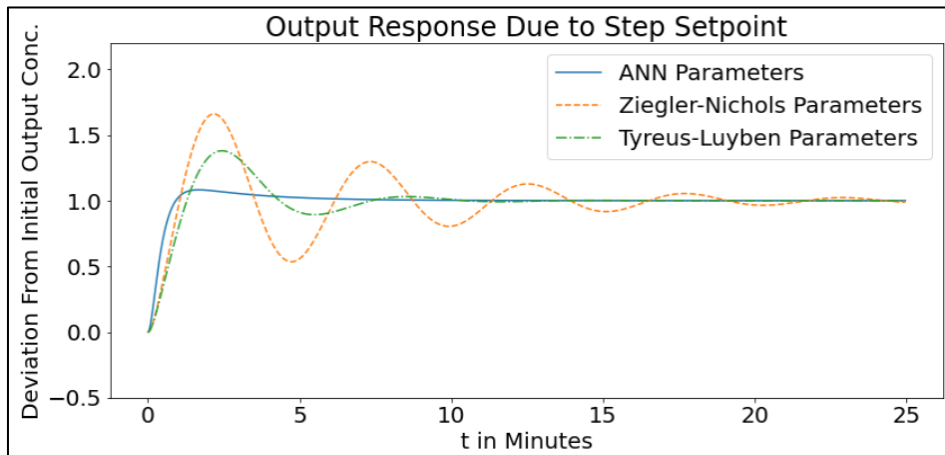


Figure 2: Measured variable response to a setpoint step-change using PID parameters predicted by the ANN, the Ziegler-Nichols method, and the Tyreus-Luyben method.

parameters was always at least 5 times larger than the percent overshoot of the system using the ANN predicted parameters; the percent overshoot from the system using the Tyreus-Luyben parameters was always at least 2.8 times larger than the percent overshoot of the system using the ANN predicted parameters; the settling time of the systems using the Ziegler-Nichols method was always at least 3.8 times larger than the settling time of the system using parameters predicted by the ANN; the settling time of the systems using the Tyreus-Luyben method was always at least 1.6 times larger than the settling time of the system using parameters predicted by the ANN. Figure 2 depicts a comparison of the response of the system using each of the three methods to obtain PID parameters. The trends which are seen in Figure 2 were consistent in all one hundred trials.

Parameter	$K_p$	$\tau_i$	$\tau_D$	Overall
R <sup>2</sup> Score	1.00	0.876	0.881	0.919

Table 1: R<sup>2</sup> score for all ANN performance on all parameters in testing set, as well as overall R<sup>2</sup> score.

## 5. Discussion

The ANN was trained using a dataset of PID controller parameters which outperformed parameters obtained from the Ziegler-Nichols method as well as those obtained from the Tyreus-Luyben method. The results verify that the ANN preserves this quality for PID controller parameters which it predicts for any system which is of the same structure as the one used to generate the training data so long as it has system parameters (space-time, inlet flowrates, and inlet concentrations) which fall within the boundaries specified by the training set. It is quite feasible that this implication extends to any system which is modelled and used to generate training data; the procedure implemented for the system detailed in this paper could be repeated for other systems, including for more complex systems. Furthermore, the procedure for generating training and testing data could be altered to meet standards for specific processes; for example, if a system requires never exceeding 1% overshoot but can tolerate a longer rise time, the ANN could be trained using parameters which meet this criteria.

It is important to note that the ANN takes inputs which capture all of the system variables which could vary across time as they appear in the transfer functions. For example, all three ANN inputs change if the disturbance variable flowrate changes. If the disturbance variable fluctuates over time, the ANN is able to predict PID controller parameters which account for the change in system dynamics. Additionally, the ANN architecture was simplified by using ratios of system variables rather than all of the system variables directly. This was possible because the system variables always appear in the transfer functions in those ratios. Specifically, the volume of tank 1 never appears in the transfer function except for when it appears in the space-time  $\tau_1$ .

Plants which run on renewable energy [16] are excellent candidates for the application of using an ANN to obtain real-time PID controller parameters. Solar and wind energy are intermittent energy sources; the amount of energy is dependent on how much sunlight or wind there is any given day. Thus, plants which run, or which will run, entirely on these energy sources do not have a consistent energy supply. While batteries may be used to store energy, the production and decommissioning of batteries is currently damaging to the environment. Having ANNs which could update tuning parameters for process setpoints would allow for adjusting process output real-time in response to energy availability. The number of batteries necessary would be considerably lowered because the process would be far more responsive to the variability of the energy sources.



The most significant implication of this work is with regard to dynamic operation or simulation. There are many scenarios in which it is preferable to account for a dynamic setpoint or time-varying inputs. Whether the inputs are inconsistent or product demand varies with time, the ability to obtain real-time PID controller parameters which meet the process performance standards could allow for a more efficient process than a static setpoint would allow. Since many processes currently use PID controllers, this implementation would not require drastic changes to the process control scheme. The computation time of ANNs is small enough that PID controller parameters could be updated every few seconds.

## 6. Conclusion

In this work, one hundred thousand variations of a third order chemical system were modelled and the parameters of the PID controller in each one were tuned to obtain previously determined acceptable output responses. These values were then used to train an ANN to predict PID controller parameters for any variation of the modelled system, and was then tested on one hundred randomly generated variations of the system. The ANN was found to predict PID controller parameters which resulted in better system performance than either the Ziegler-Nichols method or the Tyreus-Luyben method in all one hundred cases. This implies that ANNs could be used to tune PID controllers in more complex systems. Furthermore, it implies that ANNs could be used to provide real-time tuning for PID controllers in processes which could have time-varying system dynamics or setpoints.

## References

- [1] Minorsky, N. 1922. Directional Stability of Automatically Steered Bodies, *Nav. Eng. J.* 34 (2): 280-309
- [2] Ziegler, J. B. Nichols, N. B., 1942, Optimum settings for automatic controllers, *ASME Transactions*, 64 , pp. 759-768.
- [3] Maiti, D., Acharya, A., Chakraborty, M., Konar, A., Janarthanan, R., 2008, Tuning Pid and PI<sup>λ</sup>D<sup>μ</sup> Controllers Using the Integral Time Absolute Error Criterion. In Proceedings of the 2008 4th International Conference on Information and Automation for Sustainability, Colombo, Sri Lanka, 12–14 December 2008; pp. 457–462.
- [4] Fruehauf, P. S, Chien, I. L., 1994, Simplified IMC-PID Tuning Rules, 33 (1), 43-59.
- [5] Tyreus, B. D., Luyben, W. L., 1992, Tuning PI Controllers for Integrator/ Deadtime Processes. *Ind. Eng. Chem. Res.*, 31, 2625-2628.
- [6] Oyama, H., Durand, H., 2020, Interactions between control and process design under economic model predictive control, 92, 1-18.
- [7] Turman, E. M., Strasser, W., 2022, Leveraging fuzzy Logic PID controllers for accelerating chemical reactor CFD. *Chem. Eng. Sci.*, 262.
- [8] Martinez, B. V, Sanchis, J., Garcia-Nieto, S., Martinez, M., 2021, Active disturbance rejection control; a guide for design and application. *Rev Iberoam Autom In.*, 18 (3), 201-217.
- [9] Lee, Y. S., Jang, D. W., 2021, Optimization of Neural Network-Based Self-Tuning PID Controllers for Second Order Mechanical Systems. *Appl Sci-Basel*, 11 (17).
- [10] Sousa, B. S., Silva, F. V., Fileti, A. M. F., 2020, Level Control of Coupled Tank System Based on Neural Network Techniques. *Chem Prod Process Model.*, 15 (3).
- [11] Hinton, G. E., 1989, Connectionist Learning Procedures, *Artif Intell*, 40 (1-3), 185-234.
- [12] [https://scikitlearn.org/stable/modules/generated/sklearn.neural\\_network.MLPRegressor.html?highlight=neural+network](https://scikitlearn.org/stable/modules/generated/sklearn.neural_network.MLPRegressor.html?highlight=neural+network) (Accessed 2022)
- [13] Rajapakse, D., Meckstroth, J., Jantz, D. T., Camarda, K. V., Yao, Z., Leonard, K. C., 2022, Deconvoluting Kinetic Rate Constants of Catalytic Substrates from Scanning Electrochemical Approach Curves with Artificial Neural Networks, *ACS meas. sci. au*.
- [14] <https://python-control.readthedocs.io/en/0.9.2/> (Accessed 2022)
- [15] <https://keras.io/> (Accessed 2022)
- [16] Ojo, G.; Camarda, KV; Sustainable Ammonia Production via Electrolysis and Haber-Bosch Process. *ESCAPE Proc. of 2022*, vol 32, pp. 229-234

# State of the Art Flexibility Analysis for Natural Gas Monetization Production Processes Under Uncertainties

Noor Yusuf, Ahmed AlNouss and Tareq Al-Ansari

*College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Education City P.O. Box 5825, Doha, Qatar*

## Abstract

The increased demand for cleaner energy resources places a significant emphasis on monetizing natural gas into value-added products such as liquefied natural gas (LNG), ammonia and power to target different market needs. Despite the estimated growth in the demand for the different cleaner energy resources, each monetization process is subject to market uncertainties induced by global events such as pandemics, energy crises and supply chain disruptions. In this context, this study considers a promising strategic approach to investigate the economic profitability of embedding production flexibility in three natural gas monetized products (i.e., LNG, ammonia and power) subject to forecasted demand and price, and technical flexibility of each process (upper and lower production limits). The results provide decision-makers with a systematic approach to assess the economic and technical aspects of a flexible integrated LNG, ammonia, and power plant design in today's turbulent markets.

**Keywords:** Flexibility, natural gas monetization, uncertainty, LNG, ammonia.

## 1. Introduction

Natural gas (NG) is a crucial input to different processes and sectors such as household, transportation, petrochemical, agricultural fertilizers, and plastics industries. Within local markets, NG is mainly used to generate electricity using the gas-to-power (GTP) process. Furthermore, different chemical and physical monetization routes have been deployed to monetize NG into economically viable products. The monetization routes are classified based on the required final product: service-based monetization route (i.e., GTP), physical monetization route (i.e., liquefied natural gas - LNG and compressed natural gas - CNG), and the chemical monetization route (i.e., chemical conversion of natural gas into liquid products such as ammonia and methanol). The selection of a monetization route depends on factors such as international market requirements, the geographical location of gas reserves, climate in the region, gas field characteristics and gas quality.

Although renewables are expected to continue being the fastest-growing class of energy resources in the global market share, NG is expected to be the fastest-growing fossil-based energy resource due to its resilient characteristics (BP, 2014). On the other hand, energy production management has been essential to guarantee global energy sustainability, environmental sustainability, effective resources management, and economic profitability of projects. In terms of economic sustainability, producing companies must strategically decide on the final product portfolio, targeted markets, pricing, and shipping routes to optimize the project's profitability and hedge against risks. Following the various events that disrupted the global energy trade and supply chains,

such as the financial crisis of 2008, the oil markets crush of 2015, the COVID-19 pandemic in 2020, and the most recent Russian-Ukrainian crisis, the concept of production flexibility has become increasingly vital to increase the responsiveness of projects to exogenous uncertainties (Heffron *et al.*, 2021; Sarkar *et al.*, 2021; Sarkar & Seo, 2021). In a flexible production system, the decision-maker has the right, but not the obligation, to change the production capacity based on market performance to sustain profitability. In the literature, most studies focused on optimizing the production and planning of a single-product NG project subject to production flexibility based on varying demand (Bhosekar & Ierapetritou, 2020; Verleysen *et al.*, 2021; Yusuf *et al.*, 2022). There is a lack of studies that investigate the optimal production and management of integrated NG production systems. In this study, global historical and forecasted demand and price data for three NG monetization routes: LNG, power, and ammonia, are used as input to investigate the yearly optimal NG capacity allocation to the three production processes for maximum overall project profitability. Moreover, constraints in production flexibility (upper and lower limits) are considered. Fundamentally, capacity flexibility allows decision-makers to change the production level throughout the project lifetime based on market demand as a proactive response to uncontrolled exogenous uncertainties.

## 2. Methodology and Data

The methodology developed as part of this study involves performing sustainable planning within the downstream energy sector in Qatar. The designed framework serves as a decision-making tool that enables the prediction of flexible decisions for the downstream production of power, LNG, and ammonia to satisfy the local power demand and participate in the international demand for LNG and ammonia. The benefits of deploying a sustainable planning system as an allocation tool to enhance the overall economic profit are also investigated through an economic analysis of capital investment cost, operating cost, and sales revenue. This section describes the model development of the different involved processes, the tool used to execute the simulation, and the historical and forecasted data sets available to run the model.

The commercial software Aspen HYSYS is used for process modelling and simulation. The NG liquefaction process scheme evolved from condensation at constant pressure, using the Joule-Thompson method, to the energy-efficient heat exchange process using refrigeration systems. The latter is widely used in different projects worldwide. Licensors developed different technologies in the market to design cost-effective and energy-efficient processes for large-scale commercial deployment. The technologies differ based on the production capacity, refrigerants, liquefaction scheme, and equipment capacity (Al-Mutaz *et al.*, 2016; Castillo *et al.*, 2010; Danilov *et al.*, 2019). Fundamentally, the AP-X technology, an extension of the C3MR process developed by Air Products, is considered for LNG production in Qatar. While for ammonia production, NG is first reformed into syngas through steam and oxygen reforming reactions and the effluent syngas is assured of reaching the necessary 1:3 nitrogen to hydrogen ratio. After which, the syngas is cleaned from CO<sub>2</sub> and water and introduced in the ammonia synthesis loop, where ammonia is generated from the reaction of hydrogen and nitrogen at high pressures ranging from 150 to 350 bar and temperatures ranging from 430 to 480 °C. After which, the ammonia is chilled to be removed from the loop. Whereas, for power generation, the syngas generated from NG is initially conditioned before being delivered to the combustion reactor, where it burns with air. To create electricity, the effluent stream is sent into a gas turbine. The expanded gas is used to produce various forms of steam

through a series of heat exchangers, which are then utilised to create electricity utilising a steam turbine. The simulated LNG, power and ammonia models using Aspen HYSYS are assessed using a built-in economic evaluation tool. The results of the economic evaluation tool are illustrated in Table 1.

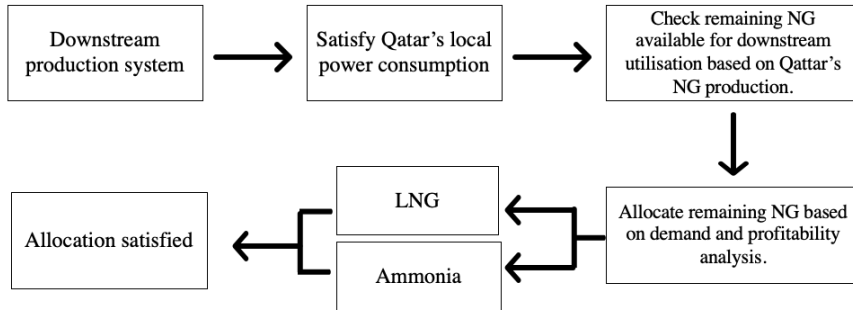


Figure 1: High-level representation of planning methodology.

Table 1. An economic evaluation of LNG, power and Ammonia using Aspen HYSYS.

Process	Power	Ammonia	LNG
<b>NG Flow (kg/h)</b>	5.40E+03	6.41E+05	1.38E+07
<b>Product Amounts</b>	Power (kWh)	Ammonia (kg/h)	LNG (kg/h)
	4.85E+08	7.26E+05	9.98E+06
	Steam (tonne/y)		
	5.21E+05		
<b>Product Prices</b>	Power (\$/kWh)	Ammonia (\$/T)	LNG (\$/T)
	0.05724	400	1785
	Steam (\$/tonne)		
	7.5		
<b>Total Capital Cost (\$)</b>	1.01E+08	1.59E+08	3.16E+08
<b>Total Operating Cost (\$/y)</b>	7.54E+06	1.41E+08	1.15E+08
<b>Total Raw Materials Cost (\$/y)</b>	5.78E+05	1.18E+09	2.54E+10
<b>Equipment Cost (\$)</b>	5.66E+07	6.06E+08	1.13E+08
<b>Total Installed Cost (\$)</b>	6.48E+07	1.03E+09	1.92E+08
<b>Total Annualized Cost (\$/y)</b>	2.89E+07	1.35E+09	2.56E+10
<b>Revenue (\$/y)</b>	3.16E+07	2.54E+09	1.56E+11

The simulation results and historical and forecasted market data are used as data sets for the planning and allocation formulated in Python using the MESA library (Mesa, 2016). The historical and forecasted demand and production data for NG, LNG, power and ammonia are illustrated in Figure 2. Whereas the trends of global prices of ammonia and LNG are illustrated in Figure 3.

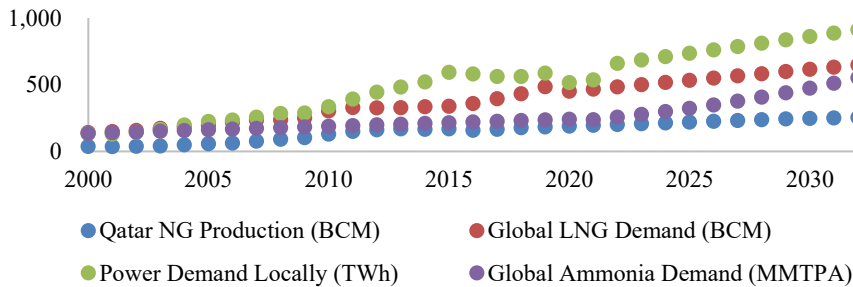


Figure 2: Historical and forecasted demand and production data.

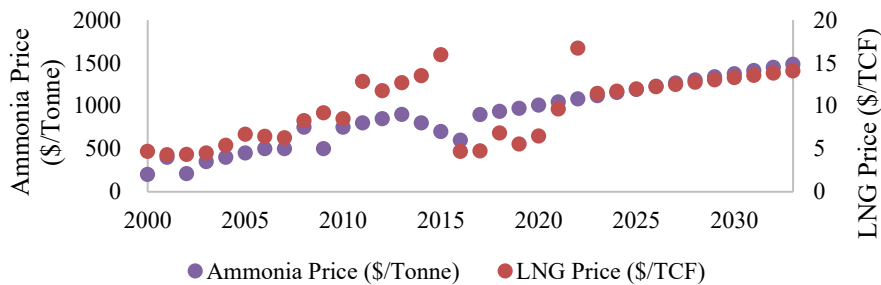


Figure 3: Trends of global prices of Ammonia and LNG.

The ammonia and LNG production must satisfy a minimum operating limit per train capacity and operational flexibility lower limits, as indicated in Table 2.

Table 2. Minimum operating limits (Cheema & Krewer, 2018; Yusuf et al., 2022).

Process	LNG (MTPA)	Ammonia (kg/h)
Lower limit Production capacity	5	708

### 3. Results and Discussion

With the influence of global events and supply chain disruptions on the planning and operation of different NG projects, the integration of various tools in the early design stages of a project is crucial to design a responsive integrated NG production system responsive to market changes. In this study, the results of simulated NG to ammonia, LNG, and power processes, along with historical and forecasted local power requirements in Qatar and global LNG and ammonia demand, are considered to investigate the optimal annual NG production capacity allocation to the three production processes using the python MESA library based on the optimal final product portfolio for maximum overall profitability. The optimal allocation is subject to primarily satisfying the local power demand and then distributing the remaining NG production capacity to the LNG and ammonia production processes to meet international demand. Moreover, lower production capacity limits for LNG and ammonia production processes must be met, wherein a production train cannot operate below the lower flexibility threshold due to technical infeasibility. As illustrated in Figure 4, the optimal NG distribution to LNG, ammonia and power processes from 2000 to 2033 is changing annually, subject to NG

production and demand for each product. The simulation indicates the common trend was the dominance of NG allocation to power and LNG.

In contrast, in the years 2017, 2019, and 2020, a shift to NG allocation for ammonia production is identified due to factors such as the decrease in local power requirements, the decline in global LNG demand in 2020 influenced by the COVID-19 pandemic, and the growth in NG production. Figure 5 illustrates the ammonia and LNG capacities produced in MTPA and Power generation in TWh. The associated annual profits and annualized costs for the optimal NG distribution to LNG, ammonia, and power are illustrated in Figures 6 and 7, respectively, which demonstrate that a significant share of the total profits is associated with international LNG and local power sales. Despite the minimal allocation for NG to ammonia production compared to the other monetization routes, the high selling price of ammonia in international markets reflected positively on the project's profitability. However, ammonia production is associated with high annualized costs due to the high costs of raw materials, equipment, and installation. Hence, annualized costs are an essential factor influencing the economic sustainability of a monetization route under turbulent markets.

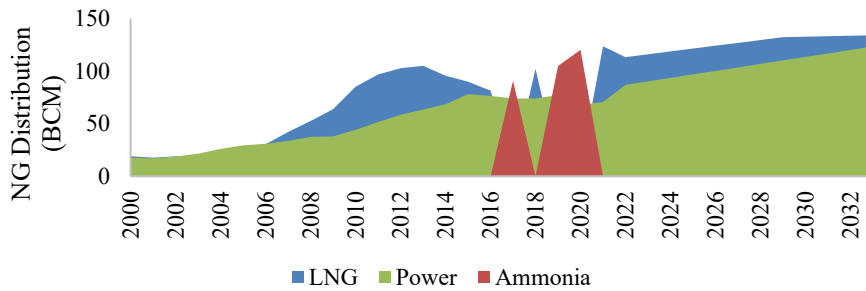


Figure 4: Annual NG distribution (BCM) to LNG, Ammonia, and Power between 2000 and 2033.

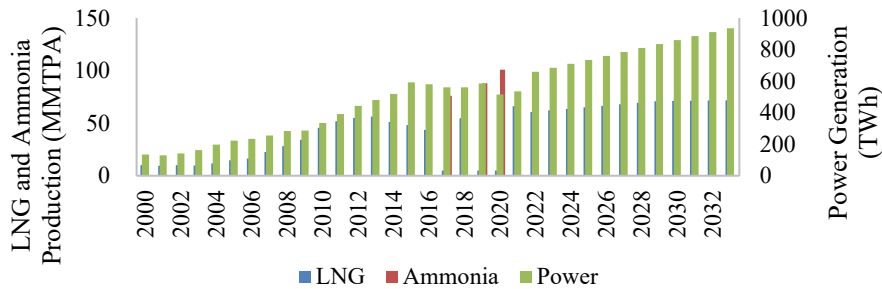


Figure 5: Annual production capacity of LNG, Ammonia, and Power between 2000 and 2033.

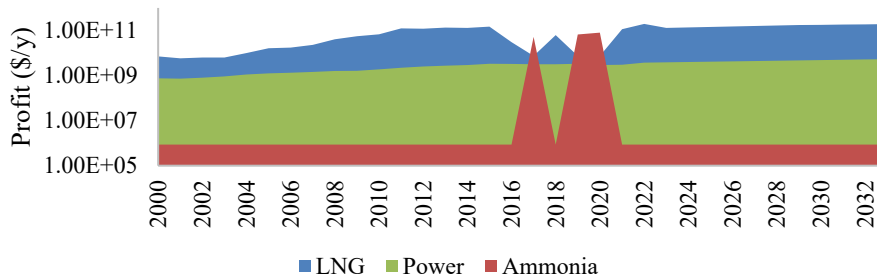


Figure 6: Annual profits for LNG, Ammonia and Power sales based on optimal NG distribution.

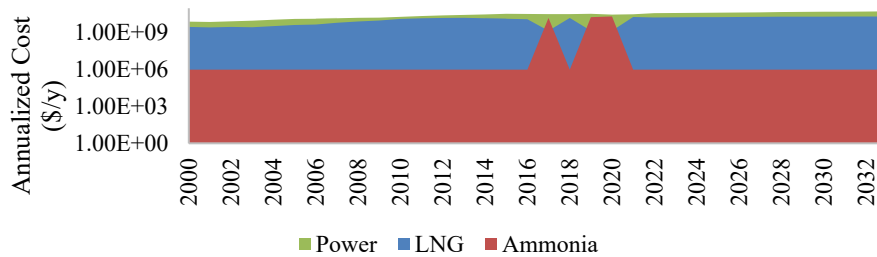


Figure 7: Annualized costs for LNG, Ammonia and Power processes.

#### 4. Conclusion

Integrating different tools for optimal NG allocation to products is crucial in designing integrated NG production systems in turbulent markets. Despite the trend for shifting to ammonia production, this study demonstrates the importance of prioritizing power and LNG production and varied production flexibility in the studied period to achieve maximum profitability and hedge against risks. The technical limitations of production were considered for optimal NG distribution to open the door for future comprehensive studies in assessing the value of flexibility associated with each monetization route.

#### References

- Al-Mutaz, I. S., Liu, X., & Mazza, G. (2016). Natural Gas Liquefaction Technologies-An Overview. *OIL GAS-EUROPEAN MAGAZINE*, 42(4), 213–218.
- Bhosekar, A., & Ierapetritou, M. (2020). Modular design optimization using machine learning-based flexibility analysis. *Journal of Process Control*, 90, 18–34.
- BP. (2014). *BP Energy Outlook 2035*.
- Castillo, L., Nadales, R., González, C., Dorao, C. A., & Vilorio, A. (2010). Technology Selection for Liquefied Natural Gas (Lng) on Baseload Plants. *XIX International Gas Convention AVPG 2010, May 24th - 26th Caracas, Venezuela*, 1–15.
- Cheema, I. I., & Krewer, U. (2018). Operating envelope of Haber–Bosch process design for power-to-ammonia. *RSC Advances*, 8(61), 34926–34936.
- Danilov, R., Arabyan, M., & Usov, D. (2019). Influence of technologies on LNG market

development. *IOP Conference Series: Materials Science and Engineering*, 537(4).  
<https://doi.org/10.1088/1757-899X/537/4/042030>

- Heffron, R., Körner, M., Schöpf, M., Wagner, J., & Weibelzahl, M. (2021). The role of flexibility in the light of the COVID-19 pandemic and beyond: Contributing to a sustainable and resilient energy future in Europe. *Renewable and Sustainable Energy Reviews*, 140, 110743.
- Mesa, P. (2016). *Mesa: agent-based modeling in python 3+*.
- Sarkar, B., Mridha, B., Pareek, S., Sarkar, M., & Thangavelu, L. (2021). A flexible biofuel and bioenergy production system with transportation disruption under a sustainable supply chain network. *Journal of Cleaner Production*, 317, 128079.
- Sarkar, M., & Seo, Y. W. (2021). Renewable energy supply chain management with flexibility and automation in a production system. *Journal of Cleaner Production*, 324, 129149.
- Verleysen, K., Parente, A., & Contino, F. (2021). How sensitive is a dynamic ammonia synthesis process? Global sensitivity analysis of a dynamic Haber-Bosch process (for flexible seasonal energy storage). *Energy*, 232, 121016.
- Yusuf, N., Govindan, R., & Al-Ansari, T. (2022). Techno-economic analysis of flexible AP-X LNG production process under risks and uncertainties. In *Computer Aided Chemical Engineering* (Vol. 51, pp. 367–372). Elsevier.





# Online Monitoring and Robust, Reliable Fault Detection of Chemical Process Systems

Zheyu Jiang<sup>a\*</sup>

<sup>a</sup>Oklahoma State University, 420 Engineering North, Stillwater, Oklahoma, USA 74078

\*Corresponding author: [zheyu.jiang@okstate.edu](mailto:zheyu.jiang@okstate.edu)

## Abstract

Nowadays, large amounts of data are continuously collected by sensors and monitored in chemical plants. Despite having access to large volumes of historical and online sensor data, industrial practitioners still face several challenges in effectively utilizing them to perform process monitoring and fault detection, because: 1) fault scenarios in chemical processes are naturally complex and cannot be exhaustively enumerated or predicted, 2) sensor measurements continuously produce massive arrays of high-dimensional big data streams that are often nonparametric and heterogeneous, and 3) the strict environmental, health, and safety requirements established in the facilities demand uncompromisingly high reliability and accuracy of any process monitoring and fault detection tool. To address these challenges, in this paper, we introduce a robust and reliable chemical process monitoring framework based on statistical process control (SPC) that can monitor nonparametric and heterogeneous high-dimensional data streams and detect process anomalies as early as possible while maintaining a pre-specified in-control average run length. Through an illustrative case study of the classical Tennessee Eastman Process, we demonstrate the effectiveness of this novel chemical process monitoring framework.

**Keywords:** Process monitoring, fault detection, statistical process control, CUSUM, big data streams

## 1. Introduction

Digitalization is transforming chemical and process industries. Modern chemical plants are equipped with sophisticated digital tools and infrastructures, including numerous sensors and advanced distributed control systems (DCSs), which continuously monitor the plants' equipment performance, manufacturing processes, and mass, energy, and information flows. Together, these sensors generate massive arrays of online data streams that are often *nonparametric* (i.e., data streams do not necessarily follow any specific distribution) and *heterogeneous* (i.e., data streams do not necessarily follow the same distribution). Over the past decades, a number of algorithmic approaches have been developed to effectively utilize the large volumes of historical and online sensor data for reliable online process monitoring and early fault detection. Among them, dimensionality reduction techniques, such as principal component analysis (Jackson and Mudholkar, 1979; Fezai et al., 2018) and partial least squares regression (Geladi and Kowalski, 1986), are the most popular ones in the literature (Russell et al., 2000). These dimensionality reduction-based approaches assume that the statistics characterizing the in-control profiles must also span the subspace defining the out-of-control states or faults (Woodall et al., 2004). In other words, to use the features (e.g., principal components) obtained from historical in-control process data (known as Phase I) for online monitoring (known as Phase II), one must ensure there is no profile shift during online monitoring in some otherwise undetectable direction. However, this assumption is not guaranteed as the

chemical process dynamics typically are quite complex and out-of-control states cannot be fully enumerated or anticipated a priori. Another shortcoming of dimensionality reduction-based methods is that operators and process engineers often have a hard time interpreting the results obtained because the features are in the reduced space, which does not have a one-to-one mapping to the original big data streams. Furthermore, since the number of possible distillation fault scenarios can be quite large, monitoring only the most significant subset of features can cause significant error, as the fault may not be noticeable in the selected features.

More recently, advancements in machine learning, such as support vector machine (Onel et al., 2018) and artificial neural network (Heo and Lee, 2018), offer new pathways toward chemical process monitoring and fault detection. Nevertheless, state-of-the-art machine learning-based approaches still face problems such as overfitting and poor predictive accuracy. For example, while most of the published machine learning methods perform well on training and validation sets, their fault detection accuracies rarely exceed 95% in test sets. Considering the strict EHS requirements on plant site and the severity of consequences in case of fault detection failure, such predictive accuracies are unacceptable and can be catastrophic. Furthermore, machine learning methods do not scale well for new fault scenarios that have not been encountered before. In summary, existing distillation process monitoring and fault detection frameworks are inadequate and unsuitable to address practical, sophisticated data stream characteristics and fault scenarios encountered in chemical process industries.

In this work, we introduce a generic chemical process monitoring and fault detection framework featuring nonparametric and heterogeneous big data streams. This framework can detect process mean shifts or anomalies as early as possible while maintaining a user-specified false alarm rate (or in-control average run length, IC-ARL). Specifically, we adopt and simplify the quantile-based statistical process control (SPC) framework that generalizes the proven and reliable multivariate cumulative sum (CUSUM) control charts for nonparametric, heterogeneous big data systems (Ye and Liu, 2022). To demonstrate its effectiveness in chemical process monitoring, we apply this new framework to the classic problem of Tennessee Eastman Process (Downs and Vogel, 1993) and compare the fault detection performance with PCA and SVM-based approaches for the first time.

## 2. Recent Advancements in Multivariate SPC

We use  $X(t) = (X_1(t), \dots, X_p(t))$  to denote the measurement of  $p$  data streams over the observation time  $t = 1, 2, \dots$ . We assume that the local statistic  $X_j(t)$  is i.i.d. across time  $t$  for every  $j$ . Note that the i.i.d. assumption of the data streams is often satisfied when  $X_j(t)$  measures the residual value (Zou et al., 2015). Also, we emphasize that the independency across different data streams is not required. To combine these individual local statistics into a single global monitoring statistic for fault detection, Tartakovsky et al. (2006) and Mei (2010) proposed using the maximum and the sum of local statistics to form the global monitoring statistic, respectively. Mei (2011) further proposed another global monitoring scheme known as the “top- $r$  approach” based on the sum of the largest  $r$  local statistics. Unfortunately, all these methods were developed under the assumption that all data streams follow a normal distribution, which is rarely encountered in chemical and process industries. To address this issue, various nonparametric multivariate CUSUM procedures (e.g., Liu et al., 2015) have been developed following the pioneering work of Qiu and Hawkins (2001, 2003). Although these multivariate CUSUM methods relax the normality assumption, they still assume that all data streams follow the same distribution,

thereby limiting its applicability to monitoring homogeneous data streams. Earlier this year, Ye and Liu (2022) proposed a quantile-based nonparametric SPC algorithm to construct the local statistic for each data stream. The idea is to incorporate the ordering information of in-control data measurements by categorizing them into a number of quantiles (e.g., 10 or 15) for each data stream in Phase I (Qiu and Li, 2011). Next, in Phase II, online measurements are categorized into the learnt quantiles to generate a quantile-based distribution, which is compared with the quantile-based distribution of in-control data for the detection of any mean shift. Thus, this new quantile-based SPC approach can effectively monitor heterogeneous data streams, and it does not require any prior knowledge about the fault scenarios. In the next section, we present a simplified formulation of the original quantile-based SPC framework of Ye and Liu (2022).

### 3. Formulation and Methodology

Due to space limitations, we only highlight the key results in this quantile-based SPC framework. First, in Phase I, for each data stream  $j = 1, \dots, p$ , in-control measurements are ordered and partitioned into  $d$  quantiles:  $I_{j,1} = (-\infty, q_{j,1}]$ ,  $I_{j,2} = (q_{j,1}, q_{j,2}]$ ,  $\dots$ ,  $I_{j,d} = (q_{j,d-1}, +\infty)$ , such that each quantile contains exactly  $\frac{1}{d}$  of the in-control measurements. Therefore, one can define cumulative intervals as  $CI_{j,i}^+ = [q_{j,i}, +\infty)$  and  $CI_{j,i}^- = (-\infty, q_{j,i}]$  for  $i = 1, \dots, d - 1$ . This information is then used in Phase II monitoring. Specifically, in Phase II, a vector  $\mathbf{Y}_j(t) = (Y_{j,1}(t), \dots, Y_{j,d}(t))$  is defined for each data stream  $j$ , where  $Y_{j,q}(t) = \mathbb{I}\{X_j(t) \in I_{j,q}\}$  and  $q = 1, \dots, d$ . Here,  $\mathbb{I}\{X_j(t) \in I_{j,q}\}$  is the indicator function that equals 1 when  $X_j(t) \in I_{j,q}$  and 0 otherwise. Correspondingly, one can further define two vectors,  $\mathbf{A}_j^+(t) = (A_{j,1}^+(t), \dots, A_{j,d-1}^+(t))$  and  $\mathbf{A}_j^-(t) = (A_{j,1}^-(t), \dots, A_{j,d-1}^-(t))$ , for each data stream  $j$ , such that  $A_{j,i}^+(t) = \mathbb{I}\{X_j(t) \in CI_{j,i}^+\}$  and  $A_{j,i}^-(t) = \mathbb{I}\{X_j(t) \in CI_{j,i}^-\}$ . One can show that  $A_{j,i}^+(t) = 1 - \sum_{k=1}^i Y_{j,k}(t)$  and  $A_{j,i}^-(t) = \sum_{k=1}^i Y_{j,k}(t)$ . And  $\mathbb{E}(A_{j,i}^+(t)) = 1 - i/d$  and  $\mathbb{E}(A_{j,i}^-(t)) = i/d$  for  $j = 1, \dots, p$  and  $i = 1, \dots, d$ . Therefore, detecting the mean shifts in the distribution of  $X_j(t)$  is equivalent to detecting shifts in the distribution of  $A_{j,i}^+(t)$  and  $A_{j,i}^-(t)$  with respect to their expected values. Specifically,  $A_{j,i}^+(t)$  (resp.  $A_{j,i}^-(t)$ ) is more sensitive to upward (resp. downward) mean shift (Ye and Liu, 2022). Thus, to detect upward (+) and downward (-) mean shifts, the multivariate CUMSUM procedure developed of Qiu and Hawkins (2001, 2003) was adopted by defining  $C_j^\pm(t)$  as:

$$C_j^\pm(t) = \left[ \left( \mathbf{S}_j^{\pm, \text{obs}}(t-1) + \mathbf{A}_j^\pm(t) \right) - \left( \mathbf{S}_j^{\pm, \text{exp}}(t-1) + \mathbb{E}(\mathbf{A}_j^\pm) \right) \right]^T \cdot \text{diag} \left( \mathbf{S}_j^{\pm, \text{exp}}(t-1) + \mathbb{E}(\mathbf{A}_j^\pm) \right)^{-1} \cdot \left[ \left( \mathbf{S}_j^{\pm, \text{obs}}(t-1) + \mathbf{A}_j^\pm(t) \right) - \left( \mathbf{S}_j^{\pm, \text{exp}}(t-1) + \mathbb{E}(\mathbf{A}_j^\pm) \right) \right], \quad (1)$$

where  $\mathbf{S}_j^{\pm, \text{obs}}(t)$  and  $\mathbf{S}_j^{\pm, \text{exp}}(t)$  are two  $(d-1)$ -dimensional vectors with  $\mathbf{S}_j^{\pm, \text{obs}}(t) = \mathbf{S}_j^{\pm, \text{exp}}(t) = \mathbf{0}$  if  $C_j^\pm(t) \leq k$ , whereas  $\mathbf{S}_j^{\pm, \text{obs}}(t) = \frac{C_j^\pm(t) - k}{C_j^\pm(t)} \left( \mathbf{S}_j^{\pm, \text{obs}}(t-1) + \mathbf{A}_j^\pm(t) \right)$  and  $\mathbf{S}_j^{\pm, \text{exp}}(t) = \frac{C_j^\pm(t) - k}{C_j^\pm(t)} \left( \mathbf{S}_j^{\pm, \text{exp}}(t-1) + \mathbb{E}(\mathbf{A}_j^\pm) \right)$  if  $C_j^\pm(t) > k$ . With this, the local statistic  $W_j^+(t)$  (resp.  $W_j^-(t)$ ) for detecting upward (resp. downward) mean shift is:  $W_j^\pm(t) = \left( \mathbf{S}_j^{\pm, \text{obs}}(t) - \mathbf{S}_j^{\pm, \text{exp}}(t) \right)^T \cdot \text{diag} \left( \mathbf{S}_j^{\pm, \text{exp}}(t) \right)^{-1} \cdot \left( \mathbf{S}_j^{\pm, \text{obs}}(t) - \mathbf{S}_j^{\pm, \text{exp}}(t) \right)$ , which is shown

to be equivalent to  $W_j^\pm(t) = \max\{0, C_j^\pm(t) - k\}$  (Qiu and Hawkins, 2001). Here,  $k$  is a pre-computed allowance parameter that restarts the CUSUM procedure by resetting the local statistic back to 0 if there is no evidence of upward or downward mean shift after a while (Xian et al., 2021). To detect both upward and downward mean shifts in a data stream  $j$ , we simply define a two-sided local statistic  $W_j(t) = \max\{W_j^+(t), W_j^-(t)\}$  (Li, 2020). And the initial condition is  $W_j(0) = W_j^+(0) = W_j^-(0) = 0$  for all  $j = 1, \dots, p$ . Finally, we ranklist  $W_j(t)$  for the current time  $t$  based on its magnitude:  $W_{(1)}(t) \geq W_{(2)}(t) \geq \dots \geq W_{(p)}(t)$ , where  $W_{(j)}(t)$  denotes the  $j^{\text{th}}$  largest estimated local statistic. With this, we generalize the top- $r$  approach (Mei, 2011) to determine the stopping time  $T$  for raising an alarm and declaring the system is out of control for monitoring heterogeneous data streams:  $T = \inf\{t > 0: \sum_{(j)=1}^r W_{(j)}(t) \geq h\}$ , where  $r$  is typically much less than  $q$  (Mei, 2011), and  $h$  is a constant threshold value related to false alarm rate (Liu and Shi, 2013). A commonly used  $h$  corresponds to the false alarm (Type-I error) rate of no more than 0.0027 (classic  $3\sigma$  limit). Overall, this quantile-based SPC framework developed by Ye and Liu (2022) offers strong statistical justifications and great flexibility as process engineers can customize the choice of  $h$  based on the severity of potential failures.

#### 4. Case Study: Tennessee Eastman Process

The Tennessee Eastman Process (TEP, see Figure 1) is an extensively used benchmark case for comparative assessment of process monitoring algorithms. It consists of a reactor, a product condenser, a separator, and a stripping column. The TEP takes four feed streams (streams 1-4) and partially converts them into desired products G and H and byproduct F. As illustrated in Figure 1, the model contains 11 manipulated and 41 measured variables, as well as 28 predefined fault scenarios to choose from. To generate in-control and out-of-control process data, we utilize the MATLAB/Simulink-based GUI developed by Anderson et al. (2022). In total, 50 hours of in-control process data were generated and collected in Phase I for quantile learning and threshold value  $h$  determination.

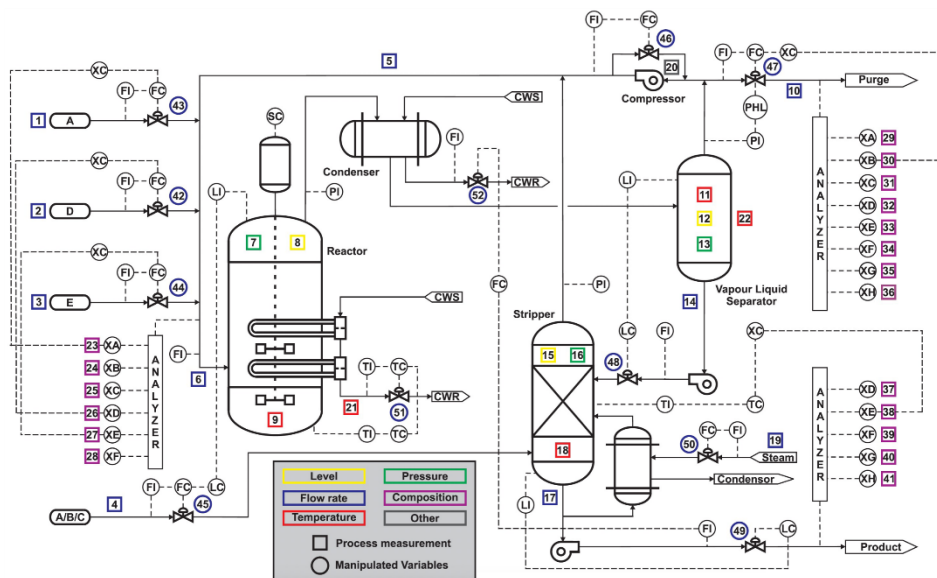


Figure 1. Schematic of Tennessee Eastman Process (source: Ma et al., 2020)

As an illustrative example, we perform process monitoring and fault detection of three representative faults, namely IDV 2, 3, and 13, as summarized in Table 1. In Phase II, we first run the process at normal operations and collect 2 hours of in-control data, followed by starting the fault and operate the process and collect 3 hours of out-of-control measurements. We compare our SPC approach with two other process monitoring algorithms commonly used in chemical industries, namely PCA and SVM. Specifically, we utilize the open source Pyphi package developed by García-Muñoz at Eli Lilly that performs multivariate PCA and Hotelling's  $T^2$  analyses (López-Negrete et al., 2010). We select and keep all principal components whose eigenvalues are greater than 1, thereby leading to a total of 15 principal components for all four fault scenarios. For SVM, we adopt the method and comparison results of Onel et al. (2018).

Table 1. Summary of fault scenarios considered in this comparison study.

Fault No.	Description	Fault Type
IDV 2	Stream 4 composition of B (with constant A/C ratio)	Step
IDV 3	Temperature in stream 2	Step
IDV 13	Reaction kinetics	Slow drift

The comparison results of fault detection speed and the corresponding false alarm rate of all three monitoring frameworks, quantified by how many additional observations (after fault is introduced) are needed for each algorithm before it realizes the process's out-of-control status and raises an alarm, are tabulated in Table 2. As we can see, among the three monitoring frameworks, quantile-based SPC framework yields the fastest fault detection speed in all three fault scenarios, while maintaining the lowest false alarm rate. Furthermore, quantile-based SPC framework only stores and uses quantile information ( $I_{j,i}$ ) for online monitoring and fault detection and is thus very computationally efficient. This result is exciting, given that a lower false alarm rate will lead to reduced fault detection speed due to more conservative monitoring behavior. While more thorough and extensive comparison studies are still ongoing, preliminary results presented in this work clearly demonstrate the effectiveness and attractiveness of this novel SPC framework in effective online monitoring of big data streams and robust, reliable fault detection.

Table 1. Summary of fault detection speed (characterized by out-of-control run length) and false alarm rate using three monitoring frameworks.

Fault No.	SPC	PCA- $T^2$	SVM
IDV 2	125 (0.27%)	216 (0.5%)	180 (0.8%)
IDV 3	95 (0.27%)	366 (0.5%)	16815 (83%)
IDV 13	128 (0.27%)	1131 (0.5%)	675 (12.7%)

## 5. Conclusion

In this work, we present a novel, powerful chemical process monitoring and fault detection framework for nonparametric, heterogeneous big data streams. In particular, the heterogeneity nature of this framework is enabled by recent advancements in SPC, such as the quantile-based multivariate CUSUM (Ye and Liu, 2022). We also compare the performance of this SPC framework with two benchmark process monitoring algorithms, PCA- $T^2$  and nonlinear SVM, in the classic example of Tennessee Eastman Process. Compared with existing dimensionality reduction or machine learning based approaches, the SPC-based framework possesses several advantages, including high reliability and accuracy with customizable, precisely controlled false alarm rate, guaranteed detection of process anomalies or mean shifts, significantly faster fault detection speed, unique

capabilities to handle nonparametric and heterogeneous big data streams, low computational costs, etc.

## References

- E.B. Andersen, I.A. Udugama, K.V. Gernaey, A.R. Khan, C. Bayer, M. KulaHCI, 2022, An easy to use GUI for simulating big data using Tennessee Eastman process, *Quality and Reliability Engineering International*, 38, 1, 264-282.
- J.J. Downs, E.F. Vogel, 1993, A plant-wide industrial process control problem, *Computers and Chemical Engineering*, 17, 3, 245-255.
- R. Fezai, M. Mansouri, O. Taouali, M.F. Harkat, N. Bouguila, 2018, Online reduced kernel principal component analysis for process monitoring. *Journal of Process Control*, 61, 1-11.
- P. Geladi, B.R. Kowalski, 1986, Partial least-squares regression: a tutorial. *Analytica Chimica Acta.*, 185, 1-17.
- S. Heo, J.H. Lee, 2018, Fault detection and classification using artificial neural networks, *IFAC-PapersOnLine*, 51, 18, 470-475.
- J.E. Jackson, G.S. Mudholkar, 1979, Control Procedures for Residuals Associated with Principal Component Analysis. *Technometrics*, 21, 341-349.
- K. Liu, J. Shi, 2013, Objective-oriented optimal sensor allocation strategy for process monitoring and diagnosis by multivariate analysis in a Bayesian network, *IIE Trans.*, 45, 6, 630-643.
- K. Liu, Y. Mei, J. Shi, 2015, An Adaptive Sampling Strategy for Online High-Dimensional Process Monitoring, *Technometrics*, 57, 3, 305-319.
- J. Li, 2020, Efficient global monitoring statistics for high-dimensional data, *Qual. Rel. Eng. Int.*, 36, 1, 18-32.
- R. López-Negrete, S. García-Muñoz, L.T. Biegler, 2010, An efficient nonlinear programming strategy for PCA models with incomplete data sets, 24, 301-311.
- L. Ma, J. Dong, K. Peng, 2020, A novel key performance indicator oriented hierarchical monitoring and propagation path identification framework for complex industrial processes, *ISA Trans.*, 96, 1-13.
- Y. Mei, 2010, Efficient scalable schemes for monitoring a large number of data streams, *Biometrika*, 97, 2, 419-433.
- Y. Mei, 2011, Quickest detection in censoring sensor networks, *Proc. IEEE Int. Symp. Inf. Theory Proc.*, 2148-2152.
- M. Onel, C.A. Kieslich, E.N. Pistikopoulos, 2018, A nonlinear support vector machine-based feature selection approach for fault detection and diagnosis: Application to the Tennessee Eastman process, *AIChE Journal*, 65, 3, 992-1005.
- P. Qiu, D. Hawkins, 2001, A Rank-Based Multivariate CUSUM Procedure, *Technometrics*, 43, 2, 120-132.
- P. Qiu, D. Hawkins, 2003, A nonparametric multivariate cumulative sum procedure for detecting shifts in all directions, *Journal of the Royal Statistical Society: Series D*, 52, 2, 151-64.
- P. Qiu, Z. Li, 2011, On nonparametric statistical process control of univariate processes, *Technometrics*, 53, 4, 390-405.
- E. Russell, L.H. Chiang, R.D. Braatz, 2000, Data-driven methods for fault detection and diagnosis in chemical processes. Springer, 10-45.
- A.G. Tartakovsky, B.L. Rozovskii, R.B. Blažek, H. Kim, 2006, Detection of intrusions in information systems by sequential change-point methods", *Statist. Methodol.*, 3, 3, 252-293.
- W.H. Woodall, D.J. Spitzner, D.C. Montgomery, S. Gupta, 2004, Using Control Charts to Monitor Process and Product Quality Profiles, *Journal of Quality Technology*, 36, 3, 309-320.
- X. Xian, C. Zhang, S. Bonk, K. Liu, 2021, Online monitoring of big data streams: A rank-based sampling algorithm by data augmentation, *Journal of Quality Technology*, 53, 2, 135-153.
- H. Ye, K. Liu, 2022, A Generic Online Nonparametric Monitoring and Sampling Strategy for High-Dimensional Heterogeneous Processes, *IEEE Trans. Auto. Sci. Eng.*, 19, 3, 1503-1516.
- C. Zou, W. Jiang, Z. Wang, X. Zi, 2015, An efficient on-line monitoring method for high-dimensional data streams, *Technometrics*, 57, 3, 374-387.

# Perspectives on New Trends of Statistical Process Monitoring for Industrial Process Safety

Cheng Ji, Jingde Wang, Wei Sun\*

*<sup>a</sup> College of Chemical Engineering, Beijing University of Chemical Technology, 100029, North Third Ring Road 15, Chaoyang District, Beijing, China*

*\*emailOfTheCorrespondingAuthor: [sunwei@mail.buct.edu.cn](mailto:sunwei@mail.buct.edu.cn)*

## Abstract

Smart manufacturing with the help of internet of things and artificial intelligence has attracted significant attention as the modern process industry is highly integrated with information technology. Among them, statistical process monitoring (SPM) is a powerful big data analytics tool for ensuring industrial process safety. In this work, developments in multivariate SPM and corresponding monitoring indicators are briefly introduced. Researches regarding new trends of SPM, the feature-based SPM methods are reviewed and discussed, including statistics pattern analysis and statistics Mahalanobis distance. The feature-based SPM has been considered as one of possible directions in next generation SPM, which shows better process monitoring performance in both sensor faults and process faults through two case studies. Meanwhile, certain limitations and challenges also can be revealed. We discuss the reasons for these challenges from the perspective of industrial process data, and suggestions are provided accordingly to make feature-based SPM better contribute the intelligent manufacturing.

**Keywords:** multivariate statistical analysis, feature selection and extraction, sensor and process fault detection, intelligent manufacturing, internet of things.

## 1. Introduction

The widespread application of distributed control systems facilitates the rapid development of data-driven process monitoring. Statistical process monitoring (SPM), as a class of highly interpretable data-enabled modeling methods, have become a powerful tool for industrial process safety. This is achieved mainly by establishing feature extraction models from historical data to support prompt decision-making regarding the operating status of a process. Statistical process control (SPC) chart, such as cumulative sum chart (CUSUM) chart and exponentially weighted moving average (EWMA) chart could be considered as the earliest SPM, in which univariate upper and lower control limits are employed to determine whether the process is remaining in a state of statistical control (MacGregor, 1994). However, with the extensive arrangement of instruments since 1980s, the limitations of univariate SPM have been revealed. The most important one is that the correlations among process variables are not considered (MacGregor, 1994), since no model is exploited in SPC. Meanwhile, multivariate SPM (MSPM) combining latent variable models and monitoring indicators appears and rapidly develops (Kresta, et al., 1991). In last thirty years, people have witnessed various successful application of MSPM, such as principal component analysis (PCA), partial least squares (PLS) and their variants (Ji and Sun, 2022, Qin, 2012). Although MSPM has become a standard for industrial applications, there are still several challenges to be addressed. The first one is that data are assumed to be multivariate normally distributed in most MSPM methods, while the assumption is usually not satisfied because the



complex features contained in industrial data, such as nonlinearity, process dynamics, and non-Gaussianity, etc(He and Wang, 2018). Although many efforts have been made to deal with this issue through kernel tricks and dynamic latent variable models, the monitoring performance still needs to be further improved. Moreover, the monitoring indicators and corresponding boundary of control region in most methods are still determined under the assumption of normal distribution. High missing alarm rate and false alarm rate will be triggered despite several improvements in monitoring indicators have been developed. On the other hand, traditional MSPM methods focus on process faults, but the incipient sensor faults, which refers to faults with tiny magnitudes(Ji, 2021), are generally not considered. Therefore, more advanced methods need to be proposed to address the above challenges. Wang and He proposed a new concept named statistics pattern analysis (SPA), in which features of process variables rather than the process variables themselves are employed as the monitoring statistics(Wang and He, 2010). Motivated by SPA, more feature-based SPM methods have been further proposed(Ji, 2021, Shang, et al., 2017). Some researchers suggest that it can be regarded as a candidate for next generation SPM(He and Wang, 2018). However, it can be observed that current feature-based SPM methods are mostly designed for incipient sensor fault, and their superiority in monitoring process faults cannot be clearly demonstrated compared to state-of-the-art MSPM. The reason could be the lack of sufficient analysis concerning the characteristics of industrial process data. Regarding this issue, this work compares common MSPM methods and feature-based SPM methods through two cases, by which the challenges in existing feature-based SPM are analyzed. Combined with the case studies, promising suggestions for overcoming these challenges are discussed, and the potential of feature-based SPM in monitoring various types of faults in industrial processes are expected.

## 2. Related Methods

In this section, common MSPM and feature-based SPM methods used in this work for comparison are briefly introduced.

### 2.1. Multivariate Statistical Process Monitoring

MSPM methods mostly employ multivariate linear projection to obtain latent features that could explain most information of the original process variables, and then monitoring indicators are established in the feature space to perform fault detection. Considering MSPM methods have been widely reviewed in literature, only PCA is briefly introduced, and the focus will be on the summary of monitoring indicators.

#### 2.1.1. Principal Component Analysis

PCA is the most commonly used dimensionality reduction method by extracting the cross-correlation among process variables, which has been widely applied in SPM. Given a set of historical data  $X_{n \times m}$  with  $n$  samples and  $m$  variables, PCA aims to find a projection  $P_{m \times k}$  that could transform  $X_{n \times m}$  into an orthogonal space, where most information of original data can be explained,

$$X_{n \times m} = TP_{m \times k}^T + E \quad (1)$$

where  $T = X_{n \times m}P_{m \times k}$  is the score matrix,  $E$  is the residue matrix, and  $k$  represents the number of principal components. To perform fault detection, monitoring indicators could be established on the principal component space and residual space respectively.

### 2.1.2. Monitoring Indicators

$T^2$  statistic is a widely applied monitoring indicator in MSPM to measure changes in the principal component space, which is calculated as follows,

$$T^2 = \sum_{i=1}^k \frac{(p_i^T x)^2}{\lambda_i} \quad (2)$$

where  $p_i$  is the projection of the  $i_{th}$  principal component direction. Comparatively, squared prediction error (SPE) is commonly used to measure the changes in the residue space, which can be calculated as follows,

$$SPE = \sum_{i=k+1}^m (p_i^T x)^2 \quad (3)$$

Moreover, the Mahalanobis distance (MD) is also widely used to measure the changes in all projection directions.

$$MD = \sqrt{\sum_{i=1}^m \frac{(p_i^T x)^2}{\lambda_i}} \quad (4)$$

Raich and Cinar further proposed a combined indicator that could consider changes in principal component space and residue space simultaneously (Raich and Cinar, 1996),

$$\varphi_1 = c \frac{T^2}{\chi^2} + (1 - c) \frac{SPE}{\delta^2} \quad (5)$$

where  $\chi^2$  and  $\delta^2$  are the control limit of  $T^2$  and SPE statistic, and  $c$  is a weight coefficient. However, a high missing alarm rate will be triggered because  $\varphi_1$  could be less than 1 even if either  $T^2$  statistic or SPE statistic exceeds its threshold (Yue and Qin, 2001). Aiming at this issue, Yue and Qin suggested using the following combined indicator (Yue and Qin, 2001),

$$\varphi_2 = \frac{T^2}{\chi^2} + \frac{SPE}{\delta^2} \quad (6)$$

and the control limit is also derived distribution of quadratic forms. Regardless of which monitoring indicator is adopted, the challenges in MSPM mentioned before still need to be faced.

## 2.2. Feature-based Statistical Process Monitoring

The feature-based SPM is a concept proposed by Wang and He (Wang and He, 2010), where the features of process variables rather than the process variables themselves are employed as the monitoring statistics. Several related methods will be introduced in this part.

### 2.2.1. Dissimilarity of Process Data

Actually, the concept of feature-based SPM could be reflected in early methods, such as the dissimilarity of process data (DISSIM) (Kano, et al., 2002). In this work, the distribution of process variables, especially the correlation among variables, which can be regarded as a kind of feature, are monitored. Therefore, this method works better than traditional MSPM on detecting changes in correlation structure. However, certain faults cannot be effectively detected since the features of variables considered by this method are still limited.

### 2.2.2. Statistics Pattern Analysis

The concept of feature-based SPM is first proposed with SPA. In SPA, process variables are divided into moving windows, and the features of the variables including means, variance, etc. in the windows are monitored. Then the PCA is established on the selected features for fault detection. In the case study, hundreds of features are considered by SPA, and therefore it is applicable to most kinds of faults.

### 2.2.3. Statistics Mahalanobis Distance

Motivated by SPA, statistics Mahalanobis distance (SMD) is newly proposed for incipient sensor fault detection (Ji, 2021). Unlike SPA, the MD is directly established on the features of process variable as the monitoring indicator, and the results show that the SMD performs perfectly in three common incipient faults by only considering mean and variance, which greatly reduces the computational costs.

## 3. Case Studies

In this section, a common numerical example and Tennessee Eastman (TE) process are used to compare the process monitoring performance of MSPM and feature-based SPM. Datasets and the implementation of all the methods are available in literature, and the codes can be obtained by request to the corresponding author.

### 3.1. A Numerical Example of Incipient Sensor Faults

A simple multivariate process given in literature is used as the first case study (Ji, 2021). In this case, a set of 100,000 samples are generated under normal conditions as the training data, and three sets containing 4000 samples each are used as test data. The three test data sets correspond to three common incipient sensor faults, which are constant bias, gain degradation, and precision degradation. The faults are introduced in the 1500<sup>th</sup> sample, and the fault detection rate (FDR) and false alarm rate (FAR) are adopted to evaluate the fault detection performance.

**Table 1 A summary of process monitoring results for incipient sensor faults**

FDR/FAR %	PCA $T^2$	PCA SPE	MD	$\varphi_1$	$\varphi_2$	SMD	DISSIM	SPA $T^2$	SPA SPE
Fault 1	1/1	1.7/1.3	4.8/0.9	2/1	0.8/0.5	<b>98.7/0.5</b>	0.3/0	61.5/1.3	97/2.7
Fault 2	0.8/1	0.8/0.8	1.6/1.1	0.7/1.1	0.4/0.5	<b>97.8/3.6</b>	7.2/0.3	1/9.7	9/1.1
Fault 3	2.6/1.2	4.4/0.6	21.9/0.9	6.2/1.1	1.9/0.6	<b>97.5/3.7</b>	99.1/0.9	76.9/4.4	98.3/0.5

The results are summarized in Table 1. It can be found that all MSPM methods fail to detect the three incipient faults because the magnitude of the fault is low. By contrast, feature-based SPM methods perform much better. Among them, DISSIM only works on fault 3. It is reasonable because only variable correlation is considered and it has been proposed for twenty years. SPA works well in fault 1 and fault 3, but the FDR of fault 2 is low because the window in which the gain degradation is most statistically evident is when the fault has just been introduced. When the window moves forward, the dissimilarity of faulty data and normal data is not that obvious, resulting in a low FDR. Moreover, the features used in SPA are too much, which may cause redundancy and decrease the sensitivity of the model to faults. Comparatively, SMD achieves a more than 95% FDR for all faults by considering only the mean and variance, which indicating its superiority to monitoring incipient sensor faults.

### 3.2. Tennessee Eastman Process

TE process is the most commonly used benchmark process in process monitoring domain. Two faults including a random variation fault and a sensor fault are taken as examples, which both are introduced at the 160<sup>th</sup> sample.

**Table 2 A summary of process monitoring results for TE process**

FDR/FAR %	PCA $T^2$	PCA SPE	MD	$\varphi_1$	$\varphi_2$	SMD	SPA $T^2$	SPA SPE
Fault 10	29.6/0	25.8/0.6	88.4/0	45.5/0.6	28.6/0	<b>100/100</b>	79.5/-	73.9/-
Fault 21	39.3/0	47.2/3.1	54.7/1.9	49.6/3.7	40.6/0	<b>100/100</b>	99.9/-	90.8/-

The results are summarized in Table 2. The SMD cannot be applied to monitor process faults because it is developed for sensor faults, and different from sensor variables, the process variables are highly influenced by random factors. As shown in Figure 1, MD performs the best among all indicators in MSPM for fault 10 because changes in all principal components are added up to obtain the MD statistic. This fault is hard to detect for its low magnitude and the response of control systems, while the SPA still performs well as long as there are features deviated from normal operating conditions. For fault 21, a valve is fixed at a constant value. As shown in Figure 2, the fault can be detected by all MSPM methods only after it has affected process variables. It can only be timely detected by SPA because the fault is reflected in the variance of the valve, which is a monitored feature of SPA. Overall, feature-based SPM represented by SPA shows obvious advantages in monitoring both sensor faults and process variables. However, there are also several challenges should be focused on, which will be discussed in next section.

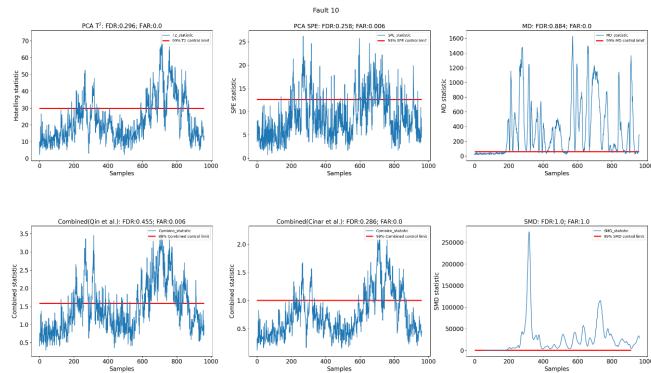


Figure 1 Process monitoring results for fault 10

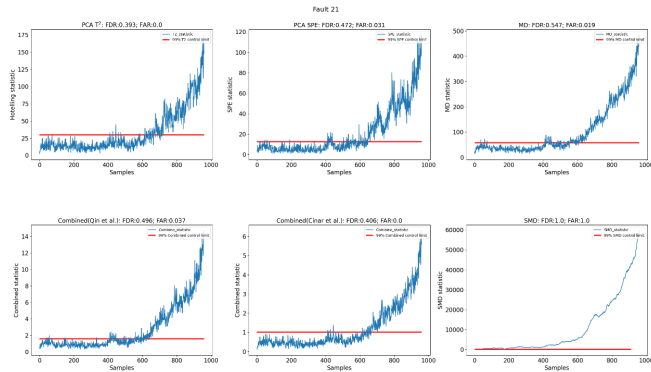


Figure 2 Process monitoring results for fault 21

## 4. Challenges and Suggestions for Feature-Based Statistical Process Monitoring

### 4.1. Redundancy of Features

The feature selection plays the most important role in feature-based SPM. The deficiency or redundancy will affect the process monitoring performance. For example, there are 265 features used for TE process in SPA. The number could be even larger for industrial processes with more variables. It can not only lead to high computational costs but also affect the sensitivity of fault detection. In particular, variable correlation, although an important monitoring statistic, will contribute at least  $m(m - 1)/2$  features. It is

suggested that features more than two orders could be excluded because they are regarding to the distribution of data, and the changes in the distribution could be reflected in other low-order statistics. Moreover, orthogonal transformation is recommended as a preprocessing to avoid the calculation of features of variable correlation.

#### 4.2. Poor Detectability

Since feature-based SPM is a class of methods using moving window technique, the detectability is an unavoidable challenge. Generally, slowly changing faults cannot be early detected by common features of the window, such as the mean and variance. To address this issue, secondary features are suggested to be employed, especially for the maximum, minimum and slope of the window.

#### 4.3. Parameter Selection

At last, the parameter selection is also a challenge that has not been effectively solved. The results of feature-based SPM are highly affected by the size of the moving window. Especially, it is not fixed, but determined according to target processes. It is suggested to be determined by a quantitative indicator obtained by information entropy or other mathematical methods in future research.

## 5. Conclusion

In this work, the performance of MSPM and feature-based SPM in monitoring both sensor faults and process faults is compared through two case studies. The results show that traditional MSPM is not applicable to monitor incipient sensor faults, while the FDR of the SMD reaches an average of 98%. Moreover, the SPA also achieves better process monitoring performance for process faults. It can be concluded that the feature-based SPM represented by SPA shows strong applicability to various types of faults in industrial processes and could be regarded as a candidate to the next generation of SPM. Existing challenges are also analyzed and corresponding solutions are discussed. Future work will focus on the development of more effective feature-based SPM method by addressing the challenges outlined in this paper.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant numbers 22278018).

## References

- Q. P. He, J. Wang, 2018, Statistical process monitoring as a big data analytics tool for smart manufacturing, *J. Process Control* 67, 35-43.
- C. Ji, W. Sun, 2022, A Review on Data-Driven Process Monitoring Methods: Characterization and Mining of Industrial Data, *Processes* 10(2), 335.
- H. Ji, 2021, Statistics Mahalanobis distance for incipient sensor fault detection and diagnosis, *Chem. Eng. Sci.* 230.
- M. Kano, S. Hasebe, I. Hashimoto, H. Ohno, 2002, Statistical process monitoring based on dissimilarity of process data, *AIChE Journal* 48(6), 1231-1240.
- J. V. Kresta, J. F. MacGregor, T. E. Marlin, 1991, Multivariate statistical monitoring of process operating performance, *The Canadian Journal of Chemical Engineering* 69(1), 35-47.

- J. F. MacGregor, 1994, Statistical Process Control of Multivariate Processes, IFAC Proceedings Volumes 27(2), 427-437.
- S. J. Qin, 2012, Survey on data-driven industrial process monitoring and diagnosis, Annual Reviews in Control 36(2), 220-234.
- A. Raich, A. Cinar, 1996, Statistical process monitoring and disturbance diagnosis in multivariable continuous processes, AIChE J. 42(4), 995-1009.
- J. Shang, M. Chen, H. Ji, D. Zhou, 2017, Recursive transformed component statistical analysis for incipient fault detection, Automatica 80, 313-327.
- J. Wang, Q. P. He, 2010, Multivariate statistical process monitoring based on statistics pattern analysis, Ind. Eng. Chem. Res. 49(17), 7858-7869.
- H. H. Yue, S. J. Qin, 2001, Reconstruction-based fault identification using a combined index, Ind. Eng. Chem. Res. 40(20), 4403-4414.



# Aeration Optimization and Control for Wastewater Treatment Processes

Ioana Naşcu<sup>a,b</sup>, Wenli Du<sup>b</sup>, Ioan Naşcu,<sup>a</sup>

<sup>a</sup>*Department of Automation Technical University of Cluj Napoca, Romania*

<sup>b</sup>*Key Laboratory of Advanced Control and Optimization for Chemical Processes, East China University of Science and Technology, Shanghai 200237, China*

[Ioana.Nascu@aut.utcluj.ro](mailto:Ioana.Nascu@aut.utcluj.ro)

## Abstract

In this work we develop model based control strategies for the optimization and control of the aeration of a municipal wastewater treatment plant, intermittently operated by switching the aeration on and off. The proposed strategies aim to reduce energy consumption as well as to further improve effluent quality. A two-layer hierarchical control structure is implemented: (i) the lower layer using a modified version of the GPC to control the dissolved oxygen concentration and minimize the aeration energy and (ii) the higher level (optimization layer) where standard GPC algorithm is used to obtain the optimal values for the setpoint of the dissolved oxygen in the bioreactor. The developed control strategies show good performances and significant energy savings in comparison with the conventional control systems.

**Keywords:** wastewater, MPC, hierarchical control, energy efficiency

## 1. Introduction

The optimal treatment and administration of industrial and municipal wastewater is critical in maintaining a safe and clean environment as well as supporting the health of the community. When dealing with Wastewater Treatment Plants (WWTPs) it is important that the plant operation can satisfy the requirements for the effluent water quality. Unfortunately this leads to a large amount of energy consumption. Therefore, another important goal for the WWTP, which has become a hot research topic during the last years, is energy efficiency optimization (Daverey et al. 2019, Cardoso et al. 2021).

The main part of the WWTPs is the biological treatment where the most used technology is the Activated Sludge Process (ASP). This is due to its cost-effectiveness, its flexibility and high reliability as well as the capability to produce effluent of high quality. Usually, the activated sludge processes are highly complex, nonlinear and challenging to control. Various works have focused on developing accurate models used to design and tune controllers (Harja and Nascu 2019, Naşcu et al. 2022). There has also been some focus on the development of multiple model based controllers as well as predictive (Nascu and Nascu 2016, Bernardelli et al. 2020), adaptive (Vlad et al. 2010), fuzzy (Yelagandula and Ginuga 2022) or fractional order PID control (Harja et al. 2015).

To optimize the performances of WWTP, real time control is a generally accepted method but one must also consider model uncertainties as well as the complexity of the process. In this work a control system with hierarchical structure is developed: a lower layer with process level control loops as well as a higher level dedicated for optimization.

Model based control (MPC) strategies have a variety of advantages when trying to control a process and are widely used in WWTP. They are robust against variable and unknown



time-delay and have good disturbance rejection properties. Moreover, they are able to incorporate constraints explicitly and handle uncertainty in process model parameters and model mismatch (Harja et al. 2016, Eagalapati et al. 2022). In this work MPC will be used in both layers of the control system.

## 2. Plant description an simulation platform

### 2.1. Plant Structure

The WWTP presented in this work is a medium municipal WWTP (maximum flow - 4500 m<sup>3</sup>/day). For this plant there are three main stages: mechanical or pre-treatment, biological treatment, and sludge treatment. In this work we will focus on the biological treatment, the most complex process used in WWTPs. The typical procedure for this treatment, is the activated sludge technology.

The biological treatment part is composed of two circular ring tanks. The inner ring is given by the secondary settler and the outer cylindrical tank represents the biological treatment tank. By switching the aeration on and off, they are operated alternately for nitrification and denitrification. In small and medium WWTP, as the one presented in this work, the bioreactor tank operates for nitrification and denitrification, alternatively. This results in having both processes for removal of nitrogen not spatially separated but sequentially in time by having the bioreactor operate once aerobic and once anaerobic in turn. To keep the contents of the bioreactor completely mixed, the bioreactor has to be equipped with submersible mixers. The aeration is performed by using fine bubble porous diffusers that are situated at the bottom part of the tank. The aeration required by the bioreactor is given through blowers with a maximum airflow of  $Q_{max}=1400$  m<sup>3</sup>/h as well as the capability to operate at variable airflows. After the bioreactor, in the secondary settler the cleaned water from the top of the settler will overflow to the pumping station of the effluent. The sludge that is settled is recirculated to the bioreactor. A small part of the sludge is evacuated to the sludge treatment stage. The aerated bioreactor has a volume of 6483m<sup>3</sup> and the volume of the secondary clarifier is 2048m<sup>3</sup>.

### 2.2. Mathematical model

Two models are used for the modeling and simulation of the WWTP processes: (i) The Activated Sludge Model No 1 (ASM1) (Henze et al. 2000) to simulate the biological processes from the aerated bioreactor and (ii) the double-exponential settling velocity model to simulate the transfer between the multiple layers in the settler. It is assumed that in the settler no biological processes occur.

The bioreactor tanks succeeded by a secondary settler represents the WWTP layout in the Benchmark Simulation Model no. 1 (BSM1)(Alex et al. 2008) where the first biological reactor tanks are anoxic well mixed and the next biological reactor tanks are aerated. To model the settler ten layers are used. The plant is developed such that, using nitrification and denitrification processes, it removes the nitrogen in wastewater. The nitrification takes place in the aerated tanks while denitrification takes place in the anoxic tanks.

#### 2.2.1. Modification of the Simulation Platform

The BSM1 simulation platform mentioned above has to be adjusted such that it accurately represents the real physical WWTP described. Hence, some bioreactors as well as the internal recirculation are excluded from the BSM1 model. An intermittent biological reactor tank having anoxic and aerated phases chronologically alternated, the secondary settler as well as the main recirculation remains in the model. The block diagram for the new Simulink model obtained from BSM1 is presented in Figure 1. The model initialization parameters are changed (according to the model calibration presented in (Harja and Naşcu 2020, Harja and Nascu 2019)) to match the real plant parameters.

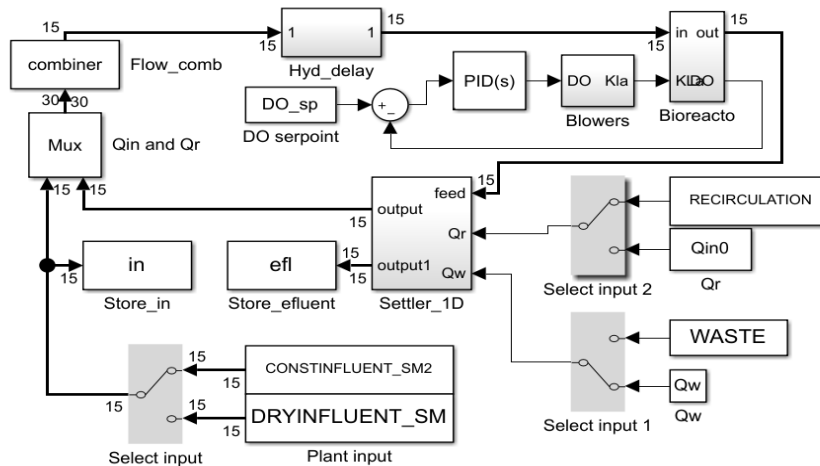


Figure 1 Modified Benchmark Simulation Model 1

To model the blowers and aeration system for the real process as accurately as possible, some modifications were performed. The variable from the BSM1 model that is used for the control of the dissolved oxygen concentration is the transfer coefficient  $KLa$ . Unfortunately, this variable cannot be used in a real plant. Therefore, to convert the aeration flow rate (further used as manipulated input for the bioreactor) into  $KLa$ , a new block is created. In technical literature, the relationship between these two variables is usually represented by an exponential model given by:

$$KLa(Q_{air}) = k_1 \cdot (1 - \exp^{-k_2 \cdot Q_{air}})$$

Using the data collected from the real process for calibration to determine the values for  $k_1$  and  $k_2$  parameters,  $k_1 = 77$  and  $k_2 = 0.0015$  were obtained.

To ensure the bioreactor with alternating nitrification and denitrification phases, the blowers are operating switching ON-OFF. To keep the DO concentration at the reference value determined by the optimization level, a control algorithm will be used.

### 3. Control Strategies

#### 3.1. Hierarchical Control Structure

A simplified diagram of the hierarchical control structure used in this paper is presented in Figure 2. For the substrate concentration control in the bioreactor, the optimization control layer comprises of a MBPC strategy. The optimal values for the DO setpoint is given by the MBPC output.

The lower layer is composed of two main control loops: (i) control of the air pressure in the air piping system and (ii) the control of the DO in the aeration tank for which the optimization level generates the setpoint. The pressure control loop design is rather straightforward and is not presented in this work, unlike the control loop for the DO concentration which poses challenging issues. The ASPs are complex processes with significant nonlinearities, large disturbances as well as uncertainty in process model parameters. These challenges will be tackled by the MBPC algorithm which is used for the control loop of the DO concentration.

In this work, two versions for the Generalized Predictive Control (GPC) method are implemented. In the optimization layer, the standard GPC algorithm is used to obtain the optimal values for the setpoint of the DO in the aeration tank. In the lower layer, a

modified version of the GPC is used to control the DO concentration and to minimize the aeration energy.

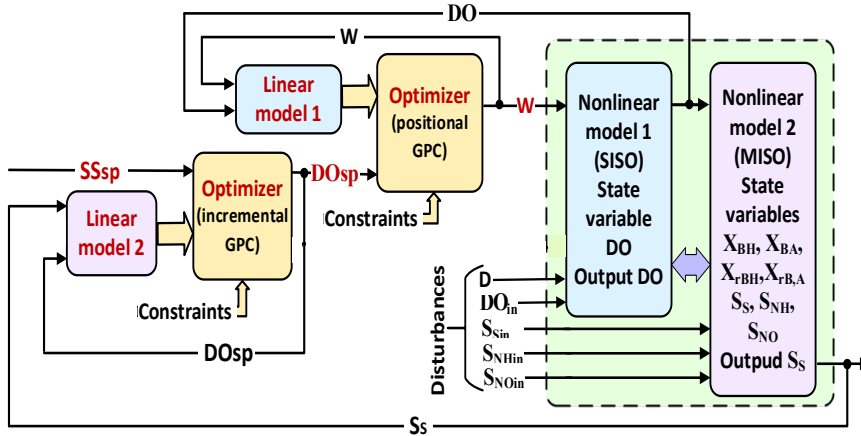


Figure 2. Hierarchical control system layers

### 3.2. GPC Algorithm

#### 3.2.1. Standard GPC Algorithm

The GPC algorithm, developed by Clarke et al is a MBPC strategy that uses the process model to compute online the process output predictions and the optimal control. The optimal control is based on the minimization of the cost function (Clarke et al. 1987):

$$J = \sum_{j=N_1}^{N_2} [y(t+j) - y_r(t+j)]^2 + \sum_{j=0}^{N_u-1} \rho \Delta u(t+j)^2 \tag{1}$$

The design parameters represent:  $N_1$  - the minimum costing horizon,  $N_2$  - the maximum costing horizon,  $N_u$  - the control horizon,  $\rho$  - the control weight parameter. The process input and output are given by  $u$  and  $y$  respectively and  $y_r$  represents the future reference sequence.

#### 3.2.2. Modified GPC Algorithm

An essential step to develop the control algorithm for DO concentration is to reparametrize the cost function (1) to include a measure of the energy that is consumed by aeration. This could lead to significant savings of energy by exploiting the fluctuation of the operating conditions. Because the aeration airflow  $W$  represents the manipulated process input given by the output, of the controller, the reparametrized cost function needs to include the controller output  $u$ , rather than the controller output increments  $\Delta u$  used in (1), to be able to minimize the airflow instead of its variations. This leads to a positional form of the cost function of the controller:

$$J = \sum_{j=N_1}^{N_2} [y(t+j) - y_r(t+j)]^2 + \sum_{j=0}^{N_u-1} \rho u(t+j)^2 \tag{2}$$

## 4. Results

The model as well as the control strategies presented in this paper are implanted in MATLAB Simulink. The model kinetic and stoichiometric parameters used in this paper are:  $Y_A = 0.24$ ,  $Y_H = 0.7$ ,  $f_P = 0.08$ ,  $i_{XB} = 0.086$ ,  $i_{XP} = 0.06$ ,  $\mu_H = 4.59$ ,  $K_S = 20.0$ ,  $K_{O,H}$

$\mu = 0.5$ ,  $K_{NO} = 0.50$ ,  $b_H = 0.4$ ,  $\eta_g = 0.6$ ,  $\eta_h = 0.8$ ,  $k_h = 2.1$ ,  $K_X = 0.025$ ,  $\mu_A = 0.5$ ,  $K_{NH} = 1.0$ ,  $b_A = 0.05$ ,  $K_{O,A} = 0.4$ ,  $k_a = 0.05$ . In one complete aeration cycle with time period  $T$ , the bioreactor is operated  $0.3T$  hours with aeration and  $0.4T$  hours without. For the design of the standard and modified GPC controllers the following tuning parameters are used: (i) for the lower layer  $N_1=1$ ,  $N_2=20$ ,  $N_u=1$ ,  $r_o=0$ , and the sampling time is  $0.01h$  (ii) for the higher layer  $N_1=1$ ,  $N_2=10$ ,  $N_u=1$ ,  $r_o=0$  and the sampling time is of  $1h$ .

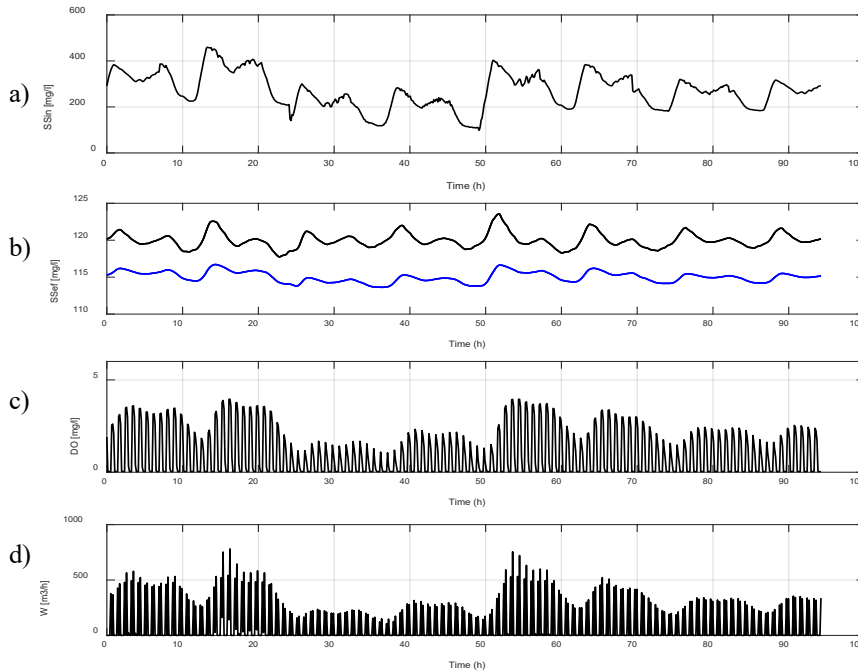


Figure 3. Simulation results. a) load disturbances, b) substrate concentration in the effluent, c) DO concentration in bioreactor, d) aeration flow

It can be observed that the developed control strategies are capable of reducing the concentration of organic substances in the effluent (SSe) in the presence of the variation of the loading of the influent with organic substance (SSin) and simultaneously aiming at a reduction of energy consumption. The variation of the loading of the influent, considered as disturbances are presented in Figure 3a. Figure 3d shows how the aeration flow changes depending on the aeration requirement. As a result, the concentration of dissolved oxygen in the bioreactor changes (Figure 3c) depending on the degree of loading of the influent so as to maintain the value of the concentration of organic substances in the effluent at the reference value (Figure 3b). Figure 3b shows the simulation results for a reference value of  $120 \text{ mg/l}$  (black line) and respectively  $150 \text{ mg/l}$  (blue line).

## 5. Conclusions

In this paper we develop model based control strategies to optimize the aeration of an intermittently operated municipal wastewater treatment plant to reduce the energy consumption as well as to further improve the quality of the effluent. A two-layer hierarchical control structure is implemented: (i) the lower layer using a modified version

of the GPC to control the dissolved oxygen concentration and minimize the aeration energy and (ii) the higher level (optimization layer) where a predictive algorithm using the standard formulation of GPC is used to obtain the optimal setpoint values of the dissolved oxygen in the bioreactor. The developed strategies show good performances, the quality performances of the WWTP are maintained, disturbances from the influent are eliminated, and the content of organic substances is maintained within the prescribed limits while the energy consumption is significantly reduced.

## Acknowledgments

Financial support from Ministry of Research, Innovation and Digitization, CCCDI - UEFISCDI, project number PN-III-P2-2.1-PED-2021-1147, within PNCDI III and with financial support from National Natural Science Fund for Distinguished Young Scholars (61725301), International (Regional) Cooperation and Exchange Project (61720106008).

## References

- Alex, J., Benedetti, L., Copp, J., Gernaey, K., Jeppsson, U., Nopens, I., Pons, M. N., Rieger, L., Rosen, C. and Steyer, J. P. (2008) 'Benchmark Simulation Model no. 1 (BSM1)', *Report by the IWA Taskgroup on Benchmarking of Control Strategies for WWTPs*.
- Bernardelli, A., Marsili-Libelli, S., Manzini, A., Stancari, S., Tardini, G., Montanari, D., Anceschi, G., Gelli, P. and Venier, S. (2020) 'Real-time model predictive control of a wastewater treatment plant based on machine learning', *WATER SCIENCE AND TECHNOLOGY*, 81(11), 2391-2400.
- Cardoso, B. J., Rodrigues, E., Gaspar, A. R. and Gomes, Á. (2021) 'Energy performance factors in wastewater treatment plants: A review', *Journal of Cleaner Production*, 322.
- Clarke, D. W., Mohtadi, C. and Tuffs, P. S. (1987) 'Generalized predictive control-Part I. The basic algorithm', *Automatica*, 23(2), 137-148.
- Daverey, A., Pandey, D., Verma, P., Verma, S., Shah, V., Dutta, K. and Arunachalam, K. (2019) 'Recent advances in energy efficient biological treatment of municipal wastewater', *Bioresource Technology Reports*, 7.
- Eagalapati, S. S. T., Sheik, A. G. and Ambati, S. R. (2022) 'Fractional order-based hierarchical controller design and evaluation with Bürger-Diehl settler model in a total nitrogen removal wastewater treatment process', *Environmental Science and Pollution Research*.
- Harja, G., Muresan, C., Nascu, I. and Vlad, G. (2015) *Fractional order PI control strategy on an activated sludge wastewater treatment process, 2015 19TH INTERNATIONAL CONFERENCE ON SYSTEM THEORY, CONTROL AND COMPUTING (ICSTCC)*.
- Harja, G. and Nascu, I. (2019) *Control of an activated sludge wastewater treatment process based on a Calibrated and modified BSM1 Model*.
- Harja, G. and Naşcu, I. (2020) *Advanced control for nitrogen removal of an intermittently operated ASWWTP*, translated by 1-6.
- Harja, G., Nascu, I., Muresan, C. and Nascu, I. (2016) 'Improvements in Dissolved Oxygen Control of an Activated Sludge Wastewater Treatment Process', *CIRCUITS SYSTEMS AND SIGNAL PROCESSING*, 35(6), 2259-2281.
- Henze, M., Gujer, W., Mino, T. and van Loosdrecht, M. (2000) *Activated Sludge Models ASM1, ASM2, ASM2D, ASM3*.
- Naşcu, I., Du, W. and Ioan, N. (2022) 'An Auto-tuning method for aeration control in activated sludge wastewater treatment processes', in *IEEE 2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, Male, 16-18 nov.2022.
- Nascu, I. and Nascu, I. (2016) 'Modelling and optimization of an activated sludge wastewater treatment process', (*ESCAPE26*), *PT A*, 1159-1164.
- Vlad, G., Crisan, R., Muresan, B., Nascu, I. and Cosmin, D. (2010) *Development and application of a predictive adaptive controller to a wastewater treatment process*.
- Yelagandula, S. and Ginuga, P. R. (2022) 'Control of a Waste Water Treatment Plant Using Fuzzy Logic Controller', *Journal of The Institution of Engineers (India): Series E*, 103(2), 167-177.

# Safe deployment of reinforcement learning using deterministic optimization over neural networks

Radu-Alexandru Burtea<sup>a</sup> and Calvin Tsay<sup>a</sup>

<sup>a</sup>*Department of Computing, Imperial College London, SW7 2AZ, United Kingdom*  
*\*c.tsay@imperial.ac.uk*

## Abstract

Enabling reinforcement learning (RL) to explicitly consider constraints is important for safe deployment in real-world process systems. This work exploits recent developments in deep RL and optimization over trained neural networks to introduce algorithms for safe training and deployment of RL agents. We show how optimization over trained neural-network state-action value functions (i.e., a critic function) can explicitly incorporate constraints and describe two corresponding RL algorithms: the first uses constrained optimization of the critic to give optimal actions for training an actor, while the second guarantees constraint satisfaction by directly implementing actions from optimizing a trained critic model. The two algorithms are tested on a supply chain case study from OR-Gym and are compared against state-of-the-art algorithms TRPO, CPO, and RCPO.

**Keywords:** Constrained reinforcement learning, Optimization of neural network surrogates, Supply chain optimization, Optimization and machine learning toolkit

## 1. Introduction

Reinforcement learning (RL) has been central to many notable successes in machine learning, such as in self-driving cars, playing games, and operating data centers (Shin et al., 2019). Compared to traditional control strategies, e.g., model predictive control, RL does not typically consider state constraints explicitly. For many practical engineering applications, simply maximizing reward without considering the appropriateness of actions can lead to undesirable consequences. For instance, in supply chain applications, an RL algorithm may direct all goods to the cheapest warehouse without considering its maximum capacity, generating infeasible policies and/or creating massive backlogs.

Given the above, it is desirable to impose constraints on the range of behavior that can be explored in RL. This has inspired research into so-called safe RL (García and Fernández, 2015). Several methods for safe RL implicitly consider constraints using stage-wise reward or penalty functions. Alternatively, some safe RL techniques use external knowledge, e.g., imitation learning, and/or risk metrics during exploration.

This work takes advantage of two recent developments to enable explicit consideration of state constraints during deployment of RL: (i) incorporation of deep neural networks (NNs) into RL algorithms, known as deep RL, and (ii) techniques for deterministic optimization of trained NNs. We present two algorithms based on this combination. The first, called OMLT-DDPG, comprises an actor-critic method, where optimization over a critic NN gives optimal actions on which to train the actor. The second, called SAFE, is a strategy for deployment, wherein directly implementing actions from optimizing a trained critic NN guarantees constraint satisfaction.

The two proposed algorithms are applied to a multi-level supply chain case study from the `or-gym` library (Hubbs et al., 2020). Computational results demonstrate that OMLT-DDPG is significantly more sample-efficient compared to other RL methods, owing to the use of deterministic optimization. The results further show that SAFE explicitly satisfies constraints during RL deployment.

## 2. Safe Reinforcement Learning (RL) Background

Markov decision processes, or MDPs, are the foundation of RL problems. A given MDP is defined by a state space  $S$ , action space  $A$ , reward function  $R: S \times A \times S \rightarrow \mathbb{R}$ , and transition probability  $P: S \times A \times S \rightarrow [0,1]$ . The goal of RL is to learn a policy  $\pi: S \rightarrow \mathcal{P}(A)$  that maximizes a performance metric  $J(\pi)$ , usually defined as the total expected reward  $R$  over an infinite time horizon, subject to a discount factor  $\gamma$ . Safe RL methods seek to additionally enforce constraints, typically by introducing some cost function(s), analogous to reward. The inclusion of cost functions results in constrained MDP, or CMDP. Recent methods for safe RL such as constrained policy optimization (Achiam et al., 2017) and interior-point policy optimization (Liu et al., 2020) can provide safety guarantees for CMDPs, but only as simple constraints on expected total discounted cost.

**Trust-Region Policy Optimization (TRPO).** TRPO (Schulman et al., 2015) serves as the basis for several safe RL algorithms: it is a policy iteration algorithm based on computing the advantage of one policy over another, i.e., the expected improvement in the performance metric between the iterations of a policy. Schulman et al. (2015) provide a method for approximating this. As these approximate updates resemble a first-order method when the policy is differentiable, the step size between policies should be constrained. This is typically done using a trust region method, such as constraining the KL divergence to the old policy when maximizing the advantage function.

**Constrained Policy Optimization (CPO).** Achiam et al. (2017) extend TRPO to CMDPs by adding constraints on the auxiliary cost functions to the above trust-region constraint on KL divergence, resulting in the following optimization problem to find policy  $\pi_{k+1}$ :

$$\begin{aligned} \max_{\pi} J(\pi) \\ \text{s.t. } J_{C_i}(\pi) \leq d_i, \forall i = 1, \dots, m \\ D_{KL}(\pi, \pi_k) \leq \delta \end{aligned} \quad (1)$$

where  $J_{C_i}$  is the total expected cost of the  $i^{th}$  constraint over an infinite time horizon and  $\delta$  limits the step size. The objective and constraints are replaced with surrogate functions, based on the algorithm hyperparameters. Again, this problem is computationally difficult and is solved using a primal-dual method after linearizing  $J(\pi)$  and  $J_{C_i}(\pi)$ , and a second-order expansion for  $D_{KL}(\pi, \pi_k)$ . This approximation motivates a small step size  $\delta$ .

**Reward Constrained Policy Optimization (RCPO).** Tessler et al. (2018) propose RCPO, which is similar to CPO but solves an unconstrained optimization problem with Lagrange multipliers. Updating the policy then becomes a bilevel optimization problem:  $\min_{\lambda \geq 0} \max_{\pi} (J(\pi) - \lambda(J_c(\pi) - d))$ . This optimization problem can be viewed on two timescales: a faster one, where the policy is optimized, and a slower one, which involves increasing  $\lambda$  until the constraint is satisfied. This timescale decomposition is achieved by selection of different step sizes for the updates to the Lagrange multipliers and the policy.

---

**Algorithm 1** OMLT-DDPG (one episode)

---

**Require:**  $|R|, \sigma \geq 0$  ▷ Replay buffer, warm-up  
1: Randomly initialize critic  $Q(s, a|\theta^Q)$ , actor  $\mu(s|\theta^\mu)$ , and target NNs  $Q'(s, a|\theta'^Q), \mu'(s|\theta'^\mu)$ .  
2: Observe:  $s_1$   
3: **for**  $t = 1, \dots, T$  **do**  
4:   Take action:  $a_t = \mu(s_t|\theta^\mu)$   
5:   Observe:  $r_t, c_t, s_{t+1}$  and store:  $R \leftarrow (s_t, a_t, r_t, c_t, s_{t+1})$   
6:   Sample  $\hat{R}$  from  $R, |\hat{R}| = N$   
7:    $y_i = r_i + \gamma Q'(s_{i+q}, \mu'(s|\theta'^\mu)|\theta'^Q), \forall i \in \hat{R}$   
8:   Update  $Q(s, a|\theta^Q)$  by minimizing MSE against  $y_i, \forall i \in \hat{R}$   
9:   **if**  $t \geq \sigma$  **then** ▷ OMLT Package  
10:     Optimize:  $a_i^* = \operatorname{argmax} Q(s_i, a_i|\theta^Q)$  s.t. constraints  $c_i$   
11:     Update  $\mu(s|\theta^\mu)$  by minimizing MSE against  $a_i^*, \forall i \in \hat{R}$   
12:   **end if**  
13:    $\theta'^Q \leftarrow \tau\theta^Q + (1 - \tau)\theta'^Q$   
14:    $\theta'^\mu \leftarrow \tau\theta^\mu + (1 - \tau)\theta'^\mu$   
15: **end for**

---

### 3. First Algorithm: OMLT-DDPG

Our approach presented in this section incorporates the Optimization and Machine Learning Toolkit (OMLT) into Deep Deterministic Policy Gradients (DDPG).

**Deep Deterministic Policy Gradients (DDPG).** DDPG (Lillicrap et al., 2016) was conceived as a continuous-space extension to the popular deep Q-learning framework (DQN). DQN cannot be directly applied to continuous action spaces, as selecting the action with maximum Q-value at a given state becomes inefficient for high-dimensional action spaces. DDPG is an off-policy, model-free algorithm that instead uses the Q-value function to estimate the policy gradient. Specifically, DDPG keeps a parameterized policy network, known as the actor, which deterministically maps states to actions. The other component is a critic network, which behaves as the Q-value function used in DQN. The actor is trained using gradients from the critic, and the critic is trained by minimizing the difference between the expected discounted rewards (following the greedy actor) and the current Q-value assigned by the critic to the state-action pair. Most implementations maintain a replay buffer to avoid “catastrophic forgetting” of previous transitions.

**Optimization and Machine Learning Toolkit (OMIT).** The challenge of selecting the action with maximum Q-value from a given neural network (NN) can be viewed as optimization over a trained neural network (the neural network parameters are fixed during an RL step). We propose to address this using the Optimization and Machine Learning Toolkit (OMLT), an open-source package for optimization over pre-trained machine learning models (Ceccon et al., 2022). OMLT enables users to easily translate learned machine learning models to optimization formulations. OMLT 1.0 supports GBTs through an ONNX (<https://github.com/onnx/onnx>) interface and NNs through both ONNX and Keras interfaces. OMLT transforms trained machine learning models into the Python-based algebraic modeling language Pyomo (<https://github.com/pyomo/pyomo>). The literature often presents different optimization formulations as competitors, but in OMLT, competing optimization formulations become alternative choices for users.

**DDPG with Deterministic Optimization of NNs.** Algorithm 1 gives pseudocode of our proposed algorithm. OMLT-DDPG preserves properties of the DDPG algorithm, but is extended for CMDPs and benefits from deterministic optimization of trained NNs.

The algorithm is initialized with actor and critic neural networks, as well as target actor and critic NNs. For each episode the agent observes the initial state and executes the



**Algorithm 2** SAFE**Require:** Trained critic  $Q(s, a | \theta^Q)$ 1: Observe:  $s_1, c_1$ 2: Optimize:  $a_1^* = \operatorname{argmax} Q(s_t, a_i | \theta^Q)$ 3: Action:  $a_1 = a_1^*$  s.t. constraints  $c_1$ ▷ OMLT Package for  $t = 2, \dots, T$  **do**

#:

**end**Optimize:  $a_t^* = \operatorname{argmax} Q(s_t, a_i | \theta^Q)$  s.t. constraints  $c_{t+1}$ 

▷ OMLT Package

5: Select action:  $a_t = a_t^*$ 

6:

policy defined by the actor NN, storing transitions in the replay buffer. The algorithm then samples a batch of transitions of length  $N$  and updates the critic NN against the target. Up to this point, our algorithm closely resembles DDPG; however, DDPG uses the negative value of the critic for a given state-action pair as the loss for training the actor. DDPG-OMLT instead uses the optimal actions obtained by optimizing the critic NN, subject to the problem constraints. The algorithm preserves the theoretical properties of the policy gradients used in DDPG, as the gradient of the loss used for training the actor is obtained by applying a linear operation on the gradient of the policy network.

Note that this involves solving one constrained optimization problem for each sample from the replay buffer. Therefore, we use the predicted action from the actor as the initial guess to expedite optimization. We hypothesize that in the initial stages of training the actions chosen in this manner will be suboptimal, as the critic is not accurate enough to judge state-action pairs accurately. Nevertheless, this may prove beneficial for exploration. Given the above, we introduce a warm-up period  $\sigma$ , where only the actor is fixed and only the critic is updated. Following the actor and critic updates, the target networks are also updated using “soft” updates, i.e., only a certain fraction  $\tau$  of the weights are updated. Lillicrap et al. (2016) found this to improve the stability of DDPG.

**4. Second Algorithm: SAFE**

The above DDPG-OMLT algorithm promotes safe exploration and exploitation by always incorporating environment constraints when optimizing over the critic network to select optimal actions. However, in deployment, the actor can still give an action that results in constraint violation. Here we describe a second algorithm, SAFE, that explicitly enforces constraints in deployment. In an actor-critic setting, the actor network usually gives the action to take (as implied by its name). OMLT (<https://github.com/cog-imperial/OMLT>) enables us to directly use the critic network to choose the optimal actions at each state. Specifically, we can optimize over the critic network to maximizing rewards, while enforcing environment constraints. If an actor NN is available, this can provide initial guesses for optimization. Algorithm 2 gives pseudocode for SAFE.

A key difference between OMLT-DDPG and SAFE is that OMLT-DDPG uses the constraint values in the previous timestep to evaluate the actor (after the action has been taken), while SAFE uses the constraint values in the current timestep to only take an action that is feasible. We note that SAFE is model-agnostic, meaning it can use any state-action value NN (supported by OMLT). While we only test SAFE with critic models from OMLT-DDPG, any Q-value NN that takes a state-action pair as input can be used.

**5. Results and Discussion**

We employ the multilevel supply chain case study from `or-gym` (Hubbs et al., 2020) to test OMLT-DDPG and SAFE, comparing against TRPO, CPO, and RCPO. The agent

must place replenishment orders at nodes throughout levels of a supply chain, subject to lead times and uncertain customer demand at retail nodes. Inventories are subject to capacity constraints, with excess incurring a penalty cost. Likewise, penalties are incurred for unmet demand. Full details can be found in Hubbs et al. (2020). A three-level supply chain is selected, and algorithms are run for 150 episodes, except for RCPO, which required more episodes to converge. Each experiment is repeated with five random starts.

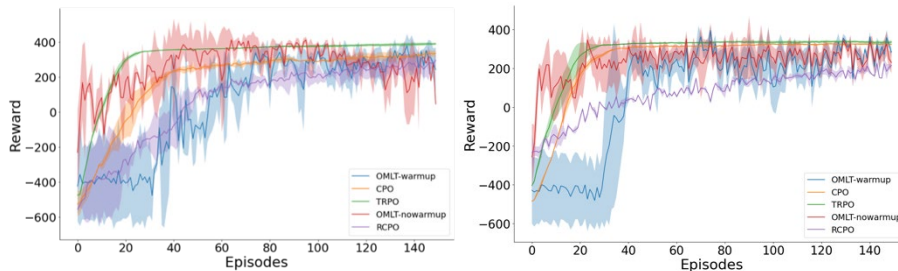


Figure 1: Rewards during training for nominal (left) and relaxed (right) supply chain.

DDPG is typically run with a lower learning rate for the actor than for the critic, allowing the two to converge on different “timescales.” However, we found OMLT-DDPG to benefit from a faster learning rate for the actor, and we set the actor and critic learning rates to, respectively, 0.005 and 0.001. We used a batch size of 8, sampled from a replay buffer size of maximum 35000. We found a strong tradeoff between performance gain and computational time related to the batch size. We use a batch size of 40000 for CPO and TRPO. To investigate the effect of the feasible region size, we consider two settings for the supply chain case study: one with the nominal constraint, and a relaxed setting, where capacity constraints are increased by 40%.

**Safe Reinforcement Learning.** Figure 1 compares the performance of the various algorithms through 150 episodes. TRPO, CPO, and RCPO exhibit more stable behavior, while OMLT-DDPG seems to fluctuate. However, for the nominal case (left), OMLT-DDPG with warm-up achieves rewards only 6% and 28% lower than CPO and TRPO, respectively, which is noteworthy as OMLT-DDPG uses  $>1000x$  fewer samples. For the relaxed case (right), OMLT-DDPG performs intermediate to CPO and RCPO, while again using significantly fewer samples. This sample efficiency may be attributed to the information gain from the use of deterministic optimization. Indeed, OMLT-DDPG achieves a reward of 200 after only four episodes, and this steep learning curve is consistent in our experiments, suggesting behavior as a “few-shot” learner.

The incorporation of the warm-up period does not seem to help OMLT-DDPG in the nominal case (left), but greatly improves performance in the relaxed case (right). In general, the lack of warm-up period results in more aggressive behavior by the agent.

**Safe Deployment.** To simulate safe deployment in production, we deploy the models trained after the 150 episodes. Specifically, we deploy the trained models in a 30-episode simulation, varying the initial state randomly. The rewards and penalties incurred by using these trained models to operate the supply chain are shown in Figure 2. We find that OMLT-DDPG outperforms the other algorithms in deployment in terms of both maximizing rewards and avoiding penalties. Compared to CPO, OMLT-DDPG achieves 21% higher rewards and 20% less penalties in deployment in the nominal case. The SAFE algorithm results in more unstable rewards, but is the only algorithm to guarantee constraint satisfaction, avoiding all penalties during deployment. This demonstrates SAFE as the most appropriate option for guaranteed safe exploitation.

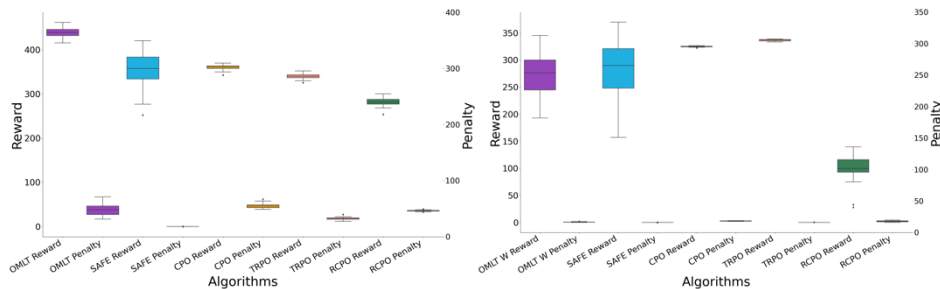


Figure 2: Deployment rewards/penalties for nominal (left) and relaxed (right) cases.

## 6. Conclusions

This paper introduces two algorithms, OMLT-DDPG and SAFE, for safe reinforcement learning and deployment, based on constrained optimization over trained critic networks. The first algorithm uses an actor-critic framework; constrained optimization over the critic network is used to provide targets on which the actor is trained. We show that this algorithm is very sample efficient, resembling behavior of a “few-shot” learner. The second algorithm uses constrained optimization over a pre-trained critic network to explicitly enforce process constraints during deployment.

## 7. Acknowledgments

The authors acknowledge support from the Engineering & Physical Sciences Research Council (EPSRC) grant EP/T001577/1 and an Imperial College Research Fellowship.

## References

- J. Achiam, D. Held, A. Tamar, P. Abbeel, 2017. Constrained policy optimization. In: International Conference on Machine Learning. PMLR, pp. 22–31.
- F. Ceccon, J. Jalving, J. Haddad, A. Thebelt, C. Tsay, C. D. Laird, R. Misener, 2022. OMLT: Optimization Machine Learning Toolkit. *Journal of Machine Learning Research* 23, 1–8.
- J. García, F. Fernández, 2015. A comprehensive survey on safe reinforcement learning. *Journal of Machine Learning Research* 16 (1), 1437–1480.
- C. D. Hubbs, H. D. Perez, O. Sarwar, N. V. Sahinidis, I. E. Grossmann, J. M. Wassick, 2020. Or-gym: A reinforcement learning library for operations research problems. arXiv:2008.06319.
- T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, D. Wierstra, 2016. Continuous control with deep reinforcement learning. In: International Conference on Learning Representations.
- Y. Liu, J. Ding, X. Liu, 2020. Ipo: Interior-point policy optimization under constraints. In: Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 34. pp. 4940–4947.
- J. Schulman, S. Levine, P. Abbeel, M. Jordan, P. Moritz, 2015. Trust region policy optimization. In: International Conference on Machine Learning. PMLR, pp. 1889–1897.
- J. Shin, T. Badgwell, K. Liu, J.H. Lee, 2019. Reinforcement learning—overview of recent progress and implications for process control. *Computers & Chemical Engineering* 127, 282–294.
- C. Tessler, D. J. Mankowitz, S. Mannor, 2018. Reward constrained policy optimization. In: International Conference on Learning Representations.

# Distributional Constrained Reinforcement Learning for Supply Chain Optimization

Jaime Sabal Bermúdez,<sup>a</sup> Antonio del Rio Chanona,<sup>b</sup> and Calvin Tsay<sup>a,b\*</sup>

<sup>a</sup>*Department of Computing, Imperial College London, SW7 2AZ, United Kingdom*

<sup>b</sup>*Sargent Centre for Process Systems Engineering, Imperial College London, SW7 2AZ, United Kingdom*

\**c.tsay@imperial.ac.uk*

## Abstract

This work studies reinforcement learning (RL) in the context of multi-period supply chains subject to constraints, e.g., on inventory. We introduce Distributional Constrained Policy Optimization (DCPO), a novel approach for reliable constraint satisfaction in RL. Our approach is based on Constrained Policy Optimization (CPO), which is subject to approximation errors that in practice lead it to converge to infeasible policies. We address this issue by incorporating aspects of distributional RL. Using a supply chain case study, we show that DCPO improves the rate at which the RL policy converges and ensures reliable constraint satisfaction by the end of training. The proposed method also greatly reduces the variance of returns between runs; this result is significant in the context of policy gradient methods, which intrinsically introduce high variance during training.

**Keywords:** Safe reinforcement learning, Process operations, Inventory management

## 1. Introduction

Recent years have highlighted the importance of efficient supply chains in modern society. The field of inventory management deals with optimization of the ordering, storing, and selling of product inventory. There are many aspects of this problem that make it challenging, and a lot of benefits in doing so that make it worth investigating. Challenges include the presence of outdated “legacy” processes or the inefficient use of available physical capital. A variety of operations research methods have been used in tackling this family of problems, including dynamic programming, linear programming, and game theory. Given complex environments, however, modern inventory management also lends itself towards optimization through reinforcement learning (RL).

In the past few decades, machine learning has gained increasing popularity for its capacity to imitate the way humans learn. RL provides a framework for artificial agents to learn by interacting with an external environment and iteratively updating the way in which they act (i.e., their policy). Specifically, each interaction moves the agent to a new state and produces some quantifiable reward; the agent seeks to maximize cumulative reward (i.e., return) from its interactions. RL has proven very effective in areas such as gaming, demonstrating superhuman performance in incredibly complex games, or robotics and control, where problems can be too complex for analytical solution (Shin et al., 2019).

Real-world environments such as supply chains, may impose physical or safety limitations on the actions an agent can take. This work introduces a framework for RL in these circumstances, combining advantages of constrained and distributional RL.

*Constrained* RL algorithms extend the objective of an RL agent to include minimizing the expected costs below some threshold defined by the constraints in an environment, in addition to maximizing expected returns. In a related vein, *distributional* RL attempts to imitate the way in which organisms consider risk in decision-making, by estimating the distribution of values for all the states in the environment. The intrinsic stochasticity of many environments motivates using probability distributions to evaluate the possible future rewards associated with an action (Petsagkourakis et al., 2022). By estimating these distributions, distributional RL simplifies the parameterization of risk-adversity in RL, giving benefits in problems where constraints should be satisfied under uncertainty.

## 2. Reinforcement Learning (RL) Background

RL often employs a Markov Decision Process (MDP). An MDP is defined by the tuple  $(\mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{P}, \gamma)$ , where  $\mathcal{S}$  is the set of states in the environment,  $\mathcal{A}$  is the set of possible actions,  $\mathcal{R}: \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow \mathbb{R}$  is a reward function,  $\mathcal{P}: \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow [0,1]$  is the transition probability function ( $\mathcal{P}(s'|s, a)$  is the probability of transitioning to state  $s'$  by taking action  $a$  from state  $s$ ), and  $\gamma$  is a discount rate for future rewards. A policy  $\pi: \mathcal{S} \rightarrow \mathcal{P}(\mathcal{A})$  maps states to probability distributions over the possible actions ( $\pi(a|s)$  is the probability of taking action  $a$  from state  $s$ ). We denote the set of all possible stationary policies as  $\Pi$ . RL seeks a policy  $\pi$  that maximizes a performance measure, often taken as  $J_R(\pi) = \mathbb{E}_{\tau \sim \pi} [\sum_t \gamma^t R(s_t, a_t, s_{t+1})]$ , where  $\tau$  is a trajectory sampled using  $\pi$ . Denoting  $\hat{R}(\tau)$  as the discounted return of a trajectory, the on-policy state value function is denoted  $V^\pi(s) = \mathbb{E}_{\tau \sim \pi} [\hat{R}(\tau) | s_0 = s]$ , with optimal policy  $\pi^* = \operatorname{argmax}_\pi V^\pi(s)$ . Analogously, the state-action value function is  $Q^\pi(s, a) = \mathbb{E}_{\tau \sim \pi} [\hat{R}(\tau) | s_0 = s, a_0 = a]$  and the advantage function of an action is  $A^\pi(s, a) = Q^\pi(s, a) - V^\pi(s)$ .

**Trust Region Policy Optimization.** Deep RL may parameterize the learned policy as a neural network (NN) with trainable parameters  $\theta$ . Policy-gradient methods train the NN using the gradient  $\nabla_\theta J(\theta) = \mathbb{E}_\pi [Q^\pi(s, a) \nabla_\theta \ln \pi_\theta(a|s)]$ . This produces an unbiased update rule, but with large variance in the gradient estimates. Trust Region Policy optimization (TRPO) (Schulman et al., 2015a) deals with this noise by limiting the Kullback-Leibler (KL) divergence between consecutive policies to a threshold  $\delta$ :

$$\max_{\theta} \mathbb{E}_{s \sim \rho^{\pi_k}, a \sim \pi_k} \left[ \frac{\pi(a|s)}{\pi_k(a|s)} A^{\pi_k}(s, a) \right] \text{ s.t. } \mathbb{E}_{s \sim \rho^{\pi_k}} [\bar{D}_{\text{KL}}(\pi_k(\cdot | s) || \pi(\cdot | s))] \leq \delta \quad (1)$$

where  $\rho^{\pi_k}$  is the state visitation probability density under  $\pi_k$ , and  $\bar{D}_{\text{KL}}$  the average KL divergence. Problem (1) is typically solved using Monte-Carlo estimates of  $Q^{\pi_k}(s, a)$ , linearization of the objective, and second-order approximation of the constraint.

**Constrained Reinforcement Learning.** Constrained RL algorithms often consider a constrained MDP (CMDP), which includes costs  $\mathcal{C}: \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow \mathbb{R}$  analogously to rewards, as well as a safety threshold  $d$ . The objective of the agent then becomes to maximize  $J_R(\pi)$  such that  $J_C(\pi) = \mathbb{E}_{\tau \sim \pi} [\sum_t \gamma^t \mathcal{C}(s_t, a_t, s_{t+1})] \leq d$ .

*Constrained Policy Optimization (CPO)* (Achiam et al., 2017) restricts policy updates similarly to TRPO, but to satisfy both the trust-region constraint from (1) and environment constraints. The proposed optimization problem can be solved efficiently using duality, where Lagrange multipliers for both the objective and environment constraints are estimated at each step. However, CPO is inherently subject to approximation errors in the estimate of the objective and constraints, which in practice lead it to converge to infeasible policies. Another way of enforcing constraints is to

augment the state-space and reshape the objective of the problem. *Safety Augmented (SAUTE) MDPs* (Sootla et al., 2022) allow for a reformulation of the problem in terms of minimization of the reshaped objective, whereby constraints are incorporated through a safety budget. This approach only modifies the environment and is compatible with most off-the-shelf RL algorithms, but the state space increases with the number of constraints, leading to scalability issues and a loss in sampling efficiency as constraints are added.

### 3. Approach: Distributional Constrained Policy Optimization

In this work we apply distributional RL to the CPO setting to balance reliable constraint satisfaction and consistent sampling efficiency, independent of the number of constraints. Distributional RL (Bellemare et al., 2017) considers the probability distribution of returns over state-action pairs, which we extend to approximate distributions of costs to manage risk aversion. Algorithm 1 shows our distributional version of CPO, which we call DCPO.

**Problem Formulation.** Consider the CPO policy update with a single constraint:

$$\begin{aligned} \pi_{k+1} &= \operatorname{argmax}_{\pi \in \Pi_\theta} \mathbb{E}_{s \sim \rho^{\pi_k}, a \sim \pi_k} [A^{\pi_k}(s, a)] \\ \text{s.t.} \quad & J_C(\pi_k) + \frac{1}{1-\gamma} \mathbb{E}_{s \sim \rho^{\pi_k}, a \sim \pi_k} [A_C^{\pi_k}(s, a)] \leq d \\ & \bar{D}_{\text{KL}}(\pi_k || \pi) \leq \delta \end{aligned} \quad (2)$$

where  $A_C^{\pi_k}$  denotes the estimated cost advantage under  $\pi_k$  for the cost function  $C$ . For small  $\delta$ , we can linearize the objective and safety constraints around  $\pi_k$ . Denoting the gradients of the objective and constraint as  $g$  and  $b$ , respectively, the Hessian of the KL-divergence as  $H$ , and defining  $c = J_C(\pi_k - d)$ , (2) can be written in terms of the step direction  $x = \theta - \theta_k$ , giving:  $x^* = \min_x g^T x$  s.t.  $\{c + b^T x \leq 0, x^T H x \leq \delta\}$ .

We note that  $g$  and  $b$  can be computed by back-propagation, and that  $H$  is always positive semi-definite, making (3) a convex optimization problem. Given a single cost function, the optimal point  $x^*$  (assuming one exists) is  $x^* = \frac{-1}{\lambda^*} H^{-1}(g - v^* b)$ , with dual variables:

$$v^* = \left( \frac{\lambda^* c - r}{s} \right); \lambda^* = \operatorname{argmax}_{\lambda \geq 0} \begin{cases} f_a(\lambda) := \frac{1}{2\lambda} \left( \frac{r^2}{s} - 1 \right) + \frac{\lambda}{2} \left( \frac{c^2}{s} - \delta \right) - \frac{rc}{s}, & \text{if } \lambda c - r > 0 \\ f_b(\lambda) := \frac{-1}{2\lambda} \left( \frac{q}{\lambda} + \lambda \delta \right), & \text{otherwise} \end{cases}$$

where  $q = g^T H^{-1} g$ ,  $r = g^T H^{-1} b$ ,  $s = b^T H^{-1} b$ . In practice  $H^{-1}$  is expensive to compute so we approximately solve for  $H^{-1} g$ ,  $H^{-1} b$  using conjugate gradients. Since  $\lambda \geq 0$ , we must restrict the optimal  $\lambda$  from the two cases above through a projection to the sets  $\Lambda_a := \{\lambda | \lambda c - r > 0, \lambda \geq 0\}$ , and  $\Lambda_b := \{\lambda | \lambda c - r \leq 0, \lambda \geq 0\}$  respectively, such that:

$$\lambda^* \in \left\{ \operatorname{Proj} \left( \sqrt{\frac{q - r^2 s^{-1}}{\delta - c^2 s}}, \Lambda_a \right), \operatorname{Proj} \left( \sqrt{\frac{q}{\delta}}, \Lambda_b \right) \right\} = \begin{cases} \lambda_a^*, & \text{if } f_a(\lambda_a^*) \geq f_b(\lambda_b^*) \\ \lambda_b^*, & \text{otherwise} \end{cases} \quad (3)$$

Note that  $H$  is computed using the quadratic constraint such that any step taken satisfies the KL-divergence constraint. In the infeasible case when  $c^2 s - \delta > 0$  and  $c > 0$ , we apply a recovery method towards constraint satisfaction,  $x^* = \sqrt{2\delta / (b^T H^{-1} b)} H^{-1} b$ .

**Distributional Value Function and Safety Baseline.** Accurate estimates of the return and safety advantages  $A^{\pi_k}$  and  $A_C^{\pi_k}$  in (2) are essential to reduce the variance associated with the policy-gradient update step. These advantage functions can be estimated, e.g.,

**Algorithm 1:** Distributional Constrained Policy Optimization (DCPO)

---

**Input:** Initial policy  $\pi_0 \in \Pi_\theta$ , safety margin  $\varepsilon$ , reshaping coefficient  $\beta$

**for**  $k = 0, 1, 2, \dots$  **do**

- Sample a set of trajectories  $\mathcal{D} = \tau \sim \pi_k = \pi(\theta_k)$
- Form sample estimates  $\hat{g}, \hat{b}, \hat{H}, \hat{c}$  using  $\mathcal{D}$ , and reshaped  $\tilde{J}_C(\pi_k), \tilde{d}$
- if** approximate CPO (3) is feasible **then**
  - Solve dual problem with reshaped constraint  $\tilde{c} = \tilde{J}_C(\pi_k) - \tilde{d}$
  - Compute new policy  $\theta^*$  using  $x^* = -\frac{1}{\lambda^*} H^{-1}(g - v^* b)$
- else**
  - Compute recovery policy  $\theta^*$  using  $x^* = -\sqrt{2\delta/(b^T H^{-1} b)} H^{-1} b$
- end**
- Obtain  $\theta_{k+1}$  through backtracking linesearch to enforce constraints in (2) over  $\mathcal{D}$

**end**

---

see Schulman et al. (2015b), and approximated with NNs. To model uncertainty, we choose to represent the return and cost value functions (used in estimating  $A^{\pi_k}$  and  $A_C^{\pi_k}$ ) using NNs whose outputs are distributions parameterized by  $N$  equally spaced quantiles in a range  $[V_{\min}, V_{\max}]$  (Bellemare et al., 2017). We employ a surrogate loss for these NNs comprising the average negative log-likelihood of obtaining the costs or returns sampled under policy  $\pi_k$ . To simplify computation of the log-likelihood, we approximate the quantile distribution as Gaussian. For a value function with parameters  $\phi$  the loss is thus  $\mathcal{L}(\phi) = -\log(p(J_R(\pi_k)|\phi))$ , which is trainable using stochastic gradient descent.

**Cost Reshaping.** To reliably satisfy constraints, we reshape the cost  $J_C(\pi_k)$  and safety threshold  $d$  based on the agent’s confidence that constraints are satisfied given a cost distribution with parameters  $\phi_c$ . Specifically, we introduce a reshaping parameter  $\rho$ , whose value is set to  $\beta \mathcal{P}(J_C^{\pi_k} > d; \phi_c)$  when  $J_C^{\pi_k} > d$  (the constraints are violated), and to  $-\beta \mathcal{P}(J_C^{\pi_k} < d; \phi_c)$  when the constraints are satisfied, with  $\beta \geq 0$ . In other words,  $\rho$  takes the value of  $\beta$  when the agent is fully confident that constraints are violated under the previous policy  $\pi_k$ , and  $-\beta$  when it is fully confident the constraints are satisfied. We then reshape  $J_C^{\pi_k}$  and  $d$  by multiplying each with  $(1 + \text{clip}(\rho, -\varepsilon, \infty))$ . The new parameter  $\varepsilon$  controls risk aversion. This reshaping only affects policy updates when constraints are satisfied, effectively avoiding large steps in directions that may increase costs. Intuitively, maximizing rewards in a constrained environment usually leads to the agent to the edges of the feasible region. We thus mitigate risk by adjusting the feasibility of the trust region depending on the confidence of constraint satisfaction.

#### 4. Results and Discussion

We consider inventory management for a multi-echelon, multi-period supply chain: the goal is to maximize profits while satisfying constraints. We use *InvManagement-v0* in the `or-gym` library (Hubbs et al., 2020). At each period  $t$ , the environment is:

$$I_{t+1}^m = I_t^m + R_{t-L_m}^m - S_t^m, \quad \forall m \in \mathcal{M}, t \in \mathcal{T} \quad (6a)$$

$$T_{t+1}^m = T_t^m - R_{t-L_m}^m + R_t^m, \quad \forall m \in \mathcal{M}, t \in \mathcal{T} \quad (6b)$$

$$S_t^m = \begin{cases} R_t^{m-1}, & \text{if } m > 0 \\ \min(I_t^0 + R_{t-L_m}^0, D_t + B_{t-1}^0), & \text{otherwise} \end{cases} \quad \forall m \in \mathcal{M}, t \in \mathcal{T} \quad (6c)$$

$$U_t^m = R_t^{m-1} - S_t^m, \quad \forall m \in \mathcal{M}, t \in \mathcal{T} \quad (6d)$$

$$P_t^m = \alpha^m (p^m S_t^m - r^m R_t^m - k^m U_t^m - h^m I_{t+1}^m), \quad \forall m \in \mathcal{M}, t \in \mathcal{T} \quad (6e)$$

Eqs. 6a–6b are mass balances for the on-hand ( $I$ ) and pipeline ( $T$ ) inventories for each stage  $m$ , where  $R$  and  $L$  denote replenishment orders and lead times, respectively.

Eq. 6c gives the sales ( $S$ ) for the retailer ( $m = 0$ ) and the rest of the pipeline ( $m > 0$ ), given demand  $D$  and backlog  $B$ . Eq. 6d describes the unfulfilled re-order quantities ( $U$ ), and Eq. 6e calculates the profit  $P$  as the sales revenue minus procurement costs, unfulfilled demand penalties, and excess inventory penalties, all subject to a discount factor  $\alpha = 0.97$ . The unit sales price  $p$ , and penalties  $k$ ,  $h$  are all known.

There are production capacity and inventory constraints on the re-order quantities  $R_t^m$ . To enforce these, we define a cost function that sums the number of constraint violations  $C(s, a) = \sum_{m \in \mathcal{M}, t \in \mathcal{T}} C(I_t^{m+1}, R_t^m)$ , where  $C(I_t^{m+1}, R_t^m)$  is the number of constraints violated at node  $m$  at time  $t$ . We use this indicator cost function rather than a continuous one in order to limit the number of constraint violations, regardless of their degree.

**Training DCPO.** All of the algorithms use separate feed-forward NNs for the policy, return value distribution, and cost value distribution of size (64, 64) with tanh activations. Running DCPO ([github.com/jaimessabalimperial/jaisalab](https://github.com/jaimessabalimperial/jaisalab)) with  $N = 102$  quantiles allowed for convergence of the learned distributions, as shown in Figure 1. Intuitively, at earlier periods the agent is more uncertain about the cost/value of a state (epistemic uncertainty). As samples are observed, the distribution converges, and remaining uncertainty likely stems from stochasticity of the MDP (aleatoric uncertainty).

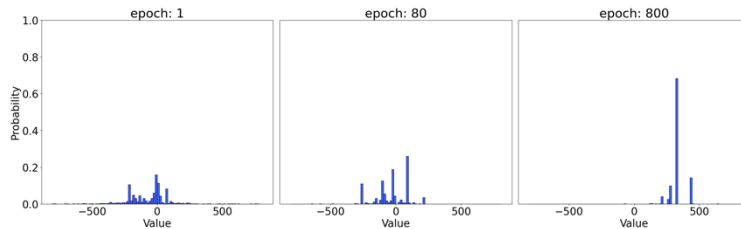


Figure 1: Learned quantiles ( $N=102$ ) for the return distribution of the initial state  $s_0$ .

**Importance of Cost Reshaping.** Figure 2 shows the costs and returns incurred during training for CPO, DCPO, and ‘ablation,’ or DCPO without cost reshaping. We observe that using parametric return and cost value distributions improves the initial stages of training, with the ‘ablation’ agent learning more quickly (see episodes 0–100) to both minimize costs and maximize returns compared to CPO. Moreover, the standard deviation of the returns across replications by is reduced by 40.0% and 76.6% for ‘ablation’ and DCPO, respectively, compared to CPO at the end of training.

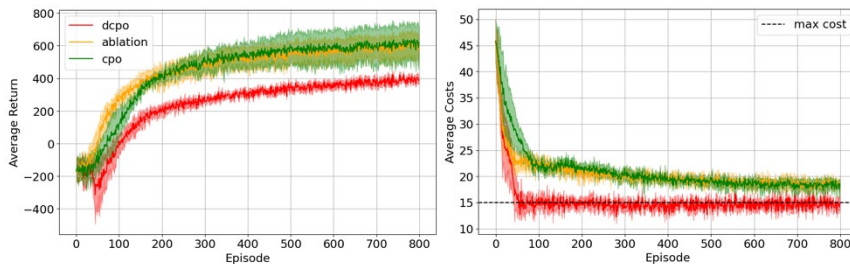


Figure 2: Returns (left) and costs (right) incurred by during training.

**Comparison to TRPO and Saute.** Figure 3 shows the costs and returns incurred during training for TRPO, Saute TRPO, CPO, and DCPO. TRPO is unconstrained and produces infeasible policies, more than doubling the permitted number of violations, but it attains



the largest returns. On the other hand, Saute TRPO incurs costs well below the limit by the end of training, with a clear trade-off in the low returns. DCPO seems to balance the intrinsic trade-off between returns and costs, with mean costs converging to the defined limit. Moreover, DCPO showed a reduction of 70.8% in the standard deviation of the returns relative to TRPO by the end of training. Finally, CPO converged to intermediate values for both costs and returns, potentially owing to approximation errors.

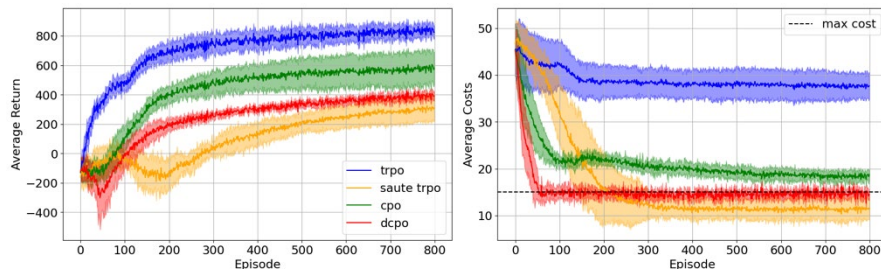


Figure 3: Returns (left) and costs (right) for algorithms over ten runs.

## 5. Conclusions

This paper introduces an approach for trust region optimization in CPO, by reshaping episodic discounted costs and the maximum allowed value using the agent’s confidence in constraint satisfaction, in turn given by an approximated cost distribution. The developed algorithm, which we call DCPO, balances the trade-off between maximizing returns and minimizing costs in a supply chain case study, such that constraints are satisfied (within one standard deviation), while reducing run-to-run variance.

## 6. Acknowledgments

The authors acknowledge support from the Engineering & Physical Sciences Research Council (EPSRC) grant EP/T001577/1 and an Imperial College Research Fellowship.

## References

- J. Achiam, D. Held, A. Tamar, P. Abbeel, 2017. Constrained policy optimization. In: International Conference on Machine Learning. PMLR, pp. 22–31.
- M. G. Bellemare, W. Dabney, R. Munos, 2017. A distributional perspective on reinforcement learning. In: International Conference on Machine Learning. PMLR, pp. 449–458.
- C. D. Hubbs, H. D. Perez, O. Sarwar, N. V. Sahinidis, I. E. Grossmann, J. M. Wassick, 2020. Or-gym: A reinforcement learning library for operations research problems. arXiv:2008.06319.
- P. Petsagkourakis, I. O. Sandoval, E. Bradford, F. Galvanin, D. Zhang, E. A. del Rio-Chanona, 2022. Chance constrained policy optimization for process control and optimization. *Journal of Process Control* 111, 35–45.
- J. Schulman, S. Levine, P. Abbeel, M. Jordan, P. Moritz, 2015a. Trust region policy optimization. In: International Conference on Machine Learning. PMLR, pp. 1889–1897.
- J. Schulman, P. Moritz, S. Levine, M. Jordan, P. Abbeel, 2015b. High-dimensional continuous control using generalized advantage estimation. arXiv:1506.02438.
- J. Shin, T. Badgwell, K. Liu, J.H. Lee, 2019. Reinforcement learning—overview of recent progress and implications for process control. *Computers & Chemical Engineering* 127, 282–294.
- A. Sootla, A. I. Cowen-Rivers, T. Jafferjee, Z. Wang, D. H. Mguni, J. Wang, H. Ammar, 2022. Sauté RL: Almost surely safe reinforcement learning using state augmentation. In: International Conference on Machine Learning. PMLR, pp. 20423–20443.

# Machine learning-based decomposition for complex supply chains

Niki Triantafyllou<sup>a</sup>, Stavros Papaiakovou<sup>b</sup>, Andrea Bernardi<sup>a</sup>, Matthew Lakelin<sup>c</sup>,  
Nilay Shah<sup>a</sup>, Antonis Kokossis<sup>b</sup>, Maria M. Papathanasiou<sup>a,\*</sup>

<sup>a</sup>*The Sargent Centre for Process Systems Engineering, Imperial College London,  
London, United Kingdom, SW72AZ*

<sup>b</sup>*School of Chemical Engineering, National Technical University of Athens, 9, Iroon  
Polytechniou St, Athens, Greece, 15780*

<sup>c</sup>*TrakCel Limited, 10/11 Raleigh Walk, Cardiff, CF10 4LN UK*

\**maria.papathanasiou11@imperial.ac.uk*

## Abstract

Personalised medicine products represent a novel category of therapeutics often characterised by bespoke manufacturing lines and dedicated distribution nodes. An example of such products is Chimeric Antigen Receptor (CAR) T-cells, whose manufacturing poses challenges to volumetric scale-up, leading to increased production and supply chain costs. From a modelling perspective, such networks lead to complex large-scale supply chain models that grow exponentially as the demand increases and more therapies are tracked simultaneously throughout the supply chain. In this work, we present a hybrid model that utilizes the potential of machine learning for strategic planning by forecasting optimal supply chain structures and Mixed Integer Linear Programming (MILP) for detailed scheduling. The proposed model is robust to uncertain demand patterns and can reduce the number of linear constraints and binary variables in the original MILP by more than 64.7%.

**Keywords:** supply chain optimisation, MILP, artificial neural networks, hybrid models, personalised medicine

## 1. Introduction

Personalised medicine products are often characterised by 1:1 business models that consider parallel manufacturing lines and dedicated distribution routes. Such models are also adopted by personalised cancer therapies that use the patient's own cells as the starting material (Papathanasiou *et al.*, 2020). Chimeric Antigen Receptor (CAR) T-cells are an indicative example of patient-specific therapies that owing to their promising clinical results have pioneered the approval of cell and gene therapy products by the regulatory authorities (Young *et al.*, 2022). The commercially available CAR T-cell products are offered at high list prices ranging between \$300,000 and \$475,000. These prices can be attributed to the high manufacturing, distribution and therapy administration costs (Spink *et al.*, 2018). In addition, there is a pressing need to coordinate the manufacturing and distribution lifecycle with the patient schedule. Depending on the clinical condition and location of each patient, in-time delivery is of utmost importance for the patient's prognosis. Future scenarios indicate that cell and gene therapy products could be treating up to 60,000 patients per year by 2030 (Quinn *et al.*, 2019). Planning and scheduling the manufacturing and distribution of such therapeutics is therefore a challenging task, currently relying on white-glove logistics, which might not always

ensure a responsive and resilient supply chain network. Process Systems Engineering tools can offer a systematic approach to assist decision-making.

Previous works on the optimization of CAR T-cell therapies via Mixed Integer Linear Programming (MILP) (Triantafyllou *et al.*, 2022a), heuristics-based decomposition algorithms (Triantafyllou *et al.*, 2022b) and metaheuristics-based algorithms (Karakostas *et al.*, 2020) highlight the patient-specific nature of these therapeutics, where the full-scale model may comprise complex sets of constraints and binary variables, translating into computationally expensive optimisation problems. In this space, methodologies that rely on Machine Learning (ML) can enhance supply chain planning in the area of Operations Research (Goettsch *et al.*, 2020; Abbasi *et al.*, 2020).

In this work, we present a hybrid model that describes the end-to-end supply chain of personalised therapeutics based on Mixed Integer Linear Programming (MILP) and Artificial Neural Networks (ANNs) (Figure 1). To reduce the computational complexity of the full-space MILP model, we harness the potential of ML for strategic planning in the supply chain by forecasting the optimal supply chain network structure.

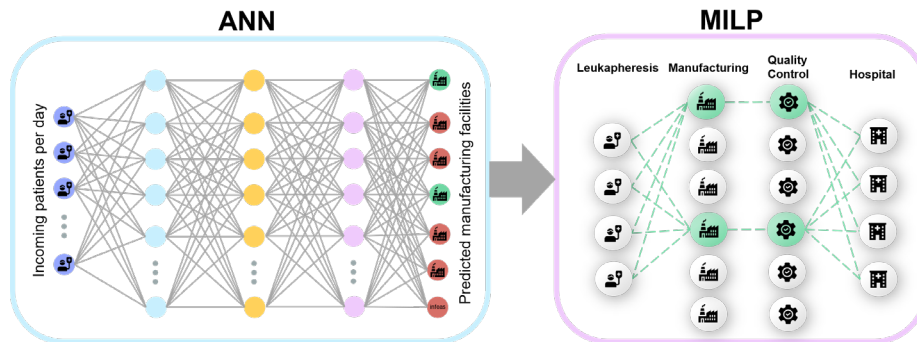


Figure 1. Proposed hybrid model for the CAR T-cell supply chain network with 4 nodes: leukapheresis centres, (b) manufacturing sites, (c) quality control sites, and (d) hospitals.

## 2. Materials & Methods

### 2.1. Mixed Integer Linear Programming (MILP) model

This study uses an in-house MILP model that describes the CAR T-cell therapy supply chain, and it is used for the identification of optimal supply chain network structures (Triantafyllou *et al.* 2022a). The supply chain superstructure consists of 4 nodes—leukapheresis centre, manufacturing site, Quality Control (QC), and hospital. CAR T-cell manufacturing starts at the leukapheresis centre, where T-cells are isolated from the patient’s bloodstream and are cryopreserved before further processing. The leukapheresis sample is then shipped to the manufacturing facility, where it is genetically engineered to express the tumour-associated antigen receptor. Lastly, the therapy undergoes in-house Quality Control (QC) and once product quality and safety are ensured, it is shipped back to the hospital for administration to the patient. We consider 4 leukapheresis sites and 4 hospitals in the UK and 6 manufacturing sites located in the UK, US and Europe. The manufacturing facilities  $m$  have a capacity of 4 ( $m_1$  and  $m_4$ ), 10 ( $m_3$  and  $m_6$ ), or 31 ( $m_2$  and  $m_5$ ) parallel lines, and a forward-looking scenario of 8 days of manufacturing is considered. The model considers demand uncertainty, manufacturing capacity limitations, patient-specificity, time and location constraints, whilst the objective is to

minimize the total supply chain cost. Finally, the total turnaround time is expressed as a non-monetary supply chain metric modelled as a constraint.

## 2.2. Hybrid model

The data-driven part of the model is responsible for strategic planning by forecasting the optimal supply chain network configuration based on the annualized demand. Specifically, the ANN model predicts the number and the location of the manufacturing facilities to optimally satisfy the demand. The mechanistic MILP model is then solved considering only the subset of manufacturing facilities chosen by the ANN, becoming a subproblem of the original planning and scheduling model. Therefore, the MILP accounts only for detailed scheduling in the supply chain, which entails optimal transport modes for the node-to-node connections, optimal allocation of patient samples in the manufacturing facilities and hospitals, and the optimal utilisation of the available parallel lines in the manufacturing sites with the scope of minimising the therapy cost and return time. The hybrid model can significantly reduce computational complexity by eliminating branches of the search space. The model's performance is assessed for different randomized annual demand scenarios (100, 200, 500, 1000, 1500 and 2000 patients).

### 2.2.1 ANN training

The ANN is trained based on sets of feasible and optimal solutions obtained from the MILP for 760 randomized demand profiles. The input features considered for the ANN are the total daily demands for a quarter of a year (90 days), assuming a recurrent demand profile per trimester. Based on three different probability distributions, namely uniform distribution, left triangular distribution, and right triangular distribution, 660 randomized demand profiles were generated for demands between 50 and 2000 patients per year (Figure 2). For each leukapheresis site  $c$ , we assume a daily capacity of 8 patients. Different probability distributions are used to account for the uncertainty in demand and lead to robust solutions as the demand distribution highly affects supply chain structures. To generate the training labels, the MILP model was solved for each demand scenario using CPLEX. All inputs were normalized.

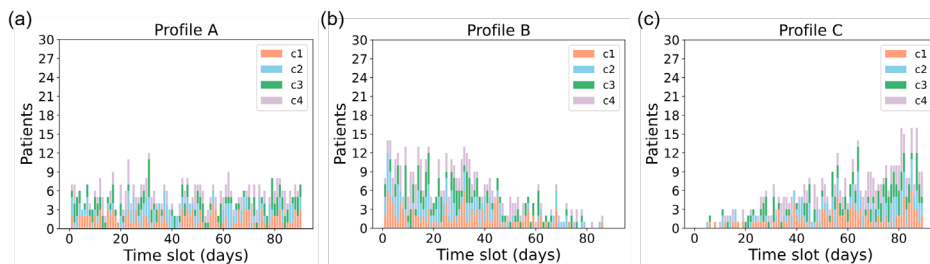


Figure 2. Randomized demand profiles for the 2000 patients' case for 3 different probability distributions: (a) uniform distribution (b) left triangular distribution (c) right triangular distribution.

A feed-forward ANN model is built in PyTorch for multi-label classification, where the labels are the 6 manufacturing facilities and the possibility of infeasible solutions due to limited facility capacity as seen in Figure 1. Multi-label classification is used so that more than one facility can be selected for each instance. However, because the original dataset was imbalanced, with classes  $m_6$  and infeasible being underrepresented, extra 100 randomized demand scenarios were generated and equally distributed for the 350, 500 and 2000 patients per year cases to balance the data and capture more instances where the

underrepresented classes are the solution. It should be noted that oversampling techniques for imbalanced classification datasets such as SMOTE do not apply to multi-label datasets. The data set is split into 80% for training and 20% for testing. The topology of the ANN was tuned using Grid Search and parameter tuning. To configure the feed-forward ANN, 90 input features, 3 hidden layers, 256 neurons at each hidden layer, and 7 output layers are used. Because multi-label classification is based on binary classification, the sigmoid activation function is used at each output layer node to predict a probability for each class in the output. The threshold that separates positive and negative classes is set at 0.5. Finally, the ReLU function is used in the hidden layers and the model is trained using the binary cross-entropy loss function and the Adam solver.

### 3. Results and discussion

The overall performance of the multi-label classification ANN model is measured by *accuracy* and additional extracted metrics from the *confusion matrix*, i.e. *precision*, *recall*, and *F-score* for each class and *micro*, *macro*, and *weighted* average of all classes (Table 2). Accuracy provides the number of correct predictions to the total number of labels (predicted and true). Precision and recall quantify the proportion of correctly predicted labels to the total number of true labels and the total number of predicted labels, respectively. On the other hand, the F1-score leads to a harmonic balance between precision and recall as it considers both. The ordinary 2-dimensional confusion matrix used in multi-class classification is undefined when it comes to multi-label classification, where each instance can be labelled with more than one class. To overcome this, the multi-label confusion matrix with one extra row (No True Label-NTL) and one extra column (No Predicted Label-NPL) is used (Heydarian *et al.*, 2022).

Table 1. Overall performance metrics for the ANN.

Class	Precision	Recall	F1-score	weight
m <sub>1</sub>	0.84	0.76	0.80	50
m <sub>2</sub>	0.92	0.88	0.90	67
m <sub>3</sub>	0.81	0.81	0.81	47
m <sub>4</sub>	0.85	0.73	0.79	15
m <sub>5</sub>	0.88	0.88	0.88	25
m <sub>6</sub>	0.78	0.90	0.84	20
infeasible	1	0.93	0.96	29
micro avg	0.84	0.84	0.84	253
macro avg	0.87	0.84	0.85	253
weighted avg	0.87	0.84	0.86	253

The resulting ANN has a train set overall accuracy of 98.5% and a test set overall accuracy of 95%. The rest of the performance metrics for each class are presented in Table 1. The ANN performs relatively well, as the average precision, recall and F1-score are above 84%. It should be noted that the ANN model is trained with three different demand distributions to better cope with demand uncertainty. This can lead to different supply chain structures for the same yearly demands. Hence, there is a possible trade-off between model accuracy and robustness to demand uncertainty.

It is observed that class m<sub>4</sub> has the lowest F1-score (0.79). The precision and recall for this class are 0.85 and 0.73, respectively, indicating more false negatives (*FN*) for this class than false positives (*FP*). By looking at the confusion matrix in Figure 3, it can be seen that the model confuses facility m<sub>4</sub> (4 parallel lines) with facility m<sub>3</sub> (10 parallel

lines) and with the no-label prediction. Specifically, while the model predicts correctly in 73% of the instances that  $m_4$  belongs to the optimal supply chain configuration, in 13% of the instances it confuses facilities  $m_4$  and  $m_3$  and in 13% of the instances, it fails to predict  $m_4$  as an extra required facility. These mispredictions are sensible as in some cases the model overpredicts by choosing a slightly bigger facility to facilitate the demand and in other cases, it underpredicts by failing to add an extra facility. In the former case, the hybrid model leads to local optimum solutions, whereas in the latter case the model leads to infeasible solutions.

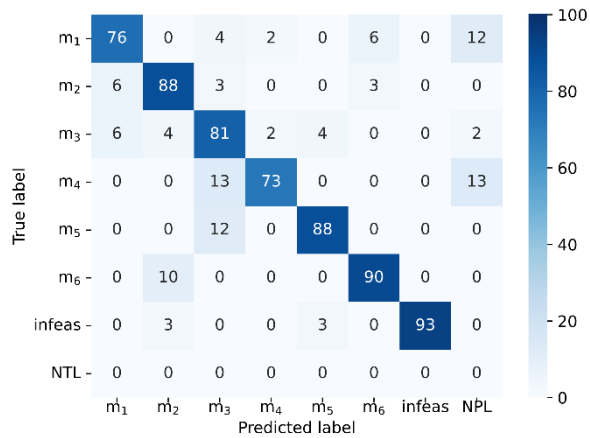


Figure 3. Normalised multi-label confusion matrix with one extra row for No True Labels (NTL) and one extra column for No Predicted Labels (NPL).

Similar behaviour is observed with facility  $m_1$ , where the recall is 0.76. Figure 3 shows that in 12% of the instances the model fails to predict that  $m_1$  is a required extra facility to satisfy the demand, thus leading to infeasible solutions. It is important to note, that the classifier is very accurate in terms of infeasibilities. The precision for infeasible solutions is 100%, which means that the classifier never predicts feasible solutions as infeasible. In rare occasions (3% of the cases) the hybrid model assumes infeasible solutions as feasible.

In those cases, the MILP part of the model will determine the infeasibility of the solutions.

In Figure 4, the performance of the hybrid model is showcased in terms of model complexity by

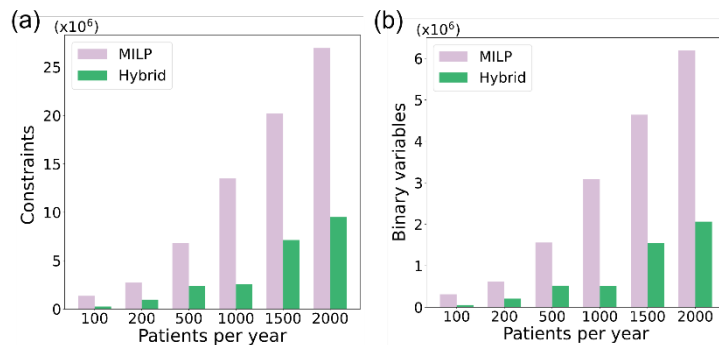


Figure 4. Comparison of the original MILP model with the hybrid model for an increasing number of annual demands in (a) the number of constraints, and (b) the number of binary variables.

evaluating the number of constraints and binary variables for both the original mechanistic model and the proposed hybrid model. The hybrid model can reliably reduce the computational burden by decreasing the total number of linear constraints by 65%–81% and the total number of binary variables by 67%–83% as shown in Figure 4. In cases of 500 and 1000 patients per year, the hybrid model is described by almost the same number of constraints and binary variables. This is attributed to the fact that in the 1000 patients per year case, the classifier predicted that only one large facility with 31 parallel lines ( $m_2$ ) can satisfy the demand, and hence the feasible search space for the MILP model was significantly reduced.

#### 4. Conclusions

In this study, a framework that integrates a multi-label classification feed-forward neural network with a MILP model in order to decompose large-scale supply chain problems is presented. The unique CAR T-cell therapy supply chain model is used as the case study. The data-driven part of the hybrid model accounts for strategic planning and the mechanistic part is responsible for detailed scheduling. The framework proposes a novel method to account for demand uncertainty using machine learning techniques, resulting in robust supply chain structures. The overall accuracy of the classifier is 95%, with just a few cases leading to suboptimal or infeasible solutions. The computational analysis exhibited significant improvements over the original MILP model. A key novelty of this work is the ability of the hybrid model to solve larger problem instances, for which the MILP problem did not manage to converge to a feasible solution. The efficiency of the hybrid model creates new possibilities to explore, such as the use of convolutional neural networks for improved classification accuracy.

#### Acknowledgements

Funding from the UK Engineering & Physical Sciences Research Council (EPSRC) for the Future Targeted Healthcare Manufacturing Hub hosted at University College London with UK university partners is gratefully acknowledged (Grant Reference: EP/P006485/1). Financial and in-kind support from the consortium of industrial users and sector organisations is also acknowledged.

#### References

- M. M. Papathanasiou, C. Stamatis, M. Lakelin, S. Farid, N. Titchener-Hooker, N. Shah, 2020, Autologous CAR T-cell therapies supply chain: challenges and opportunities?, *Cancer Gene Therapy*, 27, 799–809.
- C. M. Young, C. Quinn, M. R. Trusheim, 2022, Durable cell and gene therapy potential patient and financial impact: US projections of product approvals, patients treated, and product revenues, *Drug Discovery Today*, 27(1), 17–30.
- K. Spink, A. Steinsapir, 2018, The long road to affordability: a cost of goods analysis for an autologous CAR-T process, *Cell & Gene Therapy Insights*, 4(11), 1105–1116.
- C. Quinn, C. Young, J. Thomas, M. Trusheim, 2019, Estimating the Clinical Pipeline of Cell and Gene Therapies and Their Potential Economic Impact on the US Healthcare System, *Value in Health*, 22(6), 621–626.
- N. Triantafyllou, A. Bernardi, M. Lakelin, N. Shah, M. M. Papathanasiou, 2022, A digital platform for the design of patient-centric supply chains, *Scientific Reports*, 12, 17365.
- N. Triantafyllou, A. Bernardi, M. Lakelin, N. Shah, M. M. Papathanasiou, 2022, A bi-level decomposition approach for CAR-T cell therapies supply chain optimisation, *Computer Aided Chemical Engineering*, 49, 2197–2202.
- P. Karakostas, N. Panoskaltsis, A. Mantalaris, and M. C. Georgiadis, 2020 Optimization of CAR T-cell therapies supply chains, *Computers & Chemical Engineering*, 139, 106913.
- D. Goettsch, K. K. Castillo-Villar, M. Aranguren, 2020, Machine-Learning Methods to Select Potential Depot Locations for the Supply Chain of Biomass Co-Firing, *Energies*, 13, 6554.
- B. Abbasi, T. Babaei, Z. Hosseinifard, K. Smith-Miles, M. Dehghani, 2020, Predicting solutions of large-scale optimization problems via machine learning: A case study in blood supply chain management, *Computers & Operations Research*, 119, 104941.
- M. Heydarian, T. E. Doyle, R. Samavi, 2022, MLCM: Multi-Label Confusion Matrix, *IEEE Access*, 10, 19083–19095

# Optimal energy management in greenhouses using distributed hybrid DRL-MPC framework

Benjamin Decardi-Nelson,<sup>a</sup> Fengqi You<sup>a</sup>

<sup>a</sup>*Cornell University, Ithaca, New York, 14853, USA*

## Abstract

In this work, we propose a framework for optimal energy management of greenhouses, which considers the presence of both discrete-valued and continuous-valued actuators in the controller design. To reduce the computational burden because of having to solve mixed-integer nonlinear optimal control problems online, we propose to decompose the system model based on the type of actuators. The underlying subproblems are solved using a hybrid deep reinforcement learning (DRL)-model predictive control (MPC) framework in a distributed manner. We use a greenhouse case study to demonstrate the effectiveness of the proposed control framework.

**Keywords:** MPC, greenhouse, energy management, deep reinforcement learning.

## 1. Introduction

Greenhouses are intensified crop production systems that have a potential to improve the resilience of the food system in a sustainable way (Chen et al., 2022). One of the main drawbacks of greenhouses is the enormous energy consumption of their microclimate control activities such as lighting, heating, ventilation, and air conditioning activities (Vadiee & Martin, 2012). Because the rule-based control methods traditionally used in greenhouses are unable to optimally manage the energy usage and ensure constraint satisfaction, optimization-based model predictive control (MPC) has been proposed as a better alternative (Hu & You, 2022). Unfortunately, the effective operation of greenhouses relies on equipment with discrete-valued actuation, like ventilation window opening and lights, which complicate the design and implementation of MPC (Chen et al., 2021). This is because the underlying centralized MPC optimization problem becomes a parametric mixed-integer nonlinear (possibly non-convex) optimization problem (MINLP), which may be challenging to solve repetitively online. To facilitate the adoption of MPC in greenhouse operation, efficient control methods are needed. The use of data-driven techniques to address robustness and computational issues in MPC is on the rise (Chen and You, 2022). In particular, the use of machine learning to assist in solving mixed-integer MPC (MIMPC) problems has received considerable attention (Löhr et al., 2020; Cauligi et al., 2021). However, the approaches focus on only linear problems. Moreover, the authors assume that solutions of the original MIMPC problem can be readily obtained. Often, these solutions for large scale MIMPC optimization problems are difficult to obtain. In this work, we propose a novel distributed hybrid deep reinforcement learning (DRL)-MPC framework to address the greenhouse optimal control problem. The proposed control framework exploits the attractive features of both DRL and MPC, that is, fast discrete-valued sequential decision making, and optimal control with constraint satisfaction (Ning et al., 2021), respectively. This is achieved by first decomposing the greenhouse optimal control problem into two sub-problems based on the type of actuation, namely discrete-valued or continuous-valued. Thereafter, we train the DRL to make the discrete-valued decisions while the MPC handles the



continuous-valued decisions. Similar to the approaches used in distributed MPC (DMPC), the two controllers cooperatively determine the optimal input sequence to be sent to the greenhouse system at each time-step in a receding horizon manner. We demonstrate the effectiveness of the proposed framework using a greenhouse case study.

## 2. Greenhouse optimal control problem

In this section, we present the details of the greenhouse model and formulate the associated mixed-integer model predictive control problem.

The greenhouse model used in this work is a discrete-time version of the model used in the study by Van Henten (2003). The model is a nonlinear system of ordinary differential equations (ODEs) described by the equation

$$x_{k+1} = f(x_k, u_k, v_k) \quad (1)$$

where  $x_k$  denotes the states of the system at time-step  $k$ ,  $u_k$  denotes the continuous-valued control actions,  $v_k$  represents the discrete-valued control actions and  $f$  represents the nonlinear ODEs. The states of the system are the crop dry weight, carbon dioxide concentration, air temperature and humidity in the greenhouse. The continuous-valued control actions are the carbon dioxide enrichment, heating and cooling systems, and the (de)humidification system. The discrete-valued control actions on the other hand are the supplemental lighting and vent opening, which can either be turned on/open or off/closed. Throughout this article, the states and the inputs are constrained to be in the sets

$$x_k \in X, u_k \in U, v_k \in V = \{0,1\} \quad (2)$$

The objective of the control system is to minimize the energy consumption while ensuring a higher crop growth rate. This control problem can be formulated as a mixed-integer MPC (MIMPC) described in Equation (3).

$$\min_{u_k, v_k} \sum_{k=0}^{N-1} L(x_k, u_k, v_k) \quad (3a)$$

$$s. t. \quad x_{k+1} = f(x_k, u_k, v_k) \quad (3b)$$

$$x_0 = x(t) \quad (3c)$$

$$x_k \in X, u_k \in U, v_k \in V \quad (3d)$$

In (3), Equation (3a) is the objective function to be minimized, Equation (3b) is the nonlinear system model described in (1), Equation (3c) denotes the initial state constraint and Equation (3d) represents the state and input constraints. At each sampling time  $k$ , the MIMPC optimization problem in (3) is solved to obtain a sequence of optimal values of the continuous-valued and discrete-valued control actions. As consistent with an MPC implementation, only the first control actions ( $u_0^*, v_0^*$ ) of the optimal sequence are applied to the system in a receding horizon manner (Mayne et al., 2000). The feasibility issue is guaranteed via soft constraints (Lu et al., 2020). As mentioned earlier, the presence of both continuous and discrete decisions in the MIMPC optimization problem (3) results in having to solve an MINLP online several times. This can be computationally difficult to solve (Shang et al., 2019). In this article, we proposed a computationally efficient hybrid distributed framework that takes advantages of the improvements in deep reinforcement learning (DRL) and standard model predictive control.

## 3. Optimal control using distributed hybrid algorithm

The presence of both discrete-valued and continuous-valued actions in the MPC formulation makes it computationally demanding to implement as an online control algorithm. In this section, we present a novel distributed hybrid algorithm that considers the strength of both model predictive control and deep reinforcement learning. The central idea we employ in this framework is to use a deep Q network (DQN) policy obtained

from DRL to make the discrete-valued decisions. By fixing the discrete decisions in the original MIMPC optimization problem in (3), a simpler nonlinear problem (NLP) can easily be solved online. To achieve this, we reformulate the MIMPC optimization problem. We begin this section by first describing the reformulation. Thereafter, we present the details of the DRL. Finally, we describe and discuss the implementation details of the proposed framework. In general, it is difficult to solve mixed-integer optimization problems. This is because the complexity of the optimization problem increases exponentially in the discrete-valued decision variables. While there have been improvements in algorithms for solving such problems, using these algorithms in an online setting may be limiting. To address this, we propose to reformulate the optimization problem such that a deep reinforcement learning algorithm learns to make the discrete-valued decisions. These decisions are then fixed in the MPC optimization problem, thus resulting in having to solve online an NLP.

Let  $\tilde{v}_k$  approximate the discrete-valued decision variables in the optimal control problem described in (3) at each time-step  $k$ . If the discrete-valued variables in (3) are fixed, then the resulting optimization problem becomes an NLP which can be solved efficiently using standard NLP solvers such as IPOPT. Notice that  $N - 1$  discrete decision variables need to be determined in the optimal control problem. In contrast to the approach proposed by Löhner et al. (2020) where only the first discrete variables are fixed with the approximated value, we take a distributed approach. The reformulated optimization problem is described in (4).

$$\min_{u_k, v_k} \sum_{k=0}^{N-1} L(x_k, u_k, v_k) \quad (4a)$$

$$s. t. \quad x_{k+1} = f(x_k, u_k, v_k) \quad (4b)$$

$$x_0 = x(t) \quad (4c)$$

$$u_k \in U, v_k \in V \quad (4d)$$

$$v_k = \tilde{v}_k \quad (4e)$$

The details of the equations in (4) remain the same as that of (3). Equation (4e) ensures that the discrete valued decisions remain fixed with the approximate values from the DQN policy. To ensure feasibility of the optimization problem, the state constraint is relaxed. The objective of the DRL neural network policy is to make one-step discrete-valued decisions  $\tilde{v}_k$  given the state information  $x_k$  (Ajagekar et al., 2023). Let us denote the neural network policy as  $P$ , then the policy is described by the mapping  $P: x_k \in R^{n_x} \rightarrow \tilde{v}_k \in R^{n_v}$  where  $n_x$  and  $n_v$  denote the dimension of the state and discrete-valued action spaces respectively. Because the state variables are continuous-valued and the actions are discrete-valued, we use DQN reinforcement learning algorithm (Minh et al., 2015).

The implementation detail of the control algorithm is summarized in Figure 1. At each time-step, an estimate of the discrete actions is used in the reformulated MPC optimization problem described in (4) to obtain the state predictions and the optimal continuous-valued actions. The state predictions are sent to the DQN policy to determine an improved estimate of the discrete-valued decisions. This procedure is repeated until the stopping criteria is achieved. Notice that the proposed framework is no different from the DMPC framework, where information is repeatedly exchanged between two or more MPC-NLPs (Christofides et al., 2013). The difference here is that the communication is between an MPC-NLP and a DQN policy. Again, like DMPC algorithms, an initial guess of the subsystem solutions is required at the start of the algorithm, and a termination criterion needs to be specified. The stopping criteria employed in this work are either the

sum of the absolute difference between any two successive discrete actions from the DQN policy is less than 1 or the maximum number of iterations is reached.

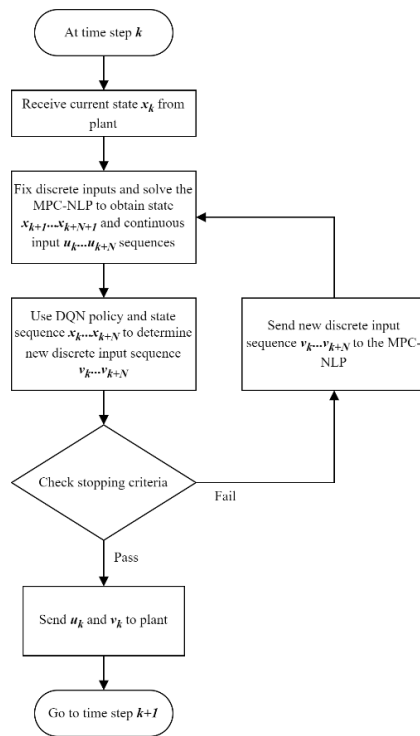


Figure 1. Implementation details of the distributed hybrid algorithm.

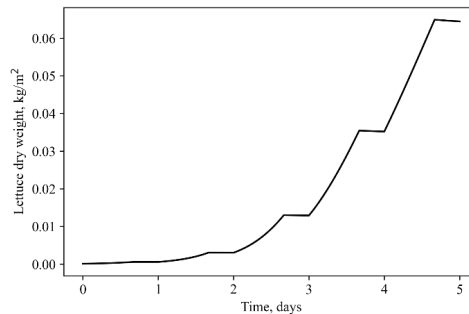


Figure 2. Crop growth profile after 5 days.

#### 4. Simulation results

In this section, we demonstrate the efficacy of the proposed hybrid algorithm using a case study. We begin this section by first presenting the simulation settings. Thereafter, we present the results of the proposed algorithm. In this work, we assume that the external weather conditions are known and constant. The values for the external climate conditions include external temperature of 20 °C, external humidity of 400 ppm, the solar radiation at 100 W/m<sup>2</sup> and off at night (Yang et al., 2022). The control objective is to track the setpoints for the CO<sub>2</sub> concentration, relative humidity, and greenhouse air temperature

which are determined as 1000 ppm, 70 % and 24 °C respectively. The photoperiod is set at 16 hours on period and 8 hours off period. The sampling time is fixed at 10 minutes and the MPC horizon is fixed at 12 which represents a two-hour prediction horizon. All computations were conducted on laptop with an Intel 2.2 GHz quadcore processor, 16 GB random access memory (RAM) and 2 GB of video RAM. The profiles of the greenhouse system under the action of the proposed control framework are presented in Figures 2 – 4. As can be observed in the profiles, the proposed control framework is able to ensure a steady crop growth rate. From Figure 2, the controller turns on the supplemental lighting in a similar fashion as the photoperiod to supplement the external light intensity entering the greenhouse. The controller optimally turns the supplemental lights off when it is not required, thus reducing energy consumption. Similarly, the controller uses the vent opening as an economizer to reduce the cooling requirements in the greenhouse. It is worth mentioning that the entire 5-day simulation took less than 5 minutes to complete.

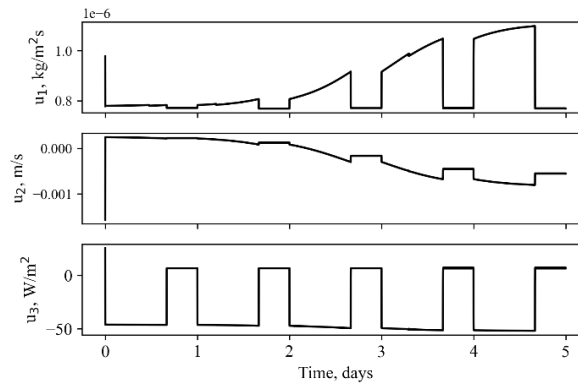


Figure 3. Control profiles for the continuous-valued actions.  $u_1$ ,  $u_2$  and  $u_3$  denote the supplemental CO<sub>2</sub> supply rate, (de)humidification rate and the cooling/heating rates respectively.

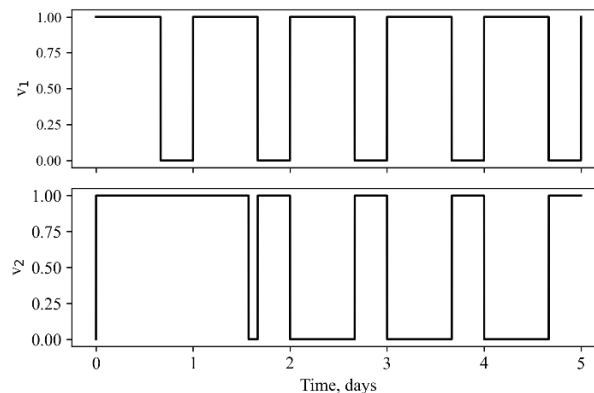


Figure 4. Control profiles for the discrete-valued actions.  $v_1$  and  $v_2$  denote the on/off state of the supplemental lighting and the open/close state of the vent opening respectively.

## 5. Concluding remarks

In this article, we have presented a novel MPC algorithm to optimally manage the energy resource in a greenhouse. A DRL neural network policy is trained to determine

approximations of the discrete decisions, which is then fixed in the original MPC optimization problem. This reduces the online computational burden in the control system implementation. A simulation case study demonstrates the efficacy of the proposed approach to ensure steady crop growth in the greenhouse.

## References

- A. Ajagekar, N.S. Mattson, et al., 2023, Energy-efficient AI-based Control of Semi-closed Greenhouses Leveraging Robust Optimization in Deep Reinforcement Learning. *Advances in Applied Energy*, 9, 100119.
- A. Ajagekar, F. You, 2023, Deep Reinforcement Learning Based Unit Commitment Scheduling under Load and Wind Power Uncertainty. *IEEE Transactions on Sustainable Energy*, DOI: 10.1109/TSTE.2022.3226106.
- A. Cauligi, P. Culbertson, E. Schmerling, et al., 2021, CoCo: Online mixed-integer control via supervised learning, *IEEE Robotics and Automation Letters*, 7, 2, 1447-1454.
- W.-H. Chen, N.S. Mattson, F. You, 2022, Intelligent control and energy optimization in controlled environment agriculture via nonlinear model predictive control of semi-closed greenhouse. *Applied Energy*, 320, 119334.
- W.-H. Chen, F. You, 2022, Semiclosed Greenhouse Climate Control Under Uncertainty via Machine Learning and Data-Driven Robust Model Predictive Control. *Ieee Transactions on Control Systems Technology*, 30, 1186-1197.
- W.-H. Chen, F. You, 2021, Smart greenhouse control under harsh climate conditions based on data-driven robust model predictive control with principal component analysis and kernel density estimation. *Journal of Process Control*, 107, 103-113.
- W.-H. Chen, F. You, 2022, Sustainable building climate control with renewable energy sources using nonlinear model predictive control. *Renewable & Sustainable Energy Reviews*, 168, 112830.
- P. D. Christofides, R. Scattolini, D. M. de la Pena, J. Liu, 2013, Distributed model predictive control: A tutorial review and future research directions, *Computers & Chemical Engineering*, 51, 21-41.
- Y. Chu, F. You, 2015, Model-based integration of control and operations: Overview, challenges, advances, and opportunities. *Computers & Chemical Engineering*, 83, 2-20.
- G. Hu and F. You, 2022, Renewable energy-powered semi-closed greenhouse for sustainable crop production using model predictive control and machine learning for Energy Management, *Renewable and Sustainable Energy Reviews*, 168, 112790.
- Y. Löhr, M. Klaučo, M. Fikar, M. Mönnigmann, 2020, Machine learning assisted solutions of mixed integer MPC on embedded platforms, *IFAC-PapersOnLine*, 53, 2, 5195-5200.
- S. Lu, J.H. Lee, et al., 2020, Soft-constrained model predictive control based on data-driven distributionally robust optimization. *AIChE Journal*, 66, e16546.
- D. Q. Mayne, J. B. Rawlings, C. V. Rao, et al., 2000, Constrained model predictive control: Stability and optimality, *Automatica*, 36, 6, 789-814.
- V. Mnih, K. Kavukcuoglu, D. Silver, et al., 2015, Human-level control through deep reinforcement learning, *Nature*, 518, 7540, 529-533.
- C. Ning, F. You, 2021, Online learning based risk-averse stochastic MPC of constrained linear uncertain systems. *Automatica*, 125, 109402.
- C. Shang, F. You, 2019, A data-driven robust optimization approach to scenario-based stochastic model predictive control. *Journal of Process Control*, 75, 24-39.
- A. Vadiee, V. Martin, 2012, Energy management in horticultural applications through the closed greenhouse concept, *Renewable and Sustainable Energy Reviews*, 16, 7, 5087-5100
- E. J. Van Henten, 2003, Sensitivity analysis of an optimal control problem in greenhouse climate management, *Biosystems Engineering*, 85, 3, 355-364
- S. Yang, H. Gao, et al., 2022, Model predictive control in phase-change-material-wallboard-enhanced building energy management considering electricity price dynamics. *Applied Energy*, 326, 120023.

# Multi-Agent Reinforcement Learning and RL-Based Adaptive PID Control of Crystallization Processes

Qingbo Meng <sup>a</sup>, Paul Danny Anandan <sup>a</sup>, Chris D. Rielly <sup>a</sup>, Brahim Benyahia <sup>a\*</sup>

<sup>a</sup>*Department of Chemical Engineering, Loughborough University, Leicestershire, UK  
B.Benyahia@lboro.ac.uk*

## Abstract

In this work, two model-based reinforcement learning (RL) control strategies are investigated namely a multi-agent RL and RL-based adaptive PID control. An off-policy deep deterministic policy gradient (DDPG) was adopted in both cases to achieve optimal trajectory tracking control of crystallization processes. Two case studies were considered to validate the new control strategies. The first is the cooling and antisolvent crystallization of aspirin in a mixture of ethanol and water, and the second is a 2-dimensional (2D) cooling crystallization of potassium dihydrogen phosphate in water. The optimal reference trajectories were identified using model-based dynamic optimization approaches which aim at maximizing the mean crystal size/minimizing the aspect ratio. Transfer Learning (TL) techniques and various reward-shaping strategies were also investigated to enhance the learning capabilities of the RL control. The results indicate that multi-agent RL saves massive training costs, compared to single agent, and RL-based adaptive PID exhibits excellent performance against state-of-the-art MPC.

**Keywords:** Multi-agent Reinforcement Learning; Transfer Learning; Adaptive PID; Semi-Batch Antisolvent and Cooling Crystallization; 2-Dimensional Crystallization Model.

## 1. Introduction and background

As one of the oldest purification technologies, crystallization has widespread applications in the pharmaceutical industry to separate and purify active pharmaceutical ingredients (API). One of the most critical features in pharmaceutical manufacturing is the stringent regulation associated with the product quality control where small deviations from the target quality profiles may significantly impact drug safety and efficacy and may lead to failures during clinical trials or even product recalls (Mascia et al., 2013). In recent years, thanks to the development of process analytical technologies (PAT), real-time measurement, monitoring and control of the crystal size and shape become achievable during the crystallization which opened new opportunities to implement advanced control strategies such as model predictive control and Artificial Intelligence based techniques such as Reinforcement Learning (Benyahia et al., 2021).

The most common control objective in crystallization processes is to maximize the mean crystal size due to its dramatic impact on the drug product's critical quality attributes, such as dissolution rate and efficacy, and on the downstream operations, such as filtration and drying (Fysikopoulos et al., 2019). Supersaturation is the fundamental driving force

in crystallization, which is typically generated by either cooling, evaporation or antisolvent addition (Liu and Benyahia, 2022). Over the last two decades, model-free and model-based techniques have been extensively investigated to control and optimize crystallization processes (Lakerveld and Benyahia, 2020). However, there is still an increasing demand for more robust techniques for real time monitoring, optimization, and control to achieve precision and simultaneous control of several critical quality attributes such as purity, crystal size and shape distribution.

Over the last few years, machine learning (ML) applications have witnessed rapid growth in many research and industrial areas. Reinforcement learning (RL) has been successfully implemented in dynamic systems particularly in robotics and automotive industries because of its ability to solve a wide range of sequential decision-making problems (Canese et al., 2021). Most of the MARL applications were developed for unmanned aerial vehicles (UAV), optimal path planning (Qie et al., 2019); resource allocation and energy efficiency (Cui et al., 2020). These recent successful applications have motivated further exploration of RL capabilities as a new control strategy for crystallization (Benyahia et al., 2021).

In this study, two crystallization case studies are presented where different types of RL model-based control methods are implemented. The first case study presents an antisolvent and cooling crystallization of aspirin (acetylsalicylic acid, ASA) in a mixture of ethanol (solvent) and water (antisolvent) with a mathematical model based on population balance equations (PBE) is used. The second case study presents a 2-Dimensional (2D) cooling crystallization process of potassium dihydrogen-phosphate (KDP) in water modelled using a 2-dimensional (2D) PBE to enable crystal shape prediction and control. Firstly, a model-based optimization approach is formulated to determine the reference trajectories that maximize the critical quality attributes. A MARL and RL-PID are developed as control approaches for both case studies respectively. Both RL agents are designed based on an off-policy deep deterministic policy gradient (DDPG) then trained to track several reference trajectories, which include process temperature, supersaturation, and particle size or/and shape. The performance of the proposed RL-based control strategies is compared against a single agent RL and model predictive control (MPC).

## 2. Problem Formulation

### 2.1 Process Modelling

The ultimate objective of this work is to exploit multi-agent RL and RL-based techniques to build responsive and efficient control strategies for batch and semi-batch crystallization processes. The mathematical model associated with the first case study was obtained from the literature (Nagy et al., 2008). The growth and nucleation kinetic parameters obtained from Barik et al., 2020. The 2-dimensional mathematical model associated with the second case study was obtained from Fysikopoulos et al., 2019. Both mathematical models were experimentally validated in the original studies. A few assumptions were considered to build the models which includes: (i) the vessel is assumed to be well-mixed; (ii) crystal breakage and agglomeration of fine crystals are negligible; (iii) crystals growth is size-independent. (iv) crystal growth and dissolution do not affect the total volume. For

the sake of brevity, the mathematical formulation and PBE for both models are not presented here.

### 2.2 Process optimization

The optimization strategy is required to ensure the best design options and optimal operational performance of the two crystallization simulation processes. It will also provide the reference trajectory profiles for control and operation purposes when applying the RL strategies. In this study, the model-based open-loop dynamic optimization method is implemented to maximize the mean crystal size/minimize the aspect ratio. The decision variables used in the first case study are the jacket temperature ( $T_j$ ) and the antisolvent flowrates ( $F_A$ ) whereas the process temperature is used in the second case study.

### 2.3 Reinforcement Learning

The development of effective two-agent RL and RL-PID are the primary objectives of this study. These objectives translate into tracking the reference trajectories that maximize crystal mean size, in case study one, and minimize the crystal aspect ratio in case study two. The reference profiles of the process temperature, supersaturation, particle size, and aspect ratio were computed using a model-based open-loop optimal control strategy. The RL agents in both case studies were developed using MATLAB Simulink and trained to reach the best control performance corresponding to the maximum rewards. Here, an off-policy actor-critic agent based on deep deterministic policy gradient (DDPG) was used to improve effective tracking control through computing the optimal control policies that maximize the total rewards (Benyahia et al., 2021).

#### 2.3.1 Two-agent reinforcement learning

Figure 1 shows the structure of the two-agent RL. Jacket temperature and antisolvent flowrate were used separately as the manipulated variables for each RL agent to control the crystallization process.

The RL objective is to maximize the rewards/minimize penalties. RL agent 1 receives one observation (process temperature) and takes one action (cooling rate) whereas RL agent 2 observes two controlled outputs, which are supersaturation and particle size, and takes one action (antisolvent flowrate). The two actions work together to maximize the instantaneous and cumulative rewards of the agents to achieve effective training.

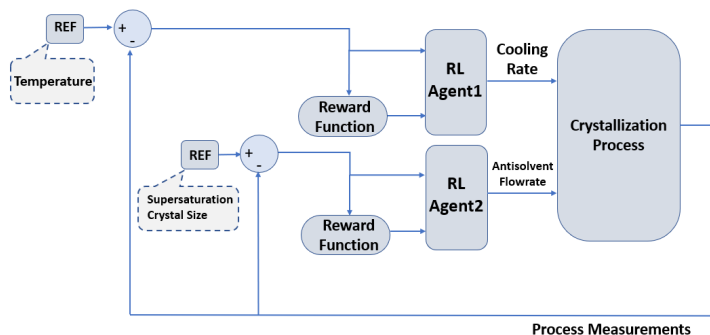


Figure 1. Block diagram of closed-loop two-agent RL.



### 2.3.2 RL-PID

The RL-PID is designed to combine an RL agent and a conventional PID control as a cascade structure to achieve effective control performance. Here, the cooling rate is used as the controlled variable and the tuning parameters of the PID,  $k_p$ ,  $k_i$  and  $k_d$ , are the agent 'actions'. The reference tracking trajectories for the RL agent are the temperature, supersaturation, crystal size and aspect ratio.

### 2.4 Reward Shaping

The reward function used in both case studies is described in figure 2 below. It is defined based on penalties only. Therefore, zero is the maximum attainable theoretical value.

$$R_t = \left\{ \begin{array}{l} \text{Penalty for temperature tracking} \\ \text{Penalty for Supersaturation tracking} \\ \text{Penalty for particle size tracking} \\ \text{Penalty for Aspect Ratio tracking} \end{array} \right\}$$

$$R_t = \left\{ \begin{array}{l} \left( 0(|T_e| \leq 0.1) \vee -10(|T_e|) \right) + \left( 0(|S_e| \leq 0.1 \times 10^{-3}) \vee -2 \times 10^4(|S_e|) \right) + \left( 0(|L_e| \leq 1) \vee -30(|L_e|) \right) + \left( 0(|AR_e| \leq 0.01) \vee -20(|AR_e|) \right) \end{array} \right\}$$

Figure 2. Simplified overview of the reward function.

### 2.5 Transfer Learning

Transfer learning (TL) is a technique which is commonly used in RL to allow the agent to learn from previous agent experiences. The main advantage of TL is to improve training performances by implementing a sequence of relatively simple and attainable training objectives to help achieve the final training target gradually (Benyahia et al., 2021). In this study, TL is utilized by training the agents with an increased number of reference trajectories sequentially. Furthermore, error tolerances were also tightened gradually to improve tracking accuracy.

## 3. Result and Discussion

The corresponding training results indicated that two-agent RL requires less training episodes compared to a single-agent RL. This key advantage may save massive time-cost in the design and implementation of advanced control systems. Figure 3 shows the validation results where the performance of the two-RL agent is compared against a single agent RL in case study one. The sum of the squared errors (SSE) is computed and shown in Figures 3d, e and f to provide more quantitative insights to compare the accuracy of both control strategies. Both control methods can achieve trajectory tracking objectives to maintain the error within the acceptable range (Figures 3a, b, and c). However, the SSE indicates that a single agent exhibits a more effective performance compared to the two-agent RL in terms of accuracy.

Figure 4 presents a comparison between RL-PID and MPC in the 2D PMB model in case study 2. Here, the MPC is also developed in MATLAB/SIMULINK using basic tuning

parameters. Overall, the RL-PID shows better performance in accuracy compared to the MPC for the 2D crystallization process of KDP.

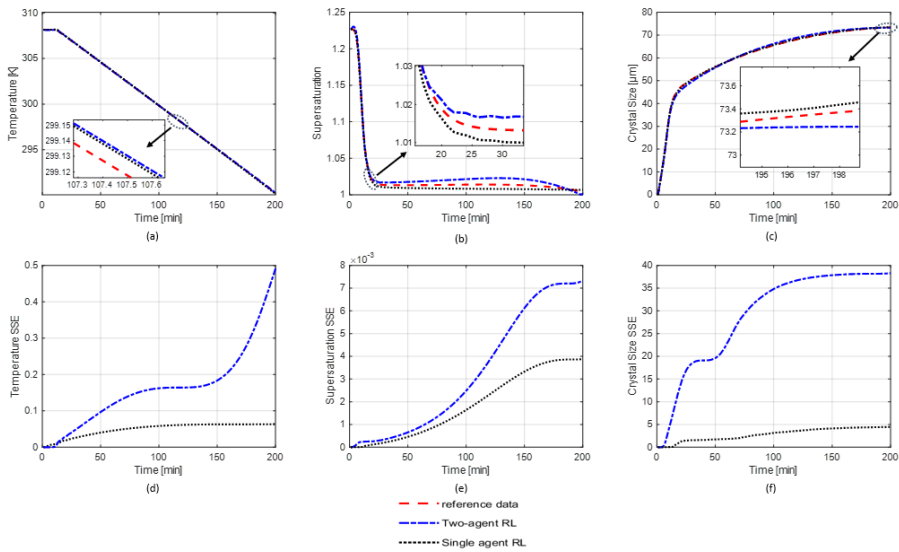


Figure 3. Validation results of the two RL agents vs single RL agent

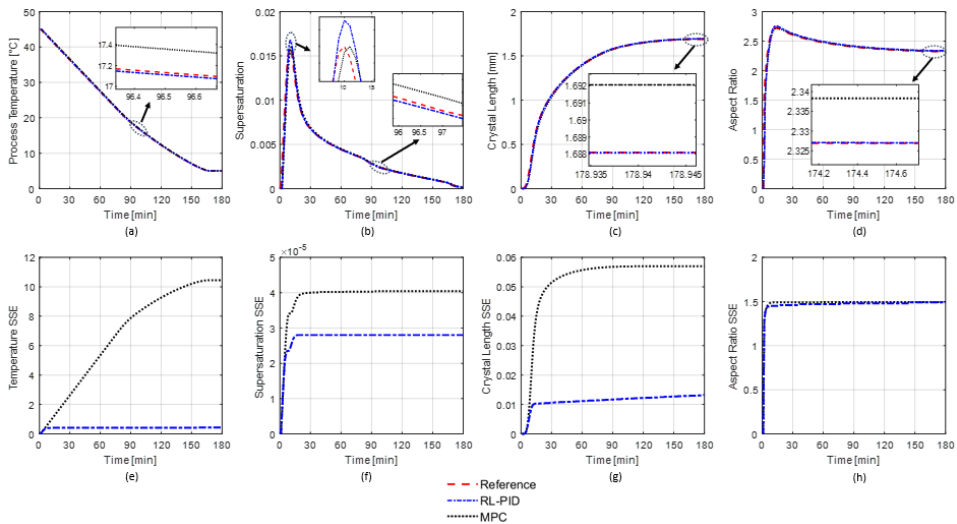


Figure 4. Overview of the comparisons between RL-PID and MPC

The simulation results depicted in Figures 4c and d show the tracking control performance of RL-PID with respect to one of the critical quality attributes, the crystal size and shape. Here, the RL-PID exhibits effectively and precise tracking performance. Figures 4e, f, g and h, provide more insights to further qualitatively investigate the performance of the RL-PID against MPC using the SSE. Overall, the RL-PID controller exhibits lower SSE compared to the MPC.

#### 4. Conclusions

A model-based RL with multiple agents and an RL-PID strategy was designed to achieve trajectory tracking control of an antisolvent cooling semi-batch and a cooling-only batch crystallization process. The proposed RL agents were based on the DDPG algorithm and were effectively trained using penalties only rewarding and transfer learning. Overall, both RL approaches exhibited excellent tracking performance. The results associated with case study one demonstrated that the two RL agents require less training time compared to the TL support single agent which reduces the training costs as a potential benefit. The results associated with case study two revealed that RL-PID can achieve more effective and accurate crystal size and shape control by tracking the optimal trajectories compared to the model predictive control.

**Acknowledgements:** This work was funded by the EPSRC (EP/R032858/1) ARTICULAR project.

#### References

- Barik, K., Prusti, P., Mohapatra, S. S. (2019). Single- and multi-objective optimisation for a combined cooling and antisolvent semi-batch crystallisation process with an ACADO toolkit, *Indian Chemical Engineer*, 62 (3), 287-300.
- Benyahia, B., Anandan, P. D., Rielly, C. (2021). Robust Model-Based Reinforcement Learning Control of a Batch Crystallization Process. *9th International Conference on Systems and Control (ICSC)*, 89-94.
- Canese, L., Cardarilli, G. C., Di Nunzio, L., Fazzolari, R., Giardino, D., Re, M., Spanò, S. (2021). Multi-Agent Reinforcement Learning: A Review of Challenges and Applications. *Appl. Sci.* 11, 4948.
- Cui, J., Liu, Y., Nallanathan, A. (2020). Multi-Agent Reinforcement Learning-Based Resource Allocation for UAV Networks. *IEEE Trans on Wire Commun*, 19, 729–743.
- Fysikopoulos, D., Benyahia, B., Borsos, A., Nagy, A. K., Rielly, C. D. (2019). A framework for model reliability and estimability analysis of crystallization processes with multi-impurity multi-dimensional population balance models, *Computers & Chemical Engineering*, 122, 275-292.
- Lakerveld, R., Benyahia, B. (2020). "CHAPTER 4: Process Control", *The Handbook of Continuous Crystallization*. The Royal Society of Chemistry, 172-218.
- Liu, J., Benyahia, B. (2022). Optimal start-up strategies of a combined cooling and antisolvent multistage continuous crystallization process, *Computers & Chemical Engineering*, 159, 107671.
- Mascia, S., Heider, P. L., Zhang, H., Lakerveld, R., Benyahia, B., Barton, P., Braatz, R. D., Cooney, C. L., Evans, J. M. B., Jamison, T. F., Jensen, K. F., Myerson, A. S., Trout, B. L. (2013). End-to-End Continuous Manufacturing of Pharmaceuticals: Integrated Synthesis, Purification, and Final Dosage Formation. *Angewandte Chemie International Edition*, 52 (47), 12359-12363.
- Nagy, Z. K., Fujiwara, M., Braatz, R. D. (2008). Modelling and control of combined cooling and antisolvent crystallization processes, *J. Process Contr*, 18(9), 856-864.
- Qie, H., Shi, D., Shen, T., Xu, X., Li, Y., Wang, L. (2019). Joint Optimization of Multi-UAV Target Assignment and Path Planning Based on Multi-Agent Reinforcement Learning. *IEEE Access*, 7, 146264–1462.

# Globally Convergent Composite-Step Trust-Region Framework for Model-Based Real-Time Optimization

Duo Zhang<sup>a</sup>, Xiang Li<sup>b</sup>, Kexin Wang<sup>a</sup>, Zhijiang Shao<sup>a</sup>

<sup>a</sup> *College of Control Science and Engineering, Zhejiang University, Hangzhou, 310027, China*

<sup>b</sup> *Department of Chemical Engineering, Queen's University, Kingston, ON K7L 3N6, Canada*

## Abstract

In model-based static real-time optimization (RTO), trust-region framework is globally convergent for unconstrained problems. However, it remains to be seen if the trust-region strategy can handle constraints directly with theoretical optimality guarantees. This paper addresses this issue and proposes a novel composite-step trust-region framework that guarantees global convergence for constrained RTO problems. In each iteration, the trial step is decomposed into a normal step that improves feasibility and a tangential step that reduces the cost function. Simulation results show that the proposed algorithm converges to the plant optimum in the presence of structural plant-model mismatches and wrong model curvatures. The proposed composite-step approach is more advantageous than the penalty approach because it needs a smaller penalty coefficient and does not suffer from ill-conditioned problems due to a large penalty coefficient.

**Keywords:** trust region, composite step, inequality constraints, global convergence, real-time optimization.

## 1. Introduction

Model-based optimization and control are important for achieving operational excellence. However, plant-model mismatches may influence practical performance and even cause convergence problems. Thus, handling mismatches by reliable optimization algorithms aroused great interest. In static model-based real-time optimization (RTO), algorithms based on gradient matching ideas can guarantee upon-convergence optimality in the presence of structural plant-model mismatch [1,2]. However, they are not globally convergent and may cause oscillation due to an inadequate model [3].

There are two ways to achieve global convergence in RTO. First, global convergence is a byproduct of feasible-side convergence. In this case, algorithms [4,5] rely on additional plant information such as Lipschitz constants or global upper bounds, which could be hard to get. Second, globalization strategies in nonlinear optimization are adopted. Among them, trust-region method [6] is the most natural for model-based RTO problems. Bunin [7] showed that the trust-region framework can be applied to unconstrained RTO problems. In each iteration, the next input is restricted to the trust region, where plant-model mismatches are small. However, constrained RTO problems have to be converted

to unconstrained ones by the penalty function to be incorporated into this framework [8]. It remains unclear if the trust-region strategy can handle constrained RTO problems directly with theoretical optimality guarantees. There are several benefits to preserving inequality constraints. On the one hand, it is more intuitive and more compatible with existing RTO algorithms. On the other hand, it is also more reliable since the physical model restricts the trial step. This paper addresses this issue and proposes a novel composite-step trust-region framework for constrained RTO problems.

## 2. Preliminaries

### 2.1. Model-Based Static Real-Time Optimization

We consider the optimization of a continuous process operating at steady states. The plant optimization problem can be stated as problem (1).  $u$  is the input variable,  $\phi$  is the cost function, and  $g$  is the constraints. The subscript  $p$  denotes the plant.

$$\begin{aligned} \min_u \phi_p(u) \\ \text{s. t. } g_p(u) \leq 0 \end{aligned} \quad (1)$$

The mathematical representation of the plant is unknown, so the model optimization problem (2) with subscript  $m$  is solved in each RTO iteration to calculate optimal  $u$ .

$$\begin{aligned} \min_u \phi_m(u) \\ \text{s. t. } g_m(u) \leq 0 \end{aligned} \quad (2)$$

However, plant-model mismatches exist, i.e.,  $\phi_p \neq \phi_m$  and  $g_p \neq g_m$ , and may lead to suboptimal or even infeasible input. Therefore, the model needs to be adapted based on the local information around  $u_k$  at each iteration, and the model optimization problem (3) then uses this adapted model, where subscript  $k$  indicates the  $k$ th RTO iteration.

$$\begin{aligned} \min_u \phi_{m,k}(u) \\ \text{s. t. } g_{m,k}(u) \leq 0 \end{aligned} \quad (3)$$

To deal with structural plant-model mismatch, the functions  $\phi_{m,k}$  and  $g_{m,k}$  in problem (3) are supposed to satisfy condition (4) given the information of plant output ( $\phi_p(u_k)$ ,  $g_p(u_k)$ ) and plant derivatives ( $\nabla\phi_p(u_k)$ ,  $\nabla g_p(u_k)$ ). Condition (4) means gradient matching between the model and the plant and leads to KKT matching upon convergence. There are various ways to do this, e.g., [1,2].

$$\phi_{m,k}(u_k) = \phi_p(u_k) \quad (4a)$$

$$g_{m,k}(u_k) = g_p(u_k) \quad (4b)$$

$$\nabla\phi_{m,k}(u_k) = \nabla\phi_p(u_k) \quad (4c)$$

$$\nabla g_{m,k}(u_k) = \nabla g_p(u_k) \quad (4d)$$

### 2.2. Trust-Region Framework for Unconstrained Problems

Algorithms that solve problem (3) repeatedly based on local models only guarantee upon-convergence optimality. Inadequate models [3] and improper algorithm parameters [2] could lead to oscillation. To address this, Algorithm 1 based on the trust-region framework was proposed [7,8].

#### **Algorithm 1, Trust-Region Algorithm for Unconstrained Problems**

**Step 0: Initialization.** Choose an initial point  $u_0$ , an initial trust-region radius  $\Delta_0 > 0$ . Choose constants  $0 < \eta_1 \leq \eta_2 < 1$ ,  $0 < \gamma_1 < 1$ .  $k \leftarrow 0$ .

**Step 1: Model adaptation.** Build a model satisfying Eqs. (4).

**Step 2: Step calculation.** Minimize the cost function of the model to obtain the step  $s_k$ .

$$\begin{aligned} s_k = \arg \min_s \phi_{m,k}(u_k + s) \\ \text{s. t. } \|s\| \leq \Delta_k \end{aligned} \quad (5)$$

Step 3: Acceptance of the trial point. Compute the ratio

$$\rho_k = \frac{\phi_p(u_k) - \phi_p(u_k + s_k)}{\phi_{m,k}(u_k) - \phi_{m,k}(u_k + s_k)} \quad (6)$$

If  $\rho_k > \eta_1$ , then define  $u_{k+1} = u_k + s_k$ ; otherwise, define  $u_{k+1} = u_k$ .

Step 4: Trust-region radius update. Set

$$\Delta_{k+1} \in \begin{cases} \min(\Delta_k, \Delta_{\max}) & , \text{ if } \rho_k > \eta_2, \\ \Delta_k & , \text{ if } \rho_k \in [\eta_1, \eta_2), \\ \gamma_1 \Delta_k & , \text{ if } \rho_k < \eta_1. \end{cases} \quad (7)$$

$k \leftarrow k + 1$ . Go back to Step 1.

In Step 2, the trial step is determined by minimizing the adapted model cost within the trust-region constraint. After applying the trial step to the plant and getting new measurements, we compare the actual improvement to the value predicted by the model in Step 3. The trial point is accepted as the next iterate if sufficient improvement is made. Otherwise, the trust region shrinks, and hopefully, the model provides a better prediction in a smaller region.

### 2.3. Penalty Method for Constrained Problems

Constrained RTO problems are more complicated because there are two aims: minimizing the cost function and reducing infeasibility. The penalty trust-region algorithm [8] converts the constrained problem to an unconstrained one by the merit functions (8), where  $c(u)$  is the infeasibility at the current point and  $\sigma > 0$  is the penalty coefficient.  $c(u)$  is calculated by Eqs. (9).

$$f_p(u, \sigma) = \phi_p(u) + \sigma c_p(u) \quad (8a)$$

$$f_{m,k}(u, \sigma) = \phi_{m,k}(u) + \sigma c_{m,k}(u) \quad (8b)$$

$$c_p(u) = \|\max(g_p(u), 0)\| \quad (9a)$$

$$c_{m,k}(u) = \|\max(g_{m,k}(u), 0)\| \quad (9b)$$

If the penalty coefficient  $\sigma$  is large enough, then Algorithm 1 with  $\phi(u)$  replaced by  $f(u, \sigma)$  is globally convergent [8]. However, determining  $\sigma$  is not easy because it depends on the specific problem, and a too large  $\sigma$  causes numerical difficulties.

## 3. Composite-Step Trust-Region Framework

In this section, we propose a novel trust-region algorithm for constrained RTO problems that deals with inequality constraints directly.

### 3.1. Algorithm Description

#### **Algorithm 2, Composite-Step Trust-Region Algorithm**

Step 0: Initialization. Choose an initial point  $u_0$ , maximum trust-region radius  $\Delta_{\max} > 0$ , an initial trust-region radius  $\Delta_0 \in (0, \Delta_{\max})$ , and a sufficiently large penalty parameter  $\sigma > 0$ . Choose constants  $0 < \eta_1 \leq \eta_2 < 1$ ,  $0 < \gamma_1 < 1$ ,  $0 < \xi < 1$ .  $k \leftarrow 0$ .

Step 1: Model adaptation. Build a model satisfying Eqs. (4).

Step 2: Normal step calculation. Move towards feasibility.

$$n_k = \arg \min_n c_{m,k}(u_k + n) \quad (10)$$

$$\text{s. t. } \|n\| \leq \xi \Delta_k$$

Step 3: Tangential step calculation. Compute a step  $t_k$  by

$$t_k = \arg \min_t \phi_{m,k}(u_k + n_k + t)$$

$$\text{s. t. } c_{m,k}(u_k + n_k + t) \leq c_{m,k}(u_k + n_k) \quad (11)$$

$$\|n_k + t\| \leq \Delta_k$$

Step 4: Acceptance of the trial point. The composite step is  $s_k = n_k + t_k$ . Compute the ratio

$$\rho_k = \frac{f_p(u_k, \sigma) - f_p(u_k + s_k, \sigma)}{f_{m,k}(u_k, \sigma) - f_{m,k}(u_k + s_k, \sigma)} \quad (12)$$

If  $\rho_k > \eta_1$ , then define  $u_{k+1} = u_k + s_k$ ; otherwise, define  $u_{k+1} = u_k$ .

Step 5: Trust-region radius update. Set  $\Delta_{k+1}$  according to Eq. (7). Go back to Step 1.

The composite step idea comes from [6]. The overall step is decomposed into a normal step that improves feasibility and a tangential step that reduces the cost function. By adjusting  $\xi$ , we change the step size for cost reduction and infeasibility reduction, and the overall step size is bounded by  $\Delta_k$ . Since the feasible set of problem (3) may not intersect with the trust region, small infeasibility has to be allowed during the iteration as long as there is a trend toward feasibility.

In Step 1, any model adaptation approach satisfying Eqs. (4) can be used, e.g., modifier adaption [1] and generalized parameter estimation [2]. In Step 2 and 3, the inequality-constrained subproblem is solved with the updated model. In Step 4, the progress towards a critical point is measured by the decrease of the merit function  $f(u, \sigma)$ .

### 3.2. Global Convergence

Global convergence is ensured because the trust-region framework enforces a monotonic decrease of the merit function. However, there are some prerequisites. First, all the subproblems are solved globally. Second,  $\sigma$  should be large enough, so the algorithm can avoid getting stuck at an infeasible point. Third, the model update approach should be consistent, i.e., it builds similar local models given similar measurements.

## 4. Case Studies

### 4.1. Simulation Settings

We illustrate the idea by a simple quadratic optimization problem (13) adapted from [9]. Input variables are bounded by  $u_1 \in [-2, 4]$  and  $u_2 \in [-4, 2]$ . It is assumed that the measurements are noise-free and that the plant gradients are available.

$$\begin{aligned} \min_u & u_1^2 + u_2^2 + u_1 u_2 \\ \text{s. t.} & 1 - u_1 + u_2^2 + 2u_2 \leq 0 \end{aligned} \quad (13)$$

Three models with different structural mismatches are considered. The first model optimization problem (14) has the right model curvature, while the other two problems (15) and (16) have the wrong curvature in the cost function and the constraint, respectively. Models with wrong curvatures are likely to be inadequate and cause oscillation.

$$\begin{aligned} \min_u & u_1^2 + u_2^2 \\ \text{s. t.} & 1 - u_1 + u_2^2 \leq 0 \end{aligned} \quad (14)$$

$$\begin{aligned} \min_u & -u_1^2 + u_2^2 \\ \text{s. t.} & 1 - u_1 + u_2^2 \leq 0 \end{aligned} \quad (15)$$

$$\begin{aligned} \min_u & u_1^2 + u_2^2 \\ \text{s. t.} & 1 - u_1 - 4u_2^2 \leq 0 \end{aligned} \quad (16)$$

The proposed algorithm is compared with modifier adaptation [1] (no trust-region constraints, with filter gain  $K = 0.1$ ) and penalty trust-region method. The simulation begins at point  $[2, -2]$ . The maximum step size for each iteration is 2. Parameters for Algorithm 1 are selected according to [6]. For the composite-step and the penalty methods,

the penalty coefficient and the initial trust radius are 1. In Step 1 of each algorithm, the model is updated in the same way as in the modifier adaptation method. For the composite-step method,  $\xi = 0.5$ . The nonsmooth maximum operator is approximated by a smooth one [8]. Global convergence of the proposed algorithm and the influence of  $\sigma$  are studied.

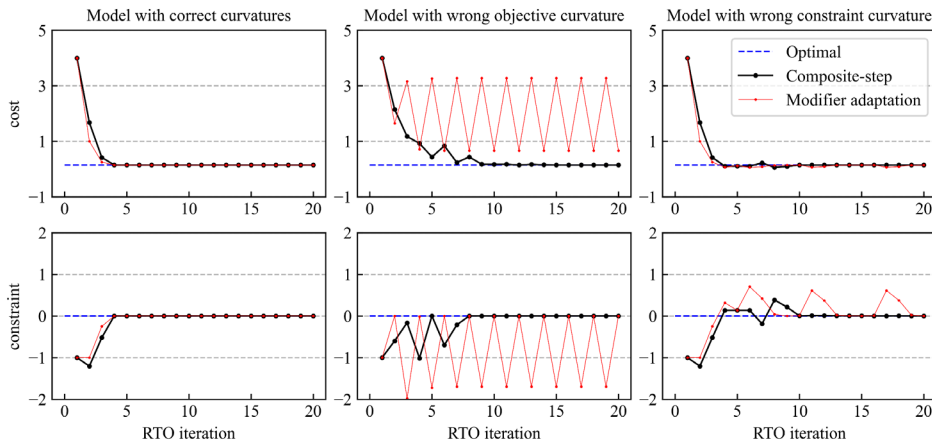


Figure 1 Evolution of plant cost and constraint using mismatched models during 20 RTO iterations. Left: simulation using model (14); middle: simulation using model (15); right: simulation using model (16). Upper: profiles of plant cost; lower: profiles of plant constraint. Black: Algorithm 2; red: modifier adaptation method [1].

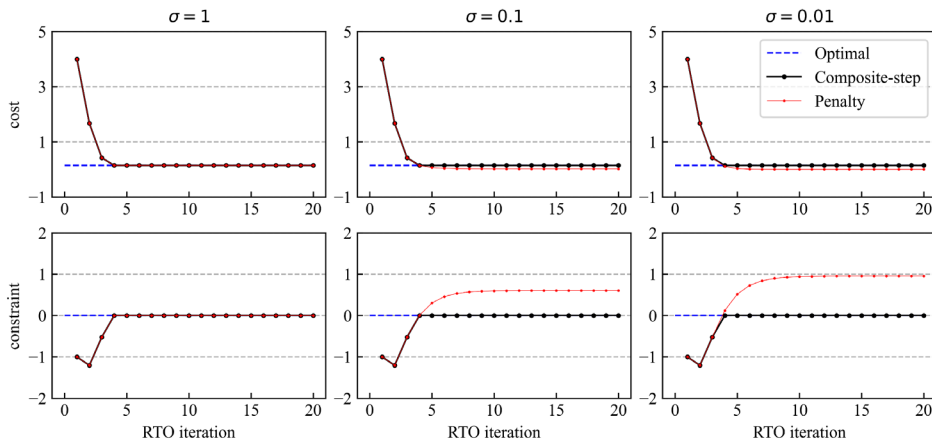


Figure 2 Evolution of plant cost and constraint with different penalty coefficients during 20 RTO iterations. Left:  $\sigma = 1$ ; middle:  $\sigma = 0.1$ ; right:  $\sigma = 0.01$ . Upper: profiles of plant cost; lower: profiles of plant constraint. Black: Algorithm 2; red: penalty trust-region method [8].

#### 4.2. Global Convergence

In this subsection, we show that the proposed algorithm converges even when objective and constraints curvatures of the model are wrong. The three model optimization problems (14)-(16) are investigated. Model adaptation method (without trust-region constraints) serves as a comparison.

As Figure 1 shows, the proposed composite-step algorithm converges to the plant optimum in all three scenarios, while the modifier adaptation method fails to find the



optimum with the wrong curvature model. For both methods, the model's KKT condition matches that of the plant. However, the second-order sufficient optimality condition may be violated for the modifier adaptation method due to the wrong curvature of the model. In other words, the trust-region framework lowers the requirement for model adequacy.

#### 4.3. Penalty Coefficient

In this subsection, we investigate the effect of a small penalty coefficient on the two trust-region algorithms. RTO based on problem (14) using the composite-step and the penalty trust-region methods with different  $\sigma$  are illustrated in Figure 2.

Theoretically, both methods could be biased if  $\sigma$  is not large enough. However, our experience shows that the proposed method is less sensitive to small  $\sigma$  than the penalty method. This is because the model constraint is explicitly imposed in Step 3.

Moreover, if a very large  $\sigma$  is adopted to avoid the bias, the optimization subproblems in the penalty method may be ill-conditioned and hard to solve. The proposed composite-step approach does not suffer from this problem. The reason is that  $\sigma$  only appears when evaluating the trial point in Eq. (12), not in solving optimization subproblems (11) or (12).

### 5. Conclusions

This paper proposes a novel composite-step trust-region algorithm that directly handles inequality RTO constraints. The composite-step approach does not suffer from numerical problems caused by a large penalty coefficient in the merit function. Numerical simulations show that the proposed method is globally convergent and reduces model adequacy requirements. Future work will be focused on the rigorous proof and methods to find a proper penalty coefficient  $\sigma$  automatically.

### Acknowledgements

This work was supported by the National Natural Science Foundation of China under grant number 62120106003, 62173301, 61873242, and China Scholarship Council under CSC NO. 202006320176.

### References

- [1] Marchetti, A., Chachuat, B. and Bonvin, D., 2009. Modifier-adaptation methodology for real-time optimization. *Industrial & engineering chemistry research*, 48(13), pp.6022-6033.
- [2] Zhang, D., Wang, K., Xu, Z., Tula, A.K., Shao, Z., Zhang, Z. and Biegler, L.T., 2022. Generalized Parameter Estimation Method for Model-Based Real-Time Optimization. *Chemical Engineering Science*, 258, p.117754.
- [3] Forbes, J.F. and Marlin, T.E., 1996. Design cost: A systematic approach to technology selection for model-based real-time optimization systems. *Computers & chemical engineering*, 20(6-7), pp.717-734.
- [4] Bunin, G.A., François, G. and Bonvin, D., 2014. Implementation techniques for the SCFO experimental optimization framework. arXiv preprint arXiv:1406.3997.
- [5] Marchetti, A.G., Faulwasser, T. and Bonvin, D., 2017. A feasible-side globally convergent modifier-adaptation scheme. *Journal of Process Control*, 54, pp.38-46.
- [6] Conn, A.R., Gould, N.I. and Toint, P.L., 2000. Trust region methods. Society for Industrial and Applied Mathematics.
- [7] Bunin, G.A., 2014. On the equivalence between the modifier-adaptation and trust-region frameworks. *Computers & chemical engineering*, 71, pp.154-157.
- [8] Biegler, L.T., Lang, Y.D. and Lin, W., 2014. Multi-scale optimization for process systems engineering. *Computers & Chemical Engineering*, 60, pp.17-30.

- [9] del Rio Chanona, E.A., Petsagkourakis, P., Bradford, E., Graciano, J.A. and Chachuat, B., 2021. Real-time optimization meets Bayesian optimization and derivative-free optimization: A tale of modifier adaptation. *Computers & Chemical Engineering*, 147, p.107249.



# Fault Detection and Diagnosis for Chemical Processes based on Deep Neural Networks with Continuous Wavelet Transform

Chinatsu Ukawa,<sup>a</sup> Yoshiyuki Yamashita,<sup>b</sup>

*<sup>a</sup>Department of Food and Energy Systems Science, Graduate School of Bio-Applications Systems Engineering, Tokyo University of Agriculture and Technology, 2-24-16 Naka-cho, Koganei, Tokyo 184-8588, JAPAN*

*<sup>b</sup>Department of Chemical Engineering, Tokyo University of Agriculture and Technology, 2-24-16 Naka-cho, Koganei, Tokyo 184-8588, JAPAN*

## Abstract

This study proposed a novel fault detection and diagnosis method using continuous wavelet transform (CWT) and a three-dimensional convolutional neural network (3D-CNN). In particular, multivariate time series data from chemical plants were divided by a time-shifting window and transformed into scalograms using CWT. These scalograms were fed to 3D-CNN to generate outputs indicating the faults that occurred in the process. We applied the proposed method to a Tennessee Eastman process dataset. The proposed method adequately captures the characteristics in the time-frequency domain and exhibited good fault detection and diagnosis performances on the simulation dataset.

**Keywords:** Fault Detection, Fault Diagnosis, Deep Learning, Wavelet Transformation

## 1. Introduction

Fault detection and diagnosis (FDD) are crucial steps in process monitoring to maintain high productivity, efficiency, and safety in a chemical plant operation. Data-based FDD methods have attracted attention in recent years due to the dramatic growth of computer and information technology. Among various data-driven methods, machine learning (ML), such as deep learning, is a promising method for capturing process dynamics and nonlinearity. ML have shown great potential in various research fields (Xiong, 2022), where numerous studies have been conducted in the process engineering field (Ge, 2021).

Data preprocessing is a critical phase in developing accurate models using ML. The appropriate selection of preprocessing methods can contribute to achieving ML exhibiting high performance. Signal processing, such as Fourier transform and wavelet transform, are typical examples of preprocessing. They are used for denoising or feature extraction in various fields. Fourier transform has been applied to numerous fields, exhibiting outstanding performance for feature extraction. However, this method misses the time information of the original signal and cannot work well when the original data includes nonlinearity and non-stationarity. In contrast, wavelet transform is a powerful tool to extract the information in the time-frequency domain of the original signal. This method can be applied to nonlinear and nonstationary signals. In particular, wavelet transform has shown great potential in fault detection, including bearing and electrocardiogram. Although ML and wavelet transform have been successfully applied in other research

fields, their implementation to obtain the time-frequency information using continuous wavelet transform (CWT) is uncommon for the field of FDD for chemical processes.

Here, we proposed a combined FDD method using CWT and a three-dimensional convolutional neural network (3D-CNN). We obtained 3D inputs called scalograms for 3D-CNN from preprocessing using CWT and classified the result outputs regarding the type of fault that occurred in the chemical process.

## 2. Method

### 2.1. Continuous wavelet transform

CWT is a widely mathematically implemented wavelet transform, which can exhibit a complete view of the signal by continuously changing the scale and translation parameters (Ramin, 2022). In particular, CWT has been used for filtering out noise and fault detection and diagnosis in various fields, such as electrocardiogram analysis (Alharbey, 2022), and the development of non-steady process simulation (Hong, 2022). CWT analyses an original signal using a mother wavelet. The equation of the mother wavelet is shown below as Equation 1, where parameters  $a$  and  $b$  denoted scaling and time-shifting.

$$W_x(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi^* \left( \frac{t - b}{a} \right) dt \quad (1)$$

In addition, CWT produces scalograms showing the feature from the original signal in the time-frequency domain. This visualization method is also used in various fields, for instance, fault detection and feature extraction of seismic data (Ali, 2020) and filtering tool for velocity signals of a thunderstorm (Brusco, 2022).

### 2.2. 3D-CNN

We applied a 3D-CNN in the proposed method for feature extraction and classification to handle multivariable datasets. Various neural network methods have been proposed in recent years, such as Long-Short Term Memory, Recurrent Neural Network, Gaussian Neural Network, and Convolutional Neural Network (CNN). Nevertheless, CNN is an already known method for feature extraction and classification methods by our group. In particular, 2-dimensional CNN is a popular method for image recognition and classification and FDD methods for chemical processes. Moreover, CNN has been widely applied to overcome various problems owing to its great performance. (Kim, 2022)

Furthermore, 3D-CNN is a logical extension of 2D-CNN. The feature extraction scheme is the same as other dimensional CNN. However, 3D-CNN uses 3D filters to convolute 3D data and passes convoluted data to pooling layers before the fully connected layers for classification. This method has been proposed for image recognition with medical 3D scans (Pintelas, 2022), or agricultural hyperspectral images (Diao, 2022). In this study, we selected 3D-CNN to deal with multivariable from a particular chemical process.

### 2.3. Proposed method

Chemical processes have non-steady and nonlinear multivariable data. These processes can be divided by a time-shifting window. Each divided data can be transformed into a scalogram to show a particular feature of data in the time-frequency domain using CWT.

We decided that the size of a scalogram was 20 x window size. This processing was applied to each process variable to obtain 3D input data (20 x window size x the number of process variables). The generated input data was fed to the 3D-CNN model to obtain outputs representing the chemical process condition.

### 3. Case study

Tennessee Eastman process (TEP) is one of the most popular benchmarks in chemical process FDD. This process contains five main units: a reactor, a condenser, a stripper, a vapor-liquid separator, and a compressor. The process flow is shown in Figure 1. Several types of TEP simulators and datasets with different types of control loops or various fault scenarios have been developed (Bathelt, 2015, Chiang, 2000D). In this study, we used an extended TEP dataset introduced in 2017 (Rieth, 2017), which includes 20 fault operation modes of all 21 modes and 52 process variables with 41 measured and 11 manipulated variables.

In this study, we set the size of the time shift window for each preprocessing method. In particular, we set 20 for CWT and 128 for Fast Fourier Transform (FFT) and Short-Time Fourier Transform (SFT). The number of selected process values was 33, including 22 measured and 11 manipulated variables.

## 4. Result

### 4.1. Metrics for evaluation

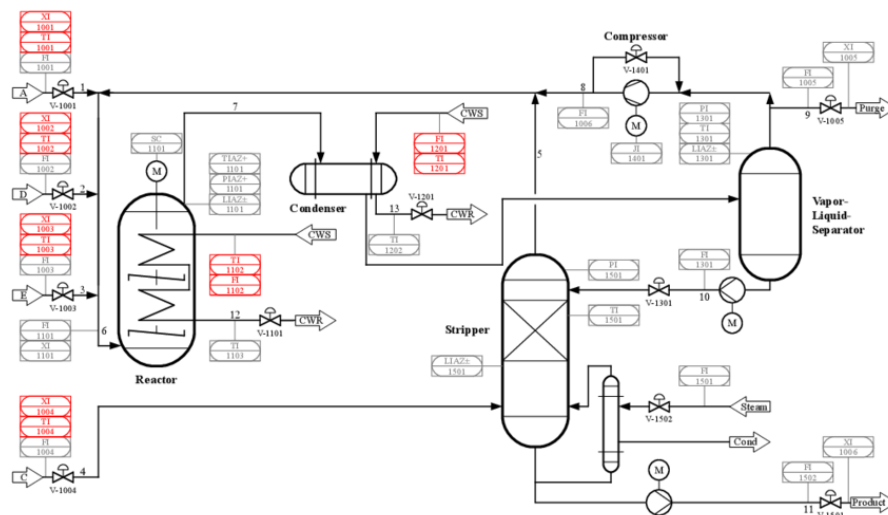


Figure 1 Process flow of TEP

We evaluated the proposed method and compared its performance using the fault detection rate (FDR) and false positive rate (FPR). TP represents true positive, and FN represents false negative.

$$\text{Fault Detection Rate (FDR)} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{False Positive Rate (FPR)} = \frac{FP}{FP + TP} \quad (3)$$

#### 4.2. Application to TEP

We compared the proposed method with a combination of 3D-CNN methods with FFT and STFT. The results of the comparison are listed in Table 1. The proposed method outperformed the other preprocessing methods.

### 5. Discussion

The result of the proposed method shows excellent performance on more than half the number of faults compared to FFT and STFT. Figure 2 shows some of the input samples for 3D-CNN. We selected measured variable 1 samples of fault 6 at normal operation and when the fault occurred. We can see the change in the time-frequency domain of this variable, indicating that the proposed model learned the feature from input samples and classifies process conditions based on it.

However, fault-free (fault 00) exhibited the worst FDR result. We verified the prediction output from 3D-CNN, representing a ratio of each fault. The ratio ranged from 0 to 1. The model estimated several fault-free data as fault data of faults 3, 9, and 15. Thus, we have to solve how to divide fault-free and faulty data in future works.

Some FDR of CWT were lower than FFT and STFT. The difference in window size between these methods might be one of the reasons for this result. We compared the effect of window size on the model's ability by setting the window size from 20 to 40. The result showed that expanded window size was valid for fault 21 even though other faults were worse detected. Our next step is establishing a method to decide on the appropriate window size for the original process data.

### 6. Conclusion

This study proposed a fault detection and diagnosis model for chemical processes using CWT and 3D-CNN. CWT was used as preprocessing method for chemical process data. This method has exhibited great potential for feature extraction for non-steady and nonlinear signals in various research fields. In the proposed method, the chemical process data was first divided into samples with a set length. After that, samples were transformed into scalograms using CWT. The 3D-CNN was applied to extract the feature embedded in the time-frequency domain and perform multivariate analysis with the chemical process variables. The structure of the 3D-CNN comprised convolutional layers, pooling layers, dropout layers, and fully connected layers. The scalograms were fed to 3D-CNN to generate the outputs, representing the faults in the chemical process. The proposed method was applied to the TEP dataset introduced by Rieth et al. in 2017. Our method could capture the characteristics in the time-frequency domain well and showed better performance on more than half the number of faults than FFT and STFT. However, in the case of fault-free, the proposed model should be improved. We will train the proposed model without fault-free data and evaluate its ability to classify faults. Additionally, we will examine the method to decide the window size.

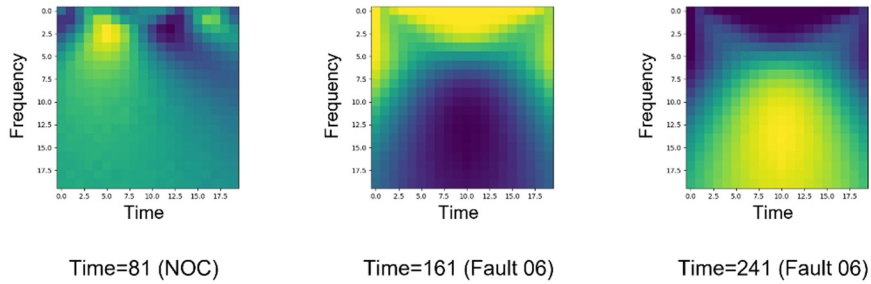


Figure 2 Input samples of fault 06

References

Table 1 Result from the proposed model and other preprocessing methods

Fault	FFT		STFT		CWT			
	FDR	FPR	FDR	FPR	window size=20		window size=40	
					FDR	FPR	FDR	FPR
0	0.063	0.343	0.168	0.767	0.139	0.458	0.138	0.534
1	<b>0.172</b>	<b>0.762</b>	<b>0.673</b>	<b>0.269</b>	<b>0.979</b>	<b>0.030</b>	<b>0.968</b>	<b>0.023</b>
2	<b>0.362</b>	<b>0.889</b>	<b>0.806</b>	<b>0.067</b>	<b>0.972</b>	<b>0.044</b>	<b>0.954</b>	<b>0.049</b>
3	<b>0.112</b>	<b>0.942</b>	<b>0.106</b>	<b>0.859</b>	<b>0.468</b>	<b>0.728</b>	<b>0.399</b>	<b>0.725</b>
4	<b>0.691</b>	<b>0.731</b>	<b>0.716</b>	<b>0.266</b>	<b>0.968</b>	<b>0.024</b>	<b>0.958</b>	<b>0.024</b>
5	<b>0.207</b>	<b>0.877</b>	<b>0.491</b>	<b>0.724</b>	<b>0.959</b>	<b>0.047</b>	<b>0.866</b>	<b>0.076</b>
6	<b>0.836</b>	<b>0.312</b>	<b>0.826</b>	<b>0.342</b>	<b>0.990</b>	<b>0.003</b>	<b>0.975</b>	<b>0.034</b>
7	<b>0.217</b>	<b>0.711</b>	<b>0.734</b>	<b>0.366</b>	<b>0.973</b>	<b>0.040</b>	<b>0.921</b>	<b>0.050</b>
8	<b>0.804</b>	<b>0.268</b>	<b>0.588</b>	<b>0.311</b>	<b>0.865</b>	<b>0.170</b>	<b>0.865</b>	<b>0.218</b>
9	<b>0.107</b>	<b>0.936</b>	<b>0.083</b>	<b>0.862</b>	<b>0.139</b>	<b>0.876</b>	<b>0.090</b>	<b>0.867</b>
10	<b>0.520</b>	<b>0.738</b>	<b>0.292</b>	<b>0.685</b>	<b>0.698</b>	<b>0.579</b>	<b>0.615</b>	<b>0.632</b>
11	0.933	0.084	0.878	0.066	0.733	0.493	0.544	0.554
12	<b>0.712</b>	<b>0.031</b>	<b>0.623</b>	<b>0.047</b>	<b>0.755</b>	<b>0.106</b>	<b>0.697</b>	<b>0.113</b>
13	<b>0.511</b>	<b>0.349</b>	<b>0.571</b>	<b>0.521</b>	<b>0.851</b>	<b>0.136</b>	<b>0.815</b>	<b>0.282</b>
14	0.955	0.019	0.954	0.015	0.864	0.371	0.863	0.355
15	0.299	0.924	0.266	0.839	0.168	0.878	0.131	0.900
16	0.828	0.274	0.740	0.277	0.667	0.356	0.523	0.599
17	<b>0.890</b>	<b>0.034</b>	<b>0.820</b>	<b>0.083</b>	<b>0.922</b>	<b>0.117</b>	<b>0.852</b>	<b>0.226</b>
18	<b>0.624</b>	<b>0.268</b>	<b>0.570</b>	<b>0.311</b>	<b>0.926</b>	<b>0.197</b>	<b>0.875</b>	<b>0.215</b>
19	0.875	0.518	0.732	0.424	0.522	0.651	0.369	0.789
20	<b>0.879</b>	<b>0.203</b>	<b>0.792</b>	<b>0.225</b>	<b>0.906</b>	<b>0.224</b>	<b>0.913</b>	<b>0.158</b>
<b>AVERAGE</b>	<b>0.552</b>	<b>0.486</b>	<b>0.592</b>	<b>0.396</b>	<b>0.736</b>	<b>0.311</b>	<b>0.682</b>	<b>0.353</b>

X. Ge, B. Wang, X. Yang, Y. Pan, B.Liu, B. Liu, 2021, Fault detection and diagnosis for reactive distillation based on convolutional neural network, Computers & Chemical Engineering, Volume 145, 107172



S. Xiong, L. Zhou, Y. Dai, X. Ji, 2022, Attention-based LSTM fully convolutional network for chemical process fault diagnosis, *Chinese Journal of Chemical Engineering* (In Press)

A. Ali, S. C. Chen, M. Shah, 2020, Continuous wavelet transformation of seismic data for feature extraction. *SN Applied Sciences*, Volume 2, 1835.

M. Ramin, M. R. Sohrabi, F. Motiee, 2022, Smart spectrophotometric methods for enhancement spectral resolution and rapid simultaneous determination of dapoxetine andildenafil in commercial tablets using fuzzy inference system and continuous wavelet transform, *Chemometrics and Intelligent Laboratory Systems*, Volume 230, 104656, ISSN 0169-7439

H. P. Hong, X.Z. Cui, D. 2022, Qiao, Simulating nonstationary non-Gaussian vector process based on continuous wavelet transform, *Mechanical Systems and Signal Processing*, Volume 165, 108340, ISSN 0888-3270

S. Brusco, G. Buresti, G. Piccardo, 2022, Thunderstorm-induced mean wind velocities and accelerations through the continuous wavelet transform, *Journal of Wind Engineering and Industrial Aerodynamics*, Volume 221, 104886, ISSN 0167-6105

C. A. Rieth, B. D. Amsel, R. Tran, M. B. Cook, 2017, Additional Tennessee Eastman process simulation data for anomaly detection evaluation", <https://doi.org/10.7910/DVN/6C3JR1>, Harvard Dataverse, V1

A. Bathelt, N. L. Ricker, M. Jelali, 2015, Revision of the Tennessee Eastman process model, *IFAC-PapersOnLine*, Volume 48, Issue 8, Pages 309-314, ISSN 2405-8963

L. H. Chiang, E. L. Russell, R. D. Braatz, 2020, Fault diagnosis in chemical processes using Fisher discriminant analysis, discriminant partial least squares, and principal component analysis, *Chemometrics and Intelligent Laboratory Systems*, Volume 50, Issue 2, Pages 243-252, ISSN 0169-7439

R.A. Alharbey, S. Alsubhi, K. Daqrouq, A. Alkhateeb, 2022, The continuous wavelet transform using for natural ECG signal arrhythmias detection by statistical parameters, *Alexandria Engineering Journal*, Volume 61, Issue 12, Pages 9243-9248, ISSN 1110-0168

E. Pintelas, P. Pintelas, 2022, A 3D-CAE-CNN model for Deep Representation Learning of 3D images, *Engineering Applications of Artificial Intelligence*, Volume 113, 104978, ISSN 0952-1976

Z. Diao, J. Yan, Z. He, S. Zhao, P. Guo, 2022, Corn seedling recognition algorithm based on hyperspectral image and lightweight-3D-CNN, *Computers and Electronics in Agriculture*, Volume 201, 107343, ISSN 0168-1699

M. I. Kim, H. S. Yoon, 2022, Geometric modification for the enhancement of an airfoil performance using deepRa CNN, *Ocean Engineering*, Volume 266, Part 4, 2022

# Anomaly detection in chemical processes with semantic knowledge graphs: an approach to reduce cause-effect diagrams

Nazanin Hamedi<sup>a</sup>, Ilhan Mutlu<sup>b</sup>, Fatima Rani<sup>b</sup>, Leon Urbas<sup>b</sup>

<sup>a</sup>*DFG Research Training Group 2323 CD-CPPS, Technische Universität Dresden, Georg-Schumann-Straße 7a, 01069 Dresden, Germany*

<sup>b</sup>*Faculty of Electrical and Computer Engineering, Chair of Process Control Systems & Process Systems Engineering Group, Technische Universität Dresden, 01069 Dresden, Germany*

[nazanin.hamedi@tu-dresden.de](mailto:nazanin.hamedi@tu-dresden.de)

## Abstract

When operators and users face an unexpected problem regarding the process, it would be easier and less time-consuming for them to narrow down the possible causes that result in the final issue. In this study, a pilot modular plant with two Process Equipment Assemblies was considered as the use case. At first, a knowledge graph representing this process was developed in Protégé software, which semantically described not only the equipment type and connectivity but also the behavior of the process. Then, the knowledge graph was imported to Python, where the sensor data were placed in their particular position in the knowledge graph. Afterward, an algorithm was developed to query the knowledge graph and verify if the relevant equipment was functioning correctly or not. Our results indicate that this approach can reduce the cause-effect diagrams in almost all scenarios. Nonetheless, there are situations where further sensor data (e.g., the temperature in a tank) is required for the algorithm to decide.

**Keywords:** Semantic knowledge graphs; Process behavior, Cause-effect graphs; Decision making.

## 1. Introduction

The advent of the computer has significantly impacted the control of chemical processes. Indeed, the regulatory control of chemical processes (e.g., opening and closing valves) is entirely automated, resulting in higher product quality and consistency, process safety, and process efficiency (Venkatasubramanian et al., 2003). Despite these advancements, the time-consuming detection of an abnormal situation, identifying its cause, and making the best decision to resolve the problem is still a manual task. Based on the statistics, minor accidents in chemical plants occur daily, leading to workers' injuries and posing huge financial pressure on the industry. Also, it has been reported that 70% of industrial accidents are due to operators' mistakes (Venkatasubramanian et al., 2003). These suggest that developing assistant systems for fault diagnosis is a pressing need.

Notable among different methodologies used for fault diagnosis are qualitative, graph-based approaches, which were comprehensively reviewed by Venkatasubramanian et al. (Venkatasubramanian, Rengaswamy and Kavuri, 2003) and Maurya et al. (Maurya, Rengaswamy and Venkatasubramanian, 2003).

A clear superiority of graph-based methodologies is that they provide a causal structure by which one can track the propagation of an error in a chemical process. However, as the processes get more sophisticated, the graphs become more complex and challenging to navigate (especially because there exist multiple causes for one problem) (Reinartz *et al.*, 2019).

This motivates us to propose an approach for reducing the cause-effect graphs, making them more traceable. In this method, the process has been represented in the form of a knowledge graph in Protégé, in which the equipment type, connectivity, and process behavior are considered. Afterward, by importing this knowledge graph to Python and developing a program that mimics the sensor data, we were able to omit some of the unnecessary branches of the cause-effect graph. It is noteworthy to mention that to examine the applicability of our approach; it has been implemented for a use-case; namely, the Safety Demonstrator (Pelzer *et al.*, 2021).

## 2. Method

### 2.1. Cause-effect graph

As mentioned before, the Safety Demonstrator was considered as a use-case in this paper. It is a modular plant with two Process Equipment Assemblies (PEAs). As its name suggests, the Safety Demonstrator aims to investigate a secure operation; and hence, many safety controllers have been taken into account in its design. Therefore, it would be helpful for the operator if he-she received information about why one of the safety control loops is active. In this study, a cause-effect graph is proposed (see figure 1), which is defined as a directed graph in which the nodes are the events (e.g., activation of one controller), and the edges illustrate that the reason for each event is actually the event on the subsequent node. Even though it is a very beneficial tool for identifying a problem in a chemical plant, it can become very complex. Thus, the idea of this research is to automatically reduce the graph based on the specific situation.

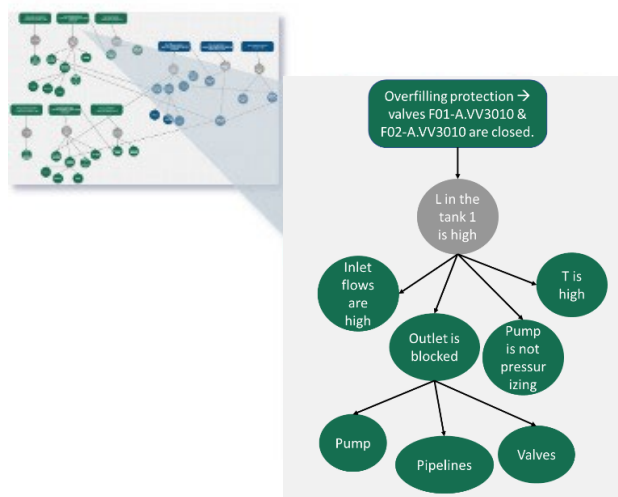


Figure 1- The cause-effect graph of the Safety Demonstrator (Gray nodes are the logics behind the controllers, rectangles show the active control system, green parts are related to PEA 1 and blue parts are related to PEA 2)

### 2.2. Semantic definition of the process

In this study, Protégé software was used to develop a knowledge graph of the investigated process. As can be seen in figure 2, it consists of 4 major classes: Equipment, Parameter, Chemicals, and Process Function. It is noteworthy to mention that the Process Function class aims to explain each equipment's phenomena semantically. To complete the definition of process function, the knowledge graph was imported from Protégé to Python with the help of owlready (a library in python for ontology descriptions) (Lamy, 2017). Then, as can be seen in figure 3 (A & D), two classes of Mass Transfer and Pressurizing and the related attributes and methods were defined. Indeed, a mass transfer takes place in an equipment when mass enters that equipment and the outlet mass is almost the same as the inlet. Also, an equipment does pressurize when the pressure of its inlet is lower than that of its outlet. These human-understandable facts are defined for the computer, as depicted in figure 3 (B).

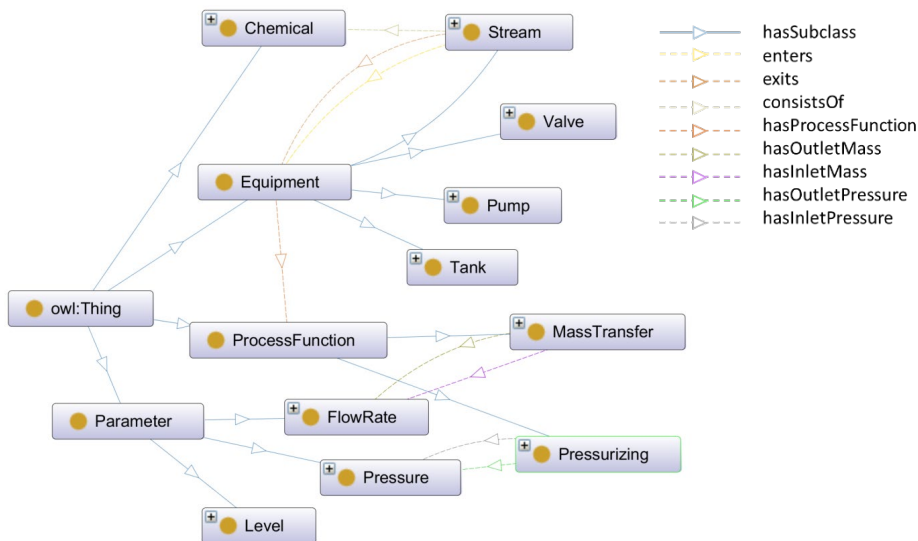


Figure 2 – Knowledge graph of the PEA1 of Safety Demonstrator developed by Protégé.

### 2.3. Algorithm

After the complete definition of the process in Python (figure 3 (A & B)), the safety control loops were also defined (figure 3 (C)). These controllers are activated when a process parameter (which has already been semantically defined) crosses the considered threshold. For instance, Overfilling-controller needs the value of the level in the tank, which has been already defined in the knowledge graph. Thus, the algorithm is aware that this value is the level of the tank, and the tank is actually connected to pump1 by stream3. The flowchart for drawing the cause-effect graph is shown in figure 3 (D). It should mention that this flowchart only considers the graph for the overfilling-controller, but a similar algorithm can be imagined for other controllers. In short, the algorithm checks whether a safety controller is active or not. If yes, then it searches for the possible reasons (based on the cause-effect graph), and if the cause is valid (based on the semantic definition of the process), it is added to the cause-effect graph.

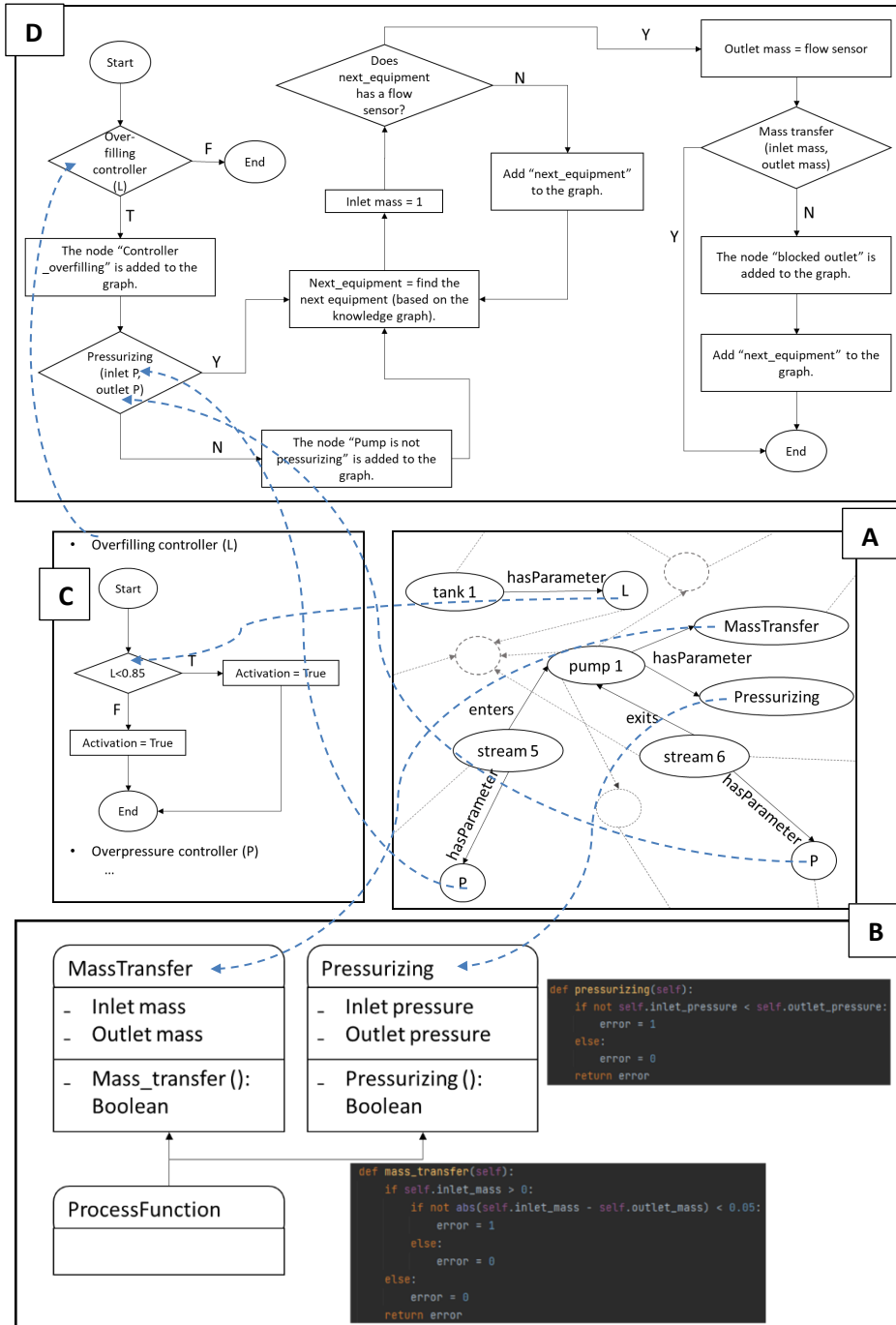


Figure 3 – A) Part of the knowledge graph from protégé (gray nodes and edges indicate that the shown nodes and edges are not the whole graph), B) Definition of process function in Python, C) Definition of the safety controllers, D) Graph development flowchart.

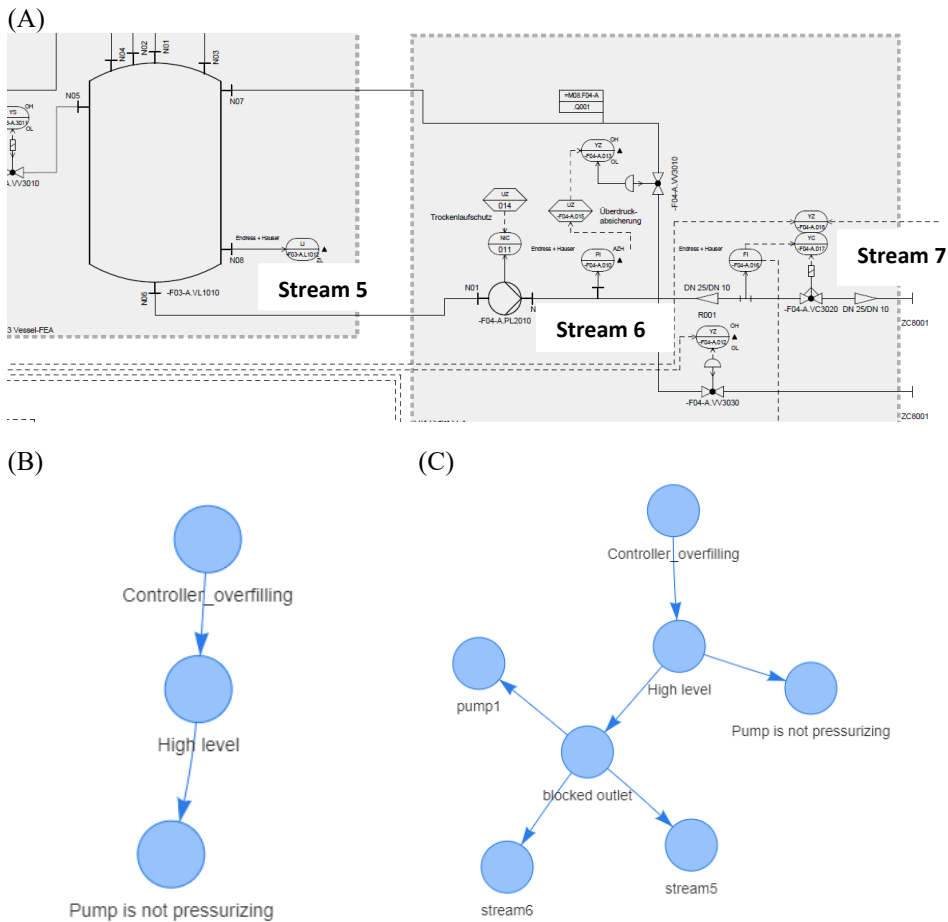


Figure 4 – A) Part of P&ID diagrams of the first PEA of Safety Demonstrator, (B) Reduced graph of a scenario in which pump is not working, (C) Reduced graph of a scenario in which pump is not working and the flow meter shows a blockage in the line.

### 3. Results and discussion

The above-mentioned method has been examined in different scenarios related to Safety Demonstrator. Because of space limitations, the scenarios investigated here are all related to PEA1 of the Safety Demonstrator (see figure 4 (A)).

- The first scenario is that pump 1 is not pressurizing the flow; i.e., the pressure after the pump (shown by the sensor) is not more than that of stream 5 (calculated from the length of the liquid in the tank), resulting in the activation of the overfilling-controller. The result of this is illustrated in figure 4 (B).
- Second scenario is that the pump is not pressurizing, and at the same time, the flow meter on stream 6 is showing a blockage. The produced graph for this scenario is depicted in figure 4 (C). Since there is no flow sensor in stream 5, our algorithm cannot precisely decide the exact point of blockage; nonetheless, it can identify it in the outlet of the tank and before the flow sensor in stream 6.

#### 4. Conclusion

The purpose of this work is to find a method to reduce cause-effect graphs, facilitating operators' problem-solving ability. To this end, a use case, namely Safety-Demonstrator, was selected, which was first semantically defined in protégé, then imported to Python, where the process behavior was defined entirely. Afterward, an algorithm was developed to communicate with the knowledge graph and, based on the process sensor data, draw the corresponding cause-effect graph. The derived results indicate that this approach is successful in considered scenarios; however, the accuracy of the results would be improved depending on the actual sensor setup on the plant side.

#### References

- J. B. Lamy, 2017, Owlready: Ontology-oriented programming in Python with automatic classification and high level constructs for biomedical ontologies, *Artificial Intelligence in Medicine*, 80, pp. 11–28.
- M. R. Maurya, R. Rengaswamy, V. Venkatasubramanian, 2003, A Systematic Framework for the Development and Analysis of Signed Digraphs for Chemical Processes. 1. Algorithms and Analysis, *Industrial & Engineering Chemistry Research*, 42, 20, pp. 4789–4810.
- F. Pelzer, et al., 2021, Safety in modular process plants: demonstration of safety concepts, *e & i Elektrotechnik und Informationstechnik*, 138, 7, pp. 462–468.
- C. Reinartz, et al., 2019, Generation of Signed Directed Graphs Using Functional Models, *IFAC-PapersOnLine*, 52, 11, pp. 37–42.
- V. Venkatasubramanian, et al., 2003, A review of process fault detection and diagnosis: Part I: Quantitative model-based methods, *Computers & Chemical Engineering*, 27, 3, pp. 293–311.
- V. Venkatasubramanian, R. Rengaswamy, and S. N. Kavuri, 2003, A review of process fault detection and diagnosis: Part II: Qualitative models and search strategies, *Computers & Chemical Engineering*, 27, 3, pp. 313–326.

**Acknowledgments:** This work was Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 319919706 /RTG2323.

# Performance analysis of three-phase fluidized bed absorber for CO<sub>2</sub> capture industrial application

Flavia-Maria Ilea, Ana-Maria Cormos, Simion Dragan, Calin-Cristian Cormos  
*Babes-Bolyai University, Faculty of Chemistry and Chemical Engineering, Chemical Engineering Department, 11 Arany Janos, RO-400028, Cluj-Napoca, Romania*

## Abstract

This work aims to analyze the performance achieved using a novel CO<sub>2</sub> capture technology, more specifically three-phase fluidized bed columns, at industrial level. For this purpose, a previously developed detailed mathematical model, validated against experimental data, was adapted to scale-up. It considers the hydrodynamics, mass transfer and chemical reaction that takes place within the process and as such is able to accurately describe the chemical absorption of CO<sub>2</sub> in three phase gas-solid-liquid systems. This system relies on the process intensification which is achieved due to the much higher value of the effective mass transfer area. This is obtained through the continuous renewal of the liquid film that forms on top of the solid particles that are in movement.

A higher scale column was considered with one to three stages, in order to see in which case, the fluidization is more favored. Five different scenarios were chosen that refer to the column configuration. The results show that due to the turbulence within the system, the division of the packing in two stages, with descending static bed height, leads to an 20% increase in carbon capture rate and a 25% decrease in pressure drop.

**Keywords:** CO<sub>2</sub> capture; Process intensification; Turbulent contact absorbers.

## 1. Introduction

### 1.1. Global context and literature overview

Due to the increasing demand of electricity that cannot be fulfilled by sustainable sources, fossil fuels remain the main source of energy in the foreseeable future. As a result, the quantity of carbon dioxide emissions continues to increase over the years, having peaked at an all-time high in 2022, despite the drop that took place in 2021 due to the global pandemic situation (IPCC reports, 2022). Taking into consideration the global warming phenomenon that takes place due to the accumulation of fossil CO<sub>2</sub> in the atmosphere, there are several ongoing European initiatives meant to work against drastic climate changes. Amongst them are some of the UN 2030 sustainability objectives and the goal for EU to be climate neutral by 2050 (UN reports, 2022). The aforementioned reasons, as well as recent events, have led to the price for pollution to increase substantially, having tripled in less than a year. The mid 2022 is represented by a price of over 90 euros per ton of carbon dioxide, which gives more reason to investing in the implementation and research regarding carbon capture technologies (Arcos-Vargas et al, 2022).

Existing capture technologies, involving the post-combustion absorption of CO<sub>2</sub> in amine solutions, still face some important disadvantages that refer not only to the energy penalty, but also to the costs implied by the regeneration and recycling of the solvents (Khosroabadi et al, 2021). As such, the search for new and improved technologies remains a top priority. Turbulent contact absorbers have been used mostly in the biochemical and petroleum fields but have yet to be analyzed regarding their process



intensification potential (Ul-Haq, 2012). This equipment unit is based on the use of low-density solid particles that are chemically inert, their purpose being the intensification of the mass transfer between the counter-current flows of gas and liquid.

### 1.2. Previous work

Previous experimental work (Ilea et al, 2022) shows that the use of three phase fluidization for carbon capture leads to an important increase of 3 – 6 times in the effective mass transfer area, which leads to a proportional increase in the CO<sub>2</sub> transferred flow. This helps to increase the gas treating capacity of the absorber, as well as it minimizes the quantity of solvent needed for the CO<sub>2</sub> capture process. Based on the experimental results, a thorough mathematical model for a pilot scale CO<sub>2</sub> capture column has been developed and validated to describe the process of turbulent contact absorption through its complex hydrodynamics and mass transfer between the liquid and the gas phases. As relevant innovative aspect, this work presents the adaptation of the previously developed model to industrial scale and analyzes several possible cases in order to find the best possible configuration and parametrization of the system. The solvent used is an aqueous solution of sodium hydroxide (as it was the solvent used in experimental work).

## 2. Scaled-up three-phase absorption column

### Experimental parametrization

The parametrization of the scaled-up column considered was done in accordance with the pilot scale column used for the experimental analysis. The ratio between the column diameter and the particle diameter needs to be above the value of 6 (in the considered cases of this study, the value varies between 10 and 20). The inlet flue gas flows considered are between 1000 and 4000 m<sup>3</sup>/h, while for the liquid spray density the values were kept as for the pilot scale column, maintain at the same the packing completely wetted. This was verified by the determination of the minimum liquid flow required to avoid flow at the periphery of the reactor. The density of the particles and the void fraction were determined experimentally using 0.075 m spherical particles. The model parameters are presented in *Table 1*.

**Table 1.** Process parameters

Parameter	Pilot scale column	Scaled-up column
Column diameter [m]	0.14	1
Total static bed height [m]	0.12	1.5
Solid particles diameter [m]	0.01	0.075
Solid particles density [kg/m <sup>3</sup> ]	170, 210, 330	10 – 100
Liquid flow [m <sup>3</sup> /h]	0.1, 0.2	10 – 50
Gas flow [m <sup>3</sup> /h]	60 – 120	1000 – 4000

## 3. Mathematical model

### 3.1. Mass and energy balance

In order to most accurately describe the process, apart from the aforementioned equations, the model includes mass and energy balance equations, as presented below:

$$F_j^e = F_j^0 \pm w_j \cdot \frac{N_{CO_2} \cdot M_{CO_2}}{\rho_j} \quad (1)$$

$$\frac{dC_{CO_2}^j}{dt} = \frac{F_j^0}{w_j} \cdot C_{CO_2}^0 - \frac{F_j^e}{w_j} \cdot C_{CO_2}^j \pm N_{CO_2} - v_r \quad (2)$$

$$\frac{dC_{NaOH}^l}{dt} = \frac{F_l^0}{w_l} \cdot C_{NaOH}^0 - \frac{F_l^e}{w_l} \cdot C_{NaOH}^l - v_r \quad (3)$$

$$\frac{dT_j}{dt} = \frac{F_j^0}{w_j} \cdot T_j^0 - \frac{F_j^e}{w_j} \cdot T_j - \frac{\Delta H_r \cdot v_r}{\rho_j \cdot c_{p_j}} \mp \frac{h \cdot a_e \cdot (T_l - T_g)}{(\rho_j \cdot c_{p_j} \cdot w_j)} \quad (4)$$

\*Note: *j* subscript refers to the gas and liquid phases respectively; *v<sub>r</sub>* refers to the reaction velocity

### 3.2. Hydrodynamics and mass transfer model

The performance investigation of the carbon capture process using turbulent contact absorbers was done using a previously developed model (Ilea et al, 2022) scaled up to industrial level. The model was updated in terms of the hydrodynamics of the process, including the comprehensive correlations that calculate the liquid holdup, column pressure drops and fluidized bed expansion. The mass transfer model is based on the two-film theory and Billet & Schultes (1993) correlations. In order to describe the performance of turbulent contact absorbers in terms of effective mass transfer area coefficient, an adjustment was performed. These correlations are presented below:

$$\frac{H_s}{H_0} = \left(\frac{D_c}{d_p}\right)^{0.8861} \cdot \left(\frac{\rho_s}{\rho_g}\right)^{-0.3346} \cdot \left(\frac{w_g}{w_{lmf}}\right)^{-0.1172} \cdot \left(\frac{w_l}{w_{lmf}}\right)^{0.2318} \quad (5)$$

$$\Delta P = \rho_l \cdot g \cdot H_0 \cdot \left(\frac{H_s}{H_0}\right)^{-1.7372} \cdot Re_g^{1.2535} \cdot Ga_l^{-0.4632} \cdot Re_l^{0.3897} \cdot We_l^{0.8760} \quad (6)$$

$$h_l = 0.001884 \cdot e^{0.237 \cdot w_g} \cdot (\rho_l \cdot w_l)^{0.616} \cdot H_s^{-0.357} \cdot d_p^{-0.411} \quad (7)$$

$$a_e = a \cdot e^{11.8967} \cdot \left(\frac{w_l^2}{d_e \cdot g}\right)^{-0.0152} \cdot Fr_g^{-0.2466} \cdot Ga_l^{-0.7174} \quad (8)$$

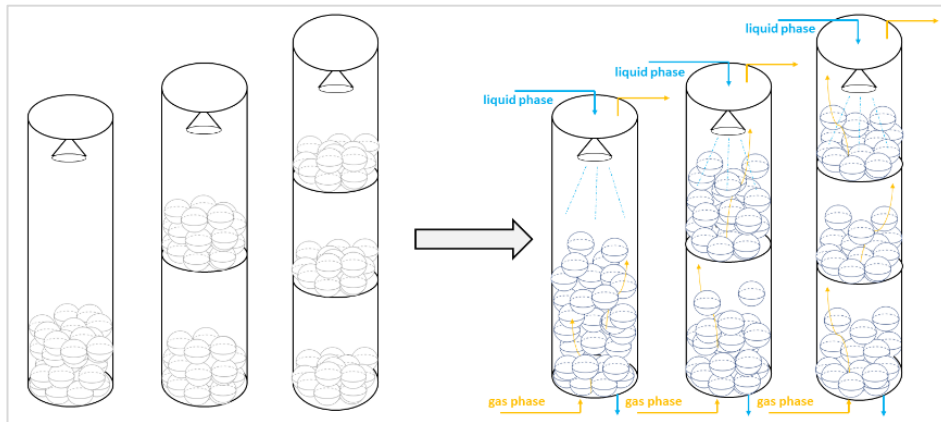
\*Note: *g* and *l* subscripts refer to the gas and liquid phases respectively

## 4. Results and discussion

The developed mathematical model was used to analyze the performance of different system configurations. The absorber was divided in either 2 or 3 sections or left undivided. In order to ensure the possibility of the fluidization of the solid phase, decreasing the static bed height with each section was also considered, as a response to the lowering of the gas velocity. The analysis was performed using Matlab/Simulink simulation software.

### 4.1. Considered cases

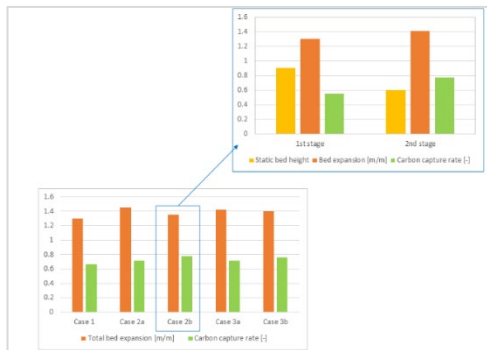
Taking into consideration the aforementioned information, 5 different cases were considered in this study. The benchmark case (Case 1) refers to a single stage column, with a static bed height of the packing of 1.5 m. The second case (Case 2) refers to a two-stage column and it is represented in two different configurations: Case 2a – in which the packing static bed heights are the same for both stages (0.75 m) and Case 2b – in which the packing static bed heights descend from the bottom to the top stage (0.9 m and 0.6 m). The third case (Case 3) refers to a three-stage column and it is represented as well in two different configurations: Case 3a – in which the packing static bed heights are the same for all stages (0.5 m) and Case 3b – in which the packing static bed heights decrease from the bottom to the top stage (0.7 m, 0.5 m and 0.3 m).



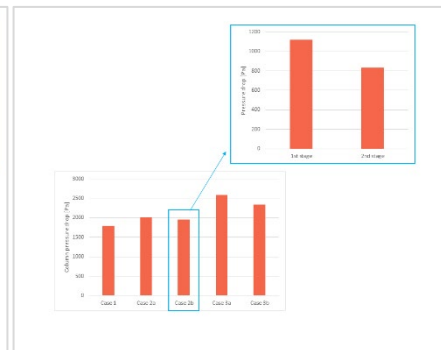
**Figure 1.** *CO<sub>2</sub> capture column configurations*

#### 4.2. Influence of bed expansion on carbon capture rate and pressure drop

The columns were designed as to have a total static bed height of 1.5 m regardless of the number of stages. In order to find the most suitable configuration of the column, the carbon capture rate was assessed for all the considered cases, dependent on the total fluidized bed expansion rate. This is calculated as the ratio between the sum of the expansions per each stage and the total static bed height. This is presented in *Figure 2*. The pressure drops for each considered case are presented in *Figure 3*.



**Figure 2.** *Carbon capture rate and bed expansion*

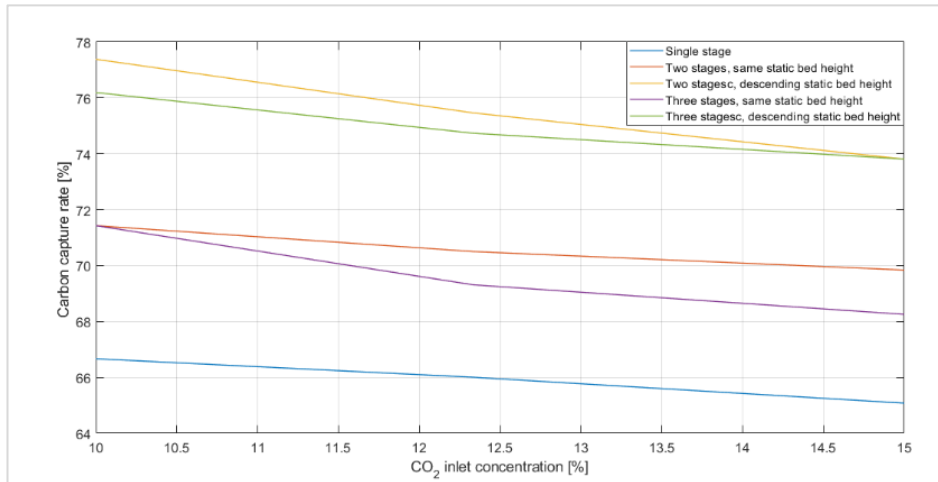


**Figure 3.** *Pressure drops*

As shown in *Figure 2*, the carbon capture rate varies between values of 65 and 80 %. For Case 2b, the highest value of capture rate is obtained (value close to the one obtained in Case 3b). Moreover, for Case 2b, a lower value of bed expansion can be observed. This proves that in order to achieve the intensification of the mass transfer between the liquid and the gas phases, a high gas velocity is not necessary. In fact, a value slightly above the value of the minimum fluidization velocity is enough to reach the turbulence level needed to get the movement of the solid particles and, as such, the renewal of the liquid film. Due to the decreasing in static bed height with the stages and compared to the single stage column, in Case 2b a lower value of the pressure drop is also achieved, it staying at values around 1850 Pa. This is significantly lower than any other packed bed absorption columns, which is an important advantage of using turbulent contact absorbers.

#### 4.3. Influence of inlet CO<sub>2</sub> concentration on carbon capture rate

The inlet carbon dioxide concentration proved to have an important impact on the carbon capture rate. An analysis was made, ranging the concentration values between 10 % and 15 %. The results are presented in *Figure 4*.

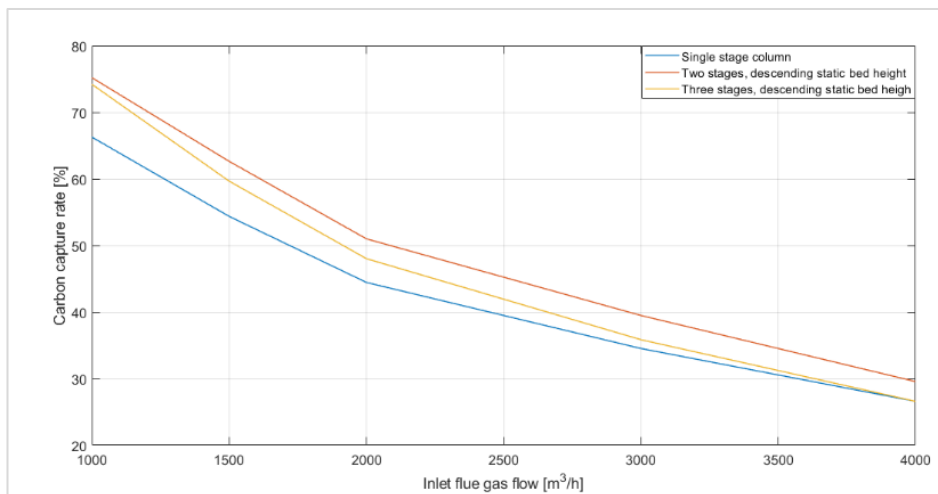


**Figure 4.** Influence of inlet CO<sub>2</sub> concentration on carbon capture rate

It can be observed that the best performance is obtained for Case 2b (two-stage column with descending static bed height). This is due to the fact that in this case, the expansion of the fluidized bed allows for the higher values of turbulence and as such, a higher value of the effective mass transfer area is obtained, leading to the intensification of the process.

#### 4.4. Influence of gas flow on carbon capture rate

The value of the flue gas flow was varied between values of 1000 and 4000 m<sup>3</sup>/h. The results in terms of carbon capture are presented in *Figure 5*, for three of the cases.



**Figure 5.** Influence of flue gas flow on carbon capture rate

The chosen cases are the one with descending static bed height per stage, as these cases shown better results in the previous analysis. The results show once again that the higher value of the carbon capture rate is obtained in Case 2b (two stage column with descending static bed height (increased by 10 %). They also show a decrease with the increase of gas flow. This is due to the increase in the void fraction of the fluidized bed, which implicitly leads to a decrease of turbulence. As such the liquid film is not as much renewed as in the case of lower gas velocities, velocities closer to the minim fluidization velocity.

## 5. Conclusions

The presented work analyses the performance of a turbulent contact absorber using a scaled up mathematical model. The absorber is configured as having one to three stages. Five cases are considered that refer to the column having one, two or three stages. The multiple staged ones are configured in two different ways, with the static bed height per stage being equal or descending from the bottom to the top section. The results show that the highest performance is achieved in Case 2b (two stage column with descending static bed height), obtaining a capture rate 20 % higher than in the benchmark case (Case 1 – one stage column). In this case, the pressure drop also decreases with about 25 % compared to the benchmark case. It is concluded that the positive results in this case are due to the high turbulence that is found within the column, which prompts the intensification of the mass transfer between the liquid and the gas phases.

## Acknowledgements

This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI – UEFISCDI, project number PN-III-P4-ID-PCE-2020-0632, within PNCDI III.

## References

- A. Arcos-Vargas, F. Núñez-Hernández, J.A. Ballesteros-Gallardo, 2022, CO<sub>2</sub> price effects on the electricity market and greenhouse gas emissions levels: an application to the Spanish market, *Clean Technologies and Environmental Policy*.
- R. Billet, M. Schultes, 1993, Predicting mass transfer in packed columns, *Chemical Engineering and Technology*, 16, 1-9.
- F.M. Ilea, A.M. Cormos, S. Dragan, C.C. Cormos, 2022, Assessment of turbulent contact absorber hydrodynamics with application in carbon capture, *Chemical Engineering Journal*, 449, 137674.
- Intergovernmental Panel on Climate Change (IPCC), 2022, Sixth Assessment Report – Mitigations on Climate Change.
- F. Khosroabadi, A. Aslani, K. Bekhrad, Z. Zolfaghari, 2021, Analysis of carbon dioxide capturing technologies and their technology developments, *Cleaner Engineering and Technology*, 5, 100279.
- A. Ul-Haq, 2012, Mass Transfer in a Turbulent Contact Absorber, PhD Thesis, Department of Chemical Engineering Pakistan Institute of Engineering and Applied Sciences Islamabad, Pakistan.
- United Nations (UN), 2022, Climate Action Reports, <https://www.un.org/en/climatechange/reports>.

# Decentralized control for optimal operation under changing active constraints

Lucas F. Bernardino,<sup>a</sup> Sigurd Skogestad<sup>a</sup>

<sup>a</sup> *Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway*  
*skoge@ntnu.no*

## Abstract

Optimal economic operation of chemical plants requires control of active constraints, which may change because of disturbances. In addition, one should ensure optimality with respect to the unconstrained degrees of freedom by driving the reduced cost gradient to zero. One solution is to use on-line optimizing control, but the preferred approach in industry is to use decentralized control and selectors whenever possible. In this paper, we consider a new framework based on identifying the cost gradient projections that can be left uncontrolled when each specific constraint becomes active, leading to a decentralized logic. The proposed framework is applied to the optimal operation of the Williams-Otto reactor, which has two degrees of freedom and two constraints. The projection matrix, which depends on the gradient of the constraints, is assumed constant, resulting in a simple control structure. The approach works well in simulations and it switches between the four active constraint regions.

**Keywords:** optimal operation, feedback optimizing control, decentralized control, selectors.

## 1. Introduction

In spite of the potential economic benefits, real-time optimization (RTO) is less used in industry than one may expect. There are several reasons for this, and one is that standard RTO applications operate on a slow time scale, such that for cases with frequent disturbances, optimal operation is not satisfactorily achieved. An appealing solution is to move some of the optimization problem into the control layer design, which operates on a much faster time scale. This means that one should find controlled variables (CVs) that, when controlled to constant setpoints, result in minimal economic loss. This is the idea of self-optimizing control (Skogestad, 2000), which consists of obtaining these CVs as a combination of the available measurements. Another class of feedback optimizing strategies aim to estimate the plant cost gradients in order to control them to zero (Krishnamoorthy et al., 2019).

One of the main challenges that such approaches face is the presence of changing active constraints (Jäschke et al., 2017), which may drastically change the operation mode of the system. For example, if a constraint becomes active due to a disturbance, not taking its control into account leads to infeasible operation. Similarly, if a constraint becomes optimally inactive, the control of such constraint should be given up. For a given set of active constraints, if the cost gradient is measured, it is known that the control of the active constraints together with the control of a projection of the cost gradient over the

nullspace of the active constraints' gradient leads to optimal operation (Jäschke and Skogestad, 2012). However, the existence of several active constraint regions may deem necessary the use of several independent control structures, each of them being able to provide near-optimal operation for their respective region of design. The switching between such control structures would also become an issue since the lack of a feedback-based switching strategy could lead to improper operation.

In practical applications of process systems, logic elements have been extensively used for reconfiguring control loops, which is often needed for attaining optimal operation (Reyes-Lúa et al., 2018). In particular, selectors have been successfully applied as a tool for automatic detection of active constraint switching for single input systems (Krishnamoorthy and Skogestad, 2020). In this work, we extend this analysis for multivariable systems, proposing a simple framework for decentralized optimal operation under changing active constraints, with the use of PID controllers and selectors.

## 2. Methodology

The main idea of this framework consists in identifying the optimal CVs for each region, and consequently proposing control loops for dealing with such CVs, in a way that reconfiguring is minimized. Consider the following convex optimization problem:

$$\begin{aligned} & \min J(u, d) \\ \text{s. t.} \quad & g_i = g_{u,i}^T u + g_{d,i}(d) \leq 0, \quad i = 1, \dots, n_g \end{aligned} \quad (1)$$

For the purposes of this work, we consider the steady-state cost gradient  $J_u$  to be known. Based on that, and assuming that  $n_u \geq n_g$  and all vectors  $g_{u,i}$  are linearly independent (LI), we can devise a simple control strategy for optimal operation. Firstly, the unconstrained degrees of freedom related to the nullspace of  $G_u = [g_{u,1} \ \dots \ g_{u,n_g}]^T$  will always be optimally controlled regardless of how many constraints are active, and therefore  $N_0^T J_u$ , with  $N_0$  being a basis for the nullspace of  $G_u$ , should always be controlled to zero. In addition to that, if all but one constraint is active, an extra unconstrained degree of freedom needs to be considered. The degree of freedom freed by the constraint  $g_i$  becoming inactive can be determined by the nullspace of the matrix  $G_{u,-i}$  comprised of all the remaining rows  $g_{u,j}^T$ ,  $j \neq i$ . By definition, this nullspace will include the space generated by  $N_0$ , and it is sufficient to pick any projection  $N_{g,i}$  that is LI from  $N_0$ . A unique solution can be obtained by picking  $N_{g,i}$  orthogonal to  $N_0$ .

With these definitions, we propose the following control strategy for a given active constraint set  $\mathcal{A}$ :

- if  $n_u > n_g$ , find  $N_0$  such that  $G_u N_0 = 0$  and control  $N_0^T J_u = 0$ ;
- for  $i = 1, \dots, n_g$ :

- find  $N_{g,i}$  such that  $\begin{bmatrix} G_{u,-i} \\ N_0^T \end{bmatrix} N_{g,i} = 0$ ;
- control  $g_i = 0$  if  $i \in \mathcal{A}$ ; else, control  $N_{g,i}^T J_u = 0$ .

It can be verified that the operating point defined by the forementioned CVs leads to the solution of (1). Furthermore, if decentralized PID control is used for every CV, the decision of controlling  $g_i = 0$  or  $N_{g,i}^T J_u = 0$  can be performed locally by comparing the corresponding control loops. This problem may be solved with selectors, such that constraint control becomes active when necessary, and the unconstrained degree of freedom is controlled whenever the constraint is not violated. If the system is such that

decentralized PID control can be used for every set of CVs, the framework will lead to optimal operation.

The proposed framework considers linear constraints with respect to the plant inputs, which is a strong assumption that is not accurate for most real systems. We will now demonstrate the effectiveness of the framework in a case study with nonlinear constraints, evaluating the loss that is obtained by the application of the proposed method.

### 3. Results

#### 3.1. Case study description

We now consider the optimal operation of the Williams-Otto reactor (Williams and Otto, 1960). The optimal operation of the reactor is described by:

$$\begin{aligned} \min_u J^{ec} &= p_A F_A + p_B F_B - (F_A + F_B)(p_E x_E + p_P(1 + \Delta p_P)x_P) \\ \text{s. t.} \quad g_1 &= x_A - 0.12 \leq 0 \\ g_2 &= x_E - 0.3 \leq 0 \end{aligned} \quad (2)$$

The degrees of freedom for operation are  $u = [F_B, T_R]$ , with  $T_R$  being the reactor temperature, and the considered disturbances are  $d = [F_A, \Delta p_P]$ . The solution of the above optimization problem as a function of the disturbances leads to the pattern shown in Figure 1, where each region correspond to the set of active constraints at the solution.

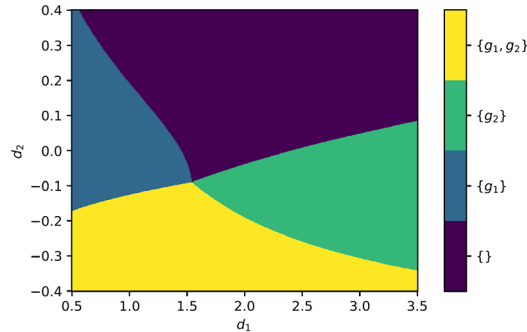


Figure 1: Optimally active constraints as function of the disturbances

#### 3.2. Control structure design

In order to apply the proposed framework, we linearize the constraints at a nominal point. With this, we assume a constant constraint gradient with respect to the inputs, and this ultimately results in constant cost gradient projections to be controlled inside each region. The nominal point was chosen to be the optimal point for  $d^* = [1.0 \ 0.0]$ , and the linearization of the constraints at this point results in:

$$\Delta g = \begin{bmatrix} g_{u,1}^T \\ g_{u,2}^T \end{bmatrix} \Delta u + g_d \Delta d = \begin{bmatrix} -0.0492 & 0.0032 \\ -0.0328 & -0.0026 \end{bmatrix} \Delta u + g_d \Delta d \quad (3)$$

Since  $n_g = n_u = 2$ , there are no unconstrained degrees of freedom that remain active at the fully constrained case, when  $\mathcal{A} = \{g_1, g_2\}$ . Therefore, no cost gradient projection  $N_0^T J_u$  is needed, and all operational degrees of freedom are filled by active constraint control. For  $\mathcal{A} = \{g_1\}$ , in addition to controlling  $g_1 = 0$ , we must control one cost



gradient projection  $N_{g,2}^T J_u = 0$  in order to fill the remaining degree of freedom, where  $N_{g,2}$  is chosen such that  $G_{u-2} N_{g,2} = g_{u,1}^T N_{g,2} = 0$ . Similarly, in the region where  $\mathcal{A} = \{g_2\}$ , the optimal CVs will be  $g_2 = 0$  and  $N_{g,1}^T J_u = 0$ , where  $N_{g,1}$  is chosen such that  $G_{u-1} N_{g,1} = g_{u,2}^T N_{g,1} = 0$ . In the unconstrained region  $\mathcal{A} = \{\}$ , the optimal CVs are all the components of the cost gradient  $J_u = 0$ . However, controlling  $N_{g,1}^T J_u = 0$  and  $N_{g,2}^T J_u = 0$  simultaneously implies in  $J_u = 0$ , as the constraints are independent and therefore  $[N_{g,1} \ N_{g,2}]$  is full rank. This means that the same gradient projections used for the partly constrained regions can be used for the fully unconstrained region.

The control structure that results from the application of the proposed methodology is presented in Figure 2, where  $K_{g_i}$  and  $K_{c_i}$  denote PID controllers related to control of constraints and gradient projections, respectively. It can be seen that control of  $N_{g,1}^T J_u$  can be optimally given up when the control of  $g_1$  becomes active, and the same happens with the pair  $N_{g,2}^T J_u$  and  $g_2$ . This pairing results in optimal operation for all possible active constraint regions, with the switching being performed by max selectors on the controller outputs.

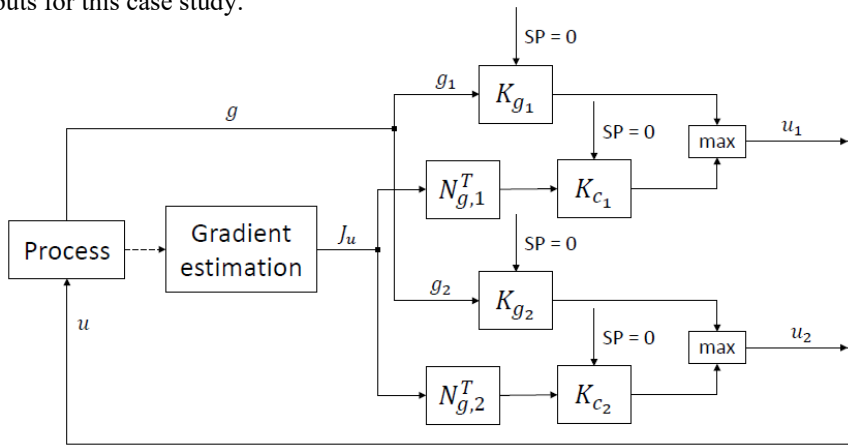


Figure 2: Control structure for optimal operation of case study

### 3.3. Simulation results

The behavior of the proposed control structure is illustrated by the simulations presented in Figure 3. The disturbance realizations were chosen such that the system operates at all active constraint regions. In the simulations,  $J_u$  was obtained using the automatic differentiation tools from CasADi (Andersson et al., 2019) with known disturbances. In practice, an estimator would be needed if the disturbances are not measured.

The system starts at the fully unconstrained region, where perfect optimal operation is possible because the cost gradient,  $J_u$ , is known. Once the system moves to the region  $\mathcal{A} = \{g_1\}$  at  $t = 4$  h, constraint violation is avoided as  $g_1 = 0$  becomes the CV chosen by the selector. Even though the inputs are not driven to their optimal value, the choice of CVs is such that low economic loss is achieved. At  $t = 6$  h, the disturbance value is equal to the nominal point, where the linearization of the constraints was performed. For this reason, the controlled cost gradient projection  $N_{g,2}^T J_u$  corresponds exactly to the optimality conditions, and perfect optimal operation is attained. At  $t = 8$  h, the system starts operating at the region  $\mathcal{A} = \{g_1, g_2\}$ . As all degrees of freedom are associated to active constraint control, perfect optimal operation is achieved. From  $t = 12$  h, the system operates at  $\mathcal{A} = \{g_2\}$ , and near-optimal operation is achieved with low economic loss.

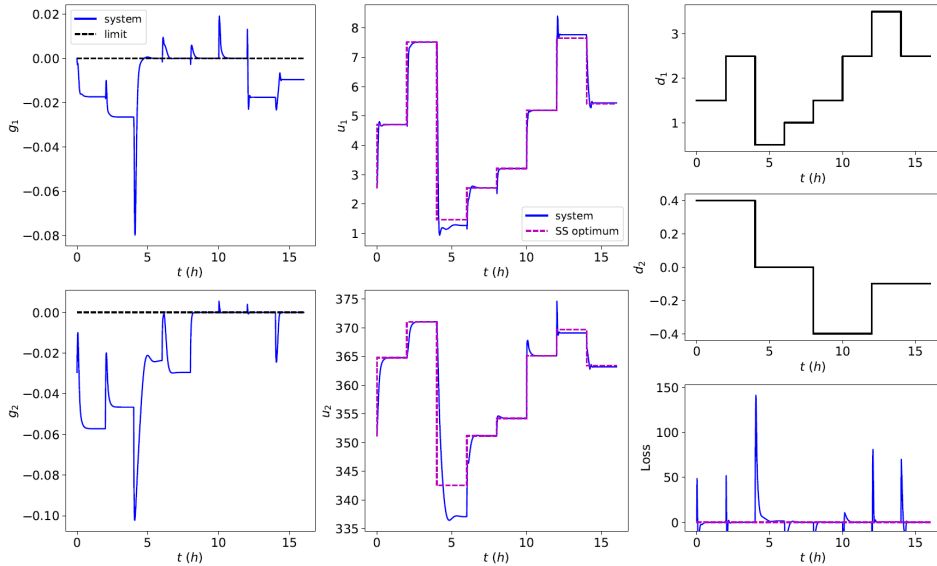


Figure 3: Dynamic simulation of proposed control structure

The steady-state economic loss for the proposed control structure is presented in Figure 4 as a function of the disturbances. Similar to what was observed in the dynamic simulations (Figure 3), the regions with nonzero loss are concentrated at the partly constrained regions. This is because the optimal CVs corresponding to the cost gradient projections change due to the nonlinearity of constraints. In contrast, the fully constrained and fully unconstrained regions mostly present zero operational loss, because the optimal CVs remain constant inside these regions, and the information about  $J_u$  is accurate. The linearization of the constraints was performed inside the region  $\mathcal{A} = \{g_1\}$ , where the operational loss is effectively zero at this reference point, and it increases as the system moves away from it. It can also be seen that optimal switching between regions is subject to errors related to the linearization of the constraints. The largest operational loss is obtained in the switching between regions  $\mathcal{A} = \{g_1, g_2\}$  and  $\mathcal{A} = \{g_2\}$ , where an inaccurate cost gradient projection is tracked, that is, controlling  $N_{g,1}^T J_u = 0$  does not lead to exact optimal operation. Therefore, the switching is not performed at the exact optimal boundary, as can be seen in the results, but it nonetheless leads to a reasonable switching policy between regions, and good resulting economic performance.

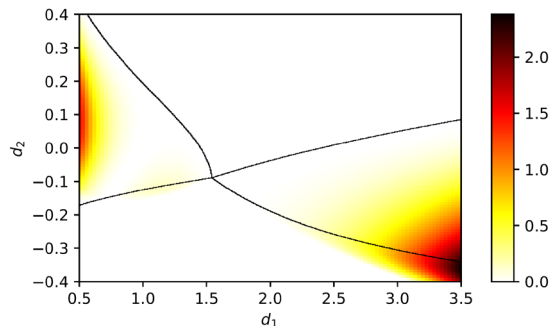


Figure 4: Steady-state economic loss of the proposed control structure as a function of disturbances (black lines represent optimal region switching)

#### 4. Discussion

The control structure resulting from the framework proposed in this work (Figure 2) makes use of selectors as simple elements for switching between operating regions. These elements are frequently used in practice for coordinating conflicting control objectives. In this work, we show how these elements can be used for the optimal operation of systems under changing active constraints in a systematic manner. The min/max nature of the selectors is ultimately related to the nature of the constraint with respect to the input (Krishnamoorthy and Skogestad, 2020).

The approach proposed in this work is based on the analysis of the partly constrained regions, where we find gradient projections that can be optimally controlled. These projected gradients are also the optimal CVs of other regions, which minimizes the number of control loops that are necessary to account for all regions. Moreover, the proposed switching operates independently for each plant input, which means that the detection of each constraint is done independently, and feasible operation is safely achieved. However, the reconfiguring of CVs done by the selectors may significantly change the interactions between loops, and therefore careful tuning of the controllers is necessary, such that a good performance is achieved regardless of which loops are active.

#### 5. Conclusion

In this work, we propose a framework for decentralized optimal operation under changing active constraints, applicable to a class of multivariable problems. Even though the approach is based on the linearization of the constraints, and therefore the quality of the linearization plays a relevant role in the economic performance, the strategy proved to be successful in a nonlinear case study, which encourages its use in other relevant problems of process systems engineering, especially when the gains from the inputs to the constraints do not change greatly in the operating range. The use of adaptive cost gradient projections would also be beneficial for improving economic performance. More theoretical aspects regarding the proposed framework will be expanded in future work.

#### References

- J.A. Andersson, J. Gillis, G. Horn, J.B. Rawlings, M. Diehl, 2019. CasADi: a software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation*, 11, pp.1-36.
- J. Jäschke, Y. Cao, V. Kariwala, 2017. Self-optimizing control—A survey. *Annual Reviews in Control*, 43, pp.199-223.
- J. Jäschke, S. Skogestad, 2012. Optimal controlled variables for polynomial systems. *Journal of Process Control*, 22, 1, pp.167-179.
- D. Krishnamoorthy, E. Jahanshahi, S. Skogestad, 2018. Feedback real-time optimization strategy using a novel steady-state gradient estimate and transient measurements. *Industrial & Engineering Chemistry Research*, 58, 1, pp.207-216.
- D. Krishnamoorthy, S. Skogestad, 2020. Systematic design of active constraint switching using selectors. *Computers & Chemical Engineering*, 143, p.107106.
- A. Reyes-Lúa, C. Zotică, S. Skogestad, 2018. Optimal operation with changing active constraint regions using classical advanced control. *IFAC-PapersOnLine*, 51, 18, pp.440-445.
- S. Skogestad, 2000. Plantwide control: The search for the self-optimizing control structure. *Journal of process control*, 10, 5, pp.487-507.
- T.J. Williams, R.E. Otto, 1960. A generalized chemical processing model for the investigation of computer control. *Transactions of the American Institute of Electrical Engineers, Part I: Communication and Electronics*, 79, 5, pp.458-473.

# Development of a Deep Learning-based Schedule-Aware Controller: Toward the Integration of Scheduling and Control

M. Abou El Qassime<sup>a,b</sup>, A. Shokry<sup>\*a</sup>, A. Espuña<sup>c</sup>, E. Moulines<sup>a,b</sup>

<sup>a</sup>*École polytechnique, Paris, France*

<sup>b</sup>*Emines School of Industrial Management, BenGuerir, Morocco*

<sup>c</sup>*Universitat Politècnica de Catalunya, Barcelona, Spain.*

[\\*ahmed.shokry@polytechnique.edu](mailto:*ahmed.shokry@polytechnique.edu)

## Abstract

Most of the existing approaches for integrating the scheduling and control layers in the process industry suffer from the solution complexity of the resulting Mixed Integer Nonlinear Programming (MINLP) program. This complexity stems from two main obstacles: i) the consideration of nonlinear closed-loop dynamics involved in the control layer and ii) the way of exchanging the information between the control and the scheduling layers. To tackle these two obstacles, this work proposes the use of a schedule-aware controller (SAC) based on deep learning (DL) models, which receives, as input, the scheduling information (i.e., production sequence, targets, and set points) along with the process conditions (values of the state, output, and control variables) to predict the optimal control actions to be applied at the next sampling periods. The method is applied to a benchmark, showing very good performance in terms of the ability to adapt the control actions in accordance with the scheduling decisions, high prediction accuracy, and reduction in the computation cost.

**Keywords:** Explicit Control- Scheduling- Integration - Machine Learning

## 1. Introduction

In process industry, the integration of decision making across different layers of the process management hierarchy is of paramount importance. In particular, integrating the scheduling and control layers can lead to significant increase in production profitability, operational reliability, and adaptability to varying environments. Most of the works that have been developed to integrate control and scheduling can be categorized into simultaneous and decoupling approaches. Flores and Grossmann (2006, 2010) proposed a simultaneous approach incorporating the nonlinear process dynamics (involved in the control layer) into the scheduling problem in the form of differential/algebraic constraints, which results in a Mixed Integer Dynamic Optimization (MIDO) problem that handles the control and scheduling simultaneously. They used a discretization method based on orthogonal collocation to transform the MIDO into a MINLP problem. Zhuge and Ierapetritou (2012) proposed a similar method able to handle online disturbances in the system dynamics. They solve the MIDO problem offline to obtain reference production sequence and control variables, which are initially implemented online. Then, once disturbances occur, the MIDO problem is updated and solved over the remaining time horizon. Conversely, the decoupling approach considers two problems: i) the scheduling

one (master), which is formulated as a MINLP and solved based on an economic objective to obtain the mid-term decisions, and ii) the control problem (primal), which is first informed by the scheduling decisions and then solved as dynamic optimization to obtain the short-term decisions. The master-primal problems are solved iteratively until convergence is achieved. Burnak et al. (2018) proposed a decoupling approach based on multi-parametric (MP) programming that consists of i) a MP scheduler, ii) a MP controller based on Model Predictive Control (MPC), and iii) a multiparametric “bridge” that connects these two layers. Martinez and Sandoval (2022) developed a nested approach that solves the scheduling problem to obtain the product switching sequence and the associated set points, which are passed to the MPC to implement them. The method is applied in a closed-loop online mode to minimize perturbations. Although both approaches demonstrated the importance of such integration for economic gain, they still have some drawbacks. In the simultaneous approach, the resulting MIDO problem is usually of large scale and can be intractable, especially in cases where the process model is of high-fidelity and/or highly nonlinear. The decoupling approach leads to more manageable optimization problems, but a significant computational effort is required to iteratively solve the control problem (Burnak et al., 2018; Martinez and Sandoval, 2022). The drawbacks of both approaches result from two main issues: i) handling the nonlinear closed-loop dynamics, and ii) the need to exchange the information between the control and the scheduling layers. To tackle the two obstacles, we propose a novel DL -based SAC that receives, as input, the med-term scheduling information (production targets, sequence, and set points) along with the instantaneous process information (state, output, and control variables) and provides the optimal control actions to be applied to the process in the future sampling period. This novel DL-based SAC represents: a) an accurate and computationally cost-effective controller (tackling the first obstacle) and, b) a controller that accommodates the scheduling decisions, adapts the control actions to pursuit these scheduling targets, and, finally, quantifies the deviation from these targets (e.g., material loss). These deviations can be sent back to the scheduler (tackling the second obstacle). Hence, this DL-based SAC could be integrated seamlessly with a higher-level scheduler.

## 2. Problem statement

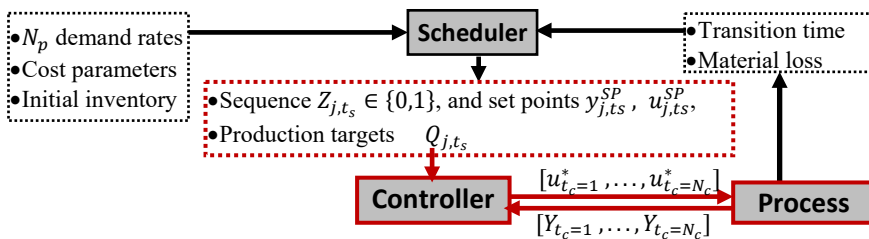


Figure 1. Representation of the framework for integrating scheduling and control.

We consider a framework for integrating the control and scheduling (Figure 1), in which both layers share information. Given  $N_p$  products to be manufactured, a demand scenario for each product over of a total time period  $T_{tot}$ , and a cost function, the scheduling problem is solved to obtain the best the production sequence  $Z_{j,t_s} \in \{0,1\}$  with the associated set points  $y_{j,t_s}^{SP} \in R^{n_y}$ ,  $u_{j,t_s}^{SP} \in R^{n_u}$  and target production rates  $Q_{j,t_s}$ , where  $j =$

$1, \dots, N_p$  and  $t_s = 1, \dots, N_s = t_{tot}/\Delta t_s$ , and  $\Delta t_s$  is the scheduling time slot. Please note that the solution of the scheduling problem itself is out of the scope of this work, so only the elements highlighted in red in Figure 1 are considered.

The scheduling decisions (bold font in the formulation below) are fed to the control layer based on MPC (Eqs. (1):(4)). We denote  $t_c = 1, \dots, N_c$  the control sampling times with  $N_c = t_{tot}/\Delta t_c$  and  $\Delta t_c \ll \Delta t_s$ . At the control level, the scheduling information take constant values within the same scheduling time slot. A nonlinear high-fidelity model,  $\mathcal{F}$  and  $\mathcal{G}$  (Eq. (2)), is used to calculate the process states  $x_{t_c} \in R^{n_x}$  and outputs  $y_{t_c} \in R^{n_y}$ , and a case-dependent constraint (Eq. (3)) is considered to correlate the control actions  $u_{t_c} \in R^{n_u}$  with the production rates advised by the scheduler,  $Q_{j,t_c}$ . Denoting  $P$ ,  $R$  and  $R_1$  positive definite matrices, and  $N_{ph}$ ,  $M_{ch}$  the prediction and control horizons, respectively, the control problem can be described as follows:

$$\min_{u_{t_c+1}, \dots, u_{t_c+N_{ph}}} J = \sum_{k=1}^{N_{ph}} (y_{t_c+k} - y_{t_c+k}^{SP})^T P (y_{t_c} - y_{t_c}^{SP}) + \sum_{k=1}^{M_{ch}} (u_{t_c+k} - u_{t_c+k}^{SP})^T R (u_{t_c+k} - u_{t_c+k}^{SP}) + \sum_{k=1}^{M_{ch}} \Delta u_{t_c+k}^T R_1 \Delta u_{t_c+k} \quad (1)$$

s. t.

$$x_{t_c+k+1} = \mathcal{F}(x_{t_c+k}, u_{t_c+k}, d_{t_c+k}^T), \quad y_{t_c+k} = \mathcal{G}(x_{t_c+k}, u_{t_c+k}, d_{t_c+k}^T) \quad (2)$$

$$u_{t_c+k} \leq \mathcal{H}(Q_{j,t_c+k}, x_{t_c+k}, u_{t_c+k}, d_{t_c+k}^T) \quad (3)$$

$$x_{min} \leq x_{t_c+k} \leq x_{max}, \quad y_{min} \leq y_{t_c+k} \leq y_{max}, \quad u_{min} \leq u_{t_c+k} \leq u_{max} \quad (4)$$

$$\Delta u_{t_c+k} = u_{t_c+k} - u_{t_c+k-1}$$

At each time  $t_c$ , the measured values of the state and output variables ( $x_{t_c}$  and  $y_{t_c}$ ) are used as initial conditions for the dynamic optimization problem (Eqs. (1):(4)), which is solved to find the optimal trajectory of the control variables  $[u_{t_c+1}^*, \dots, u_{t_c+N_{ph}}^*]$ . Only the optimal control values corresponding to the first instance,  $u_{t_c+1}^*$ , are implemented in the process. The final solution of the MPC problem are the closed-loop control profiles along the entire production horizon  $[u_{t_c=1}^*, \dots, u_{t_c=N_c}^*]$  and the resulting state and output variables  $[x_{t_c=0}, \dots, x_{t_c=N_c-1}]$  and  $[y_{t_c=0}, \dots, y_{t_c=N_c-1}]$ . Since the solution complexity of this schedule-aware MPC problem could hinder the integration of the scheduling and control layers, we propose a DL-based SAC to cheaply and accurately predict the closed-loop control actions as functions of the scheduling decisions and the process conditions, as in Eq. (5), where  $l$  is a time delay and  $\mathcal{f}$  is deep learning model.

$$\hat{u}_{t_c+1}^* = \mathcal{f}(Q_{j,t_c}, \dots, Q_{j,t_c-l}, y_{j,t_c}^{SP}, \dots, y_{j,t_c-l}^{SP}, u_{j,t_c}^{SP}, \dots, u_{j,t_c-l}^{SP}, x_{t_c}, \dots, x_{t_c-l}, y_{t_c}, \dots, y_{t_c-l}, u_{t_c}, \dots, u_{t_c-l}) \quad (5)$$

### 3. Methodology steps and techniques

#### 3.1. Sampling of the scheduler outcomes

In the first step,  $N_{sc.tr}$  different scheduling scenarios are randomly synthesized over the scheduling slots  $t_s = 1, \dots, N_s$ . For each scenario  $i$ ,  $i = 1, \dots, N_{sc.tr}$  a different production sequence  $Z_{j,t_s,i} \in \{0,1\}$  (with the associated set points  $y_{j,t_s,i}^{SP}$ ,  $u_{j,t_s,i}^{SP}$ ) and/or production rates  $Q_{j,t_s,i}$  (Figure 2) is defined. The generated scenarios should contain most of the expected transitions between the  $N_p$  products and cover the full range of variability in production targets.

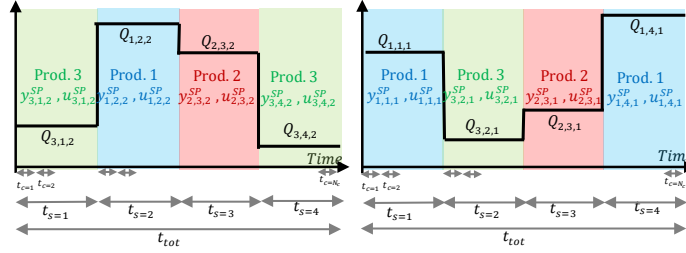


Figure 2. Example of two different scheduling scenarios where  $N_p = 3$  and  $N_s = 4$ .

### 3.2. Data generation: solving the MPC problem

The scheduling scenarios  $[Q_{j,t_s,i}]$ ,  $[y_{j,t_s,i}^{SP}, u_{j,t_s,i}^{SP}]$  are resampled to the control time scale to obtain the required values at the finer time grid  $t_c = 1, \dots, N_c$ , i.e.,  $[Q_{j,t_c=1,i}, \dots, Q_{j,t_c=N_c,i}]$ ,  $[y_{j,t_c=1,i}, \dots, y_{j,t_c=N_c,i}^{SP}]$ ,  $[u_{j,t_c=1,i}, \dots, u_{j,t_c=N_c,i}^{SP}]$ . Then, for each of the  $N_{sc.tr}$  scenario, the control problem (Section 2 and Eqs. (1):(4)) is solved to obtain the closed-loop control profiles along the entire production horizon  $[u_{t_c=1,i}^*, \dots, u_{t_c=N_c,i}^*]$  and the resulting state variables  $[x_{t_c=0,i}, \dots, x_{t_c=N_c-1,i}]$  and output variables  $[y_{t_c=0,i}, \dots, y_{t_c=N_c-1,i}]$ . The obtained trajectories of the state, output and control variables are unfolded together with the production targets and set points to an input-output data set for training the DL model, depending on the required input lag  $l$ . For example, in case of  $l = 0$ , the input-output dataset will be in the form  $[x_{t_c,j}, y_{t_c,j}, u_{t_c,j}^*, y_{j,t_c,j}^{SP}, u_{j,t_c,j}^{SP}, Q_{j,t_c,j}] - [u_{t_c+1,j}^*]$ ,  $j = 1, \dots, N_{sc.tr} \times N_c$ . Following the same steps, a testing dataset with  $N_{sc.tst}$  additional scheduling scenarios is also generated.

### 3.3. Development of DL-based SAC

The unfolded training data is used to fit the DL-SAC in Eq. (5), where  $\mathcal{f}$  is selected as a feedforward deep neural network (DNN), whose architecture (number of hidden layers, hidden neurons, activation functions), learning parameters, and training algorithm are chosen by a cut and try procedure that balance the DNN complexity and accuracy.

### 3.4. Testing of the DL-SAC via closed-loop deployment

In the final step, the developed DL-based SAC is used to predict the closed-loop control trajectories corresponding to the scheduling scenarios in the testing set. In each scenario, the DL-based SAC is used in a recursive way as follows:

- i. if  $t_c = 1$  consider the given initial values of the state, output, and control variables, else-if  $t_c > 1$  measure the current values of the state, output, and control variables from the process (high-fidelity model, in our case). Combine them with the target production rates and set points values (scheduling decisions), and use them as input for the SAC to predict the optimal values of the control variables at the next sampling period,
- ii. implement the predicted control variables values in the process (i.e., in the model),
- iii. if  $t_c = N_c$  stop, otherwise back to step i.

The accuracy of DL-based SAC is quantified by comparing the predicted control trajectories and the resulting state and output variables with their exact values. Normalized Root Mean Square Error (NRMSE) is used as accuracy metric.

#### 4. Application: Multi-product CSTR

As a benchmark, we consider an isothermal CSTR (Figure 3-(a)) which is fed by three reactants ( $R_i, i = 1,2,3$ ) and is operated to manufacture three products ( $P_j, j = A, B, C$ ) on a single production line (Flores and Grossmann, 2006; Burnak et al., 2018). In the CSTR, three irreversible reactions take place in parallel, see Figure 3-(b).

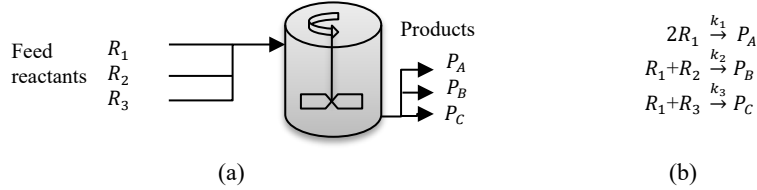


Figure 3: Representation of the multi-product CSTR (a) and its irreversible reactions (b).

The dynamics of the CSTR can be described by six differential equations, as follows:

$$\frac{dC_{R_i}}{dt} = \frac{q_{R_i} C_{R_i}^f - q_{tot} C_{R_i}}{V} + \mathcal{R}_{R_i} \quad \text{and} \quad \frac{dC_{P_j}}{dt} = \frac{q_{tot} (C_{P_j}^f - C_{P_j})}{V} + \mathcal{R}_{P_j} \quad (6)$$

Where  $C_{R_i}$  (state variables) are the concentrations of the reactants inside the CSTR (mol/L),  $C_{P_j}$  (output variables) are the products concentrations (mol/L), and  $q_{R_i}$  (control variables) are the reactants feed flowrates (L/min). Notice that  $q_{tot} = \sum_{i=1}^3 q_{R_i}$  is the total feed flowrate,  $V$  is the reactor volume (L), while  $C_{P_j}^f$  and  $C_{R_i}^f$  represent the concentrations in the feed-streams. To manufacture each of the three products, the scheduler determines different set points that the controller must track by manipulating  $q_{R_i}$ . The set points are  $C_{A,t_c}^{SP} = [0.3, 0.0, 0.0]$ ,  $C_{B,t_c}^{SP} = [0.0, 0.18, 0.0]$ , and  $C_{C,t_c}^{SP} = [0.0, 0.0, 0.2]$ .

We assume a scheduling scenario which consists of  $N_s = 4$  time slots with  $\Delta t_s = 120$  min ( $T_{tot} = 480$  min). At the control level,  $\Delta t_c = 2$  min therefore  $t_c = 1, \dots, N_c = 240$  instances. The MPC problem is casted as in Eqs. (1):(4) considering the following parameters:  $P = 10^2 \times [100, 0, 0; 0, 1, 0; 0, 0, 1]$ ,  $R = 0$  since no set points on the control variables,  $R_1 = 10^{-6} \times [1, 0, 0; 0, 1, 0; 0, 0, 1]$ ,  $N_{ph} = 6$ , and  $M_{ch} = 2$ . The problem is subjected to: i) a bound constraint on the inlet flowrates  $0 \leq q_{R_i, t_c} \leq 500$  L/min, and ii) an equality constraint  $\sum_{i=1}^3 q_{R_i, t_c} = Q_{tot, t_c}$ , which correlates the control variables with the production targets suggested by the scheduler, i.e. equivalent to Eq. (3). A dataset including  $N_{sc, tr} = 45$  scenarios is generated (see Sections 3.1 and 3.2) and used for training the DNN-based SAC (Eq. (6)). The DNN consists of three hidden layers with 30, 50, and 5 neurons, and is trained using Bayesian regularization backpropagation algorithm. Another dataset of  $N_{sc, tst} = 10$  scenarios is created to test the fitted DNN-based SAC. The time for generating the training data, testing data, and fitting the DNN equals to 269, 199.3, and 15.3 min, respectively.

$$\begin{aligned} & \hat{q}_{1, t_c+1}, \hat{q}_{2, t_c+1}, \hat{q}_{3, t_c+1} = \\ & = \mathcal{f}(C_{R1, t_c}, C_{R2, t_c}, C_{R3, t_c}, C_{A, t_c}, C_{B, t_c}, C_{C, t_c}, q_{1, t_c}, q_{2, t_c}, q_{3, t_c}, Q_{tot, t_c}, C_{A, t_c}^{SP}, C_{B, t_c}^{SP}, C_{C, t_c}^{SP}) \end{aligned} \quad (7)$$

Table 1 reports the performance of the DNN-based SAC over the 10 testing scenarios, in terms of i) low NRMSE of the predicted closed-loop control variables and the resulting state and output variables, and ii) a huge savings in the online computational cost. As example, Figure 4 shows the closed-loop profiles of one inlet flowrate ( $\hat{q}_2$ ), predicted by the DL-based SAC and two of the resulting outputs ( $C_C$  and  $C_B$ ) compared with their exact values obtained by MPC during specific two time slots representing the production



transition from product C to product B in scenario 4. The figure further confirms the very good capabilities of the proposed DL-based SAC.

Table 1. Accuracy and online computational time of the DL-based SAC based on the testing set.

Accuracy (NRMSE %)									CPU time *	Saved time %
$\hat{q}_1$	$\hat{q}_2$	$\hat{q}_3$	$\hat{C}_{R1}$	$\hat{C}_{R2}$	$\hat{C}_{R3}$	$\hat{C}_A$	$\hat{C}_B$	$\hat{C}_C$	0.52	99.7=(199.3-0.52)/1.99.3
0.15	0.23	0.3	0.73	1.8	1.1	0.85	2.26	1.01		

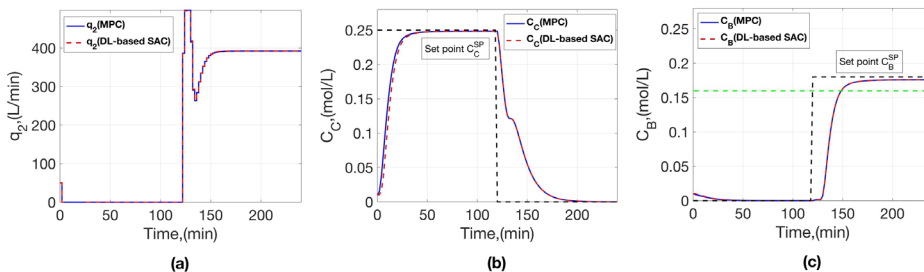


Figure 4. Predicted (red) versus exact (blue) values of the control variable  $q_2$  (a) and the resulting output variables  $C_C$  (b) and  $C_B$  (c) during transition  $C \Rightarrow B$  in the testing scenario no. 4. The green dashed line in subplot (c) represents the minimum quality threshold for product B.

The material loss and transition time during a certain transition can be calculated by comparing the output concentration with a minimum quality threshold. For example, in transition  $C \Rightarrow B$ , the output concentration  $C_B$  is compared to the quality threshold  $C_{B,\min} = 0.16$  mol/L (dashed green line in Figure 4-(c)).

## 5. Conclusions

In this paper, we present a DL-based SAC that is effectively able to receive, as input, the mid-term scheduling decisions along with the process conditions and predict the future control measures that track the scheduling targets. The results of application to a benchmark case study proves the effectiveness of the proposed DL-based SAC in terms of accuracy and very low computation time. A future research line will exploit the DL-based SAC and test its performance when integrated into scheduling problem/layer.

## References

- A. Flores-Tlacuahuac, I. Grossmann. Simultaneous cyclic scheduling and control of a multiproduct CSTR. *Ind. Eng. Chem. Res.* 2006, 45, 6698-6712.
- A. Flores-Tlacuahuac, I. Grossmann. Simultaneous scheduling and control of multiproduct continuous parallel lines. *Ind. Eng. Chem. Res.* 2010, 49, 7909-7921.
- B. Burnak, J. Katz, N. A. Diangelakis, E. Pistikopoulos, 2018. Simultaneous Process Scheduling and Control: A MP-Based Approach. *Ind. Eng. Chem. Res.* 57, 3963-3976
- J. Zhuge and M. Ierapetritou. Integration of Scheduling and Control with Closed Loop Implementation. *Ind. Eng. Chem. Res.* 2012, 51, 8550-8565
- O. A.-Martinez, L. R.-Sandoval. A nested online scheduling and nonlinear model predictive control framework for multi-product continuous systems. *AIChE journal* 68, 2022, 17665

# Robust (explicit) optimization and control via Mixed Integer Programming

Nikolaos A. Diangelakis<sup>a\*</sup>, Iosif Pappas<sup>b,c</sup>, Efstratios N. Pistikopoulos<sup>b,c</sup>

<sup>a</sup>*School of Chemical and Environmental Engineering, Technical University of Crete, Chania, 73100, Greece*

<sup>b</sup>*Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX, 77845, USA*

<sup>c</sup>*Texas A&M Energy Institute, Texas A&M University, College Station, TX, 77845, USA*

## Abstract

The presence of endogenous and exogenous uncertainty in process systems is one of the key reasons for deviation from set operation policies, resulting in suboptimal or even infeasible operation. In robust rolling horizon optimization (RHO) of discrete-time systems, such as MPC, the derivation of robust optimization policies is mainly characterized by the challenge associated with multiplicative unmeasurable uncertainty. Measurable uncertainty can be treated either “online” with the successive solution of optimization problems or “offline” via multi-parametric programming. We consider the development of closed-loop (explicit) robust rolling horizon solutions that correspond to MPC problems. The unmeasurable multiplicative uncertainty is reformulated into its deterministic mixed integer equivalent form. Furthermore, measurable uncertainty is treated either via online optimization or via multi-parametric programming. With this approach, we show (i) how the need for a multi-stage optimization approach can be alleviated while (ii) guaranteeing the robust nature of the final solution. We present the developments through an example of constrained LQR.

**Keywords:** robust optimization; robust control; mixed-integer programming; multi-parametric programming

## 1. Introduction

Model predictive control (MPC) is the established method for the advanced control of multivariable systems, extensively studied and applied by both the academic and industrial communities (D. Q. Mayne, 2014). An MPC problem is most commonly a convex quadratic mathematical optimization problem, whose solution is the optimal vector of inputs to regulate the operation of a system. By utilizing a rolling horizon optimization approach (RHO), the optimal vector of inputs that minimize a performance index while considering (i) the process model, (ii) the constraint set and (iii) a finite prediction horizon is calculated, however only the first control input of the sequence is applied to the system. As soon as the next measurement (or its estimate) is made available, the horizon is rolled over by a single time step, and the optimization problem is resolved. Being model-based, MPC approaches inevitably depend on the accuracy of the underlying model. In practical applications, model uncertainty is prevalent due to the mismatch between the real plant and the model used for the derivation of the control law (endogenous uncertainty) and/or unmeasured disturbances (exogenous uncertainty)

especially when multiple time scales are considered (N.A. Diangelakis et al. (2017), B. Burnak et al. (2020)). Consequently, the closed-loop behavior of the system is not as desired and constraint violations may occur. In an effort to address this challenge strategies that account for additive and multiplicative uncertainty have been developed (e.g. B. Kouvaritakis, et al, (2016) and references therein). Furthermore, robust optimization has received significant attention through the years (e.g. A. Ben-Tal, et. al, 1998, Q. Zhang et al. 2022 and references therein).

Robust explicit MPC aims to derive the explicit solution of the MPC problem by additionally guaranteeing that all sources of uncertainty are taken into account, and at least feasibility is satisfied. In this respect, multiple research efforts have been contributed that tackle the case where the uncertainty source is added to the future prediction (additive uncertainty) and the case where the future prediction is multiplicatively affected by it (multiplicative uncertainty). V. Sakizlis et al. (2004) included constraints in the design phase of the controller that guarantee that for the worst case of the additive uncertainty, the system is feasible. A. Bemporad et al. (2003) proposed a min-max approach where the solution is found for problems with a linear objective function and linear constraints for multiplicative uncertainty. K. Kouramas et al. (2013) tackled the case of explicit MPC problems with a quadratic objective function, linear constraints, and multiplicative uncertainty by employing dynamic programming and robust optimization. More recently, I. Pappas et al. (2022) demonstrated that dynamic programming can be avoided and extended it to hybrid systems by performing projections of the feasible space into the future. An open question in the robust explicit MPC field is how can the robust solution of an explicit MPC problem with a quadratic cost, linear constraints, and multiplicative uncertainty be developed, by using a single multiparametric optimization formulation. This contribution aims to provide a strategy for the later via mixed-integer and multiparametric quadratic optimization.

## 2. Problem formulation

Consider a linear discrete time dynamic model of the following form:

$$x_{T+1} = Ax_T + Bu_T \quad (1)$$

where  $x_T \in \mathcal{R}^m$  and  $u_T \in \mathcal{R}^n$  are the state and control input vectors respectively at time instant  $T$ . Instead of considering that the system and input matrices  $A \in \mathcal{R}^{m \times m}$  and  $B \in \mathcal{R}^{m \times n}$  are constant, in this study we assume that the model is uncertain and described by box uncertainty. Specifically:

$$A = A_0 + \Delta A \quad (2)$$

$$B = B_0 + \Delta B \quad (3)$$

$$\Delta A \in \mathcal{A} = \{\Delta A \in \mathcal{R}^{m \times m}: -\epsilon_A |A_0| \leq \Delta A \leq \epsilon_A |A_0|\} \quad (3)$$

$$\Delta B \in \mathcal{B} = \{\Delta B \in \mathcal{R}^{m \times n}: -\epsilon_B |B_0| \leq \Delta B \leq \epsilon_B |B_0|\} \quad (4)$$

$A_0$  and  $B_0$  are the nominal matrices of the model, while  $\Delta A$  and  $\Delta B$  are their uncertain component. This element-wise deviation from the nominal matrix value is described by  $\epsilon_A$  and  $\epsilon_B$  which are scalars. Here, in a slight abuse of terminology,  $|A_0|$  and  $|B_0|$  do not represent a norm or determinant of  $A_0$  and  $B_0$ , respectively, but their element-wise absolute values, i.e. if the element of  $A_0$  in the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column is  $a_0^{ij}$  then the element of  $|A_0|$  in the same row and column is  $|a_0^{ij}|$ .

Assuming a common prediction and control horizon  $N$ , the following linear quadratic regulator problem (LQR) problem can be formulated:

$$\begin{aligned}
\min_{u_0, u_1, \dots, u_{N-1}} \quad & x_N^T P x_N + \sum_{k=0}^N x_k^T Q x_k + \sum_{k=0}^{N-1} u_k^T R u_k \\
\text{s. t.} \quad & x_{T+1} = A x_T + B u_T, \forall T \in \{0, \dots, N-1\} \\
& \underline{x} \leq x_T \leq \bar{x}, \forall T \in \{0, \dots, N-1\} \\
& \underline{u} \leq u_T \leq \bar{u}, \forall T \in \{0, \dots, N-1\} \\
& x_N \in T_s
\end{aligned} \tag{5}$$

Its objective is to find the control inputs  $u_T$ , that will drive the system to the origin in the presence of the uncertainty. The lower bounds on the inputs and states are implied to be negative and the upper bounds positive. The system and input matrices in eqn. 5 are defined as in eqns. 2-4. The weights on the states and inputs are  $Q \in \mathcal{R}^{m \times m}$  and  $R \in \mathcal{R}^{n \times n}$  respectively, while  $P \in \mathcal{R}^{m \times m}$  is the terminal cost matrix derived from the solution of the discrete-time algebraic Riccati equation for the nominal system and input matrices. The states at the end of the prediction are required to belong to the terminal set  $T_s$  (Blanchini, 1999). In the methodology section, the consideration of multiplicative uncertainty within the formulation of eqn. 5 will be shown to provide an optimization based robust counterpart for the LQR.

### 3. Methodology

The methodology presented in this section aims to ensure feasibility of the system for any value of  $\Delta A$  and  $\Delta B$  within their bounds. In other words, it aims to provide a formulation that ensures that for any realization of the multiplicative uncertainty, if the initial values of the states of the system ( $x_0$ ) lay within a feasible domain, there exists a sequence of  $u_T$  that will (a) keep the system feasible for all  $T$  and (b) will try to bring the system to the origin. It is understood that through this robustification methodology the feasible domain of  $x_0$  for which (a) can happen will be significantly reduced compared to a nominal case where no multiplicative uncertainty occurs. Conservatism of the approach is not addressed in the current work, only strict feasibility of the system under uncertainty.

The methodology can be broken down into four steps, namely: (i) Robustification, (ii) Reformulation, (iii) Reduction and (iv) Solution. Note that in the following sections the bounds on  $x_T$  and  $u_T$  are omitted for brevity but are implied and the formulations are valid for  $\forall T \in \{0, \dots, N-1\}$ .

#### 3.1. Robustification

The Robustification of the problem defined by eqns. 2-5 happens successively starting from the initial state to state  $N$  following the approach by (A. Ben Tal, et al. (1998)). Given that the states  $x_T$  and inputs  $u_T$  of the system are bounded for any time-step  $T$  we have from eqns. 2-5:

$$\begin{aligned}
\underline{x} \leq x_{T+1} = A x_T + B u_T \leq \bar{x} \\
\text{Eqns. 2-4}
\end{aligned} \tag{6}$$

The robust counterpart for a matrix  $G$  of appropriate dimensions and a constraint of the form of eqn. 7 can be expressed as in eqn. 8:

$$G x_{T+1} \leq g, \forall T \in \{0, \dots, N-1\} \tag{7}$$

$$G A_0 x_T + \epsilon_A |G| |A_0| |x_T| + G B_0 u_T + \epsilon_B |G| |B_0| |u_T| \leq g \tag{8}$$

### 3.2. Reformulation

The main challenge with eqn. (8) is the absolute values of the state and input variables in every time step. We apply a reformulation based on eqn. 8 as follows:

$$GA_0x_T + \epsilon_A|G||A_0|z_T + GB_0u_T + \epsilon_B|G||B_0|v_T \leq g \quad (9.1)$$

$$-z_T \leq x_T \leq z_T \quad (9.2)$$

$$0 \leq z_T \leq \max\{|\bar{x}|, |\underline{x}|\} \quad (9.3)$$

$$y_T \underline{u} \leq u_T \leq (1 - y_T)\bar{u}, y_T \in \{0,1\} \quad (9.4)$$

$$w_T = y_T u_T \quad (9.5)$$

$$v_T = u_T - 2w_T \quad (9.6)$$

$$0 \leq u_T - 2w_T \leq \max\{|\bar{u}|, |\underline{u}|\} \quad (9.7)$$

Eqn. (9.2-9.3) corresponds to the standard relaxation of the absolute term of  $x_T$  and its bounds, eqn. (9.4) introduces an auxiliary binary variable  $y_T$  which is equal to 0 when the input action  $u_T$  is positive and 1 when it is negative. A continuous auxiliary variable  $w_T$  is introduced in eqn. (9.5) and it represents a bilinear term between the input and auxiliary  $y_T$ . Eqn. (9.6) is used to replace  $v_T$  in eqn. (9.1) while preserving its bounds by (9.7). The bilinear term introduced can be exactly reformulated with a single McCormick relaxation since it is a product between a binary and a continuous variable, as follows:

$$w_T \geq y_T \underline{u} \quad (9.5.1)$$

$$w_T \geq u_T + y_T \bar{u} - \bar{u} \quad (9.5.2)$$

$$w_T \leq \underline{u} y_T + u_T - \underline{u} \quad (9.5.3)$$

$$w_T \leq \bar{u} y_T \quad (9.5.4)$$

By successively robustifying as in 3.1 for all time step in the output/control horizon, reformulating as in eqn. (9.x), utilizing eqn. (9.5.x) instead of (9.5) and substituting  $v_T$  with eqn. (9.6) we derive a robust counterpart of the problem in eqn. (5) that involves the original terms  $x_0$ ,  $u_T$  and the auxiliary terms  $w_T$ ,  $z_T$  and  $y_T$ .

### 3.3. Reduction

The continuous auxiliary variables  $w_T$  and  $z_T$  can be a source of primal degeneracy as described in E. N. Pistikopoulos, et. al (2020), therefore a procedure is needed to project the problem onto a space where the aforementioned variables do not exist but without affecting the feasible space of the original variables. Here, we utilize a single Fourier-Motzkin elimination step per continuous auxiliary variable. For two linear constraints of the form of eqn. (10), assuming without loss of generality  $a_1, a_2 > 0$ :

$$a_1 z_T + l_1(x_0, u_T, y_T, w_T) \leq b_1 \Leftrightarrow z_T \leq \frac{1}{a_1} (b_1 - l_1(x_0, u_T, y_T, w_T)) \quad (10.1)$$

$$a_2 z_T + l_2(x_0, u_T, y_T, w_T) \geq b_2 \Leftrightarrow z_T \geq \frac{1}{a_2} (b_2 - l_2(x_0, u_T, y_T, w_T)) \quad (10.2)$$

where  $l_1$  and  $l_2$  are linear functions of their variables we can eliminate  $z_T$  by substituting eqn. (10) with:

$$\frac{1}{a_2} (b_2 - l_2(x_0, u_T, y_T, w_T)) \leq \frac{1}{a_1} (b_1 - l_1(x_0, u_T, y_T, w_T)) \quad (11)$$

Note that in the LQR problem, given the proposed reformulation, equations of the form of eqn. (10.x) will always be present for all continuous auxiliary variables as they are bounded by definition. Also note that the procedure is repeated for all  $z_T$  and  $w_T$  thus creating a problem formulation in  $u_T$ ,  $z_0$  and  $y_T$ , i.e. a mixed-integer program. Lastly, note that the reformulation presented here is based on reformulating the auxiliary variable

$v_T$  of eqn. (9.1). There exists an equivalent reformulation for the term  $z_T$  involving binary variables which is more challenging, outside the scope of the current work but within the scope of overall research associated with this approach and currently under development.

### 3.4. Solution

The reduced problem is symbolically presented below:

$$\begin{aligned} \min_{\mathbf{u}, \mathbf{y}} \quad & [\mathbf{u}^T, \mathbf{y}^T] \mathbf{H} [\mathbf{u}^T, \mathbf{y}^T]^T + [\mathbf{u}^T, \mathbf{y}^T] \mathbf{Z} \mathbf{x}_0 \\ \text{s. t.} \quad & \mathbf{G} \mathbf{u} + \mathbf{E} \mathbf{y} \leq \mathbf{b} + \mathbf{F} \mathbf{x}_0 \\ & \mathbf{C} \mathbf{R}_A \mathbf{x}_0 \leq \mathbf{C} \mathbf{R}_b \\ & \mathbf{y} \in \mathbf{0}, \mathbf{1}^n, \mathbf{u} \in \mathcal{R}^n, \mathbf{x}_0 \in \mathcal{R}^m \end{aligned} \quad (12)$$

The matrix  $\mathbf{H}$  can be shown to be  $\mathbf{H}$  a  $\{2n \times 2n\}$  positive semi-definite matrix given that the matrices defining the LQR in eqn. (5) are positive definite. Bilinear terms between continuous and binary terms or purely binary terms that may arise due to  $\mathbf{H}$  can be exactly reformulated similar to eqn. (9.5.x). The program described by eqn. (12) can be solved online as soon as values for  $\mathbf{x}_0$  are available with the use of any local MIQP algorithm (i.e. CPLEX, Gurobi, etc.) or explicitly with the use of multi-parametric programming, treating  $\mathbf{x}_0$  as uncertain but measurable parameters, as described in E. N. Pistikopoulos, et. al (2020).

## 4. Uncertain constrained LQR example

Here we consider an LQR of a double integrator system described by eqn. (5) where  $Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ,  $R = 0.01$ ,  $P = \begin{bmatrix} 2.6235 & 1.6296 \\ 1.6296 & 2.6457 \end{bmatrix}$ ,  $A_0 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ ,  $B_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ ,  $\epsilon_A = \epsilon_B = 0.1$ . The system bounds are defined as  $\bar{\mathbf{x}} = -\underline{\mathbf{x}} = [10, 10]^T$  and  $\bar{\mathbf{u}} = -\underline{\mathbf{u}} = 2$ .

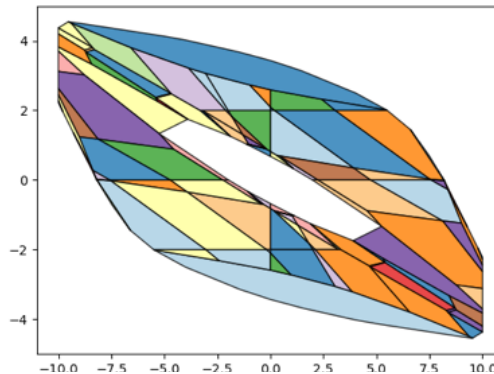


Figure 1: Critical Regions of mpMIQP. The axes represent initial state values

An horizon of 2 time-steps is used. The problem is treated as discussed in section 3 and eqn. 12 is solved via multi-parametric Mixed Integer Quadratic Programming, while preserving the envelopes of solution as described in Pistikopoulos et al. (2020). The solution comprises of 108 Critical Regions (shown in Figure 1) and is able to stir the problem back to the origin from various starting points for a plethora of multiplicative uncertainty realizations. This is presented in Figure 2 where the system returns to the origin for various, random, multiplicative uncertainty realizations and two different starting points for the states, without constraint violations, an indication of successful robustification.

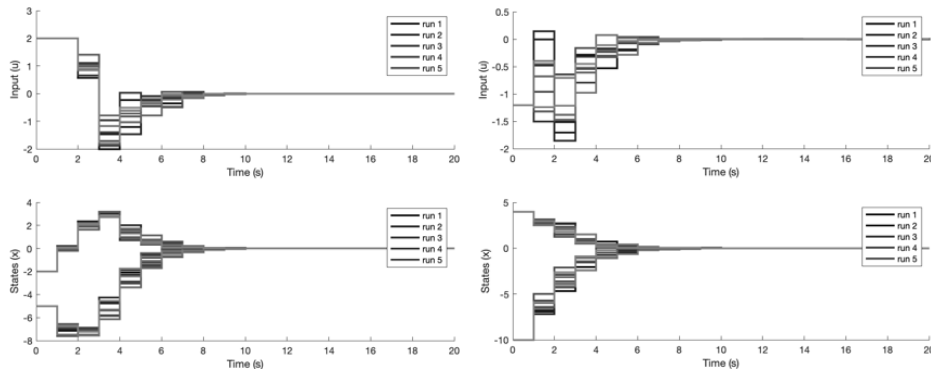


Figure 2: Optimal input and system response of the example for various values of multiplicative uncertainty and two starting points ( $x_0 = [-2, -4.4]$ : left,  $x_0 = [-10, 4.4]$ : right)

## 5. Concluding remarks

In this work, we presented an algorithm to solve robust explicit RHO problems via mpMIQP. Our approach is based on the (i) robustification, (ii) reformulation and (iii) reduction prior to the (iv) solution via multi-parametric programming. We demonstrated our findings on a constrained LQR case study where we exhibited that the system can be driven to the origin for any arbitrary, bounded value of the multiplicative uncertainty. Our next steps include the extension of step (ii) to other auxiliary variables and the extension to other classes of rolling horizon optimization problems.

## References

- D. Q. Mayne (2014), Model predictive control: Recent developments and future promise. *Automatica*, 50 (12), 2967-2986.
- N.A. Diangelakis, B. Burnak, J. Katz, E.N. Pistikopoulos (2017), Process design and control optimization: A simultaneous approach by multi-parametric programming. *AIChE Journal*, 63 (11), 4827-4846.
- B. Burnak, N. A. Diangelakis, E. N. Pistikopoulos (2020), *Integrated Process Design and Operational Optimization via Multiparametric Programming*, Synthesis Lectures on Engineering, Science, and Technology, Morgan & Claypool Publishers.
- B.Kouvaritakis, M. Cannon (2016), *Model Predictive Control*. Switzerland: Springer International Publishing, 38.
- D. Q. Mayne, J. B. Rawlings, C. V. Rao, P. O. M. Scokaert (2000) Constrained model predictive control: Stability and optimality. *Automatica*, 36 (6), 789-814.
- A. Ben-Tal, A. Nemirovski (1998), Robust convex optimization, *Mathematics of Operations Research*, 23 (4), 769-805.
- Q. Zhang, C. E. Gounaris (2022), Methodology and applications of robust optimization. *Optimization and Engineering* 23, 1761-1764
- V. Sakizlis, N. M. Kakalis, V. Dua, J. D. Perkins, E. N. Pistikopoulos, (2004), Design of robust model-based controllers via parametric programming. *Automatica*, 40, 2, 189-201.
- A. Bemporad, F. Borelli, M. Morari (2003), Min-max control of constrained uncertain discrete time linear systems. *IEEE Transactions on automatic control*, 48(9), 1600-1606.
- K. Kouramas, C. Panos, N. P. Faisca, E. N. Pistikopoulos (2013), An algorithm for robust explicit/multi-parametric model predictive control. *Automatica*, 49, 2, 381-389.
- I. Pappas, N. A. Diangelakis, R. Oberdieck, E. N. Pistikopoulos (2022), A Robust Optimization Strategy for Explicit Model Predictive Control, *CACE*, 49, 409-414.
- F. Blanchini (1999), Set invariance in control. *Automatica*, 35 (11), 1747-1767.
- E. N. Pistikopoulos, N. A. Diangelakis, R. Oberdieck (2020), *Multi-parametric Optimization and Control*, John Wiley & Sons, 2020.

# A process data prediction method for chemical process based on the frozen pretrained transformer model

Yiming Bai<sup>a</sup>, Jinsong Zhao<sup>a,b</sup>

<sup>a</sup>*Department of Chemical Engineering, Tsinghua University, Beijing 100084, China*

<sup>b</sup>*Beijing Key Laboratory of Industrial Big Data System and Application, Tsinghua University, Beijing, 100084, China*  
*jinsongzhao@mail.tsinghua.edu.cn*

## Abstract

With the reform of industrial production digitization and automation, process fault detection and diagnosis (FDD) has been an indispensable technical method to realize the safe and efficient production of chemical process, even if there may be some delay and inaccuracy. Fault prognosis could predict the occurrence of faults in advance, which would give operators more time and reduce the impact of faults. Frozen Pretrained Transformer (FPT) has shown language-pretrained transformers can obtain strong performance on a variety of simple non-language classification tasks. Inspired by the pre-trained models, we proposed a chemical process data prediction method based on FPT models. Specifically, we take a transformer model pretrained on natural language data and finetune some parameters on the chemical process data. And the Tennessee Eastman process (TEP) was used to demonstrate the validity of the method. Finally, we discussed the relevant issues in applying language-pretrained models to process data.

**Keywords:** Process data prognosis; Chemical process; Transformer model; Frozen pretrained model.

## 1. Introduction

In recent decades, with the reform of industrial production digitization and automation, various industrial processes have generated massive amounts of data. In order to fully process and apply these data, the computing power of machines has reached an unimaginable level (Rehman et al. 2019). In the modern chemical industry, the development of artificial intelligence technology has made outstanding contributions to industrial intelligence, but the promotion of automation and intelligence has also made modern chemical plants becoming more and more large-scale and highly complex than ever.

As the complexity of chemical engineering systems has increased, chemical process safety has been challenged to a large extent. The faults of chemical processes may cause economic losses and casualties, which pose a great threat to chemical safety (Venkatasubramanian et al. 2003). FDD has been an indispensable technical method to realize the safe and efficient production of chemical processes and has substantial theoretical and application value in the chemical industry (Bi et al. 2022). However, FDD may be not sufficient in some cases, and it is necessary to prognose the fault ahead of time. Fault prognosis mainly aims at potential and future faults, which means the research



object of fault prognosis is a future uncertain event (Smith and Randall 2015; Zhang et al. 2011). Bai et al. (Bai and Zhao 2023) explained the relationship between fault prognosis and FDD, which is shown in Fig. 1. In the process monitoring loop, the fault prognosis, located before FDD, attempts to identify that the system is in an abnormal state in advance, which would greatly reduce the probability of abnormal conditions.

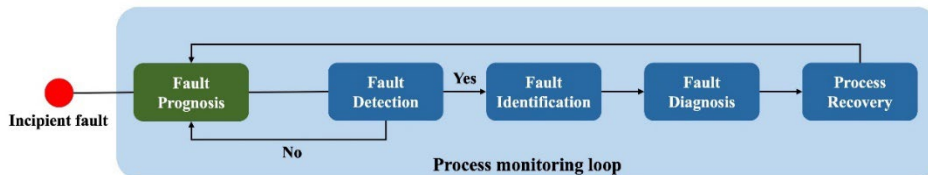


Fig. 1. Relationship between fault prognosis and process monitoring (Bai and Zhao 2023).

Current fault prognosis methods can be generally classified into three types: model-based methods, knowledge-based methods, and data-based methods (Zhong, Han, and Han 2020). Data-based methods determine the status of the system within a certain period by analyzing previously observed data. With the development of automation and digitization, data-based models have been playing an increasing role in the process industry.

Fault prediction methods have evolved with the changing characteristics of industrial data. As data complexity increases, many data-driven methods are flexibly applied in nonlinear scenarios, such as Artificial Neural Networks, autoencoder, Support Vector Machine, Radial Basis Function network, Recurrent Neural Network (RNN) and long short term memory (LSTM) (Yuan et al. 2020). Transformer architecture has widespread applications, particularly in Natural Language Processing (NLP) and computer vision, which also triggered huge interests in the time series data (Wen et al. 2022). It has become common practice to train large models on unsupervised or weakly supervised objectives before finetuning or evaluating zero-shot generalization on a downstream task. Lu et al. found that pre-training in natural language can improve the performance and computational efficiency of non-verbal downstream tasks, and proposed the Frozen Pretrained Transformer (FPT) model (Lu et al. 2021), which inspired the pre-trained models for time series data.

In this paper, we propose a process data prediction method for chemical process based on FPT models. We take a transformer model pretrained on natural language data and finetune some parameters on the chemical process data. And the Tennessee Eastman process (TEP) was used to demonstrate the validity of the method. The results show that the proposed method obtain the strong performance and less consumption on the chemical process fault prognosis task. Additionally, we discussed the relevant issues in applying language-pretrained models to process data, which verified further validity of the proposed method.

The remainder of this paper is organized as follows. Section 2 reviews the basic knowledge of Transformer models and FPT model. Section 3 includes the descriptions of the proposed process data prediction method. The applications and discussions are revealed in Section 4. Section 5 gives summary.

## 2. Methodology

### 2.1. Self-attention

The attention mechanism could overcome this weakness by learning the weights of input information at different times. We show a version of a commonly used attention in Fig. 2, which was proposed as self-attention (Vaswani et al. 2017).

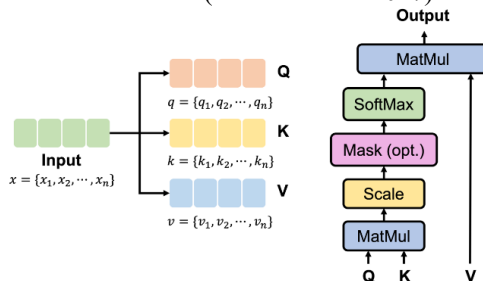


Fig. 2. A schematic diagram of self-attention.

Firstly, given an input sequence  $x = \{x_1, x_2, \dots, x_n\}$ ,  $x_i \in R^k$ , query, key and value vectors can be obtained by mapped with linear transformations and  $q_i, k_i, v_i \in R^{d_k}$  are examples to show the calculation process of each element.

$$q_i = W^q x_i, \quad k_i = W^k x_i, \quad v_i = W^v x_i, \quad i = 1, 2, \dots, n \quad (1)$$

Where  $W^q, W^k, W^v \in R^{d_k \times k}$  are learnable matrices.

Then, the similarity score of  $q_i$  and  $k_j$ , which are chosen as examples to show calculation, is calculated through a scaled dot-product function, and the SoftMax function is then used to normalize the scores and obtain the weights.

$$\alpha_{i,j} = \frac{q_i \cdot k_j}{\sqrt{d_k}} \quad (2)$$

$$\hat{\alpha}_{i,j} = \frac{\exp(\alpha_{i,j})}{\sum_{t=1}^n \exp(\alpha_{i,t})} \quad (3)$$

Finally, the output is calculated as the weighted sum of values.

$$z_i = \sum_{j=1}^n \hat{\alpha}_{i,j} v_j \quad (4)$$

### 2.2. FPT model

To investigate this hypothesis, we take a transformer model, GPT-2, pretrained on natural language data (Radford et al. 2019). GPT-2 is a transformers model pretrained on a very large corpus of English data in a self-supervised fashion. Crucially, all communication between tokens in the model are frozen. The data in each datapoint is chunked into discrete tokens, and can only reference each other via the frozen attention connections, which are not trained. Our key investigation is to analyze the performance of the model in process data prediction, which is already inherent in the language model. On the basis of the pretrained model, the model is finetuned with a minimal amount of computation on the downstream modality. The FPT model used in process data utilizes the feature extraction part of the pretrained model, and adds two fully connected layers to handle the

input process data and the output features. There are five types of parameters in the finetune process, which are input layer, output layer, layer normalization, positional embedding and input embedding.

### 3. The proposed method

The proposed process data prediction method is shown in the Fig. 3. Firstly, process data can be directly used as model input without embedding. The sequence structure of the process data can reduce the difficulty of position coding, and we can use one-hot matrix to indicate the order of input. Then the input layer expands the dimensionality of input data as needed for GPT-2. And the features of the process data are extracted by the pretrained model by NLP. Finally, the output layer gets the prediction results of the process data by the features. The dimension of the input layer is the sum of the dimension of process variables and the time dimension. The dimension of the output layer is the dimension of process variables. The number of the dimension in the GPT-2 is 762 and the number of frozen attention block is 12.

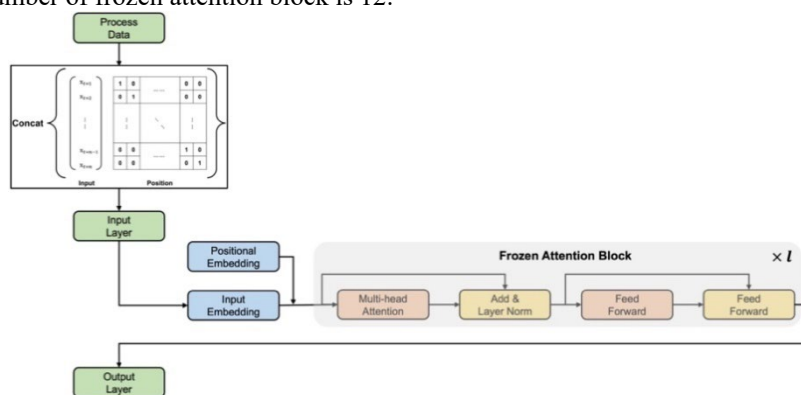


Fig. 3. The diagram of the proposed process data prediction method.

## 4. Applications and discussions

### 4.1. Applications on the TEP

TEP is a classical model benchmark for simulating chemical processes, which has been widely used as a study case for FDD and fault prognosis. It is mainly composed of 5 unit operations, including a reactor, a condenser, a recycle compressor, a vapor-liquid separator, and a stripper. The initial version of TEP simulation involves 12 manipulated variables, 41 measured variables and 20 different types of faults. In 2015, Bathelt revised the original TE model and used MATLAB's Simulink to build a new one, adding 8 measurement variables, 24 component variables, and 8 new process disturbances (Bathelt, Ricker, and Jelali 2015). In our work, the applications are based on the revised version of Bathelt, which is shown in the Fig. 4.

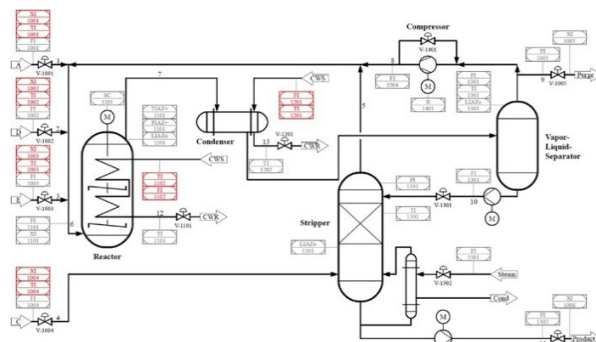


Fig. 4. P&ID of the revised TEP model (Bathelt, Ricker, and Jelali 2015).

Since the sampling frequency of component analysis measurement is much lower, we only choose the process manipulated variables and continuous process measurements except variables related to valves for dataset building. Among the common faults in the TEP, most fault types are step or random changes, which bring significant changes to the system within a short time of the fault occurrence. The process variables prediction and fault prognosis are with little necessarily in the cases, and these cases do little help for validating the effectiveness of the prediction methods in chemical process. According to the application results of the FDD methods in the abnormal state of TEP, we choose fault 13 as the target of the study, which is the fault of reaction kinetics drifting slowly.

A simulation dataset is built to assess the fault detection and identification performance of all the methods. During the simulation, 1 sample is collected every 1 min. The datasets include 10 sets of data for fault 13. Each set of data includes 480 normal samples followed by 600 fault samples, corresponding to 8h of normal data and 10h of fault data. The training dataset includes 8 sets of data and the validation dataset and the testing dataset each includes a set of data. For the selected time window, the goal of the model is to predict the changes in the 30 process variables over a future time window. And MSE of all the process variables in the time window is selected to be the loss function. We chose RNN, LSTM, BiLSTM and full GPT2 to compare the predictive performance. And the full GPT2 means that we trained all the parameters of the GPT2 model in the training stage. The prediction results of different methods with different time window are shown in the following Table. 1.

Table. 1. Prediction results of different methods.

Method	Time widow = 8	Time widow = 16	Time widow =32
RNN	0.826	1.355	<b>2.648</b>
LSTM	0.714	1.147	2.883
BiLSTM	0.737	1.157	3.842
FPT	<b>0.140</b>	<b>0.510</b>	2.774
Full GPT2	0.647	0.999	4.016

The results shown that the FPT had the best performance when the time window is 8 or 16. Theoretically, the result of full GPT2 should be a lower bound of FPT. However, due to the limited training datasets of process variables, the MSE of full GPT2 is larger than that of FPT. And it indicated that to some extent better prediction results could be achieved in the process variables using a pretrained model than a fully trained model. To the time window of 32, FPT didn't get the better performance.

When only the parameters of input layer and output layer were finetuned, the MSE was 7.808. On this basis, we only finetuned some selected parameters to find out which

parameters were most important in the GPT2 model for process variables prediction. Firstly, we only finetuned the parameters of the layer normalization, and the MSE reduced to 0.563, which indicated that finetuning the parameters of the layer normalization could improve effectively prediction results. Then we added the parameters of the position embedding to the finetuned parameters, and the MSE reduced to 0.556. Then we added the parameters of the input embedding to the finetuned parameters, and the MSE reduced to 0.551. These two increases in finetuned parameters did not effectively reduce the MSE. The results shown that the parameters of the layer normalization played the most important role in the process variables prediction.

Studies have shown that the number of layers of neural networks for process data is less than that for NLP or CV problems. In our work, the number of layers could be measured by the number of used frozen attention blocks. We finetuned the model with different numbers of frozen attention blocks to learn the most suitable model for the chemical process variables prediction. We found that the number of used frozen attention blocks was selected to be 2 to show the best results in the applications. And the MSE with the time window of 8, 16 and 32 reduced to 0.109, 0.287 and 1.667, which were better than other methods.

## 5. Summary

In the article, we proposed a process data prediction method for chemical process based on FPT models. In TEP the proposed method shown the best validity with the smallest MSE of process variables. And the FPT method could show the strong performance and less consumption on the chemical process variables prediction task. We found that the parameters of the layer normalization played the most important role in the results. And the number of used frozen attention blocks is selected to be 2, which was shown to have better results in the application of chemical process variable prediction. In the future, the application of pretrained models to process variables prediction and process fault prognosis remains widely open with many challenges left to solve.

## References

- Bai, Yiming, and Jinsong Zhao. 2023. "A Novel Transformer-Based Multi-Variable Multi-Step Prediction Method for Chemical Process Fault Prognosis." *Process Safety and Environmental Protection* 169 (January): 937–947.
- Bathelt, Andreas, N. Lawrence Ricker, and Mohieddine Jelali. 2015. "Revision of the Tennessee Eastman Process Model." *IFAC-PapersOnLine*, 9th IFAC Symposium on Advanced Control of Chemical Processes ADCHEM 2015, 48 (8): 309–314.
- Bi, Xiaotian, Ruoshi Qin, Deyang Wu, Shaodong Zheng, and Jinsong Zhao. 2022. "One Step Forward for Smart Chemical Process Fault Detection and Diagnosis." *Computers & Chemical Engineering*, June, 107884.
- Lu, Kevin, Aditya Grover, Pieter Abbeel, and Igor Mordatch. 2021. "Pretrained Transformers as Universal Computation Engines." arXiv.
- Radford, Alec, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, and others. 2019. "Language Models Are Unsupervised Multitask Learners." *OpenAI Blog* 1 (8): 9.

- Rehman, Muhammad Habib ur, Ibrar Yaqoob, Khaled Salah, Muhammad Imran, Prem Prakash Jayaraman, and Charith Perera. 2019. "The Role of Big Data Analytics in Industrial Internet of Things." arXiv.
- Smith, Wade A., and Robert B. Randall. 2015. "Rolling Element Bearing Diagnostics Using the Case Western Reserve University Data: A Benchmark Study." *Mechanical Systems and Signal Processing* 64–65 (December): 100–131.
- Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. "Attention Is All You Need." In *Advances in Neural Information Processing Systems*. Vol. 30. Curran Associates, Inc.
- Venkatasubramanian, Venkat, Raghunathan Rengaswamy, Kewen Yin, and Surya N. Kavuri. 2003. "A Review of Process Fault Detection and Diagnosis: Part I: Quantitative Model-Based Methods." *Computers & Chemical Engineering* 27 (3): 293–311.
- Wen, Qingsong, Tian Zhou, Chaoli Zhang, Weiqi Chen, Ziqing Ma, Junchi Yan, and Liang Sun. 2022. "Transformers in Time Series: A Survey." arXiv. doi:10.48550/arXiv.2202.07125.
- Yuan, Xiaofeng, Chen Ou, Yalin Wang, Chunhua Yang, and Weihua Gui. 2020. "A Novel Semi-Supervised Pre-Training Strategy for Deep Networks and Its Application for Quality Variable Prediction in Industrial Processes." *Chemical Engineering Science* 217 (May): 115509.
- Zhang, Bin, Chris Sconyers, Carl Byington, Romano Patrick, Marcos E. Orchard, and George Vachtsevanos. 2011. "A Probabilistic Fault Detection Approach: Application to Bearing Fault Detection." *IEEE Transactions on Industrial Electronics* 58 (5): 2011–2018.
- Zhong, Kai, Min Han, and Bing Han. 2020. "Data-Driven Based Fault Prognosis for Industrial Systems: A Concise Overview." *IEEE-CAA JOURNAL OF AUTOMATICA SINICA* 7 (2): 330–345.



# Sensor bias detection, isolation, and estimation in a subsea pump system.

Rafael D. Oliveira<sup>a</sup>, Halvor A. Krog<sup>a</sup>, Galo A. C. Le Roux<sup>b</sup>, Johannes Jäschke<sup>a</sup>

<sup>a</sup> *Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway*

<sup>b</sup> *Department of Chemical Engineering, University of São Paulo, São Paulo, Brazil*

[johannes.jaschke@ntnu.no](mailto:johannes.jaschke@ntnu.no)

## Abstract

Subsea pump systems are subject to many possible faults and monitoring typically relies on the measurements of key process variables. However, the measurements themselves may be subject to faults like e.g., bias. In this work, a method that combines fault detection and isolation, observability analysis and state estimation is presented and tested on a simulated pump system. The results indicate that a state estimator with correct information about which sensor has drifted gives lower estimation error than a state estimator that does not consider bias. Furthermore, we show that it is impossible to estimate the bias on all sensors simultaneously due to observability issues of the augmented system.

**Keywords:** Sensor bias, Fault detection, Subsea pump, Fault isolation, Sensor drift.

## 1. Introduction

Subsea pumping systems operate in harsh environments for long periods, making the maintenance of these pumps challenging and costly. Condition monitoring is crucial to keep these subsea pumps running for long periods and to reduce the possibility of accidents. The pumps are subject to several potential faults, and it is desirable to use measurements to detect if something is going wrong. However, the sensors that provide these measures are also susceptible to faults (e.g., gross errors). These gross errors arise from sensors malfunctioning and can significantly impact the monitoring and controlling of the subsea pump.

Given the long operation times of subsea pumps, the long-term faults (e.g., bias) are more challenging to detect than other abrupt changes. It is desirable to detect and isolate these biases as soon as possible and, if practicable, estimate their magnitude. Approaches to deal with these sensor biases can be classified as passive or active. The passive methods apply robust estimators like M-estimators, robust Kalman Filter (KF), and robust Moving Horizon Estimation (MHE) (Siddhartha et al., 2022). Passive methods efficiently handle gross errors, but no information about the bias's location and size is generated. The location of the gross errors is a valuable information for maintenance and process monitoring. On the other hand, active methods try to estimate these biases, usually representing the bias as a new time-dependent continuous parameter on the state estimation problem (Gatzke & Doyle III, 2002). However, observability problems usually arise (Gatzke & Doyle III, 2002).



In this paper, we propose to combine fault detection and isolation (FDI) methods, observability analysis and state estimators in the context of a subsea pump system. In our methodology, i) FDI is used to identify sensors with bias. ii) Then an observability analysis is performed to check if it is possible to estimate that bias, and iii) then the state estimator is extended with a bias parameter to be estimated. The FDI reduces the number of biases to be estimated, and the observability analysis assures that the bias can be estimated without worsening the state estimator's performance. This illustrates how attempting to estimate sensor bias without any *a priori* information of which sensors are failing is impractical and how combining FDI with state estimators can improve the estimation significantly. Furthermore, a Monte Carlo analysis is performed to show the advantages and limitations of that approach in multiple scenarios.

## 2. Case study: Subsea pump system

We consider a simple subsea booster pump system, as shown in Figure 1. A mixture of oil and water comes from the reservoir at pressure  $p_1$ . The pump runs with a variable speed drive (VSD) at  $\omega$  rpm, which produces a specific head,  $H$ . Given the density of the fluid,  $\rho$ , one can calculate the pressure at the outlet of the pump,  $p_2$ . The pressure  $p_3$  is determined by downstream facilities (oil platform). The choke valve has a flow rate of  $q$  and opening of  $Z$  %. The three pressures are measured, and a Venturi flowmeter measures the pressure drop over an internal orifice (not shown).

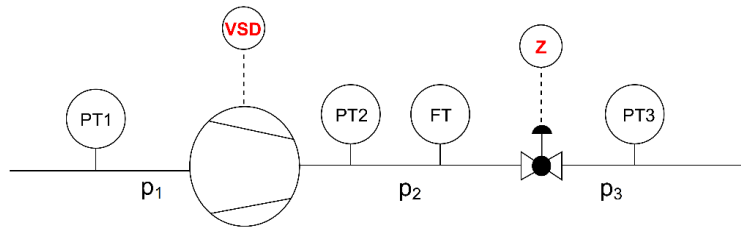


Figure 1: The subsea booster pump with its installed sensors. FT is a Venturi flowmeter.

### 2.1. Pump model

Sensor bias is a phenomenon that typically takes place in the timescale of months and years. The pressure and density changes in the reservoir are on the same timescale, and we assume the pressure on the oil platform has the same trend as the reservoir pressure. On the other hand, phenomena such as transient flow regimes happen almost instantly in the timescale of seconds. This motivates the use of a quasi-steady state model, where changes in the reservoir and  $p_3$  are modelled dynamically and the pump is modeled in steady-state (and therefore the time-subscript is  $k + 1$  on both sides of the equality sign in (4)-(6)). It is assumed that  $p_1$  and  $p_3$  decreases at the same rate, hence, they have the same parameter  $\theta_{p,1}$ . The faultless system is described by:

$$p_{1,k+1} = p_{1,k} + \theta_{p,1}\Delta t + w_{p_{1,k}} \quad (1)$$

$$\rho_{k+1} = \rho_k + \theta_{\rho}\Delta t + w_{\rho,k} \quad (2)$$

$$p_{3,k+1} = p_{3,k} + \theta_{p,1}\Delta t + w_{p_{3,k}} \quad (3)$$

$$p_{2,k+1} = p_{1,k+1} + \rho_{k+1}gH_{k+1} + w_{p_{2,k}} \quad (4)$$

$$q_{k+1} = C_v Z_{k+1} \sqrt{\frac{p_{2,k+1} - p_{3,k+1}}{\rho_{k+1}/\rho_{water}}} + w_{q,k} \quad (5)$$

where the head at time  $k$ ,  $H_k$ , is found by using the so-called pump curve.

$$H_{k+1} = \theta_{H_0} \omega_{k+1} + \theta_{H_1} q_{k+1} \omega_{k+1} + \theta_{H_2} q_{k+1}^2 \quad (6)$$

The states are  $\mathbf{x}_k = [p_{1,k}, p_{2,k}, p_{3,k}, \rho_k, q_k]^T$  and the state-propagation equations (1)-(5) can be written in the form of  $\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_{k+1}, \Delta t) + \mathbf{w}_k$ . The stochastic process noise is defined as  $\mathbf{w}_k = [w_{p_1,k}, w_{p_2,k}, w_{p_3,k}, w_{\rho,k}, w_{q,k}]^T \sim (\mathbf{0}, \mathbf{Q}_k)$ . The process noise distribution is assumed to be constant at  $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$  where  $\mathbf{Q} = 10^{-5} \mathbf{I}$ . The parameter values are  $\boldsymbol{\theta} = [\theta_{p,1}, \theta_p, \theta_{H_0}, \theta_{H_1}, \theta_{H_2}, g, C_v] = [-5.0, 5.0, 3.51 \times 10^{-5}, 3.29 \times 10^{-4}, -0.01, 9.8 \text{ m/s}^2, 47.50 \text{ m}^3/\text{h}]^T$ , where  $g$  is the gravitational constant and  $C_v$  is the valve constant. The remaining parameters describe the evolution of pressure and density in time. The initial conditions for all the simulations were  $\mathbf{x}_0 = [50.0 \text{ bar}, 64.5 \text{ bar}, 51.0 \text{ bar}, 700 \text{ kg/m}^3, 136.6 \text{ m}^3/\text{h}]^T$  and  $\mathbf{u}_0 = [65 \%, 3500 \text{ rpm}]^T$ . When there is no bias in the sensors, the sensor equations are:

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k = \left[ p_{1,k}, p_{2,k}, p_{3,k}, \frac{\rho_k(1 - \beta^4)}{2} \left( \frac{4q_k}{\pi d^2 C \epsilon} \right)^2 \right]^T + \mathbf{v}_k \quad (7)$$

where  $C = 1.01$  is the flow coefficient,  $\epsilon = 1$  is the expansibility factor,  $d = 0,0525 \text{ mm}$  is the Venturi throat diameter,  $D = 0,154 \text{ m}$  is the internal pipe diameter and  $\beta = d/D$ . The measurement noise is distributed as  $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ , where  $\mathbf{R} = \text{diag}([0.1035, 0.1035, 0.1035, R_{venturi}])$ .

## 2.2. Pressure sensor bias

Sensor bias is a constant systematic error in measurement and sensor drift is a time-varying bias in the sensor. Hence, the measurement equation (7) must account for these biases and drift in the following manner, where  $\mathbf{y}_k^d$  is the bias/drift value at time  $k$  and  $\mathbf{y}'_k$  is the measurement equation with bias/drift:

$$\mathbf{y}'_k = \mathbf{y}_k + \mathbf{y}_k^d = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k + \mathbf{y}_k^d \quad (8)$$

To estimate the bias/drift, we augment the state vector and model the bias/drift as a random walk. The augmented model is now given by

$$\mathbf{x}_{k+1}^a = [\mathbf{x}_{k+1}, \mathbf{y}_{k+1}^d]^T = [\mathbf{f}(\mathbf{x}_k, \mathbf{u}_{k+1}, \Delta t), \mathbf{y}_k^d]^T + [\mathbf{w}_k, \mathbf{w}_{y^d,k}]^T \quad (9)$$

where  $\mathbf{w}_{y^d,k} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_y)$  is the distribution describing the steps of the random walk.

## 3. Combining fault detection and observability analysis to state and bias estimation.

The simultaneous estimation of states and biases in all the sensors simultaneously can make the system non-observable. To avoid these problems, this work combines FDI methods and observability analysis with state estimation. Figure 2 illustrates the methodology. The first step consists of the residual generation. The measurement ( $\mathbf{y}'_k$ ), which may have a bias or drift, is compared to the model's predicted measurement ( $\hat{\mathbf{y}}_k$ ). Then this residual is evaluated by a statistical test that is aimed in deciding if the system is fault-free or which sensor has drifted. After that, an observability analysis is performed to evaluate if the identified bias can be estimated. In the positive

case, a new term is included on the state estimator as described in section 2.2 to track the bias/ drift on the failing sensor. Otherwise, the standard state estimation is performed without adding additional bias terms.

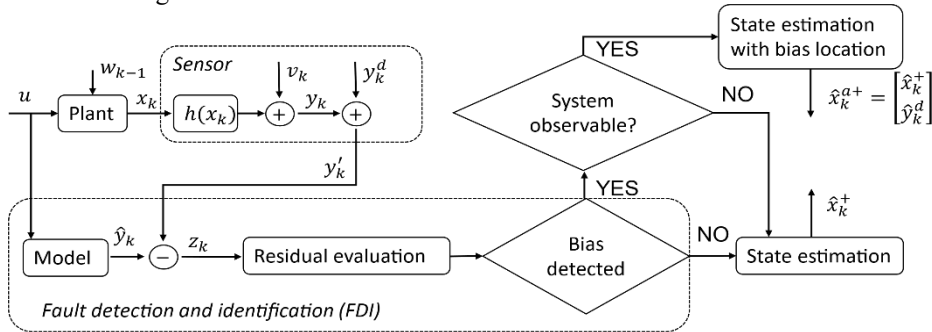


Figure 2: Conceptual flowsheet of how FDI-technique is used to update the structure of the state estimator.

### 3.1. Residual evaluation for bias detection

Three residuals are made, one for each pressure sensor, which consists of the difference between the model prediction and the measurement. Assuming a perfect model, it is possible to identify a bias in the pressure measurements since these residuals should be zero mean in a normal operating condition and different from zero when a bias occurs. To evaluate if the residuals are different from zero, the statistical method called cumulative sum control chart (CUSUM) was used (Blanke, 2016). A Gaussian distribution was assumed to represent the residual distribution ( $\mathcal{N}(\mathbf{0}, \mathbf{R})$ ), and any deviation higher than five times the standard deviation would correspond to a fault.

### 3.2. State estimation and Observability analysis

We used the Unscented Kalman filter (UKF) to obtain the optimal estimate of the true state and bias (Simon, 2006, ch. 14). The bias terms were included on the UKF as time-dependent continuous parameter as described in section 2.2. The concept of observability tells us if we can uniquely determine the state given some measurements, which is a requirement for state estimators to work. For linear systems, observability can be checked by verifying if the observability matrix has full rank. See Simon (2006, ch. 1.7.2) for a definition of the observability matrix and a discussion about observability for linear systems. In our work, we linearized the model and checked the rank of the observability matrix for the linearized model. If the (augmented) system is not observable, we did *not* include the bias term in the state estimator and used the fault-free estimator as Figure 2 shows.

## 4. Results

The results section is divided into two parts. First, in section 4.1, an illustrative simulation shows how a previous bias detection and isolation step can improve the state estimation process. Then, in section 4.2, a Monte Carlo analysis is performed to evaluate the method's performance in different scenarios.

#### 4.1. Estimator with information of bias location

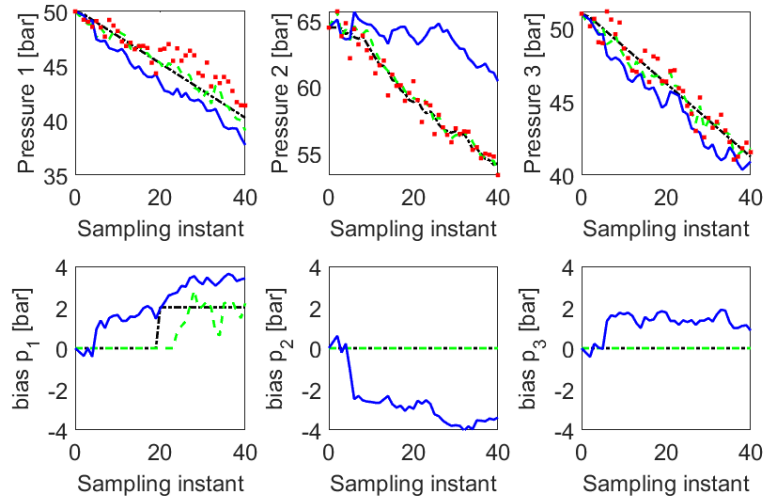


Figure 3- Comparison of true state (black), pressure measurements (red) and the estimates using  $UKF_b$  (blue) and  $UKF_{fdb}$  (green). Bias was introduced on the measurements of sensor PT1.

The model described in Section 2 was solved for a time window of 40 samplings. After 20 samples, a constant bias of 2 bar is added to the pressure sensor one (PT1). Two methodologies are compared in this scenario; first, the UKF with bias estimation ( $UKF_b$ ), where the estimators try to estimate the bias on the three sensors since the beginning of the window. Hence, the augmented state dimension is  $5 + 3 = 8$  through the entire time.

Second, our proposed method described in Section 3 is applied ( $UKF_{fdb}$ ). Here, the state dimension is 5 before sampling time 20 and after a successful FDI and observability procedure it is 6. The results of the simulation is presented in Figure 3. The results indicate that the observability problems of adding new bias parameters to be estimated arise even for a small system like the subsea pump.  $UKF_b$  fails to estimate the states and bias in all three sensors, even before the bias is added to the measurements on sensor PT1. In fact, the observability analysis showed that the states are non-observable when three bias terms are included on the estimator.

However, when just one bias term is included on the estimator, the states and the bias are observable. This explains why the  $UKF_{fdb}$  could obtain a good estimate of the states and the bias on sensor PT1.  $UKF_{fdb}$  had a slight delay in detecting the bias on sensor PT1. After that, the observability is verified, and the information on which sensors are failing is sent to the state estimator.

#### 4.2. Monte Carlo analysis

In order to perform a comprehensive analysis of the methodology, a Monte Carlo analysis was performed. This study compared three approaches:  $UKF_b$ ,  $UKF_{fdb}$  and the “standard” UKF without bias estimation. All the simulations started from the same initial condition, with 100 samplings as a time window. Three hundred random samples of different simulation features were used: time to insert the bias (between 10 and 80 samples), the magnitude of the bias (-2 bar to 2 bar), control profile (pump speed and valve opening), sensor location of the bias (sensor 1, 2 or 3) and bias or drift. The results of the simulations are summarized in Table 1.

The results show that the UKF<sub>b</sub> has poorer performance than the UKF. The system's observability must be checked before including new bias parameters to be estimated. The standard UKF is the best option in cases of non-observable states. However, when a fault detection and isolation method is applied, there is no need to include a bias term in the UKF for each sensor. Instead, only after the bias is detected a new term can be added to the UKF. The UKF<sub>fdb</sub> had a better performance in terms of state estimation, with only a slightly larger error on  $q$  and  $H$ . Besides that, it can also provide a reasonable estimate of the bias in the sensors, which is a valuable information for pump system monitoring.

Table 1: Root-mean-square error states and bias over 300 simulations. \* is “sampling instances”.

	Bias correctly Located [%]	Average delay [*]	RMSE						
			bias [bar]	$p_1$ [bar]	$p_2$ [bar]	$p_3$ [bar]	$\rho$ [kg/m <sup>3</sup> ]	$q$ [m <sup>3</sup> /h]	$H$ [m]
UKF	-	-	-	0.74	0.77	0.75	24.97	2.54	3.15
UKF <sub>b</sub>	-	-	4.89	2.32	3.57	1.92	273.61	2.40	3.08
UKF <sub>fdb</sub>	86.00	14.96	0.52	0.42	0.41	0.44	10.86	2.58	3.21

## 5. Conclusion

An approach to monitor and estimate sensor bias on subsea booster pump system was numerically evaluated. The results suggest that estimating sensor bias without any *a priori* information about which sensors are failing is impractical. Therefore, combining fault detection and observability analysis techniques with sensor estimators can improve the estimation significantly and provide helpful information for system monitoring. A more thorough structural and local observability check should be applied in future works. Also, the performance of the methods for multiple simultaneous faults should be evaluated.

**Acknowledgement:** The authors acknowledge financial support from the Norwegian Research Council through the AutoPRO Project. RDO and GACLR acknowledges financial support from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

## References

- Blanke, M.; Kinnaert, M.; Lunze, J.; Staroswiecki, M. *Diagnosis and Fault-Tolerant Control*. Springer Berlin Heidelberg: Berlin, Heidelberg, 2016.
- Rangegowda, P. H.; Valluru, J.; Patwardhan, S. C.; Biegler, L. T.; Mukhopadhyay, S. 2022. “Development of a Robust Receding-Horizon Nonlinear Kalman Filter Using M-Estimators.” *Industrial & Engineering Chemistry Research* 61, 1808–29.
- Gatzke, E. P.; Doyle III, F. J. 2002 “Use of Multiple Models and Qualitative Knowledge for On-Line Moving Horizon Disturbance Estimation and Fault Diagnosis.” *Journal of Process Control* 12, no. 2: 339–52.
- Simon, D. 2006. *Optimal state estimation : Kalman, H [infinity] and nonlinear approaches*. Hoboken, N.J.: Wiley-Interscience.

# Fusion of pupil and gaze-based features to estimate cognitive workload of control room operators

Mohd Umair Iqbal,<sup>a,b</sup> Babji Srinivasan,<sup>b,c,\*</sup> Rajagopalan Srinivasan,<sup>b,c,\*</sup>

<sup>a</sup>*Indian Institute of Technology Gandhinagar, Gujarat, 382355, India*

<sup>b</sup>*Indian Institute of Technology Madras, Tamil Nadu, 600036, India*

<sup>c</sup>*American Express Lab for Data Analytics, Risk & Technology, Indian Institute of Technology Madras, Tamil Nadu, 600036, India*

[\\*babji.srinivasan@iitm.ac.in](mailto:babji.srinivasan@iitm.ac.in), [\\*raj@iitm.ac.in](mailto:raj@iitm.ac.in)

## Abstract

Process industries are highly hazardous, and these hazards often lead to accidents. Over 70% of these accidents are attributed to human errors. With the advancements in technology and changing role of operators to the one involving an emphasis on cognitive aspects, most of these errors occur due to limitations in cognitive performance. One of the major constructs to understand cognitive performance is the cognitive workload. An increase in cognitive workload often leads to degradation in performance. Eye tracking has been used in several domains to assess cognitive workload. In this work, we propose a methodology to assess cognitive workload of control room operators during tasks that involve tackling process abnormalities. The methodology employs the fusion of metrics obtained from pupil and gaze data. Our results reveal that fusion of metrics provides better accuracies of classifying cognitive workload at three levels—low, medium and high workload.

**Keywords:** human errors, operator performance, cognitive workload, eye-tracking, decision trees

## 1. Introduction

Chemical and process industries are highly hazardous, and these hazards frequently result in disastrous accidents. Even though there have been tremendous increases in safety due to better process design and technical safety approaches, mishaps still happen. A significant majority of these accidents can be traced back to human errors; human errors contribute to around 60-80% of accidents (Jung et al., 2020). Nonetheless, humans form an essential layer of protection, especially in the wake of abnormalities in the process that are often outside the purview of control systems. However, with the advancements in technology in the form of digitalization, and concomitant automation (Khan et al., 2021), the role of human operators has changed more towards monitoring (during regular operation), diagnosis, and prognosis (during abnormal situations). To maintain adequate situational awareness, operators must prioritize, acquire, segregate, and integrate information from multiple sources in distributed control system (DCS). An abundance of information due to digitalization, coupled with the increased complexity of modern plants, makes these tasks cognitively challenging for an operator, especially if they are not highly skilled. Further, the increased automation which often keeps the operator out-of-loop with the system makes it even more challenging for them to develop correct mental models (understanding of the process). As a result, operators' cognitive workload can drastically increase if they are not adequately trained and assessed. An increase in

cognitive workload is known to result in the degradation of performance (Iqbal et al., 2020; Islam et al., 2017). There is a strong association between cognitive workload and human errors. While human errors are a result of mismatch between process demands and human capabilities high cognitive workload, at the most general level, is a result of mismatch between mental resources demanded by the process and those supplied by an operator (Parasuraman, 2008).

Traditional approaches of performance assessment are primarily based on process and operator action-based measures such as number of successes and failures, reaction times, deviations from standard operating procedures, and time taken to complete a task (Iqbal et al., 2021). These ignore the cognitive aspects of performance, such as cognitive workload, decision-making, information acquisition pattern, and operators' mental states while they perform a task. Hence, these also fail to unravel the underlying reasons for shortcomings in operator performance. While subjective assessment-based methods such as National Aeronautics Space Administration Task Load Index, or NASA-TLX (Hart, 2006) are employed for assessing cognitive workload, these are often criticized for being prone to subjective-bias. These can also not be used to assess changes in cognitive workload in real time, and thus have limited practical applicability. Therefore, there is a need to assess cognitive workload of control room operators objectively.

Advancements in sensory technology such as eye tracking, electroencephalography (EEG), electrocardiography (ECG) and electrodermal activity (EDA) have made it possible to tap on physiological measures as surrogates of cognitive performance (especially cognitive workload). Physiological measures are intrinsic and provide critical insights into covert factors that shape operator performance. In this paper, we discuss how fusion of gaze and pupil-based metrics obtained using eye-tracking can be employed to better characterize cognitive workload of control room operators while they control process abnormalities.

## 2. Literature

Researchers in safety critical domains such as aviation, driving and healthcare industries have attempted to understand cognitive workload using eye-tracking due to its non-invasive nature and ease of use. Eye tracking involves tracing the eyes of a person with respect to their head in order to obtain the coordinates of their gaze on an human machine interface (HMI). Eye tracking provides critical insights into visual behavior. Since about 50% of the brain's neural pathways are used for visual processing (Woo, 2019), visual behavior is important to understand cognitive aspects of performance. For instance, during stressful emergency accident situations, expert nuclear power plant operators have a different eye gaze pattern, marked by higher total fixation durations compared to novices (Liu et al., 2019); higher fixation duration reflecting the ability to extract more relevant information. Pupil diameter obtained using eye-tracking also provides crucial information about human cognitive performance. For instance, during simulated surgical tasks, it was found that higher pupil dilations corresponded to increase in complexity of the tasks (Zheng et al., 2015). Experiments conducted to assess air traffic controllers' workload revealed that increase in subjective workload translated into an increase in the amplitude of pupil diameter (Rodríguez et al., 2015). Another study revealed that changes in mental workload can be captured by frequency domain analysis of pupil diameter measurements (Peysakhovich et al., 2015).

Various physiological measures have also been fused for use in machine learning models for classifying operator mental states. For instance, eye gaze and EEG based features have been fused to develop models for differentiating between groups of learners: those who could successfully resolve a problem-solving task involving medical diagnosis via an HMI, and those who could not (Khedher et al., 2019). The study revealed that fusion of features resulted in an improved accuracy of 75% in identifying learners and non-learners. It was observed that using only set of features resulted in poor performance (18% using eye features only and 0% using EEG only) in correctly classifying failure cases (non-learners). Ding et al. (2020) found that fusion of sensors corresponding to ECG and EDA resulted in an accuracy of around 58% for classification of mental workload using neural networks. Similarly, four types of feature sets based on EDA, ECG, respiration, and blood oxygen saturation were used in a study to discriminate between high and low levels of workload corresponding to two different levels of difficulty induced during three different cognitive activities (Game, Numerical test, and Raven's test) (Hirachan et al., 2022). The authors report that respiration and SpO2 based features resulted in individual accuracies of 60% and 56% respectively. There have been a few studies focusing on multi-level classification of workload as well. For instance, McDonald et al., (2020) reported an accuracy in the range of 24% to 56% using several machine learning algorithms for classifying driver distraction based on physiological data. The focus of these studies has primarily been on simpler tasks which do not evolve dynamically.

Process industry operator's role involves dealing with tasks which evolve dynamically and the demands on the operator also vary in response to their actions. The disturbances in chemical processes primarily originate from within the plant rather than because of external environment such as nature of numeric calculations or traffic conditions. This makes the cognitive requirements of control room operators unique (Müller and Oehm, 2019) and makes it challenging to characterize the same. Our previous research on reveals that gaze analysis and pupil diameter can provide critical insights into cognitive aspects of control room operator's performance. For instance, gaze entropy can capture the differences in eye-gaze behaviour between novices and experts (Iqbal et al., 2018). The study found that with learning gaze entropy decreases reflecting increase in expertise owing to decreased cognitive workload. Pupillometry studies indicate that the diameter of the pupil correlates closely with operator's ability to successfully perform the tasks involving bringing process within safe limits (Bhavsar et al., 2015). Likewise, fixation duration and saccade duration can capture changes in operator mental states (decision making) during such tasks (Das et al., 2018). In this paper, we develop a decision tree model that can estimate cognitive workload of control room operators, *at three levels*, using fusion of gaze-based and pupil-based features.

### **3. Experimental Methodology**

We used a simulated ethanol process as an experimental testbed for conducting human subject studies. The HMI via which the operators interact with the process is shown in Fig. 1. Ten participants were involved in the study, each carrying out several repetitive trials. During a typical trial, a participant carried out six tasks (each corresponding to a disturbance scenario). The role of the operator is to bring the process variables within acceptable limits in case of a disturbance which is notified to the operator through an alarm and change in color of the tags (of disturbed variables). During these tasks, participants' eye gaze data and pupil diameter are recorded using an eye-tracker (Tobii 300). In addition, process data, alarm information, and operator actions are also recorded.



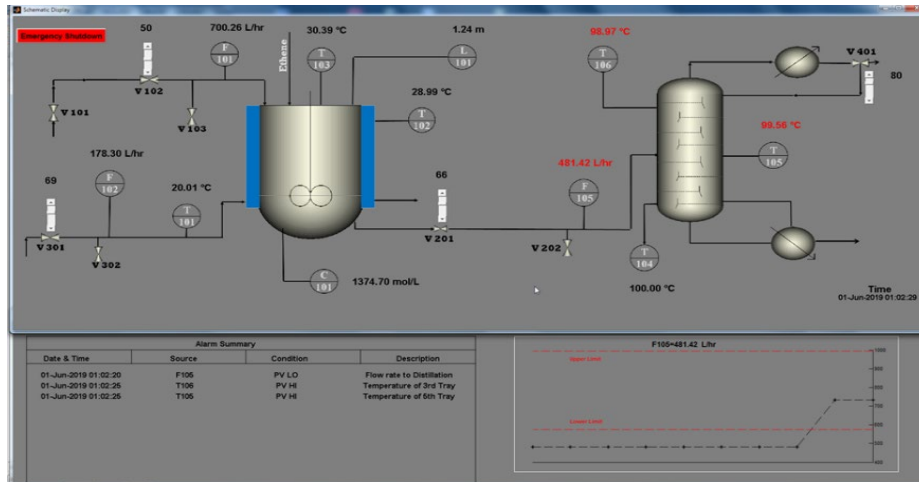


Fig. 1 Human Machine Interface (HMI) of ethanol process

From this information, we calculate integrated absolute abnormality (IAA) which measures the deviation of process variables from acceptable safe limits. We use IAA along with success/failure in a task as a surrogate to classify operators' cognitive workload into three levels- low, medium, and high. Low values of IAA correspond to low workload as these are successfully completed by the operators. On the other hand, if IAA during a task was much higher, operators failed in most of those tasks. Intermediate values of IAA correspond to low workload. Next, pupil and eye gaze-based metrics are extracted for each task (after alarm occurrence) to populate a feature matrix. These metrics for pupil include power spectral densities of pupil diameter and average pupil diameter. For gaze, these include various fixation-based metrics. A summary of these metrics is provided in Table 1. We trained decision tree models based on i) pupil-based features only ii) gaze-based features only and, iii) fusion of pupil and gaze-based features to estimate the cognitive workload of operators as well as assess the impact of fusion.

Table 1. Summary of feature categories

S. NO.	Feature	Feature Description
1	Pupil	P1: PSD of pupil diameter signal in the range 0 – 0.5Hz P2: PSD of pupil diameter signal in the range 0.5 – 1 Hz P3: PSD of pupil diameter signal in the range 1 – 1.5 Hz P4: PSD of pupil diameter signal in the range 1.5 – 2 Hz P5: Average pupil diameter
2	Gaze	G1: Longest Fixation duration/shortest Fixation duration G2: Normalized Fixation Number i.e., No. of fixations/time G3: Normalized Fixation Duration: Total fixation duration/time G4: Standard deviation of fixation duration G5: Mean of fixation duration

## 4. Results

Consider an operator O1 who performs eight trials of the experiment resulting in a total of 48 tasks. For two of the tasks, some data was missing, and therefore are not considered. Each task is labelled based on the value of integral absolute abnormality (IAA) calculated for post alarm phase of the task as discussed in Section 3.

Table 2. Cognitive workload classification accuracies of trained decision tree model

S. NO.	Feature Combination	Overall Accuracy (%)
1	Pupil only	60.87
2	Gaze only	50.00
3	Pupil + Gaze	65.22

As per the methodology discussed in Section 3, eye tracking based features are obtained for all the tasks that this operator performed. These features are then used as an input to train a decision tree and the model is tested using a 10-fold cross validation technique. It involves randomly dividing the data into 10 parts and then using 9 parts to train the model and test it on the remaining 10%. The procedure is repeated 10 times each time reserving a different one tenth of data for testing. The overall accuracies achieved using different combinations of features for operator O1 are shown in Table 2, and the results clearly highlight the benefits of the fusion of gaze and pupil-based features. Fusion of features results in an increase in classification accuracies to about 65%.

A representative decision tree for O1 trained using fusion of features is shown in Fig. 2(a). The topmost node is the root node and consists of pupil and gaze features for all the 46 tasks that the operator performed i.e., it contains a mix of features which correspond to all the levels of workload (0,1, and 2). The first split is made based on feature P4 (Power spectral density of pupil diameter in the range 1.5 to 2 Hz). This indicates that power spectral density of pupil diameter in range 1.5 to 2Hz is an important predictor of cognitive workload for this participant. For the observations for which P4 is less than 0.55 they are allocated to node 2 (on the left) otherwise to node 3. At node 2, the split is made based on G2 (Normalized Fixation Number i.e., No. of fixations/time). For the observations for which G2 is less than 0.44 they are allocated to leaf node representing workload level of 2 (medium workload). If G2 is greater than 0.22, the observations are allocated to leaf node representing workload level of 3. In a similar way, data at node 3 is split based on the learnt rules till leaf nodes are arrived. Leaf nodes represent the predictions of the target variable which in this case is cognitive workload. The confusion matrix for the decision tree trained using fusion of gaze and pupil-based features is shown in Fig. 2(b). The confusion matrix reveals that decision tree trained on fusion of pupil and gaze-based features can classify high workload levels (class 3) with an accuracy of 75%. Thus, the algorithm performs well in identifying poor performances that can pose safety challenges, and can be used to gauge adequacy of training.

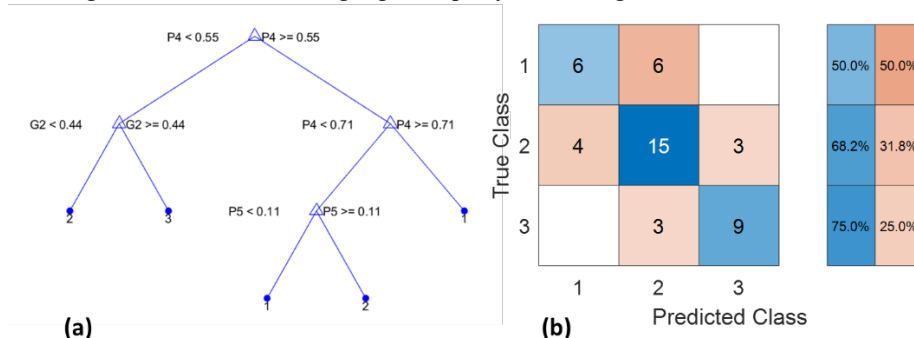


Fig. 2 a) Decision tree trained for operator O1 trained using both pupil and gaze-based features b) Confusion matrix for three levels of cognitive workload

## 5. Conclusion

In this work, we demonstrated how fusion of different categories of features (gaze and pupil) can improve classification accuracies of control room operator's performance as measured by cognitive workload—a precursor to sub-optimal performance. This is the first work of its kind, as it involves multi-level classification of process industry operator performance. The overall achieved accuracy of the model is 65.22% while it can identify poor operator performance (highest workload) with an accuracy of 75% (critical for ensuring adequacy of training). These accuracies are fairly good given the complexities and dynamic aspects of process industry tasks where operator performance is governed by the correctness of the mental models. Similar accuracies (57% for generic model) have been reported for nuclear operators (Braarud et al., 2021). This work has the potential to evaluate operator's expertise, and is thus crucial for ensuring safety in process industries. It also has the potential to guide the application of human factors by providing objective information about operator's expertise and (indirectly) impact of HMI design. In future, we plan to increase the robustness of our model by employing fusion of eye-tracking with electroencephalography (EEG) based features, and identifying other precursors of suboptimal performance. Our previous studies (Iqbal et al., 2020) reveal that EEG based-metrics have the potential to capture performance of control room operators.

## References

- Braarud, P. Ø., Bodal, T., Hulsund, J. E., Louka, M. N., Nihlwing, C., Nystad, E., ... & Wingstedt, E. (2021). An investigation of speech features, plant system Alarms, and operator–system interaction for the classification of operator cognitive workload during dynamic work. *Human factors*, 63(5), 736-756.
- Das, L., Iqbal, M. U., Bhavsar, P., Srinivasan, B., & Srinivasan, R. (2018). Toward preventing accidents in process industries by inferring the cognitive state of control room operators through eye tracking. *ACS Sustainable Chemistry & Engineering*, 6(2), 2517-2528.
- Iqbal, M. U., Srinivasan, B., & Srinivasan, R. (2018). Towards obviating human errors in real-time through eye tracking. In *Computer Aided Chemical Engineering* (Vol. 43, pp. 1189-1194). Elsevier.
- Iqbal, M. U., Srinivasan, B., & Srinivasan, R. (2020). Dynamic assessment of control room operator's cognitive workload using Electroencephalography (EEG). *Computers & Chemical Engineering*, 141, 106726.
- Jung, S., Woo, J., & Kang, C. (2020). Analysis of severe industrial accidents caused by hazardous chemicals in South Korea from January 2008 to June 2018. *Saf. Sci.* 124, 104580
- Khan, F., Amyotte, P., & Adedigba, S. (2021). Process safety concerns in process system digitalization. *Education for Chemical Engineers*, 34, 33-46.
- Liu, Z., Zhao, Q., Zhang, L., Zhang, X., Fan, J., Wang, Q., & Wu, P. (2021). Quantitative Evaluation on the Effect of Experience Under Emergency Situations in NPP Main Control Room Based on Multimodal Data. *Nuclear Technology*, 207(4), 575-581.
- Müller, R., & Oehm, L. (2019). Process industries versus discrete processing: How system characteristics affect operator tasks. *Cognition, Technology & Work*, 21(2), 337-356.
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making*, 2(2), 140-160.
- Peysakhovich, V., Causse, M., Scannella, S. and Dehais, F., 2015. Frequency analysis of a task-evoked pupillary response: Luminance-independent measure of mental effort. *International Journal of Psychophysiology*, 97(1), pp.30-37.
- Rodríguez, S., Sánchez, L., López, P., & Cañas, J. J. (2015, September). Pupillometry to assess Air Traffic Controller workload through the Mental Workload Model. In *Proceedings of the 5th International Conference on Application and Theory of Automation in Command and Control Systems* (pp. 95-104). ACM.

# Quantification of Process Operability Using Flexibility Analysis

Fei Zhao,<sup>a\*</sup> Chentao Mei,<sup>b</sup> Lingyu Zhu,<sup>b</sup> Xi Chen<sup>a</sup>

<sup>a</sup>*State Key Laboratory of Industrial Control Technology, College of Control Science and Engineering, Zhejiang University, Hangzhou, Zhejiang, 310027, China*

<sup>b</sup>*College of Chemical Engineering, Zhejiang University of Technology, Hangzhou, Zhejiang, 310014, China*  
*zhaofeizju@zju.edu.cn*

## Abstract

Industrial plants are commonly subject to uncertainties during operation. Operability and flexibility are two key properties for design models with uncertain parameters. In this study, flexibility analysis is first considered, where a derivative-free optimization method based on flexibility index is designed to identify the largest rectangular flexibility space, while the optimal nominal point of uncertain parameters can be obtained. Then, a novel operability analysis method is proposed to quantify the entire system operability. Based on the critical points of uncertainties in the flexibility space, a line projection method is adopted to locate boundary points of the operational space. The volume of the operational space is calculated by the Delaunay Triangulation method, and the ratio to the physical operational ranges can be used to assess the system operability. The case study shows the proposed method is capable of dealing with the nonconvex operational spaces.

**Keywords:** Process operability; flexibility analysis; flexibility index; derivative-free optimization.

## 1. Introduction

Uncertainties are process disturbances related to variability of process parameters, which represent the degrees of freedom or manipulated variables in an industrial process, such as feed flow rates and temperatures. The uncertainties can lead to mismatch between a plant and the corresponding design model. Flexibility analysis method was proposed as a quantitative framework for measuring the capability of feasible operations (Floudas, et al., 2001; Zhao, et al., 2022), where the uncertain parameters are measured; the control variables can be manipulated to compensate any variation of uncertain parameters during operation. Flexibility index (Grossmann and Floudas, 1987) is a well-known quantitative indicator, which can define a flexibility space with a specific shape, such as a rectangle. Apart from the process flexibility, analyzing the process operability (Gazzaneo, et al., 2020) is another indispensable part of process design. Lima et al. (2010) summarized the similarities and differences between the process flexibility and process operability in great detail. The traditional process operability focuses on the examination of whether a process controller might be able to meet all the quality requirements due to limitations of the process design, in other words, whether all quality outputs are operable in the presence of disturbances; by comparison, the process flexibility is concerned with whether the process is feasible for all the desirable quality outputs due to the uncertainties. Flexibility and operability are two closely linked concepts, which have been recognized to represent two important features in the production process. In a word, with the presence

of uncertainties, the process flexibility quantifies the ability of ensuring the feasibility of quality constraints, and the process operability represents the system implementability of the corresponding flexibility. The process flexibility and operability should be assessed together. The motivation of this study is to quantify the entire system operability.

## 2. Flexibility space description through the DFO solution strategy

For a design model, the flexibility constraint is defined by the following logic expression.

$$\forall \boldsymbol{\theta} \in FS(\boldsymbol{\theta}) \{ \exists \mathbf{z} \in T_z \{ \forall j \in J [g_j(\boldsymbol{\theta}, \mathbf{z}, \mathbf{x}) \leq 0], \forall i \in I [h_i(\boldsymbol{\theta}, \mathbf{z}, \mathbf{x}) = 0] \} \} \quad (1)$$

where  $\boldsymbol{\theta}$  and  $\mathbf{x}$  represent uncertain parameters and state variables, respectively, and the control variables  $\mathbf{z}$  represents the degrees of freedom that can be adjusted during operation to offset different realization of  $\boldsymbol{\theta}$ . The specified set of uncertain parameters and the given physically adjustable ranges of control variables are described by lower and upper bounds. The expected deviations along the negative and positive directions,  $\Delta\boldsymbol{\theta}^-$  and  $\Delta\boldsymbol{\theta}^+$ , can be defined as  $(\boldsymbol{\theta}^U - \boldsymbol{\theta}^L)/2$ . The flexibility index  $F$  can be calculated by the maximum value of  $\delta$  in the following semi-infinite programming problem.

$$\begin{aligned} F(\boldsymbol{\theta}^N) &= \max_{\delta \in \mathbb{R}^+} \delta \\ \text{s. t. } & \max_{\boldsymbol{\theta}} \min_z \max_{j \in J} g_j(\boldsymbol{\theta}, \mathbf{z}, \mathbf{x}) \leq 0 \\ & h_i(\boldsymbol{\theta}, \mathbf{z}, \mathbf{x}) = 0, \quad \forall i \in I \\ & \boldsymbol{\theta}^N - \delta \Delta\boldsymbol{\theta}^- \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}^N + \delta \Delta\boldsymbol{\theta}^+ \\ & \mathbf{z}^L \leq \mathbf{z} \leq \mathbf{z}^U \end{aligned} \quad (2)$$

Grossmann and Floudas (1987) proposed an active constraint strategy, where the lower-level problem is reformulated by the Karush-Kuhn-Tucker (KKT) conditions. Hence, the multilevel optimization problem can be converted into a single-level mixed-integer linear or nonlinear programming (MILP/MINLP) problem, which is given by,

$$\begin{aligned} F(\boldsymbol{\theta}^N) &= \min \delta \\ \text{s. t. } & g_j(\boldsymbol{\theta}, \mathbf{z}, \mathbf{x}) + s_j = 0, \quad \forall j \in J \\ & h_i(\boldsymbol{\theta}, \mathbf{z}, \mathbf{x}) = 0, \quad \forall i \in I \\ & \sum_j \lambda_j \cdot \frac{\partial g_j}{\partial x_k} + \sum_i \mu_i \cdot \frac{\partial h_i}{\partial x_k} = 0, \quad \forall k \in K \\ & \sum_j \lambda_j \cdot \frac{\partial g_j}{\partial z_k} + \sum_i \mu_i \cdot \frac{\partial h_i}{\partial z_k} = 0, \quad \forall k \in K \\ & \sum_j \lambda_j = 1 \\ & \sum_j y_j = n_z + 1 \\ & s_j - M(1 - y_j) \leq 0, \quad \forall j \in J \\ & \lambda_j - y_j \leq 0, \quad \forall j \in J \\ & \boldsymbol{\theta}^N - \delta \Delta\boldsymbol{\theta}^- \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}^N + \delta \Delta\boldsymbol{\theta}^+ \\ & \mathbf{z}^L \leq \mathbf{z} \leq \mathbf{z}^U \\ & \lambda_j \geq 0, s_j \geq 0, y_j \in \{0,1\}, \forall j \in J \end{aligned} \quad (3)$$

Moreover, if  $\boldsymbol{\theta}^N$  is unknown, the above flexibility index problem can be extended to a design centering problem shown in Equation (4), which indicates that the flexibility index model is applied for candidate nominal points,  $\boldsymbol{\theta}^N$ , which are regarded as the decision variables of the outer-level optimization problem.

$$\begin{aligned}
& \max_{\boldsymbol{\theta}^N} \delta \\
& \text{s. t. } g_j(\boldsymbol{\theta}^N, \mathbf{z}, \mathbf{x}) \leq 0, \quad \forall j \in J \\
& \quad h_i(\boldsymbol{\theta}^N, \mathbf{z}, \mathbf{x}) = 0, \quad \forall i \in I \\
& \quad \boldsymbol{\theta}^L \leq \boldsymbol{\theta}^N \leq \boldsymbol{\theta}^U \\
& \quad \text{Equation (3)}
\end{aligned} \tag{4}$$

The feasibility of  $\boldsymbol{\theta}^N$  should be guaranteed during the whole optimization process. Due to the presence of the state variables  $\mathbf{x}$  and control variables  $\mathbf{z}$ , the maximum constraint violation (MCV) of all the inequality constraints should be less than or equal to zero, i.e.,  $MCV(\boldsymbol{\theta}^N) \leq 0$ .

$$\begin{aligned}
& MCV(\boldsymbol{\theta}^N) = \min u \\
& \text{s. t. } g_j(\boldsymbol{\theta}^N, \mathbf{z}, \mathbf{x}) \leq u, \quad \forall j \in J \\
& \quad h_i(\boldsymbol{\theta}^N, \mathbf{z}, \mathbf{x}) = 0, \quad \forall i \in I \\
& \quad \mathbf{z}^L \leq \mathbf{z} \leq \mathbf{z}^U
\end{aligned} \tag{5}$$

After moving  $MCV(\boldsymbol{\theta}^N)$  to the objective function and introducing a penalty coefficient  $M$ , the DFO formulation of design centering problem can be written as

$$\begin{aligned}
& \min_{\boldsymbol{\theta}^N} -F(\boldsymbol{\theta}^N) + M \cdot \max(0, MCV(\boldsymbol{\theta}^N)) \\
& \text{s. t. } \boldsymbol{\theta}^L \leq \boldsymbol{\theta}^N \leq \boldsymbol{\theta}^U
\end{aligned} \tag{6}$$

To increase the likelihood of obtaining the largest flexibility space, the Latin hypercube sampling (LHS) strategy is applied to generate a set of points in the uncertain parameter space. Each feasible LHS point will be treated as a starting point for the DFO solution procedure. Finally, the maximum value of  $\delta$ , i.e., the flexibility index, corresponds to the final rectangular flexibility space, which can be described as

$$\theta_p^- \leq \theta_p \leq \theta_p^+, \quad \forall p \in P \tag{7}$$

where  $\theta_p^-$  and  $\theta_p^+$  represent the largest allowable range for each uncertain parameter,  $p \in P$ . Thus, the vertex set of this rectangular space contains  $2^P$  points, i.e.,

$$\text{Verts} := \{\{\theta_1^-, \theta_2^-\} \times \{\theta_1^-, \theta_2^+\} \times \cdots \times \{\theta_{p-1}^-, \theta_p^+\}\} \tag{8}$$

The optimal nominal point is  $\boldsymbol{\theta}_{op}^N := (\theta_1^N, \dots, \theta_p^N)$ , where  $\theta_p^N = (\theta_p^- + \theta_p^+)/2$ . Together with the feasible LHS points sampled in the rectangular flexibility space, denoted as FPs, we can define a critical point set  $CP$ , i.e.,

$$CP := \{\text{Verts}, \boldsymbol{\theta}_{op}^N, \text{FPs}\} \tag{9}$$

These chosen points are representative for reflecting the system flexibility, especially for the vertices, which may reflect the limitation of operations.

### 3. Quantificational description of operational space

Assuming that the control variables can be adjusted continuously within the physical ranges. For each candidate  $\boldsymbol{\theta}_k$  in the critical set  $CP$ , the corresponding operational space is the feasible region of the following system with respect to  $\mathbf{z}$ .

$$\begin{cases} g_j(\boldsymbol{\theta}_k, \mathbf{z}, \mathbf{x}) \leq 0, \quad \forall j \in J \\ h_i(\boldsymbol{\theta}_k, \mathbf{z}, \mathbf{x}) = 0, \quad \forall i \in I \\ \mathbf{z}^L \leq \mathbf{z} \leq \mathbf{z}^U \end{cases} \tag{10}$$

The solution strategy to quantifying the system operability mainly contains three steps:

(1) *Locate boundary points of operational space*

For each point of uncertain parameters in  $CP$ , given a sampling number, the LHS method is performed over the physical ranges of control variables. All the feasible LHS points that can satisfy Equation (10) form a set  $\mathbf{A}$ . In this work, we focus on exploring points on

the boundary of the operational space. A line projection method is adopted to efficiently locate boundary points. For example, through a pair of points in  $\mathbf{A}$ , e.g.,  $\mathbf{z}_1$  and  $\mathbf{z}_2$ , there are two directions,  $\mathbf{z}_1 \rightarrow \mathbf{z}_2$  and  $\mathbf{z}_2 \rightarrow \mathbf{z}_1$ . We can locate two boundary points by solving the maximization problems. The global optimization solver, e.g., GAMS/BARON, is applied to guarantee the solutions located on the boundary. After enumerating the pairs of points in the set  $\mathbf{A}$ , all the obtained boundary points define the set  $\mathbf{B}$ . The more LHS points are specified, the more boundary points we can obtain.

(2) *Delaunay triangulation for the operational space*

The boundary points illustrate the profile of the operational space, and the nonconvex parts also can be exactly captured as long as the boundary points are enough. In this work, in order to provide an effective way to quantify the operational space, a triangle meshing technique using the boundary points is proposed, which is implemented by the Delaunay triangulation method. Quickhull (Barber, 1996) is one of the most efficient methods for high-dimensional triangulations, which was developed based on the convex hull methods. As a numerical implementation of the Quickhull method in SciPy, Qhull is applied to execute the Delaunay triangulation.

(3) *Quantification of the operational space*

Since the Quickhull method creates a simplex mesh of a convex hull for a set of points, some triangles may be generated beyond the nonconvex feasible space. Thus, we should remove these infeasible triangles through checking the feasibility. In  $\mathbb{R}^d$ , the centroid of the  $k^{\text{th}}$  simplex,  $\bar{\mathbf{z}}^k$ , can be simply calculated based on its vertex coordinates.

$$\bar{\mathbf{z}}^k = \frac{1}{n+1} \sum_{i=1}^{n+1} \mathbf{x}_i^k \quad (11)$$

where  $\mathbf{x}_i^k$  is the coordinate vector for vertex  $i$  of simplex  $k$ ,  $n$  denotes the dimension. We can check the feasibility of the centroid. Thus, the nonconvex parts can be identified, and we can remove the infeasible meshes. Benefit from the Delaunay triangulation and the feasibility tests, the triangle meshing of the entire operational space is obtained. Then, we can calculate the area or volume of each simplex (triangle or polyhedron).

$$V^k = \frac{1}{n!} |\det[(\mathbf{x}_1^k - \mathbf{x}_0^k)^T (\mathbf{x}_2^k - \mathbf{x}_0^k)^T \dots (\mathbf{x}_n^k - \mathbf{x}_0^k)^T]| \quad (12)$$

where,  $k$  is the index of the simplex,  $k \in K$ . Each column of the  $n$ -by- $n$  matrix is the transpose of difference between two row vectors representing vertex  $i$  ( $i = 1, 2, \dots, n$ ) and the reference vertex  $\mathbf{x}_0^k$ , respectively. The reference vertex here defaults to the first point of each mesh. Moreover, the initial operable space is a rectangular space, and the volume can be calculated by the given adjustable ranges of the control variables, i.e.,

$$V_g = \prod (\mathbf{z}^U - \mathbf{z}^L) \quad (13)$$

Thus, we can define the system operability by the following operable space ratio.

$$FR = \frac{\sum_k^K V^k}{V_g} \quad (14)$$

#### 4. Case study

Consider the following inequalities,

$$\begin{cases} 2(z_2 - 2)(z_1 - 2) - \theta_1 \leq 0 \\ (z_2 - 2)^2(z_1 - 2)^2 - \theta_2 \leq 0 \\ 0 \leq \theta_1 \leq 2, 0 \leq \theta_2 \leq 2 \\ 0 \leq z_1 \leq 4, 0 \leq z_2 \leq 4 \end{cases}$$

where  $\theta_1$  and  $\theta_2$  can be regarded as the uncertain parameters, and  $z_1$  and  $z_2$  are the control variables. In order to evaluate the operability, the largest flexibility space of the uncertain

parameters should be first identified. According to Equation (6), the flexibility index is calculated as  $F = 1$ , which means that the flexibility space is identical to the given region, and the optimal nominal point of  $\theta_1$  and  $\theta_2$  is (1, 1).

Next, the operational space can be evaluated based on this flexibility space. Due to the nonlinearity and nonconvexity, it is generally difficult to describe the complete boundary of the operational space. According to the proposed quantification method, the profile of the operational space can be captured through the line projection approach. To fully assess the overall operability, we need describe the operational space at a critical point set of uncertain parameters. According to Equation (9), four vertices, the nominal point and two feasible LHS points are chosen as  $CP$ , i.e.,  $\{(0, 0), (0, 2), (2, 0), (2, 2), (1.0, 1.0), (1.25089, 0.70663), (0.67603, 1.61001)\}$ . Taking the point (2, 2) in  $CP$  as an example, as shown in Figure 1(a), the gray and green points are the initial and feasible LHS points in the space of control variable, respectively. Along any pair of feasible points, two boundary points can be located. Figure 1(a) also shows that the nonconvex parts can be identified. Note that the result is greatly influenced by the LHS points. The more LHS points are specified, the more likely we are to find the exact boundary. Figure 1(b) shows the feasible LHS points and all the boundary points that are identified at seven points of uncertainties in  $[z^L, z^U]$ . which indicates that seven operational spaces are overlapping, and also can verify that different uncertain parameters can correspond to different operational spaces. The outermost points depict the boundary of the entire operational space. It is worth noting that, there is only one operational point for (0, 0) and (2, 0), i.e.,  $z_1=2, z_2=2$ . Because the operational space only contains discrete points, i.e., there is no operating margin, it is very likely to lead the system uncontrollable due to unknown noise. Thus, we need avoid these uncertainties if possible.

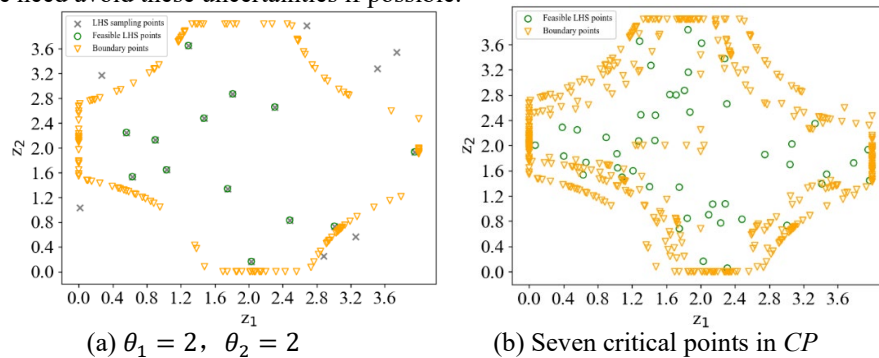


Figure 1. Boundary identification of the operational space.

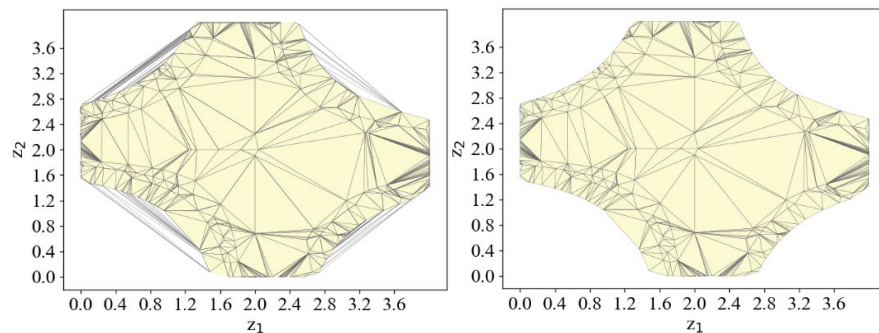


Figure 2. Delaunay triangulation of the nonconvex operational space.



Then, based on all the boundary points, the convex hull including the entire operational space can be meshed through the Delaunay triangulation method, as shown in Figure 2(a). Due to the nonconvexity, the convex hull contains many infeasible parts. Equation (11) can calculate the centroid for each triangle mesh. A feasibility check can be performed by evaluating Equation (10) at the centroid of each mesh, and the infeasible meshes can be identified and then removed, as shown in Figure 2(b). Finally, the total area of the operational space can be calculated as 10.542 by Equation (12). Based on the given ranges, the area of the given operational space  $V_g$  is 16; thus, the operable space ratio FR is 0.659.

## 5. Conclusions

Variability consistently exists in industrial process plants. Flexibility and operability are two closely linked concepts, which are recognized to represent two important features in the production process. The contribution of this work is to connect the flexibility space, the operability space and all the quality requirements closely.

In this article, a novel process operability analysis method is proposed to quantify the system operability. To assess the process flexibility and operability together, the process flexibility is designed as a prerequisite for operability analysis, which can capture the relationships between uncertain process parameters and the quality attributes, and all the specified quality constraints can be satisfied. Then, a line projection method is proposed to directly identify the boundary of the operational space, where some critical realizations of uncertainties in the rectangular flexibility space are considered. The boundary points are applied to calculate the area or volume of the operational space through the Delaunay triangulation method, and the ratio to the given physical operational ranges can assess the operability of the plant.

Challenges in analyzing the process flexibility and operability simultaneously come from the coupling relationships between the uncertain parameters and control variables. The larger flexibility and operability reflect the better performance of the design model. This work illustrates that the larger operable space ratio, the better operability will be. Thus, the proposed method has potential to guide the improvement of the process design if the operability does not meet the requirements. Once the operational space is known, we not only can provide a robust operating strategy to tackle a wide of uncertainties, but also can analyze the operating limitations. Moreover, benefit from the powerful optimization tools, the proposed method is applicable for any scales of process units or systems; however, currently only steady-state process models can be deal with, and the dynamic models are being researching.

## Acknowledgements

The authors gratefully acknowledge the financial support from the National Key Research and Development Program of China (2022YFB3305901), Zhejiang Provincial Natural Science Foundation of China (LZ21B060001), and the Fundamental Research Funds for the Central Universities (226-2022-00038 and 226-2022-00086).

## References

- C.A. Floudas, Z.H. Gümüş, M.G. Ierapetritou, 2001, Global optimization in design under uncertainty: feasibility test and flexibility index problems, *Industrial & Engineering Chemistry Research*, 40, 20, 4267-4282.

- C.B. Barber, D.P. Dobkin, H. Huhdanpaa, 1996, The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software*, 22, 4, 469-483.
- F.V. Lima, Z. Jia, M. Ierapetritou, C. Georgakis, 2010, Similarities and differences between the concepts of operability and flexibility: the steady-state case, *AIChE Journal*, 56, 3, 702-716.
- F. Zhao, M.P. Ochoa, I.E. Grossmann, S. García-Muñoz, S.D. Stamatias, 2022, Novel formulations of flexibility index and design centering for design space definition. *Computers and Chemical Engineering*, 166, 107969.
- I.E. Grossmann, C.A. Floudas, 1987, Active constraint strategy for flexibility analysis in chemical processes. *Computers and Chemical Engineering*, 11, 6, 675-693.
- V. Gazzaneo, J.C. Carrasco, D.R. Vinson, F.V. Lima, 2020, Process operability algorithms: past, present, and future developments. *Industrial & Engineering Chemistry Research*, 59, 2457-2470.



# **Coupling of support vector machine and acoustic models to characterize the droplet size distribution of emulsions using ultrasonic techniques**

Samuel V. Saraiva, Carlos A. M. Silva, Darlan Bonetti, Ana M. F. Fileti, e Flávio V. Silva.

*School of Chemical Engineering, University of Campinas, Campinas, Brazil.*

## **Abstract**

The application of ultrasonic techniques to characterize emulsions concerning the droplet size distribution (DSD) has attracted interest in recent years, mainly due to the advantage of being a non-intrusive technique able to characterize concentrated and opaque emulsions. These characterizations occur by adjusting sound attenuation spectroscopy with the acoustic models. The most used models for this purpose are simplifying more complete models, such as ECAH, mainly due to their mathematical complexity and convergence problems. However, most of these simplified models have restrictions related to the range of applications in the wave propagation regime. Therefore, the objective of this work was to couple the acoustic models to the support vector machines (SVM), which have numerous advantages linked to their successful adaptability to nonlinear data, to acting on this limitation, thus improving the determination of the emulsion DSD. This coupling proved to be satisfactory since the SVM classified all data correctly.

**Keywords:** support vector machine; acoustic models; droplet size distribution.

## **1. Introduction**

The use of ultrasonic techniques to characterize fluids is useful in a wide range of applications, both on a laboratory and industrial scale. Compared to other techniques, ultrasound has the advantages of being a non-invasive and non-intrusive technique, in addition to being able to be applied in-line for real-time measurements.

Concerning emulsion, Challis et al. (2005) present an extensive review of what has been developed in terms of cell construction and acoustic model development. The acoustic models describe the physical phenomena of ultrasound propagation in emulsions in terms of intrinsic and viscous sound dissipation and through the acoustic scattering of waves incident on the drops. The complete model developed to characterize emulsions via ultrasound was

developed by Epstein & Carhart and Allegra & Hawley (ECAH), explained in detail by Challis et al. (2005). The authors also point out that, despite being the most complete, this model has convergence problems.

Silva et al. (2022) studied the application of several simplified models to characterize concentrated water-in-edible oil emulsions, comparing the DSD results obtained using two different techniques, ultrasound, and light scattering. The results have shown a suitable convergence with those obtained from light scattering for situations involving monomodal and bimodal droplet size distributions. However, the choice of the most suitable model for the sample depended on the wave propagation regime.

Among the challenges addressed by the alternative acoustic models to the more complex ECAH model, one main is the wavelength regime identification. It is of fundamental importance for many practical industrial applications. When finding the regime, it is possible to alternate among more suitable models to characterize a particulate system. In this case, intelligent algorithms would be a convenient alternative to mitigate the model's restriction, as suggested by (Challis et al. 2005).

The use of intelligent algorithms based on machine learning models has stood out in many applications (Simon Haykin 2008) and has a good growth perspective for the coming years (Huang et al. 2015). Machine learning models stand out for having suitable adaptability to most of the nonlinear problems in the literature for regression and classification of patterns. Some authors report such machines as universal function approximators. Among these machines, we can emphasize one good performance addressed in the literature, the support vector machine (SVM). The SVM has advantages over other techniques, such as artificial neural networks. The main one is the small number of parameters adjusted in training, implying a smaller data set.

Therefore, the objective of the present work was to develop and investigate the use of empirical models based on SVM to act in the limitations of the acoustic model. The combination of these techniques was not found in the literature. The models based on SVM can assist in wave propagation regime classification to obtain real-time measurements without the prior need to choose specific acoustic models.

## **2. Theoretical foundations**

### *2.1. Acoustic models for wave propagation in emulsion*

Sound attenuation models in the emulsion can be expressed as  $\alpha(w, r)$ , generally in  $Np/m$  unit. In which  $w$  and  $r$  are the angular frequency of sound

(Hz) and droplet radius of the emulsion, respectively (m). In this context, an important dimensionless number in this analysis is the relative wavelength:

$$ka = 2\pi r/\lambda \quad (1)$$

in which  $r$  and  $\lambda$  are the radius of the particle and the wavelength of the ultrasound, respectively. For small particles ( $ka \ll 1$ ), the acoustic scattering can be neglected. This regime is defined as a long-wave regime (LWR). On the other hand, for large particles, values of  $ka \geq 1$ , the effect of sound absorption is neglected, being defined as a short-wave regime (SWR). According to the results obtained from (Silva et al. 2022), the best model that was used in the present work for LWR was coupled phase model (Evans and Attenborough 1997), while the best model for SWR was the elastic scattering model (Feuillade and Clay 1999).

The experimental data must fit an acoustic model to obtain the DSD of the emulsion, and because of these adjustments, the distribution parameters can be calculated. This adjustment will occur by minimizing the sum of squares of errors (Equation 2).

$$\text{Minimize } J = \sum_{w_{min}}^{w_{max}} \left( \alpha_{experimental} - \int_{r_{min}}^{r_{max}} \alpha(w, r) * P(r) dr \right)^2 \quad (2)$$

$P(r)$  is the droplet radius distribution function. Some of the most used distribution functions are unimodal and bimodal. According to some authors (Silva et al. 2022), these distributions represent most particulate systems processes. Due to greater simplicity (fewer variables to be adjusted), unimodal functions are used more in ultrasonic sensor applications. Therefore,  $P(r)$  is the log-normal distribution of the droplet radius.

## *2.2. Support vector machine*

SVM is an intelligent technique used for both classification and regression. Developed by Vapnik (2000), SVM is based on the theory of statistical learning and the principle of minimizing the risk structure, reaching the global solution by transforming the regression or classification problem within quadratic programming. As highlighted in the literature, SVM has a sound theoretical basis and can achieve high performance for complex practical problems (Huang et al. 2015).

In a classification task, the objective of the SVM is to linearly separate the original data in a space with a larger dimensional characteristic. It was then solving a linear separation problem in this space. Therefore, given a data set  $T = \{\mathbf{x}_i, y_i\}_{i=1}^l$ , ( $X \in Y$ )<sup>l</sup>, where  $\mathbf{x}_i \in X = R^N$  is the input vector and  $y_i \in$

$Y = R^N$  the corresponding output, and  $l$  the total number of data pairs, the SVM classification is defined by the hyperplane given by Equation (3).

$$y = f(x) = \mathbf{w}^T \varphi(x) + b \quad (3)$$

In which  $\varphi$  is the mapping function,  $\mathbf{w}$  and  $b$  represent the weight vectors and the term bias, respectively. These parameters are estimated by minimizing the regularized risk function. More details about the algorithm can be found in (Vapnik 2000).

### 3. Experimental procedure and apparatus

The emulsion used in the analyzes carried out in this study was distilled water in sunflower oil with 5% w/w of PGPR emulsifier. Two sets of emulsions were prepared.

The first set of emulsions was prepared to produce a mono-modal distribution with a droplet radius close to 2  $\mu\text{m}$  (set 1). The second set of emulsions was produced to produce an emulsion with droplets above 10  $\mu\text{m}$  (set 2). Details of the procedure used in both production and the physical-chemical properties of the oil and water used in the study can be found in Silva et al. (2022).

The DSD's quantitative analysis was performed using the light Scattering equipment (Mastersizer 2000 - Malvern Instruments Ltd., Malvern, U.K.) that calculates the distribution using the laser diffraction principle.

The system responsible for the characterization of the emulsion through ultrasound was composed of a Parametrics square wave pulse generator (5077PR model), a pair of 6mm-diameter-Parametrics ultrasonic transducers (model V113) with a 15 MHz center frequency, and a PXI-5105 measurement platform from National Instruments. Furthermore, the cell where the emulsion sample was pumped to flow was constructed by Silva et al. (2022).

### 4. Implemented routine and SVM coupling algorithm with acoustic models

The SVM architecture developed to classify the wave propagation regime (output) presented the dispersed phase volumetric fraction  $\phi$  and the area of excess acoustic attenuation (Equation (4)) as inputs. In this case, the input is defined as  $\{\phi, Area\}$  and the output is defined as the wavelength regime classification (LWR or SWR). This architecture was chosen because it presented the best performance in relation to several different architectures studied for input variables in the SVM. Emphasizing that there is a clear relationship between the excess of attenuation and the wave propagation regime for different dispersed phase fractions.

$$Area = \int_{f_{min}}^{f_{max}} (\alpha_{emulsion}(f) - \alpha_{oil}(f))df \quad (4)$$

The SVM will act previously identifying the wave propagation regime and then choosing the model that will be applied. Thus, if the SVM classifies the data as LWR, the model that presented the best performance for this regime will be chosen similarly for SWR. The strategies for coupling the SVM to the acoustic models are summarized in the flowchart of the implemented algorithm (Figure 1).

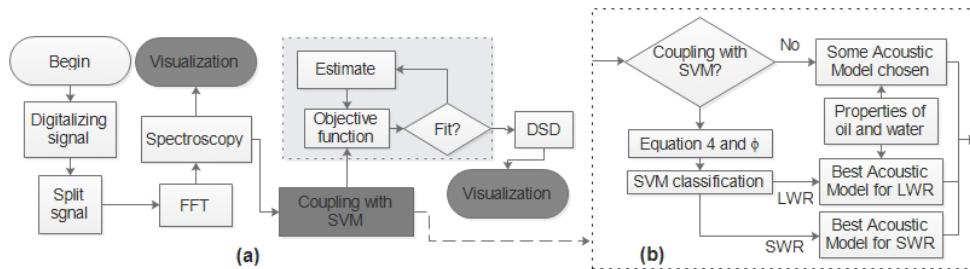


Figure 1 - Flowchart of the strategies of applying SVM.

Following the scheme shown in Figure 1, the ultrasonic signals from the transducers are digitized with the sampling rate and resolution specified previously. The signals are clipped to extract the echoes of interest. The echoes are transformed into the frequency domain, and the spectrum is extracted according to the procedure described by Silva et al. (2020). The spectrum goes to the SVM coupling flow (Figure 1b) and the optimization environment. After the adjustment, the answers are in DSD. The procedure and computational routine were developed in Python.

## 5. Results

### 5.1. Wavelength regime classification using SVM

The attenuation and experimental DSD data were subjected to calculations of the excess attenuation area (Equation 4) as a function of the dispersed phase fraction  $\phi$ . It was observed from Figure 2 that there was a linear separation between the classes of wave propagation regimes. Therefore, the kernel function in the SVM to classify the regime was linear. It can be highlighted that the classification by the SVM of the wave propagation regime proved to be an adequate strategy since the regimes for this data set were separated. The results highlight that the effects of sound energy dissipation in the LRW are more attenuating than in the SWR for a fixed dispersed phase fraction, considering the water-in-oil emulsion.



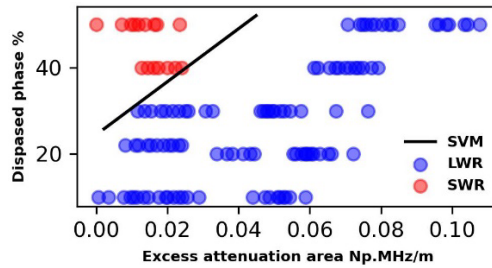


Figure 2 - Results of applying the SVM to classify the emulsions according to the wave propagation regime as a function of the excess attenuation area and the dispersed phase fraction.

### 5.2. Coupling of acoustic and SVM models

Figure 3 shows the development application. In this figure, three curves are shown. The first curvature refers to using the sum of the phase-coupled and Anderson models without coupling with the SVM. The second curve refers to the use of models with SVM coupling. And finally, the last curvature refers to the analyzes made through the Light Scattering.

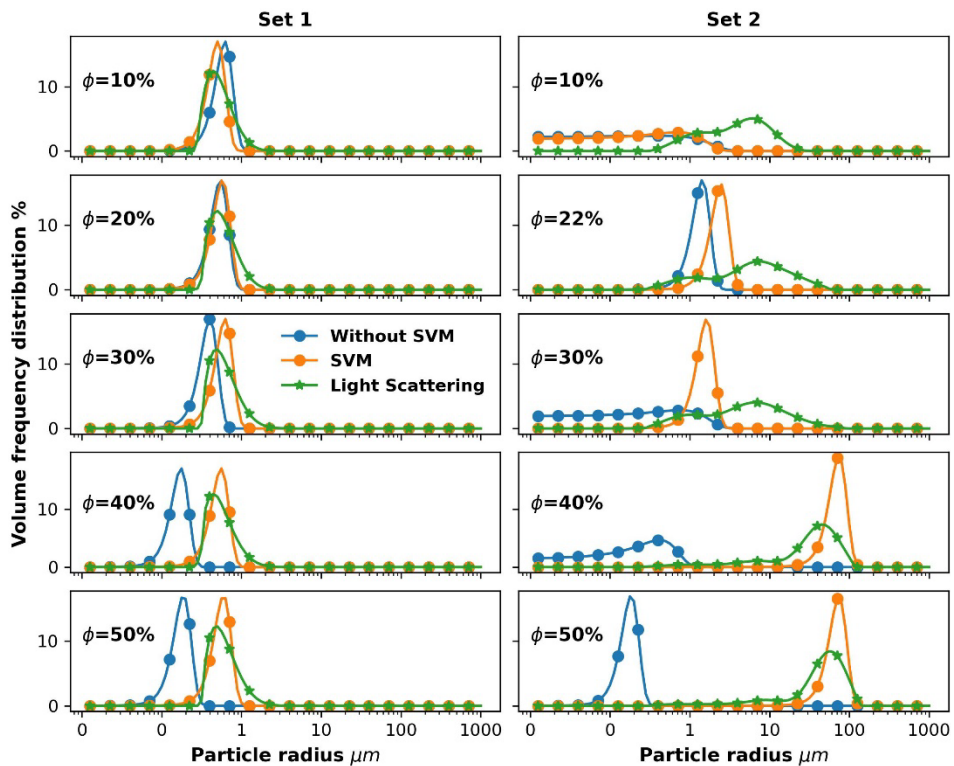


Figure 3 - DSD fitted curves from attenuation spectroscopy, and acoustic models for different phase dispersed concentrations.

## **6. Conclusion**

The present work allowed a distinct and quick procedure for analyzing the ultrasonic signal. Distinct because it does not require prior knowledge of the wave propagation regime since the SVM will act in this limitation. Fast because it requires less computational effort to converge the model in each iteration of the optimization algorithm when compared to the ECAH model. We conclude that these strategies can be implemented in an ultrasonic sensor, provided that the physical parameters, including the volumetric fraction of each emulsion phase, are previously known.

## **Acknowledgment**

We would like to thank Petrobras for the research financial support (Sigitec processes: 2018/00181-9 and 2017/00789-4).

## **References**

- Challis, R. E., M. J.W. Povey, M. L. Mather, and A. K. Holmes. 2005. "Ultrasound Techniques for Characterizing Colloidal Dispersions." *Reports on Progress in Physics* 68(7): 1541–1637.
- Evans, J. M., and K. Attenborough. 1997. "Coupled Phase Theory for Sound Propagation in Emulsions." *The Journal of the Acoustical Society of America* 102(1): 278–82.
- Feuillade, C., and C. S. Clay. 1999. "Anderson (1950) Revisited." *The Journal of the Acoustical Society of America* 106(2): 553–64.
- Huang, Gao, Guang-Bin Huang, Shiji Song, and Keyou You. 2015. "Trends in Extreme Learning Machines: A Review." *Neural Networks* 61: 32–48.
- Silva, Carlos A.M. et al. 2020. "Application of Acoustic Models for Polydisperse Emulsion Characterization Using Ultrasonic Spectroscopy in the Long Wavelength Regime." *Colloids and Surfaces A: Physicochemical and Engineering Aspects* 602.
- . 2022. "Measurements of Bimodal Droplet Size Distribution of Emulsions Using Ultrasonic Spectroscopy in the Long and Intermediate Wavelength Regimes." *Chemical Engineering Science* 252: 117274.
- Simon Haykin. 2008. *Neural Networks and Learning Machines*. 3rd ed. Pearson.

Vapnik, Vladimir N. 2000. *The Nature of Statistical Learning Theory*.  
New York, NY: Springer New York.

# An Approach to Dependable Hybrid Modeling with Application to an Industrial Distillation Column

Mohamed Elsheikh<sup>a</sup>, Yak Ortmanns<sup>b</sup>, Felix Hecht<sup>b</sup>, Volker Roßmann<sup>b</sup>, Stefan Krämer<sup>a,b</sup>, Sebastian Engell<sup>a</sup>

<sup>a</sup>*Process Dynamics and Operations Group, TU Dortmund, 44227 Dortmund, Germany.*

<sup>b</sup>*Bayer AG, 51368 Leverkusen, Germany.*

*Mohamed.elsheikh@tu-dortmund.de*

## Abstract

The use of model-based control approaches can improve the operation of many complex processes significantly but necessitates the availability of accurate dynamic process models, and the development of such models is time-consuming which is a major bottleneck for the application of advanced control. In this paper, a hybrid modeling approach is proposed, where a simple mechanistic model is augmented by a machine-learning model to compensate the plant-model mismatch based on observed data. As the data-based models cannot be trusted outside the region in which sufficiently dense training data was available, we propose to adapt the contribution of the data-based model component based on the closeness to its estimated domain of validity. We demonstrate the benefits of the novel hybrid modeling approach via the case study of an industrial distillation column for solvent recovery.

**Keywords:** hybrid modeling, dynamic modeling, validity domain, distillation column

## 1. Introduction

The stable and efficient operation of chemical processes is of high importance which is even increasing due to the trend towards more dynamic and flexible operations caused by fluctuating markets and prices and availabilities of raw materials and energy inputs. Model-based control can improve the operation significantly, especially in the case of frequent changes of the operating points. However, for the successful application of model-based control and optimization, accurate models are needed. The development of first-principles-based models requires a significant effort and can lead to large differential-algebraic equations (DAE) models that may be difficult to converge. It is therefore of interest to reduce the modeling effort and the computation times while still accurately describing the processes. An approach to reduce the modeling effort is to adapt model parameters online, e.g. by combined state and parameter estimation, to improve the model fit locally. However this is only feasible for a small number of uncertain parameters. An alternative is the use of hybrid models, i.e. combinations of mechanistic and data-based model elements. One option is to embed surrogate models in the mechanistic model to replace complex calculations e.g. of thermodynamic equilibria, see e.g. Nentwich et al. (2019). Winz and Engell (2022) proposed a systematic method to obtain embedded data-based components of hybrid models from measured data. Another frequently employed approach to hybrid modeling is to add data-based elements to mechanistic models in a parallel structure, see e.g. Bock et al. (2021) and Azadi et al. (2022). An overview of hybrid modeling techniques is given in (Von Stosch et al., 2014). A key problem in using data-based models in online applications is the trustworthiness of such models. In general, data-based models can be trusted only in regions that are dense

in training data, called the domain of validity. Several works proposed to represent the domain of validity as box constraints of the data-based model inputs. However, these approaches tend to overestimate the domain of validity. Other approaches are based on estimating the domain of validity as the convex hull of the training data. This has the shortcoming of the inability to exclude gaps in the training data and to represent nonconvex boundaries in the training data. Schweidtmann et al. (2022) suggested a novel approach to estimate the domain of validity by training a one-class support vector machine (SVM). This approach can overcome the mentioned shortcomings.

In this work, we develop a hybrid dynamic model where we combine a simple mechanistic model and a parallel data-based model that represents the deviations of the observed plant dynamics from the mechanistic model. A one-class SVM is trained on low-dimensional projections of the training data that were obtained by principal component analysis (PCA). The output of the one-class SVM is used to weight the contribution of the data-based model to the overall model, achieving a smooth transition between the full hybrid model and the simplified mechanistic model. We investigate the potential of the novel hybrid modeling approach by the application to modeling of the dynamic behavior of an industrial mother liquor distillation column.

## 2. Hybrid model with a weighted contribution of the data-based model

We propose a novel approach for obtaining reliable hybrid process models that consist of a simple mechanistic model and a parallel data-driven model for capturing the deficiencies of the mechanistic model. The developed modeling approach weights the contribution of the data-based model based on the domain of validity of the data-based model. Inside the domain of validity, the full hybrid model is used and outside this domain a smooth fading out of the contribution of the data-based element is implemented. The proposed hybrid process model is defined by:

$$x_{k+1} = f_{diff}(x_k, z_k, u_k, p_k), \quad (1)$$

$$\bar{0} = f_{alg}(x_k, z_k, u_k, p_k), \quad (2)$$

$$y_{k+1} = g(x_{k+1}, z_{k+1}), \quad (3)$$

$$e_{k+1} = f_{ML}(h_k), \quad (4)$$

$$\tilde{y}_{k+1} = y_{k+1} + \gamma_{k+1} e_{k+1}, \quad (5)$$

where  $x_k$  denotes the state vector of the simplified mechanistic model,  $z_k$  represents the vector of the algebraic states of the simplified model,  $u_k$  represents the vector of control inputs,  $p_k$  the vector of varying parameters,  $y_k$  the output vector of the mechanistic model,  $h_k$  the input vector of the data-based model,  $e_k$  the vector of the mismatches between the outputs of the full model and of the simplified model and  $\tilde{y}_k$  the output vector of the hybrid model.  $\gamma_k$  is a scalar time-dependent variable that denotes the weighting factor of the data-based model,  $f_{diff}$ ,  $f_{alg}$ ,  $g$ , and  $f_{ML}$  are nonlinear functions.

### 2.1. The domain of validity of the data-based model component

For the representation of the domain of validity, a one-class SVM is trained on low-dimensional projections of the training data of the input variables of the data-based model. These projections are obtained using principal component analysis (PCA) (Jolliffe and Cadima, 2016). The projection  $T$  of the data is defined as  $T = XW$ , where  $T \in \mathbb{R}^{N \times l}$ ,  $X \in \mathbb{R}^{N \times p}$  is the training data,  $W \in \mathbb{R}^{p \times l}$  is the loading matrix,  $N$  is the number of training samples,  $p$  is the number of input variables in training data and  $l$  is the number of the used principal components. The training data is centered and scaled before applying PCA. One-class SVM is a special case of classical SVMs that deals with the case of having only one class of data, and a maximum-margin hyperplane is sought to separate the training

data from the origin in the feature space (Schölkopf et al., 1999). Schweidtmann et al. (2022) proposed to use the decision function  $f_{DF}(h)$  of the one-class SVM to determine if a data point  $h$  belongs to the domain of validity of the data-based model by checking the inequality  $f_{DF}(h) \geq 0$ .

2.2. Weighting of the data-based model component

We propose to weight the data-based element of the hybrid model based on the function  $f_{DF}$  that indicates the closeness to the domain of validity such that the contribution of the data-based model to the overall model fades out in the regions without sufficient training data. The contribution factor  $\gamma_k$  at time step  $k$  is computed according to

$$\gamma_k = \zeta \Gamma(h_k^{s,c}) + (1 - \zeta) \gamma_{k-1}, \tag{6}$$

where  $\Gamma$  is a sigmoidal function,  $h_k^{s,c} \in \mathbb{R}^{1 \times p}$  is the vector of the scaled and centered inputs of the data-based model,  $\zeta$  is a smoothing factor, and  $0 < \zeta < 1$ . In order to ensure smooth transitions, the weighting factor  $\gamma_k$  at each time step  $k$  is exponentially smoothed. The weighting function  $\Gamma$  is designed as a sigmoidal function:

$$\Gamma(h_k^{s,c}) = \frac{1}{1 + \exp(c \times f_{DF}(h_k^{s,c} W) + \ln(\frac{\epsilon}{1 - \epsilon}))}, \quad c = \frac{2 \ln(\frac{1 - \epsilon}{\epsilon})}{z_{cut\ off}}, \tag{7}$$

where  $z_{cut\ off} \in \mathbb{R}$  and  $\epsilon \in \mathbb{R}$  are the tuning parameters of  $\Gamma$ . The weighting function  $\Gamma$  is chosen to avoid discontinuities in the process model, which can lead to numerical problems as well as abrupt control moves and chattering control behaviour.

3. Case study: Mother liquor distillation column

The considered plant is a mother liquor distillation column that is operated continuously with a varying feed due to the material fed from upstream batch processes. Figure 1 shows a schematic of the distillation column. The feed mixture consists of a low-boiling solvent, a high-boiling component, water, and other impurities. The product streams of the column are the distillate stream that consists of the low-boiling solvent with mass fraction  $w_{lb}$  and water, and the side stream that contains water of mass fraction  $w_w$  and solvent.  $w_{lb}$  and  $w_w$  must be above specified values. The sump stream is a waste stream that contains water, the high-boiling component, and impurities. The mass fraction of the high boiling component  $w_{hb}$  in the bottom stream should be below a certain limit to avoid fouling in the reboiler and the column sump. Level controllers maintain the levels of the distillate receiver and of the column sump. In addition, flow controllers are used to control the flowrates of the reflux stream, the side stream, and the steam stream. A very detailed model that was developed in Unisim Design (Honeywell, 2016) is available and is used as a virtual plant in this paper.

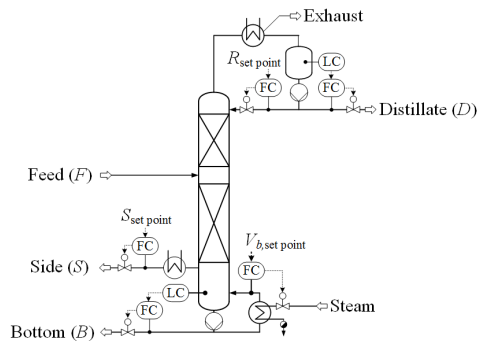


Figure 1: Distillation column schematic

4. Results

The hybrid modeling methodology explained in section 2 is applied to the mother liquor distillation column from section 3. In the simplified mechanistic model we assume that

the manipulated variables (the reflux flowrate  $R$ , the side flowrate  $S$  and the boilup flowrate  $V_b$ ) are perfectly controlled, so the behavior of the corresponding flow controllers is not included in the model. Similarly, the model does not incorporate the level controllers of the distillate receiver and the column sump. The holdup of each equilibrium stage is assumed to be constant. The model is simplified further by neglecting the energy balances of the equilibrium stages. Thus the simplified model contains only the dynamic mass and component balance equations for each equilibrium stage, the condenser and the reboiler. The vapor phase equilibrium is based on the modified Raoult's law where the liquid activity coefficient of each component is computed by the nonrandom two-liquid (NRTL) model. The saturated vapor pressure of each component at each stage is calculated by the Antoine equation. The equilibrium temperature  $T_i$  of each stage  $i$  is obtained by solving the summation equation with respect to  $T_i$ . This results in a semi-explicit index-1 DAE system, consisting of 108 differential equations and 27 algebraic equations. The DAE system is discretized in time using orthogonal collocation. Due to the simplifications performed in deriving the first-principles-based model, there is a structural mismatch between the true dynamics of the plant and the dynamics of the mechanistic model. A nonlinear autoregressive network with exogenous inputs (NARX) is used to predict and to compensate this plant model mismatch. The feedforward element of the NARX model is a feedforward neural network (FNN) consisting of 2 hidden layers, each formed of 30 neurons. The NARX model has 2 outputs which are the predicted errors of the mass fraction of the low boiling component in the top product and of the mass fraction of the high boiling component in the bottom stream  $\{e_{k+1}^{w_{lb}}, e_{k+1}^{w_{hb}}\}$ . The mass fraction of water  $w_w$  in the side stream is not corrected as the prediction of  $w_w$  by the mechanistic model is accurate enough. The NARX model can be represented as

$$e_{k+1} = f_{ML}(e_k, \dots, e_{k-d_e}, \bar{y}_k, \dots, \bar{y}_{k-d_y}, u_k, \dots, u_{k-d_u}, F_k, \dots, F_{k-d_p}), \quad (8)$$

where  $e_k$  is the vector of the mismatches  $\{e_k^{w_{lb}}, e_k^{w_{hb}}\}$ ,  $\bar{y}_k$  is the part of the output vector of the mechanistic model consisting of the mass fractions  $\{w_{lb,k}, w_{hb,k}\}$ ,  $u_k$  is the vector of the manipulated variables  $\{R_k, S_k, V_{b,k}\}$  and  $F_k$  is the feed rate. The numbers of delays  $d_e, d_y, d_u$  and  $d_p$  are chosen to be equal to 10.

The training data was generated by a reinforcement learning agent to ensure a sufficient excitation of the plant (Ortmanns et al., 2023). With a sampling time of 3 minutes, the outputs of the virtual plant were recorded over 14 (simulated) days. The training data is divided into equally sized batches, and the training objective function is the sum of the mean squared errors of the multi-step ahead predictions and of the single-step predictions over a training batch. Figure 2 shows the results of the multi-step ahead predictions of the hybrid model with the full contribution of the data-based model in the training and validation datasets. The mechanistic model has slower dynamics than the true system

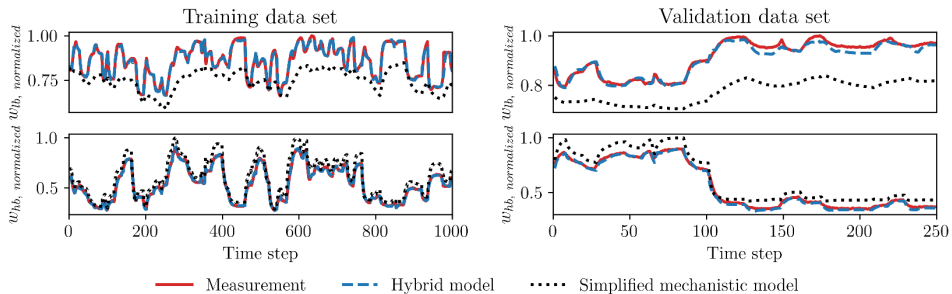


Figure 2: Results of the hybrid model with full contribution of the data-based model.

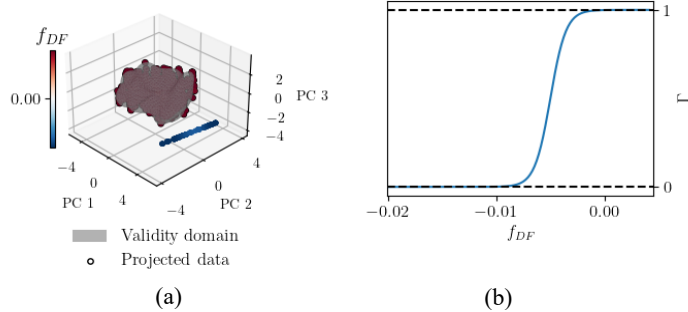


Figure 3: (a) The domain of validity of the NARX model. (b) The weighting function of the NARX model.

especially for the composition for the low boiling solvent in the distillate stream  $w_{lb}$ , while the data-based model compensates the mismatches in  $w_{lb}$  and  $w_{hb}$  very well. The manipulated variables and the feed rate were found to be the most influential inputs of the data-based model via one-at-a-time sensitivity analysis. The first three principal components explain 95% of the variation in the training data of the influential input variables. A one-class SVM was then trained on the three-dimensional projection of the training data to estimate the domain of validity of the data-based model as shown in Figure 3a. The hyperparameters of the one-class SVM were chosen such that  $f_{DF}$  is greater than zero for the projection of the training data. Figure 3a shows the projected training points in red, indicating positive values of  $f_{DF}$  according to the colour map. The projected test data points are shown in blue, yielding negative values of  $f_{DF}$  as the feed and boilup flow rates in the test data are outside the range of the training data. The relation between the weighting function  $\Gamma$  and the one-class SVM decision function  $f_{DF}$  is represented in Figure 3b. The tuning parameters of  $\Gamma$  are  $\epsilon = 10^{-3}$  and  $z_{cut\ off} = -0.01$ , so  $\Gamma$  is equal to  $10^{-3}$  when  $f_{DF} = z_{cut\ off}$ , and the value of  $\Gamma$  is 0.999 when  $f_{DF} = 0$ . The smoothing factor  $\zeta$  was chosen as 0.5.

Figure 4 shows the smooth switching from the full hybrid model to the simplified mechanistic model. As the feed rate changes to the outside of the domain of validity at the 6<sup>th</sup> time step, the weighting factor  $\gamma$  fades out leading to a better prediction quality of the mechanistic model than that of the full hybrid model without adaptation as the predictions of the data-based component get worse outside the domain of validity. Based on this modeling approach, Elsheikh et al. (2023) proposed a nonlinear model predictive controller for the considered distillation column where the domain of validity

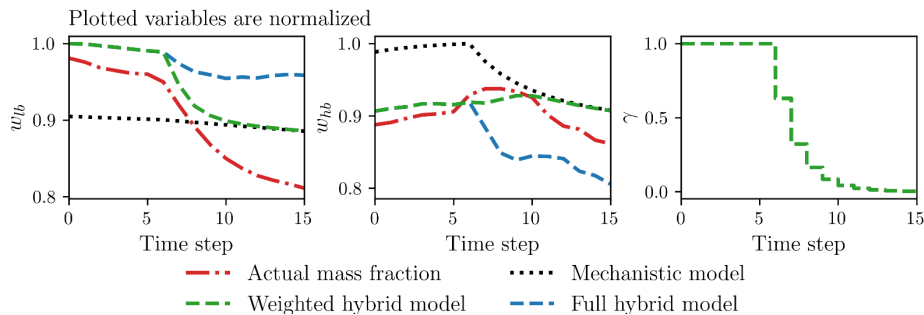


Figure 4: Smooth decay of the contribution of the data-based model.



is in addition extended online based on the observed performance of the data-based component.

## 5. Conclusion

In this paper, the use of a hybrid model is proposed which is formed of a simple mechanistic model and a data-based model which compensates for the plant dynamics that are not represented by the mechanistic model. A one-class SVM is trained on low-dimensional projections of the training data to characterize the domain of validity of the data-based model component. The contribution of the data-based model is faded out based on the output of the SVM. For the case study of an industrial mother liquor distillation column, it is shown that the data-based component of the hybrid model can accurately compensate the deficiencies of the mechanistic model. The proposed weighting of the data-based model improves the prediction of the hybrid model, as the contribution of the data-based model diminishes smoothly in the areas with insufficient training data.

## Acknowledgment

This research has been supported by the project "KI-Inkubator-Labore in der Prozessindustrie - KEEN", funded by the Bundesministerium für Wirtschaft und Klimaschutz (BMWK) under grant number 1MK20014T. This support is gratefully acknowledged.

## References

- P. Azadi, J. Winz, E. Leo, R. Klock, and S. Engell, 2022, "A hybrid dynamic model for the prediction of molten iron and slag quality indices of a large-scale blast furnace", *Computers & Chemical Engineering*, 156, 107573.
- F. Bock, S. Keller, N. Huber, and B. Klusemann, 2021, "Hybrid Modelling by Machine Learning Corrections of Analytical Model Predictions towards High-Fidelity Simulation Solutions", *Materials*, 14(8).
- M. Elsheikh, Y. Ortmanns, F. Hecht, V. Roßmann, S. Krämer, and S. Engell, 2023, "Model Predictive Control of an Industrial Distillation Column Based on a Hybrid Model: Adapting the Domain of Validity", *IFAC- PapersOnLine*, Submitted.
- Honeywell, 2016, "Unisim design simulation basis", Technical report, Honeywell, London.
- I.T. Jollie and J. Cadima, 2016, "Principal component analysis: a review and recent developments", *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065).
- C. Nentwich, J. Winz, and S. Engell, 2019, "Surrogate Modeling of Fugacity Coefficients Using Adaptive Sampling", *Industrial and Engineering Chemistry Research*, 58(40), 18703–18716.
- Y. Ortmanns, M. Elsheikh, S. Engell, J. C. Schulze, R. Hille, F. Hecht, V. Rößmann, and S. Krämer, 2023, "Reinforcement Learning for Process Control: Reducing the Training Duration with Pretraining", *IFAC- PapersOnLine*, Submitted.
- B. Schölkopf, R. C. Williamson, A. Smola, J. Shawe-Taylor, and J. Platt, 1999, "Support Vector Method for Novelty Detection", *Advances in Neural Information Processing Systems*, vol 12.
- A. M. Schweidtmann, J. M. Weber, C. Wende, L. Netze, and A. Mitsos, 2022, "Obey validity limits of data-driven models through topological data analysis and one-class classification", *Optimization and Engineering*, 23(2), 855–876.
- M. von Stosch, R. Oliveira, J. Peres, and S. Feye de Azevedo, 2014, "Hybrid semi-parametric modeling in process systems engineering: Past, present and future", *Computers & Chemical Engineering*, vol 60, 86–101.
- J. Winz and S. Engell, 2022, "A methodology for gray-box modeling of nonlinear ODE systems", *32nd European Symposium on Computer Aided Process Engineering*, vol 51, 1483–1488.

# Integrating mineral mining and metallurgical supply chains: a *qualogistics* approach

Mohammed Yaqot, Brenno C. Menezes

*Division of Engineering Management and Decision Sciences, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar Foundation, Qatar*  
*bmenezes@hbku.edu.qa*

## Abstract

The unbalance between supply and demand of raw materials and products worldwide culminates in colossal challenges in global supply chains. Engineering management that relies on lean design and operation is essential for the long-term viability of the value chain to keep enterprises' operations within an acceptable profit margin at the production level in both natural resource exploration and processing transformation sites. In this work, quantity and quality dimensions of mineral raw material feedstocks are considered from a logistics perspective for an integrated mining and metallurgical supply chain. Such a complete producer-processor integration of raw materials, amounts, and properties of the concentrated mineral diet to be fed in metallurgical furnaces, follows a sequence of mined-to-fed resource management with logistics and quality constraints to be met. This includes (1) exploration and processing in the mining sites, (2) shipping of concentrated mineral pellets, and (3) logistics and mixtures of materials before the furnace feeding in metallurgical sites. In this operational decision-making set, mixed-integer solutions for resource management include the following sequence. First, it involves the assignment of equipment to be shared in different quality exploration fields (e.g., trucks and conveyor-belts). Second, conveying the crushed-ore to be mixed-stocked using shuttle-conveyor-belts in simultaneous positioning to match the qualities of the stockpiles to be ground, floated, and processed into concentrated pellets. Finally, shipping of the blended mineral pellets and metallurgical site logistics and blending mixtures before the furnace feeding. In this work, an enterprise-wide optimization of a unified mining and metallurgical of raw materials is proposed for an improved stockpiling and blending in both supply chains. Thus, their integration can be achieved to maximise the ratio of ores (in the mining field) to metals (in the metallurgical site) output while minimising operational and capital costs.

**Keywords:** Mining fields, metallurgical sites, quality-logistics.

## 1. Introduction

Most production mining operations aim for a round-the-clock operation, seven days a week, 365 days a year. In practice, semi-autogenous milling facilities run at about 95% capacity (run-time) with an efficient crushing plant running at around 85% (Jenike and Johanson, 2022). These figures should thus be extremely near to, if not greater, the actual percentage of time that the plant works after the start-up phase. However, plant availability is just one factor determining plant run-time; coarse ore's availability (in terms of quantity, but also quality) to feed the semi-autogenous grinding (SAG) milling plant or the fine crushing plant is equally crucial importance.

Stockpile management is a critical concept in the bulk materials handling supply chain at both resource exploration sites and processing plants. The port-based stockyard can

represent a mega project investment ( $\geq 1$  billion USD) with large-scale machinery and automation, e.g., autonomous haul trucks and conveyor belts, etc. Given such a significant investment in infrastructure and technology, the more efficiently raw materials can be processed through the stockyards, the less operational cost and asset installation, the faster the payback, and the higher the operating profits will be. In accordance with the industry-wide aim of optimising asset utilisation, businesses are examining how stockpile management might enhance critical key performance indicators (KPIs) such as utilisation against availability ratios. In order to preserve market dominance, manufacturers are enhancing their stockpiling strategies, taking use of technological advancements in equipment and control systems to renew stockyard infrastructure, and strengthening their capacity to do so through increased material flow and quality control (Morley and Arvidson, 2017). Automation of all or a portion of stockpile operations, replacement of older stacker-reclaimer configurations with newer frameworks, and faster, more accurate stockpile accounting are among the upgrading aims (Menezes et al., 2019). Certain mines with unique processing needs have implemented or are developing bulk ore sorting technology to pre-concentrate low-grade ore from stockpiles, bringing it to a level that is economically viable for milling (Russell, 2018).

The stockyards in both production and metallurgical sites serve as material buffers, reserves, or blendstocks between the mining and processing of the crude-ore or mineral pellets in smelting furnaces. Mining companies are faced with the challenge of balancing the benefits provided by maintaining stockpiles, e.g., guarding against supply fluctuations while also managing costs associated with stockyard management. Ore blending is an essential technique in today's large production mines. It is necessary for the mine's life cycle to be extended and dilute various harmful elements (e.g., sulfur, silica, etc.) found in different ores from the same mine. In this sense, it is also used to combine desired qualities from various ores to provide a high quality and quantity feed. Ore blending reduces input material variability and ensures a consistent flow into the processing unit. It is consequently critical that mines maintain a precise blending strategy that meets the demands of both the mine and the metallurgical processing plants.

In this work, an integrated planning for a unified mining and metallurgical of raw materials by an improved stockpiling and blending design is discussed. Thus, additional safety, stability, and predictability can be achieved to maximise the ratio of ores (in the mining field) to metals (in the metallurgical site) output while minimising the operating and capital costs. The tangible and intangible results of this integration can be applied to fuel and petrochemical industries, food processing, mining-metallurgical parks, etc., that integrates information to be included in strategic, tactical, and operational planning studies between upstream and downstream assets design and operations.

## **2. Ore blend scheduling at mining processing sites**

Kelly and Menezes (2019) have developed a smart conveying and stockpiling flow process that combines process optimisation to leverage control and automation of a cyber-physical system that allows the inventory level control of multiple stockpiles per belt. The model can optimally manage and automate the inbound recipe (by the synchronised mixtures of solid streams) and outbound delivery of mined material including, coal, iron ore, bauxite, etc. Mineral processing facilities are typically complicated systems that must contend with internal and external disturbances, a large number of cyclical loads, and a wide range of interactions. Furthermore, the reducing ore grades as well as the continual variation of raw materials are pushing the boundaries of processing facilities to their limits even further. Through mining supply chain, each process characteristics, which miners

are oriented with at each stage, offer opportunities for improvement in terms of reducing costs, energy intensity, and solve the problem of a shortage of qualified and experienced instrumentation technicians, operators, and metallurgists (Rockwell Automation, 2021).

Such control application considers advancements in manufacturing toward the Industry 4.0 (I4.0) mandate involving information and computing technology (ICT) and mechatronics (MEC) evolving together with advances in modelling and solving algorithms (MSA). For such, recent advances in network (flowsheet) optimisation as well as in solving algorithms and computer-aided resources allow fast solutions of complex discrete decision propositions. Moreover, the ICT expansion can provide the demanded velocity of the communication of the online data measurement of complete process networks for the massive volume and variety of information from the plant, which permits the viability of the addressed control strategy. Typically, the objectives of mine planning are to establish a production schedule that culminates in the selection of ore bodies that ensures the timely supply of ore of the required grade quantity and quality to the processing plant. An appropriate ore stockpiling design can significantly reduce milling time, enhance flotation processes, and raise concentrators' production efficiency and economic performance. A complete integrated smart manufacturing system among MSA, ICT, and MEC through sensing, optimising, and actuating cycles can be achieved to precisely control stockpiling process. The I4.0 deployment facilitates a wide range of control strategies for the quantities and qualities of coarse-ore in conveyor belts. This results in a total free of human-being decision, control, calculation, actuation, etc., and moving from one to multiple stockpiles per belt. Figure 1a illustrates a single shuttle-conveyor tripper car apparatus design considering quantity-logic (QL) only, optimised in Kelly and Menezes (2019) and demonstrated in a proof-of-concept demonstration in a laboratory scale in Yaqot et al. (2022). Figure 1b shows two synchronised robotic apparatus considering a blend of both quantity logic quality phenomena (QLQP).

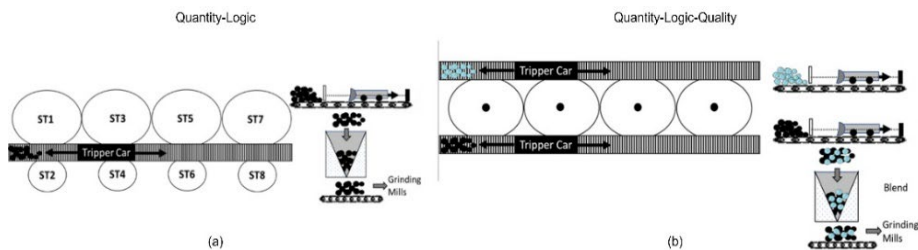


Figure 1. a) Process network with semi-continuous feeding and segregation of solids, from Kelly and Menezes (2019). b) Process network with semi-continuous feeding and blending of solids.

### 3. Integrating mineral mining and metallurgical supply chains

In light of copper's present and potential future significance in the *electrification of everything*, as an example of one metallurgical industry, there is a need to optimise the scheduling and feed mix blending of concentrates in a copper refinery integrated to the mining operations and controls. The first stage aims to integrate metallurgical planning by means of improving operational and control strategies at the mining site, which in turn potentially yields a lean production by reducing needs of mixtures of streams in the metallurgical sites and, therefore, super dimensioned infrastructure as in today's design. Due to the non-integrated optimisation between supply and demand of raw materials and products in the mining and metallurgical supply chain, besides the super dimensioned infrastructure, remaining blending steps are necessary in metallurgical sites. The

flowsheet layout of the copper-refining plant (see Song et al., 2018) is comprised of 2 ships, 6 port piles containing initial inventories or holdups, 1 pre-blender, 3 non-concentrate materials, 6 bins, and 1 smelter, as seen in Figure 2.

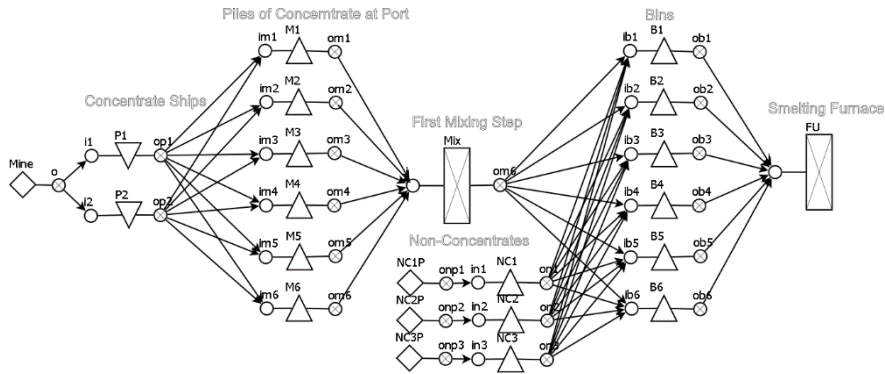


Figure 2. Network sheet for blend-scheduling of a copper-refining plant.

Therefore, quantity and quality constraints of mineral raw material feedstocks must be considered from a *qualogistics* perspective for an integrated mining and metallurgical supply chain. For this producer-processor integration of raw materials, both amounts and properties of the concentrated mineral diet to be fed in metallurgical furnaces follows a sequence of mined-to-fed resource management with logistics and quality constraints to be met. As seen in Figure 3a, this includes: (1) exploration and processing in the mining sites, (2) shipping of concentrated mineral pellets, and (3) logistics and mixtures of materials before the furnace feeding in metallurgical sites. Figure 3b shows the concatenation of solutions to be integrated as proposed, as example, in Menezes et al. (2017) for an enterprise-wide optimization connecting raw materials operations and their processing considering quality throughout the crude-oil supply chain.

In this operational decision-making set integrating mining and metallurgical supply chains, mixed-integer solutions for resource management include the following sequence. First, it involves the assignment of equipment to be shared in different quality exploration fields (e.g., trucks and conveyor-belts). Second, it demands conveying the crushed-ore raw material to be mixed-stocked using shuttle-conveyor-belts in simultaneous positioning (as illustrated in Figure 1a) to match the qualities of the stockpiles to be ground, floated, and processed into concentrated pellets. Finally, it depends in shipping and metallurgical site logistics and blending mixtures before the furnace feeding as illustrated in Figure 2.

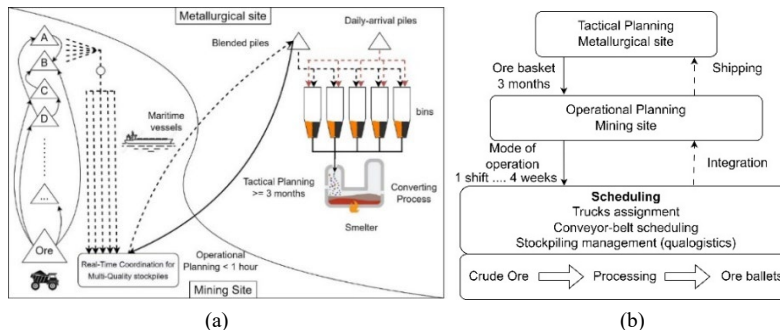


Figure 3. Integrating scheduling of stockpiles in mining-metallurgical sites.

#### 4. Tactical metallurgical and operational mining problems

As proposed in Figure 3a, the tactical planning determines the recipe of segregated crushed-ore raw materials from the mines, whereby it is considered 4 different components (Cu, S, Fe, Rest). In the furnace (FU), there are amounts of quality balances from Kelly et al. (2018) as a linear approximation to find the P1 to P5 (crushed-ore raw materials) and NC1 (non-concentrated or recycled material) that are made of minerals (A to F) or alloys (NC1) in which the results of the FU feed components made by 4 different elements are given as inputs of the operational mining example in Figure 5, whereby the profiles of the level of services of the equipment are shown in Figure 6.

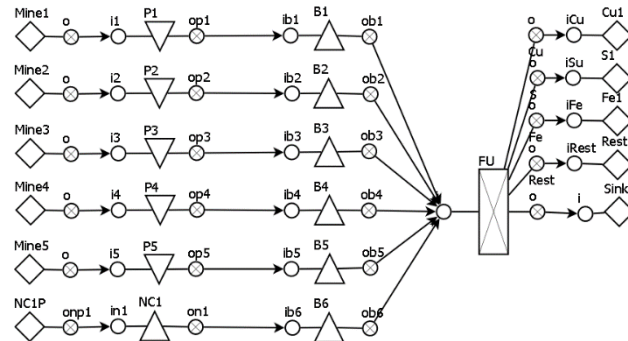


Figure 4. Tactical metallurgical planning to determine FU feed recipes in the furnace.

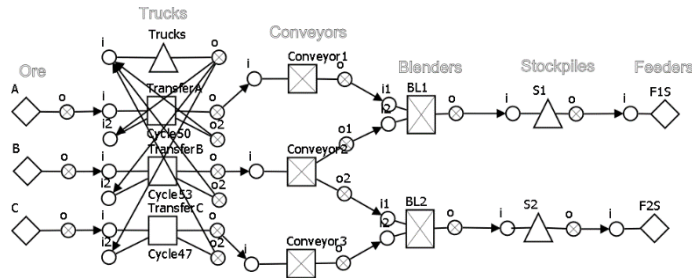


Figure 5. Operational mining scheduling for exploration and process of crushed ore.

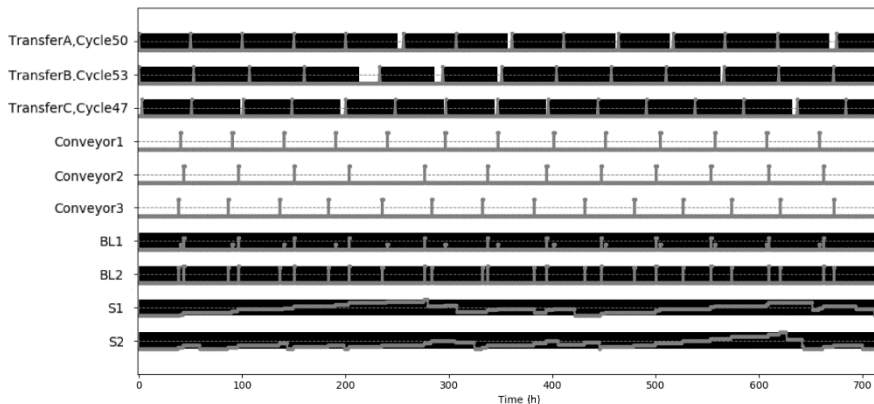


Figure 6. Level of services of the equipment in the operational mining.

## 5. Conclusion

The disparity between the global supply and demand for raw materials and finished goods creates enormous hurdles for global supply networks. In order to maintain an appropriate profit margin at the production level in both resource exploration and processing transformation sites, it is crucial that businesses use key performance indicators that are based on lean design and operation. As a result of recent developments in the coordination of integrated operations, a comprehensive view of integrated supply and demand throughout the value chain is now attainable, from resource reserves all the way to custom-made feedstocks of blended components intended for final products. This concept, which integrates information to be included in strategic, tactical, and operational planning, scheduling, and control studies between upstream and downstream assets' design and operations, is also used in the fuel and petrochemical sectors, as well as the food processing industry, mining-metallurgical parks, and other similar areas. From a logistics point of view, this study looks at the quantity and quality of mineral raw material feedstocks for an integrated mining and metallurgical supply chain. A mined-to-fed resource management procedure with logistical and quality constraints must be satisfied for such a fully integrated producer-processor raw material, quantity, and property of the concentrated minerals diet to be fed in metallurgical furnaces. Therefore, maximising the ore-to-metal ratio in the mining field and metallurgical site while lowering operational and capital expenses necessitates a higher level of safety, stability, and predictability.

## References

- Jenike and Johanson, 2022, Importance of coarse-ore stockpile design in mining mega-projects. Bulk Material Engineering. Accessed 20/11/2022 at <https://jenike.com/importance-of-coarse-ore-stockpile-design-in-mining-mega-projects/>.
- Kelly, J.D., Menezes, B.C., 2019, Automating a shuttle-conveyor for multi-stockpile level control, Computer Aided Chemical Engineering, 46, pp. 1153-1158.
- Kelly, J.D., 2005, July. The unit-operation-stock superstructure (UOSS) and the quantity-logic-quality paradigm (QLQP) for production scheduling in the process industries. In Multidisciplinary International Scheduling Conference Proceedings: New York, United States. 327, pp. 327-333
- Kelly, J.D., Menezes, B.C., Grossmann, I.E., 2018. Successive LP approximation for nonconvex blending in MILP scheduling optimization using factors for qualities in the process industry. Industrial and Engineering Chemistry Research, 57(32), pp.11076-11093.
- Menezes, B.C., Grossmann, I.E., Kelly, J.D., 2017, Enterprise-wide optimization for operations of crude-oil refineries: closing the procurement and scheduling gap, Computer Aided Chemical Engineering, 40, pp.1249-1254.
- Menezes, B.C., Kelly, J.D., Leal, A.G., 2019. Identification and design of industry 4.0 opportunities in manufacturing: examples from mature industries to laboratory level systems. IFAC-PapersOnLine, 52(13), pp.2494-2500.
- Morley, C., Arvidson, H., 2017, Mine value chain reconciliation-demonstrating value through best practice. In Proceedings of the Tenth International Mining Geology Conference, pp. 279-292.
- Rockwell Automation, 2021. Model Predictive Control for Mining: Process Optimization Solutions. Accessed 20/11/2022 at <https://www.rockwellautomation.com/>.
- Russell A. Carter, 2018, Staying on Top of Stockpile Management: Smarter, quicker solutions emerge for measuring and controlling stockpile size and quality. Engineering and Mining Journal. Accessed 20/11/2022 at <http://www.womp-int.com/story/2018vol12/story026.htm>.
- Song, Y., Menezes, B.C., Garcia-Herreros, P., Grossmann, I.E., 2018. Scheduling and feed quality optimization of concentrate raw materials in the copper refining industry. Industrial & Engineering Chemistry Research, 57(34), pp.11686-11701.
- Yaqot, M., Franzoi, R. E., Islam A., Menezes, B. C., 2022, Cyber-Physical System Demonstration of an Automated Shuttle-Conveyor-Belt Operation for Inventory Control of Multiple Stockpiles: A Proof of Concept, IEEE Access, 10, pp. 127636-127653.

# A simple two-parameter steady-state detection algorithm: Concept and experimental validation

Evren Mert Turan<sup>a</sup>, Johannes Jäschke<sup>a</sup>,

<sup>a</sup> *Department of Chemical Engineering, Norwegian University of Science and Technology, Sem Sælandsvei 4, 7491, Trondheim, Norway*  
*johannes.jaschke@ntnu.no*

## Abstract

Automatic detection of steady state periods is a necessary step for many tasks related to real time operation of a process. Based on the insight that the time series of a controlled process variable at steady state resembles a mean reverting process, we fit a first order auto-regressive model to a window of data and use the Dickey-Fuller test to test for this mean reverting property. We compare the proposed approach on two synthetic data sets and one experimental data set and find that the method performs well in comparison to methods in the literature. As the method is computationally inexpensive, only requires two parameters, and is interpretable, we suggest that this is an effective steady state detection tool.

**Keywords:** Steady state detection, Process Monitoring.

## 1. Introduction

The identification of steady state periods in industrial process operations is an essential initial step for many typical tasks, e.g. parameter estimation, abnormal event identification, and real time optimization. Incorrect identification of steady state can severely impact operation, e.g., updating a steady state model with dynamic data can lead to reduced profits and/or unstable operation. Although steady state can be reliably identified by human operators (in low-dimensional data), the automatic identification of steady state, especially with substantial amounts of online data in high dimensions, is still challenging. In this work we use the Dickey-Fuller test to identify steady state time windows, and demonstrate its effectiveness in comparison to other methods in the literature based on their performance on experimental and synthetic data.

Most approaches in for steady state detection in the process engineering literature use 1) a statistical test, or 2) a classification based machine learning method. In this work we focus on the former due to their interpretability, and simple transfer between applications due to their relatively small amount of tuning parameters. Methods relying on statistical tests either tend to perform this test on characteristic values of the system or parameters of a model fitted to online data (Rhinehart, 2013). A key difference between such methods is in the construction of the statistical test, i.e. whether steady state is the null or alternate hypothesis. Although most methods have steady state as the null hypothesis, identification of steady state by the alternate hypothesis is a statistically stronger claim (Rhinehart, 2013).

In this work we fit an auto-regressive (AR) model and use the Dickey-Fuller test to test the null hypothesis that the process is transient. The method is simple to tune, having only



two parameters: the size of the time window and the significance level used in the statistical test. In process engineering, the Dickey-Fuller test has been used in fault detection, however to the authors knowledge it has not been applied to steady state detection.

## 2. Background literature

We briefly outline three methods from the steady state detection literature that make use of a statistical test. In all these methods some quantity based on recent measurements is determined (e.g. a slope), and a statistical test is performed on this quantity (Rhinehart, 2013). Furthermore, in the first two methods we consider a rolling time-window of  $N$  recent measurements.

### 2.1 Method 1: Slope of a line

The simplest approach to steady state detection is to fit a linear model to the data:

$$y_k = mt_k + c \quad (1)$$

where  $y$  is a measured process variable,  $t_k$  is the  $k^{\text{th}}$  time point,  $m$  is a slope, and  $c$  the intercept. After fitting this line by linear least squares, one can check the null-hypothesis of steady state by testing if the slope is zero using a t-test at some significance level. However, this approach violates some assumptions, in particular the  $y$  variables at different time points are not independent, i.e.  $y_{k+1}$  depends on  $y_k$ , and  $y_{k-1}$  and so on.

### 2.2 Method 2: Kelly and Hedengren

In this approach, we test if the process can be described as being at some value, and subjected to independent, identically distributed white noise (Kelly and Hedengren, 2013). As before, we consider the linear model in equation 1. To estimate the parameters averages of the intercept,  $c$ , and gradient,  $m$ , are used. Note that if there is gradient is zero then the intercept is the mean value. If the data is sampled uniformly in time by  $\Delta t$ , then these are given by:

$$m = \frac{1}{N\Delta t} \sum_{k=2}^N y_k - y_{k-1} \quad (2)$$

$$c = \frac{1}{N} \left( \sum_k^N y_k - m\Delta t \right) \quad (3)$$

The standard deviation of the noise,  $\sigma_n$ , can then be estimated as:

$$\sigma_n = \sqrt{\frac{\sum_k^N (y_k - mt_k - c)^2}{n - 2}} \quad (4)$$

Lastly, steady state identification is performed pointwise by checking if the deviation of the data point from the intercept is within a factor of the standard deviation of the noise. This cut-off point is given by a Student's critical value,  $t_{crit}$ , at some significance level and degree of freedom  $N$ :

$$s_k = \begin{cases} 1 & \text{if } |y_k - c| \leq t_{crit} \sigma_n \\ 0 & \text{else} \end{cases} \quad (5)$$

$$S = \frac{1}{N} \sum_{k=1}^N s_k \quad (6)$$

where  $S$  represents the fraction of time in the window “at steady state”. If  $S \geq 0.5$ , then the entire window is at steady state.

### 2.3 Method 3: Cao and Rhinehart

This method relies on the idea that the ratio of two different estimates of the time series’ variance at the same time point should be close to one when the process is at steady state (Cao & Rhinehart, 1995). The first estimate of the variance at  $t_k$ ,  $\sigma_{1,k}^2$ , is given by the filtered squared difference between the process measurement, and a filtered value of the measurement,  $y_{f,k}$ . First the filtered measurement is calculated, and used to calculate a filter of the mean square deviation,  $v_k^2$ :

$$y_{f,k} = \lambda_1 y_k + (1 - \lambda_1) y_{f,k-1} \quad (7)$$

$$v_k^2 = \lambda_2 (y_k - y_{f,k})^2 + (1 - \lambda_2) v_k^2 \quad (8)$$

Assuming the process is at steady state and the measurements are uncorrelated, then the variance,  $\sigma_{1,k}^2$ , can be directly estimated as (Cao & Rhinehart, 1995):

$$\sigma_{1,k}^2 = \frac{2 - \lambda_1}{2} v_k^2 \quad (9)$$

The second estimate of the variance,  $\sigma_{2,k}^2$ , is given by filtering the difference between two consecutive measurements:

$$2\sigma_{2,k}^2 = \lambda_3 (y_k - y_{k-1})^2 + (1 - \lambda_3) 2\sigma_{2,k-1}^2 \quad (10)$$

Lastly, the ratio of these two variances are taken to calculate an R statistic:

$$R_k = \frac{\sigma_{1,k}^2}{\sigma_{2,k}^2} \quad (11)$$

The null hypothesis is that the process is steady, corresponding to a ratio close to 1. If  $R_k$  is greater than some critical value,  $R_{crit}$  then the null hypothesis is rejected.

## 3. Proposed method

The proposed method is based on the insight that the time series of a controlled process variable at steady state, subject to stochastic disturbances, resembles that of a *mean reverting process*, i.e. it tends to some mean point despite the stochasticity of the system. In the example of a controlled process unit, this would be due to the relevant controllers rejecting disturbances, or the system settling to a new operation point.

The first step in the approach is to find the mean value in the time window,  $\mu$ , and subtract this from each data point:

$$\tilde{y}_k = y_k - \mu \quad (12)$$

Then we consider the first order auto-regressive (AR) model:

$$\tilde{y}_{k+1} = p\tilde{y}_k + n_k \quad (13)$$

where  $\tilde{y}_k$  is the transformed measurement at time  $t_k$ ,  $n_k$  is some random variable with mean zero and finite variance, and  $p$  is a variable estimated by linear least squares on the data in the time window. A value of  $|p| < 1$  means that the process is mean reverting, as as the absolute disturbance from the mean will deterministically decrease at each iteration. After estimating  $\hat{p}$ , we then perform a one-sided confidence test on the null-hypothesis that  $|p| = 1$ , i.e. the process is in a transient state, by calculating the test statistic  $\tau$ :

$$\tau = (\hat{p} - 1) \sqrt{\frac{(N - 2) \sum_{k=2}^N \tilde{y}_k^2}{\sum_{k=2}^N (\tilde{y}_k - \hat{p} \tilde{y}_{k-1})^2}} \quad (14)$$

If the test statistic is less than the critical value from the Dickey-Fuller distribution then the null hypothesis is rejected, i.e. the process is at steady state. The critical values can be found from response surfaces fit to the distribution (MacKinnon 2010). The asymptotic results of this test do not require assumptions of normality or homoscedasticity (MacKinnon 2010). Lastly, note that if the process variable is controlled off-set free to some set-point, then one could use the set-point instead of calculating the mean value.

#### 4. Results

We compare the proposed method (DF) against the line slope, Cao & Rhinehart (CR, 1995) and Kelly & Hedengren (KH, 2013) methods on two synthetic data sets and one experimental data set.. For brevity these acronyms are used throughout this section. For these data sets we label steady state and transient periods based on 1) knowledge of when the system inputs are changing and 2) inspection of the data in the time windows. For the methods with time windows and significance level as hyper-parameters we use a time window of 30 seconds, and significance level of 95%. For the method of Cao & Rhinehart (1995) we use the recommended values of  $\lambda_1 = 0.2, \lambda_2 = \lambda_3 = 0.1, R_{crit} = 2$ .

##### 4.1 Low level of normally distributed noise

We first consider a synthetic example with a low level of Gaussian noise. This is an easier task than the others and therefore serves to give a baseline of the methods performance. The results are summarized in Figure 1. From a visual inspection, the proposed DF method performs the most consistently, with the worst performance by the slope method. For all the methods the primary source of error is due to the delay that occurs when the system goes from transient to steady state (around 180s, 360s, and 540s in Figure 1). In contrast, the transition from steady to transient state is captured without a significant delay. The transition to steady state is harder to detect due to the presence of the transient in the time window, as this heavily influences the fit of the models and the filtered variance. Note that if the DF method is used with a set-point instead of the mean then this delay would be reduced, i.e. the delay for this approach comes from taking the mean of the data.

##### 4.2 Moderate level of t-distributed noise

In this comparison we consider the same process as in section 4.1, but now use t-distributed noise, with a larger variance, instead of Gaussian noise. The t-distribution has heavier tails than the normal distribution leading to a high probability of “high-magnitude” noise, leading to a more challenging classification. The experiment is summarized in Figure 2, with “spikes” due to change in distribution clearly shown in Figure 2a. The aim of using this distribution is 1) to test the methods against higher noise levels, 2) to test against non-normally distributed noise.

Apart from the general decrease in performance, the clearest difference when comparing Figure 2b vs 1b is that DF, KH, and CR have a higher rate of false identification of steady state. This does not occur with slope method, as this approach is extremely sensitive to noise. As in section 4.1 the DF method has the best performance.

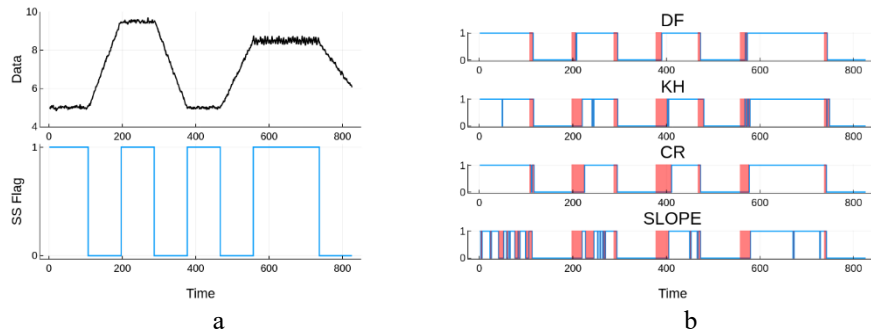


Figure 1: Synthetic data with low levels of Gaussian noise and identified steady state periods (A), and steady state predictions (B). In A and B, a steady state flag of 1 indicates steady state. The shaded red regions in B indicate a mismatch between the predicted and true system state

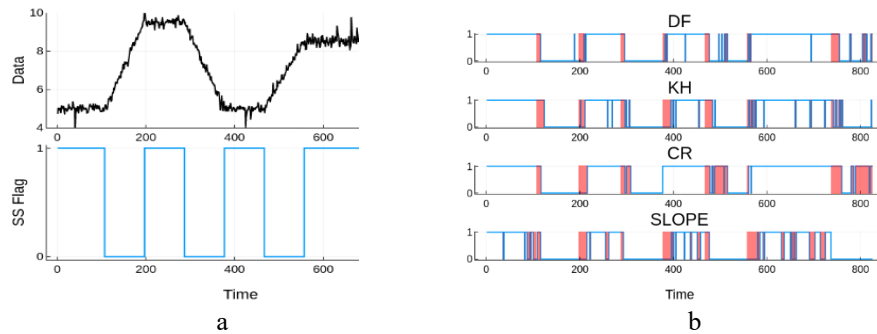


Figure 2: Synthetic data with moderate levels of t-distributed noise and steady state periods (A), and steady state predictions (B).

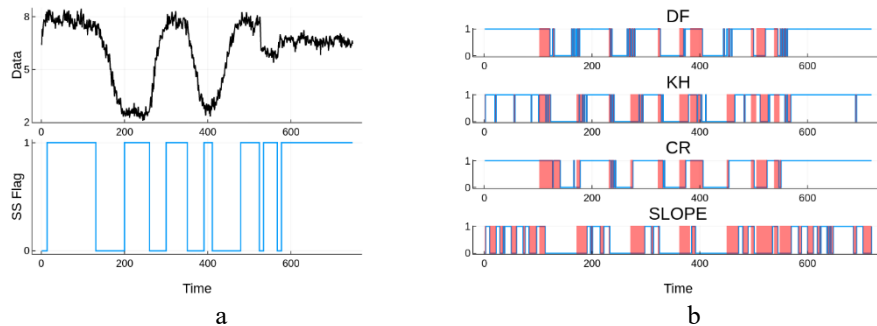


Figure 3: Experimental data with steady state periods (A), and steady state predictions (B)

### 4.3 Experimental data

Lastly, we consider the steady state algorithms applied to real experimental measurements from a lab rig with the experiment summarized in Figure 3, and Table 1.

Table 1: Summary statistics of the steady state detection algorithms applied to the experimental data, shown in Figure 3. The bolded entries indicate the best method for a statistic

STATISTIC	DF	KH	CR	SLOPE
PRECISION	0.85	0.83	0.81	<b>0.86</b>
RECALL	<b>0.90</b>	0.78	0.89	0.44
F1 SCORE	<b>0.87</b>	0.81	0.85	0.58
$\phi$ COEFFICIENT	<b>0.60</b>	0.46	0.51	0.29

This is a real system with a relatively large amount of noise, leading to worse performance compared to the synthetic results by all the methods. In addition there are two step changes between 500 and 600 seconds, which results poorer performance in this section, as shown in Figure 3. A positive aspect of the CR and DF methods are that they are more consistent as they have much less false flags within the SS and transient periods.

Summary statistics of the methods on this data set are shown in Table 1 and serves as quantitative evidence of the visually better performance of the DF method in Figure 3. Precision is the fraction of correct SS predictions over all SS predictions, while recall is the fraction of correct SS predictions over all true SS periods. The F1 score is the harmonic mean of precision and recall. Lastly the  $\phi$  coefficient is a balanced metric that requires reliable performance in both SS and TS prediction. In general, the proposed DF method performs the best, except for the precision metric where the simple line slope method performs similarly due to rarely identifying steady state, as shown in Figure 3.

## 5. Conclusions

In conclusion, we present the application of the Dickey-Fuller test for use in steady state detection and compare it with other methods in the literature on a range of examples. Based on summary statistics of these methods, the Dickey-Fuller test performs the best overall. The method is simple to implement and only requires two-hyperparameters. Further work could extend the approach to multivariate system using approaches suggested in the literature (Rhinehart, 2013, Kelly and Hedengren, 2013).

## Acknowledgments

The authors acknowledge the support of the Norwegian Research Council through the AutoPRO project.

## References

- S. Cao and R.R. Rhinehart, 1995, An efficient method for on-line identification of steady state, *Journal of Process Control*, 5(6), pp.363-374.
- D.A. Dickey and W.A. Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a), pp.427-431.
- J.D. Kelly, and J.D. Hedengren, 2013, A steady-state detection (SSD) algorithm to detect non-stationary drifts in processes, *Journal of Process Control*, 23(3), pp.326-331.
- J.G. MacKinnon, 2010, Critical values for cointegration tests, Queen's Economics Department Working Paper, No. 1227.
- R.R. Rhinehart, 2013, Automated steady and transient state identification in noisy processes, *American Control Conference*, pp.4477-4493.

# Residence Time Distribution characterization in a Continuous Manufacturing tableting line using PCA and PLS-DA modeling

Pau Lapiedra Carrasquer<sup>a</sup>, Satyajeet S. Bhonsale<sup>a</sup>, Liang Li<sup>b</sup>, Jan F.M. Van Impe<sup>a</sup>

<sup>a</sup> *BioTeC+, KU Leuven, Gebroeders De Smetstraat 1, Gent 9000, Belgium*

[jan.vanimpe@kuleuven.be](mailto:jan.vanimpe@kuleuven.be)

<sup>b</sup> *Janssen Pharmaceutica, Turnhoutseweg 30, Beerse Antwerp 2340, Belgium*

[lli110@its.jnj.com](mailto:lli110@its.jnj.com)

## Abstract

In this work, a new approach to characterize the residence time distribution (RTD) using supervised and unsupervised learning techniques was investigated. The NIR spectral data from the feed frame was used to capture the dynamic response of a continuous tableting line.

The data-driven approaches investigated in this study using MPCA and PLS-DA were able to successfully capture the main variation of the step change performed during the RTD experiments. Its good performance compared with the traditional RTD characterization methodology offers the potential to be used to predict the RTD parameters. This could contribute to making the RTD determination procedure more efficient and reduce the need to make offline measurements.

**Keywords:** data-driven modeling, NIR, MPCA, continuous manufacturing, tableting

## 1. Introduction

The progressing adoption of continuous manufacturing in the pharmaceutical industry still faces a series of challenges to achieving its full potential. Tracing any disturbance along the line in real-time is essential to guarantee the quality of the final tablets, as it contributes to significantly reducing the amount of product to be discarded, making continuous operation a more sustainable option compared to the current batch manufacturing. The Residence Time Distribution (RTD) is a widely used method that defines the probability distribution of the time that a particle stays inside a system. The RTD can be determined experimentally by measuring the concentration of the product over time using process analytical tools (PAT) such as Near Infrared (NIR) (Bhalode et al.).

Most continuous tableting lines have NIR probes located in the feed frame of the tablet press. The spectral data obtained from these measurements are normally used to measure the blend uniformity of the powder material right before being compressed. This probe operates in-line, and it can measure every second. This high measurement frequency offers the possibility to extract more information about the dynamic response of the process compared with the information obtained from the API concentration measurements of the limited number of samples extracted at the end of the line. The use of in-line NIR measurements has the potential to provide a real-time prediction of the concentration of the drug substance in the tablet (Hetrick et al.). In this study, a novel approach was explored to exploit all the information from the spectral data to capture a more detailed dynamic response to a step change experiment. We propose to use the NIR

spectral data to predict the dynamic response of a step change in the drug substance concentration to characterize the RTD. To achieve that, supervised and unsupervised data-driven approaches were used and compared. The unsupervised learning method, a variation of PCA called Multiway Principal Component Analysis (MPCA) and the supervised method, Partial Least Squares Discriminant Analysis (PLS-DA) were applied to the NIR spectral data. Both techniques were compared to see which one can better capture the dynamic response of the step change to be used for RTD modeling.

## 2. Theory and Methods

### 2.1. RTD model

Determining the RTD process model is a critical step in product and process development, and it is an essential part of the control strategy to detect non-conforming material at the outlet of the CM line. Developing a clear connection between the different inline measurements (i.e., process parameters and PAT data) and the product characteristics at specific locations is a key aspect. When a drop in the product quality is detected at a certain point, the affected fraction of the product needs to be discarded. This means that the travel time of the product through the line, and the amount of material needs to be known to accurately discard the affected material. This information is obtained from the RTD characteristics of the material in the specific production line. Traditional models that describe RTD are based on the exit-age distribution  $E(t)$ , described by Equation (1), where  $C(t)$  is the time-dependent concentration of a tracer at the outlet of the unit operation.

$$E(t) = \frac{C(t)}{\int_0^{\infty} C(t) dt} \quad (1)$$

The RTD was generally modelled as an ideal PFR + 2 CSTRs in series. Its expression can be written as in Equation (2). Where  $\theta$  is the time delay, and  $\tau_1$  and  $\tau_2$  represents time constants.

$$E(t) = \frac{e^{-\frac{t-\theta}{\tau_1}} - e^{-\frac{t-\theta}{\tau_2}}}{\tau_1 - \tau_2} \quad (2)$$

The calibrated RTD model is product specific and consists of two parts: determination of RTD model parameters and determination of RTD model uncertainty. Once calibrated, the RTD model parameters can be used to calculate the % potency at the tablet press outlet on the CM line. Additionally, the RTD model uncertainty is used to calculate the out-of-specification limits for tablet quality. A series of experiments are performed at different conditions to calibrate the RTD model for a given product.

### 2.2. Experimental data

The dataset used in this study was obtained from 13 RTD characterization runs for a Direct Compression setup. For each run, a step change was applied to the API mass flow and its response was measured in terms of API concentration (%). This concentration is what is usually used to fit the RTD model in Equation (2). To obtain these measurements, a number of tablets are sampled, and their concentration is measured offline. To develop and test the new approach, NIR spectral data of the same 13 runs obtained from the probe located in the feed frame of the tablet press was used. For every run, the measurement of absorbance over a certain wavelength range was measured during the duration of the step change study. This results in a three-dimensional data structure: runs (or batches), measurements over time, and variables (in this case the wavelengths), which can be written in the form of a three-way array.

Table 1. Step changes applied to the API mass flow for every run

Run name	Step change	Run name	Step change	Run name	Step change
Run 1	100%-->85%	Run 6	100%-->130%	Run 11	100%-->70%
Run 2	85%-->100%	Run 7	100%--> 85%	Run 12	100%-->115%
Run 3	100%-->70%	Run 8	100%-->85%	Run 13	115%-->100%
Run 4	100%-->115%	Run 9	100%-->85%		
Run 5	115%-->100%	Run 10	85%-->100%		

2.3. MPCA and PLS-DA

Principal component analysis (PCA) is a commonly used unsupervised data-driven method for dimensionality reduction. PCA transforms the original data into a latent variable model and a residual error. In this model, the latent variables are oriented in the direction that gives the greatest variance of the scores. In other words, it offers a way to summarize the main trends of the data in a reduced number of variables. Due to the three-dimensional nature of the spectral data, a variation of the classic PCA called Multiway PCA (MPCA) was chosen. Commonly used for batch process monitoring, MPCA, unfolds the three-way array into a large 2-dimensional matrix. Due to different durations of the RTD runs, a variable-wise unfolding was chosen (Figure 1). In this way, PCA can be used for the resulting 2D unfolded matrix regardless of the run duration.

The MPCA method was compared to the PLS-DA. PLS is a supervised learning data-driven method that also achieves dimensionality reduction through the latent variable space. The main difference with PCA is that in PLS the latent variables are extracted by obtaining the directions of the largest covariance between input and output. The discriminant analysis variation of PLS is used for classification and the output vector is defined as the class labels to be classified. In this study, the class labels corresponded to the setpoint concentrations before and after applying the step change (Pedersen et al.).

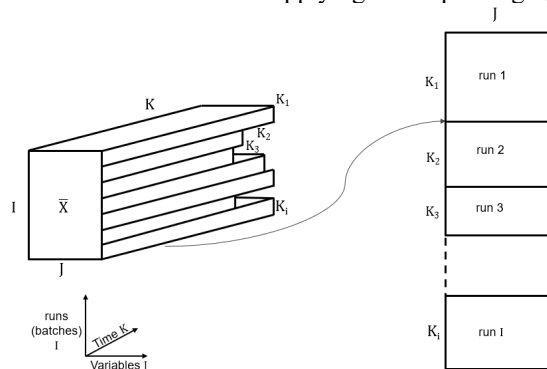


Figure 1. MPCA variable-wise unfolding of a dataset with unequal experiment duration

3. Results and discussion

3.1. MPCA and PLS-DA to NIR data

3.1.1. Pre-treatment

The first step in the data pre-treatment was to omit the measurements in which the line was stopped. To filter out the noise in the spectra data and find the most relevant information for the given product, extreme wavelength values were not considered.

A standard normal variate (SNV) was applied to the dataset. It is a common pre-processing step in spectral data analysis that neutralizes the offset of NIR signals and corrects scattering effects and makes all spectra from different runs comparable. In this



method, a normalization of the data is performed by subtracting each spectrum by its own mean and dividing it by its own standard deviation.

Prior to applying MPCA and PLS-DA, and after applying the SNV for each run, mean centering and scaling were applied to normalize the data along the different runs.

### 3.1.2. MPCA vs PLS-DA

The MPCA model was applied to the pre-treated data. Following the Wold's R criteria, the optimal rank of the model was found at 3. However, with only 2 latent variables, the model was able to explain 94% of the system variation. Similarly, PLS-DA model was trained and cross-validated using only the data corresponding to the steady state, so the model could find the correlations between the data and the label vector. The optimal rank was found to be 2. Once the rank of the PLS model was fixed at 2, the model was applied to the whole spectral data. The results of both models can be visualized by plotting the 1<sup>st</sup> latent variable against the 2<sup>nd</sup> (Figure 2).

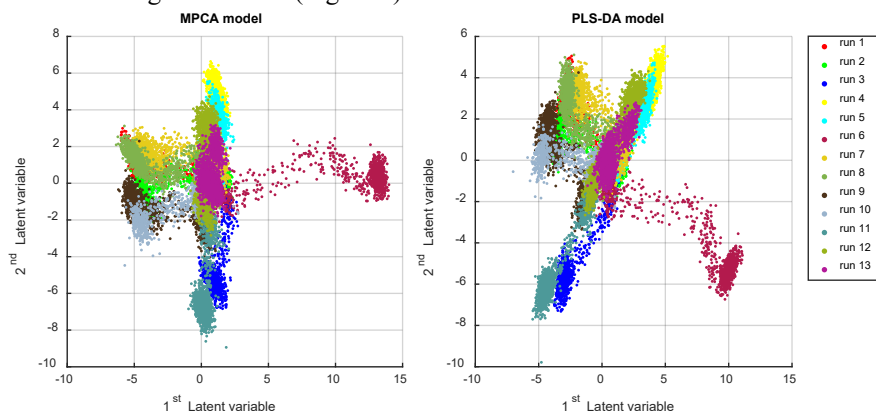


Figure 2. Scores of 1<sup>st</sup> LV against 2<sup>nd</sup> LV for both MPCA (left) and PLS-DA (right) models. The long shape of the clusters corresponds to the time evolution of the data during the step change, meaning that the edges of every cluster group the time points during steady state (before and after the step change) and the intermediate points correspond to the moment of the dynamic step change

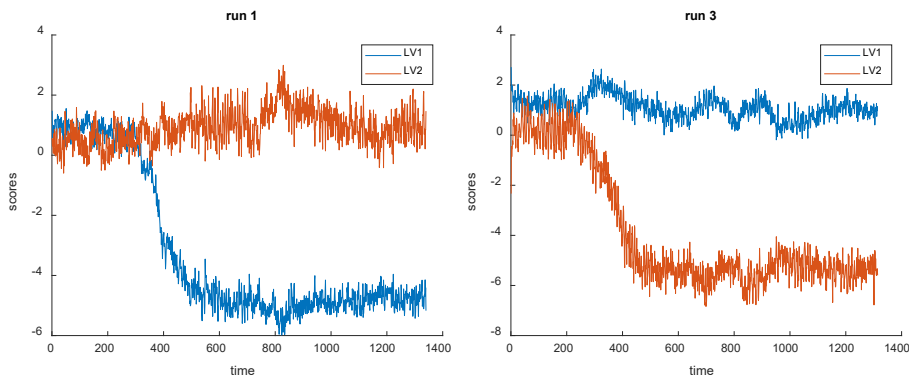


Figure 3. Scores over time for MPCA model of LV1 and LV2 in two cases: (left) when LV1 mostly summarizes the variation caused by the step change and (right) when it is LV2 that captures this variation

In both cases, the same phenomenon is observed: some of the run's variation is captured in a perpendicular direction from the other runs. However, MPCA makes a better job in

capturing these variations aligned with each LV. For this reason, MPCA scores were chosen as the input signal for the RTD model fit. Figure 3 shows the concentration profiles of two of the runs, which exemplify the two behaviors identified in Figure 2. For runs 1, 2, 6, 7, 8, 9 and 10 it is LV1's scores that best capture the dynamic step change, for the other runs it is LV2 that best captures this variation.

3.2. RTD model fit

The scores of the different runs were used to fit the RTD model and it was compared with the model obtained using the traditional methodology using the experimental concentration data. The data was fit to the RTD model in terms of concentration  $C(t)$  by combining Equation (1) to Equation (2) using the fmincon MATLAB function. For every run and approach a set of three RTD parameters were obtained:  $\theta$ ,  $\tau_1$  and  $\tau_2$ .

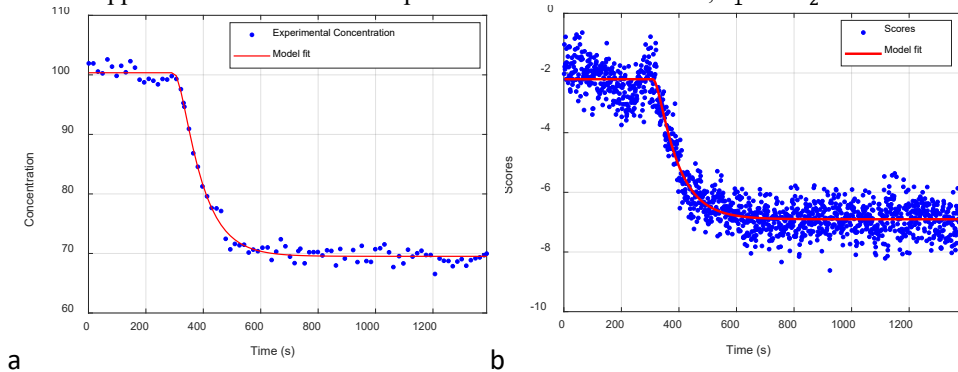


Figure 4. Example of input data and RTD model fit for experimental concentration data (a) and for scores data (b)

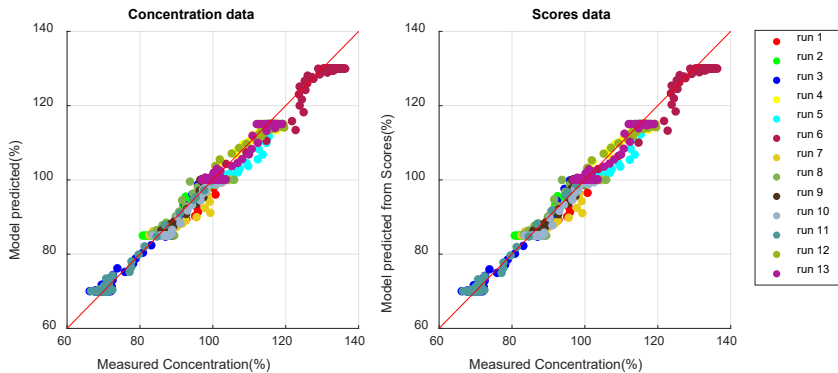


Figure 5. Parity plot of measured vs predicted tablet concentration using the RTD model for the traditional approach using measurements from tablet samples (left) and the new approach using the scores from the feed frame spectral data (right)

Using the set of RTD parameters of the run at target conditions (in this study, run 11) the predicted concentration was calculated with the RTD model. To assess the performance of the model prediction in this study, the root mean square error of prediction (RMSEP) was calculated for every run and for the two approaches using Equation (3), where  $C_{RTD\ model}$  represents the tablet concentration calculated with the RTD model,  $C_{measured}$  is the measured concentration, and  $N$  is the number of samples for every run (Hurley et al.). In the case of the experimental concentration data, the average RMSEP was 1.868

and for the scores data approach, the average RMSEP of all the runs was 1.873. A parity plot is shown in Figure 5 for both approaches, where the  $R^2$  for the experimental concentration data and scores data is 0.9877 and 0.9875, respectively. This indicates that the new approach using the scores of the MPCA performs at the same level as using the measured tablet concentration data, and both approaches do a good job in predicting the tablet concentration.

$$RMSEP = \sqrt{\frac{C_{RTD\ model} - C_{measured}}{N}} \quad (3)$$

#### 4. Conclusion

The goal of this study was related to exploiting the spectral information from the PAT data. The raw spectral data from the NIR probe located in the feed frame was used to extract the dynamic response during the RTD studies in the tableting process. An unsupervised and supervised data-driven method was applied to the spectral data to summarize the dynamic information. MPCA and PLS-DA methods were able to successfully capture the dynamic response of the step change in a single latent variable for every run. This latent variable was used as the input signal for the RTD model fit, and its performance in determining the RTD parameters was compared to the current method using tablet concentration measurements. This new approach was able to approximate the RTD parameters for every run. Moreover, the model prediction given by the scores data performs equally well as when using the tablet concentration data. These results indicate the new approach may provide an alternative to approximate the quality of the tablets using the NIR measurements from the feed frame. This could contribute to making the RTD determination procedure more efficient and lowering the need to make offline measurements.

#### Acknowledgement

The authors want to acknowledge the financial support from the Flanders Innovation & Entrepreneurship (VLAIO) and Janssen Pharmaceutica NV in Belgium (Grant No. HBC.2022.0177 and HBC.2018.0355).

#### References

- P.Bhalode et al., 2021, Using residence time distribution in pharmaceutical solid dose manufacturing – A critical review, *International Journal of Pharmaceutics*
- E.Hetrick et al. 2017, Development of Near Infrared Spectroscopy-based Process Monitoring Methodology for Pharmaceutical Continuous Manufacturing Using an Offline Calibration Approach, *Analytical Chemistry*
- T.Pedersen et al., 2021, Determination of Residence Time Distribution in a Continuous Powder Mixing Process With Supervised and Unsupervised Modeling of In-line Near Infrared (NIR) Spectroscopic Data, *Journal of Pharmaceutical Sciences*
- N. Velez 2022, Challenges, opportunities and recent advances in near infrared spectroscopy applications for monitoring blend uniformity in the continuous manufacturing of solid oral dosage forms, *International Journal of Pharmaceutics*
- S. Hurley et al., 2022, Development and Use of a Residence Time Distribution (RTD) Model Control Strategy for a Continuous Manufacturing Drug Product Pharmaceutical Process, *Pharmaceutics*

## Optimal power distribution in a P2A plant

Joachim Weel Rosbo<sup>a</sup>, Tobias K. S. Ritschel<sup>b</sup>, Steen Hørsholt<sup>c</sup>, Anker D. Jensen<sup>a</sup>, John Bagterp Jørgensen<sup>b</sup>, Jakob K. Huusom<sup>a</sup>

<sup>a</sup>*Dept. of Chemical and Biochemical Engineering, Technical University of Denmark,*

<sup>b</sup>*Dept. of Applied Mathematics and Computer Science, Technical University of Denmark,* <sup>c</sup>*2-control ApS*

### Abstract

In Power-to-Ammonia (P2A) plants the ammonia reactor is required to operate over a wide operating window between 20% to 125% of the nominal load. We formulate a rigorous model for an ammonia synthesis loop with a three-bed adiabatic quench-cooled ammonia reactor. Given a steady-state solution to the ammonia synthesis loop, we evaluate the total power input for the P2A plant composed of the electrical utility for H<sub>2</sub> and N<sub>2</sub> production, and compression work. The total power input and the electrical efficiency for storing electrical energy in NH<sub>3</sub> are evaluated over a wide range of reactor H<sub>2</sub> flow, N<sub>2</sub>/H<sub>2</sub>-ratio and recycle-ratio of the reactants. In this way, we can identify the optimal reactor flow for a given power input to the P2A plant. The optimum is revealed to be at a significantly greater N<sub>2</sub>/H<sub>2</sub>- and recycle-ratio than used in conventional ammonia plants. This reflects the very expensive production of H<sub>2</sub> via electrolysis in P2A compared to conventional production via steam-methane reforming.

**Keywords:** Power-to-Ammonia, Optimization, Flexible operation.

### 1. Introduction

Ammonia is perceived as one of the most promising mediums for chemically based storage of renewable energy sources, e.g. wind and solar power. In P2A, electrical energy from renewable sources is used to generate the reactants, hydrogen via electrolysis and nitrogen via air separation. Due to the intermittent nature of renewable energy sources, P2A plants are required to operate over an operating window from 20% to 125% of the nominal load (Armijo & Philibert, 2020). This defines entirely new operation requirements for ammonia reactors, which are conventionally operated with a stable and reliable supply of reactants from steam-methane reforming. Due to the highly energy-intensive process of hydrogen production, ammonia plants are traditionally located in areas where natural gas is cheap and readily available. But in P2A the hydrogen is produced from electrical energy, which makes the hydrogen more costly to produce. In this study, we investigate the optimal operation of an ammonia synthesis loop in P2A with varying power inputs.

### 2. System description and specifications

Figure 1 shows a schematic illustration of the P2A plant considered in this work. The ammonia is produced by the Haber-Bosch process in an adiabatic quench cooled reactor (AQCR), which is a commonly used reactor in the industry (Inamuddin et al., 2020). Due to equilibrium limitations, the single-pass conversion of the reactor is around 25 %. Therefore, a relatively large recycling of reactants is required. Rosbo et al. (2022) describe mathematical models of the units in the synthesis loop (reactor, heat exchanger, purge, and separator). We use the case study defined by Rosbo et al. (2022) based on a

100MW plant with operating pressure, temperatures, and reactor flowrates scaled from a conventional ammonia reactor.

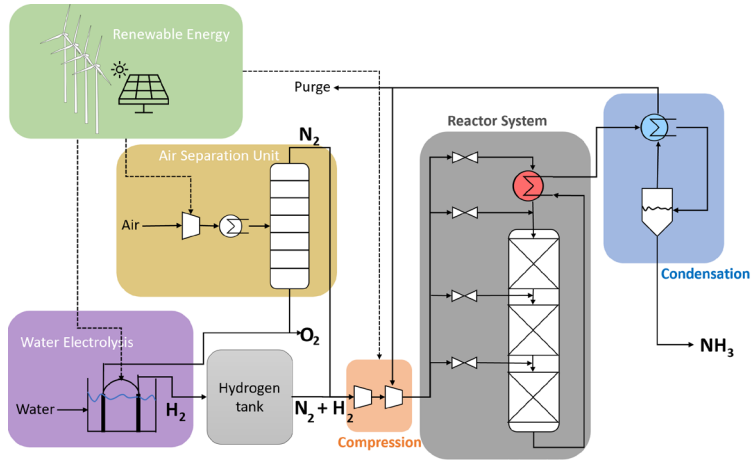


Figure 1: Schematic illustration of the main units in a P2A plant. Electricity is supplied from renewable sources and used in the electrolyser, air separation unit (ASU), and compressors.

### 2.1.1. Pressure drop

The reactor bed is assumed at constant pressure in the mathematical model as the pressure drop is not significant. However, the work of recycle compression relates to the pressure drop over the reactor. Thus, we assess the pressure drop over a bed via Ergun's equation,

$$\frac{dP}{dl} = \frac{180\mu^g (1 - \epsilon)^2}{d_p^2 \epsilon^3} v_s + \frac{1.75\rho^g (1 - \epsilon)}{d_p \epsilon^3} v_s^2$$

in which  $P$  is the pressure,  $l$  is bed length,  $\mu^g$  is the gas viscosity,  $\rho^g$  is the gas density,  $d_p$  is the catalyst particles hydraulic diameter and  $v_s$  is the superficial gas velocity.

## 3. Energy requirements for synthesis loop

### 3.1. Electrolyser

The production of hydrogen from water consumes most of the power input of the P2A plant. Hydrogen is produced in the electrolyser at a cost of 48 kWh/kg corresponding to an electrolyser efficiency of 67% compared with the low heating value of hydrogen (Armijo & Philibert, 2020).

### 3.2. Air separation unit

The nitrogen feed stream is produced in an air separation unit by cryogenic distillation. The power requirement of the air separation unit (ASU) is mainly from the air compression train. The electrical utility for air separation is set to 0.115 kWh/kg nitrogen (Morgan et al., 2014).

### 3.3. Compressors

The synthesis feed gas requires extensive compression from a few bars up to the operating pressure of the ammonia reactor at 200 bars. The work of an isentropic compression from pressure  $P_{in}$  to  $P_{out}$  is

$$W_{comp.}^{isen.} = H(P_{out}, T_{out}^{isen.}, N) - H(P_{in}, T_{in}, N) \quad (1)$$

The outlet temperature for the isentropic compression,  $T_{out}^{isen.}$ , can be found from solving

$$S(P_{out}, T_{out}^{isen.}, N) - S(P_{in}, T_{in}, N) = 0 \quad (2)$$

which is done using Newton's method. We use the thermodynamic tool ThermoLib for evaluating the state functions  $H$  and  $S$  (Ritschel et al., 2016). The actual compressor work,  $W_{comp.}$ , can be found from

$$W_{comp.} = \frac{W_s}{\eta_{isen}} \quad (3)$$

where  $\eta_{isen}$  is the isentropic efficiency ( $\eta_{isen} = 0.75$ ). The compression of the feed gas is carried out by multistage compression with a maximum compression ratio of 4.

### 3.4. Total power input and energy efficiency of the P2A plant

The total power is given as the sum of the individual contributions from  $H_2$  production,  $W_{H_2}$ ,  $N_2$  production,  $W_{N_2}$ , and compressors

$$W_{tot} = W_{H_2} + W_{N_2} + W_{comp.} \quad (4)$$

We define the energy efficiency as

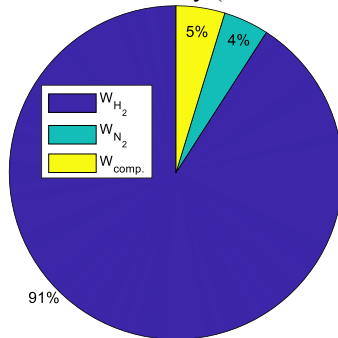
$$\eta_E = \frac{LHV_{NH_3} \dot{m}_{NH_3}}{W_{tot}} \quad (5)$$

Where  $LHV_{NH_3}$  is the lower heating value of  $NH_3$ . The energy efficiency,  $\eta_E$ , describes how well the electrical energy is stored as chemical energy in ammonia.

## 4. Results

### 4.1. Nominal case with conventional reactor feed stoichiometry

For the nominal case, based on a conventional reactor configuration, the total power requirement for the P2A plant is 104.2 MW and the electricity to ammonia efficiency is 50.35%. Figure 2 shows the division of the electrical utility between  $H_2$  and  $N_2$  production, and compressor work. The electrolyser is consuming the largest fraction of the total power input (91%), while the ASU and compressor consumes relatively equal amounts of electricity (4% and 5%).



Case	$W_{tot}$ [MW]	$\eta_E$
Nominal	104.2	50.35%
Optimal	118.1	50.52%
Opt. norm	104.2	51.29%

Left: Figure 2: Fractional energy requirements in nominal P2A plant.

Right: Table 1: Total work and energy efficiency for the nominal and optimised reactor configuration.

The reactor valve settings are optimised using the optimisation algorithm presented by Rosbo et al. (2022). Table 1 shows the total power input and energy efficiency with optimised reactor quench flows. The conversion is about 10% higher for the optimised reactor and, consequently, more reactants are consumed. Therefore, the optimised configuration uses more power at 118.1 MW. Compared to the nominal case, the energy efficiency is increased to 50.52%. For comparison at an equal power input the optimised solution is scaled down to match the power utility of the nominal case. In this way, we obtain an energy efficiency of 51.29%.

### 4.2. Operating at larger $N_2/H_2$ -ratio

By optimising the reactor system, we increase the energy efficiency by almost 1%. This was achieved by keeping  $N_2/H_2$  at the stoichiometric ratio in the reactor feed, which reflects operation in a conventional ammonia plant. However, from the pie chart in Fig-

ure 2 it is evident that in P2A the hydrogen is very costly to produce. Thus, in P2A it may be more energy efficient to operate at a  $N_2/H_2$ -ratio larger than the stoichiometric ratio. By operating with  $N_2$  in excess we facilitate greater conversion of  $H_2$ . We define the stoichiometry factor,  $S_{N_2/H_2}$ , based on the  $N_2/H_2$ -ratio in the reactor feed,  $N_{Rf}$ , as

$$S_{N_2/H_2} = \frac{3N_{Rf,N_2}}{N_{Rf,H_2}}$$

Figure 3a displays the  $NH_3$  production rate as a function of the hydrogen reactor feed rate relative to the nominal case and stoichiometry factor,  $S_{H_2/N_2}$ . The total power input,  $W_{tot}$ , is overlaid by black contours.

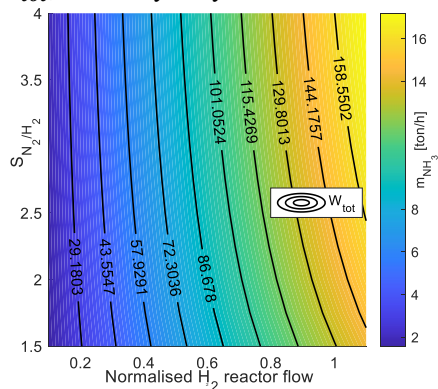


Figure 2a:  $NH_3$  production rate overlaid with contours of total power input.

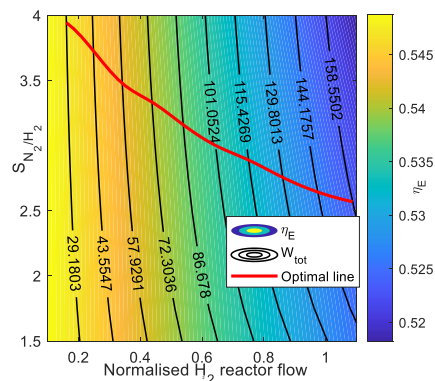


Figure 3b: Energy efficiency,  $\eta_E$  overlaid with contours of total power input.

Figure 3b shows the energy efficiency in the plane of relative  $H_2$  reactor feed flow and  $S_{N_2/H_2}$ . Observe that the optimal stoichiometry factor is as high as 2.7 to 3.8 over the power range. Naturally, the energy efficiency is increasing for lower power input as the reactor is converting the reactants more efficiently at lower throughput. The red curve identifies the maximum energy efficiency for a given power. The optimal curve is moving towards higher stoichiometry factor for lower power input. At lower reactor throughput the reactor is slightly oversized, and it is advantageous to operate with relatively more nitrogen. Operating with excessive flow of  $N_2$  yields a significantly improved energy efficiency at approximately 53.5% for a power input around 104 MW. This clearly illustrates the difference in operation of an ammonia reactor for P2A compared with conventional operation. Figure 4a and 4b show accumulative graphs of the electrical power consumption and reactor flowrate, respectively. Although the reactor is operated at close to equal flowrate of  $N_2$  and  $H_2$ , the power input for the ASU and compressors is still only a small fraction of the total power consumption.

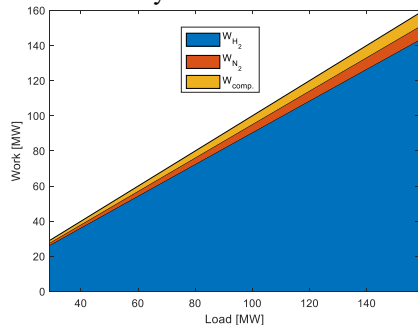


Figure 4a Accumulated energy.

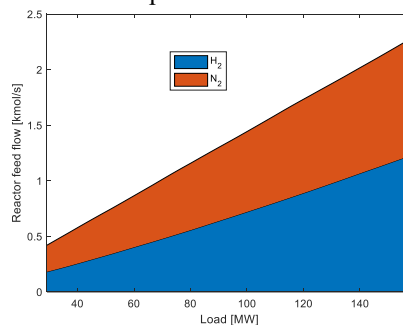


Figure 4b Accumulated flow rate.

### 4.3. Operating at a larger recycle-ratio

In the above analysis we kept the flow of argon in the reactor constant. From Figure 4a it is observed that the compressor work is relatively small compared to the amount of power used for hydrogen production. By reducing the purge fraction, we can facilitate more overall conversion of hydrogen. This is at the cost of a higher recycle stream and more compression work. In practise we vary the recycle by changing the reactor flowrate of argon. Figure 5 shows surfaces of constant power input as a function of normalised H<sub>2</sub> flowrate,  $S_{N_2/H_2}$  and Ar/N<sub>2</sub>-ratio. The red curve illustrates the optimal reactor inlet flows for a given power consumption. The argon to nitration ratio in the reactor feed is ranging from 0.6 to 1.8, which is significantly higher than a conventional ratio of 0.13. Figure 5 shows, we can obtain an energy efficiency of around 56% by operating with a large recycle-ratio even for power inputs above 100 MW.

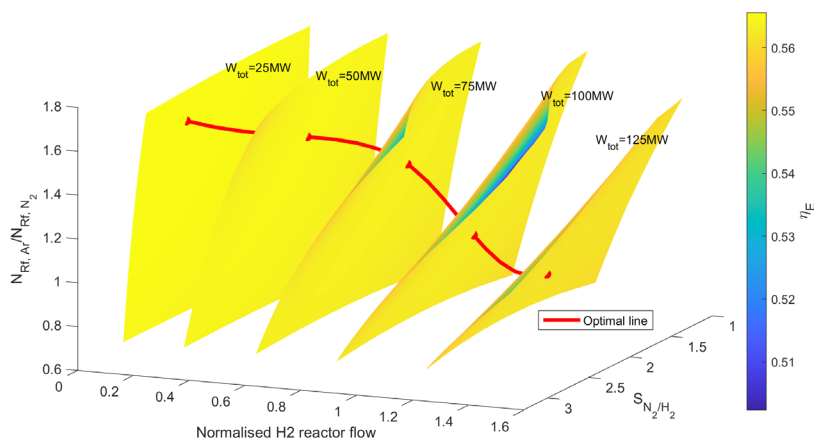


Figure 5 Surfaces of constant power input as a function of normalised H<sub>2</sub> flowrate,  $S_{N_2/H_2}$  and Ar/N<sub>2</sub>-ratio. The surfaces color depicts the electrical efficiency.

Figure 6a displays the electrical utility of the main P2A units for achieving optimal energy efficiency. Note the compression work is increased and the ASU power input decreased compared with Figure 4a. The larger recycle-ratio causes excess N<sub>2</sub> to quickly build up, which lowers the requirement of fresh feed N<sub>2</sub> production. Figure 6b shows the reactor feed flow of H<sub>2</sub>, N<sub>2</sub> and Ar achieving maximum energy efficiency. Clearly, a large amount of inert Argon is built up in the synthesis loop.

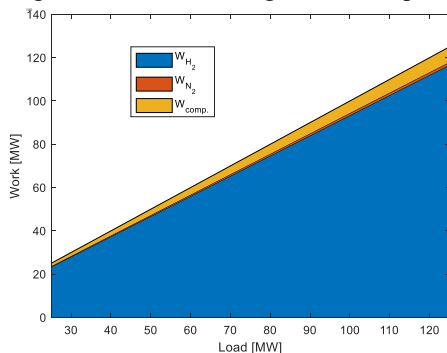


Figure 6a Accumulated energy.

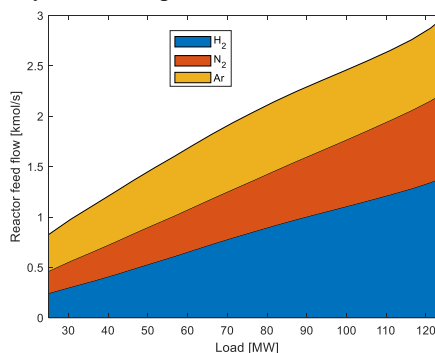


Figure 6b Accumulated flowrate.



## 5. Discussion

This study suggests that in P2A the synthesis reactor should be operated at significantly higher  $N_2/H_2$ -ratio than stoichiometric to obtain maximum energy efficiency. Additionally, we found that larger recycle of reactants compared to conventional operation increased the energy efficiency. We assume that the effectiveness factor of the catalyst particles is independent of the gas concentrations. Most literature on modelling of ammonia reactors use either the Temkin-Pyzhev rate expression or Dyson's equation (Dyson & Simon, 1968). These rate expressions from 1940 and 1968 are naturally developed at conventional reactor conditions with a stoichiometric  $N_2/H_2$ -ratio. Therefore, the validity of these rate expressions is unclear when operating at large excess of  $N_2$  and inert argon. Nielsen et al. (1964) investigated the reaction rate of ammonia synthesis on iron catalyst under various temperatures, pressures, flow rates and even five runs with changing  $N_2/H_2$ -ratio. These results suggested that the reaction rate may be negatively affected at lower flowrates, but the Temkin-Pyzhev equation yielded relatively accurate results at more industrially relevant flow velocities. Traditionally, argon is not mentioned as a poisonous species for ammonia catalysts and no literature seems to discuss the effect of argon on catalyst activity. Clearly, more experimental research needs to be conducted on how operating at large excess of  $N_2$  and inert argon affects the ammonia catalyst. But this study indicates that better utilisation of the electrical energy in P2A can be achieved by operating the ammonia reactor at non-traditional conditions.

## 6. Conclusion

This paper has investigated the optimal operation of an ammonia synthesis loop for P2A using steady state optimisation. We identified the optimal distribution of electrical energy between  $H_2$  and  $N_2$  production, and compression work for power inputs ranging from around 20% to 125% of the nominal load. We found the optimum to be at significantly larger  $N_2/H_2$ - and recycle-ratios compared with conventional ammonia plants. Operating with large excess of  $N_2$  and recycle-ratio ensures a high conversion of  $H_2$ , which is extremely costly to produce in P2A. However, the effect of operating at large excess of  $N_2$  and inert argon on the ammonia catalyst needs to be further investigated experimentally.

## References

- Armijo, J., & Philibert, C. (2020). Flexible production of green hydrogen and ammonia from variable solar and wind energy: Case study of Chile and Argentina. *International Journal of Hydrogen Energy*, 45(3), 1541–1558.
- Dyson, D. C., & Simon, J. M. (1968). A kinetic expression with diffusion correction for ammonia synthesis on industrial catalyst. *Industrial and Engineering Chemistry Fundamentals*, 7(4), 4–10.
- Inamuddin, R. B., & Asiri, A. (2020). Sustainable Ammonia Production. In *Green Energy and Technology*. Springer Nature Switzerland.
- Morgan, E., Manwell, J., & McGowan, J. (2014). Wind-powered ammonia fuel production for remote islands: A case study. *Renewable Energy*, 72, 51–61.
- Nielsen, A., Kjaer, J., & Hansen, B. (1964). Rate equation and mechanism of ammonia synthesis at industrial conditions. *Journal of Catalysis*, 3(1), 68–79.
- Ritschel, T. K. S., Gaspar, J., Capolei, A., & Jørgensen, J. B. (2016). *An open-source thermodynamic software library. DTU Compute-Technical Report-2016 No. 12*.
- Rosbo, J. W., Ritschel, T. K. S., Hørsholt, S., Huusom, J. K., & Jørgensen, J. B. (2022). Flexible operation, optimisation and stabilising control of a quench cooled ammonia reactor for Power-to-Ammonia. *Submitted to Computers & Chemical Engineering*.

# Blend-scheduling optimisation for continuous and batch mixtures: modelling and solving algorithms

Mahmoud A. Ahmednooh,<sup>a,b</sup> Brenno C. Menezes,<sup>b</sup> Mohammed Yaqot,<sup>b</sup> Jeffrey D. Kelly<sup>c</sup>

<sup>a</sup>*Division of Engineering Management and Decision Sciences, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar Foundation, Qatar*

<sup>b</sup>*Division of Production Planning and Scheduling, Um Said Refinery, Qatar Energy, Doha, Qatar*

<sup>c</sup>*Industrial Algorithms Ltd., 15 St. Andrews Road, Toronto M1P 4C3, Canada  
bmenezes@hbku.edu.qa*

## Abstract

Simultaneous blending and scheduling optimisation represents a mixed-integer nonlinear programming (MINLP) problem, whereby the binary variable relaxation that forms a nonlinear programming (NLP) in the first stage of a full-space algorithm (without dropping any decision variable) may lead to convergence issues. Furthermore, there is no guarantee to reach a global optimal solution since it solves an NLP problem to then address the mixed-integer linear programming (MILP) problem in outer-approximation algorithms. On the other hand, by neglecting quality variables and constraint of the blends or mixtures in the MILP problem, in a decomposed or non-full space problem (when dropping nonlinear constraints), infeasibilities or local optimal solutions may be found in the second stage NLP programs. In this case, there will be a change that from the MILP resulted assignments of the components to be blended, quality to specify the final blended material in the NLP problem may not be suffice. Or even if so, local optimal solution may occur. To skip these issues, an MILP-NLP decomposition is tailored to solve, in the first stage, the MILP logistics problem of the blend-scheduling optimisation whereby the blending relationship for each property quality is approached as amounts of quality balances. In such constraints, the blended quality variable is replaced by the bounds of the specifications plus the slack variable or less the surplus variable to close the linear quality balance. However, this MILP-NLP decomposition with a blending approximation in the MILP still cannot avoid convergence issues or local optimal inherent to the NLP stage and is dependent on the volume- or mass-based material flow and blended property governing rules. Then, we apply an optimisation-simulation algorithm to converge the MILP solution to a global optimal by considering a substitution of the current blending error successively in the next MILP solution. Nevertheless, this algorithm can be applied only in continuous mixtures, but we introduce the novel strategy to solve batch mixtures by using component quantities found in a continuous mixture topology as inputs.

**Keywords:** blending operations, scheduling optimisation, MILP, NLP, decomposition.

## 1. Introduction

Blending and scheduling processes are considered crucial steps in crude-oil and metallurgical refineries, food processing sites, pulp and paper productions, to name a few. The main goal in these processing industries is to transform different raw materials into

useful products such as gasoline, diesel, and kerosene (from crude-oils) and copper alloys (from minerals). Stricter constraints from environmental regulations and lower profit margins from commodity falling prices have been considered for the blending of these streams to achieve better efficiency in the operations and to adhere to specifications (Lotero et al., 2016). The blending process is a significant stage that helps the processing sites to convert low-value raw materials to higher added-value products, with consideration to eliminate giveaway or excess of qualities and off-specification products.

Most (if not all) of the blend-scheduling solutions in the market are either simulation-based or simple linear programming (LP) approaches. These tools are limited in terms of capability since the simulation-based ones are entirely based on trial-and-error of the binary variables, which is time-consuming and difficult to change or readjust. Also, linear programming software is limited to a fixed or given operation, in general, addressing a single-period problem. A proposed solution that can provide multi-period solution with automated decision-making in a blend-scheduling perspective is a breakthrough in the quantity-logic-quality decisions in the aforementioned industries. This becomes explore in optimisation problems since the blending of streams presents a nonlinear (NLP) behavior, which brings difficulties on convergence and cannot avoid local optimal solutions. For an optimisation of a blendshop or blend-scheduling problem, the blending and assignment of equipment (and their modes) and of material streams represents a mixed-integer nonlinear programming (MINLP), whereby, to avoid MINLP convergence issues when scaling-up blend-scheduling problems, we proposed a successive substitution approach based on optimisation-simulation methods to iteratively reduce the error inherent by using factor-flows or amounts of quality balances from Kelly et al. (2018). In equality balances of input-output amounts of qualities, the outputted amounts of qualities are calculated by the quality specification multiplied by the total flow of material plus the slack variable or less the surplus variable to close the linear quality balance, whereby this additional variable, known as factor-flows, close the balance of each property.

## 2. Problem statement

In the continuous mixture of the blend-scheduling represented in Figure 1, it is produced diesel by blending light diesel gasoil in R1 and light cycle oil in R2. Both intermediate components have similar specific gravity (SG) and sulfur concentration (SC) to produce diesel, whereby specifications on the final diesel qualities should be maintained within bounds. Their qualities are: R1 with 0.80 of SG and 0.90 of SC (at \$ 85 cost); R2 with 0.95 of SG and 1.20 of SC (at \$ 70 cost); and D1 with 0.875 of SG and 1.05 of SC to specify (at \$ 100 as selling price per barrel). There are four cases to be addressed in this work to avoid an MINLP inherent to blend-scheduling problems. Cases 1 to 3 are applied to continuous blending (Figure 1) and case 4 to batch mixtures (Figure 1, small topology).

1) *MILP-NLP decomposition*: nonlinear quality balances are neglected in the MILP problem, then the assignments of the S1 to S4 tanks to the blender are fixed to 0 or 1 in the following NLP solution.

2) *MILP-NLP decomposition*: using factor-flow balances where the nonlinear quality constraints are approximated in the MILP problem as amounts of qualities (Kelly et al. 2018), whereby the blended quality variable is replaced by the bounds of the specifications plus the slack variable or less the surplus variable to close the linear quality

balance. Then, the assignments of the S1 to S4 tanks to the blender are fixed to 0 or 1 in the following NLP solution.

3) *SS-MILP iteration*: successive substitution of an optimisation-simulation method considering the error between the factor-flow approximated property calculated by the quantities of the components found in the MILP and the volume- or mass-based blended property governing rules (from the simulation step) for the same amounts of the components. The quality error is used in the next MILP optimisation until convergence.

4) *SS-MILP iteration and MILP-NLP batch*: the amounts or quantity flows from the tanks S1 to S4 to the blenders converged in the SS-MILP algorithm are fixed in the blend-scheduling case without the blender (as seen in Figure 1, small topology)

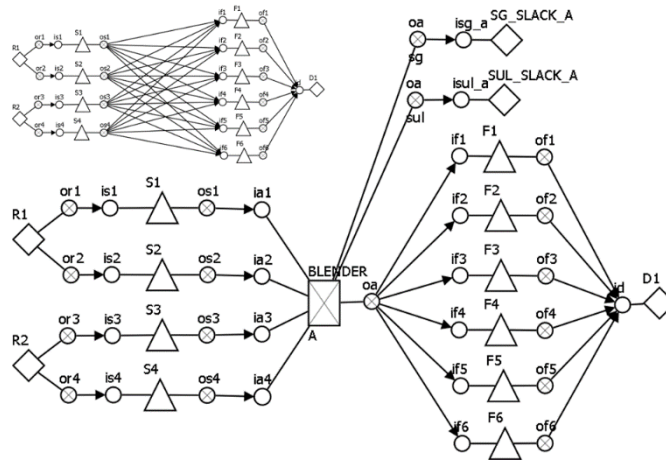


Figure 1: Blend-scheduling using a continuous and batch (small) mixture topology.

### 3. Mathematical modelling

The objective function in Eq.(1) maximises the diesel production, where  $x_{j,i,t}$  represent the selling product and consumed components in the diesel production at time  $t$ . Following the unit-operation-port-state superstructure (UOPSS) formulation from Kelly (2005), in the MILP problem, all flows and holdups are governed by semi-continuous constraints (relating binary and continuous variables) of the shapes to themselves, such as  $\bar{x}_{j,i,t}^L y_{j,i,t} \leq x_{j,i,t} \leq \bar{x}_{j,i,t}^U y_{j,i,t} \forall (j, i) \in JI, t$  considering outlet port set  $j$  to inlet set  $i$ . The sets  $I$  and  $J$  represent in- and out-ports, respectively, while the set  $JI$  defines connecting flows between out- and in-ports. The semi-continuous constraints for throughputs  $x_{m,t}$  of the unit-operations  $m$  for the blender are controlled by  $\bar{x}_{m,t}^L y_{m,t} \leq x_{m,t} \leq \bar{x}_{m,t}^U y_{m,t}$  and the holdup or inventory level  $xh_{m,t}$  by  $\bar{x}h_{m,t}^L y_{m,t} \leq xh_{m,t} \leq \bar{x}h_{m,t}^U y_{m,t}$ . In the indices in the summations from Eq.(1) to (7), the subsets of the  $I, J$ , and  $JI$  follow the flowsheet in Figure 1, considering  $i \in I_{do}$  as inlets of downstream unit-operations and  $j \in J_{up}$  as outlets of upstream unit-operations. For  $x_{j,i,t}, x_{m,t}, xh_{m,t} \geq 0; y_{j,i,t}, y_{m,t} = \{0,1\}; zsu_{m,t} = (0,1)$ :

$$Max Z = \sum_t \left( \sum_{product} x_{j,i=D1,t} - \sum_{components} x_{j=R1 \text{ or } R2,i,t} \right) \quad (1)$$

$$\frac{1}{\bar{x}_{m,t}^U} \sum_{j \in J_{up}} x_{j,i,t} \leq y_{m,t} \leq \frac{1}{\bar{x}_{m,t}^L} \sum_{j \in J_{up}} x_{j,i,t} \quad \forall (i, m) \in (M_{Blender}, M_{Product}), t \quad (2)$$

$$\frac{1}{\bar{x}_{m,t}^U} \sum_{i \in I_{do}} x_{j,i,t} \leq y_{m,t} \leq \frac{1}{\bar{x}_{m,t}^L} \sum_{i \in I_{do}} x_{j,i,t} \quad \forall (m, j) \in (M_{Blender}, M_{Component}), t \quad (3)$$

$$\frac{1}{\bar{x}_{m,t}^U} \sum_{j \in J_{up}} x_{j,i,t} \leq x_{m,t} \leq \frac{1}{\bar{x}_{m,t}^L} \sum_{j \in J_{up}} x_{j,i,t} \quad \forall (i, m) \in M_{Blender}, t \quad (4)$$

$$\frac{1}{\bar{x}_{m,t}^U} \sum_{i \in I_{do}} x_{j,i,t} \leq x_{m,t} \leq \frac{1}{\bar{x}_{m,t}^L} \sum_{i \in I_{do}} x_{j,i,t} \quad \forall (m, j) \in M_{Blender}, t \quad (5)$$

$$xh_{m,t} = xh_{m,t-1} + \sum_{j_{up} \in J} x_{j_{up},i,t} - \sum_{i_{do} \in I} x_{j,i_{do},t} \quad \forall (i, m, j) \in M_{TK}, t \quad (6)$$

$$\sum_i \sum_{j_{up}} x_{j_{up},i,t} = \sum_{i_{do}} x_{j,i_{do},t} \quad \forall ((j_{up}, i), (j, i_{do}), t), m \in M_{Blender} \quad (7)$$

$$mu_{j,t}^L y_{m,t} \leq \sum_{i_{do}} y_{j,i_{do},t} \leq mu_{j,t}^U y_{m,t} \quad \forall (m, j) \in M_{Blender}, t \quad (8)$$

$$mu_{i,t}^L y_{m,t} \leq \sum_{j_{up}} y_{j_{up},i,t} \leq mu_{i,t}^U y_{m,t} \quad \forall (i, m) \in M_{Product}, t \quad (9)$$

$$y_{m_{up},t} + y_{m,t} \geq 2y_{j_{up},i,t} \quad \forall (m_{up}, j_{up}, i, m), t \quad (10)$$

$$y_{m,t} - y_{m,t-1} - zsu_{m,t} + zsd_{m,t} = 0 \quad \forall m \in M_{Blender}, t \quad (11)$$

$$y_{m,t} + y_{m,t-1} - zsu_{m,t} - zsd_{m,t} - 2zsw_{m,t} = 0 \quad \forall m \in M_{Blender}, t \quad (12)$$

$$zsu_{m,t} + zsd_{m,t} + zsw_{m,t} \leq 1 \quad \forall m \in M_{Blender}, t \quad (13)$$

Equations (2) and (3) represent the semi-continuous constraints relating neighboring flows of the  $JI$  connections to the throughputs of unit-operation setups they are connected to. Different from the semi-continuous constraints to itself (of the shapes to themselves), Eq.(2) and Eq.(3) allow the flows throughout the connections only if the nodes (the unit-operations) are true. On the other hand, Eq.(4) and Eq.(5) balance the in-out material flows of the connections to the throughputs of the unit-operations, whereby there is a process transformation such as in the blender. The unit-operations' inventory or holdup quantity balance of tanks are calculated in Eq.(6). Equation (7) is the material flow balance in the blender. Multi-use constraints control the use of the blender's outlet in Eq.(8) and the inlet of the product demand point  $DI$  in Eq.(9).

Equation (10) represent the constraint for the structural transitions that allow the setup  $y_{m,t}$  of connected out-port-states  $j$  and in-port-states  $i$  unit-operations. When the setup of unit-operations  $m$  and  $m'$  is equal to the unitary, by implication, the setup variable of the arrow stream  $y_{j,i,t}$  between the neighbor unit-operations must be true, and vice-and-versa. These logic valid cuts reduce the tree search in branch-and-bound methods. The temporal transition in Equations (11) and (12) control the operations for semi-continuous blenders from Kelly and Zyngier (2007). The binary variable  $y_{m,t}$  manages the start-up ( $zsu_{m,t}$ ) switch-over ( $zsw_{m,t}$ ) and shut-down variables ( $zsd_{m,t}$ ), which are relaxed in the interval  $[0,1]$ . Equation (13) guarantees the integrality of the relaxed variables.

For the NLP problem, the binary variables found in the MILP are fixed. The quality balance constraints in the blender are defined in Eq.(14) for the volume-based property  $v_{j,p,t}$  for  $p \in P_v$  (as in SG) and in Eq. (15) for the mass-based property  $w_{j,p,t}$  for  $p \in P_w$  (as in SC). Equations (16) and (17) represent the quality balances for volume- and mass-based properties in tanks. The quality variable for the out-ports of a tank unit-operation ( $m \in M_{TK}$ ) is the quality of the blend within the tank  $v_{j,p,t} = v_{m,p,t}$  and  $w_{j,p,t} = w_{m,p,t}$ .

$$v_{j,p,t} \sum_{j_{up}} x_{j_{up},i,t} = \sum_{j_{up}} v_{j_{up},p,t} x_{j_{up},i,t} \quad \forall j \in M_{Blender}, p \in P_v, t \quad (14)$$

$$w_{j,p,t} \sum_{j_{up}} v_{j_{up}=sg,p,t} x_{j_{up},i,t} = \sum_{j_{up}} w_{j_{up},p,t} v_{j_{up}=sg,p,t} x_{j_{up},i,t} \quad \forall j \in M_{Blender}, p \in P_w, t \quad (15)$$

$$v_{m,p,t} xh_{m,t} = v_{m,p,t-1} xh_{m,t-1} + \sum_{j_{up}} v_{j_{up},p,t} x_{j_{up},i,t} - v_{m,p,t} \sum_{i_{do}} x_{j,i_{do},t} \quad (16)$$

$$\forall (i, m, j) \in M_{TK}, p \in P_v, t$$

$$w_{m,p,t} v_{m,p=sg,t} xh_{m,t} = w_{m,p,t-1} v_{m,p=sg,t-1} xh_{m,t-1} + \sum_{j_{up}} w_{j_{up},p,t} v_{j_{up},p=sg,t} x_{j_{up},i,t} -$$

$$w_{m,p,t} v_{m,p=sg,t} \sum_{i_{do}} x_{j,i_{do},t} \quad \forall (i, m, j) \in M_{TK}, p \in P_w, t \quad (17)$$

#### 4. Results

The optimisation for the proposed problems (to produce 100 K barrels per day) for 10 days of time-horizon and 1 day as time-step uses an Intel Core i7 machine at 2.2 GHz (6 threads) with 32 GB of RAM. In the case 1, for the MILP, there are 552 constraints (122 equality) for 312 continuous variables and 210 binary variables with 400 degrees-of-freedom (variables minus equality constraints). For the NLP, there are 531 constraints (491 equality) for 621 continuous variables with 130 degrees-of-freedom. The results using GUROBI 10.0 and CPLEX 20.0 for the MILP blend-scheduling cases are solved within 1 second. And the NLP are solved by linking these solvers to sequential linear programming quadratic engine (SLPQPE) in the modelling platform Industrial Modeling and Programming Language (IMPL) used in this work (see Kelly and Menezes, 2019).

1) *MILP-NLP decomposition*: the MILP solution yields \$ 440 and it gets infeasible in the NLP. Since the nonlinear quality balances are neglected in the MILP problem, the assignments to produce diesel selects the S3 or S4 tanks to the blender at a reduced cost. However, then the NLP problem is solved it is not possible to meet the specification since better SG and SC qualities from S1 or S2 tanks are needed.

2) *MILP-NLP decomposition*: the MILP solution yields \$ 380 and the NLP \$ 320. By using factor-flow balances with the nonlinear quality constraints approximated in the MILP as amounts of qualities, whereby the blended quality variable is replaced by the bounds of the specifications plus the slack variable or less the surplus, S1 or S2 tanks are assigned. Then, for the binary variables fixed in NLP problem, it is possible to meet the specifications of SG and SC qualities. However, as seen in the solution in Table 1, the flows are not optimal since in several days the blend-shop uses only S1 or S2 tanks.

3) *SS-MILP iteration*: successive substitution of an optimisation-simulation method considering the error between the factor-flow approximated property calculated by the quantities of the components found in the MILP and the volume- or mass-based ad hoc blended property governing rules (from the simulation step) for the same amounts of the components. The quality error is used in the next MILP optimisation until convergence. It reaches the global optimum solution within two SS iterations

4) *SS-MILP iteration and MILP-NLP batch*: the amounts or quantity flows from the tanks S1 to S4 to the blenders, converged in the SS-MILP algorithm, are fixed in the blend-scheduling case without blender. In the MILP-NLP batch case, with the SS-MILP quantities fixed in the outlets of the S1-S4 and inlets of the F1-F6 tanks (within an adjusted time-step that permits one batch flow at a time), the solution yields a feasible and optimal results with 0% gap between the MILP and NLP problems.

Table 1. Results of the cases in the blend-scheduling problems.

day	Case 2-MILP (380 K USD)				Case 2-NLP (320 K USD)				Case 3 (371.6 K USD)			
	S1	S2	S3	S4	S1	S2	S3	S4	S1	S2	S3	S4
1	50.0		50.0		58.7		41.3		54.3		45.7	
2	50.0		50.0		100.0				54.3		45.7	
3		50.0		50.0		100.0				54.3		45.7
4		50.0		50.0		100.0				54.3		45.7
5	50.0		50.0		100.0				54.3		45.7	
6		50.0		50.0		100.0				54.3		45.7
7		50.0	50.0			100.0				54.3	45.7	
8		50.0	50.0			91.3	8.7			54.3	45.7	
9		50.0		50.0		92.4		7.6		54.3		45.7
10	50.0		50.0		57.6			42.4	54.3			45.7

## 5. Conclusion

The importance of approximating an NLP blending as an LP factor-flow approach in the MILP stage is considered in this work, whereby factor-flow variables by equalizing in-out amounts of qualities in the blender allow global optimal solutions since it avoids NLP problems in a second stage of the MILP stage. For batch mixture topology, the drawback of the factor-flow approach by considering a convergent and continuous point of mixture can be circumvented when fixing the material flows of the components found in the SS-MILP (case 3) as inputs in the MILP-NLP batch case, adjusting of needed the time-step.

## References

- I. Lotero, F. Trespacios, I.E. Grossmann, D.J. Papageorgiou, M.-S. Cheon, (2016). An MILP-MINLP decomposition method for the global optimization of a source based model of the multiperiod blending problem, *Computers & Chemical Engineering*, 87, 13–35.
- J.D. Kelly, B.C. Menezes, I.E Grossmann, 2018, Successive LP approximation for nonconvex blending in MILP scheduling optimization using factors for qualities in the process industry, *Industrial & Engineering Chemistry Research*, 57(32), 11076-11093.
- J.D. Kelly, 2005, The Unit-Operation-Stock Superstructure (UOSS) and the Quantity-Logic-Quality Paradigm (QLQP) for production scheduling in the process industries, In *Multidisciplinary International Scheduling Conference Proceedings: New York, United States*, 327-333.
- J.D. Kelly, B.C. Menezes, 2019, Industrial Modeling and Programming Language (IMPL) for off- and on-line optimization and estimation applications. In: Fathi M., Khakifirooz M., Pardalos P. (eds) *Optimization in Large Scale Problems*. Springer Optimization and Its Applications, 152, 75-96.

# A machine learning dynamic modelling scheme for wastewater treatment plants using cooperative particle swarm optimization and neural networks

Teo Protoulis,<sup>a</sup> Ioannis Kalogeropoulos,<sup>b,c</sup> Ioannis Kordatos,<sup>a</sup> Haralambos Sarimveis,<sup>b</sup> Alex Alexandridis<sup>a</sup>

<sup>a</sup>*Department of Electrical and Electronic Engineering, University of West Attica, Ancient Olive Grove Campus, Thivon 250 & P. Ralli, Aigaleo, 12244, Greece*

<sup>b</sup>*School of Chemical Engineering, National Technical University of Athens, Heroon Polytechniou 9, Zografou, 15780, Greece*

<sup>c</sup>*Systemica - G.Vangelatos & Co L.P., Ethnikis Antistaseos 70, 16237, Athens, Greece*

## Abstract

This work introduces a framework for developing data-driven dynamic models for wastewater treatment plants (WWTPs) by using a limited amount of data. More specifically, a system identification problem is defined that estimates a number of critical parameters included in the COST/IWA benchmark simulation model No.1 (BSM1), which is adapted to a real WWTP. Identification of these parameters is performed by formulating an optimization problem, which is solved by utilizing a novel cooperative particle swarm optimization (CPSO) method. The identified model is then used to generate big volumes of dynamic data for training radial basis function networks (RBFNs), which are suitable for being integrated in process control configurations, like model predictive control schemes. The proposed optimization method is validated using statistical testing and is shown to compare favorably against alternative methods, while evaluation metrics indicate that the developed models predict the dynamic behavior of the plant with high accuracy.

**Keywords:** Cooperative particle swarm optimization; identification; modeling; radial basis function neural networks; wastewater treatment plants

## 1. Introduction

Wastewater treatment plants (WWTPs) are considered to be a major component in circular economy due to their ability of receiving large volumes of polluted water as influent and allowing the maximum percentage of it to be reused. Developing automatic control schemes for WWTPs is a challenging task, not only due to the multiple and highly complex biological and biochemical phenomena taking place in the various processes, but also because strict environmental regulations have to be met. The performance of WWTP control systems can be improved, by incorporating mathematical models capable of accurately capturing the system dynamics. On the other hand, developing such models that are also simple enough to be incorporated in control schemes, is a challenging task.

Data-driven technologies, such as artificial neural networks (ANNs) appear as a promising solution due to the ability to capture the nonlinear dynamics of WWTPs with satisfactory accuracy (Sadeghassadi et al., 2018), and the potential to be easily integrated within advanced model predictive control (MPC) methodologies (Papadimitrakakis &



Alexandridis, 2022). On the other hand, training ANNs for WWTP modelling needs big volumes of data (Matheri et al., 2021). Unfortunately, due to the limited number of sensors available in WWTPs, the volumes of experimental data collected from the process are usually not sufficient; at the same time the collected data are of low quality due to several reasons (e.g. discontinuity, missing values, sensors failures, steady-state operation for long intervals), which lead to poor predictive performance of the developed models.

In this work, a novel approach is proposed for overcoming the problem of limited data availability in order to build highly accurate ANN models, suitable to be incorporated in MPC schemes. More specifically, a system identification problem is first defined, where the objective is to identify the critical parameters of the COST/IWA benchmark simulation model No. 1 (BSM1) (Alex et al., 2008), adapted to a real WWTP. The identification task takes the form of a complex nonlinear optimization problem, which is solved efficiently by introducing a novel cooperative particle swarm optimization (CPSO) (Kapnopoulos & Alexandridis, 2022) approach, specifically customized for WWTPs. The identified BSM1 model is used next to generate large volumes of dynamic data suitable for building high accuracy models that are based on radial basis function networks (RBFNs), trained with the fuzzy means (FM) algorithm (Alexandridis et al., 2013).

The rest of this work is structured as follows; in section 2, the formulation of the system identification problem is described and the proposed CPSO solver is presented. The next section presents the proposed method for developing dynamic WWTP models, including a brief description of the RBF architecture and the FM algorithm. The method is evaluated through a case study presented in section 4. Finally, the paper ends with concluding remarks and directions for future work.

## **2. System identification using cooperative particle swarm optimization**

### *2.1. Parameter identification of the Benchmark Simulation Model No. 1 (BSM1)*

BSM1 is a detailed WWTP model as described in (Alex et al., 2008). For the purposes of this work, the original BSM1 model has been adapted to a real WWTP, located in Greece. This procedure involved modifications to both structural and functional aspects of the original BSM1. The layout had to be adopted in terms of the number of process lines and the volume of the reactors and the secondary settlers. In addition, the capacity of aeration and pumping systems was modified. As far as the functional interventions are concerned, the average influent flow rate of the actual plant was used and correspondingly, the internal and external recirculation flow rates were also modified. Fig. 1 depicts a schematic representation of the particular WWTP.

The BSM1 model includes 14 kinetic and 5 stoichiometric parameters involved in the mass-balance equations of each reactor, the default values of which are fixed. Proper adaptation of this generic model to an actual WWTP requires an accurate estimation of these parameters. In this work, an identification scheme for estimating these parameters is proposed, using limited dynamic process data that can be directly collected from specific sensors of the WWTP, such as nitrate and nitrite concentration of the anoxic tanks, dissolved oxygen of the aerobic tanks and the ammonia and nitrite and nitrate concentration of the last aerobic tanks.

The identification process is carried out by formulating a nonlinear optimization problem, where the objective is to minimize the mean squared error (MSE) between the actual

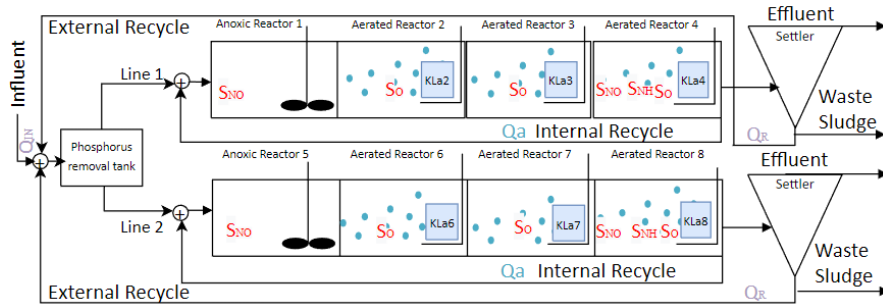


Fig. 1: Schematic Representation of the real WWTP under study

process data and the data generated at discrete time instances, in particular every 15 minutes, by imposing a limited number of steps on the manipulated variables of the identified BSM1 model. The resulting optimization problem exhibits several unwelcome properties, including high dimensionality, multimodality and noise contamination, which render the use of standard optimization methods inefficient.

### 2.2. Cooperative particle swarm optimization

Particle swarm optimization (Alexandridis et al., 2013), is a metaheuristic search method based on a population of potential solutions, called particles. Each particle explores the search space based on the personal and global best position achieved at each iteration of the algorithm. The standard PSO algorithm groups all the design variables into one swarm, optimizing them simultaneously. However, the high number of parameters involved in this particular problem makes it difficult for the solver to approach a satisfactory solution. To remedy this situation, we introduce a cooperative PSO method (CPSO), which takes into account correlations between the design variables and splits them accordingly into several groups, thus forming distinct swarms. Each swarm evolves independently in the search space while concurrently exchanging information with the rest of the swarms. In each iteration, the positions and velocities of the particles of each swarm are updated according to the following equations:

$$P_k v_{i,j}(t + 1) = w \cdot P_k v_{i,j}(t) + c_1 \cdot r_{1,i}(t) \cdot [P_k y_{i,j}(t) - P_k x_{(i,j)}(t)] + \dots + c_2 \cdot r_{2,i}(t) \cdot [P_k \hat{y}_j(t) - P_k x_{i,j}(t)]$$

$$P_k x_{i,j}(t)(t + 1) = P_k x_{i,j}(t) + P_k v_{i,j}(t + 1) \tag{1}$$

where the position of the  $i$ -th particle in  $j$ -th dimension is denoted by  $P_k x_{i,j}(t)$  for the  $k$ -th swarm. The velocity of each particle is denoted by  $P_k v_{i,j}(t)$ ,  $P_k y_{i,j}(t)$  represents the position of the best particle of each swarm and  $P_k \hat{y}_j(t)$  denotes the position of the global best particle. Coefficients  $c_1, c_2$  correspond to the local and global adjustment weights,  $w$  is the inertia weight and  $r_{1,i}, r_{2,i}$  are uniformly distributed random numbers. As each swarm now contains only a part of the design vector, the value of the objective function is calculated by creating a context vector which combines the particles of each swarm with the global best particles of the remaining swarms. A more detailed description of CPSO can be found in (Kapnopoulos & Alexandridis, 2022). This customized CPSO methodology has a number of advantages over standard PSO, including a finer-grained assignment which avoids the two steps forward - one step back problem, and the fact that each swarm evolves independently in the search space, thus, improving the diversity of the solution. These advantages ultimately lead to a highly accurate estimation of the BSM1 parameters.

### 3. Dynamic modelling using radial basis function neural network models

RBF neural networks represent a special type of ANNs used for application to various problems of diverse nature, including MPC controller design (Papadimitrakis & Alexandridis, 2022), because of their simple architecture and their high approximation capabilities. For a typical structure of an RBF neural network with  $N$  input variables,  $L$  hidden nodes and one output variable, the output of the RBF network for an observation  $t$  is calculated as:

$$\hat{y}(t) = \sum_{l=1}^L w_l \varphi_l \left( \sum_{n=1}^N (x_n(t) - \hat{x}_{l,n})^2 \right)^{1/2} \quad (2)$$

where  $\varphi$  represents the radial basis function used,  $x_n$  is the input space vector of the network,  $\hat{x}_l$  is the center of the  $l$ th hidden neuron, and  $w_l^T = [w_1, w_2, \dots, w_L]$  is the vector of synaptic weights that connect the hidden layer with the output layer. The radially symmetric function employed in this work was the thin plate spline (TPS) function:

$$\varphi(x) = x^2 \log(x) \quad (3)$$

In order to tune the structure of nodes of the hidden layer, the centers are selected employing the unsupervised FM algorithm, which has found numerous successful applications in different scientific fields (Papadimitrakis & Alexandridis, 2022). The FM algorithm uses fuzzy partitions of the input space to determine the number and locations of the hidden nodes. A detailed description of the algorithm can be found in (Alexandridis et al., 2013). After the estimation of the hidden node centers, weights are calculated by using linear least squares, implemented in matrix form:

$$w^T = Y^T \Phi (\Phi^T \Phi)^{-1} \quad (4)$$

where  $\Phi = [\varphi_1, \varphi_2, \dots, \varphi_L]$  and  $Y$  are the measured target values.

In order to produce dynamic RBF models with the aforementioned approach, past values of the process manipulated variables, the measured disturbances and the measured state variables - except from the output variable - are used as inputs to the RBF network.

## 4. Results and Discussion

### 4.1. System identification

The proposed identification scheme is employed by initially applying pseudorandom binary sequences on the BSM1 model input variables and disturbances. Data for the 12 measured state variables are collected for a 2-month period with a sampling period of  $T_s = 15 \text{ min}$ , which is a time interval suitable for controller design (Alex et al., 2008); the data are then corrupted with Gaussian noise  $\sim N(0, \sigma)$ , where  $\sigma$  is 2.5% of the mean measured value for each variable. This set of data is introduced to the CPSO identification scheme for estimating the values of the BSM1 kinetic and stoichiometric parameters.

The parameters to be estimated are separated into distinct swarms based on the different processes they affect during the operation of the plant. The first two swarms consist of the 8 and 5 parameters related to the heterotrophic and autotrophic phenomena, respectively. The 3 parameters affecting hydrolysis and the unique parameter influencing ammonification, form the 3<sup>rd</sup> and 4<sup>th</sup> swarm, respectively. The remaining parameters related to the COD in biomass and in products from the biomass, constitute the 5<sup>th</sup> swarm.

For comparison purposes, the employed CPSO solver was tested against standard PSO and a modified PSO methodology (Mezura-Montes et al., 2011). The algorithms are individually tuned using literature suggestions in conjunction with a trial-and-error procedure. Due to the inherent stochastic nature of the algorithms, 15 runs are executed for each algorithm and the superiority of CPSO is validated by running a two-sample *t*-test. Table 1 summarizes the results. As it can be seen, the performance of CPSO is superior in terms of the best or the average solution with a statistical significance higher than 90%, while it also manages to produce more consistent results, as indicated by the standard deviation values.

#### 4.2. Dynamic modeling

The proposed method can produce dynamic models for different process variables; in this work, we present the results for the ammonia mass concentration ( $S_{NH}$ , g N.m<sup>-3</sup>) in the last aerated tank (Reactor 4) in Line 1. The input features were: (A) lags for all the manipulated variables, i.e. oxygen coefficients ( $K_L a$ , d<sup>-1</sup>) in all the aerated reactors and the internal recycling flows ( $Q_a$ , m<sup>3</sup> d<sup>-1</sup>) for both lines, (B) lags for measured plant disturbances, i.e. the flow rate ( $Q_{input}$ , m<sup>3</sup> d<sup>-1</sup>) and ammonia influent ( $S_{NH_{input}}$ , g N.m<sup>-3</sup>) and (C) lags for all the measured state variables, i.e. dissolved oxygen concentration ( $S_O$ , g (-COD).m<sup>-3</sup>) in all aerated tanks, nitrite and nitrate nitrogen mass concentration ( $S_{NO}$ , g N.m<sup>-3</sup>) for all the anoxic tanks and aerated reactors 4, 8, and ammonia mass concentration ( $S_{NH}$ , g N.m<sup>-3</sup>) for reactor 8. The number of past values for each feature was decided empirically by performing step tests and observing the dynamic responses of the measured output variables. The total number of input features was equal to 46. Using the BSM1 model with parameters identified by the proposed CPSO scheme, a large dataset containing 10 years of simulated data with 15 minutes sampling was generated in order to train the RBFN with the procedure described in the previous section. To properly evaluate the RBFN model predictive abilities, a new testing dataset was produced, spanning 5 years of operation, using this time the BSM1 with the original parameters and contamination with noise; this dataset allows for an unbiased evaluation of the RBFN predictions, as it is built using the original model and thus is independent from the RBFN training data generation procedure. The mean absolute error (MAE) between RBFN predictions and the original BSM1 data and the R<sup>2</sup> coefficient were found equal to 0.88 g N.m<sup>-3</sup> and 0.99, respectively. A visual comparison between the actual and predicted values is shown in Figure 2. It can be seen that the proposed framework manages to produce RBFN models that accurately track the data produced by the original BSM1 model, despite using a very limited amount of data from it (just two months in this case).

### 5. Conclusions

This work introduces a dynamic modelling scheme to derive highly accurate data-driven predictive models for WWTPs by using only a limited amount of data. The collected data are fed to a novel CPSO-based identification scheme, to accurately estimate the values of the kinetic and stoichiometric parameters of the BSM1 model. The performance of this method is verified through comparisons against two other PSO algorithms.

*Table 1. Performance Metrics*

Algorithm	Fitness Average	Fitness standard deviation	Best Fitness	Average Function Evaluations	p-value
CPSO	0.00068	0.00014	0.00041	2348	-
Standard PSO	0.00223	0.00244	0.00045	2726	0.0311
Modified PSO	0.00106	0.00076	0.00044	2535	0.0885

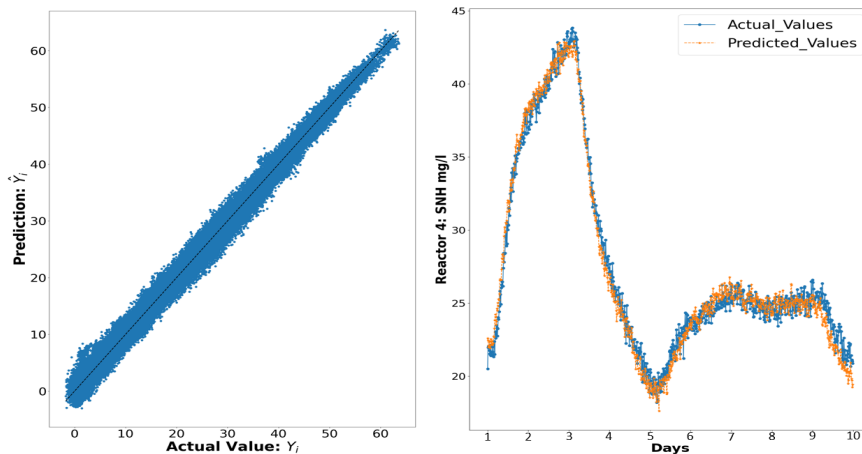


Fig. 2. Results for the RBF-based ammonia concentration model (a) actual vs predicted values for the whole testing dataset and (b) actual and predicted values vs time for an arbitrary selected window of the testing dataset

The identified model is used to produce large volumes of dynamic data for training discrete time RBF predictive models. The results indicate that the derived models are capable of predicting the dynamic behavior of the WWTP with high accuracy, while at the same time being much simpler than first-principle-based models; these properties make them ideal candidates for integration in MPC schemes, a step which will be pursued as feature research.

**Acknowledgements:** This research has been co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code: T2EAK-02191).

## References

- J. Alex, L. Benedetti, J. B. Copp, et. al, 2008, Benchmark simulation model no.1 (BSM1), Report by the IWA Task group on benchmarking of control strategies for WWTPs, 1.
- A. Alexandridis, E. Chondrodima, H. Sarimveis, 2013, Radial Basis Function network training using a non symmetric partition of the input space and particle swarm optimization, *IEEE Transactions on Neural Networks and Learning Systems*, 24(2), 219-230.
- A. Kapnopoulos, A. Alexandridis, 2022, A cooperative particle swarm optimization approach for tuning an MPC-based quadrotor trajectory tracking scheme, *Aerospace Science and Technology*, 127, 107725.
- A. N. Matheri, F. Ntuli, J. C. Ngila, T. Seodigeng, C. Zvinowanda, 2021, Performance prediction of trace metals and cod in wastewater treatment using artificial neural network, *Computers & Chemical Engineering*, 149, 107308.
- E. Mezura - Montes, C. A. Coello Coello, 2011, Constraint – handling in nature-inspired numerical optimization: Past, present and future, *Swarm and Evolutionary Computation*, 1(4), 173-194.
- M. Papadimitrakis, A. Alexandridis, 2022, Active vehicle suspension control using road preview model predictive control and radial basis function networks, *Applied Soft Computing*, 120.
- M. Sadeghassadi, C. Macnab, B. Gopaluni, D. Westwick, 2018, Application of neural networks for optimal-setpoint design and MPC control in biological wastewater treatment, *Computers & Chemical Engineering*, 115, 150-60.

# Can electrical utilities play a decisive role in generating knowledge for better use of energy in small and medium-sized industries?

José Luis Suarez Castaneda,<sup>a</sup> Mélissa Lemire,<sup>a</sup> Marc-André Richard<sup>a</sup>

*<sup>a</sup>Laboratoire des technologies de l'énergie, Institut de Recherche d'Hydro-Québec, 600 av de la Montagne, Shawinigan, Québec, G9N 7N5, Canada*

## Abstract

This paper presents two visualization tools that use meter data to facilitate power management in the industrial sector. The first determines typical daily load profiles based on shape and magnitude with a clustering algorithm. A whole year's worth of electricity consumption can then be visualized on a calendar featuring typical daily profiles on a single screen. Different clustering algorithms and scaling techniques are compared for twelve plants from four different industrial sectors. Data scaled between 0 and 1 or with the normal distribution and the k-Means algorithm showed the best performance. The second tool specifically addresses the management of maximum power demand by suggesting power shedding levels and showing their impacts on billing.

**Keywords:** electricity consumption, power demand, cluster analysis, operational control, SME energy efficiency

## 1. Introduction

To achieve net-zero emissions by 2050, electrification is the preferred option for industrial decarbonization in Québec, Canada, since the vast majority of the power generated by its public electric utility, Hydro-Québec, comes from renewable resources and it is competitively priced (Hydro-Québec, 2022). Therefore some pathways to achieve climate targets predict an increase in electricity demand of the order of two-thirds more than current levels (Dunsky, 2021). The substantial increase in demand for both energy and capacity required for this transition could limit decarbonization potential or require major investments in infrastructure, which would in turn lead to rising energy costs and delays. It is therefore essential that every industrial sector optimize its energy use and maximize its flexibility. In this context, the current role of a utility is not only to supply energy, but also to facilitate such a transition in the most efficient way possible. With limited resources and information, demand side management in small and medium-sized industries is particularly challenging (Thiede et al., 2013). The focus of programmes should be to facilitate knowledge creation among industries (Palm et al., 2020).

Previous work (Narciso and Martins, 2020; Richard et al., 2017; Zhou et al., 2013) confirmed the benefits of using machine learning strategies and visualization tools to facilitate the understanding of electricity consumption. However, most of these works focus on a single industrial activity do not confirm the method performance when analyzing different production processes. The proposed solution is a platform to help industries better understand their energy consumption and act accordingly. It comprises three modules: the first focuses on metering and public data; the second introduces

production data and the third integrates an IoT solution. This paper focuses on the first module, more specifically the description of two visualization tools. The first one is a clustering algorithm that groups typical daily load profiles of a production plant based on shape and consumption level (“amplitude”) from a year’s worth of meter data. A whole year’s electricity consumption can be visualized on a calendar featuring typical daily profiles on a single screen. Automatically generated typical profiles help industrial customers relate the shape of the profiles to their production schedule. This may stimulate a discussion about various demand side management measures, often related to operational control. Different clustering algorithms and scaling techniques are compared in Section 2 of this paper. The second visualization tool specifically addresses the management of maximum power demand by suggesting and evaluating power shedding levels and is discussed in Section 3. Finally, Section 4 presents the conclusion.

## 2. Typical daily load profiles (clustering)

### 2.1. Methodology

#### 2.1.1. Data

Contrarily to Zhou et al. (2013) who use clustering algorithms for load classification purpose, *i.e.* segmentation of electricity consumers, this work aims to analyze electricity consumption of a single production plant by identifying typical daily load profiles, and therefore, energy consumption patterns. To verify the performances of the algorithms of the proposed solution, three plants from each of four different industrial processes were considered: pharmaceutical (plants A to C), chemical (plants D to F), food (plants G to I) and plastic and rubber (plants J to L). To avoid loss of information, no features were extracted and no outliers were removed so that both good and bad consumer practices could be identified if unusual load profiles recur. However, if missing data (or non-existent data) is greater than 10% of all data for a given day, this day is removed from analysis; otherwise, missing values are interpolated. Raw data corresponds to the plant’s electrical energy consumption (kWh) for every 15 minutes for the whole year 2021. The data is organized such that each day,  $i$ , is represented by a vector with 96 dimensions ( $L_i = [x_{i1}, x_{i2}, x_{i3}, \dots, x_{iN=96}]$ ). All the data is organized in a  $365 \times 96$  matrix. However, the inclusion of a supplementary dimension representing this vector norm (as a representation of consumption level) is included as an option to evaluate the impact of this “load amplitude”. The vector norm is calculated using Eq. (1):

$$\|L_i\| = \left( \sum_{j=1}^N |x_{ij}|^2 \right)^{1/2} \quad \text{Eq. (1)}$$

As mentioned, this module aims to identify the typical consumption profile. Then, the average consumption in each time step for each cluster is considered as the expected one. Therefore, given a consumption profile and belonging to a cluster, it is possible to evaluate the consumption of a plant considering its deviation.

#### 2.1.2. Clustering algorithm

Clustering analysis is a machine learning tool to identify patterns by forming groups of data that are similar to one another but different from other groups. This technique is an unsupervised learning method because target values are not known. Most of this work has been aimed at comparing the consumption of different plants, buildings and industries (Zhou et al., 2013). However, Richard et al. (2017) and Bourdeau et al. (2021) have shown the advantages of using clustering to find the typical profiles that represent the expected

behaviors of plant/building consumption. Before selecting the specific clustering method, the data to be analyzed had to be prepared. In addition to the treatment mentioned above, three different data scaling techniques were used: based on the vector norm (each day is scaled individually), min-max of the whole matrix (between 0 and 1) and normal (or Gaussian) distribution of the whole matrix. These scaling techniques are represented in the following equations:

$$L_i' = L_i \cdot [1/\|L_i\|] \tag{Eq. (2)}$$

$$x_{ij}' = \frac{x_{ij} - \min(X)}{\max(X) - \min(X)} \tag{Eq. (3)}$$

$$x_i' = \frac{x_i - \mu}{\sigma} \tag{Eq. (4)}$$

where  $L_i'$  is the scaling vector for day,  $i$ ,  $x_{ij}'$  is the normalized value ( $j$  is the 15-minutes interval between readings) and  $\mu$  and  $\sigma$  are respectively the mean and standard deviation of all data (whole year,  $X$  matrix). When the norm is included as the last dimension, it is scaled considering all the norms of all vectors.

This paper summarizes the performance evaluation of the most common clustering methods used for load classification:  $k$ -means (Bourdeau et al., 2021),  $k$ -shape (Yang et al., 2017), Gaussian mixture model (GMM) (Wang et al., 2022), and fuzzy  $c$ -means (FCM) (Zhou et al., 2013). Each clustering method is evaluated using the three scaling techniques mentioned above. The number of clusters is obtained by maximizing the silhouette score (Rousseeuw, 1987), which is also used to compare the performance of clustering strategies. The silhouette score for an object  $i$  belonging to cluster  $A$  is defined in Eq. (5).

$$S(i) = \frac{b(i) - a(i)}{\max(a, b)} \tag{Eq. (5)}$$

where  $a$  is the average intra-cluster distance and  $b$  is the average shortest distance to another cluster Euclidian distance is used for these calculations. The range of this score is  $[-1,1]$ , therefore the value closest to 1 represents the best clustering strategy.

## 2.2. Results

In this section, we present the results of the clustering strategies. Figure 1 shows the average silhouette scores for all clustering methods and the three corresponding scaling techniques. In general,  $k$ -means, GMM and FCM algorithms performed the best. The number of clusters for GMM and  $k$ -means are higher than or equal to FCM (Table 1).

Table 1. Number of clusters for all clustering strategies (min-max scaling technique)

Plant/Algorithm	A	B	C	D	E	F	G	H	I	J	K	L
$k$ -means	3	3	2	2	2	2	3	3	6	2	3	5
$k$ -shape	4	2	2	3	2	2	2	4	2	3	2	2
GMM	3	3	2	2	3	2	2	3	6	3	3	6
FCM	3	2	2	2	2	2	3	3	5	2	3	2

A higher number of clusters allows for more information on plant performance and production behavior to be obtained. However, GMM is more sensitive to outliers, so unusual loads are identified as a cluster with one profile, which can affect the silhouette score and induce distortions. This is the case for the plant G, which explains why the number of clusters is higher for  $k$ -means than for GMM. On the other hand, it can be seen



in Figure 1 that scaling based on the vector norm is less successful. Although adding the norm as an additional dimension increases the score, it is still worse than other techniques. The supplementary dimension for the other scaling techniques does not have a significant effect.

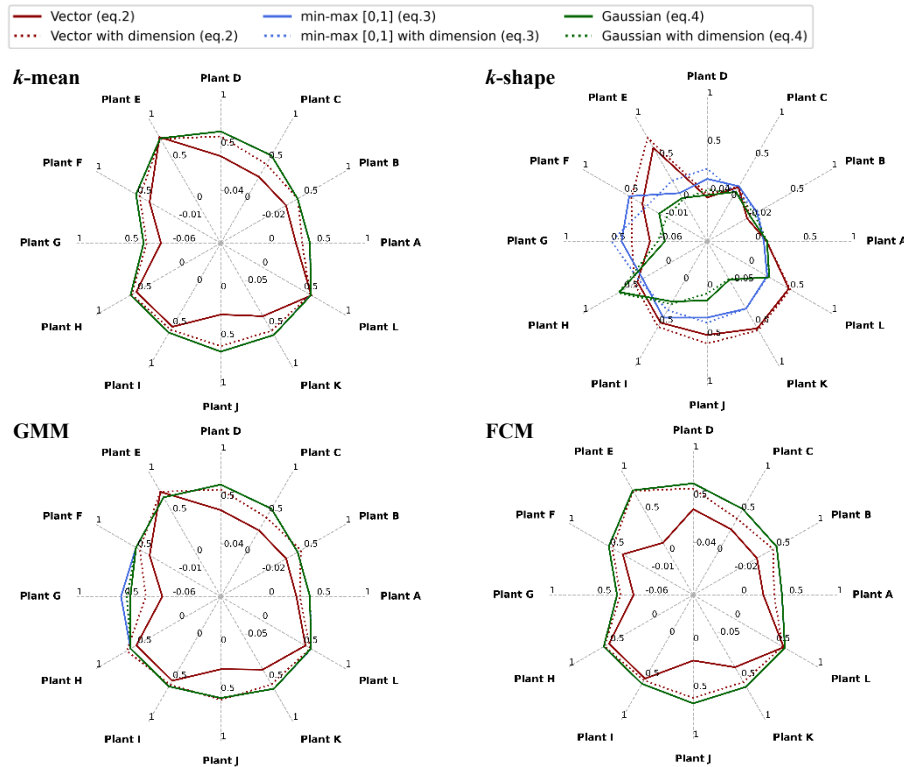


Figure 1. Average silhouette scores for all clustering strategies

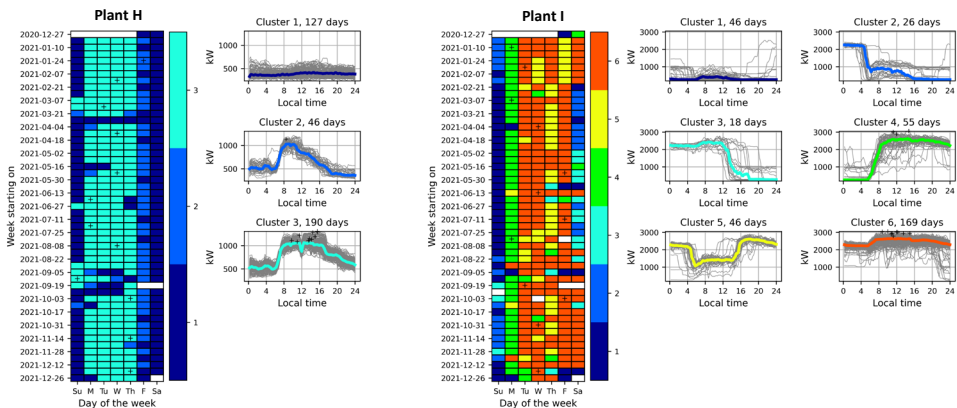


Figure 2. Cluster distribution on calendar and suggested clusters for plants H and I with *k*-means and min-max scaling technique.

Figure 2 presents the proposed display. The clustering strategy allows production days to be distinguished from non-production days and identifies the shape of some typical

production profiles. For instance, plant I cluster 5 represents days when maintenance was being performed on major equipment. The module can incite customers to stop unnecessary loads during non-production hours. For example, plant H had a nighttime load that was half of its load during the day. The day and time of the peaks that induced power demand costs are indicated with a plus sign (+).

### 3. Peak power demand management tool

#### 3.1. Methodology

The second visualization tool aims to raise interest in power demand management by comparing how three different levels of power shedding can impact electricity billing. The structure of the rate in Québec for business customers includes billing demand (CAD/kW) and energy charges (CAD/kWh) (Hydro-Québec, 2022).

The proposed indicator to demine load shedding levels is based on the marginal cost of electricity during peaks. The highest peaks typically account for a limited quantity of energy, but significantly affect billing demand. Hence, above a given level, if power demand costs are distributed over energy, the marginal energy cost increases to a few dollars per kWh. The power threshold,  $P_{Threshold}$ , (in kW), is calculated using the marginal energy,  $e_{marg}$  (in kWh) and the distributed power demand cost parameter, DPDC (in CAD/kWh), using:

$$e_{marg} = \int_{\text{Billing period}} (P(t) - P_{Threshold}) \cdot (H(P(t) - P_{Threshold})) dt \quad \text{Eq. (6)}$$

$$P_{Threshold} = P_{max} - \frac{DPDC * e_{marg}}{PDC} \quad \text{Eq. (7)}$$

where  $P_{max}$  is the maximum power demand of the billing period and PDC is the power demand cost, CAD15.154/kW (Hydro-Québec, 2022). Marginal energy is modeled by equation (6). Where  $H$  is the Heaviside step function. The proposed algorithm to resolve this problem is as follow. All power data is sorted in descending order based considering the accumulated value of time on the abscissa axis. This relationship between power and time allows us to obtain the time  $t_{Threshold}$  according to a fixed  $P_{Threshold}$ . Then, equation (8) can be used to obtain the marginal energy. These equations are solved iteratively until reaching a convergence in the value of DPDC to CAD1.00/kWh (based on experience) to determine the minimum plausible load shedding power level. The two additional levels were evenly divided based on power demand.

$$e_{marg} = \int_0^{t_{Threshold}} (P - P_{Threshold}) dt \quad \text{Eq. (8)}$$

Not only the DPDC parameter is significant to customers, but this indicator also indirectly allows for the shape of the profile and the cut-off time to be taken into account. It also facilitates price comparisons with alternative energy sources.

#### Results and discussion

Table 2 shows results of the peak power demand management tool. Figure 3 shows a month in winter for plants B and D.

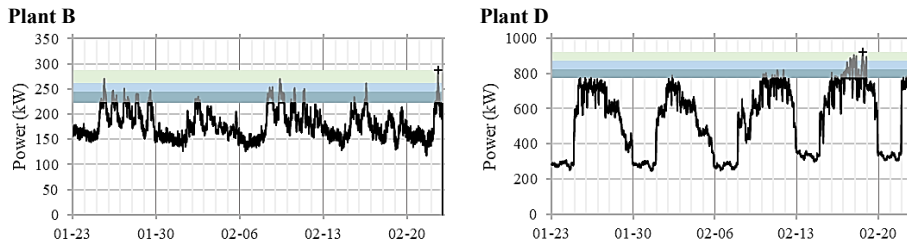


Figure 3. Peak power management tool example and impact on billing for plant B and plant D

Table 2. Summary of peak power management tool

Period		Winter month				Summer month			
Plant	Shedding	$P_{\text{threshold}} - P_{\text{max}}$ (kW)	Power demand savings (CAD)	% savings on total bill	Distributed power demand costs above threshold (CAD/kWh)	$P_{\text{threshold}} - P_{\text{max}}$ (kW)	Power demand savings (\$ <sup>CAD</sup> )	% savings on total bill	Distributed power demand costs above threshold (CAD/kWh)
B	1	263-288	397	3.47%	38.60	301-324	359	3.01%	1.56
	2	244-288	685	5.99%	5.43	298-324	406	3.40%	1.40
	3	222-288	1028	8.98%	0.99	287-324	584	4.90%	0.93
D	1	869-921	806	2.46%	8.12	871-896	389	1.17%	6.36
	2	821-921	1562	4.77%	2.20	847-896	767	2.30%	2.08
	3	772-921	2334	7.13%	1.00	823-896	1139	3.42%	0.99

#### 4. Conclusion

The combined use of clustering algorithm  $k$ -means and scaling data, either between 0 and 1 or using the normal distribution, for cluster analysis achieved the best performance in determining typical profiles in the plants analyzed, i.e., silhouette score, number of clusters, and number of profiles per cluster. For these scaling techniques, the vector norm (as a measure of consumption level) has no effect on the results. The silhouette score can clearly identify different typical profiles. However, more detailed divisions may be required to identify different groups of typical profiles with similar consumption levels but with different consumption shapes. An optional feature integrating a user choice is considered in the proposed solution. Future work could be aimed at improving and evaluating other metrics that would suggest efficient and meaningful clustering for the customer.

In this study, we have also confirmed the great variability in electricity consumption, even within the same industrial sector. This calls for further work on robust solutions to meet this challenge. In the same vein, although twelve different plants (representing four different industrial sectors) are evaluated in this study, a results validation is included in the following stages of the project to consider other types of industries, batch and continuous processes, as well as customer experience. The power demand cost savings calculated in the peak power management tool and the distributed power costs in kWh clearly illustrate the effect of shedding power and can encourage customers to manage peak power demand. They would be able to see the number of hours and quantity of the power to shed and the resulting savings. Future work will focus on the validation of the cut-off threshold and methodology as they relate to customer experience (technical and financial feasibility).

#### References

D.A.C. Narciso, F.G. Martins, 2020, Application of machine learning tools for energy efficiency in industry: A review, *Energy Reports*, 6, 1181–1199.

- Dunsky, 2021, Dunsky Advises Government of Quebec on Pathways to Achieve Climate Targets, <https://www.dunsky.com/dunsky-advises-government-of-quebec-on-pathways-to-achieve-climate-targets>
- EU Commission, 2014, European Competitiveness Report, 190-195.
- Hydro-Québec, 2022, Electricity Rates Effective April 1, 2022, 204 p. <https://www.hydroquebec.com/business/customer-space/rates/>
- H. Wang, Y. Tian, A. Li, J. Wu & G. Sun, 2022, Resident user load classification method based on improved Gaussian mixture model clustering, MATEC Web of Conferences 355, 2021 International Conference on Physics, Computing and Mathematical (ICPCM2021).
- J. Palm, F. Backman, 2020, Energy efficiency in SMEs: overcoming the communication barrier, Energy Efficiency, 13:809–821.
- J. Yang, C. Ning, C. Deb, F. Zhang, D. Cheong, S. E. Lee, C. Sekhar, K. W. Tham, 2017, k-shape clustering algorithm for building energy usage patterns analysis and forecasting model accuracy improvement, Energy and Buildings, 146:27-37.
- K.-L. Zhou, S.-L. Yang, C. Shen, 2013, A review of electric load classification in smart grid environment. Renewable and Sustainable Energy Reviews, 24:103–110.
- M.-A. Richard, H. Fortin, M.-A. Leduc, A. Poulin, M. Fournier, 2017, Daily load profiles clustering: a powerful tool for medium-sized industries demand side management, ACEEE Summer Study on Energy Efficiency in Industry 2017, Denver, Colorado.
- M. Bourdeau, P. Basset, S. Beauchêne, D. Da Silva, T. Guiot, D. Werner, E. Nefzaoui. Classification of daily electric load profiles of non-residential buildings, Energy & Buildings, 233.
- P.J. Rousseeuw, 1987, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, J Computational Appl Math, 20:53-65.
- S. Thiede, G. Posselt, C. Herrmann, 2013, SME appropriate concept for continuously improving the energy and resource efficiency in manufacturing companies, CIRP Journal of Manufacturing Science and Technology, 6:204–211



# Hybrid data-driven and first principles monitoring applied to the Tennessee Eastman process

Eduardo Iraola<sup>a,b</sup>, José M. Nogués<sup>a</sup>, Antonio Del Rio Chanona<sup>c</sup>, Lluís Batet<sup>b</sup>,  
Luis Sedano<sup>b,d</sup>

<sup>a</sup> *Inprocess Technology and Consulting Group, Carrer de Pedro i Pons, 9, 08034  
Barcelona, Spain*

<sup>b</sup> *ETSEIB, Universitat Politècnica de Catalunya, Avinguda Diagonal, 647, 08028  
Barcelona, Spain*

<sup>c</sup> *Sargent Centre for Process Systems Engineering, Imperial College London, Roderic  
Hill Building, South Kensington Campus, London SW7 2AZ, United Kingdom*

<sup>d</sup> *FUS ALLIANZ Science, Engineering & Consulting. C/ Nord 19, Àtic, 43700. El  
Vendrell, Tarragona, Spain*

*eduardo.iraola@inprocessgroup.com*

## Abstract

In this work we present a hybrid monitoring approach for fault detection using the Tennessee Eastman (TE) process. We benchmark our proposed approach against previous methods in the available literature and analyze the benefits and shortcomings. The hybrid monitoring approach contains two steps. First, from a model-based perspective, a dynamic model of the TE plant is constructed using commercial dynamic simulation software and data is generated from the TE plant and its model. In the second step, a data-driven analysis is conducted. This involves training different fault detection models with the previously obtained datasets to detect plant faults as early as possible. The results show that combining datasets can improve the traditional pure data-driven monitoring performance with only plant data. This work highlights the usefulness of combining process modeling and machine learning in the monitoring and prognostics fields when data availability from the actual process is limited.

**Keywords:** Tennessee Eastman, monitoring, Aspen HYSYS, machine learning, fault detection.

## 1. Introduction

Process monitoring and fault detection is critical in the process industry in terms of improving process performance, reducing plant halts, and therefore improving return of investment. Generally, we can approach process monitoring from a model-based or a data-driven point of view. Model-based perspectives—for instance, observers and parity equations—are based on traditional control theory and are well-suited for small systems [1]. However, this approach has not transcended to industry-scale processes due to the difficulty in adapting the high complexity of these processes to this kind of models. Data-driven methods can overcome this issue by being trained using only output data of the system. Therefore, they can describe the complexity of a system without the need to know the intrinsic process details, which can be time-consuming and impractical to identify. However, data-driven approaches do not extrapolate well to regions of the state space not

observed in training. In these cases, given judicious assumptions, model-based approaches can extrapolate and allow better interpretability. We can see that both approaches show advantages and disadvantages, and the choice will vary depending on the specific case.

In this work, we present a hybrid fault detection approach that merges data fed directly from the plant and model data derived from first principles. The goal is to evaluate if there is any benefit in using a model to boost the information we provide the fault detector. In this way, we take advantage of rigorous simulation to enhance the results of a pure data-driven approach. Early examples of this kind of hybrid approach can be found in Schwarte and Isermann (2002) [2], however, in our opinion, this research theme can be further improved.

The paper is organized as follows. Section 2 outlines a short review of data-driven fault detection methods and formalizes the hybrid approach. Section 3 describes the case study selected for this work, the Tennessee Eastman process. Then, Section 4 shows the results of the benchmark. Finally, Section 5 summarizes the conclusions of the work.

## 2. Fault detection

Fault detection is the branch of process monitoring that studies how to identify events that cause a variable or property of the system to deviate from an allowed range. Figure 1 summarizes a possible process monitoring and fault diagnostic framework based on Aldrich and Auret (2013) [3]. In this scenario, the monitoring system is trained with normal operation conditions (NOC) data only, therefore a semi-supervised approach is followed. A semi-supervised approach provides the advantage of not needing real fault data for training, which may be expensive or impractical to obtain—e.g., operating the plant to off-control states or introducing dangerous disturbances just to acquire data. Supervised classifiers may struggle with any kind of new fault that was not previously trained for; therefore, they can be easily outperformed by semi-supervised approaches.

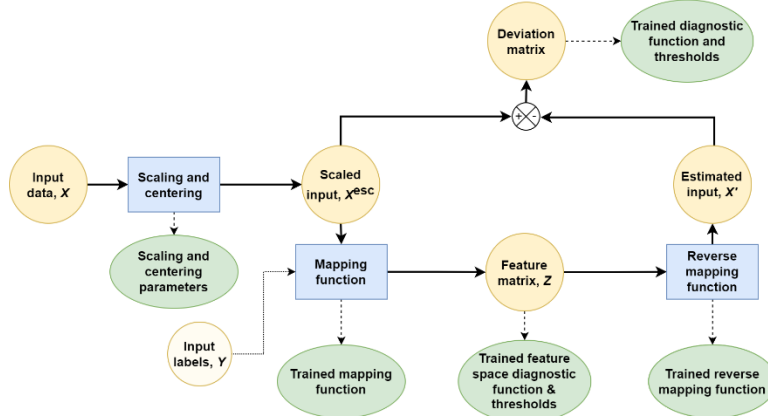


Figure 1. Semi-supervised fault detection framework, with data preprocessing, mapping, and reconstruction steps.

A monitoring system such as that in Figure 1 works as follows: it develops a series of features,  $Z \in \mathbb{R}^{n \times n_z}$ , based on a raw dataset,  $X \in \mathbb{R}^{n \times m}$ , which is usually preprocessed to improve performance, and tries to reconstruct the input into  $X' \in \mathbb{R}^{n \times m}$ .  $n$  is the number of instances of the dataset, and  $m$  and  $n_z$  are the size of the input and the latent space, respectively (typically  $m > n_z$ ). Process monitoring compares the reconstructed

input with the actual input and extracts the result of the in-control or out-of-control state of the plant.

### 2.1. Types of semi-supervised data-driven fault detection methods

One of the fundamental fault detection methods is *univariate statistical monitoring* [4], which obtains the operation limits using statistical theory and hypothesis testing. Given a level of significance,  $\alpha$ , and assuming Gaussian distribution on the input data, our null hypothesis would assume normal operation, while the alternative hypothesis would imply the existence of faults. This way, if the mean and standard deviation of the tracked variable,  $x$ , is  $\mu$  and  $\sigma$  respectively, we say that the process is in normal operation conditions (NOC) if  $x$  is inside the following range:

$$\mu - c_{\alpha/2}\sigma \leq x \leq \mu + c_{\alpha/2}\sigma, \quad (1)$$

where  $c_{\alpha/2}$  is the standard normal deviate for the  $1 - \alpha/2$  percentile. This approach considers serial relationships between a variable state in different time steps, but fails to include spatial relationships, i.e., the influence between different variables at the same instant in a complex process.

To consider this influence, we need to work with a *multivariate statistical monitoring* approach. In this case, instead, the Hotelling's  $T^2$  test should be used [4]. To this end, the statistic is  $T^2 = \mathbf{z}^T \mathbf{z}$ , where the feature vector,  $\mathbf{z} \in \mathbb{R}^{n_z}$ , is  $\mathbf{z} = \Lambda^{-1/2} V^T \mathbf{x}'$ , with  $\mathbf{x}' \in \mathbb{R}^m$ .  $\Lambda \in \mathbb{R}^{n_z \times n_z}$  is the diagonal matrix containing the  $n_z$  eigenvalues of the covariance matrix of the training dataset,  $X$ , and  $V \in \mathbb{R}^{n_z \times m}$  is the matrix of the eigenvectors of the same eigenvalue decomposition. Equation (2) shows how to compute the  $T^2$  detection threshold for an  $\alpha$  level of significance when we use the sample covariance matrix,

$$T_\alpha^2 = \frac{m(n-1)(n+1)}{n(n-m)} F_\alpha(m, n-m) \quad (2)$$

The output of this fault detection approach is a confidence ellipse for which we accept a new measurement as NOC if it satisfies  $T^2 < T_\alpha^2$ . If we cannot assume a Gaussian distribution or the input dataset is too small, we would need to consider more robust hypothesis tests [5].

The multivariate detection approach can be further improved by dimensionality reduction techniques such as Principal Component Analysis (PCA), Fisher Discriminant analysis (FDA), Partial Least Squares (PLS), or Canonical Variate Analysis (CVA) [4]. These perform the same multivariate statistical testing but now apply it to a lower dimensional dataset that is more representative of the actual trends of the system, rejecting noise and non-useful information. For instance, PCA projects the raw data into the lower dimensional space created by the  $k$  eigenvectors corresponding to the highest eigenvalues of the covariance matrix, being  $k < m$ . Then, the detection threshold can be calculated using Equation (2), with  $k$  number of features instead of the original number of features  $m$ . If several timesteps are stacked in the same time window that is fed to the PCA algorithm to also account for serial relationships, we call this method Dynamic PCA (DPCA). The added time series analysis tends to outperform standard PCA.

More advanced techniques for fault detection include the use of machine learning models. Given that, we are focusing on semi-supervised approaches in which we have only NOC data from the plant, a vanilla neural network classifier architecture does not apply to a semi-supervised approach. It is here where auto-associative networks or autoencoders appear.

Autoencoders (see Figure 2) use an encoder set of layers to transform a single input instance  $\mathbf{x}$  into a latent space representation  $\mathbf{z}$ , typically of lower dimensionality than the input data, and then use the decoder layers to return a reconstruction of the input  $\mathbf{x}'$  from the latent representation obtained by the encoder [6]. As in Figure 1, the deviations



between the actual input and the reconstruction are used to determine if the instance corresponds to normal operation or not. Unlike the previous approaches, autoencoders can capture nonlinear relationships between the different variables, therefore being useful for more complex processes. For this reason, these models are also sometimes called nonlinear PCA (NLPCA).

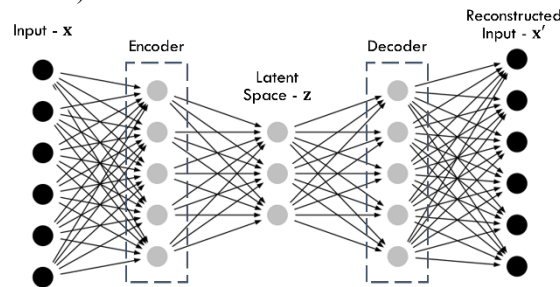


Figure 2. Autoencoder architecture.

A set of the different methods presented in this section will be tested with both pure data-driven and hybrid dataset approaches, the latter being explained next.

### 2.2. Hybrid fault detection approach

In this section, we present a hybrid fault detection approach that adds the advantages of model-based methods to data-driven approaches, while considering the tools already available in the process industry. This method leverages process simulation models to improve the dataset preprocessing step using software such as Aspen HYSYS.

Here, the plant and the model use the same high-level vector of control signals,  $\mathbf{u}$ , as inputs; and their outputs,  $\mathbf{y}$  and  $\mathbf{y}^*$ , respectively, are compared. This results in the so-called residual generation,  $\mathbf{r}$ . Residuals will make up the datasets for training and cross-validation, and therefore will be fed to the selected fault detection method: limit sensing, PCA, autoencoder, etc. Figure 3 summarizes the procedure.

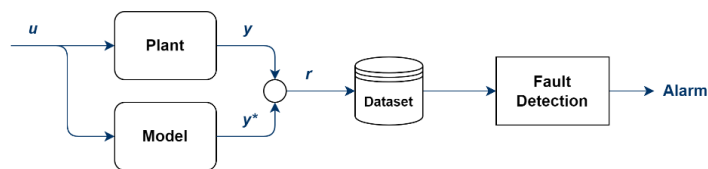


Figure 3. Hybrid fault detection framework.

The hybrid procedure is expected to show better invariant metrics with respect to faults and anomalies in the process: the plant output  $\mathbf{y}$  can be affected by external disturbances, hence intrinsic changes and its comparison with the ideal behavior represented by the model  $\mathbf{y}^*$  can be more efficient as a process monitoring indicator. Note that data harmonization between model and plant must be studied to minimize differences and account for model mismatches. This is assumed as a preliminary condition in this paper and will not be analyzed here.

## 3. Case study – The Tennessee Eastman process

The Tennessee Eastman (TE) process is a virtual plant that Downs and Vogel designed based on an Eastman Chemical Company process [7]. In this virtual plant, a series of reactants, A, C, D, and E are processed to obtain two products, G and H, as pure as possible, with an unavoidable inert component mixed in the feed, B, and a byproduct, F.

To obtain the desired products, the system is made up of five unit operations: a reactor, a condenser, a liquid-vapor phase separator, a compressor, and a stripper column. Figure 4 shows a process flow diagram of the complete system.

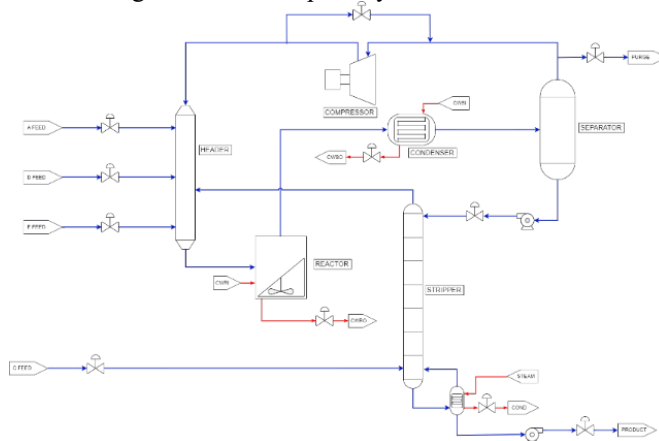


Figure 4. The Tennessee Eastman process.

In terms of the user interface, the TE process presents 41 measurable variables with noise and delay and 12 first-order dynamics actuators that allow for external interaction with the system. The process allows the activation of 20 types of disturbances at any point during the simulation. These can work as the faults to be detected by the monitoring system. The disturbances represent step changes, drifts, random noise, and component malfunctions in the supply conditions, cooling streams, and actuators of the plant.

For this work, we modeled the TE process in Aspen HYSYS, a dynamic simulator widely used in the process industry. We use the original TE script as the *plant* and the HYSYS simulation as the *model*, and collect data from the steady-state operation of both. The training dataset includes normal operation from both plant and model, and the validation dataset includes the activation of different disturbances for the plant and fault-free operation for the model. The new datasets are created according to the scheme in Figure 3, and we apply the monitoring methods summarized in Section 2 to compare them with pure data-driven approaches found in the literature.

#### 4. Results

The results of the hybrid approach over the TE process data are shown in Figure 4 using previous literature results for comparison [4]. We selected the  $F_1$  score,

$$F_1 = \frac{TP}{TP + 0.5(FN + FP)} \quad (3)$$

as the target performance metric to harmonize the different results. TP stands for the number of true positives, FN for the number of false negatives, and FP for the number of false positives predicted by the data-driven model. Therefore,  $F_1$  represents the harmonic mean of precision (the percentage of instances predicted as positive that are correctly classified) and recall (the percentage of true positive instances that are correctly classified).  $F_1$  is widely used because it works as an effective way to balance the effect of false negatives and false positives in classification models. Note that, in the framework of fault detection, positive instances correspond to faults, and negative instances correspond to normal operation.

Figure 5 shows the results. Note that all pure data-driven approaches have similar performance and the hybrid approaches tend to perform better. The DPCA pure data-driven score from the literature shows a small unexpected drop that can be explained by a high false positive ratio caused by the specific threshold calibration procedure used in [4]. The  $F_1$  score heavily penalizes unbalances between false positives and false negatives, therefore the unexpected behavior. Results strongly suggest that accounting for first principles through rigorous simulation can improve data-driven approaches.

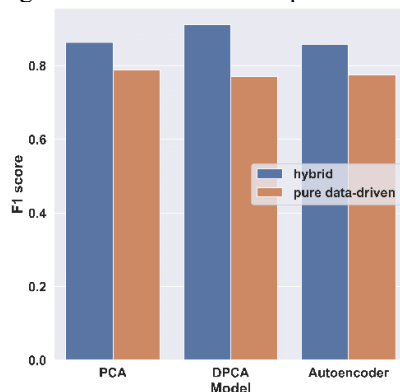


Figure 5. Comparison between pure data-driven (literature) versus hybrid dataset performance.

## 5. Conclusions

This work describes the hybrid fault detection approach and analyses its performance using the classical Tennessee Eastman process. To conduct this benchmark, the Tennessee Eastman original script is assumed to be the actual plant, and an Aspen HYSYS model is built to try to replicate it and generate a new residuals dataset. The results show that the hybrid approach is an interesting framework that tends to improve monitoring performance, even when applied to the most basic monitoring approaches. Future work will introduce more complex datasets, including dynamic plant operation, and perform fully supervised classification, taking advantage of the models to replicate faults and feed them into the fault detection methods and comparing them to the residual-based approach.

## References

- [1] R. J. Patton, P. M. Frank, and R. N. Clark, *Issues of Fault Diagnosis for Dynamic Systems*. London: Springer-Verlag, 2000.
- [2] A. Schwarte and R. Isermann, *Neural network applications for model based fault detection with parity equations*, vol. 35, no. 1. IFAC, 2002. doi: 10.3182/20020721-6-es-1901.01502.
- [3] C. Aldrich and L. Auret, *Unsupervised Process Monitoring and Fault Diagnosis with Machine Learning Methods*. London: Springer London, 2013. doi: 10.1007/978-1-4471-5185-2.
- [4] L. H. Chiang, E. L. Russell, and R. D. Braatz, *Fault Detection and Diagnosis in Industrial Systems*. London: Springer London, 2001. doi: 10.1007/978-1-4471-0347-9.
- [5] R. Wilcox, *Introduction to Robust Estimation and Hypothesis Testing*, 5th ed. Elsevier, 2021. doi: 10.1016/C2010-0-67044-1.
- [6] M. A. Kramer, "Nonlinear principal component analysis using autoassociative neural networks," *AIChE J.*, vol. 37, no. 2, pp. 233–243, 1991, doi: 10.1002/aic.690370209.
- [7] J. J. Downs and E. F. Vogel, "A plant-wide industrial process control problem," *Comput. Chem. Eng.*, vol. 17, no. 3, pp. 245–255, Mar. 1993, doi: 10.1016/0098-1354(93)80018.

# Data-Driven Process Monitoring for Knowledge Discovery: Local and Global Structures

Estelle E. Seghers, José A. Romagnoli

*Department of Chemical Engineering, Louisiana State University, Baton Rouge,  
Louisiana 70803, United States of America*

## Abstract

In industrial processes, the amount of raw data generated can add complexity in the analysis and understanding of the process dynamics. Being able to properly interpret this data can help improve plant operation. A platform is introduced for monitoring of industrial processes and optimization of the model-building process. FASTMAN-JMP (FAST MANUAL Data Manipulation implemented in JMP [1]) is a tool developed in Python to apply various data mining and machine learning techniques quickly and easily to better understand valuable patterns and hidden trends in process data. It is shown that local and global structures in the data set can be visualized and related to actual process operations through the identification of the variables responsible for the separation between selected given clusters. Furthermore, adequate comparisons of these algorithms can be difficult, having different loss functions with many parameters. We aim to decipher these algorithms, and how they work in the context of industrial data. Results are presented for an industrial case study of a pyrolysis reactor.

**Keywords:** process monitoring, data mining, knowledge discovery, machine learning, data-driven modeling.

## 1. Introduction

The digital data being created today has surpassed many projected values, reaching roughly 44 zettabytes ( $10^{21}$  bytes) generated every year as of 2020 [2]. This era of big data introduces an opportunity to improve industrial plant operation and control with increased productivity and safety. Data mining and machine learning algorithms present a way to achieve increased knowledge discovery in industrial processes, but it is also vital to understand how these methods function. Of particular interest is understanding the granularity of these analyses performed with industrial data, namely in how machine learning methods distinguish and illustrate the global and local structures present in data. In the context of industrial chemical engineering plant data, global structures generally suggest overall trends in the process (over a period of hours, or days, or longer) while local structures generally refer to short term fluctuations in certain process variables or controllers. Dimensionality reduction, also known as feature extraction, is a particularly important component within an unsupervised approach for analysing process data.

Most of the discussions in the computer science community deal with synthetic data and the main conclusion may or may not be relevant when dealing with industrial process data. In the author's opinion there is a need for a proper discussion on these issues in the context of industrial manufacturing process data. A Python-based data analysis graphical user interface, called FASTMAN-JMP (FAST MANUAL Data Manipulation implemented in JMP [1]), was developed using Python libraries (NumPy, pandas, sci-kit learn, etc.) and incorporated as an add-in into JMP, a statistical data analysis software, to facilitate the

use and visualization of different methods for data mining and knowledge discovery. With the availability of such a powerful environment, the concepts of global and local data structures are first analyzed using data from an industrial pyrolysis reactor. It will be shown that local and global structures in the data set can be visualized within the Self Organizing Maps (SOM) [3] and related to actual process operations through a subspace greedy search (SGS) algorithm [4] allowing for the identification of the variables responsible for the separation between any two given clusters. Furthermore, when running alternative dimensionality reduction algorithms, it is not clear how these results yield trustworthy representations of the original data distribution and, more importantly, if they lead to proper characterization of the process states. Even comparisons of these algorithms can be difficult having different loss functions with many parameters. Consequently, another goal of this work is to decipher these algorithms, and how they work in the context of industrial chemical engineering process data.

## **2. Understanding the Data: Global and Local Structures**

The goal of dimensionality reduction is to reduce the number of features in a dataset to obtain a compressed representation that is useful for analysis, by retaining as much “structure” in the data as possible. There are many algorithms for dimensionality reduction, and they can generally be discussed in the context of two principal archetypes: those that primarily aim to preserve global structure (overall trends in the process over a period of hours, or days, or longer) and those that primarily aim to preserve local structure (short-term fluctuations in certain variables or controllers). There is no strict definition of what constitutes global or local structure preservation, but global structure-preservation methods may be thought of as preserving overall placement of large clusters of observations, i.e., the data at the extremes in high-dimensional space are also at the extremes in low-dimensional space.

In contrast, local structure-preservation methods focus on preserving neighborhoods of observations, so that near neighbors in the high-dimensional space are still near neighbors in the low-dimensional space. If the data lie along manifolds or in clusters, that information will not necessarily be preserved when projecting to a very low number of dimensions. That is, a technique which preserves global structure does not necessarily preserve which points are close neighbors in the high-dimensional space.

## **3. Dimensionality Reduction: Global and Local Structure Preservation**

In Principal Components Analysis (PCA) [5], a very popular very popular dimensionality reduction technique, the extracted principal component vectors graphically illustrate most of the variance in the data, and therefore primarily captures the global structure of the dataset while ignoring the detailed local structure information among the process data. Since the development of PCA, a class of dimensionality reduction techniques, known as manifold learning, has been developed in the pattern recognition area, including t-SNE [6], UMAP [7], and more recently TriMAP [8]. These methods are popular in biological applications and have been used in chemical manufacturing applications ([6], [9], [10]). However, one critical drawback of these methods is that they were either designed to preserve local structure or struggle with balancing the preservation of both global and local structures. Furthermore, these methods can be sensitive to initialization, and more importantly, they can be difficult to understand and tune.

Pairwise Controlled Manifold Approximation and Projection (PaCMAP) is a dimensionality reduction technique developed by Wang et al. [11] that uses a stochastic gradient descent algorithm, boasts effective and reliable performance with preserving

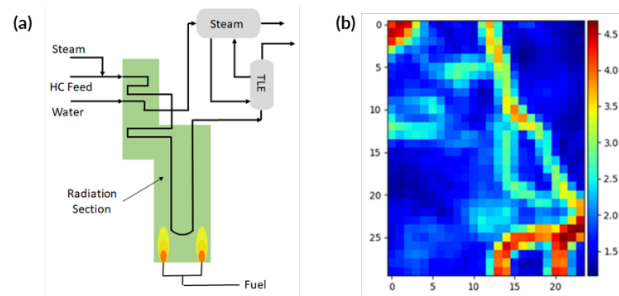
local and global structure alike, and performs significantly faster than other contemporaneous dimensionality reduction methods. Of interest to note is that none of the previously described methods are able to smoothly transition from local structure preservation to global structure preservation or vice versa through manipulation of parameters, however, PaCMAP can and does so automatically by dynamically adjusting its own parameters over a series of iterations through three distinct phases using a unique loss function. PaCMAP performs this optimization of the low-dimensional embedding by using three kinds of pairs of datapoints: neighbor pairs, mid-near pairs, and further pairs. The PaCMAP algorithm applies the initialization procedure selected to set the initial values of the low-dimensional embedding (initialization can be done randomly or using PCA). Once initialized, the values are then optimized. The algorithm uses the Adam Optimizer to reduce the loss function. The overall loss function is defined as:

$$\begin{aligned} Loss^{PaCMAP} = & w_{NB} \cdot \sum_{i,j} \frac{\tilde{d}_{ij}}{10 + \tilde{d}_{ij}} && (\text{neighbors; strong attraction}) \\ & + w_{MN} \cdot \sum_{i,k} \frac{\tilde{d}_{ik}}{10,000 + \tilde{d}_{ik}} && (\text{mid-near pairs; weak attraction}) \\ & + w_{FP} \cdot \sum_{i,l} \frac{1}{1 + \tilde{d}_{il}} && (\text{further pairs; repulsion}) \end{aligned}$$

where  $d$  is the scaled distance between a pair of observations,  $i$  and  $j$  are neighbors,  $i$  and  $k$  are mid-near pairs,  $i$  and  $l$  are further pairs, and  $w$  refers to the weights of all the pairwise interactions that change over the course of iterations. In general, the goal is to strongly attract neighbor pairs, weakly attract mid-near pairs, and repulse further pairs to properly construct the global and local structures. In the first stage of the PaCMAP algorithm's optimization, the initial embedding of points is improved to preserve primarily the global structure. In the second phase, the focus eases to improving the local structure while maintaining the developed global structure by moving to a higher weighting of neighbor pairs. Finally, by more heavily weighting the further pairs and neighbor pairs, the third phase shifts to refinement of the local structure without adjusting the global structure.

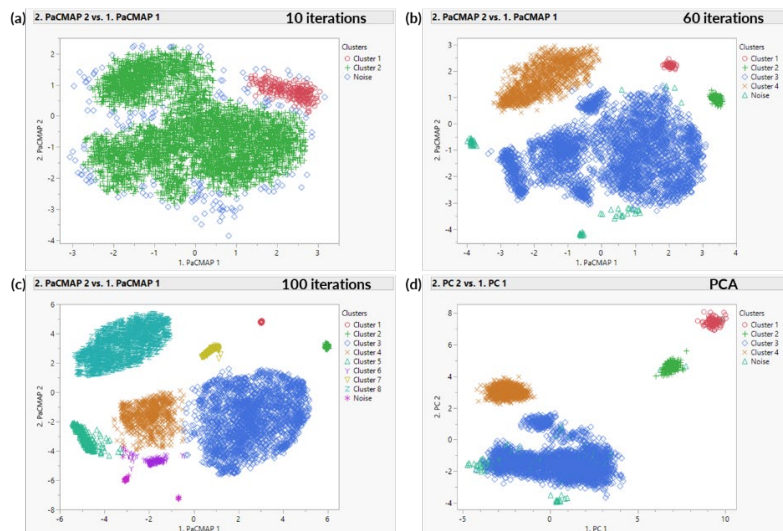
#### 4. FASTMAN-JMP Use-Case of an Industrial Pyrolysis Reactor

The implementation of this tool was tested using an industrial pyrolysis reactor case study. The use of an industrial dataset illustrates the suitability and effectiveness of using FASTMAN-JMP to apply the methodology and algorithms discussed to complex problems in the chemical engineering industry. The industrial pyrolysis reactor dataset analyzed covers a period of several months. The reactor, which is a very well-known chemical process unit, cracks heavy hydrocarbons into higher-value, lighter hydrocarbons. These datasets depict multiple modes of steady-state operation as well as process drift due to a buildup of equipment coke. The chemical reaction occurs in a fired furnace that is heated by burning fuel gas, as shown in **Figure 1**. This pyrolysis reactor takes a naphtha feed, injects it with steam, and cracks it into ethylene at extremely high temperatures. The entire process has 27 variables, including hydrocarbon flow, steam flow, cracking temperatures per coil, and coil outlet temperature. Also shown in **Figure 1** is the associated SOM plot, where the axes are the coordinates of the nodes present in the map. As illustrated by the distinct borders of contrasting colors, several main regions and subregions can be visualized in the SOM plot as an indication of the data structure i.e., global and local events. In this application the process is moving along different regions during the full cycle of operation.

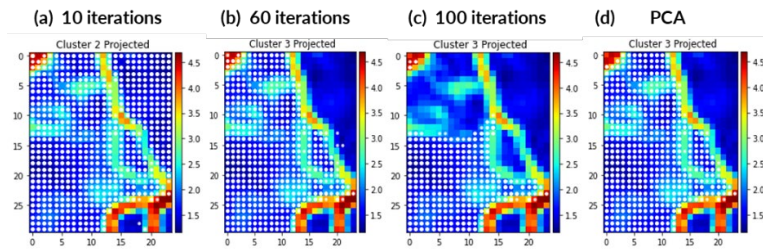


**Figure 1:** (a) Pyrolysis reactor process scheme; (b) SOM plot

**Figure 2** illustrates the 2D projection results of initial stages of the PaCMAP algorithm using PCA initialization as the number of iterations increases. These results are at the very beginning of the first stage of the PaCMAP algorithm. There are very few clusters identified initially, categorizing most of the data into one large grouping. This is also revealed in the clusters projected on the SOM plots in **Figure 3** (corresponding to the clusters identified in **Figure 2**) as indicated by the dots overlaying the SOM. In **Figure 2.a**, “Cluster 1” includes datapoints that cover an area on the SOM with very distinct divisions (the bottom right corner in **Figure 3.a**) and “Cluster 2” encompasses the entire remainder of the area on the SOM (the projected cluster shown in **Figure 3.a**), with both clusters grouping together data that would otherwise be separated based on the more granular SOM coloring. Clearly, this dimensionality reduction method initially identifies the global structure of the data (i.e., the main process events) then gradually eases its global structure optimization and begins optimization of the local structure. Finally, in the later stages, the algorithm focuses on refining the local structure. In comparison to the early stages of the PaCMAP algorithm, **Figures 2.d** and **3.d** correspond to the results using PCA projection which focuses on global data structure. The PCA results are equivalent to those after just 60 iterations of the PaCMAP algorithm.

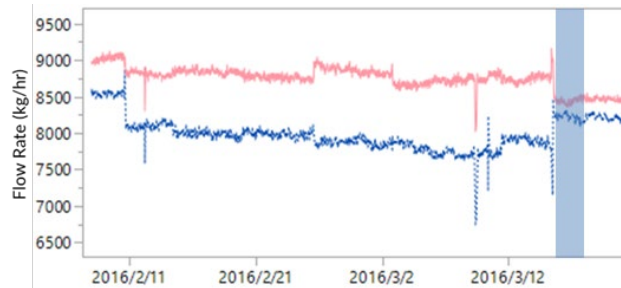


**Figure 2:** (a) PaCMAP results after 10 iterations; (b) PaCMAP results after 60 iterations; (c) PaCMAP results after 100 iterations; and (d) PCA results



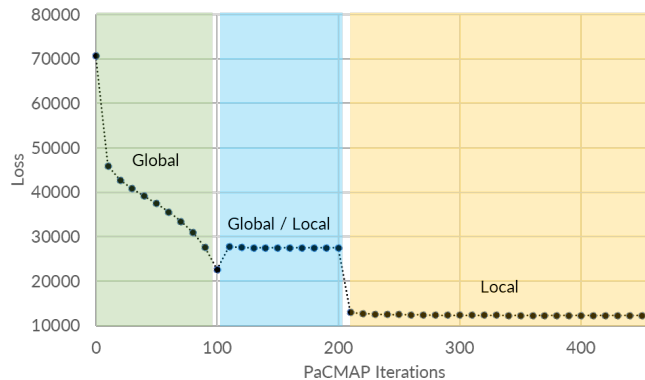
**Figure 3:** PaCMAP cluster projection on the SOM: (a) after 10 iterations; (b) after 60 iterations; (c) after 100 iterations; and (d) PCA cluster projection on the SOM

This is also corroborated in **Figure 4** where the time evolution of the hydrocarbon flow rate in coils 1 and 4 is shown (solid and dashed, respectively). The highlighted section of the data corresponds to the conditions identified as “Cluster 3” in **Figures 2.b** and **3.b**. The identification of key process variables responsible for the cluster formation is performed using the subspace greedy search (SGS) algorithm [4].



**Figure 4:** Time evolution of hydrocarbon flows for coils 1 (solid; top) and 4 (dashed; bottom)

Shown in **Figure 5** is a graph of the value of the PaCMAP loss function for the pyrolysis reactor dataset as a function of number of iterations. As discussed before, the loss function changes over the course of the PaCMAP algorithm stages based on the weights of neighbor pairs, mid-near points, and further points. As shown, the loss is slightly improved as the number of iterations increases in the second and third stages. However, the largest rate of the improvement (i.e., the most noticeable change in the embedding) occurs in the first stage, with the focus on the global structure development primarily focusing on mid-near pairs.



**Figure 5:** PaCMAP loss function evolution as a function of iterations



The sharp transitions in **Figure 5** are due to the distinct changes in the loss function weightings at the onset of the three optimization phases. Of course, since this is based on the relative amount of neighbor pairs, mid-near pairs, and further pairs present in a dataset and the distances between these points, the qualitative trend may differ between datasets.

## 5. Conclusions

The concepts of global and local data structures were analyzed using data from an industrial pyrolysis reactor. It was shown with FASTMAN-JMP that local and global structures in the data set can be visualized and related to actual process operations allowing for the identification of the variables responsible for the separation between any two given clusters. Furthermore, it was shown that PaCMAP is an effective and reliable dimensionality reduction technique that preserves both local and global structures and performs significantly faster than other methods. By building up an understanding of how these dimensionality reduction algorithms work and how to tune parameters – a process made easier with FASTMAN-JMP – we can more effectively use these powerful tools to visualize and extract knowledge from large, high-dimensional datasets in the chemical manufacturing industry.

## References

- [1] *JMP®*, Version 16.2, Cary, NC: SAS Institute Inc., 1989–2021.
- [2] Y. Shi, *Big Data and Big Data Analytics*, Singapore: Springer, 2022.
- [3] T. Kohonen, "The self-organizing map," *Proceedings to the IEEE*, vol. 78, no. 9, pp. 1464-1480, 1990.
- [4] W. Zhu, W. Sun and J. Romagnoli, "Adaptive k-Nearest-Neighbor Method for Process Monitoring," *Industrial & Engineering Chemistry Research*, vol. 57, pp. 257-258, 2018.
- [5] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *Journal of Educational Psychology*, vol. 24, pp. 417-441, 1933.
- [6] L. van der Maaten and G. Hinton, "Visualizing Data using t-SNE," *Journal of Machine Learning Research*, vol. 9, pp. 2579-2605, 2008.
- [7] L. McInnes, J. Healy and J. Melville, "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction," *arXiv:1802.03426v3 [stat.ML]*, 2020.
- [8] E. Amid and M. Warmuth, "TriMap: Large-scale Dimensionality Reduction Using Triplets," *arXiv:1910.00204v2 [cs.LG]*, 2019.
- [9] Z. Webb and J. Romagnoli, "Real-Time Chemical Process Monitoring with UMAP," *Computer Aided Chemical Engineering*, vol. 50, pp. 2077-2082, 2021.

- [10] L. Briceno-Mena, M. Nnadili, M. Benton and J. Romagnoli, "Data mining and knowledge discovery in chemical processes: Effect of alternative processing techniques," *Data-Centric Engineering*, vol. 3, p. e18, 2022.
  
- [11] Y. Wang, H. Huang, C. Rudin and Y. Shaposhnik, "Understanding How Dimension Reduction Tools Work: An Empirical Approach to Deciphering t-SNE, UMAP, TriMap, and PaCMAP for Data Visualization," *Journal of Machine Learning Research*, vol. 22, no. 201, pp. 1-73, 2021.



# Fault detection in high-pressure pumps in Low-Density Polyethylene Autoclave Reactors

Maria Giuliana F. Torraga,<sup>a</sup> Tahyna B. Fontoura,<sup>b</sup> Rodrigo M. Lima,<sup>c</sup> Luciana A. da Silva,<sup>d</sup> Tiago da Silva Osório<sup>d</sup>

<sup>a</sup>*Braskem S.A., Rua da União 765, Jardim Sonia Maria, Mauá - SP 09380-900, Brazil*

<sup>b</sup>*Braskem S.A., Rua Marumbi, 1001, Campos Elíseos, Duque de Caxias - RJ, 25221-000, Brazil*

<sup>c</sup>*Braskem S.A., Av. Pres C. e Silva, 400, Capuava - Santo André - SP, 09270-000, Brazil*

<sup>d</sup>*Braskem S.A., BR 386, Rod. Tabai, km 419, Via Oeste, Triunfo - RS, 95853-000, Brazil*  
[maria.giuliana@braskem.com](mailto:maria.giuliana@braskem.com)

## Abstract

Low-density polyethylene is a polymer obtained via free-radical polymerization at high pressures in both tubular and autoclave reactors. At high pressures, ethylene can suffer an explosive decomposition (“decomp”), leading to runaway reactions. Several disturbances that can lead to “decomp”, such as excess initiator in feed, feed impurity, controller failures, etc. The present work focus on the development of a fault detection algorithm to detect failures in high-pressure pumps in the autoclave process. These pumps inject the initiator and are prone to failures such as leaks, cylinder failures, reversal system failure and check valve failures. The fault detection algorithm was developed to early detect failures in the pumps. The methodology consisted in developing data-driven models to monitor the process. The models were developed using data from an industrial autoclave reactor unit and were tested against failures events that happened between 2019 and 2022, with at least 70% of the events successfully detected between 5 and 50 minutes prior to the failure.

**Keywords:** fault-detection, polyethylene, statistical process control, data-driven models.

## 1. Introduction

Low-density polyethylene (LDPE) is a polymer obtained via free-radical polymerization in bulk in both tubular and continuous autoclave reactors, at high pressure. Free-radical polymerization includes mainly four elementary reactions: initiation, propagation, bimolecular termination, and chain transfer reactions. In this type of polymerization, the polymer microstructure and properties are intrinsically related to the reaction. For instance, the long and short-chain branches observed in LDPE are due to chain transfer reactions to polymer species and backbiting reactions (Hamielec and Tobita, 2000). In order to generate the free radicals to initiate the polymerization, chemical initiators can be used, at low concentrations. These initiators decomposes thermally generating radicals for monomer addition. Commercial thermal initiators include azo and peroxides compounds, and in the case of LDPE, the initiation depends both on the temperature and the pressure of the system.

Free-radical polymerization is highly exothermic ( $\Delta H_{\text{poly}}$  of 24 kcal mol<sup>-1</sup> for LDPE), therefore, the heat removal and temperature control play an important role in both productivity, product quality and process safety. In industry, there are two main processes to produce LDPE, in tubular and in autoclave reactors. In both cases, the reaction is carried out at high pressures (1500-2000 atm) and high temperatures (150-230 °C). An

important reaction that also occurs at high temperature and pressure is the explosive decomposition of ethylene (Zhang et al., 1996), forming hydrogen, carbon, methane, and other low-saturated and unsaturated hydrocarbons. This undesired reaction affects both tubular and autoclave reactors and can lead to a runaway (“decomp”).

Several disturbances that can lead to “decomp”, such as excess initiator in feed, feed impurities, feed inlet temperature, controller failures, mechanical frictions, inadequate mixing, etc (Zhang et al., 1996). Ultimately, all these disturbances can cause hot spots in the reactor, which can initiate the runaway. As Zhang et al. (1996) have pointed out, the decomposition occurs in order of milliseconds, and it is irreversible once it has started. Therefore, several works have been developed to predict failures in LDPE reactors.

Sivalingam et al. (2015) developed a predictive model to detect decomposition for high-pressure ethylene/vinyl acetate copolymerization in an autoclave reactor. The energy balance error was used as a reference to normal operation, considering the heat released by reaction and the sensible heat absorbed by the feed and products. The heat of stirring and lost to the environment were not considered. A principal component analysis (PCA) on heat balance model components was performed, using data from normal operation, under steady state, with 1s time interval for calibration. Control charts based on heat balance allowed detecting impending decomposition of the reactor. Sharmin et al. (2008) also developed a data-driven model to detect failures in LDPE reactors. The methodology coupled the global energy balance terms (reaction, products and reactants enthalpy), calculated using the industrial data, and PCA method to detect decomposition in the reactor. The heat released by the reaction was obtained using the overall conversion in the reactor. The model could detect decomposition with a lead-time about 30 seconds.

The main objective of the present work was the development of a fault detection algorithm to detect failures in high-pressure pumps used in the autoclave reactors of LDPE polymerization. Particularly, the modeling approach presented here takes into account the initiator injection, which is known to be a potential cause of decomposition. Besides that, the heat released by the reaction was accounted using the definition from the free-radical polymerization. The fault detection tool was developed using plant data from an industrial autoclave reactor unit and was tested against pump leakages events that happened between 2019 and 2022.

## 2. Methodology

The methodology consisted in developing data driven models coupled with the energy balance to monitor the process. In order to do that, the following models were developed: (i) charts of the residuals for the energy balance model; (ii) control charts of a process KPI and; (iii) a PCA model using Hotelling’s T-squared statistic using the energy balance as input.

### 2.1. Historical data for calibration and validation

The fault detection tool was developed using plant data from an industrial autoclave reactor unit and was tested against pump failure events that happened between 2019 and 2022. The reactor has three pumps that inject initiator at different reactor heights and has monomer feeds along the reactor. The process variables collected to monitor the energy balance were: monomer feed mass flow rates, monomer feed temperature, initiator mass flow rate and temperature along the reaction height, with 1 second sampling time. Another process variable collected specifically related to the pump operation is the “pump output” variable, that is an indirect measurement of the number of pulses of the pump piston, also with 1 second sampling time. The dataset was cleaned, and data recorded during start-up, shut downs, and containing corrupted values were removed, and the cleaned dataset was

divided in two sets (calibration and validation). The calibration set contains the cleaned base without the 2h before and after the failure in each pump.

### 2.2. Charts of the residuals for the energy balance error model

The energy balance for each zone of the continuous autoclave reactor at steady-state operation and neglecting the heat transfer to the surroundings and the shaft work can be written as:

$$\dot{m}_{in}c_pT_{in} + \dot{m}_{feed}c_pT_{feed} - \dot{m}_{out}c_pT_{out} + R_p(-\Delta H_p)V = 0 \quad (1)$$

where,  $\dot{m}_{in}$ ,  $\dot{m}_{feed}$  and  $\dot{m}_{out}$  are the mass flow rate entering the zone (in), of the fresh monomer entering the zone (feed) and the feed leaving the zone (out);  $c_p$  is the specific heat capacity of the mixture;  $T_{in}$ ,  $T_{feed}$  and  $T_{out}$  are the temperature of the feed entering the zone (in), of the fresh monomer (feed) and the feed leaving the zone (out);  $R_p$  is the polymerization rate;  $-\Delta H_p$  is the polymerization enthalpy;  $V$  is the zone volume.

The polymerization rate,  $R_p$  can be obtained from the free-radical polymerization kinetics, assuming pseudo steady-state for the radicals:

$$R_p = k_p[M] \sqrt{\frac{2fk_i[I]}{k_t}} \quad (2)$$

where,  $k_p$ ,  $k_i$  and  $k_t$  are the propagation, initiation and termination rate constants;  $f$  the initiator efficiency;  $[M]$  the monomer concentration; and  $[I]$  the initiator concentration. Accounting the polymerization rate as a function of initiator concentration allows the model to be sensitive to changes in the initiator feed, which is associated with failures in the injection pumps. The kinetic rate constants used in the model are from the work of Zhang et al. (1996). During normal operation, the energy balance should be equal to zero at all times within reasonable limits (Sharmin et al., 2008), and when failures occurs in the pumps, the energy balance error will exceed the normal threshold limits, indicating that the system is moving away from the normal condition operation. The error balance in reactor zone  $i$  is:

$$Error_i = \dot{m}_{in}c_pT_{in} + \dot{m}_{feed}c_pT_{feed} - \dot{m}_{out}c_pT_{out} + R_p(-\Delta H_p)V \quad (3)$$

The error in the energy balance was calculated for the normal operation (calibration set) and can be applied to each reactor zone sequentially. The upper and lower control limits (UCL and LCL, respectively) for the error in each reactor zone are calculated from the average ( $\overline{Error}_i$ ) and standard deviation ( $\sigma_i$ ) of the in control dataset.

$$UCL_i/LCL_i = \overline{Error}_i \pm L\sigma_i \quad (4)$$

where  $L$  is the control limit width, typically,  $L = 3$ .

### 2.3. Monitor a process KPI

A KPI can be used to monitor the failures, using the process variables “pump output” and the temperature of the reactor zone in which the pump is injecting the initiator.

$$KPI_i = \frac{\left( \frac{T_{zone\ i}}{T_{zone\ i}} \right)}{\left( \frac{Output\ pump\ i}{Output\ pump\ i} \right)} \quad (5)$$

This KPI is normalized and sensitive to the failures in the individual pumps, once when the pump is leaking, less peroxide is entering the reactor zone and, as consequence, the reactor zone temperature decreases and the “pump output” increases, because of the temperature reactor control policy. During normal operation, the KPI should vary within the normal limits, and when failures in the pumps occur, the KPI will exceed the normal threshold limits, indicating that the system is moving away from the normal condition. The UCL and LCL for the KPI were calculated in the same way as the energy balance error.

#### 2.4. PCA model

PCA is a multivariate statistic technique commonly used to statistic quality control and fault detection. This method is ideal for handling a large number of highly correlated data and noisy process variables. PCA decomposes the observation matrix  $\mathbf{Y}$  (process data) in a linear combination:

$$\mathbf{Y} = \mathbf{TP}^T = \sum_{i=1}^q t_i p_i^T \quad (6)$$

where,  $t_i$  are the principal component score,  $p_i$  are the loading vector (eigenvalues of the covariance matrix of  $\mathbf{Y}$ ) and  $q$  is the number of measured quality variables. Usually, most of the variability in the data is retained by the first principal components (MacGregor and Kourti, 1995). The data matrix used for the PCA model contains the scaled the elements of the energy balance (Equation 3). In order to build the quality control charts based on PCA, the model was developed using historical data from steady state, when only common cause variation was present. For new measurements, the observations can be projected using the loading vectors to obtain the scores,  $t_{i,new}$ . The Hotelling's  $T^2$  is used for the multivariate control charts. The upper control limit (UCL) for  $T^2$  is obtained according to MacGregor and Kourti (1995).

### 3. Results

#### 3.1. Charts of the residuals for the energy balance error model

The energy balance error was evaluated for the calibration dataset, according to Equation 3, allowing the determination of the UCL and LCL of the control charts for each one of the three pumps present in the process. Table 1 shows the percentage of events detected, lead-time and percentage of false alarms for the model.

The lead-time is the warning time of the model before the event occur, although the average lead-time is presented in the results of Table 1, there is a distribution of warning times for the events, as depicted in the histogram of lead-time in Figure 1. Some events can be detected between 5-35 minutes before the failure, and others can be detected between 95-125 minutes before the event. The broad range of warning times is related to the different failures modes present in these pumps, as leaks, cylinder failures, reversal system failure and check valve failures. Depending on the failure mode, the fault can happen very fast, reducing the model warning time. Then, the broad warning time is associated to the inherent nature of the failures.

#### 3.2. Monitor a process KPI

The percentage of events detected, lead-time and percentage of false alarms for the model using the process KPI (Equation 5) are presented in Table 1. Compared to the energy balance error model (Table 1), the KPI monitoring has smaller warning times, once the KPI model depends only on the actual value of the process variables and is not able to represent an energy accumulation in the process, as the energy balance is. Despite of that,

the KPI model can detect more events than the energy balance error for the middle pump. Figure 2 presents the control chart for the KPI during a failure event; it is possible to observe the KPI value moving away from the control limits around 1 hour before the failure. After the pump maintenance, the KPI value returns to the normal operation limits.

Table 1: Percentage of events detected, lead-time (min) and percentage of false alarms for the models.

Method	Pump	Events detected (%)	Average lead-time (min)	False alarm (%)
Energy balance error model (L=3)	Top	42.86%	58.33	9.10%
	Middle	40.00%	13.75	5.56%
	Bottom	88.89%	6.88	6.79%
KPI model (L=3)	Top	42.9%	29.00	5.80%
	Middle	80.0%	8.38	7.70%
	Bottom	55.6%	13.40	4.50%
PCA model (alpha = 0.01)	Top	100.0%	51.43	17.35%
	Middle	70.0%	34.29	8.84%
	Bottom	88.9%	30.00	12.01%

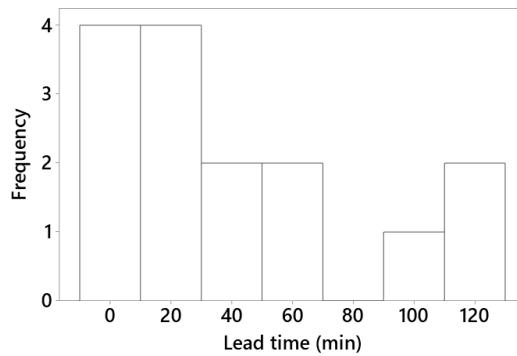


Figure 1: Histogram of lead-time (min) for the failures events for the energy balance error model.

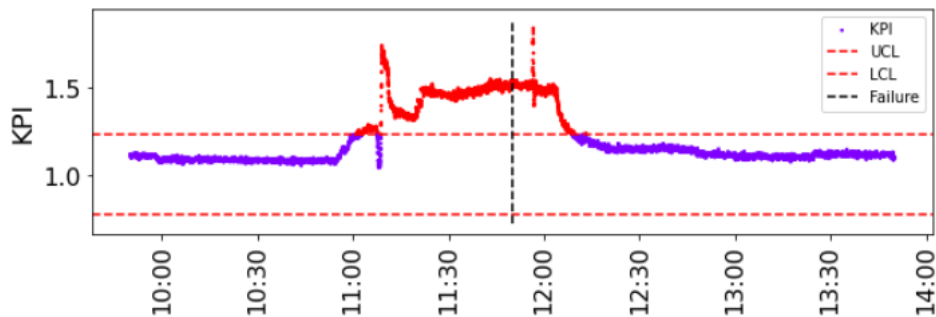


Figure 2: Control chart for the normalized KPI during a failure event. KPI (.) UCL/LCL (red --) and failure event (black --).



### 3.3. PCA model

Table 1 presents PCA model performance, using the energy balance terms as input. The results show that combining the energy balance model with the multivariate statistic approach of PCA provides a fault detection algorithm that can detect more events with increased warning times for the failures in the peroxide pumps. The  $T^2$  control chart in Figure 3 shows the validation data for a failure event in one of the pumps.

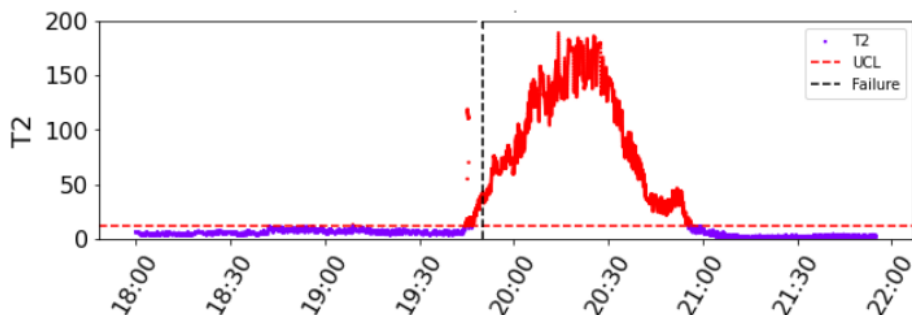


Figure 3: Control chart for the  $T^2$  statistic during a failure event.  $T^2$  (.) UCL (red --) and failure event (black --).

## 4. Conclusion

This work presented the study of fault detection models for LDPE autoclave reactors peroxide pumps. The failures in initiator pumps can lead to the explosive decomposition of ethylene and are critical for process control and safety. Data-driven models combined with phenomenological knowledge of the system play an important role to improve fault detection algorithms, as showed in this paper. The model developed was tested against failures events that happened between 2019 and 2022, with at least 70% of the events successfully detected between 5 and 50 minutes prior to the failure. The development of tools that can support the operation of this process is still a challenge, due to the non-linear and fast nature of the process.

## References

- Hamielc, A. E., & Tobita, H. (2000). Polymerization processes. Ullmann's Encyclopedia of Industrial Chemistry.
- MacGregor, J. F., & Kourti, T. (1995). Statistical process control of multivariate processes. Control engineering practice, 3(3), 403-414.
- Sharmin, R., Shah, S. L., & Sundararaj, U. (2008). A PCA based fault detection scheme for an industrial high pressure polyethylene reactor. Macromolecular Reaction Engineering, 2(1), 12-30.
- Sivalingam, G., Soni, N. J., & Vakil, S. M. (2015). Detection of decomposition for high pressure ethylene/vinyl acetate copolymerization in autoclave reactor using principal component analysis on heat balance model. The Canadian Journal of Chemical Engineering, 93(6), 1063-1075.
- Zhang, S. X., Read, N. K., & Ray, W. H. (1996). Runaway phenomena in low-density polyethylene autoclave reactors. AIChE journal, 42(10), 2911-2925.

# Using ultrasonic and process data to identify surging and gas lock phenomena during ESP operation

Tiago F. Souza<sup>a</sup>, Cáio C. S. Araújo<sup>a</sup>, Maurício M. F. Figueiredo<sup>a</sup>, Ana M. F. Fileti<sup>a</sup>

<sup>a</sup> *School of Chemical Engineering, University of Campinas - UNICAMP, Albert Einstein Avenue, 500 – Campinas/SP, 13083-852, Brazil*  
Corresponding author: [t226161@dac.unicamp.br](mailto:t226161@dac.unicamp.br)

## Abstract

Electrical Submersible Pump (ESP) has deteriorated performance due to free gas in its intake, which may result in two undesirable phenomena: surging and gas lock. In this sense, this work investigates and determines features set — based on ultrasonic and process data — to identify these undesirable phenomena. The experimental data was obtained with a fixed ESP rotational speed (3500 rpm), varying the inlet water flowrate (1 to 60 kg/min) and air flowrate (0.02 to 0.1 kg/h). The ultrasonic measurements were carried out using a 2.25 MHz transducer. Using ultrasonic echo energy and time of flight signals, and the pump discharge pressure, it was possible partitioning the experimental data into two sets using *K*-means clustering tool. One of these groups is mainly composed of ESP operating points in surging or gas lock conditions, and another predominantly consists of ESP stable conditions.

**Keywords:** Electrical Submersible Pump; surging; gas lock; ultrasonic; clustering.

## 1. Introduction

Electrical Submersible Pump (ESP) is widely applied to lift oil from offshore wells to production facilities (Takacs, 2017). The ESP performance is negatively affected due to the presence of free gas in pump intake. The free gas in ESP intake can result in two undesirable phenomena: surging and gas lock (Verde, 2016). Surging occurs when the pump head decreases as the liquid flowrate decreases. In turn, gas lock occurs when the liquid pumping stops due to blockage of free gas at ESP intake. Operating the ESP in surging or gas lock conditions can lead to severe instabilities, such as abrupt shutdowns due to fluctuations in the load of the motor coupled to the ESP.

Experimental investigations have been developed to predict the operating conditions that results in surging or gas lock during ESP operation under liquid-gas flow. Such investigations resulted in empirical or semi-empirical models to identify the ESP operating points in which surging or gas lock occurs (Turpin; Lean; Bearden, 1986; Zhu, 2017). However, these models require input variables not measured in real-time or continuously in individual oil wells equipped with ESP, such as liquid flowrate, gas flowrate, and free gas in ESP inlet. In order to develop an alternative technique, Oliveira (2014) carried out an experimental investigation whose results demonstrate the potential of ultrasonic measurement technique to identify the surging during ESP operation under liquid-gas flow.

In this context, this work aims to investigate and determine features set, based on ultrasonic and process data, capable of identifying the unstable operating conditions —

surging or gas lock — and stable operating conditions during ESP operation under liquid-gas flow.

## Material and methods

### 1.1. Test plant

The test plant was built to acquire ultrasonic and process data during ESP operation. It was composed of three flowlines: single-phase liquid flow line, single-phase gas flowline and two-phase liquid-gas flow line. An illustrative diagram of the setup is shown in Figure 1. The liquid and liquid-gas flow line are composed of a 23 mm inner diameter ASTM 240 – AISI 304 stainless steel. The gas flow line is a polyurethane tube of 8 mm inner diameter.

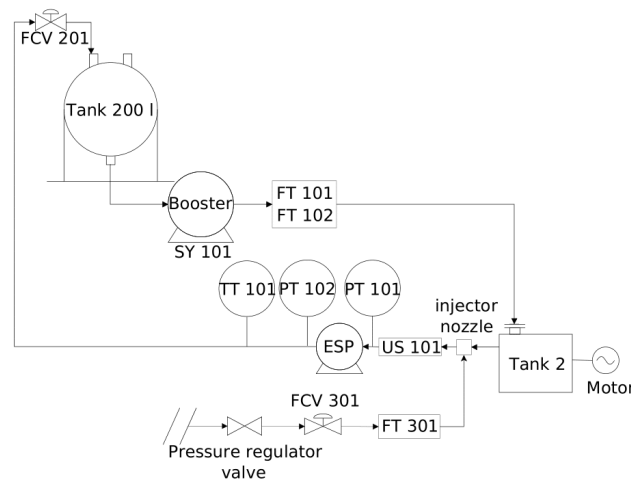


Figure 1 – Experimental setup diagram.

The liquid flow line begins at the 200 liters separator tank (Tank 200 l) which has the function of separating the gas and liquid phases from the two-phase mixture and feeds liquid to the booster pump. The booster, a centrifugal pump (KSB / Megablock 40-25-160 GG), was driven by 3 HP motor that was controlled by an inverter frequency drive (SY-101). The booster kept constant the pressure at the ESP intake (Schlumberger / D475N). After the booster, a Coriolis mass flowmeter (FT-101/FT-102) was installed to measure the liquid mass flow and liquid density. The ESP was driven by a 1 HP electric motor (WEG / W22 IR3 PREMIUM).

The gas flow line was fed by an external air supply system. The pressure regulator valve kept at the 5 bar g the air pressure upstream the control valve (FCV 301), which controlled the amount of air in ESP intake. To measure the gas mass flowrate, a Coriolis mass flowmeter (FT 301) was installed. The gas flow line ends at the injector nozzle, where the liquid-gas flow line begins.

The injector nozzle was installed at 35 cm upstream of the ESP intake. A porous medium was used as the injector nozzle to ensure uniform distribution of the gas in form of bubbles into the liquid flow. After the injector nozzle, an annular flow occurs with the outside inner diameter of 52 mm and the shaft diameter of 20 mm.

In the horizontal position, the inlet and outlet flow conditions of the ESP was monitored by one temperature transmitter (TT 101), one ultrasound structure (US 101) and two pressure sensors (PT 101 and PT 102). At the US 101, the ultrasound transducer was fixed

in a non-invasive and non-intrusive way. Besides, the FCV 201 was placed after the ESP to regulate the flowrate of liquid-gas flow. A sample rate of 4 Hz was used to acquire the process data.

### 1.2. Description of ultrasonic measurement technique

The ultrasonic measurement system consists of an ultrasound support (US 101), an ultrasonic transducer, and a device for generating and digitizing ultrasonic signals. The function of the US 101 support is to accommodate the ultrasonic transducers orthogonally to the flow in a non-invasive and non-intrusive way. This structure was made using polyacetal material, internal diameter of 52 mm, 10 mm delay line and a hole to accommodate transducers positioned perpendicular to the flow. Industrial coupling gel was used to perform the coupling of the ultrasonic transducer with the inner wall of the ultrasonic support. The ultrasonic measurements were performed using only one piezoelectric ultrasound transducer (Olympus - V106-SM) with a central frequency of 2.25 MHz.

The ultrasonic signal generation and digitization device (US Ultratek - DSPUT5000-2) was used in pulse-echo mode, that is, the system emit and capture waves through the same transducer. The device has 8-bit vertical resolution and a maximum sampling rate of 100 MHz. The parameters of the ultrasonic acquisition device were: pulse voltage of -300 V, a pulse width of 222 ns; sample rate of 25 MHz, a pulse repetition frequency of 1 kHz, a digital gain of 40 dB; damping of 620  $\Omega$ , an acquisition time of 15 s; and 720 points in depth per wave. Such values were defined according to the fundamentals of signal sampling (Holman, 2012), and the technical guidelines of the manufacturers of the ultrasonic acquisition board and ultrasonic transducers.

The Figure 2-a represents a schematic mode of the ultrasonic wave propagation through solid and liquid based medium. At the wall-flow interface, part of the wave intensity associated with the initial pulse ( $P_0$ ) is reflected in the first wall pulse reflection ( $R_{WF}$ ). The remainder is transmitted to the liquid based medium. After that, the transmitted pulse will propagate until reaches the liquid-gas ( $R_{LG}$ ) or shaft-flow ( $R_{SF}$ ) interface, in which the wave will be reflected to the ultrasound. A typical ultrasonic echo signal is illustrated in the Figure 2-b. The position of the liquid-gas ( $R_{LG}$ ), shaft-flow ( $R_{SF}$ ) interface and first wall pulse reflection ( $R_{WF}$ ) is indicated.

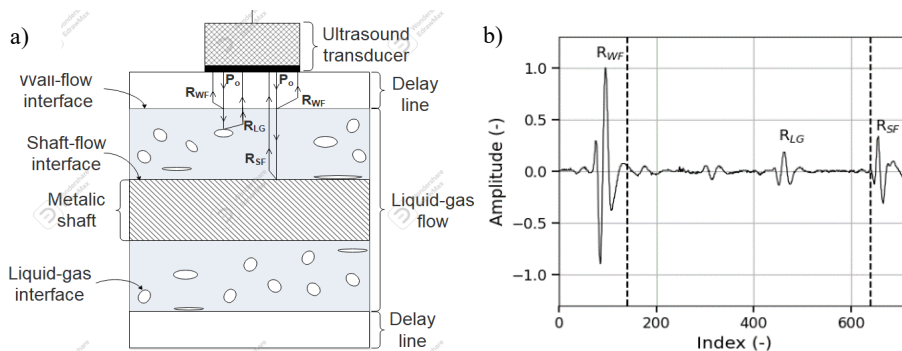


Figure 2 – Schematic ultrasonic mode of the wave propagation and reflection at the wall and two-phase flow (a). Echo signal containing two-phase flow reflection (b).

### 1.3. Experimental design and procedure

The test matrix has 9 experimental conditions for single-phase, and 26, 23 and 20 experimental conditions for two-phase with gas flowrate of 0.02 kg/h, 0.06 kg/h and 0.10

kg/h, respectively. Each experimental condition was experimentally measured in 4 replicates to evaluate the reproducibility, totalling 312 sample acquisitions. Each sample acquisition lasted 30 s for the process data under constant ESP rotation (3500 RPM) and intake pressure (100 kPa). The gas mass flowrate ( $m_G$ ) were experimentally adopted according to the ESP capacity to deal with the total amount of free gas in its intake. The experiments were carried out with water as the continuous phase and air as the dispersed phase. During the experiments, the temperature remained in the range of 24 to 26 °C.

In summary, the procedure adopted for the execution of the matrix was: a) The first acquisition was obtained from the condition in which the intake pressure was 100 kPa and the pump head was equal to zero, which corresponds to the maximum liquid flowrate in that condition; ii) afterwards, the variables of interest corresponding to the next acquisition point were adjusted, waiting 2 minutes to stabilize the flux in the new condition; iii) ultrasonic measurements and process data were acquired; iv) the new experimental condition (acquisition) was adjusted, and ultrasonic and process data were collected until the experimental points of the test matrix finished.

## 2. Results and discussion

### 2.1. ESP performance curves

Figure 3 shows the ESP performance curves under single-phase and two-phase flow conditions investigated experimentally in this work. The ESP operation under two-phase flow conditions can exhibit two phenomena: surging and gas lock. The surging phenomenon implies an abnormal behavior, the pump head decreases as the liquid flowrate decreases. The beginning of surging (surging point) corresponds to the curve peak in the ESP performance curve under two-phase flow. It is noticeable that as the gas flowrate increases, the stable region — the right side of the curve peak — became narrower due to the less capacity of the ESP to deal with free gas in its intake. The gas lock occurs when the pump head is approximately null after beginning of surging. In face of that, we can divide the two-phase performance curves in two regions: stable and unstable.

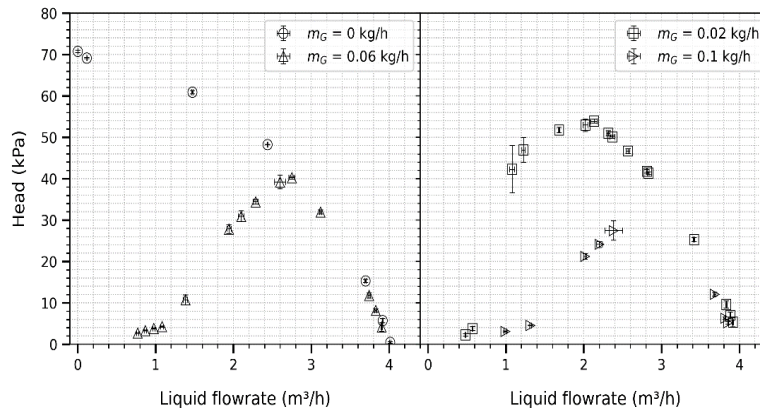


Figure 3 – ESP performance curves under single-phase and two-phase flow.

### 2.2. Echoes signals and process data - Feature extraction

The features used to cluster the ESP operating points were obtained by processing the ultrasonic echoes signals and pump discharge pressure data. Four features were extracted: arithmetic mean of echoes energy signals from ultrasonic wave reflections in liquid-gas

interface ( $\bar{E}_{LG}$ ); arithmetic mean of echoes energy signals from ultrasonic wave reflections in liquid-shaft interface ( $\bar{E}_{SF}$ ); time of flight standard deviation from ultrasonic wave reflections in liquid-gas interface ( $t_{LG}$ ); and variation coefficient of ESP pressure discharge ( $\bar{P}_D$ ). The features mathematical definitions are available in Table 1. These features were normalized between -1 and 1 before applying the  $K$ -means clustering.

Table 1 – Features definitions

$\bar{E}_{LG}$	$\bar{E}_{SF}$
$mean \left[ \left( \sum_{k=140}^{640} (A_k)^2 \right)_i \right] \quad (1)$ $i = 1, 2, \dots, 15000$	$mean \left[ \left( \sum_{k=640}^{720} (A_k)^2 \right)_i \right] \quad (2)$ $i = 1, 2, \dots, 15000$
$t_{LG}$	$\bar{P}_D$
$std[\text{argmax}(A_{120:640})_i] \quad (3)$ $i = 1, 2, \dots, 15000$	$std[(P_D)_i] / mean[(P_D)_i] \quad (4)$ $i = 1, 2, \dots, 120$

### 2.3. $K$ -means clustering

To execute  $K$ -means clustering, the method  $k$ -means++ was adopted for handling centroids initialization (Guha *et al.*, 2003). In addition, two metrics were used to assess the clustering performance: silhouette coefficient and rand index (Rousseeuw, 1987; Hubert; Arabie, 1985). Figure 4-a shows these metrics score versus the number of clusters due to  $K$ -means application to ESP experimental data. Facing these metrics scores, it is reasonable to consider clustering the data in 2 generated clusters due to the highest values of the rand index and silhouette coefficient. Figure 4-b indicates the ESP experimental data distribution into 2 clusters. The amount of 279 samples data was partitioned into two clusters. Cluster 1 has grouped 138 samples, and Cluster 2 grouped has 141 samples. There are samples of data obtained experimentally at the same operating condition (acquisition) partitioned in different clusters. These samples are associated with acquisitions 14, 15, 16, and equal to 0.02 kg/h; acquisition 10 and equal to 0.06 kg/h; acquisition 6 and equal to 0.1 kg/h. These acquisitions in performance curves are identical or closer to surging points. Despite having samples associated with identical acquisitions in different clusters, it is notable that Cluster 1 is predominantly composed of ESP experimental data in stable conditions, and Cluster 2 is majorly formed by ESP experimental data in unstable conditions. These results indicate that features used in  $k$ -means clustering can be used to identify if the ESP is operating or not in surging or gas lock conditions (unstable conditions).

### 3. Conclusions

This work suggests that is possible identifying surging or gas lock phenomena (unstable conditions) in ESP operation under two-phase flow by using the following features extracted from ultrasonic echoes signals and process data:  $\bar{E}_{LG}$ ;  $\bar{E}_{SF}$ ;  $t_{LG}$ ; and  $\bar{P}_D$ .  $K$ -means clustering was used to cluster the ESP experimental data by using the features previously mentioned. According to silhouette coefficient and rand index scores, it is reasonable to cluster the ESP experimental data in 2 clusters. The generated Cluster 1 is predominantly composed of ESP experimental data in stable conditions, and Cluster 2 is majorly formed by ESP experimental data in unstable conditions. The silhouette coefficient and rand

index scores for this clustering were 0.5 and 0.93, respectively. Thus, these results indicate that features used in  $K$ -means clustering can be used to identify the two undesirable phenomena during ESP operation.

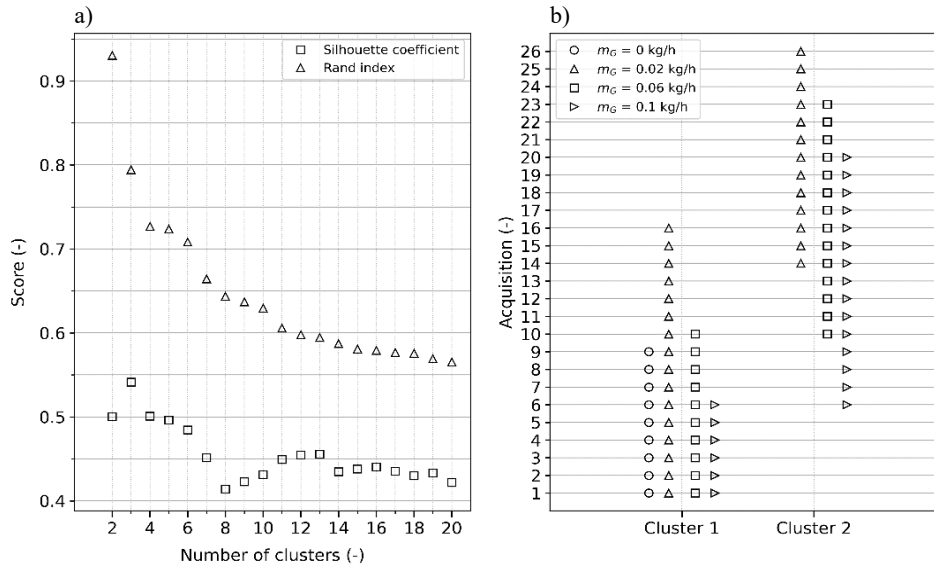


Figure 4 - Number of clusters analysis based on metrics score (a) and visual representation of the ESP experimental data grouped in only two clusters (b).

## References

- W. M. Verde et al., 2017. Experimental study of gas-liquid two-phase flow patterns within centrifugal pumps impellers, **Experimental Thermal and Fluid Science**, v. 85, p. 37-51.
- G. Takacs, 2017. **Electrical submersible pumps manual: design, operations, and maintenance**, 2. ed. Gulf professional publishing.
- J. P. Holman. **Experimental Methods for Engineers**. 8. ed. McGraw-Hill, 2012.
- J. Zhu, X. Guo, F. Liang, H. Zhang. Experimental Study and Mechanistic Modeling of Pressure Surging in Electrical Submersible Pump. **Journal of Natural Gas Science and Engineering**, v. 45, p. 625-636, 2017.
- J. L. Turpin, J.F. Lea, J.L. Bearden. Gas-Liquid Flow Through Centrifugal Pumps-Correlation of Data. In: **Proceedings of the 3rd International Pump Symposium**, p. 13-20, 1986.
- B. D. Oliveira. **Monitoramento do Desempenho de uma Bomba Centrífuga Submersa com Auxílio da Técnica Ultrassônica**, 2014.
- T. F. Souza. **Desenvolvimento de Uma Técnica Ultrassônica para a Estimativa da Fração de Vazio e Vazão da Fase Dispersa de Escoamentos Água-Ar**, 2021.
- L. Hubert, P. Arabie. Comparing partitions. **Journal of classification**, v. 2, n. 1, p. 193-218, 1985.
- J. P. Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. **Journal of computational and applied mathematics**, v. 20, p. 53-65, 1987.
- S. Guha, A. Meyerson, N. Mishra, R. Motwani, O'Callaghan, L. Clustering data streams: Theory and practice. **IEEE transactions on knowledge and data engineering**, v. 15, n. 3, p. 515-528, 2003.

# A data-driven predictive model of gas flow rate from ultrasound and process variables using artificial neural networks

Cáio C. S. Araújo<sup>a</sup>, Tiago F. Souza<sup>a</sup>, Maurício M. F. Figueiredo<sup>a</sup>, Ana M. F. Fileti<sup>a</sup>

<sup>a</sup> *School of Chemical Engineering, University of Campinas - UNICAMP, Albert Einstein. Avenue, 500 – Campinas/SP, 13083-852, Brazil*  
*Corresponding author: c228076@dac.unicamp.br*

## Abstract

The bubble flow is a flow regime present in various fields. A usual equipment that operates with the bubble flow is the bubble column. Many papers deal with determining flow regime, hold up, and mass transfer coefficient in a bubble column. However, little attention has been given to the measurement of the gas flowrate. In this context, this work presents a technique to predict the gas flowrate in a bubble column from ultrasound signals using a new set of ultrasonic variables combined with process variables using a feedforward artificial neural network (ANN). Experiments were carried out in a 2-meters-long-bubble column, with inner diameter of 52.5 mm using a single 2.25 MHz ultrasound transducer. The gas flow rate varied in the range of 0 to 4 liters per minute. The best architecture obtained for ANN has one hidden layer with 36 neurons. The test mean absolute error was 0,1 LPM.

**Keywords:** flow rate, artificial neural networks, monitoring, ultrasound, soft computing.

## 1. Introduction

The bubble flow is a flow regime present in various fields, like the pharmaceutical, biochemical, and petrochemical industries (Shu et al., 2019). A usual equipment presents in these industries, that operates with the bubble flow, is the bubble column. A bubble column can be used as a reactor that operates with a stagnant liquid phase, in which the gas flows throughout. Many papers deal with determining flow regime, bubble size, liquid hold up, and mass transfer coefficient in a bubble column using machine learning (Biessey et al., 2021; Serraa et al., 2020; Supardan, 2003). However, little attention has been given to the measurement of the gas phase flowrate in a bubble column direct from the combination of ultrasonic and process signals without intermediate needed analysis, like calculation of velocity and liquid hold up. In addition, according to the findings of Mudde & Saito (2001), it is expected that the method developed here using bubble columns can be applied in two-phase liquid-gas bubble flow and get similar results. In this context, this work proposed a new set of ultrasonic variables combined with pressure drop for gas flowrate prediction in a bubble flow using a 52.5 mm inner diameter vertical bubble column based on feedforward artificial neural network. The ultrasonic signals were acquired using just one ultrasound transducer.



## 2. Material and methods

### 2.1. Experimental setup

The experimental setup consists of a pipe with an internal diameter of 52.5 mm and a length of 2 m arranged vertically (Figure 1). The two-phase line was instrumented with two pressure sensors (PT 20 and PT 21) to measure the void fraction of the dispersed phase and a temperature sensor (TT 20) to obtain the physical properties of the two-phase mixture. Ultrasonic measurements were performed with a single inclined at  $10^\circ$  with the flow, located at support (UT 10).

The air flow line contains a pressure regulator, a rotameter (FI 10), a check valve, a pressure sensor (PT 10) and a temperature sensor (TT 10). The gas injection was performed through a porous structure positioned at the base of the bubble column (gas injection).

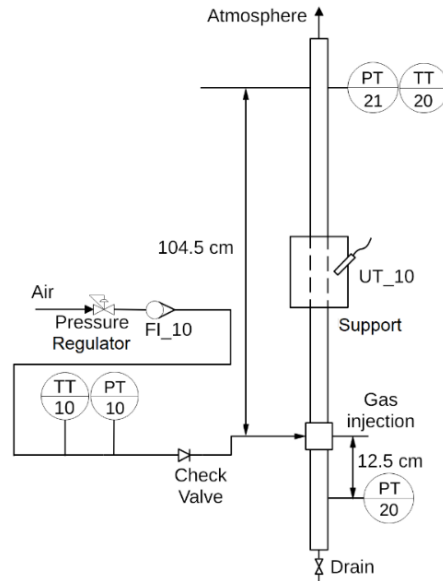


Figure 1 – P&ID diagram of the experimental setup.

### 2.2. Ultrasonic instrumentation

The ultrasonic measurement system basically comprises an ultrasonic transducer and a ultrasound device (US Ultratek/DSPUT5000-2) for generating, receiving and digitalizing ultrasonic signals (Figure 2). In all experiments, the transducer remained confined between the support and a spring, which is pressed by a cable gland. Industrial Vaseline was used to perform the acoustic coupling between the transducer and the ultrasonic spool. The support, made with Delrin polymeric material, was installed 60 cm above the air injection point. During the experiments, the transducer was installed with inclination  $\theta = 10^\circ$ . Besides, the transducer (Panametrics/ M106-SM) with a central frequency of 2.25 MHz was used.

The ultrasound device with 8 bits of vertical resolution was used for generation, reception and digitization of ultrasonic signals. The sampling frequency (50 MHz) was defined in order to respect Nyquist's theorem, which defines the minimum frequency as being at least twice the highest frequency to be observed. The damping (620  $\Omega$ ), along with the gain (60 dB), will be experimentally defined in order to return the highest amplitude values without signal saturation. The pulse frequency (2 kHz) was defined based on the ability of the system to detect the same bubble at least twice throughout the ultrasonic

measurement volume. The pulse width (222 ns) was defined based on the manual of manufacturer, based on the central frequency of the transducer. Furthermore, a transducer supply voltage of  $-300 V_{P-P}$  was used.

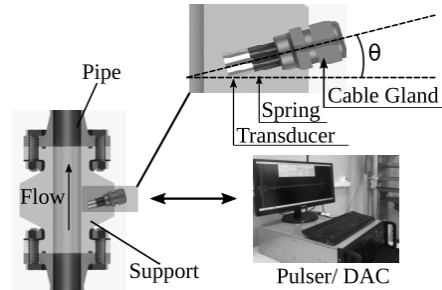


Figure 2 - Components of the ultrasonic measurement system.

### 2.3. Test matrix and procedure

The test matrix was defined based on the flow pattern map of Shah et al. (1982), considering only the bubble flow pattern in all experiments. In this way, the minimum and maximum flow limit was defined from 0 to 4 liters per minute (LPM). Experiments were carried out varying the flowrate from 0 to 4 LPM with increments of 0.5 LPM, that is, 9 measurements. Each measurement was repeated 4 times, that is, 36 measurements. The experiments were carried out with water as the continuous phase and air as the dispersed phase. The mixture temperature of the experiments was maintained in the range of 20 to 22 °C.

In each experimental condition, 15 s of ultrasonic measurements and 120 s of process data at 1 Hz of acquisition frequency were acquired. All ultrasonic echoes had 1,000 points in depth, the necessary to capture half cross-section area of the pipe, resulting in one ultrasonic echo signal per experimental measurement. In summary, the experimental procedure adopted was: i) Fill the vertical pipe with tap water until it reaches 1.2 m above the gas injection point and keep it constant; ii) define the pre-defined gas flowrate randomly; iii) wait 60 s for flow stabilization; iv) acquire ultrasonic and process data; v) return to the ii step until the final of the test matrix.

### 2.4. Topology of artificial neural network

The ANN developed and used to predict the gas flow rate was developed in Python programming using the scikit-learn library. The ANN was chosen because it has the ability map non-linear relationships between inputs and outputs. In the training step, the Adam optimization algorithm was used. The ANN input ultrasonic variables used were: mean energy signal ( $\bar{E}$ ); standard deviation of the energy signal ( $\sigma_E$ ); mean transit time ( $\bar{ToF}$ ); standard deviation of transit time ( $\sigma_{ToF}$ ). The ANN input process variable used was the mean pressure drop ( $\Delta P$ ). The ultrasonic echo signal (matrix of 30,000 rows by 1,000 columns) was processed to return energy ( $E$ ) and transit time ( $ToF$ ) signals in which were calculated the associated mean and standard deviation, according to definitions presented in Table 1. Indeed, the  $\bar{E}$ ,  $\sigma_E$ ,  $\bar{ToF}$ ,  $\sigma_{ToF}$  and  $\Delta P$  was used to feed the ANN.

The output variable was the gas flow rate. The mathematical definition of  $E$  and  $ToF$  signal associated with echoes signals are presented in Table 1. The dataset was randomly divided in two groups: i) 75% for training; and ii) 25 % for test step. In training step, 10 % of the training dataset was used for validation. The best architecture was defined using GridSearchCV optimization method (Pedregosa et al., 2011). It was tested the number of

neurons on hidden layer in the range of 1 to 50 and identity, relu, hyperbolic tangent and logistic as activation functions. Besides, the performance of the artificial neural network was evaluated using the mean square error (MSE), mean absolute error (MAE), and determination Coefficient ( $R^2$ ) of the straight line that fits predicted versus actual values of gas flow metrics.

Table 1 - Definition of energy and transit time signals.

$E_i = \sum_{j=1}^n [e_{(i,j)}]^2$	$ToF_i = \text{argmax}_{e_{(i,j)}flow} - \text{argmax}_{e_{(i,j)}wall}$
$E = [E_1, E_1, \dots, E_m]$	$ToF = [ToF_1, ToF_2, \dots, ToF_m]$

Where:  $e_{(i,j)}$  is the value of echo in row  $i$  and column  $j$  with  $m$  rows and  $n$  columns;  $E_i$  is the energy associated with row  $i$ ;  $ToF_i$  is the transit time associated with row  $i$ ; and  $\text{argmax}_{e_{(i,j)}flow}$  and  $\text{argmax}_{e_{(i,j)}wall}$  are the position of maximum echo amplitude value in the flow and wall region, respectively.

### 3. Results and discussion

#### 3.1. Ultrasonic echoes

The Figure 3 illustrates the ultrasonic echoes referring to the 2,000 ultrasonic pulses (1 s) emitted for the single-phase (0 LPM) and two-phase simulated bubble flows. Dividing the wall signal from the signal in the position of the dashed line, where from the begging to the dashed line is the wall region, it is possible to note the sensitivity of the transducer with the passage of bubbles in the ultrasonic measurement region. In addition, it is observed that the reflected signals from the liquid-gas interfaces of the bubbles are not limited to a specific region of half cross-section area, showing the relative uniformity distribution of the bubbles across the half cross sectional area of the pipe.

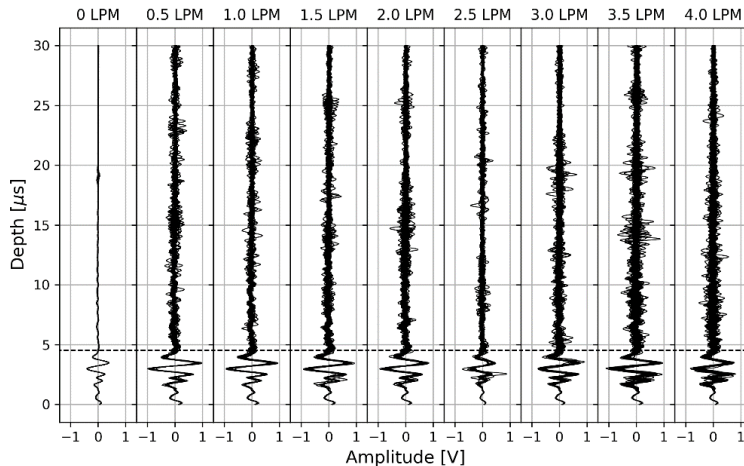


Figure 3 - Ultrasonic echoes for single-phase (0LPM) and two-phase flows in the vertical bubble flow.

The Table 2 shows the limits of values associated with ultrasonic ( $\bar{E}, \sigma_E, ToF, \sigma_{ToF}$ ) and process ( $\Delta P, Q$ ) signals. These variables were used as input, except the variable  $Q$  that was uses as output.

Table 2 - Range of the ultrasonic and process variables.

Variable	$\bar{E}$	$\sigma_E$	$\overline{ToF}$	$\sigma_{ToF}$	$\Delta P$	$Q$
Units	[V <sup>2</sup> ]	[V]	[ $\mu s$ ]	[ $\mu s$ ]	[kPa]	[LPM]
Minimum value	0.17	0.015	650.30	404.74	10.56	0
Maximum value	3.39	2.656	1018.06	495.55	12.64	4.0

3.2. ANN for prediction of gas flow rate

The ANN which presented the best prediction of gas flow rate had an architecture with 36 neurons in only one hidden layer, and rectified linear unit (relu) as activation function. The ANN predictions and absolute error for training and test set are illustrated in Figure 4 and 5, respectively. The ANN score metrics calculated for the training set are the following: MSE (0.007); MAE (0.063); and R<sup>2</sup> (1.000). In turn, the score metrics calculated for the test set are: MSE (0.013); MAE (0.078); and R<sup>2</sup> (0.998).

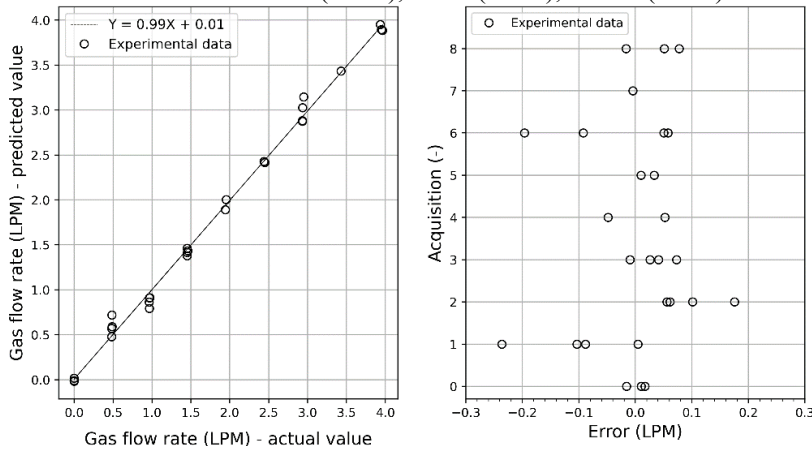


Figure 4 - ANN performance comparison for train step (left) and absolute error values between actual and prediction value of the gas flow rate for training (right).

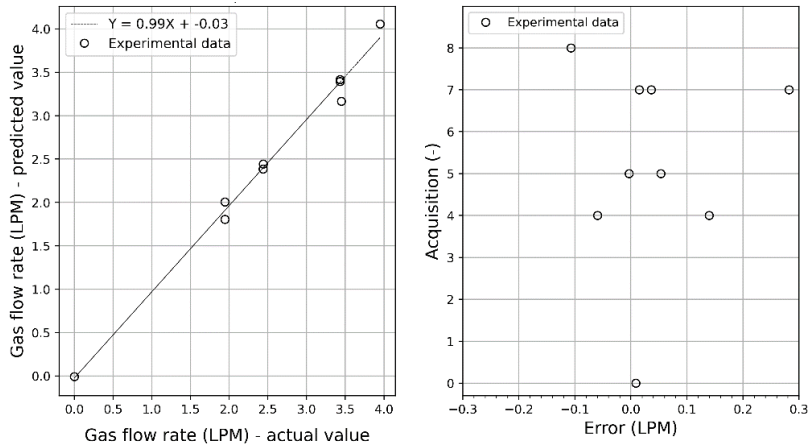


Figure 5 - ANN performance comparison for test step (left) and absolute error values between actual and prediction value of the gas flow rate for testing (right).

The results of test step, illustrated in Figure 5, showed that the distribution of the absolute error was close to the obtained in training step, which represents a good agreement between the prediction and actual gas flowrate for test and train step. As a consequence, the MAE, MSE and  $R^2$  are also in good agreement in relation to the train step results. Therefore, the metrics score relative to test set suggest a satisfactory generalization capability of ANN developed to predict gas flowrate.

The results of gas flowrate prediction also indicates that the choose ultrasonic variables ( $\bar{E}$ ,  $\sigma_E, ToF, \sigma_{ToF}$ ) together with process variable ( $\Delta P$ ) have the ability to represent the physics of the flow. The energy ( $E$ ) captures the frequency and intensity of the bubble reflection in the echoes, the transit time ( $ToF$ ) captures the location of the bubbles reflection in the ultrasonic echoes and the pressure drop ( $\Delta P$ ) infer about the concentration of the bubble in the two-phase mixture. Indeed, the physical representativity of the choose variable contribute to the good agreement in the results. Further, the good results indicate that these ultrasonic and process variables could be used together to estimate the gas flow rate in bubble columns.

#### 4. Conclusions

The results showed the ability of the ANN to predict the gas flow rate. The mean absolute error from the reference was 0.1 LPM and the majority of the error in training and test was in the range of -0.1 to 0.1 LPM. From the agreement between the predicted and actual gas flow rate, we concluded that the ANN based model proposed is promising to estimate gas flow rate in bubble flows using ultrasonic combined with process data as input. Further, the selected ultrasonic variables ( $\bar{E}$ ,  $\sigma_E, ToF, \sigma_{ToF}$ ) and the process variable ( $\Delta P$ ) showed good ability to represent the bubble flow phenomena and predict the gas flowrate. To increase the robustness of the proposed ANN based technique we proposed tests with different fluid as continuous and dispersed phase and in other pipe inner diameters.

#### References

- S. Shu, D. Vidal, F. Bertrand, J. Chaouki, 2019, Multiscale multiphase phenomena in bubble column reactors: A review, **Journal of Renewable Energy**, v. 141, p. 613-631.
- P. Biessey, H. Bayer, C. Thebeling, E. Hilbrands, and M. Grunewald, 2021, Prediction of Bubble Sizes in Bubble Columns with Machine Learning Methods. **Journal of Chemie Ingenieur Technik**, v. 93, p. 1968-1975.
- P. L. S. Serraa, P. H. F. Masottib, M. S. Rochab, D. A. de Andradeb, W. M. Torresb, R. N. de Mesquita, 2020, Two-phase flow void fraction estimation based on bubble image segmentation using Randomized Hough Transform with Neural Network (RHTN). **Journal of Progress in Nuclear Energy**, v. 118, p. 103133.
- M. D. Supardan, A. Maezawa, and S. Uchida, 2003, **Determination of Local Gas Holdup and Volumetric Mass Transfer Coefficient in a Bubble Column by Means of an Ultrasonic Method and Neural Network**. *Journal of Chemical Engineering Technology*, v. 26, p. 1080-1083.
- R. F. Mudde, T. Saito, 2001, Hydrodynamical similarities between bubble column and bubbly pipe flow, **Journal of Fluid Mechanics**, v. 437, p. 203-228.
- Y. T. Shah, B. G. Kelkar, S. P. Godbole, and W. D. Deckwer, 1982, Design parameters estimations for bubble column reactors. **AIChE Journal**, v. 28, p.353–379.
- F. Pedregosa et al, 2011, Scikit-learn: machine learning in Python. **Mach. Learn. Res Journal**.

# An NN-NMPC for Controlling a Crystallization Process in the Saturated and Undersaturated Zones

Fernando Arrais R. D. Lima<sup>a</sup>, Marcellus G. F. de Moraes<sup>b</sup>, Amaro G. Barreto<sup>a</sup>,  
Argimiro R. Secchi<sup>a,b</sup>, Maurício B. de Souza Jr.<sup>a,b</sup>

<sup>a</sup>*School of Chemistry, EPQB, Universidade Federal do Rio de Janeiro, Av. Horácio Macedo, 2030, CT, Bloco E, 21941-914, Rio de Janeiro, RJ – Brazil*

<sup>b</sup>*PEQ/COPPE – Universidade Federal do Rio de Janeiro, Av. Horácio Macedo, 2030, CT, Bloco G, G115, 21941-914, Rio de Janeiro, RJ – Brazil*  
*farrais@eq.ufrj.br*

## Abstract

The crystallization process has been applied in pharmaceutical, fine chemicals and food industries and is the main process in the pharmaceutical industry for particle formation. This process is used for separation and purification, recovering as much solute as possible and producing crystals with the desired length. In this work, a nonlinear model predictive controller based on neural networks (NN-NMPC) was proposed to control the coefficient of variation and length of crystals for the potassium sulphate batch cooling crystallization. The neural networks were designed using an optimizer to choose the best hyperparameters with simulated and experimental data. They were trained to predict the first four moments and the concentration in order to calculate the coefficient of variation and the length of crystals. The potassium sulphate crystallization process was modelled using a population balance model solved by the method of moments, considering the nucleation, growth and dissolution mechanisms. This rigorous model was used to represent the real process for the controller's performance evaluation. Differently from previous studies, the neural networks were trained considering the saturated and the undersaturated thermodynamic conditions. The controller's performance was studied for servo cases, analyzing the efficiency of the NN-NMPC to reach different set-points using the process temperature as the manipulated variable. The operation in the undersaturated and saturated conditions guarantees more degrees of freedom to control both variables (coefficient of variation and length of crystals), manipulating only the temperature. The proposed approach efficiently achieved the desired set-points. The controller also presented good performance for a case considering plant-model mismatch. The NN-NMPC performance was compared to an NMPC based on the phenomenological model, showing similar behavior with reduced computational effort, evidencing that the NN-NMPC has great potential to be applied in real time for the crystallization process.

**Keywords:** Crystallization, neural networks, NMPC, servo control, dissolution.

## 1. Introduction

The crystallization process is used for product separation and purification, generating a solid crystalline product. The regulatory requirements must be respected for an efficient crystallization process, producing crystals with the desired size distribution. Therefore, the crystallization process needs effective control to achieve this goal<sup>[1]</sup>. Another challenge of crystallization processes consists of modeling this phenomenon, which is

usually done using the population balance approach. However, this can take too much effort because of the necessity of using different models to describe the mechanisms of this process. Data-driven models, in particular, neural networks, have mainly been applied to model and control chemical processes because they can capture intrinsic nonlinearities<sup>[2]</sup>. Nevertheless, there is a lack of applications in the literature using neural networks to model the crystallization process under saturated and undersaturated conditions.

This work aims to model a crystallization process to predict the moments of the particle-size distribution and the solute concentration with neural networks. Unlike other works, this study considered data under saturated and undersaturated conditions, using only one model to predict each state variable under different conditions. Also, the current work proposes control strategies for the crystallization process different from the study of Moraes et al.<sup>[3]</sup>, which focuses on modeling it. Five neural networks were trained using experimental and simulated data, and the hyperparameters were chosen considering an optimization approach. The capacity of the neural networks was tested to predict the state variables  $n$  sampling times in the future, making good predictions for large horizons. Moreover, the proposed networks were used as an internal model in a Nonlinear Model Predictive Controller (NMPC).

## 2. Neural Networks Models

Neural networks are mathematical models that try to represent the correlation between input and output patterns. The most applied neural network in the literature consists of the feedforward network, particularly the multilayer perceptron (MLP) network. An MLP network comprises an input layer, an output layer, and one or more hidden layers<sup>[4]</sup>. In the current work, five MLP networks were developed to predict the moments and the solute concentration of a potassium sulfate batch crystallization process.

The dataset consisted of ten experimental batches in the saturated condition and five in the subsaturated condition developed by Moraes et al.<sup>[3]</sup> in LABCADS/UFRJ. In these experiments, the moments and the solute concentration were determined for different temperatures, totaling 724 patterns. Also, simulated data were generated using a population balance model for the saturated condition and a population balance model for the undersaturated state, developed by Moraes et al.<sup>[3]</sup>. Therefore, 15 batches were simulated, considering different temperatures and initial and saturation conditions, producing a total of 6800 simulated patterns. The training dataset was composed of seven experimental batches in the saturated state, four experimental batches in the subsaturated condition, and eleven simulated batches, totaling 5726 patterns. The test dataset was composed of the 1798 patterns of the simulations and experiments that were not used for the training. A sampling time of 1.25 min was adopted in the experiments and simulations, considering the same value adopted in the experiments of Moraes et al.<sup>[3]</sup>. The batches chosen to compose the training and test datasets were randomly selected, using in the test dataset experimental and simulated data for the saturated and undersaturated states with conditions not used in training.

The inputs and outputs of each NN are shown in Table 1, in which  $T$  is the process temperature,  $C/C_{eq}$  is the ratio between the solute concentration and the concentration of the solute in equilibrium,  $\mu_i$  is the  $i^{th}$ -order moment, and  $k$  is the sampling time. Different

combinations of inputs were investigated. The NNs were developed using the Scikit-learn library from Python. Furthermore, the hyperparameter of each neural network was chosen using an optimization approach from a Python library called Optuna. The optimization problem consisted of finding the activation function, number of hidden layers, number of neurons in the hidden layer, batch size, and learning rate that leads to  $R^2$  values of the test samples close to one.

**Table 1:** Inputs and outputs of the developed neural networks.

Neural Network	Inputs	Outputs
1	$T(k-1), C/C_{eq}(k-1), \mu_0(k-1),$ $T(k), C/C_{eq}(k), \mu_0(k)$	$\mu_0(k+1), \dots, \mu_0(k+n)$
2	$T(k-1), C/C_{eq}(k-1),$ $\mu_0(k-1), \mu_1(k-1),$ $T(k), C/C_{eq}(k), \mu_0(k), \mu_1(k)$	$\mu_1(k+1), \dots, \mu_1(k+n)$
3	$T(k-1), C/C_{eq}(k-1),$ $\mu_1(k-1), \mu_2(k-1),$ $T(k), C/C_{eq}(k), \mu_1(k), \mu_2(k)$	$\mu_2(k+1), \dots, \mu_2(k+n)$
4	$T(k-1), C/C_{eq}(k-1),$ $\mu_2(k-1), \mu_3(k-1),$ $T(k), C/C_{eq}(k), \mu_2(k), \mu_3(k)$	$\mu_3(k+1), \dots, \mu_3(k+n)$
5	$T(k-1), C/C_{eq}(k-1), C(k-1),$ $T(k), C/C_{eq}(k), C(k)$	$C(k+1), \dots, C(k+n)$

The hyperparameters used in the networks are presented in Table 2, and all neural networks were trained using the Adam algorithm.

**Table 2:** Hyperparameters of the neural networks.

Neural Network	Hidden Layers	Number of Neurons	Activation Function	Batch Size	Learning Rate
1	4	55; 55; 55; 55	Relu	15	Constant
2	3	45; 45; 45	Relu	5	Constant
3	4	45; 45; 45; 45	Relu	5	Constant
4	4	40; 40; 40; 40	Relu	5	Adaptive
5	1	15	Logistic	135	Constant

The  $R^2$  values for predictions two, five, ten, fifteen and twenty steps forward for the test samples are presented in Table 3. It can be observed that the neural networks could efficiently predict the moments and the solute concentration of a batch crystallization process, even for large prediction horizons. Therefore, the neural networks to predict  $\mu_0, \mu_1,$  and  $\mu_2$  were used as the internal model of an NMPC.



**Table 3:**  $R^2$  values, considering the test samples and different prediction horizons.

Neural Network	2 steps	5 steps	10 steps	15 steps	20 steps
1	0.9998	0.9996	0.9986	0.9987	0.9977
2	0.9997	0.9997	0.9989	0.9956	0.9941
3	0.9997	0.9995	0.9980	0.9963	0.9932
4	0.9998	0.9997	0.9990	0.9982	0.9970
5	0.9995	0.9987	0.9955	0.9947	0.9901

### 3. Nonlinear Model Predictive Control

The controlled variables chosen for the optimization problem were the coefficient of variation and the crystal length, which are calculated by Equations 1 and 2, respectively. The manipulated variable  $u$  is the process temperature. The MPC goal is to minimize the process performance index  $J$ , the objective function of the problem shown in Equation 3. The parameters  $\delta$  and  $\varphi$  are the weights of the outputs, and  $\gamma$  is the weight of the increment input. The parameters  $\delta$ ,  $\varphi$  and  $\gamma$  were tuned as 1, 10 and 5, respectively. The superscript  $SP$  indicates the set-point value of the controlled variable. The prediction horizon  $P$  was defined as 10, and the control horizon was tuned as 1. The optimization constraint is given by Equation 4.

$$V = \sqrt{\frac{\mu_2 \mu_0}{\mu_1^2} - 1} \quad (1)$$

$$L = \mu_1 / \mu_0 \quad (2)$$

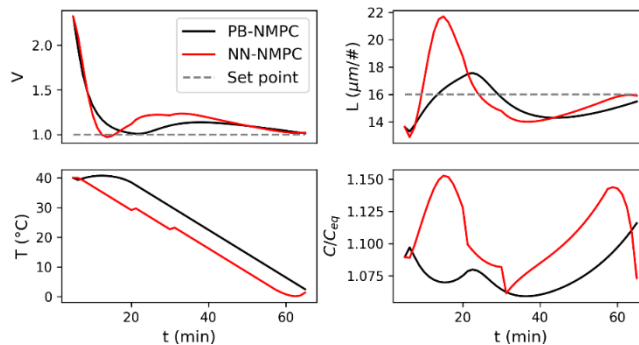
$$J = \sum_{j=1}^P \{ \delta [V(k+j) - V^{SP}(k+j)]^2 + \varphi [L(k+j) - L^{SP}(k+j)]^2 \} + \gamma [\Delta u(k)]^2 \quad (3)$$

$$-1 \leq \Delta u(k) \leq 1 \quad (4)$$

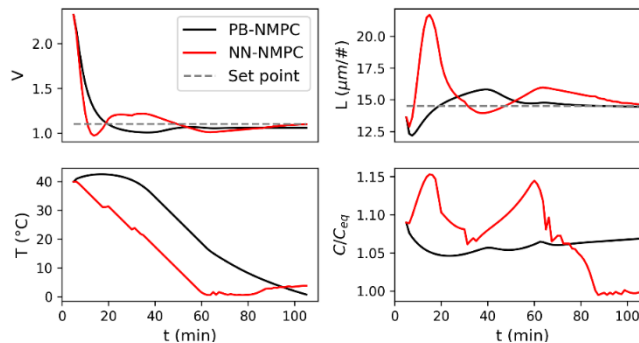
In the control loop, the process was simulated based on the model developed in the study of Moraes et al.<sup>[3]</sup>, and a sampling time of 1.25 min was adopted. This optimization problem was solved using Successive Quadratic Programming (SQP) algorithm with absolute accuracy of  $10^{-10}$  and relative accuracy of  $10^{-8}$ . The performance of the controller based on the neural networks was compared to a controller based on the phenomenological model. For the NMPC based on the population balance model, the optimization problem was solved using the IPOPT solver, defining a tolerance of  $10^{-8}$ .

The controller was tested for set-points equal to 16 and 14.5  $\mu\text{m}$  for the crystal's length and 1.0 and 1.1 for the coefficient of variation. The results of the simulations for both controllers are shown in Figures 1 and 2, in which PB-NMPC is the NMPC based on the phenomenological model developed by Moraes et al.<sup>[3]</sup>, and NN-NMPC is the NMPC based on the neural networks. In the first case, the PB-NMPC reached the set-point for the coefficient of variation but could not get the set-point for the length. However, the NN-NMPC could achieve both set-points during this simulation. In the second case, the

NN-NMPC could reach both set-points, but the PB-NMPC presented a small offset for the coefficient of variance. Even with the advantage of the NN-NMPC, the PB-NMPC presented a mean squared error (MSE) equal to 1.7437 for the crystal's length and 0.0787 for the coefficient of variation in the first scenario, while these values were 0.5731 and 0.0526 in the second case studied. In contrast, the NN-NMPC shown an MSE equal to 4.9396 for the crystal's length and 0.0970 for the coefficient of variation in the first case, while these values were 4.1703 and 0.0428 in the second scenario. In both cases, the ratio  $C/C_{eq}$  was higher than one almost all the time, indicating that the simulations happened basically in the saturated condition. This situation was expected because the process does not stay too long in the subsaturated condition and rapidly tries to reach equilibrium ( $C/C_{eq}$  equals to one).

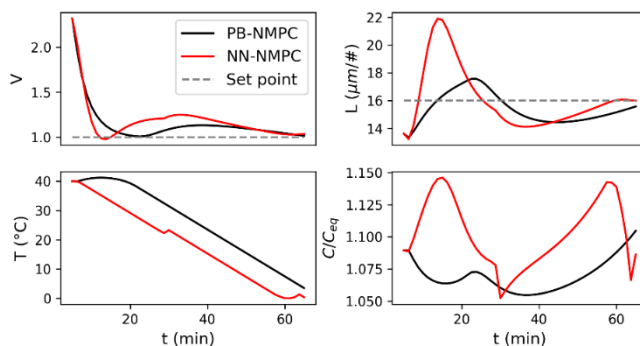


**Figure 1:** Control loop results for set-points equal to 16  $\mu\text{m}$  and 1.0.



**Figure 2:** Control loop results for set-points equal to 14.5  $\mu\text{m}$  and 1.1.

The performance of the controllers was tested to achieve the same set-points of 16  $\mu\text{m}$  and 1.0 but considering mismatches in the growth coefficient and the dissolution coefficient, using a difference of 10% on their original values. The results for this case are presented in Figure 3, showing similar results to the previous case. The NN-NMPC could reach both set-points once again, while the PB-NMPC could not reach the set-point for the crystal's length. The MSE values were 4.9370 for the crystal's length and 0.0921 for the coefficient of variation considering the NN-NMPC, while these values were 1.5551 and 0.0792 for the PB-NMPC.



**Figure 3:** Control loop results considering mismatches.

#### 4. Conclusion and Prospects

In this study, neural network models were developed to predict the first four moments and the solute concentration of a batch crystallization process. The hyperparameters of the neural networks were defined using an optimization approach. The capacity of the neural networks was tested to predict these values from two to twenty sampling times ahead, verifying  $R^2$  greater than 99% for the test samples in all cases analyzed. Moreover, the neural networks have the advantage to use only one model to make predictions of a particular state variable under the saturated and undersaturated conditions. The neural network models were used as the internal model of an NMPC to control the crystal's length and the coefficient of variation manipulating the temperature. The proposed controller could efficiently maintain the controlled variables in the set-points, showing similar results to an NMPC based on the phenomenological model. The simulations remained mostly in the supersaturation condition, since the process remains quickly in the undersaturation region and seeks to reach equilibrium. Moreover, the controller could efficiently handle mismatches in the model. The proposed strategy reached the set-points efficiently and with low computational cost, showing a potential to be used in real crystallization processes. A similar application could be made to a continuous crystallization process for future work.

#### Acknowledgements

CAPES for providing scholarship (Finance Code 001), the financial support from CNPq (Grants No. 303587/2020-2 and 311153/2021-6), and FAPERJ (Grant No. E-26/201.148/2022).

#### References

1. Ahn B., Bosetti L., Mazzotti M., 2021. Secondary Nucleation by Interparticle Energies. II. Kinetics. *Cryst. Growth Des.*, v. 22, p. 74–86.
2. Lima, F. A. R. D., Moraes, M. G. F., Secchi, A. R., Souza Jr., M. B., 2022. Development of a recurrent neural networks-based NMPC for controlling the concentration of a crystallization process. *Digital Chemical Engineering*, v. 5, 100052.
3. Moraes, M. G. F., Secchi, A. R., Vieira, G. M., Barreto Jr., A. G., Souza Jr., M. B., 2019. Measurement and Modeling of Crystal Size and Shape Distributions of Potassium Sulfate Through Dynamic Image Analysis in Batch Cooling Crystallization. *I Brazilian Congress on Process Systems Engineering (PSEBR 2019)*, p. 170.
4. Himmelblau D. M., 2008. Accounts of Experiences in the Application of Artificial Neural Networks in Chemical Engineering. *Ind. Eng. Chem. Res.*, v. 47, p. 5782–5796.

# A novel neural network bounds-tightening procedure for multiparametric programming and control

Dustin Kenefake,<sup>a,b</sup> Rahul Kakaodkar,<sup>a,b</sup> Moustafa Ali,<sup>a,b</sup> Efstratios N. Pistikopoulos,<sup>a,b</sup>

<sup>a</sup>Texas A&M Energy Institute, Texas A&M University, College Station, United States

<sup>b</sup>Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, United States

## Abstract

Recently, it has been shown that rectified linear unit (ReLU) based neural networks (NN) are mixed integer linear representable and therefore can be incorporated into mixed integer linear programming frameworks. In this work we propose a novel tightening procedure that is based on a multiparametric programming formulation of the corresponding ReLU-reformulated optimization problem. The bounding procedure features 1) the generation of valid tight bounds on the individual auxiliary variables introduced from the ReLU NN reformulation, and 2) bounds on groups of input and output variables to these ReLU NNs. The tightened bounds are valid for all parameter realizations, and thus can be reused between different problem realizations. This tightening only needs to be computed a single time and thus the overall procedure is once-and-offline. As this bounds-tightening procedure is applied offline, more computationally expensive methods for formulation tightening can then be applied to the online case. We demonstrate the effectiveness of this method in a case study of model predictive control of a nonlinear chemostat where the dynamics are approximated with a ReLU NN.

**Keywords:** Multiparametric Programming, Model Predictive Control, Neural Network

## 1. Introduction

Multiparametric programming has garnered much interest in the control literature since it was shown that Model Predictive Control (MPC) problems can be formulated as multiparametric programs (Bemporad et al., 2000; Pistikopoulos et al., 2021) where these restated MPCs are typically referred to as multiparametric MPCs (mpMPC). The feature of multiparametric programming often used in the control literature is that it allows for the generation of explicit algebraic solutions to the MPC problem ahead of time that are functions of the uncertain parameters,  $\theta$  (Pappas et al., 2021; Pistikopoulos et al., 2021). Another feature of the multiparametric program formulation is that it represents the set of all possible optimization problems given a parametric uncertainty set  $\Theta$ , and it is this aspect that will be explored in this work. A prototypical multiparametric program is shown in eqn. 1. With  $x \in \mathbb{X}$  being the set of decision variables, and  $\theta \in \Theta$  being the set of uncertain parameters. Here,  $f(x, \theta)$  is the objective function,  $g(x, \theta) = 0$  is the set of equality constraints, and  $h(x, \theta) \leq 0$  are the inequality constraints.

$$\begin{aligned}
\min_x \quad & f(x, \theta) \\
\text{s.t.} \quad & g(x, \theta) = 0 \\
& h(x, \theta) \leq 0 \\
& x \in \mathbb{X}, \theta \in \Theta
\end{aligned} \tag{1}$$

It has been shown that ReLU-based NNs are mixed integer linear representable using one of the various reformulation strategies that have been proposed in the literature, such as the partition-based reformulation approach by Tsay et al. (2021) and the Big-M reformulation as proposed by Grimstad and Andersson (2019) and Fischetti and Jo (2018). Various methods have been explored to accelerate the solution of these optimization problems, such as deriving tighter bounds and reformulations (Tjandraatmadja et al., 2020; Tsay et al., 2021; Grimstad and Andersson, 2019) and finding the explicit solution of the resulting multiparametric program (Katz et al., 2020). Here, bounds-tightening refers to finding tight ranges for the auxiliary variables introduced in the reformulation of NNs. In the case of bounds-tightening, the literature has focused on finding tighter bounds for individual optimization problems and has not thoroughly explored the case where one can tighten and reuse bounds over sets of optimization problems, such as those that can be represented by multiparametric programming formulations, e.g., eqn. 1.

In this work, we propose a multiparametric programming-based strategy for generating tight bounds for ReLU-based NN reformulations that can be reused for all realizations of optimization problems,  $\theta \in \Theta$ . These are tight bounds that are valid for all possible problem realizations and thus can be effectively reused between parameter realizations,  $\theta \in \Theta$ . This reusing of tight bounds between problems allows for the bounds-tightening to be calculated once-and-offline for any particular multiparametric program with ReLU-based NN constraints. The effectiveness of this approach is displayed on a ReLU-based NN MPC of an unstable nonlinear chemostat process.

## 2. Method

This section is separated into two subsections. The first subsection briefly describes the Big-M MILP reformulation of ReLU-based NN. The second subsection gives an overview of the proposed bounds-tightening procedure.

### 2.1. ReLU Neural Network Reformulation

A ReLU-based NN is in essence a continuous piecewise affine function and thus can be represented via mixed integer linear constraints, in the same way that any other continuous piecewise affine function can be. With ReLU-based NN, the only non-linearity that must be addressed is the ReLU activation function itself. This is a rectification function where if the input is less than 0, then it is rectified to 0; otherwise, the input is not altered by the activation function. This can be modeled with the max function,  $\max(0, x)$ . This can be effectively linearized by introducing auxiliary variables.

The Big-M reformulation for layer  $k$  of a ReLU-based NN is shown in eqn. 2. Where the input to hidden layer  $k$  is  $z^{k+1}$ , the weight and bias values of the layer are  $W^k$  and  $b^k$  respectively, and the output of the hidden layer is  $z^{k+1}$ . Auxiliary slack variables  $s^k$  and selection variables  $y^k$  are introduced to linearize the max function exactly. In the case  $y_i^k = 0$ , then  $(W^k z^k + b^k)_i > 0$  and thus does not to be rectified so  $s_i^k = 0$ ; similarly, in the case of  $y_i^k = 1$ , then  $(W^k z^k + b^k)_i < 0$  and thus  $z_i^{k+1}$  needs to be rectified to zero

via  $s_i^k$ . Initial bounds for  $M^{s_i^k}$  and  $M^{z_i^k}$  are typically found via interval arithmetic (Grimstad and Andersson (2019)). As long as the Big-M values are significantly large, i.e. provide a true upper bound to the auxiliary variables, the resulting reformulation is valid. However, if these values are made overly large, the computational performance of the resulting optimization problem will be affected. More details on this reformulation can be found in the works of Grimstad and Andersson (2019), Fischetti and Jo (2018), and Tjandraatmadja et al. (2020).

$$\begin{aligned}
 z^{k+1} = \max(0, W^k z^k + b^k) \quad & \rightarrow \quad \begin{aligned} z^{k+1} &= W^k z^k + b^k + s^k \\ z^{k+1} &\geq 0 \end{aligned} \\
 z^{k+1} \geq 0 \quad & \rightarrow \quad \begin{aligned} 0 &\leq s_i^k \leq M^{s_i^k} y_i^k \\ 0 &\leq z_i^{k+1} \leq M^{z_i^k} (1 - y_i^k) \\ y_i^k &\in \{0,1\} \end{aligned} \end{aligned} \quad (2)$$

### 2.2. Tightening Procedure

While the Big-M reformulation of the ReLU-based NN typically allows for the use of standard MIP solvers, such as Gurobi or CPLEX, the formulation does not typically lead to a tight mathematical program (Grimstad and Andersson, 2019; Tsay et al., 2021). Instead of proposing a bounds-tightening procedure for a particular problem realization, we propose tightening over all possible problem realizations,  $\theta \in \Theta$ .

By having an explicit parameterization of an optimization problem, as represented in eqn. 1, the parametric uncertainty space,  $\Theta$ , can directly be used to find tight valid bounds on all optimization variables. By promoting the parameters,  $\theta$ , into decision variables in the reformulated optimization problem, tight worst-case bounds can be found for each variable via MILP optimization, such as can be seen in eqn. 3. These bounds are valid for every possible parameter realization in the uncertainty space and allow for these bounds to be shared between optimization problems with different parameter realizations,  $\theta \in \Theta$ .

$$\begin{aligned}
 \min_{x,z,s,y,\theta} \quad & x_i \\
 \text{s.t.} \quad & g_{\text{reformulated}}(x, z^1, \dots, z^M, s^1, \dots, s^M, y^1, \dots, y^M, \theta) = 0 \\
 & h_{\text{reformulated}}(x, z^1, \dots, z^M, s^1, \dots, s^M, y^1, \dots, y^M, \theta) \leq 0 \\
 & x \in \mathbb{X}, \quad \theta \in \Theta \end{aligned} \quad (3)$$

A similar optimization problem, as shown in eqn. 3, is solved for every variable in the formulation, e.g.,  $s_i^k$ ,  $z_i^k$ ,  $y_i^k$  and  $\theta_i$ , for both the upper and lower bounds of each variable given the parameters,  $\theta \in \Theta$ .

### 3. Case Study: Unstable Nonlinear Chemostat

The proposed methodology is applied to a ReLU-based NN MPC for the unstable nonlinear chemostat process adopted by Eaton and Rawlings (1992). The model equations of the considered process are shown in eqn. 4, describing the dynamics of the biomass,  $x$ , and the substrate,  $s$ , with time along with the substrate inhibition model,  $\mu(s)$ . In this process the specific parameters are  $s_f = 4$ ,  $y = 0.4$ ,  $\mu_{\max} = 0.53$ ,  $k_m = 0.12$ ,  $k_1 = 0.4545$ . The dilution rate,  $D$  is the input variable that can be manipulated to change the trajectory of the states,  $x$  and  $s$ , of the process.

$$\begin{aligned}
\frac{dx}{dt} &= (\mu(s(t)) - D(t))x(t) \\
\frac{ds}{dt} &= (s_f - s(t))D(t) - \frac{\mu(s(t))x(t)}{y} \\
\mu(s) &= \frac{\mu_{\max}s(t)}{k_m + s(t) + k_1s(t)^2}
\end{aligned} \tag{4}$$

The dynamics of the chemostat process are then approximated by a ReLU-based NN model with 3 hidden layers with 15 nodes in each hidden layer. The NN was generated via the following procedure. A data set was constructed by simulating 1,000,000 randomly generated initial states and input actions inside of the operating window,  $(x, s, D)$ , for a single time step to build an input-output data set,  $(x(t), s(t), D(t)) \rightarrow (x(t + \Delta t), s(t + \Delta t))$ . The input action,  $D(t)$ , was held constant for the entire time interval,  $\Delta t$ , in a zero-order old fashion. This input-output data set was split into a train and test set with a 90/10 ratio. The NN model was then trained via stochastic gradient decent (sgd) to a fit performance of  $9.16 \cdot 10^{-5}$  MSE, and  $6.9 \cdot 10^{-3}$  MAE. This model was then validated on the test set with a performance of  $9.3 \cdot 10^{-5}$  MSE and  $6.3 \cdot 10^{-3}$  MAE, indicating that a predictive model was generated that has acceptable performance for this application.

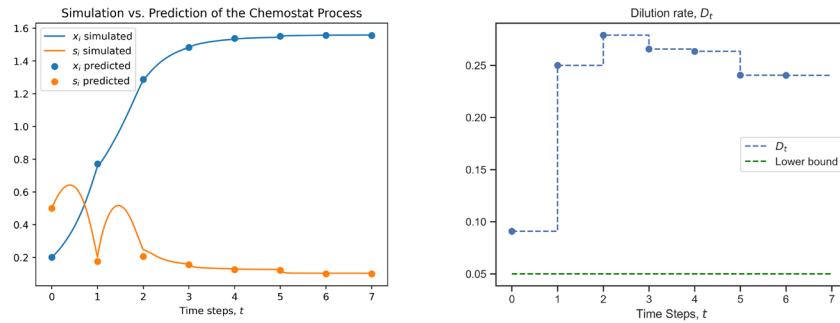
The multiparametric MPC (mpMPC) for this process can be seen in eqn. 5. Here the objective is to maximize the biomass production at each time period,  $i$ . The predictive model of this MPC is based on the NN that was generated in the previous step. The states of the system are bounded above and below, as is the input action. For this case study, the control and output horizon are 7-time steps.

$$\begin{aligned}
&\max_{x,s,d} \sum_{i=0}^N x_i \\
\text{s.t. } &x_{i+1}, s_{t+1} = \text{NN}(x_t, s_t, D_i), \quad \forall i \in \{0, \dots, N-1\} \\
&0.1 \leq x_i \leq 2.0, \quad \forall i \in \{0, \dots, N\} \\
&0.1 \leq s_i \leq 2.0, \quad \forall i \in \{0, \dots, N\} \\
&0.05 \leq D_i \leq 0.6, \quad \forall i \in \{0, \dots, N-1\} \\
&x_0 = \theta_0, s_0 = \theta_1, \theta \in \theta
\end{aligned} \tag{5}$$

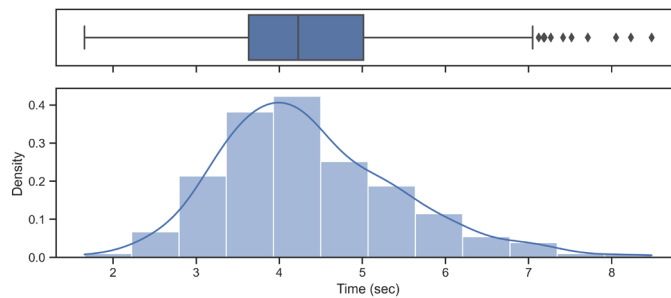
Firstly, the proposed bounds-tightening procedure is applied to the NN separate from the overall optimization problem. Here tight upper and lower bounds on the auxiliary variables were found that are valid for all problem realizations,  $\theta \in \theta$ . This tightening of the individual NN took only 11.5 seconds. These tightened bounds were then used when generating the Big-M reformulation of the mpMPC. It was found that 12 of the 45 auxiliary binary variables in each NN were only feasible at either 0 or 1 and thus could effectively be removed from every problem instance.

The MILP-reformulated mpMPC has 701 continuous variables, 315 binary variables, and 2,221 constraints. Tight upper and lower bounds were then found for  $x_i$ ,  $s_i$ , and  $D_i$ , as well as sums of these variables, e.g.,  $\sum x_i$ , which are valid for any problem realization,  $\theta \in \theta$ . These bounds were added as constraints to the reformulated mpMPC. This once-and-offline bounds-tightening process took 2,941 seconds.

The effectiveness of the once-and-offline bounds-tightening procedure is shown in a control case study where the computational burden of the ReLU NN-based MPC is compared with and without the proposed bounds-tightening procedure. For these computational experiments, we used Gurobi 10.0 as the MILP solver and used a computer with an Intel i7-12700k and 32GB DDR4 RAM. Here, we take the system at the initial state realization of  $\theta = (0.2, 0.5) = (x_0, s_0)$ . Without the proposed bounds-tightening procedure applied, the MPC problem hit the time-out limit of 8,000 seconds, explored 21,976,708 nodes, and performed 334,045,961 simplex iterations with a gap of 9.9%. Utilizing the proposed bounds-tightening procedure where valid tight bounds are generated for all problem realizations ahead of time, the ReLU-based NN MPC problem was solved to proven optimality in only 5.36 seconds, explored 14,046 nodes, and performed 205,700 simplex iterations. The predictive performance of the NN model and the effect of applying the optimal inputs,  $D_t$ , that were generated by the MPC to the process can be seen in Figure 1.



*Figure 1. Evolution of the chemostat process with time given the dilution rates,  $D_t$ , from the MPC compared to the prediction of the NN inside the MPC.*



*Figure 2. Histogram and box plot of time to solve each problem in the problem set.*

Another computational study was performed, where 500 random initial state realizations of the chemostat system,  $\theta = (x_0, s_0) \in \Theta$  were generated. The corresponding MPC problems for each initial state were solved utilizing the calculated tight bounds. Each problem was solved to optimality; if the resulting problem was found to be infeasible, the time to solve was reported as the time to prove infeasibility. The average time to solve the pre-tightened MPC problems was 4.37 seconds, with a maximum of 8.48 seconds.



The distribution of the solve times for these control problems is presented in Figure 2. As a comparison, 5 problems were attempted without the improved bounds; however, they all timed out at 8,000 seconds without a proven optimal solution. As shown, this proposed procedure leads to significant reductions in the time to solve, especially when compared to the case where these precomputed bounds are not provided.

#### 4. Conclusions

In this work, a once-and-offline bounds-tightening procedure for ReLU-based NN multiparametric programming problems was developed. This method finds tight bounds for every auxiliary variable introduced via the Big-M reformulation that is valid for every realization of the parametric uncertainty set,  $\theta \in \Theta$ , allowing for the reuse of these bounds between problems. In addition, bounds on the input and output variables of the NN were computed, removing the need to generate new bounds for each optimization problem. The tightened bounds on the auxiliary, input and output variables furnish a much more computationally tractable optimization problem, and this was demonstrated in the case study where a ReLU-based NN MPC was implemented for a chemostat process. The proposed method made the controller computationally feasible, taking only 5.36 seconds compared to the untightened case which timed out at 8,000 seconds.

In the future, we plan to apply the proposed tightening procedure to the partition reformulation of ReLU-based NNs as proposed by Tsay et al. (2021) to analyze if similar computational acceleration can be realized in those ReLU-based NN reformulations. We plan to apply this to other realizations of multiparametric mixed integer programming formulations. Hierarchical MPC, which consists of layers of MPC and is typically stated as a multilevel optimization problem, is of keen interest to the authors. In the case of two controllers where the lower-level controller is a convex Quadratic Program, this can be reformulated as a multiparametric mixed integer program that could have the same tightening procedure applied to it.

#### References

- A. Bemporad, M. Morari, V. Dua, E. N. Pistikopoulos, 2000. The explicit solution of model predictive control via multiparametric quadratic programming 2, 872–876 vol.2.
- J. W. Eaton, J. B. Rawlings, 1992. Model-predictive control of chemical processes. *Chemical Engineering Science* 47 (4), 705–720.
- M. Fischetti, J. Jo, 2018. Deep neural networks and mixed integer linear optimization. *Constraints* 23 (3), 296–309.
- B. Grimstad, H. Andersson, 2019. Relu networks as surrogate models in mixed-integer linear programs. *Computers & Chemical Engineering* 131, 106580.
- J. Katz, I. Pappas, S. Avraamidou, E. N. Pistikopoulos, 2020. Integrating deep learning models and multiparametric programming. *Computers & Chemical Engineering* 136, 106801.
- I. Pappas, D. Kenefake, B. Burnak, S. Avraamidou, H. S. Ganesh, J. Katz, N. A. Diangelakis, E. N. Pistikopoulos, 2021. Multiparametric programming in process systems engineering: Recent developments and path forward. *Frontiers in Chemical Engineering* 2, 32.
- E. N. Pistikopoulos, N. A. Diangelakis, R. Oberdieck, 2021. *Multi-parametric optimization and Control*. Wiley.
- C. Tjandraatmadja, R. Anderson, J. Huchette, W. Ma, K. K. Patel, J. P. Vielma, 2020. The convex relaxation barrier, revisited: Tightened single-neuron relaxations for neural network verification. *Advances in Neural Information Processing Systems* 33, 21675–21686.
- C. Tsay, J. Kronqvist, A. Thebelt, R. Misener, 2021. Partition-based formulations for mixed-integer optimization of trained relu neural networks. *Advances in Neural Information Processing Systems* 34, 3068–3080.

# Multistage Economic NMPC for Gas Pipeline Networks with Uncertainty

Sakshi Naik,<sup>a,\*</sup> Robert Parker,<sup>a</sup> Lorenz T. Biegler,<sup>a</sup>

<sup>a</sup>*Carnegie Mellon University, Pittsburgh, PA 15289, USA*

\**ssnaik@andrew.cmu.edu*

## Abstract

Gas pipelines form complex highly integrated networks to transport natural gas with dynamic operation due to time varying demands, composition, and ambient conditions. These can be modeled and optimized through non-linear optimal control problems with model equations, operation bounds and uncertainty descriptions. In this study, we use a multistage Economic Nonlinear Model Predictive Controller (eNMPC) to find optimal operational policies for networks with dynamic demands and uncertain parameters. This approach relies on constructing a scenario tree by generating extreme cases of the uncertain parameters with separate control sequences to address constraint violations in each case. For demonstration, we assume that the efficiency of the compressors is uncertain. Under this uncertainty we show that standard eNMPC, designed with a nominal value of compressor efficiency, violates constraints in the plant. We demonstrate that multistage eNMPC prevents constraint violations in all uncertain realizations and provides a promising robust control strategy for these networks.

## 1. Introduction

Natural gas is the cornerstone of the electric grid and significant increases in consumption are projected in the coming years. Gas is transported from suppliers to consumers using large pipeline networks, which operate under time varying conditions such as temperature, customer demands and gas composition. Pipeline networks typically consist of supply nodes, demand nodes, compressor stations and individual lines of various diameters and lengths. Compressor stations are used to supply boost pressure to the gas to compensate for friction losses, and this is a major operating cost (Liu et al. (2020)). Moreover, a major issue in pipeline operation is the presence of uncertainty. Liu et al. (2020) used a robust optimization approach to optimize gas networks with demand and composition uncertainty. Also, two-stage stochastic optimal control for pipeline networks has been performed by Zavala (2014) with a focus on uncertainty in gas demands. However, this work assumes that uncertain profiles will resolve in time after a fixed horizon, but it was not implemented in a moving horizon approach. In contrast, we consider a multistage MPC in which the value of the uncertain parameter is revealed at each sampling time. Lin et al. (2022) demonstrated the effectiveness of multistage NMPC to handle uncertainty in pumping treatment in hydraulic fracturing. The advantage of using an eNMPC over tracking NMPC for optimal pipeline operation was demonstrated in a deterministic setting by Gopalakrishnan and Biegler (2013). In this paper, we extend this approach through multistage eNMPC for optimal control of gas-pipeline networks under uncertainty.

There are multiple uncertain parameters that can be considered in the operation of gas pipelines such as uncertainty in demand, gas composition, compressor efficiency, etc.

Although the concept of multistage eNMPC can be generalized to any uncertain parameter, we choose the uncertainty in compressor efficiency to demonstrate the results of this work. The uncertainty in compressor efficiency is due to time-varying flow of gas through the compressors and ambient conditions of operation such as temperature and relative humidity. Moreover, compression unit failures are extremely challenging for gas transmission systems and a significant issue for gas operators (Tran et al. (2018)). Using multistage eNMPC to control gas-pipeline networks with uncertainty is therefore a promising direction for further research.

## 2. Model Formulation

We first define the pipeline momentum and mass balances, compressor equations, node balances and equations that describe the network interconnections (Zavala (2014)). We assume isothermal gas flow through a horizontal pipe. The momentum balance and mass balance equation for a given line  $l \in L$  with a given value of parametric uncertainty  $w \in \Omega: \{1, 2, \dots, N_\Omega\}$  is given by:

$$\frac{1}{A_l} \frac{dF_l(t, x, w)}{dt} + \frac{dP_l(t, x, w)}{dx} + \frac{8\lambda_l c^2 F_l(t, x, w) |F_l(t, x, w)|}{\pi^2 D_l^5 P_l(t, x, w)} = 0 \quad (1) \quad \frac{dP_l(t, x, w)}{dt} + \frac{c^2}{A_l} \frac{dF_l(t, x, w)}{dx} = 0 \quad (2)$$

where  $t \in T$  and  $x \in X$  are points in time and space domains respectively.  $P_l$  and  $F_l$  refer to the pressures and flows in line  $l$  and  $c$  is the speed of sound of the gas.  $A_l$  is the cross-sectional area of the pipeline and  $D_l$  is the diameter of the pipeline  $l$ . Area and friction factor ( $\lambda_l$ ) equations are:

$$A_l = \frac{\pi D_l^2}{4} \quad \forall l \in L \quad (3) \quad \lambda_l = \left[ 2 \log_{10} \left( \frac{3.7 D_l (mm)}{(\text{rugosity})} \right) \right]^{-2} \quad (4)$$

In the pipe network we consider the node set  $\mathcal{N}$  and line set  $L$ . For each node  $n \in \mathcal{N}$  we define a set of supplies  $S_n$  and demands  $D_n$ .  $L^{in}$  is the subset of lines bringing gas to node  $n$  and  $L^{out}$  is the subset of links which take out gas from node  $n$ .  $s_i(t, w)$  represents the actual supply flowrate from supply  $i$  and  $d_j(t, w)$  represents the actual demand flowrate from demand  $j$ . The flow balances at the nodes are given by:

$$\sum_{i \in S_n} s_i(t, w) + \sum_{l \in L_n^{in}} F_l^{in}(t, w) = \sum_{j \in D_n} d_j(t, w) + \sum_{l \in L_n^{out}} F_l^{out}(t, w) \quad \forall n \in \mathcal{N} \quad (5)$$

$F_l^{in}$  and  $P_l^{in}$  are dummy flows linked to the pipeline flows by the boundary equations:

$$F_l(t, \bar{L}_l, w) = F_l^{in}(t, w) \quad (6) \quad F_l(t, 0, w) = F_l^{out}(t, w) \quad \forall l \in L, w \in \Omega, t \in T \quad (7)$$

where  $\bar{L}_l$  is the length of line  $l$ . In addition to the pipelines and the nodes, we have compressor stations associated with some pipelines. We define a set of pipelines which have a compressor associated with them as  $L_a$  (active lines) and a set of pipelines which don't have a compressor as  $L_p$  (passive lines). The power consumed in the active lines  $L_a$  is given by:

$$\text{Power}_l(t, w) = T_{in}(t, w) c_p \frac{F_l(t, 0, w)}{\eta_l(t, w)} \left[ \left[ \frac{P_{in,l}(t, w) + \Delta P_l(t, w)}{P_{in,l}(t, w)} \right]^\beta - 1 \right] \quad \forall l \in L_a, w \in \Omega, t \in T \quad (8)$$

$T_{in}$  is the temperature of the inlet stream.  $\eta_l(t, w)$  is the efficiency of the compressor in line  $l$  which is an uncertain parameter.  $\beta$  is the compression coefficient which is calculated as  $\beta = \frac{c_p - c_v}{c_p}$ . The boundary conditions for the pipeline pressures are:

$$P_n^{rec}(t, w) = P_1(t, \bar{L}_1, w) \quad \forall l \in L, t \in T, w \in \Omega \quad (9)$$

$$P_1(t, 0, w) = P_{in,1}(t, w) \quad \forall l \in L_p, t \in T, w \in \Omega \quad (10)$$

$$P_1(t, 0, w) = P_{in,1}(t, w) + \Delta P_1(t, w) \quad \forall l \in L_a, t \in T, w \in \Omega \quad (11)$$

$P^{rec}$  are the pressures received at the nodes from the pipelines. The time domain  $t$  is discretized using an implicit Euler method with equally spaced points each having  $\Delta t$  length. We represent each point in the discretized time set by the index  $\tau$ . The space domain along the length of each line is discretized using a forward discretization scheme with  $N_x$  equally spaced points with each segment having a length of  $\Delta x$ . We define  $\bar{X} = \{1, 2, \dots, N_x - 1\}$ . Spatial and temporal discretizations were chosen based on simulation results obtained from the model. The differential states from the resulting differential-algebraic equations (DAEs) correspond to spatially discretized flows and pressures. The algebraic state equations include the node balances, compressor power expression and the boundary conditions. A dynamic optimization problem is formulated to minimize the boost cost of the compressor while meeting the demands without depleting the linepack inventory (gas stored in the pipelines at the end of operation). Fixing the supply pressure and demand flows leaves us with the compressor boost pressures as the control variables.

### 2.1. eNMPC Formulation

eNMPC provides a sequence of controls for a given horizon. The first control input is implemented in the plant. The new state of the plant is then sampled and acts as a new starting point for the controller. The horizon of the controller is then advanced by one sample time and the process repeats again. We consider an eNMPC horizon of 24 hours and a total time of 48 hours. The points on the discretized horizon are defined by the set

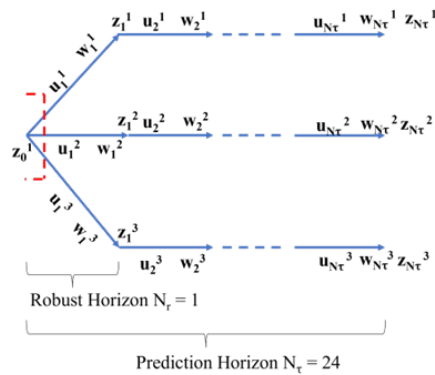


Figure 1. Scenario tree with a robust horizon  $N_r = 1$  for multistage eNMPC

the linepack inventory, we use a lower bound on linepack flow and pressure. The optimization problem for each control horizon is formulated as:

$$\begin{aligned}
 \min_{\Delta P_{l,\tau}(w)} \quad & \sum_{l \in L_a} \sum_{\tau \in \hat{T}} C_{pr} \times Power_{l,\tau}(w) \Delta t \\
 d_{j,\tau}^{target} - d_{j,\tau}(w) = 0 \quad & \forall \tau \in \hat{T}, \forall j \in D_n, \forall n \in \mathcal{N}, \forall w \in \Omega \\
 P_{l,N_r,k}(w) \geq P_{LB} \text{ bar} \quad & \forall l \in L, \forall k \in \{2, 3, \dots, N_x\}, \forall w \in \Omega \\
 F_{l,N_r,k}(w) \geq F_{LB} \text{ kmol/hr} \quad & \forall l \in L, \forall k \in \bar{X}, \forall w \in \Omega \\
 s_{LB} \leq s_{i,\tau}(w) \leq s_{UB} \quad & \forall i \in S_n, \forall n \in \mathcal{N}, \forall \tau \in \hat{T}, \forall w \in \Omega \\
 P_{LB}^{rec} \leq P_{j,\tau}^{rec}(w) \leq P_{UB}^{rec} \quad & \forall j \in D_n, \forall n \in \mathcal{N}, \forall \tau \in \hat{T}, \forall w \in \Omega \\
 0 \leq Power_{l,\tau}(w) \leq Power_{UB} \quad & \forall \tau \in \hat{T}, \forall w \in \Omega
 \end{aligned}
 \quad \left. \begin{array}{l} \text{Demand} \\ \text{constraints} \\ \text{Terminal} \\ \text{bounds} \\ \text{Variable} \\ \text{bounds} \end{array} \right\}$$

Constraints (1)-(12), discretized in  $\hat{T}$  and  $\bar{X}$

$\hat{T} = \{1, 2, \dots, N_\tau\}$ . The goal of the optimization problem is to minimize the boost cost of the compressor while meeting the demands, without completely depleting the inventory. The boost cost is calculated based on the power of the compressor and the cost of electricity. Assuming the cost of electricity is  $C_{pr}$  \$/kWh, the boost cost can be calculated as:

$$J(w) = \sum_{l \in L_a} \sum_{\tau \in \hat{T}} C_{pr} \times Power_{l,\tau}(w) \Delta t \quad (12)$$

To meet the demands we have equality constraints between the target demands and the actual demands. To avoid depletion of

The above constraints are labelled as (13). This formulation holds for the dynamic optimization of a pipeline network where the value of the uncertain parameter is known. In case of uncertain efficiency it is useful to implement Multistage eNMPC as it considers explicit realization of uncertainty by considering a scenario tree. In accordance with the standard setup of multistage eNMPC (Lucia et al. (2013)) we choose three realizations of the uncertain parameter: minimum, nominal and maximum in set  $\Omega = \{w^{min}, w^{nom}, w^{max}\}$ . Non-anticipativity constraints enforce equal values of the control variables for branches  $b$  and  $\bar{b}$  with the same parent node. Figure 1 shows a scenario tree with a robust horizon  $N_r = 1$ . The level of uncertainty stays the same after the robust horizon. This allows us to have only two non-anticipativity constraints and the problem size remains tractable. Since the efficiency at each time point is an uncertain parameter, we have  $N_\tau$  uncertain parameters. The optimization problem for a multistage NMPC is then formulated as:

$$\min_{\Delta P(w_\tau^b)_{\{\tau \in T\}}} \sum_{w_\tau^b \in \Omega} \frac{1}{|\Omega|} J(w_\tau^b)$$

$$\Delta P_\tau^b = \Delta P_\tau^{\bar{b}} \text{ if } , z_{\tau-1}^b = z_{\tau-1}^{\bar{b}} \quad \forall b, \bar{b} \in B, \forall \tau \in \{1, \dots, N_r\} \quad \text{Nonanticipativity constraints}$$

$$w_\tau^b = w_{\tau+1}^b \quad \forall \tau \in \{N_r+1, \dots, N_\tau-1\}$$

Where B denotes the set of all branches in the scenario tree.  $z_\tau^b, \Delta P_\tau^b, w_\tau^b$  are the states, controls, and uncertain parameters at stage  $\tau$  in branch  $b$ . The objective function is the average of the objectives of the nominal NMPC. The system is modeled in IDAES (Lee et al. (2021)), with Pyomo DAE (Nicholson et al. (2018), Bynum et al. (2021)) and is solved using IPOPT version 3.13.2 (Wächter and Biegler (2006)).

### 3. Model Formulation

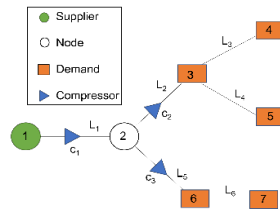


Figure 2(a) Gas pipeline network

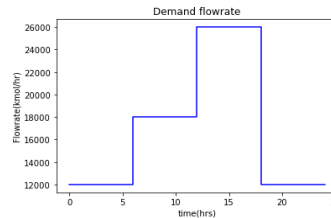


Figure 2(b) Dynamic demand profile at demand nodes

A pipeline network with five demands, one supply and three compressor stations is considered to demonstrate the proposed approach (Liu et al. (2020)). Figure 2(a) shows the network structure while the gas flow demand for each demand node is shown Figure 2(b). It is assumed that the profile repeats over a period of 24 hours. Node 1 is the supply node with a fixed gas supply pressure of 100 bar. Node 2 is a junction node with only inlet and outlet pipeline with no demand or supply. It is assumed that at the initial state, compressor  $c_1$  provides a boost pressure of 20 bar. This is necessary to obtain a consistent initial condition for our dynamic system. We use an upper bound of 60 MW on

compressor power. A standard NMPC is used on this network to find the optimal control

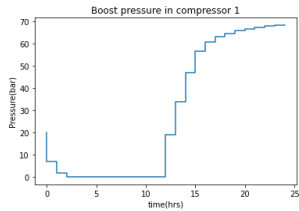


Figure 3(a) Boost pressure in compressor  $c_1$

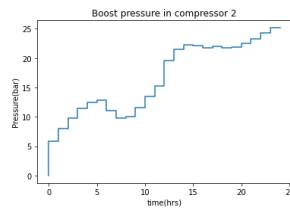


Figure 3(b) Boost pressure in compressor  $c_2$

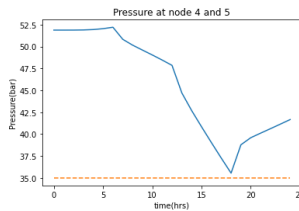


Figure 3(c) Pressure at node 4 & 5

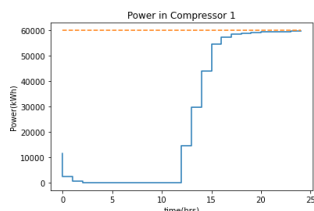


Figure 3(d) Power profile of compressor  $c_1$

actions with no plant-model mismatch. The NMPC has a horizon of 24 hrs with a sampling time of 1 hr. Under deterministic conditions, the eNMPC controller finds the optimal boost pressure profiles compressors  $c_1$  and  $c_2$  in the network as shown in Figure 3(a) and Figure 3(b), respectively. The pressures at nodes 4 and 5 are also at their lower bounds as shown in Figure 3(c). For the given set of initial and

boundary conditions, there is no boost pressure required in compressor  $c_3$ . Figure 3(d) shows that the power in compressor  $c_1$  is at its bound and is likely to experience bound violations under uncertainty in compressor efficiency.

### 3.1. Introducing uncertainty in plant

Now we consider uncertainty in the efficiencies of all three compressors. Five randomly generated efficiency profiles are applied, with efficiency at each time point taking one of the three extreme values  $\{0.55, 0.7, 0.85\}$ . The standard eNMPC modeled with nominal compressor efficiency of 0.7 is unaware of this uncertainty and therefore is likely to violate the power bound in multiple efficiency realizations. It can be seen from Figure 4(a) that the power bound in compressor  $c_1$  is violated beyond 15 hours with standard

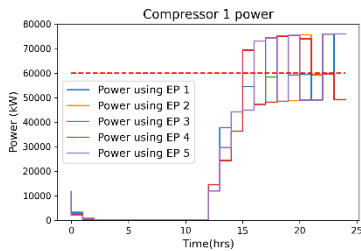


Figure 4(a) Compressor  $c_1$  power profiles using standard eNMPC with plant-model mismatch

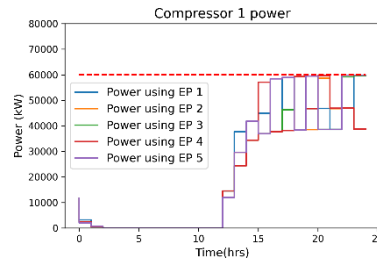


Figure 4(b) Compressor  $c_1$  power profiles using multistage eNMPC with plant-model mismatch

eNMPC. (The different colors indicate the power profiles obtained in each uncertain scenario). Multistage eNMPC as shown in Figure 4(b) doesn't violate the power bound in any case. The average cost obtained using standard eNMPC was US\$37885 while a multistage eNMPC predicted a cost of US\$42963. The optimal solution obtained using

the multistage eNMPC is 13.4% more expensive compared to the standard eNMPC, this represents the cost of handling uncertainty. The multistage eNMPC finds alternate control

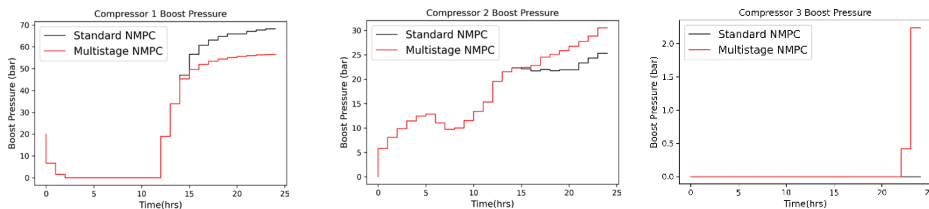


Figure 5. Boost pressure profiles obtained using standard eNMPC vs Multistage eNMPC

profiles which are robust to uncertainty and don't lead to bound violations in the plant. Figure(5) shows a comparison between the controls found by the multistage eNMPC and the standard eNMPC. We see that multistage eNMPC predicts the same control sequence as a standard eNMPC until 15 hrs. Beyond 15 hrs, multistage eNMPC suggests that a lower boost pressure should be used in compressor  $c_1$  to avoid the power bound violation. But, to supply gas above the minimum demand pressure at nodes 4 and 5, a higher boost pressure is used in compressor  $c_2$ . Also, unlike the deterministic case, Compressor  $c_3$  provides boost pressure, which is required to satisfy the lower bounds on terminal pressure.

#### 4. Conclusions and Future Work

Standard eNMPC with fixed, nominal compressor efficiency finds the optimal control sequence without violating any bounds. On the other hand, when standard eNMPC encounters parametric uncertainty, we observe bound violations due to plant-model mismatch. The multistage eNMPC approach results in a control policy which does not violate any bounds in the uncertain plant. In the future, we plan to apply these eNMPC strategies to larger, more complex problems and investigate how they scale up to real world GASLIB networks (Schmidt et al. (2017)).

#### References

- M.L. Bynum, 2021, Pyomo—optimization modeling in python, 3rd Edition. Vol. 67. Springer Science & Business
- A. Gopalakrishnan, 2013. Economic Nonlinear Model Predictive Control for periodic optimal operation of gas pipeline networks. *Computers & Chemical Engineering* 52.
- A. Lee, 2021, The IDAES process modeling framework and model library-Flexibility for process simulation and optimization. *Journal of Advanced Manufacturing and Processing* 3 (3), e10095
- K.-H. Lin, 2022, Multistage nonlinear model predictive control for pumping treatment in hydraulic fracturing. *AIChE Journal* 68 (3), e17537,
- K. Liu, 2020, Dynamic optimization of natural gas pipeline networks with demand and composition uncertainty. *Chemical Engineering Science* 215, 115449.
- S. Lucia, 2013. Multi-stage Nonlinear Model Predictive Control Applied to a Semi-batch Polymerization Reactor under Uncertainty. *Journal of Process Control* 23, 1306–1319.
- B. Nicholson, 2018, Pyomo.DAE: a modeling and automatic discretization framework for optimization with differential and algebraic equations. *Mathematical Programming Computation* 10 (2), 187–223.
- M. Schmidt, 2017. GasLib—A Library of Gas Network Instances. *Data* 2 (4), 40, number: 4 Publisher: Multidisciplinary Digital Publishing Institute.
- T. H. Tran, 2018. Linepack planning models for gas transmission network under uncertainty. *European Journal of Operational Research* 268 (2), 688–702.
- A. Wachter, 2006. On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. *Mathematical Programming* 106 (1), 25–57.
- V. M. Zavala, 2014. Stochastic optimal control model for natural gas networks. *Computers & Chemical Engineering* 64, 103–113.

## **A new time-bucket MILP formulation for optimal lot-sizing and scheduling of an industrial make-and-fill process**

Roderich Wallrath <sup>a,b</sup>, Florian Seanner <sup>c</sup>, Matthias Lampe <sup>a</sup>, Meik Franke <sup>b</sup>

<sup>a</sup>*Bayer AG, Kaiser-Wilhelm Allee 1, 51368 Leverkusen, Germany*

<sup>b</sup>*University of Twente, Faculty of Science and Technology, Sustainable Process Technology, Process Design and Optimization, Drienerlolaan 5, 7522 NB Enschede, The Netherlands*

<sup>c</sup>*SimPlan Systems GmbH, Sophie-Scholl-Platz 6, 63452 Hanau, Germany*

[r.wallrath@utwente.nl](mailto:r.wallrath@utwente.nl)

### **Abstract**

We propose a new time-bucket MILP model for lot-sizing and scheduling problems arising in multistage production processes. The time-bucket model benefits from advantages of both continuous and discrete time representations while overcoming their shortcomings. In particular, we show how the time-bucket model allows to easily include a variety of typical, important real-world constraints, can be solved with moderate computer effort, and thus promotes MILP for large-scale, industrial problems. To illustrate that, we apply the time-bucket approach to the flow shop problem of a batch formulation and filling process from an industrial pesticide production. We reconcile the MILP solution with a validated discrete event simulation (DES) model of the process to obtain optimal and real-world feasible results. A comparison of the MILP-DES solution to a manually optimized solution for a one-month production data set shows that more than 17% of production capacity can be freed up and significant improvement in on-time delivery.

**Keywords:** Production planning and scheduling; mixed-integer linear programming; discrete event systems in manufacturing; modeling of manufacturing operations.

### **1. Introduction**

Production scheduling of large-scale, industrial processes is widely recognized as a challenging optimization task both by academic and industrial researchers. With an increasing availability of real-time process data, ever-growing computational capabilities, and the advance of digital twins, optimal production scheduling has gained new relevance. However, there is still a large gap between industrial reality and academic research. As pointed out in Harjunkski (2016), practitioners are confronted with many different challenges, when building their solutions in an industrial environment. While mixed-integer programming (MIP) is a well-established optimization technique in asset-heavy industries such as petrochemicals Castro et al. (2018), and discrete manufacturing industries such as semiconductors Qin et al. (2019) and automotive Gnoni et al. (2003), it is relatively new to industries such as specialty chemicals Borisovsky et al. (2019), pharma Sarin et al. (2014); Costa (2015), and consumer goods Sel et al. (2015); Clark et al. (2011). Alternatives such as discrete event simulation (DES) address natural drawbacks of MILP such as the curse of dimensionality but do not provide bounds on the solutions. Despite research efforts with MIP and DES Frazzon et al. (2016); Nikolopoulou and Ierapetritou (2012), real-world scheduling problems with dozens of



products, multiple production lines, changeover costs, and lot-sizing still can only be solved to sub-optimality or require massive computational resources. Herein we propose a new time-bucket mixed-integer linear programming (MILP) model for lot-sizing and scheduling problems arising in multistage production processes. The main idea behind time-bucket models based on Fleischmann and Meyr (1997) is to partition time into fixed-length macro-periods and flexible length micro-periods that lie within the macro periods. This model formulation allows to easily include a variety of typical, important real-world constraints and can be solved with moderate computer resources. We apply our approach to a real-world case study of a large-scale industrial make-and-fill process of an agrochemical production plant, and show that our modelling approach is able to capture the different characteristics of the problem efficiently. We analyse the evolution of the bounds to ensure that the exact search procedure of a MIP-solver can be truncated after short time with near-optimal solutions. In this way, a 1-month production data set can be solved with a moderate computational resources and within 1 hour solution time.

## 2. Time-bucket MILP formulation

### 2.1 Time-bucket approach

In the time-bucket MILP model mass balances, resource allocations, shift schedules, and due dates are modeled in the fixed time grid, while production quantities and changeover times are modeled in the flexible time grid. The fixed time grid consists of macroperiods, while the flexible time grid consists of microperiods. Macroperiods are fixed time intervals with uniform length, that are defined globally. As shown in Figure 2, there is a configurable number of microperiods  $s \in \mathcal{S}_t$  in each macroperiod  $t \in \mathcal{T}$ . Microperiods are flexible time intervals with variable length, that are defined per processing resource. Binary changeover variables  $y_{pp'ls}$  indicate changeovers from product  $p$  to  $p'$  on processing resource  $l$  in microperiod  $s$ . Each microperiod consists of a fixed sequence of 4 phases, set-up start, production, idle, and set-up end as shown in Figure 1.



Figure 1: Microperiods consist of setup start, production, idle and setup end phase.

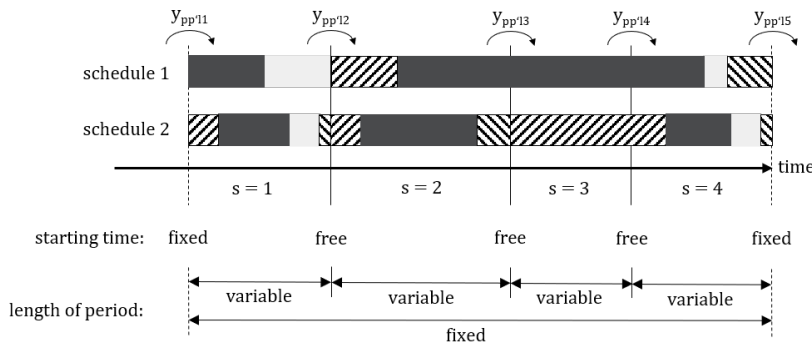


Figure 2: 2 possible schedules on a resource  $l$  in one macroperiod with 4 microperiods  $s$ .

The length of phase production is determined by the production quantity  $q_{pls}$  of a product  $p$  in microperiod  $s$  on a processing resource  $l$ . The lengths of phases set-up start and set-up end are determined by the sequence-dependent changeover times before and after the

production. The phase idle provides an optional interval for idle time. *Example.* We consider a processing resource  $l$  in a macroperiod with 4 microperiods as shown in Figure 2. In *schedule1* two production phases take place. The same time structure also allows to model 3 production phases as shown in *schedule2*. The changeover times must be taken into account using the setup start and setup end phases. If no changeover time is required the production phase can cover full microperiods. Similarly, long changeover times and idle times can be modeled. Since the microperiods are defined per processing resource  $l$  and have variable lengths the time bucket formulation provides a highly flexible time representation for MILP models.

### 2.2 MILP formulation

A crucial variable of the MILP model are production orders  $q_{p|s}$ . They express that a quantity of a product  $p \in P$  is produced on processing resource  $l \in L$  during the microperiod  $s \in S_t$  of macroperiod  $t \in T$ . Production orders seize processing resources for a minimum duration which depends on (1) the product  $p$  because of product dependent processing parameters (for example viscosity), (2) the resource  $l$  because of resource dependent processing parameters (for example milling efficiency) and (3) the quantity of the production order, which must be greater than a single batch size  $q_{min,l}$  at resource  $l$ . To describe (1) and (2) a production coefficient  $a_{p,l}$  is defined, which corresponds to the production rate of product  $p$  on the resource  $l$ . A changeover time  $ct_{pp'l}$  for subsequent products  $p$  and  $p'$  is defined for each resource  $l$ . The corresponding binary changeover variable  $y_{pp'ls}$  indicates whether a changeover from product  $p$  to  $p'$  takes place during the microperiod  $s$  on resource  $l$ . Furthermore, we distinguish intermediate products  $p \in IP \subset P$  and finished products  $\hat{p} \in FP \subset P$ . Since an intermediate product can be converted into different finished products, bill of material coefficients  $b_{p\hat{p}}$  are defined as indicator parameters. Demands  $D_{\hat{p}t}$  are defined with respect to finished product  $\hat{p}$  in macroperiod  $t$ . Similarly, the backlog quantities  $B_{\hat{p}t}$  are defined for finished product  $\hat{p}$  in macroperiod  $t$ .

### 2.3 Objective function

We suggest a 2-step, monolithical solution approach with warm start. In the first optimization step, we define the minimization of all backlog quantities  $B_{\hat{p}t}$  as the objective as shown in Equation 1.

$$\min \sum_{\hat{p} \in FP} \sum_{t \in T} B_{\hat{p}t} \tag{1}$$

In the second optimization step, the binary changeover variables  $y_{pp'ls}$  are initialized with the result of the first step. The objective of the second step is the minimization of the total changeover time  $y_{pp'ls} ct_{pp'l}$  as shown in Equation 2.

$$\min \sum_{\substack{s \in S_t \\ t \in T}} \sum_{\substack{p, p' \in P \\ l \in L}} y_{pp'ls} ct_{pp'l} \tag{2}$$

## 3. Case Study

We consider the scheduling problem of a flexible flow shop in an industrial manufacturing plant for crop protection chemicals as shown in Figure 3. The problem originates from the last two stages in the manufacturing process of pesticide chemicals, which are also known as product formulation and filling. The formulation and filling process takes place in a multi-product batch plant with 2 stages. In the first stage mixing, milling and reactions of raw materials take place, which results in bulk product. In the

second stage, bulk product is filled into final containers. The first and second stage each consist of 7 individual processing lines. We consider 50 unpackaged products and 83 finished products. We assume product-sequence and line-specific changeover costs, and minimum lot-sizes. There are 7 optional buffer tanks to decouple the two stages and maximize their utilization. Different shift schedules for the two stages determine how many lines can be operated simultaneously. The bottleneck of the process varies with production demand and the shift schedules.

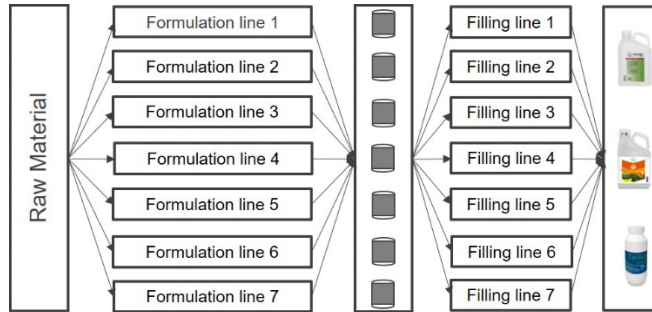


Figure 3: Schematic overview of the formulation and filling process.

## 4. Results

### 4.1 MILP Solution

We optimize the formulation and filling process for a 1-month production period. For the original demand data, we find a solution with no production backlog and 110 hours of changeover time with a remaining gap of 27.3% after a total solution time of 1.22 hours as further specified in Table 1. Optimization step 1 shows that the 1-month production demand  $D_{pt}$  can be satisfied without backlog in a makespan less than 31 days after 805 seconds solution time. Consequently a no-backlog constraint  $\Sigma B_{pt} = 0$  is enforced in the second optimization step, which minimizes the total changeover time to a value of 110 hours after 3600 seconds solution time. For the alternative demand scenarios we observe similar results.

Table 1: MILP results for different 1-month demand scenarios.

Set	Changeover (h)	Gap (%)	1st (s)	2nd (s)
Org	110	27.3	805	3600
2	130	35.1	659	3600
3	134	38.8	1033	3600
4	120	26.0	739	3600
5	120	21.8	828	3600

Since the optimality gap of the second optimization step is greater than 0 for all instances, we investigate the evolution of bounds and gap in the second optimization step over the course of 4 hours. As shown in Figure 4, the remaining gap is decreasing further for all instances after 3600 seconds. However, that is mainly due to the improvement of the lower bounds. Near-optimal solutions are found relatively quickly as indicated by the upper bounds in Figure 5. Consequently, the MIP-search can be truncated early despite relatively large remaining gaps. We conclude that the proposed MILP formulation provides near optimal solutions within 30 minutes solution time, but a proof of optimality is costly due to weak lower bounds.

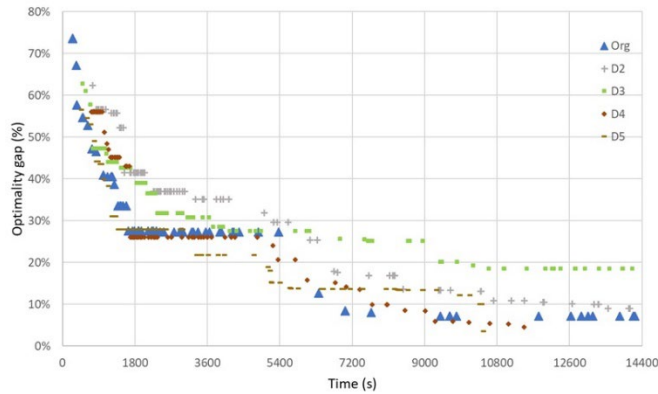


Figure 4: Evolution of the remaining gaps (%) for 5 demand scenarios.

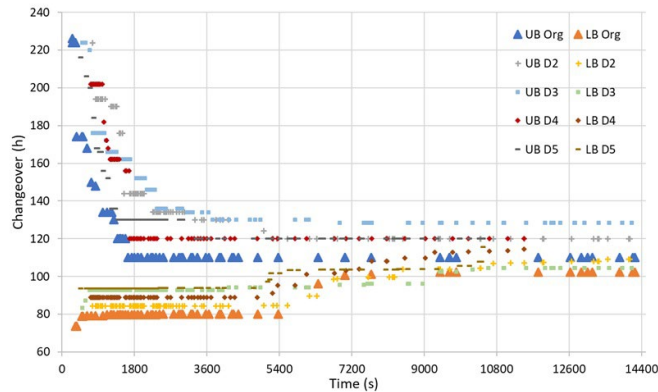


Figure 5: Evolution of the upper and lower bounds for 5 demand scenarios.

#### 4.2 DES validation

For validation the MILP result is reconciled using a validated DES model of the process. We compare the resulting MILP-DES schedule with a manually optimized DES (manual-DES) schedule. To obtain the latter, production experts generate campaign lists and refine them iteratively using the DES model. From the makespans of the MILP-DES schedule and the manual-DES in Table 2, we conclude that the MILP approach is able to generate schedules that do not violate the delivery dates at the end of the month, while the manually optimized schedule contains backlog of approximately 2%. The adherence to delivery dates and demand quantities is ensured by minimizing the total backlog in the first optimization step (see Equation 1). When comparing the occupancy times of all lines in the formulation and filling stage, we observe that MILP-DES approach reduces the number of occupancy time by a total of 1218 machine hours. Approximately 10% formulation capacity and 27% filling capacity is freed as a result of minimizing the total changeover times in the second optimization step (see Equation 2).

Table 2: DES reconciled results.

Approach	Bcklg (%)	Mkspan (d)	Form (h)	Fill (h)
MILP-DES	0	31	3309	2251
manual-DES	1.988	33.75	3674	3104

A direct comparison of the schedules shows that the MILP-DES solution in Figure 7 contains a higher number changeovers on the filling lines to synchronize with the formulation stage. As a result of this synchronization less changeover costs are incurred on the formulation lines and no backlog is produced (see Table 2). A high number of changeovers on the filling lines also is justified by the fact that they are relatively inexpensive compared to formulation changeovers because filling lines consist of a smaller number of machines and less cleaning effort is needed.

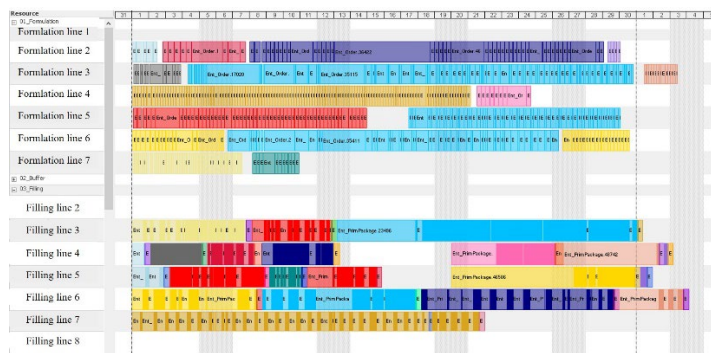


Figure 6: Manually-optimized DES schedule of 7 formulation lines (top) and 7 filling lines (bottom).

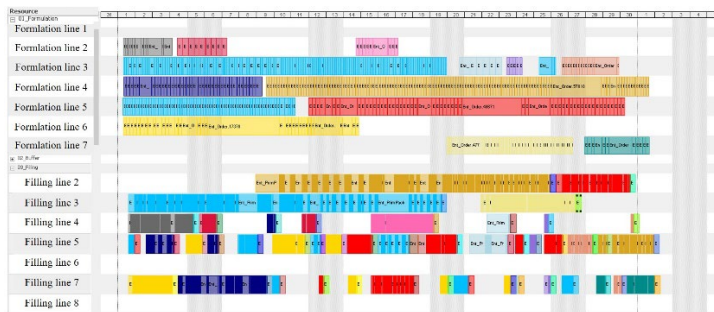


Figure 7: MILP-optimized DES schedule of 7 formulation lines (top) and 7 filling lines (bottom).

## 5. Conclusion

We showed that scheduling problems can be solved using a new time-bucket MILP formulation, which allows to include important real-world parameters such as lot-sizing, sequence dependent changeovers, buffer of intermediate bulk product, and operator schedules. We demonstrated the applicability of the time-bucket formulation with an industrial case study of a 2-stage agrochemical production process, for which we proposed a 2-step, monolithic solution procedure of the MILP model followed by a DES reconciliation step. We compared results of the MILP-DES approach with the results of

a DES-aided, manual optimization approach, and observed that counterintuitive yet better schedules are found with respect to production backlog and resource utilization. We conclude that the time-bucket formulation represents viable industrial solution, that can be adopted for similar problems.

## References

- P. Borisovsky, A. Eremeev, J. Kallrath, 06 2019. Multi-product continuous plant scheduling: combination of decomposition, genetic algorithm, and constructive heuristic. *International Journal of Production Research* 58, 1–19.
- P. M. Castro, I. E. Grossmann, Q. Zhang, 2018. Expanding scope and computational challenges in process scheduling. *Computers & Chemical Engineering* 114, 14–42, fOCAPO/CPC 2017.
- A. Clark, R. Morabito, E. A. Toso, 02 2011. Production setup-sequencing and lot-sizing at an animal nutrition plant through atsp subtour elimination and patching. *J. Scheduling* 14, 119.
- A. Costa, 2015. Hybrid genetic optimization for solving the batch-scheduling problem in a pharmaceutical industry. *Computers & Industrial Engineering* 79, 130–147. B. Fleischmann, H. Meyr, Mar. 1997. The general lotsizing and scheduling problem 19 (1), 11–21.
- E. M. Frazzon, A. Albrecht, P. A. Hurtado, 2016. Simulation-based optimization for the integrated scheduling of production and logistic systems. *IFAC-PapersOnLine* 49 (12), 1050–1055, 8th IFAC Conference on Manufacturing Modelling, Management and Control MIM 2016.
- M. Gnoni, R. Iavagnilio, G. Mossa, G. Mummolo, A. Di Leva, 2003. Production planning of a multi-site manufacturing system by hybrid modelling: A case study from the automotive industry. *International Journal of Production Economics* 85 (2), 251–262, supply Chain Management.
- I. Harjunkoski, 2016. Deploying scheduling solutions in an industrial environment. *Computers & Chemical Engineering* 91, 127–135, 12th International Symposium on Process Systems Engineering & 25th European Symposium of Computer Aided Process Engineering (PSE-2015/ESCAPE-25), 31 May - 4 June 2015, Copenhagen, Denmark.
- A. Nikolopoulou, M. Ierapetritou, 12 2012. Hybrid simulation based optimization approach for supply chain management. *Computers & Chemical Engineering* 47, 183–193.
- W. Qin, Z. Zhuang, Y. Liu, O. Tang, 2019. A two-stage ant colony algorithm for hybrid flow shop scheduling with lot sizing and calendar constraints in printed circuit board assembly. *Computers & Industrial Engineering* 138, 106115.
- S. C. Sarin, H. D. Sherali, L. Liao, 2014. Primary pharmaceutical manufacturing scheduling problem. *IIE Transactions* 46 (12), 1298–1314.
- C. Sel, B. Bilgen, J. Bloemhof-Ruwaard, J. van der Vorst, 2015. Multi-bucket optimization for integrated planning and scheduling in the perishable dairy supply chain. *Computers & Chemical Engineering* 77, 59–73.



# Study of Different Formulations for the Multiperiod Blending Problem Applied to Lithium Recovery from Produced Water

Daniel Ovalle<sup>a</sup>, Joshua L. Pulsipher<sup>a</sup>, Camilo Gomez<sup>b</sup>, Jorge M. Gomez<sup>c</sup>,  
Carl D. Laird<sup>a</sup>, Markus G. Drouven<sup>d</sup>, Ignacio E. Grossmann<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA*

<sup>b</sup>*Department of Industrial Engineering, Universidad de Los Andes, Bogota, 111711, Colombia*

<sup>c</sup>*Department of Chemical Engineering, Universidad de Los Andes, Bogota, 111711, Colombia*

<sup>d</sup>*National Energy Technology Lab, Pittsburgh, PA 15236, USA*

*grossmann@cmu.edu*

## Abstract

Oil and gas production wastewater (i.e., produced water) may contain appreciable concentrations of rare-earth elements and critical minerals (REE/CMs), such as lithium, that can be recovered. However, each individual produced water source may have insufficient concentration or volume to meet the economic and operating requirements of the recovery facility. Therefore, there is the need to appropriately blend multiple sources to meet recovery and water reuse demands. Optimal stream mixing operation planning can be posed as a multiperiod blending problem (MPBP) to provide quantity and quality guarantees over time. We present several formulations to solve the MPBP for the recovery of REE/CMs from produced water and propose a decomposition approach that leverages strategies in general disjunctive programming to enhance its performance. We compare these proposed formulations/strategies via two illustrative case studies on recovering lithium from a network of produced water sources.

**Keywords:** Multiperiod Blending, Lithium Recovery, Mixed-Integer Nonlinear Programming, Decomposition

## 1. Introduction

Wastewater streams from oil and gas production (i.e., produced water) are expected to surpass 60 million barrels of water per day by 2030 in the U.S. alone (Wright, 2022). The treatment needed to mitigate the environmental impact of these wastewater streams is expensive due to high salinity and contaminant levels (Gaustad et al., 2021). Hence, a significant portion of produced water in the U.S. is simply disposed of via underground injection.

Notably, however, produced water sources may contain appreciable concentrations of critical minerals (CMs) such as lithium which are crucial for manufacturing electronics, pharmaceuticals, batteries, renewable energy generators, and more (Quillinan et al., 2018). Moreover, establishing a sustainable REE/CM supply chain is a critical concern of many industrial and governmental stakeholders. Thus, the recovery of REE/CMs from produced water has the potential to add sufficient economic value to incentivize the use of treatment technologies that mitigate the impact of produced water on the environment.

Operating a water network for the recovery of REE/CMs from multiple produced water sources is highly complex since REE/CM concentrations vary dynamically and geographically, there are limited storage locations, and treatment/recovery facilities have strict inlet-flow requirements (Gaustad et al., 2021). Such a system can be modeled as a classic multiperiod blending problem where source



streams with time-varying compositions are blended in pooling tanks to satisfy quality and quantity demand requirements. Classic pooling formulations such as the one proposed in Haverly (1978) use bilinear terms to track compositions along the network. However, such formulations are often difficult to solve due to their inherent nonconvexity (Gounaris and Floudas, 2008). This difficulty is compounded by multiperiod models that incorporate binary variables.

In this work, we investigate several multiperiod blending problem (MPBP) formulations and solution strategies to optimally plan the delivery of lithium-rich streams for recovery and water reuse in development activities. In particular, we adapt the direct, generalized disjunctive programming, and decomposition MPBP approaches as described in Kolodziej et al. (2013) and Lotero et al. (2016) to a representative produced water network case study. It is shown that such strategies are key to guiding the treatment of a multi-enterprise produced water blending and determine whether such a system is feasible and economically viable.

## 2. Problem Formulation and Solution Strategies

### 2.1. Basic Problem Setup and Nomenclature

The multiperiod blending problem (MPBP) is defined over a set of supply tanks  $\mathcal{S}$ , blending tanks  $\mathcal{B}$ , and demand tanks  $\mathcal{D}$  interconnected by a set of edges  $\mathcal{A}$ . The problem seeks to determine the optimal mixing schedule over a discretized time horizon  $\mathcal{T}$  that maximizes the operation profit, while meeting flow and concentration specifications for each component  $q \in \mathcal{Q}$ . The initial conditions for each tank  $n \in \mathcal{N} := \mathcal{S} \cup \mathcal{B} \cup \mathcal{D}$  are specified via the initial level of every tank  $I_n^0$  and its corresponding compositions  $C_{qn}^0$ . The incoming flow  $F_{st}^{IN}$  at each supply  $s$  is known with an incoming composition  $C_{qst}^{IN}$  that varies over  $t$ . The nodal inventories and flows are restricted by tank and pipeline capacities, which are denoted  $[I_n^L, I_n^U]$  and  $[F_{nn'}^L, F_{nn'}^U]$ , respectively. Required demand varies with time and needs to satisfy specifications on the flow  $[FD_{dt}^L, FD_{dt}^U]$  and concentration  $[C_{qd}^L, C_{qd}^U]$ . The MPBP formulations considered here assume that blending tanks do not operate at steady-state. Mixing requires that the tanks be charged at one period of time and discharged at another period (Lotero et al., 2016).

### 2.2. Direct MIQCP Formulation

We model the MPBP directly as a mixed-integer quadratically constrained program (MIQCP). The MIQCP formulation follows from extending steady-state pooling formulations to incorporate multiperiod scheduling. Here, binary variables  $x_{nn't} \in \{0,1\}$  are introduced to indicate the existence of flow at each time period. This addition enables us to enforce a minimum flow in the pipeline, when the flow exists. We obtain the MIQCP formulation by adapting the formulation presented by Kolodziej et al. (2013) to handle time-varying concentration in the supply:

$$\max \sum_{t \in \mathcal{T}} \left[ \sum_{(n,d) \in \mathcal{A}} \beta_d^T F_{ndt} - \sum_{(s,n) \in \mathcal{A}} \beta_s^T F_{snt} - \sum_{(n,n') \in \mathcal{A}} (\alpha_{nn'}^N x_{nn't} + \beta_{nn'}^N F_{nn't}) \right] \quad (1a)$$

$$\text{s.t. } F_{nn'}^L x_{nn't} \leq F_{nn't} \leq F_{nn'}^U x_{nn't}, \quad (n, n') \in \mathcal{A}, t \in \mathcal{T} \quad (1b)$$

$$C_{qd}^L - M(1 - x_{bdt}) \leq C_{qbt-1} \leq C_{qd}^U + M(1 - x_{bdt}), \quad q \in \mathcal{Q}, (b, d) \in \mathcal{A}, t \in \mathcal{T} \quad (1c)$$

$$C_{qd}^L - M(1 - x_{sdt}) \leq C_{qst}^{IN} \leq C_{qd}^U + M(1 - x_{sdt}), \quad q \in \mathcal{Q}, (s, d) \in \mathcal{A}, t \in \mathcal{T} \quad (1d)$$

$$I_{st} = I_{st-1} + F_{st}^{IN} - \sum_{(s,n) \in \mathcal{A}} F_{snt}, \quad s \in \mathcal{S}, t \in \mathcal{T} \quad (1e)$$

$$I_{bt} = I_{bt-1} + \sum_{(n,b) \in \mathcal{A}} F_{nbt} - \sum_{(b,n) \in \mathcal{A}} F_{bnt}, \quad b \in \mathcal{B}, t \in \mathcal{T} \quad (1f)$$

$$I_{dt} = I_{dt-1} + \sum_{(n,d) \in \mathcal{A}} F_{ndt} - FD_{dt}, \quad d \in \mathcal{D}, t \in \mathcal{T} \quad (1g)$$

$$x_{nbt} + x_{bn't} \leq 1, \quad (n, b), (b, n') \in \mathcal{A}, t \in \mathcal{T} \quad (1h)$$

$$I_{bt}C_{qbt} = I_{bt-1}C_{qbt-1} + \sum_{(s,b) \in \mathcal{A}} F_{sbt}C_{qst}^{IN} + \sum_{(b',b) \in \mathcal{A}} F_{b'bt}C_{qb't-1} - \sum_{(b,n) \in \mathcal{A}} F_{bnt}C_{qbt-1}, \quad q \in \mathcal{Q}, b \in \mathcal{B}, t \in \mathcal{T} \quad (1i)$$

$$I_n^L \leq I_{nt} \leq I_n^U, \quad n \in \mathcal{N}, t \in \mathcal{T} \quad (1j)$$

$$C_q^L \leq C_{qbt} \leq C_q^U, \quad q \in \mathcal{Q}, b \in \mathcal{B}, t \in \mathcal{T} \quad (1k)$$

$$FD_{dt}^L \leq FD_{dt} \leq FD_{dt}^U, \quad d \in \mathcal{D}, t \in \mathcal{T} \quad (1l)$$

$$F_{nn't} \geq 0, x_{nn't} \in \{0, 1\}, \quad (n, n') \in \mathcal{A}, t \in \mathcal{T} \quad (1m)$$

Where  $\beta_d^T$  is the unit profit of demand  $d$  satisfied,  $\beta_s^T$  is the unit cost of supply stream  $s$ ,  $\beta_{nn'}^N$  is the unit transportation cost,  $\alpha_{nn'}^N$  is the fixed transportation cost,  $F_{st}^{IN}$  is the incoming flow from supply tank  $s$  at time period  $t$ ,  $F_{nn't}$  is the flow between node  $n$  and node  $n'$  at time  $t$ ,  $FD_{dt}$  is the flow directed to demand tank  $d$  at period  $t$ , and  $M \in \mathbb{R}$  is a sufficiently large big-M constant. For simplicity in presentation, the edge set  $\mathcal{A}$  can refer to a subset of edges depending on the indices being used where  $n$  refers to general nodes and  $s$ ,  $b$ , and  $d$  refer to supply, blending, and demand nodes, respectively. The formulation allows flows to go directly from the supply to the demands without blending if and only if they meet the required specifications. Also note that Equation (1h) enforces that blending tanks cannot be charged, mixed, and discharged simultaneously. The key complicating factor of Problem (1) is that Equation (1i) involves bilinear terms.

### 2.3. Generalized Disjunctive Programming Formulation

Generalized disjunctive programming (GDP) is a framework for naturally posing mathematical programs that incorporate symbolic logic relationships between the variables (Grossmann and Trespalcios, 2013). In the MPBP, the bilinear mass balance shown in (1i) only needs to be considered when tanks are charging. Hence, this nonconvex equation can be omitted when a tank is discharging. Using GDP, we can define  $Y_{bt} \in \{\text{True}, \text{False}\}$  to specify whether or not, at a given period, tank  $b$  is charging ( $Y_{bt} = \text{True}$ ) or discharging ( $Y_{bt} = \text{False}$ ). With this, we can replace Equation (1i) in Problem (1) with the following disjunction that accounts for the operational mode of the tank:

$$\left[ \begin{array}{l} I_{bt} = I_{bt-1} + \sum_{(n,b) \in \mathcal{A}} F_{nbt} \\ I_{bt}C_{qbt} = I_{bt-1}C_{qbt-1} + \sum_{(s,b) \in \mathcal{A}} F_{sbt}C_{qst}^{IN} \\ \quad + \sum_{(b',b) \in \mathcal{A}} F_{b'bt}C_{qb't-1}, \quad q \in \mathcal{Q} \end{array} \right] \vee \left[ \begin{array}{l} -Y_{bt} \\ I_{bt} = I_{bt-1} - \sum_{(b,n) \in \mathcal{A}} F_{bnt} \\ C_{qbt} = C_{qbt-1}, \quad q \in \mathcal{Q} \end{array} \right] \quad (2)$$

for each  $b \in \mathcal{B}$  and  $t \in \mathcal{T}$ . Note that bilinear terms are only considered in the left disjunct. Also, the mass balances in both disjunctions have fewer terms since they leverage the tank operational mode to only consider active flows.

Formulation (1) has other constraints that enforce logical implications on variables, which are amenable for GDP reformulation. For instance, Equation (1b) activates the flow  $F_{nn't}$  only if  $x_{nn't} = 1$  during that period. Hence we can also redefine the flow existence as a Boolean variable  $X_{nn't} \in \{\text{True}, \text{False}\}$  and rewrite Equation (1b) with the following disjunction:

$$\left[ \begin{array}{l} X_{nbt} \\ F_{nb}^L \leq F_{nbt} \leq F_{nb}^U \end{array} \right] \vee \left[ \begin{array}{l} -X_{nbt} \\ F_{nbt} = 0 \end{array} \right], \quad (n, b) \in \mathcal{A}, t \in \mathcal{T}. \quad (3)$$

Equations (1c) and (1d) can be reformulated via similar GDP disjunctions. Moreover, Equation (1h) can be reformulated using GDP logical propositions:

$$X_{nbt} \Rightarrow Y_{bt}, \quad (n, b) \in \mathcal{A}, t \in \mathcal{T} \quad (4a)$$

$$X_{bnt} \Rightarrow \neg Y_{bt}, \quad (b, n) \in \mathcal{A}, t \in \mathcal{T} \quad (4b)$$

All these GDP equations are substituted into Formulation (1) to yield the GDP-based MPBP, which generalizes the GDP formulation proposed in (Lotero et al., 2016) to account for time-varying concentrations (prevalent in produced water constituents). Note that this formulation extends the one presented in (1) to account for the tank operating mode to yield fewer bilinearities after operating modes are determined.

#### 2.4. GDP with Redundant Constraints (RC)

Typically, mixed-integer solvers employ continuous relaxations of the problem (by relaxing variable integrality) to iteratively obtain an optimal solution. Tight relaxations often provide high-quality solutions relative to the mixed-integer formulation, which can accelerate solver convergence. One way to tighten relaxations is to include additional constraints therefore redundant that provide no additional modeling information but further restrict the feasible region.

We derive redundant constraints for our GDP formulation by tracking component flow origins:

$$F_{mn't} = \sum_{r \in \mathcal{R}} \tilde{F}_{rnm't}, \quad (n, n') \in \mathcal{A}, t \in \mathcal{T} \quad (5a)$$

$$I_{bt} = \sum_{r \in \mathcal{R}} \tilde{I}_{rbt}, \quad b \in \mathcal{B}, t \in \mathcal{T} \quad (5b)$$

where  $\mathcal{R} := \mathcal{S} \cup \widehat{\mathcal{B}}$  is the set of possible initial origins,  $\widehat{\mathcal{B}}$  is the set of blending tanks with nonzero initial inventory, and  $\tilde{F}$ ,  $\tilde{I}$ , and  $\tilde{C}^0$  are the flows, inventories, and initial composition identified by origin, respectively. Following Lotero et al. (2016), we employ other redundant constraints to strengthen the GDP formulation, and we refer the reader to that work for more details and analysis on the tightness of the relaxation.

#### 2.5. Two-Stage MILP-MIQCP Decomposition

We can pose a two-stage decomposition to the above GDP formulations for complex MPBPs that incur high computational cost. We define an upper-level problem that omits the bilinear species balances to obtain a linear formulation that provides an upper bound  $UB$  on the optimal solution and candidate operating states of  $Y_{bt}$ . Then we define a lower-level problem that fixes  $Y_{bt}$  in the full formulation to the values provided by the upper-level problem and removes the bilinear terms corresponding to idle/discharging tanks. When feasible, the lower-level problem provides a lower bound on the optimal solution. Integer cuts are added to the upper-level problem after each run of the lower-level problem as shown in (Lotero et al., 2016):

$$Z \leq -(UB - Z^i) \left( \sum_{b \in \mathcal{B}, t \in \mathcal{T} | \hat{y}_{bt}^i = 1} y_{bt} - \sum_{b \in \mathcal{B}, t \in \mathcal{T} | \hat{y}_{bt}^i = 0} y_{bt} \right) + \quad (6a)$$

$$(UB - Z^i) \left( \sum_{b \in \mathcal{B}, t \in \mathcal{T}} (\hat{y}_{bt}^i) - 1 \right) + UB, \quad i \in \mathcal{I}_O$$

$$\sum_{b \in \mathcal{B}, t \in \mathcal{T} | \hat{y}_{bt}^i = 1} (1 - y_{bt}) + \sum_{b \in \mathcal{B}, t \in \mathcal{T} | \hat{y}_{bt}^i = 0} y_{bt} \geq 1, \quad i \in \mathcal{I}_F \quad (6b)$$

where  $Z$  is the objective of the upper-level problem,  $y_{bt}$  are the binary variables associated with  $Y_{bt}$ ,  $\hat{y}_{bt}$  are the fixed values of the tank modes at the  $i^{th}$  iteration, and  $\mathcal{I}_O$  and  $\mathcal{I}_F$  contain the feasible and infeasible iteration indices, respectively. Here, Equation (6a) is added based on a feasible lower-level solution, and Equation (6b) is added based on infeasible solutions. Iterating between upper- and lower-level problems leads to convergence to a globally optimal solution of the MPBP. The gap of the decomposition is calculated between the upper and lower bounds ( $LB$  and  $UB$ ) as shown in Equation (7) where  $\epsilon$  is a small tolerance. A formal analysis is provided in (Lotero et al., 2016).

$$\text{Gap} = \frac{UB - LB}{LB + \epsilon} \quad (7)$$

### 3. Case Study

We compare the formulations in Section 2 via two case studies focused on delivering a lithium rich stream from a network of produced water sources. Case 1 is a smaller system composed by a subset of the tanks considered in Case 2. Produced water with varied lithium concentrations is produced at different wells and a source of fresh water is available. Several demand nodes are considered with their respective schedules for feed quality and quantity. We choose parameters based on those reported by Dworzanowski (2019) and Figure 1 details the topology of Case 2.

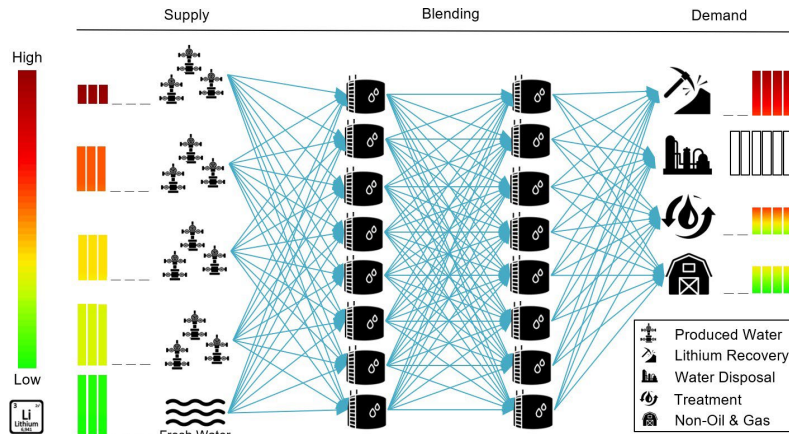


Figure 1: Topology of Case Study 2. Edges that connect sources with demands directly are considered in the case study but are not shown in the figure for simplicity.

GDP problems are often solved via big-M reformulation (BM) or Hull reformulation (HR) (Grossmann and Trespalcacios, 2013; Lee and Grossmann, 2000). Previous work conducted by Lotero et al. (2016) and Ovalle Varela et al. (2021) only use BM reformulations. We explore solving these formulations using both BM and HR to study their impact on solution performance. We implement all approaches in Pyomo.GDP on a Linux machine with 8 Intel® Xeon® Gold 6234 CPUs running at 3.30 GHz with 128 hardware threads and 1 TB of RAM with Ubuntu. We use Gurobi v9.5.1 and BARON v22.7.23 as appropriate to solve all formulations. We impose a target optimality gap of 1% and 3% for case studies 1 and 2, respectively; moreover, we set a wall-time of 3600s.

Table 1: Approach comparison for the lithium recovery case studies

Upper-level	Lower-level	Case 1		Case 2	
		Time [s]	Gap[%]	Time [s]	Gap[%]
MIQCP	-	843.87	0.77	3,600	20.02
GDP (BM)	-	3,600	61.10	3,600	47.91
GDP (HR)	-	3,600	-	3,600	-
RC (BM)	-	69.05	1.00	3,600	26.33
RC (HR)	-	3,600	1.30	3,600	-
RC (BM)	RC (BM)	160.57	<b>0.03</b>	3,600	-
RC (HR)	RC (BM)	<b>12.54</b>	0.08	<b>54.1</b>	<b>1.08</b>

Table 1 shows the solution time and final optimality gap of each solution method and case study. The columns corresponding to the levels indicate the selected formulation and GDP reformulation (BM or CH) when applicable. Rows without an entry in the lower-level column are solved directly

without the two-stage decomposition. Note that the two-stage decomposition using RC (BM) and RC (BM) was originally proposed by Lotero et al. (2016), and decomposition using RC (HR) and RC (BM) is proposed in this work. The results from Case 2 suggest that the monolithic formulations readily become intractable, which justifies the use of decomposition strategies. In fact, only our proposed decomposition is able to achieve an optimal solution for Case 2 within the wall-time. Hence, we observe that the choice of GDP reformulation strategy can significantly affect solution performance.

#### 4. Conclusions

The multiperiod blending problem is a nonconvex mixed-integer quadratically constrained program that is challenging to solve for real life applications such as lithium recovery from produced water. We observe that the previous solution approaches considered in this work are extremely expensive computationally for this problem class. Moreover, our adaptation of the two-stage decomposition approach to use hull reformulations significantly enhances scalability. These results motivate further investigation into how combinations of possible formulations, GDP solution techniques, fixing strategies, and solver tunings can accelerate the convergence of the two-stage decomposition strategy considered in this work. Future work also includes integrating these formulations within the PARETO framework (Drouven et al., 2022) for REE-CM recovery extensions.

**Acknowledgements:** We gratefully acknowledge support from the U.S. Department of Energy, Office of Fossil Energy and Carbon Management, through the Environmental Prudent Stewardship Program.

**Disclaimers:** This project was funded by the United States Department of Energy, National Energy Technology Laboratory, in part, through a support contract. Neither the United States Government nor any agency thereof, nor any of its employees, makes any warranty, expressor implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

#### References

- M. G. Drouven, A. J. Caldéron, M. A. Zamarripa, K. Beattie, 2022. Pareto: An open-source produced water optimization framework. *Optimization and Engineering*, 1–21.
- M. Dworzanowski, 2019. Preliminary economic assessment of LANXESS smackover project. URL <https://www.sec.gov/Archives/edgar/data/1537137/000119312521204052/d194326dex991.htm>
- G. Gaustad, E. Williams, A. Leader, 2021. Rare earth metals from secondary sources: Review of potential supply from waste and byproducts. *Resources, Conservation and Recycling* 167, 105213.
- C. Gounaris, C. Floudas, 2008. Convexity of products of univariate functions and convexification transformations for geometric programming. *Journal of Optimization Theory and Applications* 138 (3), 407–427.
- I. E. Grossmann, F. Trespalacios, 2013. Systematic modeling of discrete-continuous optimization models through generalized disjunctive programming. *AIChE Journal* 59 (9), 3276–3295.
- C. A. Haverly, 1978. Studies of the behavior of recursion for the pooling problem. *ACM SIGMAP bulletin* (25), 19–28.
- S. P. Kolodziej, I. E. Grossmann, K. C. Furman, N. W. Sawaya, 2013. A discretization-based approach for the optimization of the multiperiod blend scheduling problem. *Computers & Chemical Engineering* 53, 122–142.
- S. Lee, I. E. Grossmann, 2000. New algorithms for nonlinear generalized disjunctive programming. *Computers & Chemical Engineering* 24 (9–10), 2125–2141.
- I. Lotero, F. Trespalacios, I. E. Grossmann, D. J. Papageorgiou, M.-S. Cheon, 2016. An MILP-MINLP decomposition method for the global optimization of a source based model of the multiperiod blending problem. *Computers & Chemical Engineering* 87, 13–35.
- D. Ovalle Varela, et al., 2021. Systematic solving study for the optimization of the multiperiod blending problem: a multiple mathematical approach solution guide. Universidad de los Andes.
- S. Quillinan, C. Nye, M. Engle, T. T. Bartos, G. Neupane, J. Brant, D. Bagdonas, T. McLing, J. F. McLaughlin, E. Phillips, et al., 2018. Assessing rare earth element concentrations in geothermal and oil and gas produced waters: A potential domestic source of strategic mineral commodities. Univ. of Wyoming, Laramie, WY (United States).
- B. Wright, 2022. The produced water conundrum grows across unconventional. *Journal of Petroleum Technology* 74 (01), 38–45.

# State and Parameter Estimation in Dynamic Real-time Optimization with Closed-Loop Prediction

Jose Matias,<sup>a</sup> Christopher L.E. Swartz,<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, McMaster University, 1280 Main St W,  
Hamilton, Canada, {assumpcj, swartz}@mcmaster.ca*

## Abstract

In this paper, we propose the integration of a state and parameter estimation step with closed-loop dynamic real-time optimization (CL-DRTO). Different from other dynamic real-time optimization strategies, CL-DRTO uses a dynamic model that considers the closed-loop response of the plant under the action of a controller. So far, plant feedback has been incorporated into CL-DRTO using an additive noise paradigm by a bias updating strategy. In this work, we perform the feedback step in CL-DRTO updating model states and parameters via Moving Horizon Estimation. As an additional contribution, we compare CL-DRTO with two other production optimization schemes, Steady-state Real-time Optimization (SRTO) and Real-time Optimization with Persistent Parameter Adaptation (ROPA). In the latter, a dynamic model is used for only the state and parameter estimation step in SRTO, eliminating the requirement for the plant to be at steady state prior to the economic optimization execution. We show that CL-DRTO outperforms the other two schemes in terms of economic and control performance when applied to systems under control performance limitations.

**Keywords:** real-time optimization; dynamic real-time optimization; closed-loop; state estimation; parameter estimation.

## 1. Main Text

Dynamic real-time Optimization (DRTO) is a useful tool for improving plant economic performance in an increasingly dynamic market environment. DRTO computes optimal trajectories for the plant decision variables by maximizing an economic performance index subject to the system dynamics and operational/quality constraints. The optimal decisions are then sent to the plant controllers, which are responsible for implementing the computed values. In contrast to the traditional Steady-state RTO (SRTO) (Chen and Joseph, 1987), DRTO utilizes a dynamic plant model, making it better suited for systems that exhibit slow dynamics and/or frequent transitions.

The dynamic models employed at the DRTO level are often derived under a “perfect control” assumption. That is, the control response time for disturbance rejection and setpoint tracking is assumed negligible when compared to the DRTO sampling time. However, depending on the controller tuning and the system characteristics (e.g., long dead time, inverse response, etc.), this assumption may not hold. To address this issue, Jamaludin and Swartz (2017) proposed a closed-loop DRTO (CL-DRTO) scheme in which future control actions of an underlying constrained MPC system are considered. By rigorously predicting the interaction between controllers and system response, CL-DRTO improves the system performance compared to the open-loop counterpart.

One important challenge when applying CL-DRTO is how to use plant information to update the model. In the previous CL-DRTO implementations, the plant condition is represented in the model by an additive model uncertainty. Here, the difference between the measurements at the current time and the predicted output by the model is computed as a "bias", which is then added to the model outputs and kept constant over the CL-DRTO prediction horizon (similar to the typical DMC/QDMC updating scheme). Despite presenting good results for a variety of case studies, this feedback strategy is not flexible in terms of different noise structures, which is critical for the model updating step (Kolås et al. 2008). For example, if a certain parameter drifts away from its nominal value, simply adding a bias to the model outputs may not suffice to rigorously represent the plant condition in the model.

Therefore, we propose the use of a Moving Horizon Estimator (MHE) for updating the model in CL-DRTO. With this model updating formulation, we can represent the model uncertainty using time-varying parameters by augmenting the states with a parameter vector. Therefore, we are not limited to an additive noise representation, which may lead to an improper model updating depending on the nature of the model uncertainty (Kolås et al. 2008). To benchmark this new model updating strategy, we compare CL-DRTO + MHE with two other production optimization schemes that also deal with model uncertainty via online parameter estimation, namely Steady-state Real-time Optimization (SRTO) and Real-time Optimization with Persistent Parameter Adaptation (ROPA) (Matias and Le Roux, 2018). In ROPA, a dynamic model is used for only the state and parameter estimation step in a SRTO, which eliminates the requirement for the plant to be at steady state prior to an RTO execution. We show that the combination of CL-DRTO and MHE yields a superior performance over the other two schemes in terms of economic and control performance when the controllers are detuned or tuned too aggressively.

## 2. Case Study: Distillation Column

The distillation column described in Skogestad (1997) is used as the case study. The column has 40 theoretical stages and the goal is to separate an ideal binary mixture and obtain a top product with 99% purity. A diagram of the distillation column and of the

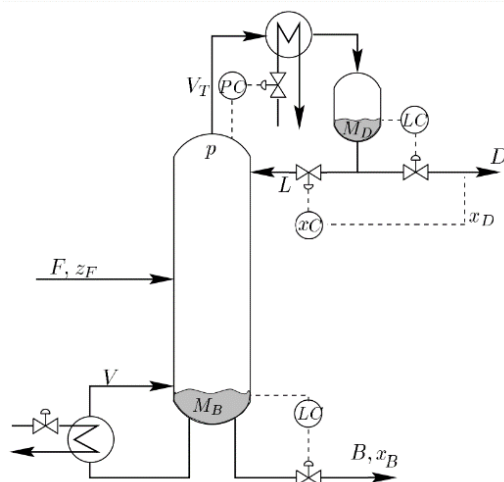


Figure 1: Distillation column diagram adapted from Skogestad (1997)

control structure is shown Figure 1. The model follows these assumptions: binary mixture; constant relative volatility; equilibrium on all stages; total condenser; constant molar flows; no vapor holdup; and linearized liquid dynamics.

For the control structure, we assume perfect pressure control, i.e. constant pressure, and we use the LV-configuration, where the reboiler and condenser levels are controlled using the bottoms (B) and distillate (D) flows, respectively. For composition control, we use the one-point control strategy, where the reflux (L) is used to control the top composition  $x_D$  and

the bottom composition is left uncontrolled. Hence, the boilup ( $V$ ) is the only degree of freedom (independent variable) of the system.

Ideally, these controllers should be tuned such that they yield a fast control loop response. However, in practice, this fast-responding tuning may lead to aggressive behavior to disturbances and setpoint changes. Consequently, the controlled variable may oscillate and overshoot before settling around the setpoint, which is undesired from a practical point of view. One alternative is to detune the controllers and change the gains to obtain a slower and smooth response, i.e., less oscillatory, and more tolerant to changes in process conditions. The problem is that this strategy goes against the “perfect control” assumption at the production optimization layer.

### 3. Simulation Setup

We tuned the sump and condenser level controllers for tight control. On the other hand, the parameters of the top composition control were chosen to achieve a smooth and slow response subject to acceptable disturbance rejection. The controller sampling time is set to  $\Delta t_{cont} = 5$  s.

The plant is affected by disturbances (Figure 2) in the feed fraction ( $z_F$ ), which we assume as unknown, and in the relative volatility ( $\alpha$ ) of the mixture. The rationale behind the disturbances is the following: the feed fraction may vary depending on the unit load and its actual value may be hard to determine exactly due to an upstream tank layering, or the variations of the raw material. The oscillation in  $\alpha$  may be connected to the vapor-liquid equilibrium and caused by changes in the column pressure setpoint.

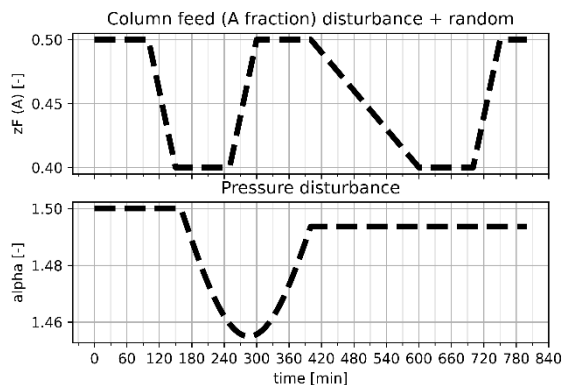


Figure 2: Disturbance scenarios

For the economic optimization problem, we want to maximize the difference between the product revenues, considering their purity, and the utility costs (steam for the reboiler, and cooling water for the condenser). The heating and cooling costs are assumed to be proportional to the flows. The objective function can then be written as:

$$\Phi = c_1 D x_D + c_2 B x_B - (c_3 (D + L) + c_4 V) \quad (1)$$

in which,  $c_i$  are the respective prices and costs in  $[\$/\text{mol}]$ , and  $c = [1.5, 0.2, 0.03, 0.02] \frac{\$}{\text{kmol}}$ . Since the top stream contains the product of interest, we specify  $c_1 \gg c_2$ . For the economic optimization problem, we also want to consider a constraint on the top purity ( $x_D \geq 0.99$ ).



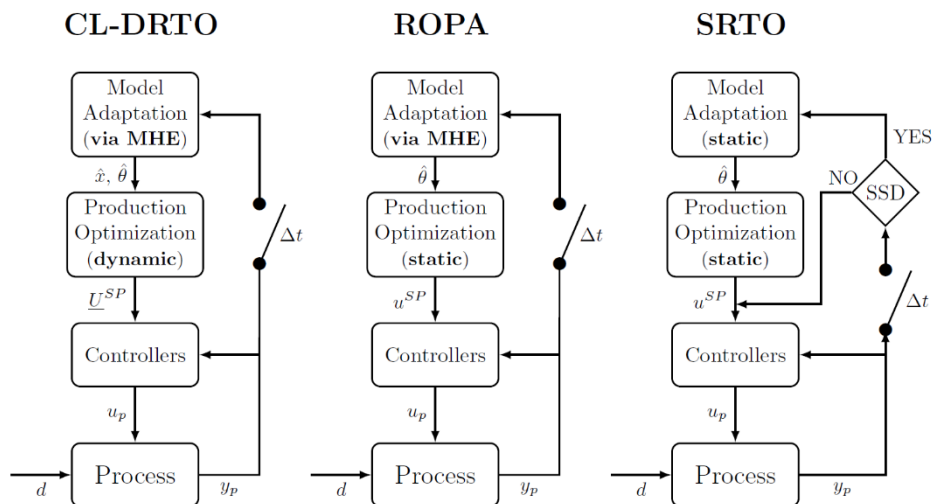


Figure 3: Block diagram of the production optimization methods

#### 4. Production Optimization Methods

The production optimization methods studied in this paper are implemented as shown in Figure 3. All methods are executed at the same rate ( $\Delta t = 15$  min), which is approximately the open-loop composition response time to changes in internal flows. Then, SRTO and ROPA optimize the future steady-state behavior of the plant and send this information to the controllers as the optimal setpoints ( $u^{SP}$ ), which are kept constant until the next cycle execution. On the other hand, CL-DRTO optimizes the behavior of the system (plant + controllers) for the next 75 minutes, discretized in intervals of 3 minutes. Next, a sequence of 5 setpoints ( $\underline{U}^{SP}$ ) is sent to the controllers. Similar to the other two methods,  $\underline{U}^{SP}$  is updated every 15 minutes.

In the SRTO case, the steady state detection (SSD) was performed based on the comparison of total variance of a signal in a fixed length window of measurements with the total variance of subsequent measurements within the same window length (see Tool A in Camara et al. (2016)). If this ratio of variances (R-statistics) is below a chosen threshold, we assume that the two variances come from the same distribution and, consequently, the process is at steady-state. We use a composition in the rectifying section as a steady-state representative measurement. The other system measurements  $y_p$  are 4 compositions along the distillation column and the sump and reboiler levels. For the formulation of the Production Optimization (dynamic and static) and Model Adaptation (static) block, please refer to Matias et al. (2022). The model adaptation via MHE is shown next.

##### 4.1. Dynamic State and Parameter Estimation via Moving Horizon Estimator (MHE)

We assume that the uncertainty structure is known, and the plant-model mismatch can be compensated by estimating the model parameters  $\theta$  (here,  $z_F$  and  $\alpha$ ), which represent the disturbances  $d$  from Figure 2. Hence, the MHE formulation can be written as:

$$\begin{aligned}
& \min_{\theta_k, x_k} \left( \begin{bmatrix} x_{t_0 - N_e} \\ \theta_{t_0 - N_e} \end{bmatrix} - \begin{bmatrix} \hat{x}_{t_0 - N_e} \\ \hat{\theta}_{t_0 - N_e} \end{bmatrix} \right)^T \Pi_{x, \theta} \left( \begin{bmatrix} x_{t_0 - N_e} \\ \theta_{t_0 - N_e} \end{bmatrix} - \begin{bmatrix} \hat{x}_{t_0 - N_e} \\ \hat{\theta}_{t_0 - N_e} \end{bmatrix} \right) + \\
& \sum_{k=t_0 - N_e}^{t_0} (y_{p,k} - y_k)^T Q_y (y_{p,k} - y_k) + \sum_{k=t_0 - N_e + 1}^{t_0} (\theta_k - \theta_{k-1})^T Q_\theta (\theta_k - \theta_{k-1}) \\
& \text{subjected to:} \\
& x_{k+1} = F(x_k, u_{p,k}, \theta_k) \quad k = \{t_0 - N_e, \dots, t_0 - 1\} \\
& y_k = Hx_k \quad k = \{t_0 - N_e, \dots, t_0\} \\
& x_k \in \mathcal{X}, \quad \theta_k \in \Theta \quad k = \{t_0 - N_e, \dots, t_0\}
\end{aligned} \tag{2}$$

where, the subscript  $k$  represents a given sampling time,  $t_0$  is the current time, and  $N_e$  is the estimation window length.  $x$  and  $y$  are the model states and outputs.  $u_p$  and  $y_p$  are the inputs implemented in the plant and plant measurements.  $F(\cdot; \cdot)$  is the discretized dynamic model, which considers the underlying PI controllers.  $H$  is a selector matrix.  $Q_y$  and  $Q_\theta$  are symmetric positive definite tuning matrices.  $\mathcal{X}$  and  $\Theta$  are the feasible state and parameter set.  $\hat{x}_{t_0 - N_e}$  and  $\hat{\theta}_{t_0 - N_e}$  are the most likely prior value of states and parameters at the beginning of the estimation window, and  $\Pi_{x, \theta}$  is the arrival cost matrix. The latter is used for summarizing the information before the estimation window and we use an extended Kalman Filter to compute it. Note that in this MHE formulation we assume that the parameters change within a sampling time following a random walk model.

## 5. Results and Discussion

The results are shown in Figure 4. Regarding the system disturbances, Figure 4a shows that MHE is able to estimate the uncertain parameters/disturbances accurately, even though they drift away from their nominal value with time. Moreover, we note that SRTO is seldom executed (executions are indicated by red circles) because the plant is at steady-state only for short periods due to the nature of the disturbances. This influences how often the boilup rate  $V$  (the free manipulated variable) is updated in the SRTO simulation. In Figure 4b, we note that  $V$  is updated only twice, which decreases SRTO economic benefits. On the other hand, the values of  $V$  computed by ROPA are a good approximation of the CL-DRTO, where the transients are rigorously optimized.

The main benefits of CL-DRTO, i.e., predicting the underlying behavior of the controllers, can be seen in Figures 4c and Figure 4d. In Figure 4c, we show the profile of the distillate rate determined by the condenser level controller, which was tuned to obtain fast level control. The consequence is the highly oscillatory behavior of the control variable  $D$  in ROPA simulations. On the other hand, CL-DRTO takes the aggressive control actions into account and the resulting behavior of  $D$  presents only minor oscillations, which are necessary to maintain the level constant due to the relatively small holdup of the condenser. For the top constraint controller (Figure 4d), we see the detrimental effect of the slow control tuning in ROPA. Even though the manipulated variable profile  $V$  is relatively well approximated as shown in Figure 4b, the slow composition control is not able to maintain  $x_D$  above 0.99. Moreover, the implementation of the setpoints computed by ROPA and SRTO overpurify the top stream, which causes a negative effect on the economic performance. The cumulative profit achieved by the methods in [\$] are  $\Phi_{cumm.SRTO} = 451$ ,  $\Phi_{cumm.ROPA} = 446$ , and  $\Phi_{cumm.CL-DRTO} = 455$ . Hence, ROPA and SRTO lead to smaller cumulative profits than CL-DRTO.

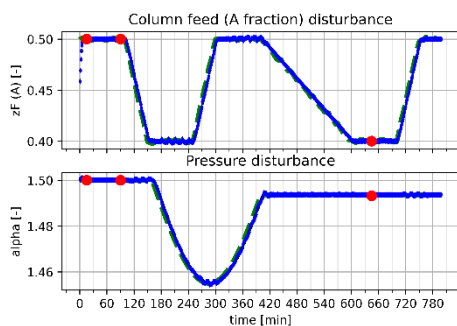


Figure 4a: Parameters  
Distillate Rate

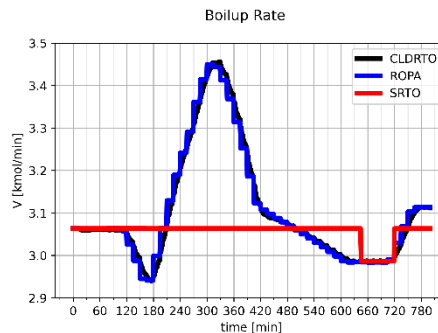


Figure 4b: Boilup Rate (V)  
Top Constraint

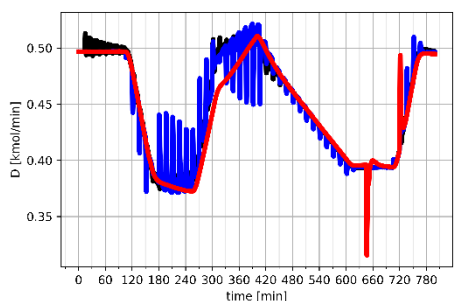


Figure 4c: Distillate Rate (D)

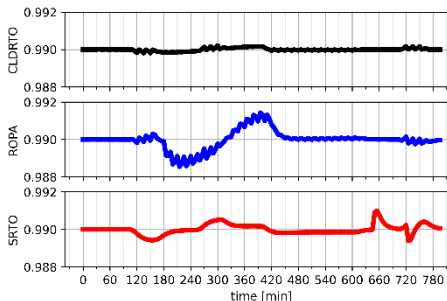


Figure 4d: Top Composition ( $x_D$ )

Figure 4: Simulation Results. CL-DRTO, ROPA and SRTO results are represented by black, blue, and red lines, respectively. In Figure 4a, the dashed line represents the true parameter values.

## 6. Conclusion

We explored the use of state and parameter estimation in Closed-loop Dynamic Real-time Optimization and evaluated the benefits of using this plant feedback strategy when compared to two other production optimization methods that incorporate plant information via online parameter estimation: Steady-state Real-time Optimization and Real-time Optimization with Persistent Parameter Adaptation. The simulations were conducted using a distillation system case study. The results show that the performance limitations caused by detuning or tightly tuning the controllers can be overcome by CL-DRTO, while they have a detrimental effect on the performance of the other two methods.

## References

- M. Câmara, et al., 2016. Performance evaluation of real industrial rto systems. *Processes* 4 (4), 44.
- C. Y. Chen, B. Joseph, 1987. On-line optimization using a two-phase approach: An application study. *Industrial & engineering chemistry research* 26 (9), 1924–1930.
- M. Z. Jamaludin, C. L. Swartz, 2017. Dynamic real-time optimization with closed-loop prediction. *AIChE Journal* 63 (9), 3896–3911.
- S. Kolás et al., 2008. State estimation is the real challenge in nmpc. In: *International workshop on assessment and future directions of nonlinear model predictive control*.
- J. Matias, G. A. Le Roux, 2018. Real-time optimization with persistent parameter adaptation using online parameter estimation. *Journal of Process Control* 68, 195–204.
- J. Matias, et al., 2022. Steady-state real-time optimization using transient measurements on an experimental rig. *Journal of Process Control* 115, 181–196.
- S. Skogestad, 1997. Dynamics and control of distillation columns: A tutorial introduction. *Chemical Engineering Research and Design* 75 (6), 539–562.

# Unsupervised anomaly detection model for diesel off-spec color change triggered by flooding

A. Eren Vedin<sup>a</sup>, Sadik Odemis<sup>b</sup>, Aysegul Sener<sup>b</sup>, Gizem Kayar<sup>b</sup>, Mammad Aliyev<sup>b</sup>

<sup>a</sup>*SOCAR Turkey, Vadistanbul, Istanbul 34485, Turkey*

<sup>b</sup>*SOCAR STAR Oil Refinery, Aliaga, Izmir 35800, Turkey*

## Abstract

Crude oil feed instability can cause significant problems in refineries, where various types of crude oils are processed and frequent feed changes are taken place. In particular, if crude oil changes are performed in a short time, it can create various operational disruptions in the Crude Distillation Unit (CDU). In case of a sudden increase or decrease of light ends in the CDU feed, pressure fluctuation or flooding problems can occur in the atmospheric distillation column. In order to prevent potential upset cases caused by crude changes, an unsupervised machine-learning model was developed. The model which is presented in this article analyzes the high-correlated relationship between pressure of column flash zone, upper and lower diesel trays in steady-state operation. Principal Component Analysis (PCA) is performed with these three inputs. The PCA reconstruction error is calculated with new process values in every minute and the pressure profile in the column is instantly detected by the error. A soft sensor was developed by applying the unsupervised anomaly detection method to detect off-spec color.

**Keywords:** Crude oil distillation, diesel off-spec color, unsupervised anomaly detection, Principal component analysis

## 1. Introduction

The crude distillation unit separates the components in crude oil by utilizing the difference in boiling points (McCaffrey, 2021). Product yields differ depending on the choice and origin of crude oil. Due to market conditions, refineries today prefer to process crude oil from different sources. However, operational conditions like the blending of different crude oils, rapid crude-oil tank changes, or operation capacity maximization can lead to challenges for column operation. These challenges make the operation more susceptible to potential operational faults or other process disruptions. In such cases, flooding and weeping problems can be seen in the atmospheric column. In this article, diesel off-spec color change triggered by flooding prediction is submitted. Although there are other reasons causing off-spec color like pH change or preheat-train heat exchanger leaks, these conditions are not taken into account for prediction (Kister, 2006).

A flooding prediction model was developed because there is no online ASTM color analyzer system in the process that automates the color measurements (ASTM D1500-12(2017)). When a color change of the product is detected by routine laboratory samples or by visual controls, the product has already become off-spec. Therefore, an analytical

model development was considered necessary. Since flooding is a rare anomaly in normal operation, historical data obtained from the process becomes also imbalanced. Overfitting and inaccurate results can be observed in traditional supervised approaches which apply train-validation separation in the dataset, and these models can be inadequate against the problem. For this reason, an unsupervised anomaly detection model was adopted to overcome this weakness.

It is known in the literature that the temperature diagram, fraction yields, or column pump around flows can play a role in flooding identification. Since these parameters are also intervened in steady state operating conditions, flash zone pressure, diesel lower and upper stage pressures were used instead of these process parameters. An unsupervised anomaly detection model based on Principle Component Analysis (PCA) and PCA Reconstruction Error was deployed on Distributed Control System (DCS) for flooding. The model was developed by Python programming language and Python Anomaly Detection Tool Kit (ADTK) library was used for PCA method. After this study, the model detected a real diesel off-spec color case triggered by flooding 100 minutes before happening. With the early warning system, operators took action to adjust the operation to prohibit color changes in products

## 2. Methodology

### 2.1. Flooding and Off-Spec Mechanism

The objective of the crude unit (CDU) is to process a crude oil blend into its fractions by means of boiling point differences. In distillation process, flooding and weeping are common problems that can occur and cause loss of separation. The flooding problem is a complicated abnormality whose mechanism is based on liquid or excessive vapor phase accumulation and liquid-vapor separation inability. Although flooding reduces separation efficiency in the column dramatically, there can be additional and more crucial effects in the process like feed rate reduction, operational interruption, or off-spec product (Kister and Olsson, 2019). As can be seen both in the refinery incidents and in the literature, flooding causes a sudden increase in delta pressure in the column and can cause off-spec color diesel which is also called as black diesel. (Bird et al., 2018).

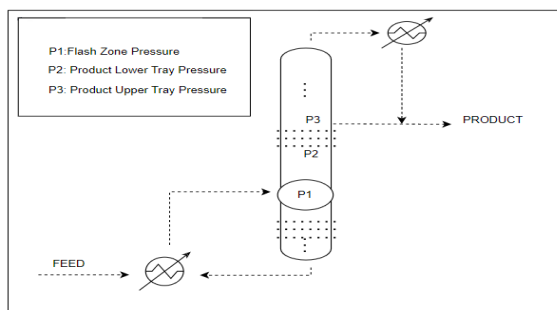


Figure 1: Typical Distillation Column Schema

A typical distillation column scheme was given in Figure 1, where P1 refers to flash zone pressure, P2 and P3 refer to diesel lower tray pressure and diesel upper tray pressure. In case rapid pressure difference changes between P1 and P2 or P3, product draw-off amounts from column trays are also changed. This causes heavy components to come

along with the bottom stream and turns the diesel color black (Bird et al., 2018). Fortunately, flooding and off-spec color prediction are possible by examining the pressure at these stages. This study gives a chance to the operator sufficient time to eliminate the risk of flooding.

### *2.2. PCA Analysis and Reconstruction Error*

Principle Component Analysis (PCA) is a classical statistical method that creates principle components that refer to the data set. Firstly, data is converted to a matrix, and vectors are generated by the decomposition of this matrix. As a second step, the original data is reconstructed by the first principle components, and reconstruction error is calculated with the difference between the new reconstructed data and the original data. The difference between the new matrix and the original data matrix is compared at each model run and gives an error. Calculated reconstruction error is examined by anomaly detection method, to expose at which level of error there is abnormal or not. With the increase in error, anomalies that occur in the process are detected (McCaffrey, 2021). Although there is no supervised learning included in this method, both high correlation in the dataset and strong coherence in the time series, generate a natural knowledge for the model. Therefore, the help of reconstruction error rise in value can detect anomalies in the data effectively. Unsupervised Anomaly Detection Machine Learning models based on PCA analysis are widely applied to industrial problems (Takeishi, 2019).

Another benefit that enables PCA analysis valuable in the industry is dimension reduction. Today, the management and control of any distillation unit are operated more efficiently and stable by engineers and operators with the help of transmitters and process control systems. The presence of many transmitters such as temperature, pressure, and flow, allows the system to be controlled with more assurance. However, having many measurement points, which are highly correlated to each other, makes data analysis more complicated and model data larger. The PCA algorithm eliminates this situation, which is known as multicollinearity. In PCA method, new independent orthogonal variables are created. At this point, PCA analysis is used as a rapid and efficient multivariate statistical method with ease of PCA reconstruction error that can be successfully applied in industry to detect abnormal situations or deviations in the process (Thomas et al., 1996).

### *2.3. Data Flow from Business Network to DCS and Application*

Model outputs were integrated into the Industrial control system (ICS) so that control operators can use the results. Process Data History (PHD) in PHD Historian is operated to collect and pull the live process data and the results are calculated with Python code. After calculation, results are sent to Distributed Control System (DCS) for operator control. Purdue Enterprise Reference Architecture (PERA) model by ISA-99 was used as a concept model for ICS network segmentation (ANSI/ISA-TR99.00.01-2007) Deployed anomaly model was observed to predict flooding 100 minutes in advance. By adding an alarm to the tags created in the DCS system, audible and written warnings are created for the operator to take action. When the operator receives an alarm from the system, they try to reduce the risk of flooding by interfering with the crude oil column product drafts, pump around flows, or return temperatures.

### 3. Results and Discussion

Pressure differentiation is not the only reason for product color change. Incorrect process control tunings, abnormal column temperatures, abnormal operation of condenser or furnaces, physically damaged inner-column parts, leaks in exchangers or high nitrogen levels in the crude mix may also cause darker color. Each of these parameters were examined one by one before the study. Since sudden pressure changes and sudden increase in column level was detected during off-spec product, it was estimated that the problem was caused by flooding and a pressure-induced model can help to detect off-spec color. An unsupervised anomaly detection model based on PCA analysis was studied over the pressure parameters in the CDU column to predict the flooding abnormality and off-spec diesel color. Pressure transmitters at the column flash zone, diesel stage lower tray, and upper tray were used with the help of their highly correlated interactions.

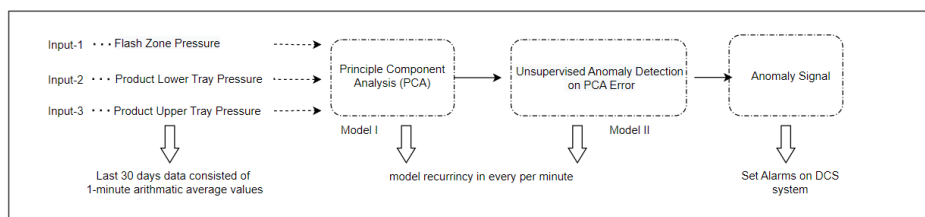


Figure 2: Data structure in the model

Not only early flooding detection but also taking fast action is significant for unit operation. Therefore, the model was scheduled to run once per minute. By processing data from the last 30 days to the model at each run, the normal fluctuation and noise level in the steady state process are understood by the model despite unsupervised learning. The data structure was created with 1-minute average values to increase sensitivity against small new changes in operation. The data architecture drawing was given in Figure 2.

A real diesel off-spec color change case, which was detected by the analytical model, is seen in Figure 3. Pressure parameters used in the model, atmospheric distillation column level, PCA reconstruction error, and delta pressure are the process parameters shown in the figure after rescaling between 0 and 1. Hairline I indicates that the model alarmed at 5:45 (see figure 3c) when incipient flooding occurred, and hairline II indicates runaway flooding occurred at 7:30 with column level increase (see figure 3 b).

Before hairline I, there was a high correlate movement of flash zone pressure, diesel lower and upper tray pressures, but the movement seems to have ceased just before the onset of the flooding initiation (see figure 3a). PCA analysis used this behavior between model parameters. The deterioration of the relationship between the parameters caused the increase in PCA error. However, it is seen that there was no increase in column level until hairline II, where flooding increased its severity. With this time difference, It was aimed to give the operator the necessary time to act.

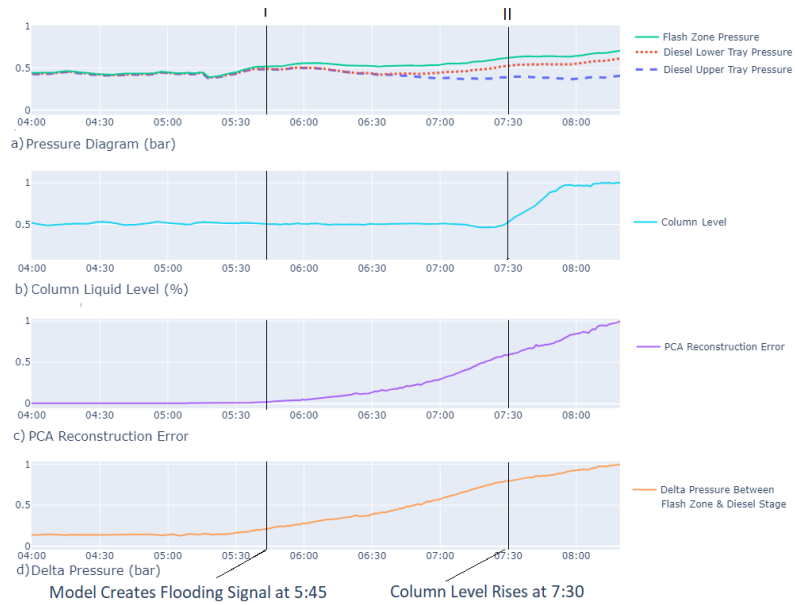


Figure 3: Real case off-spec color triggered by flooding

In literature knowledge, delta pressure between product tray and flash zone is used as another method. In this method, delta pressure is used as a stand-alone indicator to detect flooding. (Deviation, 2009). During flooding, responses of delta pressure and PCA error can be seen as similar to each other (see figure 3d). However, this method was not preferred in the study. There are always continuous and small deviations in delta pressure (DP). Although a rapid increase of these deviations can be used to create an alarm condition, DP is usually manually controlled and changed by operators during ordinary operations too. This means, deviations coming from operational set point changes or process deviations can be confused by the flooding model and this can create false alarms.

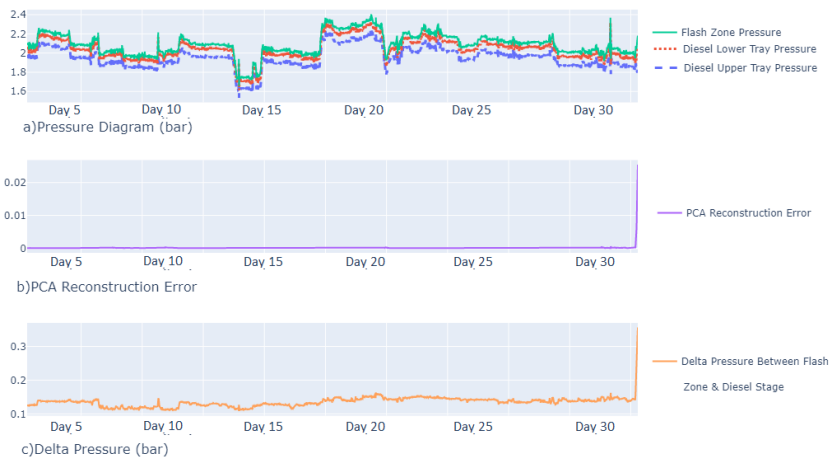


Figure 4: Historical data prior to flooding. a) Pressure diagram, b) PCA reconstruction error, c) Delta pressure between flash zone and diesel stage



One-month pressure changes of diesel flash zone and diesel trays are given in figure 4a. The operational changes and process noises were seen to create bigger noises during both flooding case and steady-state operation period (see figure 4c). On the other hand, PCA reconstruction error seems more stable during steady-state period and started a rapid increase during only flooding (see figure 4b). For this reason, PCA analysis was preferred as an alternative model in this application. Instead of using the delta pressure alone, a PCA error-based unsupervised anomaly detection model was developed with flash zone, upper and lower diesel pressures.

Nevertheless, delta pressure calculation was also used for understanding if there was a positive or negative differentiation in the column. Inferential was developed based on PCA analysis, and DP was added as a supportive decision mechanism for the operator to see to filter flooding conditions among all abnormal conditions like weeping.

#### 4. Conclusion

Crude oil feed instabilities can create instability in atmospheric distillation column and can cause upsets like flooding and weeping. In the article unsupervised anomaly detection model based on PCA analysis was used for early flooding detection which can cause off-spec color changes. It has been seen that the developed model predicted a real flooding and color change case 100 minutes in advance. The off-spec color formation was seen by a diesel laboratory sample. The study was proven and confirmed by users. Any off-spec problem has not been observed in the refinery for one year.

#### References

- ANSI/ISA-TR99.00.01-2007 Security Technologies for Industrial Automation and Control Systems.
- ASTM D1500-12(2017), DOI: <https://dx.doi.org/10.1520/D1500-12R17>.
- Bird, J., Brown Burns, J., Racette, Y.,J, Beaulieu, 2018, Use advanced predictive analytics for early detection and warning of column flooding events, *Hydrocarbon Processing, June 2018*.
- Deviation, D. S. (2009). Distillation Column Flooding Diagnostics with Intelligent Differential Pressure Transmitter.
- Henry Z. Kister,2006, Distillation Troubleshooting,John Wiley & Sons, Inc.,576.
- Kister, Z., Olsson, M. (2019). An Investigation of Premature Flooding in a Distillation Column,Chemical Engineering Journal,<https://www.chemengonline.com/investigation-premature-flooding-distillation-column>
- McCaffrey, J. (2021). Anomaly Detection Using Principal Component Analysis (PCA). Visual Studio Magazine, <https://visualstudiomagazine.com/articles/2021/10/20/anomaly-detection-pca.aspx>.
- Takeishi, N.,2019, Shapley values of reconstruction errors of pca for explaining anomaly detection. In 2019 international conference on data mining workshops (icdmw) (pp. 793-798). IEEE.
- Thomas, C., Wada, T., & Seborg, D. E. (1996). Principal component analysis applied to Process monitoring of an industrial distillation column. IFAC Proceedings Volumes, 29(1), 5859-5864.

# Artificial neural networks-based identification of the WWTP DO sensor types of faults

Norbert-Botond Mihály<sup>a</sup>, Alexandra-Veronica Luca<sup>a</sup>, Vasile Mircea Cristea<sup>a\*</sup>

<sup>a</sup>*Babes-Bolyai University of Cluj-Napoca, 1 Mihail Kogalniceanu Street, 400028 Cluj-Napoca, Romania*

## Abstract

Efficient and sustainable operation of the wastewater treatment plants using the activated sludge technology strongly depends on the aeration control. The role of the Dissolved Oxygen (DO) sensor for achieving the desired performance of the control loop is critical and the faults of the sensor have to be identified with promptitude. The paper presents an effective artificial neural networks (ANN)-based diagnosis and identification methodology for seven types of DO sensor faults, revealing the advantages of the proposed identification approach. The classification ANNs trained with the Bayesian regularization method showed high accuracy in identifying the fault classes, with an overall accuracy of 99.5% for the testing dataset. The developed ANNs allow the efficient and early detection of the sensor fault presence, while performing the correct identification of the sensor fault type.

**Keywords:** Artificial Neural Networks, fault type diagnosis, Dissolved Oxygen sensor.

## 1. Introduction

Increased awareness of the eutrophication negative impact on water quality and advances in environmental protection technologies have resulted in more stringent wastewater treatment requirements. Regulations developed during recent decades have followed this trend. Tightening treatment regulations are driving the addition of new processing units and renewal of existing ones in the wastewater treatment plants. The subsequent increase in operational and management investments, primarily related to energy consumption and sustainable development, encouraged modern WWTPs to meet the challenges of preserving and enhancing effluent quality while ensuring efficient and safe operation, associated to cost optimization (Haimi et al. 2013).

For comprehensive process monitoring and efficient control, accurate measurement of the composition, flow rate and temperature of the influent, effluent and process streams are obviously required. Furthermore, satisfying the needs for safety and effective process control asks for the ability to detect sensor faults as early as possible. Monitoring and control at the WWTP are focused on specific equipment, such as biological reactors, settlers, pumps, air blowers, control valves, diffusers, and the pipeline system (Morera et al. 2017). Composition measurements make use of sensors that monitor Dissolved Oxygen (DO), nitrates and nitrites, ammonia, phosphorous, suspended solids, and organic matter (Mamandipoor et al. 2020). The precise information provided by these sensors is essential for the efficient and secure operation of the entire plant.

Development of numerous methods have been aimed for fault detection. Due to their inherent detection capacity, multivariate statistical methods have the ability to reveal faults in time-varying, ill-defined, and non-linear systems (Li and Yan 2019). The role of the Dissolved Oxygen sensor for achieving the desired performance of the nitrification

control loop is critical and the faults of the sensor have to be identified with promptitude. Few studies have been conducted to investigate the issue of DO sensor failures. PCA methodology was used to detect deviations due to sensor clogging (Samuelsson et al. 2018), while another study implemented PCA for data driven detection of oxygen sensor faults (Luca et al. 2021).

Artificial intelligence (AI) techniques have been aimed to perform sensor fault detection and error type diagnosis in a variety of fields. They span over different applications, such as error diagnostic strategies for vapor compression refrigeration system (Kocyigit 2015), on-line fault detection of a sensor in a nuclear reactor core (Messai et al. 2015), fault detection and isolation in sensors of an internal combustion engine (Cervantes-Bobadilla et al. 2023), and fault diagnosis of vibration signals (Zhang et al. 2022).

AI was proven to be an effective tool, intended to target various implementations in WWTPs for: prediction of process parameters (Zaghloul and Achari 2022), prediction of plant performance indices and operation optimization by finding the control loop setpoints (Mihály et al. 2022a), optimization of aeration in the aerobic reactor (Mihály and Cristea 2022b), and optimization of recycle flowrates for better effluent quality and energy savings (Mihály et al. 2022c). One fault diagnosis using ANNs was successfully applied to sludge bulking measurements with the intention of ensuring process safety and WWTP effluent quality. The developed model could differentiate five types of sludge bulking faults (Han et al. 2019). However, artificial neural networks (ANNs) were scarcely applied to the identification of WWTP sensor faults.

The scope of the current paper is to design, train and test the performance of an ANN based tool that is capable of detection and identification of seven types of Dissolved Oxygen sensor faults. The novelty consists in the application of proposed ANN diagnosis tool for the municipal A<sup>2</sup>O WWTP, having implemented the two main control loops, one for nitrification and another for denitrification.

## 2. Methodology

### 2.1. Data generation

The data implied by the development of the ANN fault identification tool were obtained using a previously calibrated first-principle Activated Sludge Model no.1 for the Anaerobic-Anoxic-Oxic reactors configuration (Várhelyi et al. 2019) of the WWTP considered as case study. Seven types of sensor faults were simulated for the Dissolved Oxygen sensor: bias, also known as shift or off-set of the signal; complete failure, which can be maximum or minimum and occurs when the measured value is either the highest or the lowest value of the sensor calibration range; drift, which is an irregularly varying deviation in time of the measured value, compared to the true DO one; fixed value, when the DO signal value is constant; loss of accuracy, meaning that the value provided by the sensor is affected by imprecision around the true value; wrong gain, known also as a typical calibration error (Luca et al. 2021). They were denoted with class numbers ranging from two to eight, while normal fault-free case was considered as class number one.

The first 100 days of the simulations were performed in normal (nominal) operation and were considered for obtaining the regular (quasi-stationary) state of the WWTP. The simulation continued in normal operation mode until day 140, where the different types of DO sensor faults were implemented until the 160<sup>th</sup> day of operation. Data obtained during these 20 days of either normal or faulty operation periods, with a sampling time of 15 minutes, was used in the development of the classification ANNs. The same data generation approach, for a different fault identification methodology, was implemented in a previous study and where additional information can be found (Luca et al. 2021).

### 2.2. ANN development

A set of process parameters were utilized in the development of the ANNs. They were the clean effluent concentration of the variables: total nitrogen, total Kjeldahl nitrogen, chemical oxygen demand, nitrate and nitrite nitrogen, free and saline ammonia, and total suspended solids. The following concentration variables of the bottom effluent from the secondary clarifier were also taken into consideration: nitrate and nitrite nitrogen, free and saline ammonia, total suspended solids, slowly biodegradable substrate, heterotrophic biomass, autotrophic biomass, inert particulate products, particulate biodegradable organic nitrogen, soluble biodegradable organic nitrogen, Dissolved Oxygen concentration, readily biodegradable substrate, alkalinity, associated to waste flow rate, and the temperature variables. For the identification of the different DO sensor faults an ANN with two layers was considered, the first hidden layer with tangent-sigmoid transfer function, while the second (classification) layer with SoftMax as transfer function. The implementation of the first-principle model and development of ANNs were carried out in MATLAB & Simulink software environment. The hidden layer size (number of hidden neurons) was optimized by a trial-and-error process, searching around the number of ANN inputs, while the training efficiency was studied with both Levenberg-Marquardt and Bayesian regularization backpropagation training algorithms with variable learning rates.

### 3. Results and discussion

The dataset was divided in the following parts when the Levenberg-Marquardt training algorithm was employed: 70% for training, 15% for validation, and 15% for testing. The training results showed good results, with the exception of the highest obtained confusion value of 58.7%, due to the confusion between the complete failure maximum (class 3) and fixed value (class 6) fault types, characterized by comparable effects. This can be observed for the confusion matrix of the testing dataset presented in Figure 1.

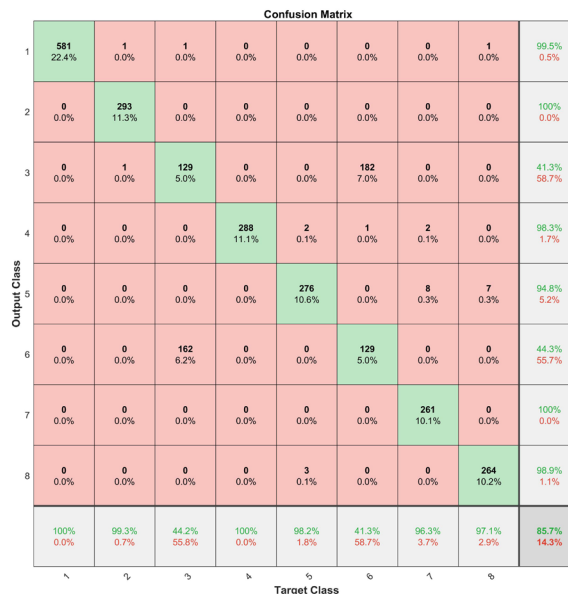


Figure 1. Confusion matrix of the testing dataset for ANN trained with Levenberg-Marquardt algorithm

However, the trained classification ANN properly detects the different fault classes, only in 0.5% of the cases classifying the faulty operation as normal (class 1). At the same time, the identification is fairly accurate except the previously mentioned confusion, while the next highest confusion percentage being of 5.2%, when loss of accuracy (class 7) and wrong gain (class 8) fault types are wrongly classified as drift faults (class 5).

In comparison to the previous training approach, the classification performed by the ANN trained with the Bayesian regularization algorithm showed better results. Results obtained for the testing data set are presented in Figure 2. This method of training did not require a validation dataset. Consequently, 85% of the data were used for training, while the remaining 15% constituted the testing dataset. In terms of training, the highest confusion value was of 1.1% in case of the wrong gain fault, where out of 1920 data points only 21 were classified as other fault classes. It is worth mentioning that neither at training, nor at testing were any of the faulty data points classified as normal (faulty-free) operation, and the previously confused complete maximum and fixed value sensor faults were correctly classified by this latter ANN.

Output Class \ Target Class	1	2	3	4	5	6	7	8	Accuracy
1	600 23.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
2	0 0.0%	260 10.0%	0 0.0%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	0 0.0%	99.6% 0.4%
3	0 0.0%	0 0.0%	269 10.4%	0 0.0%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	99.6% 0.4%
4	0 0.0%	1 0.0%	0 0.0%	221 10.8%	1 0.0%	0 0.0%	0 0.0%	0 0.0%	99.3% 0.7%
5	0 0.0%	0 0.0%	0 0.0%	0 0.0%	280 10.8%	0 0.0%	3 0.1%	3 0.1%	97.9% 2.1%
6	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	287 11.1%	0 0.0%	0 0.0%	100% 0.0%
7	0 0.0%	0 0.0%	0 0.0%	1 0.0%	1 0.0%	0 0.0%	298 11.8%	1 0.0%	99.0% 1.0%
8	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	304 11.7%	100% 0.0%
Overall	100% 0.0%	99.6% 0.4%	100% 0.0%	99.6% 0.4%	99.3% 0.7%	99.7% 0.3%	98.7% 1.3%	98.7% 1.3%	99.5% 0.5%

Figure 2. Confusion matrix of testing dataset for ANN trained with Bayesian regularization algorithm

Similar to the previous ANN, at the testing dataset we can observe that loss of accuracy and wrong gain fault types (classes 7 and 8) are the toughest to identify correctly, exhibiting the highest confusion percentage revealed by the reduced value of 2.1%. Regarding the training time of the two ANNs, the Levenberg-Marquardt algorithm was less time-consuming, as it required 1.3 minutes to run, while the training with Bayesian regularization lasted for 15.8 minutes. Even though there is a significant difference in computational resources needed between the two training algorithms, the superior classification accuracy offered by the second ANN leads to its ranking as the best and recommended classification model.

The classification with the second ANN for the testing data set and case of fault class 7 (loss of accuracy), which was collectively the most difficult to identify, are presented in Figure 3. The faulty operation of the WWTP is detected without time lag, from the very first failure point of the testing data (following day 140, when the sensor fault begins). The classification network quickly and correctly identified the different fault types in: 3 hours for bias, 2.5 hours in case of complete fail maximum, 2.5 hours at complete fail minimum, 3 hours for drift, 3.5 hours for fixed, 4 hours at loss of accuracy, and the longest identification of 11 hours in case of wrong gain also shown in Figure 3.

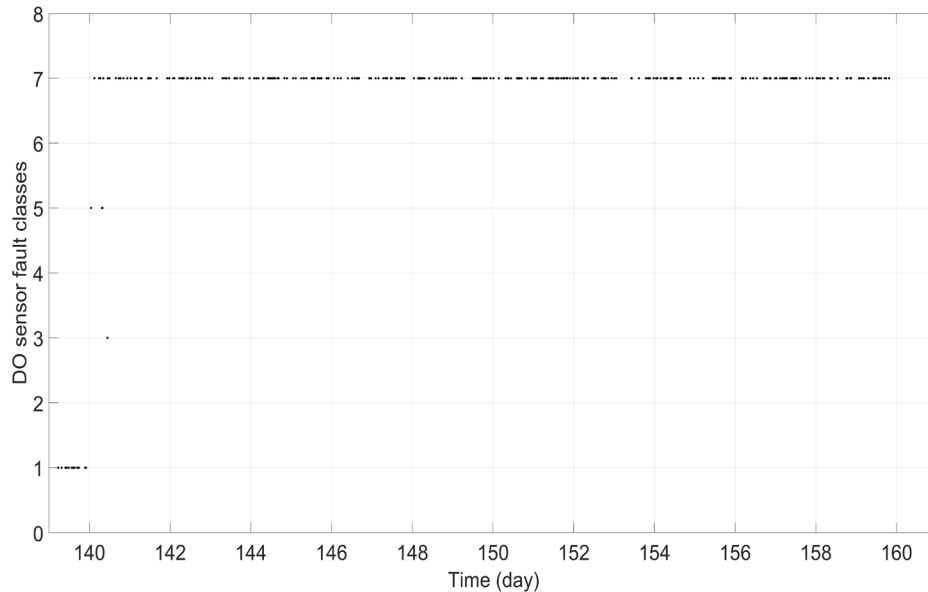


Figure 3. Classification of the testing dataset for loss of accuracy DO sensor fault

The practical implementation of the ANN based identification tools relies on the dedicated measuring instrumentation required to obtain in real time the variables used as ANN inputs. The classification offered by this ANN is of great importance for the efficient management of the WWTP and mitigation of both the environmental and economic negative impact that a faulty sensor causes by the late detection and identification of the fault. Depending on the type of the sensor fault, the CO<sub>2</sub> and N<sub>2</sub>O green House Gases emissions released during the faulty sensor operation may show increased values, characterized by factors reaching one order of magnitude higher than the values of the normal WWTP operation. They are also associated to the significant degradation of the discharged effluent water quality and harmfully affect the downstream river basin. The proposed ANN diagnosis for DO sensor fault types is useful for metrological maintenance of measuring equipment, producers of intelligent sensors and their integration in digital twins.

#### 4. Conclusions

In this study, the ANN-based sensor fault type identification method was proposed for diagnosing seven types of DO sensor faults, typical to the operation of a municipal case study WWTP. The DO measured values are of vary large importance for the aeration control loop, and they significantly influence the effluent quality, energy consumption

and sustainable operation of the plant. The investigated fault types (bias, complete failure maximum, complete failure minimum, drift, fixed value, loss of accuracy and wrong gain) were successfully identified by the classification ANN. Assessed for the testing datasets, a confidence level of 85.7% was achieved in case of the ANN trained with Levenberg-Marquardt algorithm, while this value was of 99.5% for the ANN trained with Bayesian regularization algorithm. The potential of the methodologies had been demonstrated by the efficiency of the ANN-based fault identification approach. Major practical importance can be attributed to the developed classification ANNs for the accurate monitoring and high-performing control of the WWTP complex and non-linear intrinsic processes, in support of their safe and efficient operation.

## References

- M. Cervantes-Bobadilla, J. García-Morales, Y.I. Saavedra-Benítez, J.A. Hernández-Pérez, M. Adam-Medina, G.V. Guerrero-Ramírez, R.F. Escobar-Jimenez, 2023, Multiple fault detection and isolation using artificial neural networks in sensors of an internal combustion engine, *Eng Appl Artif Intell*, 117:105524.
- H. Haimi, M. Mulas, F. Corona, R. Vahala, 2013, Data-derived soft-sensors for biological wastewater treatment plants: An overview, *Environ Model Softw*, 47:88–107.
- H-G. Han, H-X. Liu, Z. Liu, J-F. Qiao, 2019, Fault detection of sludge bulking using a self-organizing type-2 fuzzy-neural-network, *Control Eng Pract*, 90:27–37.
- N. Kocuyigit, 2015, Fault and sensor error diagnostic strategies for a vapor compression refrigeration system by using fuzzy inference systems and artificial neural network, *Int J Refrig*, 50:69–79.
- Z. Li, X. Yan, 2019, Ensemble model of wastewater treatment plant based on rich diversity of principal component determining by genetic algorithm for status monitoring, *Control Eng Pract*, 88(130):38–51.
- A.V. Luca, M. Simon-Várhelyi, N.B. Mihály, V.M. Cristea, 2021, Data driven detection of different dissolved oxygen sensor faults for improving operation of the wwtp control system, *Processes*, 9(9):1633.
- B. Mamandipoor, M. Majd, S. Sheikhalishahi, C. Modena, V. Osmani, 2020, Monitoring and detecting faults in wastewater treatment plants using deep learning, *Environ Monit Assess*, 192(2):148.
- A. Messai, A. Mellit, I. Abdellani, A. Massi Pavan, 2015, On-line fault detection of a fuel rod temperature measurement sensor in a nuclear reactor core using ANNs, *Prog Nucl Energy*, 79:8–21.
- N.B. Mihály, M. Simon-Várhelyi, V.M. Cristea, 2022a, Data-driven modelling based on artificial neural networks for predicting energy and effluent quality indices and wastewater treatment plant optimization, *Optim Eng*, 23:2235–2259.
- N-B. Mihály, V.M. Cristea, 2022b, Optimization of the Wastewater Treatment Plant Aeration Using Artificial Neural Networks Models, *Computer Aided Chemical Engineering*, 51:1375–1380.
- N-B. Mihály, M. Simon-Várhelyi, A.V. Luca, V.M. Cristea, 2022c, Optimization of the Wastewater Treatment Plant Recycle Flowrates Using Artificial Neural Networks, 2022, *IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR)*, 1–6.
- S. Morera, L. Corominas, M. Rigola, M. Poch, J. Comas, 2017, Using a detailed inventory of a large wastewater treatment plant to estimate the relative importance of construction to the overall environmental impacts, *Water Res*, 122:614–623.
- O. Samuelsson, A. Björk, J. Zambrano, B. Carlsson, 2018, Fault signatures and bias progression in dissolved oxygen sensors, *Water Sci Technol*, 78(5):1034–1044.
- M. Várhelyi, V.M. Cristea, M. Brehar, E.D. Nemeş, A. Nair, 2019, WWTP model calibration based on different optimization approaches, *Environ Eng Manag J*, 18(8):1657–1670.
- M.S. Zaghoul, G. Achari, 2022, Application of machine learning techniques to model a full-scale wastewater treatment plant with biological nutrient removal, *J Environ Chem Eng*, 10(3):107430.
- Y. Zhang, L. He, G. Cheng, 2022, MLPC-CNN: A multi-sensor vibration signal fault diagnosis method under less computing resources, *Measurement*, 188:110407.

# Energy Out-of-distribution Based Fault Detection of Multivariate Time-series Data

Umang Goswami,<sup>a</sup> Jyoti Rani,<sup>a</sup> Deepak Kumar,<sup>a</sup> Hariprasad Kodamana<sup>a,b</sup>  
Manojkumar Ramteke<sup>a,b</sup>

<sup>a</sup>*Department of Chemical Engineering, Indian Institute of Technology, New Delhi-110016, India*

<sup>b</sup>*Yardi School of Artificial Intelligence, Indian Institute of Technology, New Delhi-110016, India*

## Abstract

A major challenge faced by the chemical process industry is carrying out operations safely and safely. The proposed work entails a fault detection approach for a multivariate time series dataset by utilizing the energy scores instead of the traditional approach. This work proposes a loss function which utilizes the concept of in-distribution and out of the distribution of data. Energy scores are more theoretically aligned with the probability density of the inputs and can be used as a scoring function. For a pre-trained neural network, energy can be utilized as a scoring function and can also be used as a trainable cost function. The concept of out-of-distribution is similar to that of any outlier identification method. Similarly, for energy out of distribution, an energy value which falls below a certain threshold can be considered an outlier and is addressed as out-of-distribution. The values within the range are in-distribution. Higher energy values imply a lower likelihood of occurrence and vice versa. The proposed approach is compared with different deep learning approaches like Auto-encoders (AEs), LSTMs and LSTM-AEs that are traditionally used for anomaly detection and utilize the softmax scores. The Proposed methodology is also compared with some state-of-the-art fault detection methods, such as the PCA and DPCA and returns encouraging results. Energy based out of distribution is coupled with various deep learning methods to identify faulty and normal points. When teamed with the Auto-encoder network, energy-based scoring proved to be of significant dominance compared to other methods. The study was validated for the benchmark Tennessee Eastman data for fault detection.

**Keywords:** Fault detection, softmax score, energy score, outliers, Out-of distribution

## 1. Introduction

Anomalies or Outliers constitute a significant hindrance in the smooth functioning of the process industry. Fault detection or anomaly identification is one of the exercises that need to be carried out not only for safety purposes but also from a financial standpoint. Abnormal behaviour of the systems is a significant challenge not only for the chemical process plant but also for manufacturing, production and pharmaceutical industries. (Kodamana; 2017). Process models can be used to identify such faults, which are obtained either by first-principles-based models or data-based models. In systems with inherent non-linearity and non-Gaussianity, first-principle models are tougher to build. On the other hand, data-based models can account for all the difficulties faced by the first-principles models and can be applied due to their flexibility.



Over the years, several data-based methods have been developed for fault detection. Various machine learning approaches, such as Principal Component Analysis and Dynamic Principal Component Analysis (Rani, 2023), have been thoroughly used by the Process industries. Similarly, deep learning frameworks have also found their feet in the fault detection domain. Neural network architectures such as Auto-encoders (Sakurada, 2014), long short-term memory (LSTMs) (Nguyen, 2021; Yang, 2019) and an amalgamation of LSTM-AEs have been widely used. Deep learning architectures utilize neural networks to understand and learn a pattern for identifying faults. This work proposes an approach where the traditional neural networks are combined with the energy scores for fault classification. The loss function of the energy-based out-of-distribution is utilized for fault identification such that the problem of vanishing gradients does not occur in the system.

Various researchers have proposed out of distribution-based methods (Du, 2019) with different scoring functions. Scoring methods such as softmax score, Mahalanobis distance score, OE (Hendrycks, 2019) and ODIN (Liang, 2017) score are some of the distribution methods that the researchers have proposed. The energy-based scoring method proposed by (Liu, 2020) uses an energy-based score to identify the out-of-distribution data. In contrast, little to no work is done concerning the multi-variate time-series data for identifying faults using energy out-of-distribution methods. Compared to the softmax scores, energy scores are more theoretically aligned with the probability density of the inputs. The higher dimensional data can be mapped to a single dimension data with the help of energy scores. Based on the application, energy can be used as a trainable cost function for fine-tuning the classification model. A high energy value corresponds to a lower likelihood of occurrence and vice versa. Values exceeding a particular threshold can be considered under the umbrella of out-of-distribution and can be regarded as a fault. This serves as a classification method for faulty and normal points. The present work is carried out for Multivariate Time-series data and validated for the Tennessee Eastman Dataset benchmark. Energy-based scoring functions, when coupled with Deep learning frameworks, produced promising results. The remainder of the article is structured as follows: Section 2 contains the proposed methodology, Section 3 deals with the results and discussions, and Section 4 holds the conclusion and the significance of the proposed work.

## **2. Proposed Methodology**

### *2.1. Deep-Learning Frameworks for Fault Detection*

As discussed in the previous section, deep-learning-based fault detection methods have traditionally been used for fault detection. These architectures differ from each other in the way the neural network is designed for each of them. Different layers and number of neurons for each of the framework is used. Auto-encoder is a reconstruction-based operator. On the other hand, LSTM is not a simple feed-forward network but also has feedback connections for the neurons, and LSTM-AE utilizes the architecture of Auto-encoders in an LSTM model. The proposed methodology is presented in figure 1, where an auto-encoder network is considered, wherein an energy-based scoring method is used to detect the faults. Similarly, this methodology also considers different neural networks, as presented in the sub-sections below.

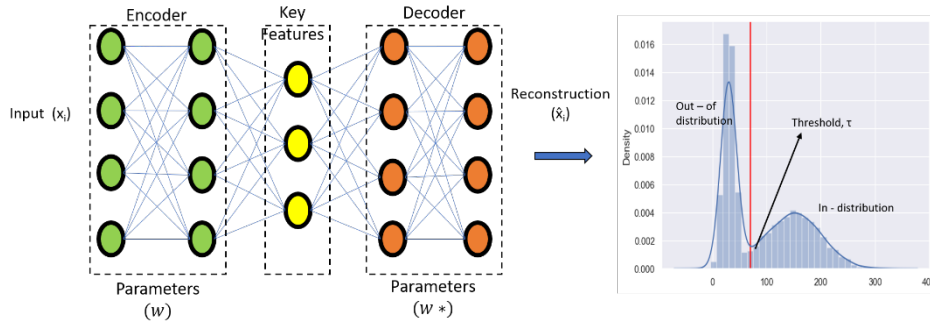


Figure 1: A schematic of the proposed Methodology.

### 2.1.2. Auto-Encoder (AE)

Auto-encoder network is a feed-forward neural network which has an encoder and decoder entailed it. The encoder tries to learn the data and converts it to a hidden space of lower dimension, and the decoder then captures the hidden space and reconstructs the original data. In a traditional auto-encoder, the mean squared error is used as a loss function, and the loss function is computed between the network output and the original input.

### 2.1.3. Long-Short Terme Memory (LSTM) Networks

LSTM networks are an extension of the Recurrent Neural networks (RNN), which can handle long-term dependencies. LSTM is not a simple feed-forward network; instead, it is a particular type of neural network that considers the previous neurons and takes inputs from them. LSTM considers sequential data and returns output for a specific window of time steps. It has three significant parts; in the first part, the model is trained to forget the irrelevant information, the second part learns the new information, and the third part returns the updated information from the current time step to the next time step.

### 2.1.4. LSTM-AE

LSTM-AE considers the encoder-decoder structure for an input sequence. The salient feature of the auto-encoder network is to reconstruct the input, and the LSTM-AE uses this to tackle the sequential data.

### 2.1.5. Energy Function

The main objective of Energy-based models is to map higher dimensional data to a single non-probabilistic scalar quantity called energy. The function  $E(x): R^D \rightarrow R$  maps each input  $x$  to a single value, and collection of these can be turned into a probability distribution function through Gibbs distribution. The probability density  $P(x)$  can be represented as:

$$P(y | x) = \frac{e^{-E(x,y)/T}}{e^{-E(x)/T}} \quad (1)$$

The parameter  $T$  is a temperature function, and its value can be considered according to the data. The energy-based model can be linked to machine learning, especially in any discriminative model where the high dimensional data, converted to a single dimensional value, called logits. Softmax function can be used to derive the categorical distribution of logits, which can be seen in equation (2).

$$P(y | x) = \frac{e^{f_y(x)/T}}{\sum_i^K e^{f_i(x)/T}} \quad (2)$$

Combining equations (1) and (2), it can be inferred that the free energy function  $E(x;f)$  can be written as:

$$E(x: f) = -T. \log \sum_i^K e^{f_i(x)/T} \quad (3)$$

In this work, energy could not serve as a sole function that could help identifying the distribution. Energy scores can then be used as trainable cost function and can be treated as a regularization term along with the cross-entropy loss function. It helps to learn better the in-distribution and out-distribution data, and an energy gap can be created where the out-of distribution data is segregated from the in-distribution data. In equation (4), the energy regularization term is added to the existing cross-entropy loss.  $F_y(x)$  is the traditional softmax scoring function and  $\alpha$  is the regularization parameter. Equation (5) which represents the cost function with two squared hinge loss terms which penalize any value exceeding the in-distribution and any value below the out-distribution margin., where are marginal values for the in-distribution and out-distribution, respectively.

$$\min_w E_{(x,y) \sim D_{in}^{train}} [-\log F_y(x)] + \alpha. L_{energy} \quad (4)$$

$$L_{energy} = E_{(x_{in},y) \sim D_{in}^{train}} (\max(0.01 * (E(x_{in}) - n_{in}), E(x_{in}) - n_{in}))^2 + E_{x_{out} \sim D_{in}^{train}} (\max(0.01 * (n_{out} - E(x_{out})), n_{out} - E(x_{out})))^2 \quad (5)$$

In this work, we have updated the regularization term in the loss function (Liu, 2020) to avoid the vanishing gradient problem. We have used a  $0.01 * (E(x_{in}) - n_{in})$  term in both the hinge losses on the lines of leaky ReLU. The loss function now penalizes the samples in energy values in the range of  $n_{in}$  and  $n_{out}$ .

### 3. Numerical Simulation

#### 3.1. Datasets

The benchmark Tennessee Eastman dataset is used for testing the proposed algorithm for fault detection. There are 41 measured and 12 controlled variables in the benchmark Tennessee Eastman dataset and twenty-one faults are available for fault detection. In particular, we have considered Faults 6, 9, 15 and 19 for this study.

#### 3.2. Simulation

For training the neural networks, we have employed the novel energy loss coupled with the cross-entropy loss. For the encoder, 64 neurons were considered in the first layer, followed by 32,16 and 8 neurons in the subsequent layers. Similarly, 16,32 and 64 neurons were considered for the decoder. Adam optimizer was used for the auto-encoder network, and the newly framed loss function was used. For the auto-encoder network, the energy threshold was found to be 0.030, and any value above the threshold value was considered a fault. Similarly, for the LSTM network, which was used for prediction in a sliding window approach, 40 neurons were used in the first layer, 20 in the second layer and the Adam optimizer was used. A dropout of 0.2 was utilized. LSTM-AE had one hundred neurons, followed by 50 neurons for the encoder section and 50, and 100 for the decoder section with a dropout of 0.5 and Equation (5) was used as the loss function. Similarly, the proposed methodology was also compared to the existing machine learning models such as the PCA and DPCA for the faults considered below.

### 3.3. Results and Discussion

From the parameters taken into consideration, it can be seen from Table 1 and Table 2 that the Auto-encoder network with energy out of distribution (EOD) performed better as compared to other fault detection methods. Table 1 represents the data for Faults 6 and 9 (average of both) for the Tennessee Eastman dataset, and Table 2 is the average of the faults 15 and 19 considered for the Tennessee Eastman dataset. Faults 6, 9, 15 and 19 were considered for the study, as they capture different types of faults in the Tennessee Eastman data. Faults 1-7 represent the step type, 8-12 are random variations, 13 is slow drift, 14 and 15 are sticking types, whereas 16-20 are unknown. Fault 21 is a constant position fault. In our work, we have tried to incorporate the step, random variations, sticking and unknown fault types.

Table 1: Performance Metrics for Faults 6 and 9 using the proposed loss function.

Methodology	Recall	Precision	F1-Score	AUC	FP	FN
<b>Auto-Encoder EOD</b>	<b>0.9405</b>	<b>0.98</b>	<b>0.9598</b>	<b>0.944</b>	<b>5</b>	<b>8</b>
<b>LSTM EOD</b>	0.8695	0.93	0.8988	0.862	12	18
<b>LSTM-AE EOD</b>	0.9008	0.921	0.9107	0.918	8	13
<b>PCA</b>	0.961	0.71	0.8166	0.831	27	3
<b>DPCA</b>	0.985	0.69	0.816	0.845	31	2

Table 2: Performance Metrics for Fault 15 and 19 of Tennessee Eastman dataset using the proposed loss function.

Methodology	Recall	Precision	F1-Score	AUC	FP	FN
<b>Auto-Encoder EOD</b>	<b>0.918</b>	<b>0.96</b>	<b>0.9385</b>	<b>0.941</b>	<b>7</b>	<b>10</b>
<b>LSTM EOD</b>	0.81	0.88	0.843	0.828	14	22
<b>LSTM-AE EOD</b>	0.87	0.92	0.8943	0.871	10	14
<b>PCA</b>	0.912	0.49	0.6374	0.656	51	7
<b>DPCA</b>	0.978	0.53	0.6928	0.765	47	3

Auto-encoder EOD, the network coupled with the proposed loss function had a higher Recall and Precision value, thus accounting to lower False Positives and False Negatives, which is a desirable result. Auto-encoder networks in the traditional environment perform better when compared to the other networks. LSTM-AE, a combination of two frameworks, turns out to be the second best as indicated in Table 1 and 2.

#### 4. Conclusion

The proposed work utilizes the concept of energy out-of-distribution for fault detection. To this extent, higher dimensional data is converted to a vector of a single dimension. This data mapping helps to formulate the energy values, which are then identified as in-distribution and out-distribution based on a suitable threshold. The energy values exceeding the threshold are classified as faults. Traditional methods such as Auto-encoders, LSTMs and LSTM-AEs were coupled with the proposed loss function for fault identification. The methodology was also validated on the benchmark Industrial Tennessee Eastman Dataset. The results indicate that the proposed approach, Auto-encoder EOD, when compared to the other deep learning and machine learning techniques yielded significant results.

#### Acknowledgements

The authors would like to gratefully thank the Indian Institute of Technology, Delhi, for providing the necessary research facilities to carry out the proposed work.

#### References

- Kodamana, H., Raveendran, R., & Huang, B. (2017). Mixtures of probabilistic PCA with common structure latent bases for process monitoring. *IEEE Transactions on Control Systems Technology*, 27(2), 838-846.
- Liu, W., Wang, X., Owens, J., & Li, Y. (2020). Energy-based out-of-distribution detection. *Advances in Neural Information Processing Systems*, 33, 21464-21475.
- Du, Y., & Mordatch, I. (2019). Implicit generation and generalization in energy-based models. *arXiv preprint arXiv:1903.08689*.
- Sakurada, M. and Yairi, T., 2014, December. Anomaly detection using autoencoders with nonlinear dimensionality reduction. In *Proceedings of the MLSDA 2014 2nd workshop on machine learning for sensory data analysis* (pp. 4-11).
- Nguyen, H.D., Tran, K.P., Thomassey, S. and Hamad, M., 2021. Forecasting and Anomaly Detection approaches using LSTM and LSTM Autoencoder techniques with the applications in supply chain management. *International Journal of Information Management*, 57, p.102282.
- Yang, J., Guo, Y. and Zhao, W., 2019. Long short-term memory neural network based fault detection and isolation for electro-mechanical actuators. *Neurocomputing*, 360, pp.85-96.
- Liang, S., Li, Y., & Srikant, R. (2017). Enhancing the reliability of out-of-distribution image detection in neural networks. *arXiv preprint arXiv:1706.02690*.
- Hendrycks, D., Mazeika, M., & Dietterich, T. (2018). Deep anomaly detection with outlier exposure. *arXiv preprint arXiv:1812.04606*.
- Rani, J., Roy, A. A., Kodamana, H., & Tamboli, P. K. (2023). Fault detection of pressurized heavy water nuclear reactors with steady state and dynamic characteristics using data-driven techniques. *Progress in Nuclear Energy*, 156, 104516.

# Development of a Centralized Classifier for Decentralized Decision Making

Marco S. Reis,<sup>a,\*</sup> Eugeniu Strelet,<sup>a</sup> Joel Sansana,<sup>a</sup> Margarida J. Quina,<sup>a</sup> Licínio M. Gando-Ferreira,<sup>a</sup> Tiago J. Rato,<sup>a</sup>

<sup>a</sup>*Univ Coimbra, CIEPQPF, Department of Chemical Engineering, Rua Silvio Lima, Pólo II – Pinhal de Marrocos, 3030-790 Coimbra, Portugal*

\**marco@eq.uc.pt*

## Abstract

With the current development pace of IoT technology and analytical instrumentation, the acquisition of information from devices or operations that are geographically distributed is becoming increasingly common. In some applications, these devices/operations are essentially alike but executed with slightly distinct machines/measurement systems. Such differences have been handled by constructing local models, in a fully decentralized and independent way. However, it is now possible to conceive improved centralized schemes to derive a unified model that takes advantage of the entire data lake. In this work, we demonstrate how such an endeavor can be achieved in the scope of a real case study of predicting an important property (coagulation) of waste lubricant oil (WLO) using FTIR spectra collected at different locations (laboratories). The unified classifier uses a compound mapping of Partial Least Squares for Discriminant Analysis (PLS-DA) and the Bayesian linear classifier. The best prediction models were selected by screening over 36 potential models. The best candidate models to be implemented in practice showed an accuracy greater than 98 %, while retaining less than four latent variables.

**Keywords:** Decentralized classification; Waste lubricating oil; Standardization methods; Partial Least Squares for Discriminant Analysis.

## 1. Introduction

The development of analytical solutions for decision-making and operational management using data collected from geographically distributed devices is becoming increasingly common (Marr, 2015, Qin, 2014, Reis et al., 2021). The conventional approach adopted consists of deriving local models to handle the specificities of the devices or operations. However, this poses several problems of scalability, maintenance, and comparability. Therefore, new centralized methodologies have been developed. The overarching goal is to integrate the information collected from multiple sources and derive a unified model that takes advantage of the entire data lake. A variety of model transfer approaches were developed for handling spectroscopic data (Mishra et al., 2021, Nikzad-Langerodi et al., 2021). Techniques such as Partial Least Squares (PLS) model inversion (Jaeckle et al., 2000), Joint-Y PLS (García Muñoz et al., 2005), calibration transfer (Diaz et al., 2022), and more recently transfer learning (Mishra et al., 2021, Nikzad-Langerodi et al., 2021) and domain adaptation (Pan et al., 2011), offer different paths to achieve the aforementioned goal. In general, they assume the existence of a master or reference device, whose model needs to be transferred to the remaining devices. However, we will consider here the case where there is no such master device or

operation, all of them having symmetrical roles and importance, but where a common model still needs to be found for the final classification task. We demonstrate how such an endeavor can be achieved in the scope of a real case study, briefly described below. Sogilub is the national organization that manages the entire supply chain for the regeneration of waste lubricant oil (WLO) in Portugal. WLOs are collected from different locations but should comply with certain specifications to proceed for regeneration. Currently, WLO characterization is made at four different accredited laboratories, using equipment that, even though similar (Fourier-Transform Infrared spectrometers), show some differences, that originate slightly different spectra even for the same samples. The goal of the present work is to develop a centralized classifier that all laboratories can use, despite the local differences on the devices. In other words, instead of developing a classifier for each laboratory, with much less data and representativeness of the entire population of WLO, we aim at constructing a unified, centralized classifier, that benefits from the information collected in all laboratories. The proposed centralized approach is based on a combination of data pre-processing, a projection operation to a common space (enabling data integration) and finally modeling of the projected data.

The rest of this article is organized as follows. In Sections 2 the standardization techniques, pre-processing techniques and modeling methodologies considered in this work are briefly introduced. Afterwards, in Section 3, the results are presented. Finally, a summary of the conclusions is provided in Section 4.

## 2. Methodology

The proposed methodology aims to build a centralized classifier based on FTIR spectra of WLO to predict a key property in multiple locations (laboratories). The methodology is composed by three main stages addressing (Stage 1) spectral standardization, (Stage 2) spectral pre-processing and (Stage 3) training of the unified classification model as described in the following subsections. A schematic representation of the processed procedure is presented in Figure 1.

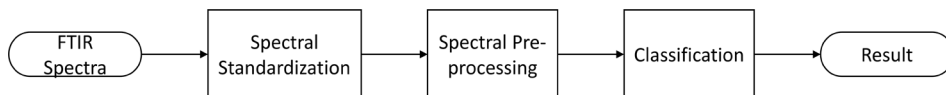


Figure 1 Diagram of the main stages of the centralized classifier.

### 2.1. Stage 1: Spectral Standardization

In this framework, samples are collected from the same population, but present distinct correlation structures due to the use of different measurement equipment. Therefore, they must first be brought to a common space so that the spectra become source-independent. To achieve this, standardization techniques are used to relate the spectra measured in a reference equipment ( $\mathbf{S}_1 \in \mathbb{R}^{N \times M_1}$ , where  $N$  is the number of samples and  $M_i$  is the number of wavelengths measured in equipment  $i$ ) to the spectra of the same samples obtained on a secondary equipment ( $\mathbf{S}_2 \in \mathbb{R}^{N \times M_2}$ ) by a transformation matrix ( $\mathbf{F} \in \mathbb{R}^{M_2 \times M_1}$ ) composed of linear relationships, such that  $\mathbf{S}_1 = \mathbf{S}_2\mathbf{F}$ . In this work two standardization techniques were considered to determine the transformation matrix:

- Direct Standardization (DS) (Wang et al., 1991);
- Piecewise Direct Standardization (PDS) (Feudale et al., 2002).

### 2.2. Stage 2: Spectral Pre-Processing

After projecting the spectra to a common space, it is necessary to correct for undesirable artifacts and enhance the relevant features in the spectra. To this end, several pre-

processing techniques are currently available. In this work, four pre-processing techniques were considered:

- Mean Centering (MC);
- Multiplicative Scatter Correction (MSC) (Martens et al., 1983);
- Standard Normal Variate (SNV) (Barnes et al., 1989);
- Savitzky-Golay Differentiation (SGD) (Savitzky et al., 1964).

Different combinations of these pre-processing techniques, as well as different parameterizations of SGD (derivative order, “D#”; window size, “W#”; and polynomial order, “O#”) were examined (see Table 2), leading to 12 distinct pre-processing variations.

### *2.3. Stage 3: Classification in the Common Space*

The final stage of the proposed procedure comprises the training of a classification model. In this framework, classification was made through a two stages procedure by application of Partial Least Squares - Discriminant Analysis (PLS-DA) (Barker et al., 2003), followed by Fisher Discriminant Analysis (FDA) (Hastie et al., 2009). In this case, the response variable of PLS-DA is an indicator variable that codifies the class labels. Afterwards, PLS is applied to extract the latent variables with greater discriminative power. Next, the extracted latent variables are fed to FDA to obtain the final classifier. As the performance of the classifier may be affected by the wavelengths included in the model, three PLS-based approaches were considered to select the relevant wavelength intervals:

- PLS (Wold et al., 2001);
- Forward interval PLS (FiPLS) (Xiaobo et al., 2007);
- Backward interval PLS (BiPLS) (Xiaobo et al., 2007).

## **3. Results**

### *3.1. Case Study*

The case study concerns the regeneration process of WLO conducted in four facilities with different geographical localizations. Before processing, the coagulation potential of the collected WLO is determined through a coagulation test using an alkaline treatment with KOH. If a certain WLO sample gives a positive result in the coagulation test, the associated lot must follow an alternative treatment other than regeneration. Otherwise, the WLO may clog the equipment, causing a premature stop of the process for cleaning and subsequent disposal of the entire production. Therefore, it is highly desirable to have a fast and reliable prediction of the coagulation potential to direct the WLO to a suitable treatment. To achieve this, it is proposed to predict the coagulation potential using the FTIR spectra of WLO samples.

This study uses two datasets: (i) a dataset composed by WLO samples (spectra) analyzed in four laboratories, which is used for selecting the standardization technique and (ii) another dataset composed by WLO samples (spectra and the respective coagulation label) used to train and test the calibration models. In all laboratories, the spectra were collected in the transmittance mode for 700 wavenumbers in the range of 4000 to 500  $\text{cm}^{-1}$  at a resolution of 5  $\text{cm}^{-1}$ .

### *3.2. Setting up the Spectral Standardization to the Common Space*

The first dataset is composed of six samples of WLO whose FTIR spectra were measured in all four laboratories with three replicates each. By applying Principal Component Analysis (PCA) (Jolliffe et al., 2016) to the spectra, the differences between laboratories can be inspected through analysis of the scatterplot of the first two principal components (explaining 92 % of the variability in the spectra) represented in Figure 2. In Figure 2 (a)



it is verified that even though the spectra are from the same set of samples, there are systematic deviations in the spectra due to the use of different equipment. This motivates the use of a standardization technique to bring the spectra to a common space. In this regard, Figure 2 (b) shows a significant improvement in homogeneity between the different laboratories after standardization.

In this work, DS and PDS were adopted using Laboratory II as the reference laboratory. To avoid overfitting, the transformation matrices of DS and PDS were obtained by PLS. To determine the best standardization technique (DS or PDS), their quality was assessed using the root mean squared error (RMSE) between the standardized and reference spectrum of each sample. These results are shown Table 1 along with a two-sided paired  $t$ -test. The results show that DS tends to produce standardized spectrum with lower RMSE. Furthermore, visual inspection of the standardized spectra shows that PDS introduced some artifacts in the transformed spectra. Thus, DS was selected as the best standardization technique for the current case study.

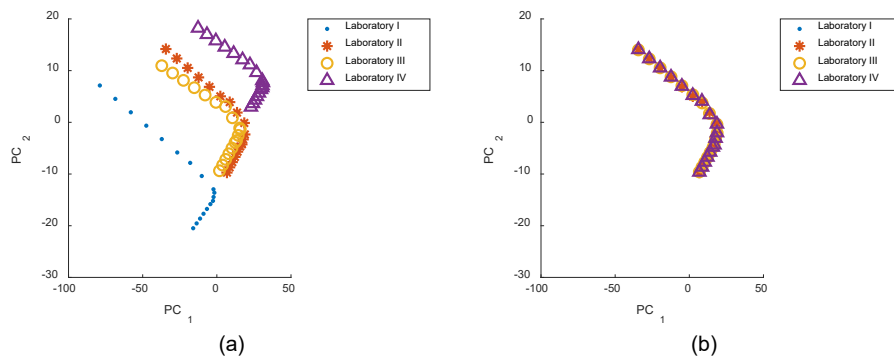


Figure 2 Scatterplot of the first and second principal components of samples measured in four different laboratories: (a) original spectra; (b) spectra after DS standardization.

Table 1 Comparison of the RMSE of the standardization techniques in each laboratory using a two-sided paired  $t$ -test.

Laboratory	RMSE		RMSE Comparison (DS – PDS)	
	DS	PDS	Test statistic ( $t$ )	p-value
I	0.2324	0.3814	-4.1525	<0.001
III	0.3559	0.4317	-0.5264	0.6054
IV	0.4727	0.4934	-1.8912	0.0758

### 3.3. Setting up and Selection of Calibration Model

The second dataset is used to train and test the centralized classification model. This dataset is composed of 120 samples (16 samples from Laboratory I; 58 samples from Laboratory II; 30 samples from Laboratory III; and 16 samples from Laboratory IV). Among these samples, 60 of them are from WLO that coagulate, and the other 60 samples are from WLO that do not coagulate.

Following the previous results, the samples were standardized by DS into a common space defined by Laboratory II. The samples were then used to train and select the best classification model from a pool of 36 distinct models (resulting from the combination of 12 pre-processing techniques and three modeling approaches). For FiPLS and BiPLS the spectra were divided into 15 intervals with approximately 47 wavenumbers each.

The raw dataset was randomly split into a training dataset with 80 % of the samples and a test dataset with the remaining 20 % of the samples. The models were then trained on

the training dataset using Monte Carlo Cross-Validation (MCCV) for tuning the model's hyperparameters. Afterwards, their performance was assessed on the test dataset. In both cases, the key performance indicator (KPI) to assess the model's performance was the prediction accuracy given by,

$$H = \frac{TP + TN}{n} \quad (1)$$

where  $TP$  is the number of true positives,  $TN$  is the number of true negatives and  $n$  is the number of samples in the training/test dataset. The accuracy varies from 0 to 1, where values closer to 1 relate to better classification capabilities.

The accuracy and number of retained latent variables of each model under consideration is presented in Table 2. The results show that all models attain high accuracy. However, several models also retain a considerably large number of latent variables, indicating that some overfitting may be taking place. Nevertheless, a few models present a good balance between the number of retained latent variables and accuracy. For instance, SGD-D2-W9-O2 in combination with FiPLS has an accuracy of 100 % with four latent variables, while SNV-SGD-D1-W7-O2 in combination with BiPLS has an accuracy of 98 % with only two latent variables.

Table 2 Accuracy and number of retained latent variables of each model.

Pre-Processing Technique	Accuracy			Number of Retained Latent Variables		
	FiPLS	BiPLS	PLS	FiPLS	BiPLS	PLS
MC	0.9960	1.0000	0.9710	5	9	4
SNV	0.9360	0.9757	0.9757	4	5	5
MSC	0.9998	0.9707	0.9640	5	4	4
SGD-D1-W7-O2	1.0000	1.0000	1.0000	7	8	8
SGD-D1-W15-O2	0.9452	0.9525	0.9527	7	8	8
SGD-D2-W9-O2	1.0000	0.9993	0.9400	4	7	3
SNV-SGD-D1-W7-O2	0.9843	0.9836	0.9990	6	2	2
SNV-SGD-D1-W15-O2	1.0000	0.9998	0.9827	6	5	4
SNV-SGD-D2-W9-O2	0.9683	0.9571	0.9788	6	4	4
MSC-SGD-D1-W7-O2	0.9843	1.0000	0.9610	4	7	4
MSC-SGD-D1-W15-O2	0.8866	0.8736	0.8818	5	2	2
MSC-SGD-D2-W9-O2	0.9752	0.9995	0.9948	6	5	4

#### 4. Conclusions

A centralized classification model was developed to predict the coagulation potential of WLO during the recovery process. Predictions are made from FTIR spectra collected in four laboratories, which introduces laboratory-specific variability in the data. The proposed methodology was able to handle this heterogeneity, by projecting all sources to a common space. In this regard, DS provided the best results. The best prediction model was then selected by screening over 36 potential models. The results showed that FiPLS and BiPLS in combination with Savitzky-Golay Differentiation led to prediction accuracies greater than 98 %, while retaining a low number of latent variables. Thus, these models are the best candidates to be implemented in practice. In future work, we will evaluate the performance of the centralized classifier over time and assess the impact of instrument and/or sample degradation on model quality.

## Acknowledgments

The authors gratefully acknowledge the financial support of SOGILUB – Sociedade de Gestão Integrada de Óleos Lubrificantes Usados, Lda. The authors also acknowledge the support from the Chemical Process Engineering and Forest Products Research Centre (CIEPQPF), which is financed by national funds from FCT/MCTES (reference UIDB/EQU/00102/2020).

## References

- M. Barker, W. Rayens, 2003, Partial least squares for discrimination, *Journal of Chemometrics*, 17, 3, 166-173.
- R. J. Barnes, M. S. Dhanoa, S. J. Lister, 1989, Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra, *Applied Spectroscopy*, 43, 5, 772-777.
- V. F. Diaz, P. Mishra, J.-M. Roger, W. Saeys, 2022, Domain invariant covariate selection (Di-CovSel) for selecting generalized features across domains, *Chemometrics and Intelligent Laboratory Systems*, 222, 104499.
- R. N. Feudale, N. A. Woody, H. Tan, A. J. Myles, S. D. Brown, J. Ferré, 2002, Transfer of multivariate calibration models: a review, *Chemometrics and Intelligent Laboratory Systems*, 64, 2, 181-192.
- S. García Muñoz, J. F. MacGregor, T. Kourti, 2005, Product transfer between sites using Joint-Y PLS, *Chemometrics and Intelligent Laboratory Systems*, 79, 1, 101-114.
- T. Hastie, R. Tibshirani, J. Friedman, 2009, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, New York, NY, Springer.
- C. Jaecle, J. F. MacGregor, 2000, Product Transfer Between Plants Using Historical Process Data, *AIChE Journal*, 46, 10, 1989-1997.
- I. T. Jolliffe, J. Cadima, 2016, Principal component analysis: a review and recent developments, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374, 2065, 20150202.
- B. Marr, 2015, *Big Data - Using Smart Big Data Analytics to Make Better Decisions and Improve Performance*, Chichester, Wiley.
- H. Martens, S. A. Jensen, P. Geladi, 1983, In *Proc. Nordic Symp. Applied Statistics* (Ed, Christie, O. H. J.) Stokkland Forlag, Stavanger, Norway, pp. 205-234.
- P. Mishra, D. Passos, 2021, Deep calibration transfer: Transferring deep learning models between infrared spectroscopy instruments, *Infrared Physics & Technology*, 117, 103863.
- R. Nikzad-Langerodi, E. Andries, 2021, A chemometrician's guide to transfer learning, *Journal of Chemometrics*, 35, 11, e3373.
- S. J. Pan, I. W. Tsang, J. T. Kwok, Q. Yang, 2011, Domain Adaptation via Transfer Component Analysis, *IEEE Transactions on Neural Networks*, 22, 2, 199-210.
- S. J. Qin, 2014, Process data Analytics in the Era of Big Data, *AIChE Journal*, 60, 9, 3092-3100.
- M. S. Reis, P. M. Saraiva, 2021, Data-Centric Process Systems Engineering: a Push Towards PSE 4.0, *Computers & Chemical Engineering*, 155, 107529.
- A. Savitzky, M. J. E. Golay, 1964, Smoothing and Differentiation of Data by Simplified Least Squares Procedures, *Analytical Chemistry*, 36, 8, 1627-1639.
- Y. Wang, D. J. Veltkamp, B. R. Kowalski, 1991, Multivariate Instrument Standardization, *Analytical Chemistry*, 63, 23, 2750-2756.
- S. Wold, M. Sjöström, L. Eriksson, 2001, PLS-Regression: A Basic Tool of Chemometrics, *Chemometrics and Intelligent Laboratory Systems*, 58, 109-130.
- Z. Xiaobo, Z. Jiewen, L. Yanxiao, 2007, Selection of the efficient wavelength regions in FT-NIR spectroscopy for determination of SSC of 'Fuji' apple based on BiPLS and FiPLS models, *Vibrational Spectroscopy*, 44, 2, 220-227.

## Fault detection using Fourier neural operator

Jyoti Rani<sup>a</sup>, Tapas Tripura<sup>c</sup>, Umang Goswami<sup>a</sup>, Hariprasad Kodamana<sup>a,b</sup>, Souvik Chakraborty<sup>b,c</sup>

<sup>a</sup> *Department of Chemical Engineering, Indian Institute of Technology Delhi, New Delhi-110016, India*

<sup>b</sup> *Yardi School of Artificial Intelligence, Indian Institute of Technology Delhi, New Delhi-110016, India*

<sup>c</sup> *Department of Applied Mechanics, Indian Institute of Technology Delhi, New Delhi-110016, India*

### Abstract

In order to generate higher-quality products and increase process efficiency, there has been a strong push in the processing and manufacturing sectors. This has called for the creation of methods to identify and fix faults to ensure optimal performance. As a result, it is essential to develop monitoring systems that can effectively detect and identify these faults so that operators can quickly resolve them. This article proposes a novel fault detection method that adopts a deep learning approach using a Fourier neural operator (FNO) in a probabilistic way, an operator learning model that aims to learn the distribution of multivariate process data and apply them for fault detection. Herein, the historical data under normal process conditions were first utilized to construct a multivariate statistical model; after that, the model was used to monitor the process and detect faults online. The proposed FNO combines the integral kernel with Fourier transformation in a probabilistic way. As the Fourier transform helps in the time-frequency localization of time series, FNO takes advantage of them to discover the complex time-frequency characteristics underlying multivariate datasets. On the benchmark Tennessee Eastman process (TEP), a real-world chemical manufacturing dataset, the performance of the proposed method was demonstrated and compared to that of the widely used fault detection methods.

**Keywords:** Fault detection; Fourier Transform; Neural operator; Probability distribution

### 1. Introduction

A fault in industrial manufacturing processes is any anomalous deviation from normal operating conditions (NOC). Faults are a reason for concern since even minor flaws in a industrial system can set off a cascade that reduces efficiency and dependability. On the other hand, numerous cutting-edge fault detection techniques are accessible to ascertain whether the issue has occurred in the system (also known as anomaly detection in other applications). In recent decades, multivariate statistical techniques for process monitoring, especially fault detection, have developed rapidly (Kodamana et al., 2019). The principal component analysis (PCA) is mainly used in linear data processing algorithms and has proven to be highly effective in fault detection. Dynamic principal component analysis (DPCA) has been suggested in the literature to cope with serially correlated multivariate observations. Both PCA and DPCA have a linear structural model and correlation constrain while dealing with in-process dynamics (Rani et al., 2023). However, extremely non-linear and highly correlated process data pose challenges. Simultaneously, Neural network (NN) based techniques have gained considerable attention due to their flexibility and capability in modelling complex structures and temporal dynamics. Researchers have applied deep learning to the field of process monitoring, using techniques such as long short-term memory (LSTM), autoencoder (AE), Recurrent neural network (RNN), and LSTM-based AE (Lee et al., 2019). It is challenging for data-driven neural networks (NNs) to generalize beyond the training data. As an efficient and successful strategy, the trained networks should generalize to unseen inputs and predict the output precisely to

address these challenges; neural operators (NOs) have been developed recently. Since NOs only need to be trained once, they can learn the mapping among two infinite-dimensional function spaces (Kovachki et al., 2021). After being trained, they can be applied to predict the outcome of any particular input function.

The global integral operators are passed through the non-linear local activation functions via NOs, similar to NNs, to learn the complex non-linear dynamics of process data. Among the recently developed NOs, the DeepONet architecture is the first to learn such infinite-dimensional function spaces (Lu et al., 2022). Similarly, a graph neural operator (GNO) is proposed that gave new ideas for using graphs but became unstable with the increased number of hidden layers. In order to learn the network parameters in Fourier space, the Fourier neural operator (FNO) is invented, simultaneously (Li et al., 2020). The fast Fourier transform (FFT) spectral decomposition of the input and computing the integral convolution kernel in the Fourier space constitute the basis of the FNO (Li et al., 2020). The present work extends the FNO in a probabilistic sense and proposes the Fourier neural operator (FNO) for learning the distribution of finite-dimensional multivariate input. The proposed FNO is then applied for fault detection.

## 2. Methodology

### 2.1. Proposed Framework

This paper presents the FNO algorithm for learning the distribution of provided time series. To learn the distribution of the framework, the output space data was created by splitting the input data set. The input-output pairs were denoted as  $\mathbf{Y} \in \mathcal{U}$  and  $\bar{\mathbf{Y}} \in \mathcal{U}$ , where  $\mathcal{U} \in \mathbb{R}^N$  is a completely normed vector space. The distribution of the input was characterized by using the FNO, and a new channel was created by augmenting an  $m$ -dimensional random variable  $\mathbf{Z} = \{\xi_1, \dots, \xi_k, \dots, \xi_T\}$  with  $\xi_k \in \mathbb{R}^N$ . The purpose of this framework is to develop a network that can utilise the information to learn about the distribution employing the operator.

To do this, first, lift the inputs  $a(y_t) \in \mathcal{A}$  to a high-dimensional space via a local transformation  $M(a(y_t)): \mathbb{R}^{d_a} \mapsto \mathbb{R}^{d_v}$  and denote the high-dimensional space as  $v_0(y_t) = M(a(y_t))$ , where  $v_0(y_t)$  has values in  $\mathbb{R}^{d_v}$ . In a neural network framework, the local transformation  $M(a(y_t))$  can be accomplished by building a shallow, fully connected neural network (FNN). In addition, it was assumed that a total of  $n$  steps are necessary to achieve convergence. After that,  $n$  updates  $v_{j+1} = F(v_j); \forall j = 1, \dots, n$  was applied to  $v_0(y_t)$ , where the function  $F$  returns a value in  $\mathbb{R}^{d_v}$ . Once all of the  $n$ -iterations were performed, then another local transformation  $G: \{v_n(y_t) \in \mathbb{R}^{d_v}\} \rightarrow \mathbb{R}^{d_u}$  was applied using an FNN to transform the lifted space into the output space  $\bar{\mathbf{Y}} \in \mathcal{U} = G(v_n(y_t))$ . The step-wise update is given in Equation 1.

$$v_{j+1}(y_t) = Q\left(\left(\mathcal{K}(a; \phi) * v_j\right)(y_t) + \check{\partial}v_j(y_t)\right); t \in [0, T] \quad (1)$$

where  $Q(\cdot)$  denotes a non-linear activation function,  $\check{\partial}: \mathbb{R}^{d_v} \rightarrow \mathbb{R}^{d_v}$  is a linear transformation,  $*$  is the convolution operator, and  $\mathcal{K}$  denotes the wavelet integral operator parameterized by  $\phi \in \Theta_{NN}$ . The convolution in the integral kernel operator is represented as,

$$\left(\mathcal{K}(\phi \in \Theta_{NN}) * v_j\right)(y_t) = \int_T \kappa(y_t - \tau; \phi)v_j(\tau)d\tau; t \in [0, T] \quad (2)$$

where  $\kappa(\cdot, \cdot)$  is the kernel of the neural network parameterized by  $\phi \in \Theta_{NN}$  in the Fourier domain. The overarching framework of the FNO consisting of the Fourier integral layer (FIL) is illustrated in Figure 1, which is further discussed in sub-section 2.1.

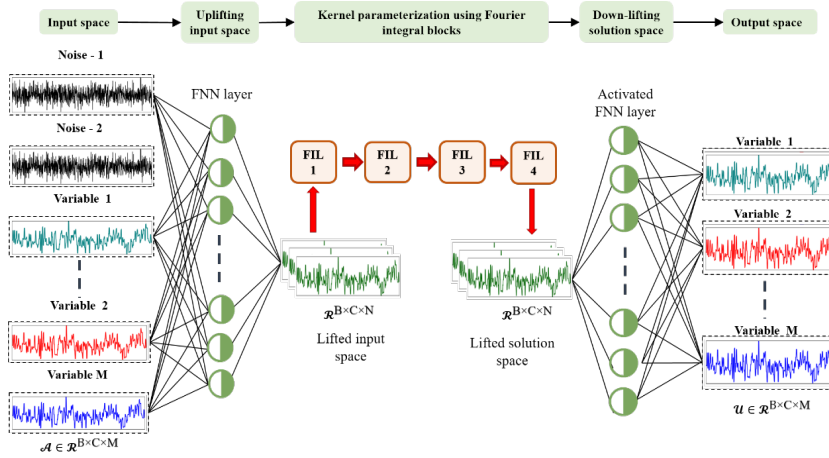


Figure 1: Architecture of Probabilistic Fourier Neural Operator. It has  $M+2$  input consisting of  $M$  variables and 2-time series noise inputs. First, a higher dimensional projection of the inputs was made by utilizing an FNN space. After this, 4 Fourier internal blocks were applied to the projected data. The results from the other FNN were then used to transform the fourth Fourier integral block back to the original domain. This way, it will learn the distribution, and at the output, it will reconstruct only the  $M$  original variable without the noise.

## 2.2. Fourier Integral Block

In this section, the kernel was discovered in the frequency domain by constructing a finite-dimensional space with a Fourier transform. After that, the convolution operator was substituted in Equation 2, defined in Fourier space for the kernel integral operator. Assume  $\mathcal{W}$  represent the Fourier transform (Equation 3) of a function  $f \in \mathbb{R}^{d_v}$  and  $\mathcal{W}^{-1}$  denotes the inverse, then

$$(\mathcal{W}f)_l(\omega) = \int_D f_l(y) e^{-2i\pi(\tau, \omega)} d\tau, \quad (\mathcal{W}^{-1}f)_l(\tau) = \int_D f_l(\omega) e^{2i\pi(\tau, \omega)} d\omega \quad (3)$$

for  $l = 1, \dots, d_v$  where  $i = \sqrt{-1}$  is the imaginary unit. By letting  $\kappa_\phi(y, p, a(y), a(p)) = \kappa_\phi(y - p)$  in Equation 2 and applying the convolution theorem, the convolution was defined in the Fourier domain as Equation 4.

$$(\mathcal{K}(\phi)v_j)(\tau) = \mathcal{W}^{-1} \left( \mathcal{W}(\kappa_\phi) \cdot (\mathcal{W}v_j) \right) (\tau) \quad (4)$$

As parameterizing of  $\kappa_\phi$  occurs directly in Fourier space, so introducing a weight tensor within the Fourier domain makes better sense as  $R_\phi = \mathcal{W}(\kappa_\phi)$ . With the new information, Fourier integral layer is defined in Equation 5.

$$(\mathcal{K}(\phi)v_j)(\tau) = \mathcal{W}^{-1} \left( R_\phi \cdot (\mathcal{W}v_j) \right) (\tau) \quad (5)$$

In the frequency mode,  $\omega$  have  $(\mathcal{W}v_j)(\omega) \in \mathbb{C}^{d_v}$  and  $R_\phi(\omega) \in \mathbb{C}^{d_v \times d_v}$ , and discrete modes  $\mathbb{Z}^d$  can be used since  $\kappa$  is periodic, as it admits. Thus, a Fourier series expansion and continuous modes become more redundant and computationally costly. Therefore, a finite-dimensional parameterization was selected by truncating the Fourier series at the maximum number of modes,  $\omega_{\max} = |Z_{\omega_{\max}}| = |\{\omega \in \mathbb{Z}^d: |\omega_l| \leq \omega_{\max, l}, \text{ for } l = 1, \dots, d\}|$ . After this, the parameterization of  $R_\phi$  occurs directly as a set of complex-valued  $(\omega_{\max} \times d_v \times d_v)$ -tensors made up of truncated Fourier modes as a result. The  $\ell_1$ -norm of  $\omega \in \mathbb{Z}^d$  is typically given an upper bound in order to characterize the low-

frequency modes. Thus, computational and memory bottleneck was overcome by deciding  $Z_{\omega_{\max}}$  since it enables a practical implementation. In this work, the Fourier integral layers were constructed by using  $Z_{\omega_{\max}} = 16$ .

### 2.3. FNO-based Process Monitoring

Once the FNO model is constructed, an FNO provides a reconstructed version of the input together with the distribution mean and standard deviation. Firstly, define the likelihood distribution  $p_{\mathbf{X}|\bar{\mathbf{Y}}}(Y | \bar{Y})$  as Gaussian distribution with sample mean  $\mu_s$  and sample standard deviation  $\sigma_s$ . From the Gaussian representation of the likelihood function, a closed form for the reconstruction loss (Equation 6) was derived in terms of the log-likelihood as follows,

$$\log(p_{\mathbf{Y}|\bar{\mathbf{Y}}}(Y | \bar{Y})) = -\left(\log(\sigma_s) + \frac{1}{2}\left(\frac{y-\mu_s}{\sigma_s}\right)^2\right) \quad (6)$$

where  $\mu_s$  and  $\sigma_s$  are the mean and sample standard deviation of  $\bar{\mathbf{Y}}$ . The details steps leading to Eq. 6 are omitted for brevity. Further, a Monte Carlo framework was adopted to learn the parameters of the input distribution. For  $\bar{\mathbf{Y}}^{(l)}$  where  $l = 1, \dots, L$  were computed from the samples of  $P_{\theta}(\bar{y} | y)$ . After this, the standard deviation and mean were estimated for all variables. In the end, the presence of a fault in the data was tested using the condition standard deviation greater than the suitable threshold.

### 2.4. Datasets used in Validation Studies

Tennessee Eastman Process (TEP) dataset is a chemical simulation model based on a real-world chemical production process, and it has been widely implemented in the fields of process control, process optimization, and fault detection. There are five significant unit operations: a reactor, a product condenser, a recycle compressor, a vapor-liquid separator, and a product stripper. The TE simulator can simulate 21 different fault operation conditions in addition to the typical operating environment. Twenty-two continuous process variables, 19 composition variables, and 12 controlled variables make up the data set. In this investigation, 52 variables were monitored with fault IDs named ID1, ID6, ID7, ID8, ID16, ID18, and ID19, as all show different behaviour.

## 3. Simulation Studies

### 3.1. FNO Architecture

The four FNO layers that make up the FNO architecture were all triggered using the GeLU activation function. In order to improve the network settings, the ADAM optimizer was used. For this optimizer, the weight decay factor was set at  $10^{-6}$ , and the initial learning rate was set at 0.001. Depending on the data loader's underlying technology, the batch size ranges from 10 to 25. The numerical experiments were carried out on a Quadro RTX5000 32 GB GPU. Each dataset was normalized using Z-score and Min-Max scaling before training the models to prevent training bias toward metrics with higher values.

### 3.2. Fault Detection Results

The FNO model established in the preceding section was used in the suggested process monitoring technique. The process data under normal operating conditions were gathered and used as a training dataset after normalizing to zero mean and unit variance. Variability in data reconstruction was monitored to determine the probability that the online data have a distribution similar to that of normal data, and it was calculated by employing standard deviation. The monitoring index was computed for each process data in online monitoring, and the control limit was compared using normal data distribution. The process was declared out of control when a high abnormality occurred in the data distribution, and because the standard deviation exceeded the limit, the process was considered faulty. Firstly, compare the multivariate data samples produced by FNO to the

real samples supplied for the stable data to evaluate the model's effectiveness. As shown in Figure 2, the FNO could learn the distribution and accurately reconstruct TEP data after a sufficient number of epochs.

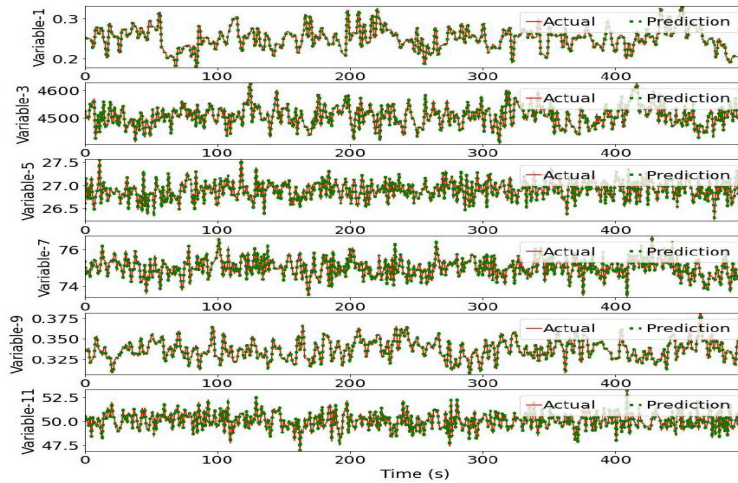


Figure 2: Reconstruction plot of TEP data for normal conditions using FNO

The standard deviation plot showed different behaviour when the fault was introduced within the dataset. The contribution plots of the factors contributing to the standard deviation at samples were computed to identify the variable responsible for any anomalous event, as shown in Figure 3. Therefore, this method aids in locating the variable also responsible for the unusual behaviour. In accordance with that, precision, recall, accuracy, and F1-score were calculated.

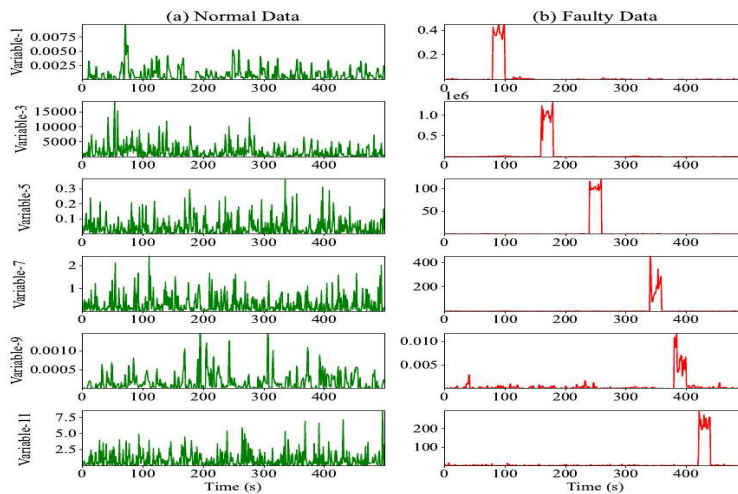


Figure 3: Standard deviation plot of normal data vs. faulty data for the Tennessee Eastman Process dataset

### 3.3. Performance Comparison of FNO with Baseline Methods

The fault detection performance of FNO was compared with other popular models such as DPCA, LSTM, and AE. When comparing the same number of noise levels among the other algorithms for each metric in Table 1, we see that the FNO has the best performance. Overall, FNO outperformed AE, LSTM, and other fault detection algorithms for the TEP dataset, as depicted in Table 1.



Table 1: Average performance metrics for Tennessee Eastman data set in the presence of the fault

Methods	Recall	Precision	F1-Score	AUC	FP	FN
AE	0.93	1.0	0.96	0.94	2	11
LSTM	0.80	0.87	0.83	0.81	11	21
LSTM-AE	0.86	0.91	0.88	0.87	7	16
DPCA	0.96	0.52	0.68	0.74	4	8
FNO	0.98	1.0	0.99	0.98	0	6

#### 4. Conclusion

In this work, a novel fault detection method using FNO architecture was developed and evaluated on an enlarged TEP dataset. A broad range of recurrent and convolutional architectures such as LSTM, AE, LSTM-based AE, and DPCA was compared to establish the baseline using F1-score and Accuracy values. The novel Fourier neural operator (FNO) technique showed its effectiveness in distribution learning and anomaly detection in multivariable time series data by showing F1 score value as 0.99, AUC value of 0.98 with no FP values which show the superiority of suggested method. The proposed model brings in threefold advantages. (i) It accurately learns the operator dynamics, (ii) the introduction of Fourier transform brings in time to frequency localization, and (iii) the probabilistic framework helps to work in a noisy environment.

#### 5. Acknowledgement

The authors acknowledge the Indian Institute of Technology, Delhi(IIT Delhi) for providing computational resources and a place to carry out this work.

#### References

- Kovachki, N., Li, Z., Liu, B., Azizzadenesheli, K., Bhattacharya, K., Stuart, A., & Anandkumar, A. (2021). Neural operator: Learning maps between function spaces. *ArXiv Preprint ArXiv:2108.08481*.
- Lee, S., Kwak, M., Tsui, K.-L., & Kim, S. B. (2019). Process monitoring using variational autoencoder for high-dimensional nonlinear processes. *Engineering Applications of Artificial Intelligence*, 83, 13–27. <https://doi.org/https://doi.org/10.1016/j.engappai.2019.04.013>
- Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., & Anandkumar, A. (2020). Fourier neural operator for parametric partial differential equations. *ArXiv Preprint ArXiv:2010.08895*.
- Lu, L., Meng, X., Cai, S., Mao, Z., Goswami, S., Zhang, Z., & Karniadakis, G. E. (2022). A comprehensive and fair comparison of two neural operators (with practical extensions) based on FAIR data. *Computer Methods in Applied Mechanics and Engineering*, 393, 114778. <https://doi.org/10.1016/j.cma.2022.114778>
- Kodamana, H., Raveendran, R., & Huang, B. (2019). Mixtures of Probabilistic PCA With Common Structure Latent Bases for Process Monitoring. *IEEE Transactions on Control Systems Technology*, 27(2), 838–846. <https://doi.org/10.1109/TCST.2017.2778691>
- Rani, J., Roy, A. A., Kodamana, H., & Tamboli, P. K. (2023). Fault detection of pressurized heavy water nuclear reactors with steady state and dynamic characteristics using data-driven techniques. *Progress in Nuclear Energy*, 156, 104516. <https://doi.org/https://doi.org/10.1016/j.pnucene.2022.104516>

# Proximal policy optimization for the control of mAB production

Nikita Gupta,<sup>a</sup> Shikhar Anand,<sup>a</sup> Deepak Kumar,<sup>a</sup> Manojkumar Ramteke,  
<sup>a,b</sup>Hariprasad Kodamana<sup>a,b\*</sup>

<sup>a</sup>*Department of Chemical Engineering, Indian Institute of Technology, New Delhi-110016, India*

<sup>b</sup>*Yardi School of Artificial Intelligence, Indian Institute of Technology, New Delhi-110016, India*

\**kodamana@iitd.ac.in*

## Abstract

The increasing demand of monoclonal antibodies (mAB) has been the motivation to develop methods to improve its yield of its production. The product yield for any chemical process highly depends on the fidelity of the process automation used. To this extent, deployment of model-based controllers results in heavy computational load, and its performance highly dependent on the model's accuracy. To address some of these issues, the role of model-free controllers like reinforcement learning (RL) based controllers seems to be promising. Proximal policy optimization (PPO) is one of the stochastic deep RL methods that can be used to achieve this task. In this work, we have studied the PPO algorithm for the control of mAB concentration in the bioreactor. In this process, mAB concentration is the control variable and feed flowrate is the manipulated variable. Our simulation study shows that by application of the PPO algorithm for bioreactor, we are able to track the end-point concentration of mAB, by providing the optimal flowrate strategy.

**Keywords:** Reinforcement Learning (RL), Proximal Policy Optimization (PPO), monoclonal antibodies (mAB)

## 1. Introduction

For many years monoclonal antibodies (mAB) have been used to treat several complex human diseases like arthritis, cancers, Alzheimer's, COVID-19, and HIV as they have less antagonistic effects. Due to these benefits, research has been going on to improve the yield of the mAB. Therefore, a reasonable control strategy needs to be developed for this purpose. The controller performance is highly affected even by small changes in process parameters. As stated in the literature, many model-based control strategies like iterative learning control (ILC) and model predictive control (MPC) proved to overcome these issues (Gupta et al.,2022, Kern et al.,2015, Lee et al.,2007). Model-based controllers are highly dependent on the model and moreover, they require more computational time as the optimization problem has to be performed at each time step. However, many data-driven modeling strategies for chemical processes to overcome the limitations of developing a first-principle-based model (Jiang et al.,2019). Data-driven techniques, such as model-free reinforcement learning (RL), technique tackle the major drawbacks of model-based controllers (Joshi et al.,2021, Singh et al.,2020, Nikita et al., 2021). In contrast to conventional controllers, a controller based on RL allows the agent, which is equivalent to the controller, to learn the best course of action to take in response to input from the environment (Yoo et al., 2021, Bao et al., 2021, Sutton et al., 2018). One of the promising algorithms that belong to the category of deep reinforcement learning for is

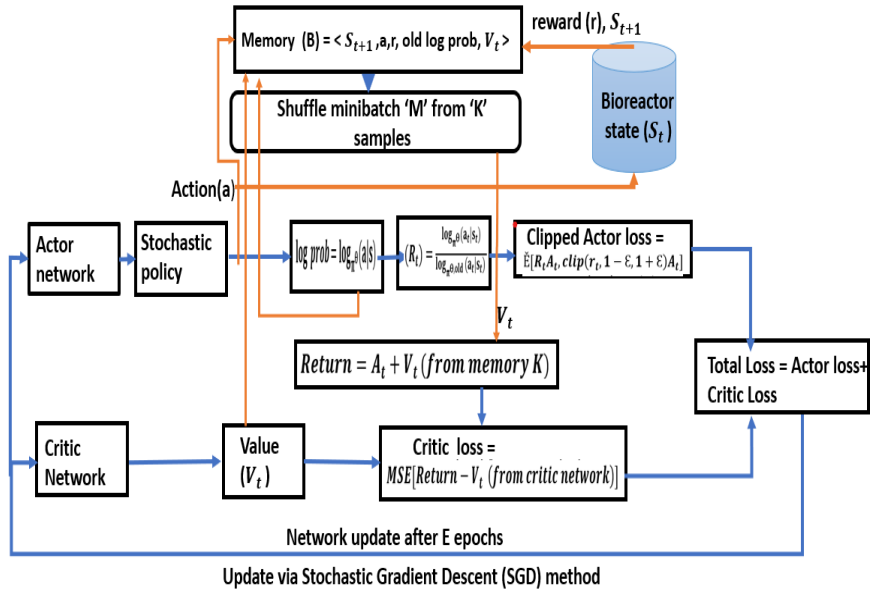
proximal policy optimization (PPO). PPO greatly reduces the complexity by adopting the clipping mechanism. The objective function uses specific clipping which prevents new policy to diverge away from the old policy (Schulman et al., 2017).

84% of all mAbs (monoclonal antibodies) are generated from hamster ovary tissue. These cells have been altered to grow in a media that has been chemically produced in order to produce the target protein of interest. The principal source of carbon and energy for cultured cells is glucose, which they primarily utilize to maintain the cell's general homeostasis (Kumar et al., 2021). That's why varying glucose concentration in media affects the growth of mammalian cell culture performance. Inlet glucose concentration can be varied by changing the inlet flow rate. Temperature also affects overall mammalian cell culture. Therefore, temperature or inlet glucose concentration or flowrate can be considered as a manipulated variable to enhance the protein (mAb) concentration, which is the control variable. But in the present work, only feed flowrate has been used as a manipulated variable for the same. In the present work, the PPO algorithm has been applied to identify the optimal operating points that enhances the antibody production in the bioreactor.

## 2. Methodology

### 2.1 Process description

The fermentation process is carried out in a bioreactor in which the inlet stream consists of feed and media having glucose as raw material. Glucose, dead cells, lactate, viable cells, and protein (monoclonal antibodies) are the constituents of the outlet stream (Kumar et al., 2021). The six states are viable cells ( $X_v$ ), monoclonal antibodies (mAB), glucose, lactate, dead cell density ( $X_d$ ) and viability. The detailed schematics of the process can be shown below:



**Figure 1: Schematic of the process**

In this process, our main objective is to maximize the concentration of monoclonal antibodies (mAB) by controlling the glucose concentration by varying the inlet flowrate.

Therefore, for this process mAB concentration is our state (S) and flowrate is our action (a). Here we have used proximal policy optimization (PPO), one of the RL techniques to obtain the optimal control policy for obtaining the desired concentration. PPO is an actor-critic method, which acts as an agent and provide optimal action ( $a_t$ ) to the bioreactor i.e., our environment at current state ( $S_t$ ) and provide us with next state ( $S_{t+1}$ ) and reward ( $r$ ). Based on the reward earned, the agent learns on continuous interaction with the environment to provide optimal control strategy as indicated in Figure 1. In this schematic, we have two networks, actor and critic network. Actor networks take state (s) as input and provide distribution of actions and one action is sampled from it. Critic network also take state (s) as input to give value function. This value function evaluates, weather the action is good or bad.  $R_t$  is evaluated as ratio of old log of probability of actions and new log probability of actions. If the value of ratio is 1, then current action policy is better but if it is between 0 and 1, then current action policy is not good. Advantage function is also a good measure of if action is good or bad. If value of advantage is positive, then probability of selecting those set of actions is good and if value of advantage is negative, then those actions are discarded. Product of ratio and advantage is very good optimization function. This ratio is clipped between a minimum and maximum value, so that new policy will not diverge much away from the old policy. This clipped function is termed as actor loss. Critic loss is evaluated using mean square error of Return and value (i.e., output from the critic network). Total loss is the sum of critic loss and actor loss, which is used to update the networks based on stochastic gradient descent method. We shuffle the minibatch size of 'M' from the 'K' samples selected from the memory 'B'. The networks are updated after every 'E' epochs, then memory 'K' is discarded. This process is repeated till the old policy weights converges with new policy weights.

## 2.2 Proximal policy optimization (PPO) algorithm

Reinforcement learning (RL) is one of the branches of machine learning (ML) in which the agent (analogous to the controller) learns to give optimal actions by maximizing the reward through interaction with the environment (bioreactor). RL-based algorithms can be divided into categories, value-based method in which agent learn action based on learned value function and policy gradient method in which policy is optimized using gradient descent method. PPO is an on-policy method that can be implemented in an actor-critic style. Here critic can be implemented as the value function approximator, and the actor can be implemented as a policy approximator. Therefore, performance of this method proved better than the plain policy gradient method (Byun et al.,2020, Wang et al.,2020).

PPO performs multiple gradient descent to update the neural network for each iteration. The stability of the system is ensured by calculating the ratio between old policy and the new policy. Here the main objective is to reduce the divergence between old and the new policy (Hämäläinen et al.,2020). The basic algorithm steps are stated below:

- 1 There are two networks i.e., actor network ( $\pi^\theta(s_t)$ ) and critic network ( $Q^\phi(s_t)$ ) with  $\theta$  and  $\phi$  as respective weights for current state ' $s_t$ '.
- 2 Initialize the memory B.
- 3 Collect set of tuples of <state ( $s_t$ ), action ( $a_t$ ), reward ( $r_t$ ), log probs, values ( $V_t$ )> at time instant 't' from memory B. Here 'log probs' is the logarithmic values of the probability of actions and values are the output from the critic network.
 
$$\log prob = \log_{\pi^\theta}(a|s) \quad (12)$$
- 4 Sample the set of 'K' samples from the memory 'B'.
- 5 Shuffle the minibatch size of 'M' from the 'K' samples collected form 'B' and perform 'E' epochs to update the network for each batch.
- 6 Calculate ratio ( $R_t$ ) =  $\frac{\log_{\pi^\theta}(a_t|s_t)}{\log_{\pi^{\theta,old}}(a_t|s_t)}$  i.e., ratio of log of probability of actions from memory and old log probabilities from the memory.

- 7 Update the actor network based on following clipped loss function:

$$L_{\pi\theta} = \mathbb{E}[R_t A_t, \text{clip}(r_t, 1 - \epsilon, 1 + \epsilon) A_t] \quad (13)$$

$$A_t = r + \beta V(s_{t+1}) - V(s_t) \quad (14)$$

Here  $A_t$  is the advantage, which considers for the goodness of state,  $\epsilon$  (epsilon) is the clipping parameter and  $\beta$  is the smoothing parameter.

- 8 Critic network is updated based on mean square error (MSE) value:

$$\text{Return} = A_t + V_t \text{ (from memory } K) \quad (15)$$

$$L_{Q\phi} = \text{MSE}[\text{Return} - V_t \text{ (from critic network)}] \quad (16)$$

- 9 Total loss = Actor loss + Critic loss

- 10 Update weights of the actor and critic networks ( $\Theta$  and  $\phi$  respectively) via stochastic gradient descent (SGD) method.

- 11 Perform iteration till the  $\Theta_{\text{old}}$  converges to  $\Theta$ .

### 3. Results and Discussions

Proximal policy optimization (PPO) based control have been applied to the bioreactor and their simulation results has been shown in this section. The efficacy of the algorithm can be shown by the root mean square error (RMSE), reward, and tracking plots. The feed flowrate mainly affects the desired protein concentration (mAB). Therefore, protein concentration is our control variable, and flowrate (ml/min) is our manipulated variable which is analogous to action. Python 3.7 version with Pytorch package has been used to train the deep neural network (DNN) for the application of the PPO algorithm. The actor and critic networks are DNN with 3 hidden layers and 256 hidden nodes each. A ReLU activation function has been used as an activation function for both actor and critic networks. Here the reference mAB concentration ( $\text{mAB}_{\text{ref}}$ ) i.e., end-point concentration is 590 mg/l. Our main objective is to reduce the error ( $e = \text{mAB} - \text{mAB}_{\text{ref}}$ ) with the help of the reward function as given below:

$$\text{Reward} = -|e| \quad (15)$$

The tracking performance of PPO can be shown in Figure 1(a) and the corresponding control policy (flowrate) can be seen in Figure 1(b). The PPO-based controller was trained for 200 episodes in presence of noise of 20 % in the input concentration of viable density. The controller performance can be evaluated with the help of average rewards (average of last 10 episodic rewards) as shown in Figure 2(a) and root mean square error (RMSE) values plot as shown in Figure 2(b). The results indicate that the proposed PPO based RL agent can effectively track the desired operating point.

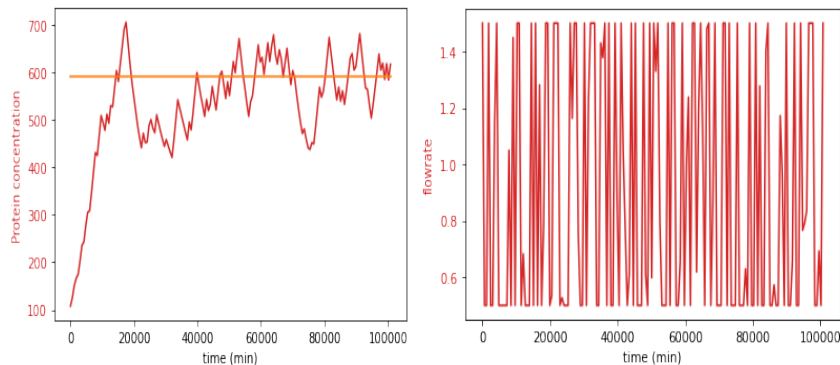


Figure 1 (a) Controller tracking performance (b) Optimal control strategy (flowrate)

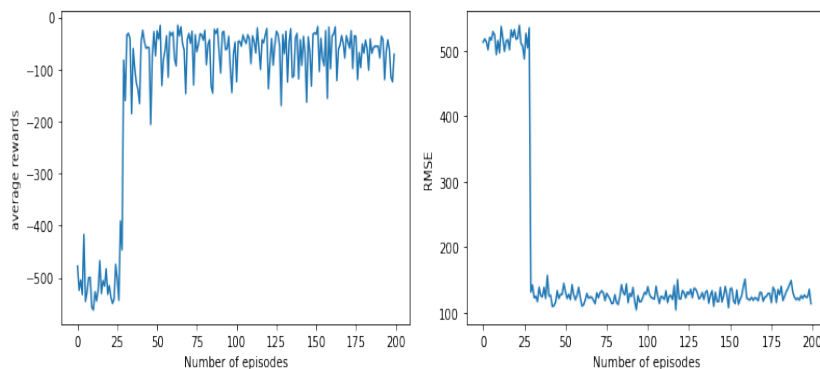


Figure 2 (a) Average reward plot (b) RMSE plot

#### 4. Conclusions

Proximal policy optimization (PPO), an actor-critic RL-based algorithm has been implemented for the bioreactor system (environment) to control the protein concentration (mAB) to obtain the optimal flowrate (action) to achieve the desired results. PPO can track the desired mAB concentration which is 590 mg/l. The average rewards and RMSE plot show that the controller has performed well. Therefore, PPO can be a good candidate for control the complex dynamics of the bio reactor process.

#### 5. Acknowledgement

The authors gratefully acknowledge Indian Institute of Technology, Delhi and DST-SERB India with file number CRG/2022/003722, SERB, INDIA for providing us resources to carry out experiments and simulations.

#### 6. References

- [1] Kern, Richard, and Yogendra Shastri. "Advanced control with parameter estimation of batch transesterification reactor." *Journal of Process Control* 33 (2015): 127-139. <https://doi.org/10.1016/j.jprocont.2015.06.006>
- [2] Lee, Jay H., and Kwang S. Lee. "Iterative learning control applied to batch processes: An overview." *IFAC Proceedings Volumes* 39, no. 2 (2006): 1037-1046. <https://doi.org/10.1016/j.conengprac.2006.11.013>
- [3] Jiang, Qingchao, Xuefeng Yan, Hui Yi, and Furong Gao. "Data-driven batch-end quality modeling and monitoring based on optimized sparse partial least squares." *IEEE Transactions on Industrial Electronics* 67, no. 5 (2019): 4098-4107. [10.1109/TIE.2019.2922941](https://doi.org/10.1109/TIE.2019.2922941)
- [4] Yoo, Ha Eun, Bo Eun Kim, Jong Woo Kim, and Jay H. Lee. "Reinforcement learning based optimal control of batch processes using Monte-Carlo deep deterministic policy gradient with phase segmentation." *Computers & Chemical Engineering* 144 (2021): 107133. <https://doi.org/10.1016/j.compchemeng.2020.107133>

- [5] Bao, Yaoyao, Yuanming Zhu, and Feng Qian. "A deep reinforcement learning approach to improve the learning performance in process control." *Industrial & Engineering Chemistry Research* 60, no. 15 (2021): 5504-5515.  
<https://doi.org/10.1021/acs.iecr.0c05678>
- [6] Sutton, Richard S., and Andrew G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018.
- [7] Schulman, John, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. "Proximal policy optimization algorithms." *arXiv preprint arXiv:1707.06347* (2017).  
<https://doi.org/10.48550/arXiv.1707.06347>
- [8] Wang, Yuhui, Hao He, and Xiaoyang Tan. "Truly proximal policy optimization." In *Uncertainty in Artificial Intelligence*, pp. 113-122. PMLR, 2020.
- [10] Kumar, Deepak, Neelesh Gangwar, Anurag S. Rathore, and Manojkumar Ramteke. "Multi-objective optimization of monoclonal antibody production in bioreactor." *Chemical Engineering and Processing-Process Intensification* (2021): 108720.  
<https://doi.org/10.1016/j.cep.2021.108720>
- [11] Byun, Ha-Eun, Boeun Kim, and Jay H. Lee. "Robust dual control of batch processes with parametric uncertainty using proximal policy optimization." In *2020 59th IEEE Conference on Decision and Control (CDC)*, pp. 3016-3021. IEEE, 2020.  
10.1109/CDC42340.2020.9304019
- [12] Hämmäläinen, Perttu, Amin Babadi, Xiaoxiao Ma, and Jaakko Lehtinen. "PPO-CMA: Proximal policy optimization with covariance matrix adaptation." In *2020 IEEE 30th International Workshop on Machine Learning for Signal Processing (MLSP)*, pp. 1-6. IEEE, 2020.  
10.1109/MLSP49062.2020.9231618
- [13] Gupta, Nikita, Riju De, Hariprasad Kodamana, and Sharad Bhartiya. "Batch-to-Batch Adaptive Iterative Learning Control— Explicit Model Predictive Control Two-Tier Framework for the Control of Batch Transesterification Process." *ACS omega* 7, no. 45 (2022): 41001-41012.  
<https://doi.org/10.1021/acsomega.2c04255>
- [14] Joshi, Tanuja, Shikhar Makker, Hariprasad Kodamana, and Harikumar Kandath. "Twin actor twin delayed deep deterministic policy gradient (TATD3) learning for batch process control." *Computers & Chemical Engineering* 155 (2021): 107527.  
<https://doi.org/10.1016/j.compchemeng.2021.107527>
- [15] Singh, Vikas, and Hariprasad Kodamana. "Reinforcement learning based control of batch polymerisation processes." *IFAC-PapersOnLine* 53, no. 1 (2020): 667-672.  
<https://doi.org/10.1016/j.ifacol.2020.06.111>
- [16] Nikita, Saxena, Anamika Tiwari, Deepak Sonawat, Hariprasad Kodamana, and Anurag S. Rathore. "Reinforcement learning based optimization of process chromatography for continuous processing of biopharmaceuticals." *Chemical Engineering Science* 230 (2021): 116171.  
<https://doi.org/10.1016/j.ces.2020.116171>

# Process design and economic evaluation of biomass-based negative emission technologies

Wei Wu<sup>a\*</sup>, Rasa Supankanok<sup>a,b</sup>, Walairat Chandra-Ambhorn<sup>b\*\*</sup>, Nattapat Pongboriboon<sup>a,b</sup>

<sup>a</sup>*Department of Chemical Engineering, National Cheng Kung University, Tainan 70101, Taiwan, ROC*

<sup>b</sup>*Department of Chemical Engineering, School of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520, Thailand*  
*Corresponding weiwu@gs.ncku.edu.tw*

## Abstract

A palm oil-based polygeneration system (POPS) is simulated to produce the main product of high-purity green diesel as well as the liquefied petroleum gas (LPG) as a by-product. The CO<sub>2</sub>-negative design includes approaches of (i) a series of cryogenic separators for the recovery of approximately 65.5% of hydrogen feedstock, (ii) the evacuated tube solar collector (ETSC) for reducing 35% flue gas from the furnace, (iii) the amine-based CO<sub>2</sub> capture process for pursuing the high-purity CO<sub>2</sub> product. The economic analysis of POPS shows that the process becomes economically attractive if the diesel price and crude palm oil should be around 1.98 and 0.47 \$kg<sup>-1</sup>, respectively.

**Keywords:** Palm oil; Carbon capture; Polygeneration; Negative CO<sub>2</sub> emission; Economic analysis

## 1. Introduction

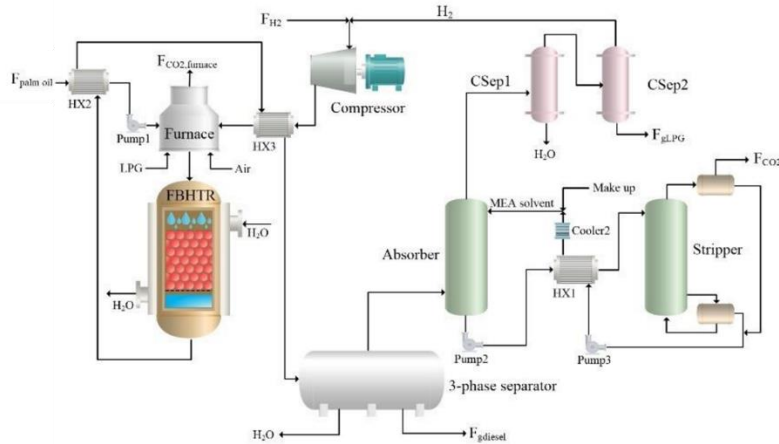
Oil palm is the major source of vegetable oil in the world, mainly produced in Southeast Asia. Palm oil not only contains triglycerides and many free fatty acids, but also it is a proper feedstock for green diesel production in Thailand due to Thailand being currently the third-largest palm oil producer in the world. A palm fatty acid distillate (PFAD) is a by-product generated from refining crude palm oil, so it can be treated as the feedstock of the green diesel production process. A study showed that the green diesel produced from PFAD achieved 84% GHG emissions reduction as compared to the GHG emissions of petroleum diesel (Alhajji and Demirel, 2016). In the process simulation of the production of green diesel from the hydrotreatment of vegetable oil, a model of the hydrotreatment of soybean oil for producing green diesel was validated according to experimental data, and the green diesel produced from Brazilian soybean oil could meet quality standards (Cavalcanti et al., 2022). The negative emissions technologies will become more important in the next decades due to pricing the carbon emissions. The bio-energy with carbon capture and storage technology was one of the primary negative carbon emission methods. A study showed that the process design of a biomass-based polygeneration system for power generation and ammonia synthesis could achieve negative CO<sub>2</sub> emissions (Xu et al., 2021).



## 2. Study Design

### 2.1 Palm oil-based polygeneration system (POPS)

A palm oil-based polygeneration system (POPS) named Design 1 is depicted in **Fig. 1** which mainly consists of a furnace, an FBHTR, a three-phase separator, and a series of cryogenic separators (C Sep1, C Sep2) to produce green diesel and green LPG (propane), notably the CO<sub>2</sub>-rich gases are emitted from the furnace and C Sep1, respectively. Two feedstocks of palm oil and hydrogen are heated up to 324°C by a furnace and a combination of a heater and a compressor, respectively, and they are fed into the fixed bed hydrotreating reactor (FBHTR) at the prescribed operating pressure. Notably, 65.5% of the feedstock of hydrogen is recovered through a series of cryogenic separators in Design 1, and the FBHTR is covered with a cooling jacket where the recycled water is produced from the three-phase separator. The outlet stream of FBHTR is cooled down and fed into a specific three-phase separator. The vapor-phase stream is fed into a sequence of cryogenic separators (C Sep1 and C Sep2) at the prescribed operating temperature and pressure to separate green LPG, CO<sub>2</sub>, and hydrogen, where the hydrogen flow is fully recycled and mixed with fresh hydrogen flow. The simulation results show that the purity of green diesel is up to 98.3wt% and the yield of green diesel is around 84.4%, but the purities of CO<sub>2</sub> and green LPG from C Sep1 and C Sep2 are close to 80.6% and 84.6%, respectively. Besides water and hydrogen being fully recycled, green LPG is used to meet the LPG fuel demand of the furnace.



**Figure 1** Palm oil-based polygeneration system

### 2.2 Performance evaluation

To address the performances of the POPS design, the indicators like net energy efficiency and net CO<sub>2</sub> emissions are addressed as follows.

(i) Net energy efficiency

$$\eta_{net} = \frac{E_{net}}{F_{palm\ oil} \times HHV_{palm\ oil} + F_{H_2} \times HHV_{H_2}} \quad (1)$$

and  $E_{net}$  is the net energy shown by

$$E_{net} = F_{gdiesel} \times x_{gdiesel} \times LHV_{gdiesel} + (F_{LPG} \times x_{LPG} \times LHV_{LPG} - W_{tot}) \quad (2)$$

where  $F_{palm\ oil}$  and  $F_{H_2}$  are the inlet mass flowrates of palm oil and hydrogen, respectively.  $F_{gdiesel}$  and  $F_{LPG}$  are the outlet mass flowrates of green diesel and green

LPG, respectively.  $x_{gdiesel}$  and  $x_{LPG}$  represent mass fractions of green diesel and green LPG in the outlet streams of  $F_{gdiesel}$  and  $F_{LPG}$ , respectively.

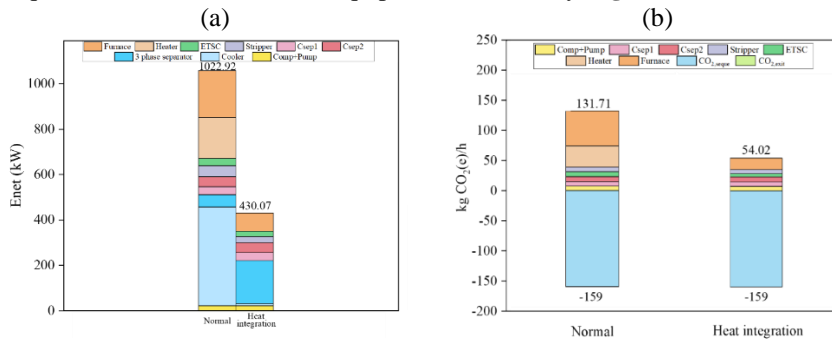
(ii) Net CO<sub>2</sub> emissions

The net CO<sub>2</sub> (equivalent) emissions ( $CO_{2,net}$ ) of each design including the total CO<sub>2</sub> emissions ( $CO_{2,tot}$ ) of all utilities are used in each design and net CO<sub>2</sub> sequestration ( $CO_{2,seque}$ ) of producing palm oil from oil palm plantations. It is noted that CO<sub>2</sub> (equivalent) emissions of equipment are evaluated by using the carbon tracking tool in Aspen Plus®. The detailed calculations have been shown by the previous work (Wu et al, 2023). **Table 1** shows that  $\eta_{net}$  is up to 68.14%, and the corresponding  $CO_{2,net}$  is negative.

**Table 1** Energy efficiency and CO<sub>2</sub> emissions

$E_{net}$	kW	8696.47
$\eta_{net}$	%	68.14
$CO_{2,net}$	kg/h	-113.39

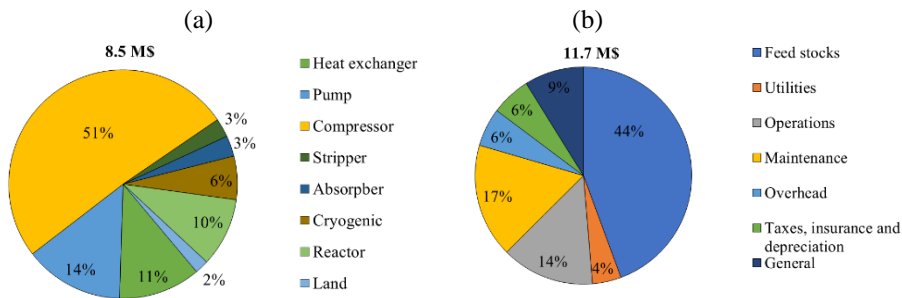
**Fig. 2** shows the combination of the duty of each unit in the POPS, the corresponding CO<sub>2</sub> (equivalent) emissions of each equipment is shown by **Fig. 3**.



**Figure 2** POPS in terms of (a)  $E_{tot}$  and (b) CO<sub>2</sub> (e) emissions

**2.3 Economic evaluation**

The economic evaluation is based on a combination of capital expenditure, operating expenditure, and carbon revenue. Through the Aspen Process Economic Analyzer (APEA), the capital expenditure (CAPEX) and operating expenditures (OPEX) of POPS are 8.5 M\$ and 11.7M\$/year, respectively.



**Figure 3** Pie chart of POPS in terms of (a) CAPEX and (b)OPEX

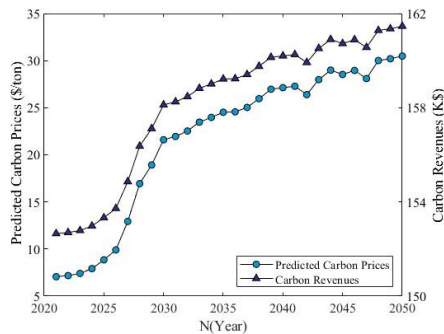
**Fig. 3(a)** shows that the equipment cost of the compressor dominates the total investment cost due to high hydrogen partial pressure, and **Fig. 3(b)** shows that the palm oil price dominates the total operating cost since it takes up to 44%.

### Carbon Revenue

Referring to the emission trading scheme (ETS) for estimating CO<sub>2</sub>price, the carbon revenues ( $\Delta P_i$ ) is based on the following calculation (Wu et al., 2023)

$$\Delta P_i = D_i \times P_{t,i} + B_i \quad (3)$$

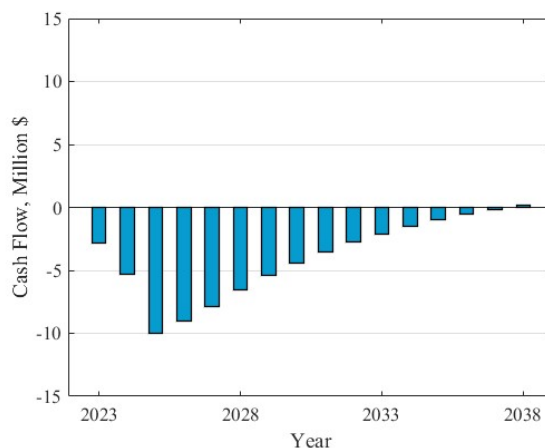
where  $D_i$  represents the amount of CO<sub>2</sub> reduction through the CO<sub>2</sub> capture and utilization (CCU) approaches.  $B_i$  represents the government incentives for the CCU, which here equals to 15000 USD. Moreover, the profile of annual carbon revenues of the POPS are shown in **Fig. 4** according to the carbon price from 2021 to 2050.



**Figure 4** Carbon revenue of the POPS

### Net present value

To evaluate the cumulative discounted cash flow of this project, it is noted (i) an interest rate of 15%, (ii) depreciation in five years, (iii) the plant construction for the first three years, and (iv) the working capital included in the third year. The cumulative discounted cash flow is shown in **Fig. 5**, notably, the net present value (NPV) remains negative until 2037. The payback period is more than 12 years from the start of plant operation.



**Figure 5** Cumulative discounted cash flow

Assuming an internal rate of return (IRR) of 15% and. The actual diesel and crude palm oil (CPO) prices are approximately 1.13 and 0.88  $\text{\$/kg}$ . In September 2019, the purchase price of CPO was set at 16.2 THB per kilogram (0.47  $\text{\$/kg}$ ). In order to make this process economically attractive, the diesel price and CPO should become 1.98 and 0.47  $\text{\$/kg}$ , respectively.

### 3. Conclusions

The process design of the green diesel produced from the hydrotreating reaction of palm oil is addressed by using Aspen Plus<sup>®</sup>. The outlet compositions of green diesel, water, propane, carbon dioxide, and hydrogen from the FBHTR are effectively separated and purified through a three-phase separator and a series of cryogenic separators. The amine-based CO<sub>2</sub> capture process integrated with two cryogenic separators in the post-separation framework not only increases two high-purity (over 99.9%) products of carbon dioxide and propane but also it ensures the maximum negative net CO<sub>2</sub> emissions. The economic scenarios show that the prices of crude palm oil and carbon revenue are key factors to promote the commercial potential of the POPS.

### References

- Alhajji, M., Y.Demirel, Energy intensity and environmental impact metrics of the back-end separation of ethylene plant by thermodynamic analysis, *Int J Energy Environ Eng.* 7 (2016) 45–59. <https://doi.org/10.1007/s40095-015-0194-9>.
- Cavalcanti, CJS, Ravagnani, MASS, Stragevitch, L, Carvalho, FR, PimentelMF. Simulation of the soybean oil hydrotreating process for green diesel production. *Clean Chem Eng* 2022;1:100004.
- Xu, C, Y.Liu, Q.Zhang, T.Xin, R.Zhao, M.Wang, W.Saw, S.Lu, Thermodynamic analysis of a novel biomass polygeneration system for ammonia synthesis and power generation using Allam power cycle, *Energy Convers Manag.* 2021;247:114746.

Wu, W, Supankanok, R, Chandra-Ambhorn, W, Taipabu, MI, Novel CO<sub>2</sub>-negative design of palm oil-based polygeneration systems. *Renewable Energy* 2023;203:622-633.

Wu, W, Xu, H, Shi, B, Kuo, PC, Techno-economic analysis of plastic wastes-based polygeneration processes. *Chemical Engineering & Processing: Process Intensification* 2023;184:109297.

# Design and intensification of the production process of propylene glycol as a high value-added glycerol derivative

Jahaziel Alberto Sánchez-Gómez <sup>a</sup>, Fernando Israel Gómez-Castro <sup>a</sup>, Salvador Hernández <sup>a</sup>

<sup>a</sup> *Departamento de Ingeniería Química, División de Ciencias Naturales y Exactas, Campus Guanajuato, Universidad de Guanajuato, Noria Alta S/N, Guanajuato, Gto. 36050, Mexico. [fgomez@ugto.mx](mailto:fgomez@ugto.mx).*

## Abstract

Glycerol is a by-product obtained in the biodiesel production. Such by-product can be used as raw material to obtain high added value derivatives. However, the reduction in energy requirements and capital costs for the conversion routes is necessary to enhance the economic indicators. This can be achieved through the intensification of the conventional glycerol processing routes. This work proposes the design and intensification of processes to produce propylene glycol (PG) from crude glycerol. Four different PG production schemes are analyzed. The conversion processes are intensified by means of thermal coupling and reactive distillation systems. Each system is designed by rigorous simulation in Aspen Plus<sup>TM</sup>. As results, it has been determined that implementing a thermally coupled system in the PG purification section, a clear decrease in the required reboiler duty (37%) is obtained.

**Keywords:** process intensification, crude glycerol valorization, process simulation, propylene glycol production.

## 1. Introduction

In the biodiesel production, the transesterification reaction generates large amounts of glycerol; for instance, in 2020, 46 billion liters of biodiesel were produced worldwide, with an estimated generation of more than 4.6 billion liters of crude glycerol (REN 21, 2021). The increasing glycerol excess has motivated the development of chemical processes which employ it as the main raw material. Among the different processes for the valorization of glycerol, oxidation, hydrogenolysis, dehydration, esterification, polymerization, carboxylation and pyrolysis can be mentioned (Pagliaro and Rossi, 2010; Monteiro et al., 2018). In the oxidation reaction, the main chemical derivatives to be obtained from glycerol are glyceric acid, dihydroxyacetone, lactic acid, mesoxalic acid and oxalic acid. Some intermediate products are also obtained, such as glyceraldehyde and glycolic acid (Monteiro et al., 2018). On the other hand, the reduction of glycerol generates mainly 1,2-propylene glycol, 1,3-propylene glycol, ethylene glycol and other by-products such as propanol, methanol and carbon dioxide (Muraza, 2019).

Propylene glycol (PG) is an important chemical product that can be used as surfactant, humectant, antifreeze, solvent, and preservative. Conventionally, PG is obtained by hydrolysis of propylene oxide, which is generated from petrochemicals. Alternatively, PG can be produced by a more environmentally friendly route by using glycerol as a

carbon and hydrogen source, requiring fewer fossil sources. Although cost competitiveness remains a major constraint, the conversion of glycerol to PG is a promising route. Evidence of this is the recent announcement by ORLEN Południe about the opening by the end of 2021 of a commercial green PG plant with a production capacity of 30,000 tons per year (ORLEN Południe, n.d.).

A strategy to enhance the economic indicators on the PG production process with glycerol as raw material implies applying process intensification strategies. Process intensification (PI) is a branch of chemical engineering by means of which it is possible to obtain more efficient processes. The main objective of PI is to achieve a technological improvement in a process, so that it occupies less space, it is more efficient with energy management and is cleaner and safer (Kiss, 2019). To achieve this objective, the aims are to reduce the size of the equipment and/or the number of basic operations of the process using multifunctional equipment. In this way, PI emerges as a strategy for generating highly efficient processes.

The present work has as main objective, the simulation, intensification and assessment of the hydrodeoxygenation process to produce high purity PG (99% mol) from glycerol using thermally coupled distillation columns and reactive distillation. The performance of the conventional and intensified processes is evaluated in terms of the energy consumption and the size of the equipment, measured through the required number of stages.

## 2. Methodology

### 2.1. Thermodynamic and kinetic modelling

Considering the polar nature of glycerol, the presence of hydrogen in the reaction systems, and the presence of organic compounds in a large amount of water, the NRTL model turns out to be the most suitable approach to describe the liquid-vapor equilibrium. Also, this model has been used in the literature to describe similar mixtures (Sun et al., 2022). The binary interaction parameters for the NRTL model were obtained from the Aspen Plus database.

The production of propylene glycol ( $C_3H_8O_2$ ) from glycerol ( $C_3H_8O_3$ ) has as by-products ethylene glycol ( $C_2H_6O_2$ ), methanol ( $CH_3OH$ ) and water ( $H_2O$ ). The reactions involved in the generation of PG are shown in Eqs. 1 and 2.



The hydrodeoxygenation of glycerol to produce PG occurs in the liquid phase, using Cu/ZrO<sub>2</sub> as catalyst. For this catalytic reaction, the kinetic parameters for a first-order power-law kinetic model obtained by Gabrysch et al. (2019) have been used; such kinetic model is given by the following equations:

$$r_1 = k_1 c_{H_2} \quad (3)$$

$$r_2 = k_2 c_{H_2} \quad (4)$$

Variation of the kinetic constant ( $k_i$ ) with temperature is given by the Arrhenius equation:

$$k_I = k_{0,I} e^{-E_{A,I}/RT} \quad (3)$$

where  $k_{0,I}$  is the pre-exponential factor, and  $E_{A,I}$  is the activation energy. The parameters required for the calculation of reaction rates are summarized in Table 1.

**Table 1.-** Pre-exponential factor and activation energy for hydrodeoxygenation reactions.

<b>I</b>	<b><math>k_{0,I}</math></b>	<b><math>E_{A,I}</math> (kJ/mol)</b>
1	$2.3 \times 10^{10} \text{ L mol}^{-1} \text{ s}^{-1}$	106
2	$9 \times 10^7 \text{ L mol}^{-1} \text{ s}^{-1}$	97

## 2.2. Process design and simulation

The analyzed processing schemes for PG production from glycerol are the conventional scheme (CS), a thermally coupled scheme (TCS), a reactive distillation scheme (RDS) and a thermally coupled reactive distillation scheme (TCRDS). The studied configurations are shown in Fig 1. The conventional process scheme (CS) is considered as the base case for comparison with the proposed intensified processes. The schemes have been rigorously simulated in Aspen Plus™, and the PG production process alternatives were jointly evaluated using the energy consumption (associated with operation costs) and the number of stages (associated with capital costs) as decision criteria.

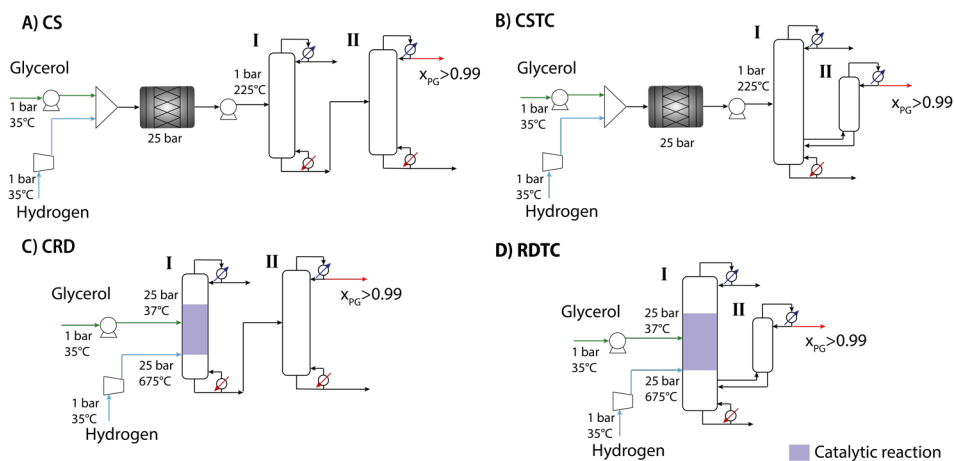
Most of the components have a very low boiling point compared to the product of interest and the rate reactions are fast, then it may be appropriate to use heuristic strategies for the design of the intensified process schemes (Keller, 2014). The feed stream has a flow rate of 100 kmol/h of glycerol and water (1:1 wt.). The distillation columns operate with a pressure drop of 10 psia (0.68 atm) and the reactions are defined in Aspen Plus™ as REAC-DIS type reactions, with residence times for the gas phase of 0.8 min for the hydrodeoxygenation reaction. The number of stages is determined through a sensitivity analysis, aiming to obtain the lowest number of stages necessary to obtain PG with a minimum purity of 99% in mol.

In the design of the thermally coupled alternatives, a sensitivity analysis in the interconnection streams was carried out. This procedure consists in the manipulation of the interconnection mole flowrate (vapor or liquid) to reduce the energy requirements of such schemes while the purity of PG is maintained at 99 % mol. Additionally, in the simulation of the proposed processing schemes, it has been necessary to fix the molar flowrate at the bottoms, since, based on a previous study, it was observed that these operating variables have a greater influence on the purity of PG in the output stream. In addition, by manipulating the reflux ratio it is possible to control the energy load in the reboiler.

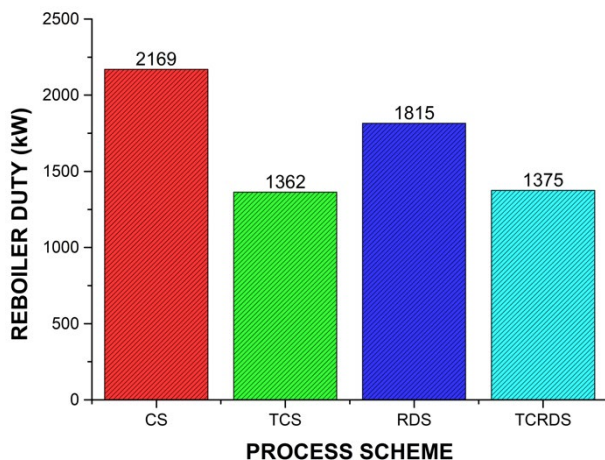


### 3. Results and discussion

Figure 2 shows the required reboiler duty for the conventional and intensified processes. It is worth to mention that all the designs are still non optimal. However, all the designs satisfy the requirements in purity and recovery of PG. The results indicate that the conventional scheme (CS) have energy consumption of around 2169 kW. Additionally, the direct thermal coupling of CS to TCS represents a more significant decrease in reboiler duty than the intensification of the process using the reactive distillation system RDS. In the case of the TCS, this process shows the lowest energy consumption (1362 kW). However, the reactive distillation thermally coupled (TCRDS) has a very similar energy consumption to the TCS. The difference is only 13 kW.



**Fig 1.-** Schemes for the process of propylene glycol production from glycerol, a) conventional scheme (CS), b) thermally coupled scheme (TCS), c) reactive distillation scheme (RDS), d) thermally coupled reactive distillation scheme (TCRDS).



**Fig 2.-** Reboiler duty required in the processes for the production and purification of PG.

To complement the analysis, Table 2 presents all the design specifications, including the number of stages, the reboiler duty and the obtained savings. It has been found that for a similar value of energy requirements (1375 kW), the TCRDS requires a lower number of total stages (15) than the TCS scheme. Moreover, the TCS scheme requires a separate reactor, which would increase the equipment cost. In this sense, the TCRDS scheme would be presented as the most promissory processing scheme for PG production and purification, since it would represent the scheme associated with the lowest operating and capital costs.

**Table 2.-** Design specification for the conventional and intensified processes schemes.

<b>Scheme</b>	<b>CS</b>	<b>TCS</b>	<b>RDS</b>	<b>TCRDS</b>
Number of Stages				
Column I	20	40	20	35
Column II	40	20	35	10
<b>Total Stages</b>	<b>60</b>	<b>60</b>	<b>55</b>	<b>45</b>
Feed Stage				
Column I	5	5,20	11,19	9,21,22
Column II	20	20	5	10
Reflux Ratio				
Column I	0.5	0.21	0.1	0.1
Column II	2.6	0.35	2.6	0.15
Distillate Flow, kmol/h				
Column I	99.56	99.75	100	99.73
Column II	15	15.15	15	15.27
Bottoms Flow, kmol/h				
Column I	16.44	1.1	16	1
Column II	1.44	4.85	1	2.13
Reboiler Duty, kW	2169	1362	1815	1375
<b>Savings, %</b>	<b>-</b>	<b>37.21</b>	<b>16.32</b>	<b>36.61</b>

#### 4. Conclusions

This work presents different alternatives to produce and purify PG using glycerol as raw material. Based on the topology of the conventional process, different intensified alternatives are proposed in the reaction and purification zones. The intensified processes have been contrasted with the conventional scheme. The results indicate that a maximum reduction of 37% in energy requirement can be achieved in comparison with the conventional scheme. The reduction in energy requirements lead to a more sustainable process and minor production costs of PG from glycerol. Additionally, the TCRDS shows important reductions in the number and size of processing equipment.

#### References

A.A. Kiss, 2019, Novel catalytic reactive distillation processes for a sustainable chemical industry, *Top. Catal.*, 62, 1132–1148.

M.R. Monteiro, C.L. Kugelmeier, R.S. Pinheiro, M.O. Batalha, 2018, Glycerol from biodiesel production: Technological paths for sustainability, *Renew. Sustain. Energy Rev.*, 88, 109–122.

M. Pagliaro, M. Rossi, 2010, *The Future of Glycerol*, RSC Publishing, 2nd Edition, Cambridge.

O. Muraza, 2019. Peculiarities of glycerol conversion to chemicals over zeolite-based catalysts, *Front. Chem.*, 7, 1–11.

ORLEN Południe, n.d., in <https://www.orlenpoludnie.pl/en/media/news/Pages/ORLEN-Poludnie-finalises-its-investment-in-green-glycol.aspx>.

P Sun, W Zhang, X Yu, J Zhang, N Xu, Z Zhang, M Liu, D Zhang, G Zhan, Z Liu, C Yang, W Yan and X Jin, 2022, Hydrogenolysis of Glycerol to Propylene Glycol: Energy, Tech Economic, and Environmental Studies; *Front. Chem.*, 9:778579.

REN 21, 2021, *Renewables 2021 Global Status Report - REN21* in [https://www.ren21.net/wp-content/uploads/2019/05/GSR2021\\_Full\\_Report.pdf](https://www.ren21.net/wp-content/uploads/2019/05/GSR2021_Full_Report.pdf).

T. Gabrysch, M. Muhler, B. Peng, 2019, The kinetics of glycerol hydrodeoxygenation to 1,2-propanediol over Cu/ZrO<sub>2</sub> in the aqueous phase, *Appl. Catal. A Gen.*, 576, 47–53.

T. Keller, 2014. *Reactive Distillation*, in A. Górak, Z. Olujić (Eds.), *Distillation: Equipment and Processes*, Academic Press, Boston, 261–294.

# Optimal Design of Heat Integrated Hybrid Dividing Wall Columns by Vapor Recompression

Dian Ning Chia, Fanyi Duanmu, Eva Sorensen\*

*Department of Chemical Engineering, University College London (UCL),  
Torrington Place, London WC1E 7JE, United Kingdom*

*\*e.sorensen@ucl.ac.uk*

## Abstract

A hybrid dividing wall column (HDWC) is a highly integrated design that combines a dividing wall column with a hybrid distillation-membrane process into a single process. It has previously been shown that HDWC can reduce the total annualized cost compared to its conventional DWC counterpart for the separation of an azeotropic mixture. The energy efficiency of HDWC can be further enhanced by vapor recompression, which is studied in this work. Since the HDWC has two reboilers, three different variations of heat integration are investigated, with the heat exchanger in the compression cycle replacing the reboiler in either the prefractionator, the main column, and for both columns, respectively. For the case study considered, it was found that the total annualized cost for the heat integrated HDWC can be up to 35% lower than the HDWC design, however, although heat integration always reduces operating costs, it may not always reduce total annualized costs.

**Keywords:** Process Intensification, Distillation, Hybrid Dividing Wall Column, Vapor Recompression, Optimization

## 1. Introduction

Process intensification (PI), the concept of integrating two or more processing units or methods within a single unit to enhance operation, is an important approach when designing more energy-intensified processes. Two of the most prominent examples of PI separations are the dividing wall column (DWC), where a ternary separation takes place within a single column shell, and the hybrid distillation-membrane process, where a membrane stage is incorporated into the design to help separate a close-boiling or azeotropic mixture. The DWC achieves energy savings through the reduction of mixing/remixing effects due to the central wall, while the incorporation of the membrane in a hybrid process provides savings by separating the mixture via a mechanism that is not based on VLE (e.g., solution-diffusion). Recently, Chia et al. (2022) introduced a hybrid dividing wall column (HDWC) that integrates both of these PI alternatives. In their work, the HDWC was compared with various standard hybrid distillation-pervaporation processes, and it was shown that the HDWC could separate an azeotropic mixture with a lower total annualized cost (TAC). The HDWC could be further improved by applying suitable heat integration methods to reduce energy consumption and TAC even further. It was also shown that heating and cooling requirements in the column part were about four times higher than that of the total energy required in the membrane part of the process. Therefore, further enhancement of energy efficiency is mainly dependent on the distillation column. There is no guarantee that heat integration methods will benefit the design (i.e., reduce TAC), as this depends on the process and the mixture, however,

Duanmu and Sorensen (2022) showed that vapor recompression can have great success in reducing both energy consumption and TAC for DWCs, thus vapor recompression will be considered and studied in this work. The HDWC process is here used to separate a binary minimum-boiling azeotropic mixture rather than a ternary mixture as for “regular” DWC, and the DWC has the central wall extending all the way to the bottom. The DWC is considered in a Petlyuk arrangement, i.e. a prefractionator and a main column.

## 2. Methodology

To apply vapor recompression heat integration, the condenser of the column is removed and the vapor distillate is pressurized by a compressor. A superheater before the compressor is required to maintain the vapor phase even after increasing the stream pressure. The pressurized high-temperature vapor is then utilized to vaporize the boil-up flow by replacing the reboiler with a heat exchanger. The standard hybrid dividing wall column (HDWC) is shown in Figure 1a. The design incorporates one condenser and two reboilers, which means there are a few variations that can be considered for vapor recompression-assisted HDWC. The high-temperature vapor can be used to vaporize the boil-up flow from the prefractionator (VR-HDWC-Pre; Figure 1b); from the main column (VR-HDWC-Main; Figure 1c); or from both columns (VR-HDWC-Both; Figure 1d). In the following, these three variations will be developed, optimized, and compared with the base case without heat integration (HDWC). All four structures shown in Figure 1 contain two separation units, a DWC (wall extended from the middle to the bottom) and a membrane network. For the DWC, the feed, with composition assumed to be below the azeotropic point, enters the prefractionator. A stream with composition close to the azeotropic point will leave from the top of the prefractionator (i.e., top of the wall) and enters the main column as the thermal coupling stream, while the remaining component (almost pure heavy component) will leave the system from the bottom of the prefractionator as product. The azeotropic stream will leave the DWC and pass to the membrane network. Unlike in the standard hybrid distillation-pervaporation process, where the purpose of the membrane is to purify the azeotropic mixture to achieve the product specification, here the purpose of the membrane is to assist in crossing the azeotropic point. Next, the retentate (with composition shifted to the other side of the azeotropic point compared to the feed composition) will be further separated in the main column section and almost pure light component will leave from the bottom of the main column as product.

The membrane network model proposed by Marriott and Sorensen (2003) is applied, where the membrane network contains  $n$  membrane stages connected in series, and in each membrane stage there are  $i$  membrane modules connected in parallel. It should be noted that the number of membrane modules in parallel in each membrane stage can be different and is determined through optimization. The procedure for constructing this membrane network and the overall hybrid distillation-pervaporation process can be found in Chia and Sorensen (2022). Although not shown in this work, before each membrane stage, a heater may be installed to improve the membrane separation performance and the requirement of each membrane stage heater can be optimized.

For the vapor recompression assisted HDWCs shown in Figures 1b to 1d, the condenser is replaced with a series of superheater, condenser, and trim cooler. The superheater and condenser create a pressurized vapor with high temperature, which allows heat exchange between the vapor and the liquid bottom flow from the distillation column. Three VR-

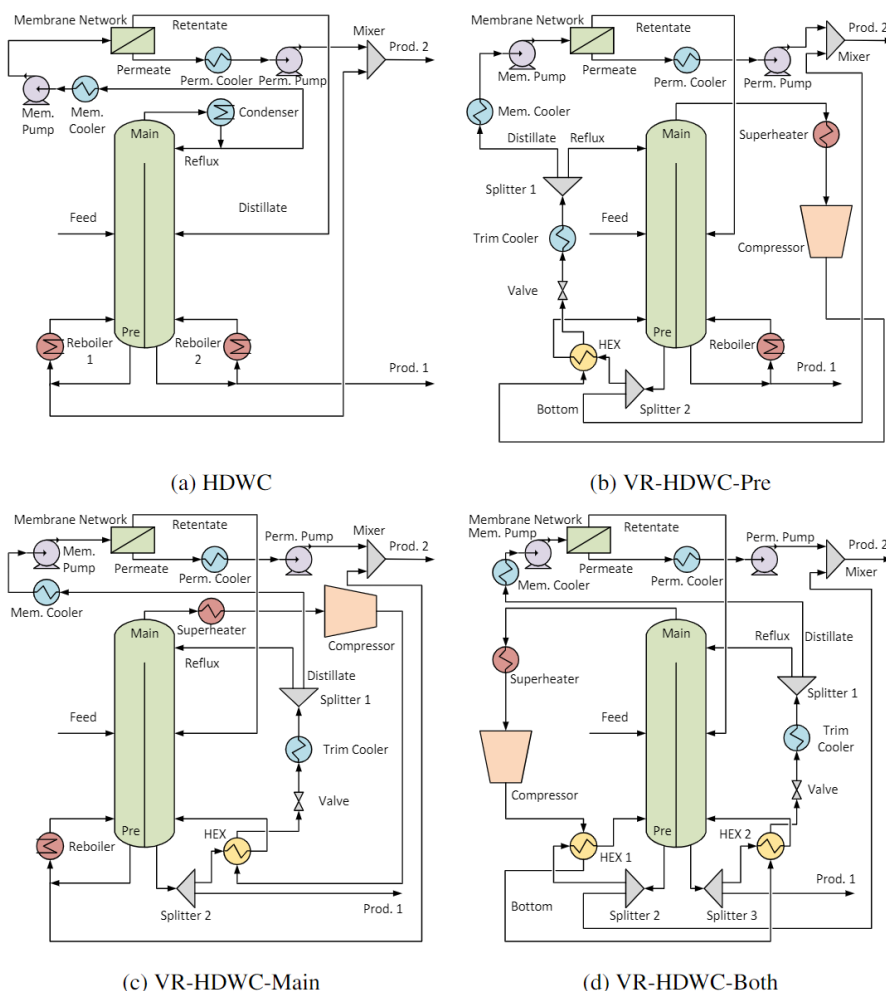


Figure 1. Flowsheets of (a) the base hybrid dividing wall column (HDWC) and vapor recompression assisted HDWCs with the heat exchanger replacing the reboiler in (b) prefractionator, (c) main column, and (d) both columns.

HDWC structures are named based on the location of the heat exchanger, i.e., VR-HDWC-Pre is the system with the reboiler in the prefractionator replaced with a heat exchanger, and VD-HDWC-Main where the heat exchanger is replacing the reboiler on the main column side. In VR-HDWC-Both, the pressurized vapor heats up the liquid bottom stream from the prefractionator first (as it has a higher boiling point compared to that of the main column), and then heats up the liquid bottom stream from the main column. For all, the hot outlet from the last heat exchanger passes through a valve to reduce its pressure back to the column operating pressure. Then, the stream is cooled down to the saturated liquid as both reflux and the feed to the pervaporation membrane need to be in the liquid phase.

### 3. Case Studies

The case study considered is the separation of  $0.2/0.8 \text{ mol mol}^{-1}$  of ethyl acetate and

ethanol. The feed is assumed to be saturated liquid at 1 atm, and is supplied at  $200 \text{ kmol h}^{-1}$ . The product purity specifications are 0.99 mol%. UNIQUAC is used to describe the liquid phase thermodynamics, while ideal gas is assumed for the vapor phase. The minimum temperature approach is set as 10 K in the heat exchangers.

All models (rigorous equilibrium-based) are built in gPROMS Process (Process Systems Enterprise, 2021). As there are no available libraries for dividing wall columns, the corresponding Petlyuk structure (not shown) is constructed using the built-in distillation column library, while the pervaporation membrane is modeled using a user-defined model. The cost equations for the distillation column can be found in Duanmu et al. (2022b), while the membrane costs can be found in González and Ortiz (2002). The optimizer chosen is the Particle Swarm Optimization (PSO), and the details of the settings can be found in Duanmu et al. (2022a). The constraints for the optimization include the product purities, using the same number of stages on both sides of the wall, a minimum approach temperature in the heat exchanger, and no phase change in the compressor (i.e., the vapor fraction in the stream after the compressor equal to one).

The key results for each optimal design are shown in Table 1. For the column design, all designs have a similar total number of stages for both the prefractionator (25-28) and the main column (27-29). The feed locations in HDWC are slightly lower compared with the VR-HDWC designs. There are no clear reasons for this, as there are no clear mixing effects observed from the liquid composition profiles in each design (not shown), indicating that the feed locations are reasonable. The changes in the feed locations may be dominated by other factors, which will be discussed later. Moving to the specifications in the main column, it is clear that the side liquid flow rate in the main column (i.e., the liquid thermal coupling stream at the top of the wall) for both VR-HDWC-Pre and VR-HDWC-Both is higher compared with the other two designs, and this is a result of the reboiler in the prefractionator being replaced with a heat exchanger. It can be seen that the trim cooler duty in both VR-HDWC-Pre and VR-HDWC-Main ( $-1164 \text{ kW}$  and  $-971 \text{ kW}$ , respectively) are large, indicating that the energy capacity of the pressurized vapor is much higher than the energy required in the heat exchanger. Therefore, increasing the side liquid flowrate will increase the flow rate of the cold liquid in the heat exchanger, which will, in turn, maximize the energy transfer between the pressurized vapor and cold liquid, and at the same time develops a better composition profile in the prefractionator (i.e., the top vapor compositions from the prefractionator are closer to the azeotropic point than the other two designs).

Next, the smaller number of membrane modules required in VR-HDWC-Main and VR-HDWC-Both can easily be explained. In order to minimize the energy wasted (i.e., to minimize the trim cooler duty), VR-HDWC-Main and VR-HDWC-Both tend to maximize the energy exchange in the heat exchanger by prioritizing the separation of the pure component from the azeotrope in the main section of the distillation column rather than in the membrane, thus the number of membrane modules required are fewer. In other words, the sloppier membrane separation will cause the heat duty required in the heat exchanger to increase, thus utilizing more of the energy from the pressurized vapor (i.e., decreasing the trim cooler duty) and, at the same time, reducing the membrane demand, which is a win-win situation for VR-HDWC-Main and VR-HDWC-Both. This is reflected in the higher boil-up flow rate and lower pure retentate composition in VR-HDWC-Main and VR-HDWC-Both when compared to the other two designs (not shown). Considering the energy consumption, it is clear that VR-HDWC-Both has the

Table 1. Optimal results for the base hybrid dividing wall column (HDWC) and the vapor recompression assisted HDWC with the heat exchanger replacing the reboiler in the prefractionator (VR-HDWC-Pre), main column (VR-HDWC-Main) and both columns (VR-HDWC-Both).

Items	HDWC	VR-HDWC			Units
		Pre	Main	Both	
<b>Prefractionator</b>					
Total stages	26	26	28	25	-
Feed stage	11	14	14	15	-
Bottom product *	118	128	125	128	$kmol\ h^{-1}$
Boil-up flow	201	231	214	237	$kmol\ h^{-1}$
<b>Main Column</b>					
Total stages	28	29	29	27	-
Feed stage (from pre.)	2	2	2	2	-
Feed stage (from mem.)	12	17	18	18	-
Side liq. stage	2	2	2	2	-
Side liq. flow rate	142	186	164	194	$kmol\ h^{-1}$
Reflux *	230	268	286	322	$kmol\ h^{-1}$
Bottom product *	39	39	39	39	$kmol\ h^{-1}$
Boil-up flow	63	54	98	105	$kmol\ h^{-1}$
<b>Vapor Recompression System</b>					
Superheater temp.	-	345.25	345.20	345.16	<i>K</i>
Compressor pressure	-	1.74	1.68	1.74	<i>bar</i>
<b>Total Duty/Power (Excluding Membrane Network)</b>					
Reboiler (pre)	2113	-	2252	-	<i>kW</i>
Reboiler (main)	567	488	-	-	<i>kW</i>
Condenser / Trim cooler	-3456	-1164	-3209	-971	<i>kW</i>
Compressor	-	221	234	271	<i>kW</i>
Heat exchangers (total)	-	2429	875	3442	<i>kW</i>
<b>Membrane Network</b>					
No. membrane stages †	4	4	5	4	-
Total no. modules	87	77	60	58	-
Total membrane area	522	462	360	348	$m^2$
Total heater duty	359	292	281	232	<i>kW</i>
Permeate cooling duty	-602	-463	-505	-462	<i>kW</i>
<b>Capital Cost</b>					
Column	1.21	1.35	1.38	1.31	<i>M</i> \$
Reboiler / Superheater	0.66	0.34	0.58	0.13	<i>M</i> \$
Condenser / Trim cooler	0.52	0.24	0.49	0.22	<i>M</i> \$
Heat exchanger	-	0.36	0.20	0.57	<i>M</i> \$
Compressor	-	3.56	3.62	3.78	<i>M</i> \$
Membrane	2.86	2.53	1.97	1.91	<i>M</i> \$
Others	0.81	0.79	0.92	0.79	<i>M</i> \$
<b>Operating Cost</b>					
Heating	1.29	0.33	1.08	0.10	$M\ \$y^{-1}$
Cooling	0.28	0.20	0.23	0.19	$M\ \$y^{-1}$
Electricity	0.00	0.11	0.12	0.14	$M\ \$y^{-1}$
Mem. replacement	0.08	0.07	0.06	0.05	$M\ \$y^{-1}$
<b>Overall</b>					
Capital cost	6.06	9.17	9.16	8.70	<i>M</i> \$
Annualized capital cost	0.76	1.15	1.14	1.09	$M\ \$y^{-1}$
Operating cost	1.65	0.71	1.49	0.48	$M\ \$y^{-1}$
Total annualized cost	2.40	1.86	2.63	1.57	$M\ \$y^{-1}$

\* For VR-HDWCs, the bottom products are controlled by Splitters 2 and 3, and the reflux is controlled by Splitter 1.

† Upper bound for the number of membrane stages is 8.



highest duty in the heat exchanger and lowest duty in the trim cooler, indicating a higher energy efficiency of the heat integration structure.

Comparing the economic performances of each design, the capital cost (CAPEX) is much lower for HDWC (about 34 % reduction compared with VR-HDWC-Pre), and the main reason is the high cost of installing the compressor in VR-HDWCs (contributing to about 40% of the CAPEX). It should be noted that the membrane cost in VR-HDWC-Main and VR-HDWC-Both on the other hand drops significantly (up to 33%) compared with HDWC due to the smaller membrane network required, as explained earlier. Considering the operating cost (OPEX), it is not surprising that all three VR-HDWCs have lower OPEX due to the energy saving in the reboiler heating cost, especially for VR-HDWC-Both. In terms of the TAC, the objective function in the optimisation, VR-HDWC-Both is the cheapest design with about 35% TAC reduction compared to HDWC. VR-HDWC-Main, however, is about 10 % more expensive than HDWC. This indicates that the VR-HDWC designs may not always yield a lower TAC, as the VR-HDWC designs are closely related to the trade-off between the capital cost of the compressor and the operating cost. The TAC is also strongly influenced by individual costs (not shown).

#### 4. Conclusions

In conclusion, this work proposed three vapor recompression assisted hybrid dividing wall column designs (VR-HDWC) to further improve energy efficiency compared to just a hybrid dividing wall column (HDWC). Compared with the base HDWC design, all VR-HDWC designs save operating costs, although not always total annualized costs (TAC) as the results of the vapor recompression assisted design relies heavily on the trade-off between the more expensive compressor capital cost vs the cheaper electricity cost.

#### References

- D. N. Chia, F. Duanmu, E. Sorensen, 2022. Optimal Design of Hybrid Dividing Wall Columns for Azeotropic Separations. In: The 12th International Conference Distillation & Absorption 2022.
- D. N. Chia, E. Sorensen, 2022. Optimal Design of Hybrid Distillation/ Pervaporation Processes. In: Y. Ya-mashita, M. Kano (Eds.), Proceedings of the 14th International Symposium on Process Systems Engineering - PSE2021+. Elsevier B.V., pp. 313–318.
- F. Duanmu, D. N. Chia, E. Sorensen, 2022a. A Combined Particle Swarm Optimization and Outer Approximation Optimization Strategy for the Optimal Design of Distillation Systems. In: Y. Yamashita, M. Kano (Eds.), Proceedings of the 14th International Symposium on Process Systems Engineering - PSE2021+. Elsevier B.V., pp. 1315–1320.
- F. Duanmu, D. N. Chia, E. Sorensen, 2022b. A shortcut design method for complex distillation structures. *Chemical Engineering Research and Design* 180, 346–368.
- F. Duanmu, E. Sorensen, 2022. Optimal Design of Heat Integrated Reduced Vapor Transfer Dividing Wall Columns. In: Y. Yamashita, M. Kano (Eds.), Proceedings of the 14th International Symposium on Process Systems Engineering - PSE2021+. Elsevier B.V., pp. 175–180.
- B. González, I. Ortiz, 2002. Modelling and simulation of a hybrid process (pervaporation-distillation) for the separation of azeotropic mixtures of alcohol-ether. *Journal of Chemical Technology & Biotechnology* 77 (1), 29–42.
- J. Marriott, E. Sorensen, 2003. A general approach to modelling membrane modules. *Chemical Engineering Science* 58 (22), 4975–4990.
- Process Systems Enterprise, 2021. gPROMS Process version 2.2.

## Efficient use of energy in distillation sequences.

José A. Caballero,<sup>a</sup> Juan A. Labarta,<sup>a</sup> Laureano Jiménez,<sup>b</sup>

<sup>a</sup>*Institute of Chemical Process Engineering, University of Alicante, Carretera de S. Vicente s.n. 03690, Alicante, Spain.*

<sup>b</sup>*Department of Chemical Engineering, Universitat Rovira I Virgili, Av. Països Catalans 26, 43007, Tarragona, Spain.*

### Abstract

In this work, we show that the systematic and simultaneous implementation of some heat integration alternatives like direct reboiler condenser heat exchange, implementation of vapor (re)compression cycles, leverage of heat in high-pressure steam utilities at lower pressure/temperatures, multiple effect distillation, etc. can eventually produce considerably reductions in energy and total annualized cost. While none of these alternatives is new, they are rarely simultaneously implemented in a distillation sequence, however, we show that there is great potential in the synergic effects of implementing simultaneously some of them. We illustrate the procedure with the separation of a mixture of acetone and chloroform using pressure swing distillation. Results show that we can get impressive reductions in energy and total annualized costs.

**Keywords:** add three to five keywords here.

### 1. Introduction

Even though in the chemical industry there are a good number of alternatives to separate components, distillation accounts for around 90-95% of all the separations and purification (Humphrey and Keller, 1997), and it seems that this situation will not change shortly. Notwithstanding, it is not uncommon the claim that distillation must be substituted by other, more efficient separation technologies (National Academies of Sciences, 2019). The main reason is based on the fact that distillation vaporizes a mixture in the reboiler and then liquifies a vapor stream in the condenser. While this is true in a single distillation column, Agrawal and Tumbalam Gooty, (2020) have shown that energy required in distillation is not necessarily proportional to the reboiler heat duty and that distillation can be much more efficient than what intuition seems to suggest. Even more, they presented an industrial separation example -separation of propylene from propane using distillation and membranes- that contradicts this believe. They showed that, in this case, distillation needs only a fraction of the energy requirement for membrane separation.

Different alternatives have been proposed to increase the energy efficiency in distillation. If we are dealing with a single separation (i.e. a single distillation column) we can consider alternatives like multi-effect distillation, and intermediate heat exchangers either in the rectifying or in the stripping section, -the optimal location, which can be done using column grand composite curves-, (Bandyopadhyay et al., 1998), pre-fractionations, internally heat integrated distillation columns (HIDiC) and heat pump assisted distillation (Vapor recompression, Mechanical or Thermal Vapor recompression -VRC- or absorption heat pumps) (Kiss et al., 2012) and combinations of some of the previous

alternatives. Kiss et al., (2012) also gave an algorithm for selecting a technology for heat integration in binary distillation.

In multicomponent distillation, the number of alternatives increases because we can integrate the alternatives for heat integration between different columns with those of a single column and the alternatives for column sequencing inherent in multi-component distillation. Therefore, if the number of possible sequences is large, the first stage must be selecting a sequence of distillation columns with inherent good energy performance and then systematically checking for the different alternatives to increase energy efficiency.

In the case of zeotropic distillation, the structural characteristics of the search space were defined by different authors around 10-15 years ago (Caballero and Grossmann, 2006; Giridhar and Agrawal, 2010) Those authors, established the basis, and the structural considerations to take into account for generating valid sequences of columns that range from sequences based on conventional distillation columns (a column with one condenser, one reboiler, and two products) to fully thermally coupled distillation sequences (a single reboiler and a single condenser for all the separation sequence). From these sequences, it is possible to systematically generate intensified alternatives like Divided Wall Columns (DWC), and Kaibel Columns – which separate 4 components in a single column-, among others. A recent review of the advances in intensification in distillation can be found in the work by Jiang and Agrawal, (2019).

In the case of separations involving azeotropes, the number of alternatives is considerably lower than that of the zeotropic counterpart, and there are other considerations to take into account. For example, in extractive distillation, the selection of the extractant is likely the decision with the highest influence on the energy consumption of the system.

In multi-component distillation, the direct heat integration between the condenser of a column and the reboiler of another, modifying the pressure(s) in the column(s) if necessary is the most obvious alternative for saving energy, and different optimization strategies have been proposed since the 70s of the last century.

However, while there are a good number of alternatives for improving the energy efficiency in distillation it is uncommon to implement two or more alternatives simultaneously. In this work, we show that the simultaneous combination of different alternatives can produce important savings in energy (and in total annual cost). Besides we present a framework to combine all those alternatives, and of course, the new ones that could eventually appear.

## **2. An algorithm for efficient energy integration in distillation.**

The implementation of an optimization model (i.e. through a superstructure) to simultaneously design/optimize the distillation sequence and each column in it together with all the considered alternatives for heat integration, is likely to produce very large and complex models, that could be, in the best case very hard to solve. Instead, we propose an approach that sequentially tries to improve efficiency in the selected configuration:

1. Select the best sequence of distillation columns. If we are considering a zeotropic mixture we can use any of the approaches based on mathematical programming to get the optimal configuration. For example, (Nallasivam et al., 2016) used an algorithm that ensures the global optimal solution. In those cases, it is also common that there are some alternatives with similar performance, so a priori we should select a sub-set of alternatives. In azeotropic (extractive, reactive...) distillation, the starting point is an optimized feasible separation sequence.

2. In the first place, we must consider the possibility of heat integration with the rest of the process. In some cases, there is a waste heat stream, that can be used in distillation and no further action is needed. (We do not consider such a case in this paper)
3. For each column in the sequence, we determine the lower and upper pressure limits. This can be based on thermodynamics, product stability, security or operational issues, etc. The pressure determines which are the highest and lowest temperatures in each column in the condenser and reboiler.
  - 3.a. The change of temperature (pressure) allows the development of strategies for direct heat integration between condensers and reboilers of different columns by manipulating the pressure of the columns if necessary. In the literature, there are different models to do this step rigorously (Caballero and Grossmann, 2006; Floudas and Paules, 1988)
  - 3.b. The limits in pressure and temperature allow for determining the possibility of multi-effect distillation which has a special interest in difficult separations.
4. When we stack two or more columns -this is the case of divided wall columns and some thermally coupled sections- it is usually necessary to increase the internal flows in some column sections with the corresponding extra energy consumption. In those cases, an intermediate heat exchanger can help to recover the system's optimal operating conditions, and in some cases, it is possible to implement a VRC that reduces, even more, the column duty (Navarro-Amorós et al., 2013)
5. If the sequence requires heat at very different temperatures, in some situations it is possible to recover energy from the exhausted high-temperature utility to use them at lower temperatures. For example, from high-pressure steam (say at 250 °C, 40 bar) that produces 100 kW when condensates in a reboiler, it is possible to recover 18.9 kW at 180 °C by expanding the exhaust water up to 10 bar and 34 kW if we expand to 1.2 bar (~120 °C).
6. The use of intermediate heat exchangers (reboilers in the stripping section and condensers in the rectifying section) can, in some cases, reduce the condenser/reboiler heat duties, which can be substituted with cheaper utilities. Besides, these new heat exchangers can eventually be used in the heat integration strategy. Column Grand Composite curves (CGCC) can be used to determine the optimal location of those heat exchangers (Bandyopadhyay et al., 1998)
7. Heat pump-assisted configurations (Kiss et al., 2012) (vapor compression, mechanical or thermal vapor recompression, absorption heat pumps, bottom flashing, etc.) can be implemented not only between the condenser and reboiler of a given column but between different heat exchangers (source and sinks of heat) whose temperature difference is not too large (typically no more than 30°C)

The systematic application of all those alternatives can yield a rich space of alternatives with an important potential for energy reduction.

Finally, considering the intensification of the sequence of columns could allow extra savings in investment. This last point is out of the scope of this work, but a review of different alternatives can be found in Jiang and Agrawal, (2019).

### 3. Discussion on Results

To illustrate some of the characteristics of the proposed algorithm we consider the separation of a mixture of Acetone and Chloroform using pressure swing distillation (PSD). The objective is to obtain each product with a purity of at least 0.995 mol fraction. Table 1 shows the data for the example.

Table 1. Data for the example.

Components	Composition (mol fraction)	Feed Flow	Pressure
Acetone	0.75	80 kmol/h (5872 kg/h)	101.3 kPa
Chloroform	0.25		
Cold Utilities	Cost (\$/kW·y)	Hot Utilities	Cost (\$/kW·y)
water (20-15 °C)	11.4	LP Steam (~2 bar 120 °C)	277.5
		HP Steam (~10 bar 180 °C)	292.18
Thermodynamics NRTL (default Aspen-HYSYS parameters)			
Electricity 0.067 \$/kWh			
interest = 10% in 10 years			
Cost estimation based on correlations by (Turton et al., 2013)			

The base case distillation sequence is shown in Figure 1. The simulation of the PSD separation was done in Aspen-Hysys V.11, including the sizing of the columns. The base case has a total module cost of k\$1,710. The total cost of utilities is 2,901.7 k\$/year with a cost of manufacturing (COM) of 3,900.3 k\$/year, which yields a TAC (total annualized cost) of 4,198.7 k\$/y. The feeds to all the columns are forced to be saturated liquids, so it is necessary to include some heat exchangers in all the streams entering the distillation columns.

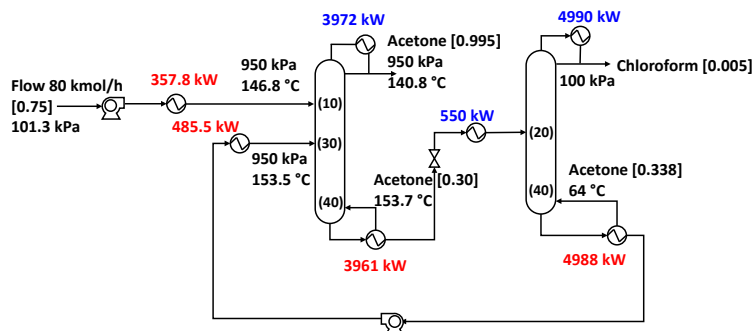


Figure 1. Separation of Acetone and Chloroform by PSD. Base case for comparison. Numbers in the square brackets refer to acetone mole fractions, and numbers in brackets refer to tray number.

Even though there is some degree of freedom in the selection of the pressures of both columns, in this case, the pressures are mainly determined by the variation of the azeotrope composition, so column pressures will be fixed in all the alternatives.

It is possible to leverage part of the energy of the exhausted MPS (water around 10 bar and 180 °C) by expanding it to 1.4 bar (~110°C). If we separate the vapor fraction it is possible to recover 593.8 kW by condensing this new low-pressure steam than can be used in the low-pressure column. Figure 2a shows the resulting process. The utility cost is reduced by 5.8 %.

Figure 2a. suggests that it is possible to increase energy efficiency and the manufacturing cost by the heat integration of the different hot and cold streams (heaters, coolers, condensers, and reboilers.) The most obvious is the integration of the condenser of the high-pressure column with the reboiler of the low-pressure column -Figure 2b-. A pinch analysis shows that it is possible to reduce the cost of utilities to 1,433 k\$/year with a cost

of manufacturing of 2,094.7 k\$/year and a TAC of 2,392.8 k\$/y. This is a reduction in the cost of utilities of 50.1% and a reduction in TAC of around 43%.

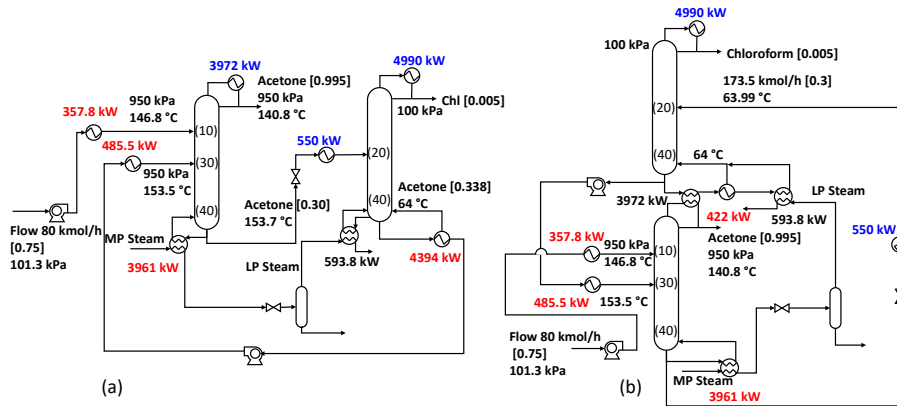


Figure 2. (a) PSD acetone and Chloroform in which the exhausted MP Steam is expanded to 1.4 bar to recover a fraction of energy in form of LP-Steam that is used in the reboiler of the LP-column. (b) include the integration between the condenser of the HP column and the reboiler of the LP column.

The temperature difference between the reboiler and condenser in the HP column is lower than 15 °C in both columns. In these conditions, a VRC usually allows for reducing energy costs, even though the capital cost can increase considerably. In this case, it is possible to implement a VRC in each column and at the same time maintain the possibility of heat integration between the condenser or HP column and the reboiler of the LP column (Figure 3).

Capital cost increases to 4,040 k\$, around 120%, but energy cost decreases also considerably. A pinch analysis shows that total energy cost is reduced to 533.0 k\$/y, which yields a COM of 1,382.8 k\$/y and a TAC of 2,037 k\$/y. Savings in energy costs and TAC are around 81% and 51.5 % respectively compared to the base case and around 63% and 14.9 % compared to the best heat-integrated sequence without VRCs.

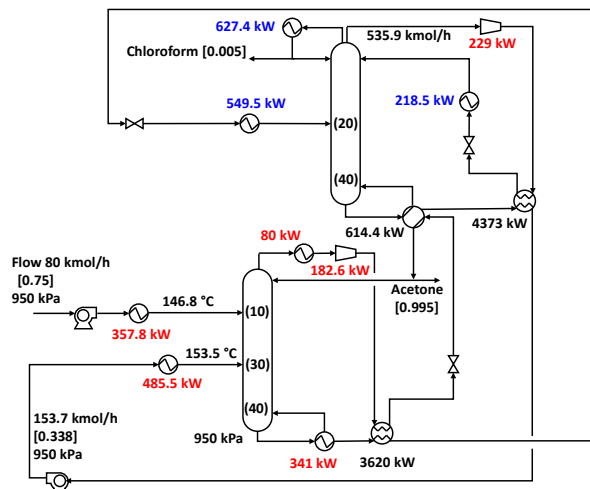


Figure 3. Integration of a Vapor Recompression Cycle in each column. Note that heat integration between the condenser in the HP column and the reboiler in the LP column is still possible.

#### 4. Conclusions

In this paper, we have shown that it is possible to increase the efficiency of distillation columns by sequentially and iteratively adding different alternatives of heat integration. While no one of those alternatives is new, the simultaneous consideration of all/some of them can produce an impressive increase in energy efficiency that is reflected in a considerable reduction in total costs.

We have illustrated the procedure with the separation of a mixture of acetone and chloroform using pressure swing distillation. The systematic consideration of different available alternatives allows for reducing the energy cost by an impressive 81% when compared with a non-heat integrated based case or 51.5% compared with a base case using only direct heat integration. In terms of TAC, reductions are around 63% and 14.9% respectively.

#### 5. Acknowledgments

The authors acknowledge financial support to the Spanish «Ministerio de Ciencia e Innovación» under project PID2021-124139NB-C21

#### References

- Agrawal, R., Tumbalam Gooty, R., 2020. Misconceptions about efficiency and maturity of distillation. *AIChE J.* 66, 2–4. <https://doi.org/10.1002/aic.16294>
- Bandyopadhyay, S., Malik, R.K., Shenoy, U. V., 1998. Temperature-enthalpy curve for energy targeting of distillation columns. *Comput. Chem. Eng.* 22, 1733–1744. [https://doi.org/10.1016/S0098-1354\(98\)00250-6](https://doi.org/10.1016/S0098-1354(98)00250-6)
- Caballero, J.A., Grossmann, I.E., 2006. Structural Considerations and Modeling in the Synthesis of Heat-Integrated–Thermally Coupled Distillation Sequences. *Ind. Eng. Chem. Res.* 45, 8454–8474. <https://doi.org/10.1021/ie060030w>
- Floudas, C.A., Paules, G.E., 1988. A mixed-integer nonlinear programming formulation for the synthesis of heat-integrated distillation sequences. *Comput. Chem. Eng.* 12, 531–546. [https://doi.org/10.1016/0098-1354\(88\)87003-0](https://doi.org/10.1016/0098-1354(88)87003-0)
- Giridhar, A., Agrawal, R., 2010. Synthesis of distillation configurations: I. Characteristics of a good search space. *Comput. Chem. Eng.* 34, 73.
- Humphrey, J., Keller, G., 1997. *Separation Process Technology*. McGraw-Hill Education.
- Jiang, Z., Agrawal, R., 2019. Process intensification in multicomponent distillation: A review of recent advancements. *Chem. Eng. Res. Des.* 147, 122–145. <https://doi.org/10.1016/J.CHERD.2019.04.023>
- Kiss, A.A., Flores Landaeta, S.J., Infante Ferreira, C.A., 2012. Towards energy efficient distillation technologies – Making the right choice. *Energy* 47, 531–542. <https://doi.org/http://dx.doi.org/10.1016/j.energy.2012.09.038>
- Nallasivam, U., Shah, V.H., Shenvi, A.A., Huff, J., Tawarmalani, M., Agrawal, R., 2016. Global optimization of multicomponent distillation configurations: 2. Enumeration based global minimization algorithm. *AIChE J.* 62, 2071–2086. <https://doi.org/10.1002/aic.15204>
- National Academies of Sciences, 2019. *A Research Agenda for Transforming Separation Science, A Research Agenda for Transforming Separation Science*. <https://doi.org/10.17226/25421>
- Navarro-Amorós, M.A., Ruiz-Femenia, R., Caballero, J.A., 2013. A new technique for recovering energy in thermally coupled distillation using vapor recompression cycles. *Aiche J.* 59, 3767–3781.
- Turton, R., Bailei, R.C., Whiting, W.B., Shaeiwitz, J. A., & Bhattacharyya, D., 2013. *Analysis, Synthesis and Design of Chemical Processes.*, 4th editio. ed. Pearson Education, Inc., Upper Saddle River, NJ. USA.

# The use of Design Thinking in the innovation of artificial connected reefs

Justine Michel<sup>a</sup>, Manon Arnaud<sup>a</sup>, Adrien Roudaut<sup>a</sup>, Albane Tenneson<sup>a</sup>, Lucie Mace<sup>a</sup>, April Schnaffner<sup>a</sup>, Mame Cheikh Sow<sup>a</sup>, Nicolas Chaton<sup>a</sup>, Nader Wehbe<sup>a</sup>, Paul-Emeric Roger<sup>b</sup>, Anne Asensio<sup>b</sup>, Arnaud Coutu<sup>a</sup>

<sup>a</sup>*Institut Polytechnique UniLaSalle, Rue Pierre Wagué, 60000 Beauvais, France*

<sup>b</sup>*Dassault Systèmes, 10 Rue Marcel Dassault, 78140 Vélizy-Villacoublay, France*  
[arnaud.coutu@unilasalle.fr](mailto:arnaud.coutu@unilasalle.fr)

## Abstract

The climate changes that have occurred in recent years are modifying the characteristics of the environment. Seas and oceans swallow, each year, a part of the CO<sub>2</sub> we emit, acidifying it in consequence (*Francour and al., 2006*). To respond to this environmental alert the Dassault Systèmes Foundation launched a call for projects to UniLaSalle students to help protecting the oceans and reduce the anthropic impact. Using the Design Thinking method, the students developed solutions that links the working methods of designers and engineers. This method is a perfectly adapted approach to answer this type of problem because designers work for human beings, while engineers work for the nature. It was use in order to well understand the perimeter of the project. Its exercises allow you to ask questions about the general appearance of a project while keeping in mind how humans and/or the environment can impact it or be impacted.

**Keywords:** Artificial Reef, Design Thinking, Fluid simulation, Innovation, Modelling.

## 1. Introduction

The Dassault Systèmes Foundation had invited UniLaSalle students to participate to a call for projects. Their goal was to bring together Design and Engineering in order to mix, within the framework of science, different ways of thinking. To do that, the Foundation asked students to create a scientific project with the aim of protecting the oceans and offered them a training in Design Thinking to organize their ideas in a new way.

Climate change, which currently affects 45% of threatened marine species, is not the only stressor threatening the marine environment. It is the fifth most important factor after overfishing, transport, urban development and pollution (Wait, 2022). Therefore, it is imperative to start to focus on the preservation of the ecosystem.

During this project, Design Thinking allowed the team to highlight the importance of public awareness and the progress of research. In order to deal with this problem, the student chose to develop a project based on the creation of artificial reefs connected to sensors in order to follow the progress of the environment in which they are located. The goal was to create a database that could be used to create animations in museums or to support researchers working on specific marine environments. Step by step, the main idea switch from the development of artificial coral reefs to intelligent 'reefed' mooring posts. This change expanded the study area to regions of high anthropogenic stress and promoted the presence of fauna and flora in ports (Luff et al., 2019). On this file, it will be explained the processes of design thinking, modulization and simulation that permitted this reconversion.

## 2. Methodology

### 2.1. The Design Thinking process

*« Most people mistakenly think that design is about appearance (...)  
That's not how we define design. (...) Design is how it works»  
- Steve Jobs in 2003*



This approach, based on people, was oriented towards innovation in order to create innovative services or products. The particularity of this approach is the modification of the working method (Usabilis, 2018) by asking the right questions at the beginning of the project in order to guide its development into the right direction.

During the creation phase of the project, three workshops had been set up by the Design Studio in order to guide UniLaSalle students through the learning and use of Design Thinking. This approach made it possible to co-create projects and to determine future problems and needs by following a precise thought process (Figure 1). In this context, the supervisor prepared exercises on a remote collaboration space called Miro. This tool consists of a whiteboard that allows you to work and share your ideas with images, post-its, mind maps, tables, text boxes, and other tools.

Design Thinking exercises allow to ask questions about the general aspect of a project while keeping in mind how humans and/or the environment can impact it or be impacted. These are lighthearted questions that specify each parameter of the project, thus allowing the team to grasp most of the difficulties it will encounter in the future (Figure 2). They follow a pattern that always involves: actors, priorities, problems and proposed solutions.

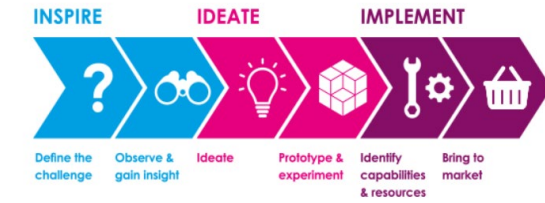


Figure 1: Three steps to simplify the process (Usabilis, 2018)

**The group work on the determination of the specific part of the opportunity with the helps of the analogy of a fortress:**

- 1 **The Drawbridge:** The digital innovation possibilities.
- 2 **The Walls:** The biggest legacy we'll have to handle.
- 3 **The Towers:** Actors who can help to leverage the project.
- 4 **The Market:** External data and information needed to trigger the project.
- 5 **The Dungeon:** The core mission of the opportunity.
- 6 **The Dragon:** What can make us fail? (Competitors, business context...)

Figure 2: Exercise questions of the fortress realized by the Design Studio

To complete these exercises, the engineering students answered the questions using post-its. This method allowed everyone to express themselves and to understand the opinions of others in order to discuss the problems of the project.

Following this step, each group made a pitch explaining their project. Three meetings were held in order to answer new questions:

- What is the problem? Context, understanding of human problems, environmental impact, opportunities and trade-offs?
- What solution is proposed? To do what? At what cost? What is needed to implement it? What will be tested to verify its relevance?

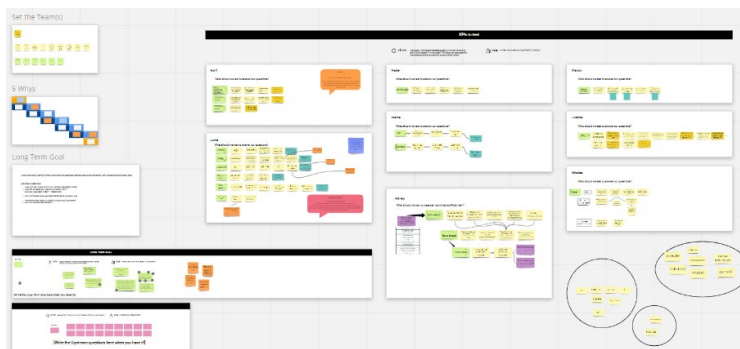


Figure 3: Example of Design Thinking work done on Miro

These pitches were used to distinguish which project would be about to be developed for an internship. The chosen subject had been the one that is presented on this paper: the "intelligent artificial reef". Therefore, the 4 members that imagined this solution constituted a team of students coming from different specialities (biology, environment, mechanics and geology) in order to develop it.

In order to create a group spirit and to communicate everyone's ideas, the new team had a formation of Design Thinking with the Design Studio supervisor. This period allowed students to bond during the important choices of the project by taking votes. Each student had post-its that they could place on whatever their choice was. The 3 choices with the most votes were then put aside to make a second round. The main objective was to make everyone agree with the development of the project by letting them discuss and debate (Figure 3).

This period also gave the students personal goals and a problem to solve. From this stage on, the students worked semi-autonomously as a small research office would. The Design Studio accompanied the group conducting weekly meetings on Friday mornings, for 30 minutes to 1 hour, to monitor the project and advise its members. These meetings began early and ended on the day of the Fondation Dassault Systèmes visit, at the end of the internship. The framing of these meetings allowed the team to refocus in this research in order to avoid misdirection and not waste time. Moreover, these project reviews allowed students to use Design Thinking by reflex, they no longer need to consult each other before a meeting because everyone follows the role of the others and knows where the project is at.

## 2.2. Digital designs and simulations

The Ocean ReGen team worked on the creation of an artificial reef and "reefed" mooring post, both composed of sensors to monitor the physico-chemical and biological criteria of the marine environment (sound, image, turbidity, temperature, pollution and population tracking). The sensors that make up the electrical system of both solutions could be modified according to the user's needs by slightly modifying the proposed system.

To carry out this project, the students used Dassault Systèmes solutions such as the 3DEXPERIENCE platform. This platform is a collaborative environment that allows the creation of innovative products using virtual experiences. It also provides a real-time view of project activity through its 3D Dashboards (Figure 4), which connect people, ideas and data.

This collaborative environment also allows the use of some Dassault Systèmes software, including Catia, Creative Experience and Drafting. The most important is the Catia software. It allowed students to model their reefs and moorings spots in 3D. These models were then used in the Drafting application to create constructions plans for each shape and then in Creative Experience to make virtual enhancement videos. In addition, these models were imported into XFlow, a fluid simulation software, also produced by Dassault Systèmes, but not included into the 3DEXPERIENCE platform. In order to optimize to the maximum their creation, numerous 3D models had been produced.

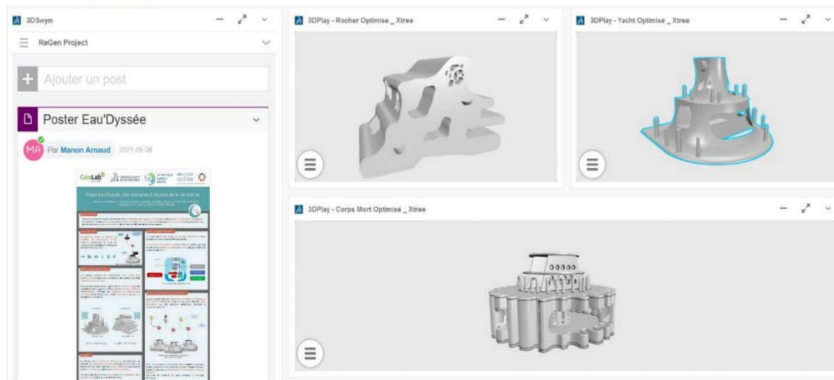


Figure 4: Example of a 3D Dashboard presenting the 3D models of the reefs made

XFlow permitted to choose the most appropriate shape for the team's needs by following some criteria... In order to attract as many species and entities as possible, reefs and moorings must have a complex shape (Charbonnel et al., 2001) and had to own shadow zones, turbidity, an upward current and a dead zone (where the current is weak or non-existent) within or behind the structure (Figure 5).

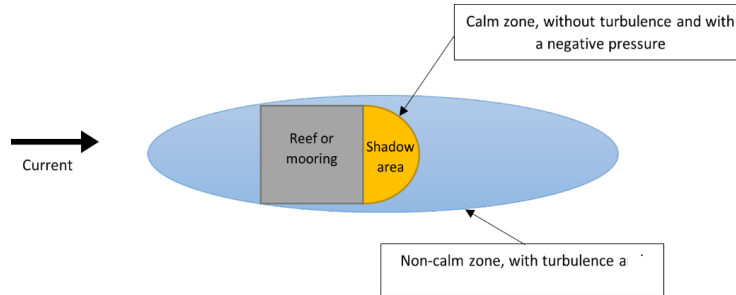


Figure 5 : Diagram of the current behavior in front of an immersed structure

Ocean ReGen students realized meetings with 6 scientists working in the domain of artificial reefs in order to have opinions on their achievements and to have new proposals. Those meetings helped the students to choose the ideal location, the shape and the materials of the reef.

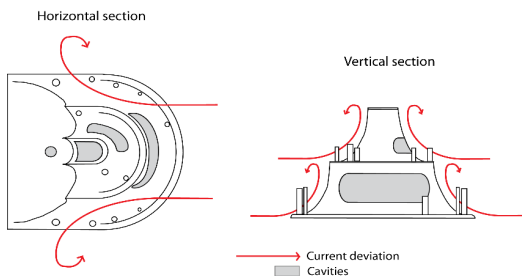


Figure 6 : Deflecting the current with an artificial reef (realised on Adobe Illustrator)

This particularly form (Fig. 6) had been developed in order to correspond to the criteria previously described in figure 5.

This form can further disperse the sand like you can see with the red arrows. This allows lateral movement of the sediment which helps slow down the waves at the back of the reef. However, this also creates an acceleration of the current on each side of the shape and creates erosion which can cause the structure to bury.

Indeed, most of the structures are made of concrete and the classic concrete weighs  $250 \text{ kg/m}^2$  for 10 cm thickness (Béton, 2019). It happens that some structures are buried by 60% in only 5 years. Unfortunately, it remains one of the biggest challenges that encounter companies or researcher that try to create new reefs and the students weren't an exception. Indeed, because of this burial rate, the structures must be removed after few years.

This burial is mainly due to their weight and shape.

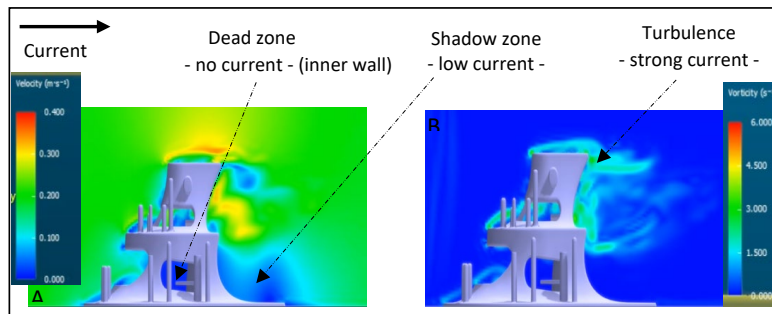


Figure 8 : Simulation of current speed (A) and turbidity (B) on XFlow® (Yacht)

Thanks to a fluid simulation software, it was possible to see that this shape respects the presence of a dead zone (without current), a shadow zone (weak current) and a zone with strong turbulences. These selection criteria were chosen because the turbulence makes noise and attracts crustaceans,

the dead zones are useful for species that like areas without current and the shadow zone at the back of the structure allows a mix between the two. The objective in gathering all these criteria is to attract as many different fish species as possible. These three 3D models are therefore the latest versions proposed by the students (Figure 7).

### 3. Analysis

#### 3.1. Project development through collaboration between Design & Engineering

The confrontation between Design & Engineering allowed the emergence of many ideas. The initial concept developed by the selected team proposed the implementation of an intelligent artificial coral reef. However, following the collective reflexion activities orchestrated by the Design Studio, the training of the different Dassault Systèmes tools and the exchanges with professionals such as Ifremer, ALR, the MIO... the project underwent many changes.

A final visit to the Fondation Dassault Systèmes permitted the Ocean ReGen project team to propose two different scenarios for the implementation of reefs. First of all, the aspect of artificial coral reefs had disappeared. Indeed, the project had split into two parts, one of which allows the implementation of an intelligent artificial reef and the other one of a reefed mooring post, also intelligent. An artificial reef would be a structure that would be voluntarily immersed in water to help the biomass to develop. A traditional mooring post is a massive concrete platform that is immersed. This structure, connected to a chain and a buoy, allows the anchoring of boats in order to avoid the ravages caused by anchors on the sea floor. The difference with the reefed mooring post would come from the shape of the structure. Unlike a massive block, the goal is to complexify its shape to create a mini reef while keeping the basic utility of a traditional mooring post. This idea came from exchanges with researchers who worked on the Port-Cros reefs immersed in 1985. Their feedback and advices guided the choices made by the group of students regarding the shape of their structures and their location.

On the other hand, the weekly design meetings also helped to guide the choice of sensors for the intelligent system while linking them to the human. To do this, the student had to answer why and to whom? In others terms: Why use these sensors and for whom would they be useful? The chosen ones are hydrophones, cameras, a tracking system, turbidity, temperature and pollution sensors. (Table 1)

Sensors		for whom?
Hydrophones	<ul style="list-style-type: none"> <li>Lack of knowledge about marine sounds.</li> <li>Identify day/night organisms.</li> <li>Locating disturbing anthropogenic sounds.</li> <li>Monitoring against illegal fishing ...</li> </ul>	<p><b>Museums:</b> diffusion of videos and aquatic sounds to show marine life to the general public</p> <p><b>Researchers:</b> ecological monitoring of the environment and scientific discoveries</p> <p><b>Authorities:</b> coastal surveillance for the protection of marine organisms by fighting against illegal fishing</p> <p><b>Swimming monitoring:</b> turbidity and pollution levels can be used to prohibit swimming and protect the public.</p>
Video cameras	<ul style="list-style-type: none"> <li>Ecological monitoring of species.</li> <li>Surveillance against illegal fishing.</li> <li>Information on the transparency of the water and the possibility of diving.</li> </ul>	
Turbidity sensors	<ul style="list-style-type: none"> <li>Detection of turbidity anomalies and thus detect certain pollutions.</li> <li>Information on the transparency of the water and the possibility of diving.</li> </ul>	
Temperature sensors	<ul style="list-style-type: none"> <li>Information on ocean warming.</li> <li>Understand the impact of temperatures on marine organisms.</li> </ul>	
Capteurs de pollution	<ul style="list-style-type: none"> <li>Monitoring of pollution levels in ports and bays.</li> <li>Monitoring the impact of pollution on marine populations</li> </ul>	
Tracing system	<ul style="list-style-type: none"> <li>Ecological monitoring of fish species (interaction and movements).</li> </ul>	

Table 1: Table of sensors use

In order to illustrate the process a little better... After a meeting with Atlantique Landes Récifs, the association expressed a need for cameras in order to obtain information on the environment around the area where their reefs are located. The usefulness would allow to know the level of transparency of the water thanks to the recordings video in order to be able to organize their diving missions. Public awareness was one of the main missions of the Ocean ReGen team. For this purpose, the data from the tracking systems, the sounds recorded by the hydrophones as well as the images from the cameras could be transmitted and used by museums or schools. The same is true for the data collected by the temperature and pollution sensors. Moreover, those sensors would also help the scientific progress and the development of nature. Human being knows little about marine sounds. It would have the possibility to increase the amount of data concerning the degradation of the biomass and the impact of the implantation of the structures that the ReGen Project team proposed.

The Design Studio's methods were reused many times by the team members. When it came to gathering everyone's ideas, Miro was the main tool used to create, deliberate and choose the best idea. This was helpful when it came to choosing the shape of the logo, the slogan phrase and the project name. The same was true for the choice of the reef shapes and mooring post. Dozens of 3D models were made and only 3 were chosen.

### 3.2. Students' feedback on the collaboration between Design & Engineering

This experience was filled with questioning, misdirection and success. The Design Thinking process allowed the group of students to see and revisit their ideas, project progress, and goals many times. Originally, the students aimed to incorporate salinity and oxygen sensors into the system in addition to temperature and turbidity... However, after a meeting with the deputy director of the MIO (Mediterranean Institute of Oceanology), instead he advised them to implement hydrophones. This technology is not very developed nowadays and marine sounds are still not well known

in the scientific world. The same is true for cameras. After discussions with ALR, the usefulness of incorporating them into the system would be beneficial to them. Consequently, after having worked for several weeks on the installation of the wrong sensor device, the mechanical pole had to radically reorient its research.

The team of 9 students managed to be quick in their decision making allowing the project to develop quickly. This was possible thanks to their motivation, the advice of external professionals and the Design Thinking methods taught by the Design Studio at the beginning of their internship. The team became more cohesive and was able to focus on the task at hand.

Design Thinking allows different domains to mix and move forward together. Indeed, this method allows to build a project link by link, by considering future problems. This way of thinking is the very source of research. Generally, the students particularly appreciated the use of Design Thinking and planned of using it for their future research projects.

#### 4. Conclusion and perspectives

Finally, the use of Design Thinking helped to direct and organize the project. The collaboration with the designers allowed the engineering students to learn a new approach to project management and the objectives of innovation. Thanks to it, the students were able to propose two scenarios that meet the needs of the environment and humans. Their work on the reefs already allows the 3D printing of the chosen shape thanks to the different construction plans made. As for the electrical system, it still needs improvements to be immersed and functional.

In order to finalize the project, UniLaSalle was considering organizing a new summer internship. This internship would focus on the development of the electronic part and would aim to realize a prototype of the reef and its intelligent system. The current team continues the exchanges with the various contacted organizations... The exchanges with ALR aimed to confirm their collaboration with the GeoLab. If this were to happen, ALR would help ReGen members with the prototyping of the electrical part and would implement it on its reefs in the Atlantic. The students were also keeping in touch with Xtree, in hopes of getting them to print the modeled reefs.

#### Acknowledgments

The authors gratefully thank the Design Studio, the educational research platform Le GéoLab and La Fondation Dassault Systèmes for the support provided for this work. The authors want to thank also Maxime Denis and Julien Duquennoy for their precious help.

#### References

- Béton Direct., 2017. Quel est le poids du béton ? Tout sur le béton Date de consultation : 04/02/2023. Disponible sur : <https://www.toutsurlebeton.fr/le-ba-ba-du-beton/quel-est-le-poids-du-beton/>.
- E. Charbonnel, D. ODY, L. LE DIREAC'H, 2001, Effet de la complexification de l'architecture des récifs artificiels du Parc National de Port-Cros (Méditerranée, France) sur les peuplements ichtyologiques | Parc national de Port-Cros et Porquerolles, date of consultation : 14/11/2022. <http://www.portcros-parcnational.fr/fr/rapports-scientifiques/effet-de-la-complexification-de-larchitecture-des-recifs-artificiels-du-parc>.
- P. Francour, J.-F. Magréau, P.-A. Mannoni, J.-M. Cottalorda, J. Gratiot, 2006, Ancrages écologiques permanents. Guide d'Aide à la Gestion des Aires Marines Protégées. Université de Nice-Sophia Antipolis & Parc National de Port-Cros, Nice : 68
- A.L. Luff, E.V. Seehan, M. Parry, N.D. Higgs, 2019, A simple mooring modification reduces impacts on seagrass meadows, date of consultation: 30/11/2022. <https://www.nature.com/articles/s41598-019-55425-y>
- Usabilis., 2018, Design Thinking Definition. Date of consultation : 10/11/2022. <https://www.usabilis.com/quest-ce-que-le-design-thinking>.
- L. Wait, 2022, Unchecked global emissions on track to initiate mass extinction of marine life. In : Princeton University, date of consultation : 30/11/2022. <https://www.princeton.edu/news/2022/04/28/unchecked-global-emissions-track-initiate-mass-extinction-marine-life>.

# Substitution of hexane in vegetable oil extraction using Computer Aided Molecular Design

Mohamad Nehmeh,<sup>a,b</sup> Ivonne Rodriguez-Donis,<sup>a</sup> Vincent Gerbaud,<sup>b</sup> and Sophie Thiebaud-Roux.<sup>a</sup>

<sup>a</sup>*Laboratoire de Chimie Agro-Industrielle, Université de Toulouse, INRAE, INP, Toulouse (Fr)*

<sup>b</sup>*Laboratoire de Génie Chimique, Université de Toulouse, CNRS, INP, UPS, Toulouse (Fr)*

## Abstract

Hexane is currently the preferred solvent for the extraction of oilseed oils. Despite a good selectivity towards lipids and an easy separation from the oil by distillation, it raises toxicity and non-renewable origin issues. Hence its replacement is investigated using Reverse Engineering methodology based on Computer Aided Molecular Design (CAMD). It starts by defining target values for selected physicochemical properties (boiling point, flash point, etc.) of the existing solvents used for oil extraction (Nehmeh et al., 2022). Then molecular structures which properties match these targets are sought: thousands of solvents were evaluated by means of our CAMD tool InBioSynSolv (IBSS) (Heintz et al., 2014). Results show that hexane is not the most performant solvent, impaired by a low flash point. Further refining of the candidates generated with IBSS is done by comparing their sigma-potential using COSMO-RS. In fine, alternative and greener solvents are identified for the substitution of hexane.

**Keywords:** Computer Aided Molecular Design, Reverse Engineering, Rapeseed oil, Oil extraction.

## 1. Introduction

The use of hexane in oil extraction is a widely spread practice. It is in fact the solvent of choice for industrial processes due to its selectivity towards lipids and ease of separation (Nehmeh et al., 2022). Nevertheless, this highly flammable solvent with an unpleasant gasoline-like odour, forms an explosive mixture with air (Pohanish, 2017). Besides, hexane is toxic to the environment (hazard statement H411) and is suspected of being reprotoxic (H361f) and fatal (H304) (Regulation (EC) No 1272/2008). Hitherto, its use in the production of foodstuff is still allowed under strict regulations, i.e. a maximum residue of 1 ppm in the oil (Directive 2009/32/EC).

The time and resource consuming basic “trial and error” methodology for finding solvent substitutes is nowadays replaced by a Reverse Engineering approach based on Computer Aided Molecular Design (CAMD). The methodology starts by defining a set of physicochemical properties related to the main solvent functionalities, along with the corresponding target values for each property. It is then followed by a search or a design of solvents meeting these requirements.

In this work, CAMD is used for the substitution of hexane in the extraction process of rapeseed (canola) oil. The solubility of the oil in the alternative solvents is first assessed using Hansen Solubility Parameters. Chem21 solvent selection guide (Prat et al., 2016) is then used to evaluate the safety, health and environmental impacts of the solvents, based on the hazard statements (GHS) available to date. A second validation is done by comparing the sigma-potential profile of the solvents and the oil, using COSMO-RS.

## 2. Methodology

Reverse Engineering is a 3-phase process. The first phase, the intelligence phase, consists of identifying the problem and defining a comprehensive list of solvent functionalities, described by target physico-chemical properties (e.g. solubility power described by Hansen Solubility Parameters). Afterwards, an individual performance for each property as well as a global performance function are defined. In the second phase, the design phase, two paths can be considered: the evaluation of solvents or the design of molecular structures using a CAMD tool. The last phase is about the choice, and deals with the selection of the best candidates, either pure components or mixtures.

### 2.1. Hansen Solubility Parameters

Grounded on the “like dissolves like” rule, the interactions between a solute and a solvent can be described using the Hansen Solubility Parameters (HSP),  $\delta D$ ,  $\delta P$  and  $\delta H$  (Hansen, 2007). The cohesive energy density  $\delta_T^2$  (Hildebrand solubility parameter) is defined as the amount of energy required to break a solvent into gas molecules, divided by its molar volume. It is further approximated as the sum of the energy densities:  $\delta D^2$ ,  $\delta P^2$  and  $\delta H^2$ , for London atomic dispersion, Keesom polar forces and hydrogen bonds. This can be visualized by plotting the parameters as coordinates into a 3D-space. The HSP values for a given solute, as well as a sphere radius encompassing solvents that solubilize the solute, can be determined experimentally by a series of solubility tests in solvents with known parameters. The sphere is constructed as to include a maximum of the solvents miscible with the solute. According to the theory, similar HSP values for a solvent and a solute indicate a higher affinity between them. Furthermore, the Relative Energy Difference RED for a solvent can be calculated as the ratio of the relative distance  $D$  between the solute and the given solvent, and the sphere radius  $R$ . A solvent “inside” the sphere, thus having a  $RED < 1$ , should dissolve the solute.

### 2.2. CAMD method by using IBSS

IBSS (InBioSynSolv) is a CAMD tool developed in the Laboratoire de Génie Chimique (LGC), based on a predictive and screening process. The prediction of physicochemical properties takes place through various implemented group contribution methods. As defined in the intelligence phase step, several physicochemical properties must be matched, each having a performance function  $Perf$  (eq.1) based on the target value for the property. For a given property  $p$ , the function compares the target value  $P$  with the predicted value  $x$ . The performance is set to 1 if the predicted value meets the target value, or is computed by a Gaussian function having a tolerance  $tol$  for a deviation  $val$ . A global performance function  $GloPerf$  (eq. 2) then transforms the multi-objective problem into a single-objective problem. During the molecular design, an optimization function  $OF$  (eq. 3) subject to  $i$  inequality and  $k$  equality is used in the aim of maximizing the  $GloPerf$ . The optimization variables are the molecular graph structure  $MG_i$  of the component  $i$  and the conditions  $cond_j$  (conditions under which the properties calculated e.g. temperature, pressure) (Heintz et al., 2014).

$$Perf_p = [\ln(val)] \times \left(\frac{P-x}{tol}\right)^2 \quad (\text{eq. 1})$$

$$GloPerf = \frac{\sum_{p=1}^{np} w_p \times Perf_p}{\sum_{p=1}^{np} w_p} \quad (\text{eq. 2})$$

$$OF = \max(GloPerf(MG_i, cond_j)) \quad (\text{eq. 3})$$

$$\text{Such that } P_k(MG_i, cond_j) = P_{k, fixed}$$

$$P_{i, lowerbound} \leq P_i(MG_i, cond_j) \leq P_{i, upperbound}$$

Table 1: Solvent properties, target values and used models

Functionality	Property	Target value	Parameters	Models and individual weights
Solency power	$\delta D$	$12.6 < \delta D < 20$ Based on the values of the test solvents that solubilized the oil	$\delta D < 12.6$ Val=0.8, Tol=0.5 $\delta D > 20$ Val'=0.5, Tol'=2.0	MB2010: 1.5 HSKASG2012:1.5
	$\delta P$	$\delta P < 12$ Based on the values of the test solvents that solubilized the oil	$\delta P > 12$ Val=0.7, Tol=2	MB2010: 1.5 HSKASG2012: 1.5
	$\delta H$	$\delta H < 19$ Based on the values of the test solvents that solubilized the oil	$\delta H > 19$ Val=0.5, Tol=2.0	MB2010: 1.5 HSKASG2012:1.5
	RED	RED < 1 A solvent with a RED < 1 should theoretically solubilize the oil	RED > 1 Val=0.4, Tol=0.5	MB2010: 4
Liquid state	Melting point	Melting point = 0 °C to keep solvent in the liquid form in normal temperature conditions	MP > 0 Val=0.8, Tol=2.0	HSKASG2012: 1 JR1987: 1
	Vapor pressure	Vapor pressure < $1 \times 10^{-4}$ bar This value corresponds to the limit of what is a VOC (Directive 1999/13/EC)	VP > $1 \times 10^{-4}$ Val=0.7, Tol=0.005	Riedel1954: 0.5
	Acentric factor	Acentric Factor < 1 Based on the values of the test solvents that solubilized the oil	AF > 1 Val= 0.8, Tol= 0.2	CDJ1993: 0.5 HSKASG2012: 0.5
Low flammability	Flash point	FP > 23 °C Solvents with Flash Point <23°C are said highly inflammable	FP < 23 Val=0.6, Tol=2.0	CPN2006: 1.5 HSKASG2012: 1.5
Energy related property	Boiling point	$60 \text{ °C} < BP < 150 \text{ °C}$ These values are in the set of specifications of the project	BP < 60 Val=0.5, Tol=10 BP > 150 Val'=0.5, Tol'=20	HSKASG2012: 1 MG2001: 1 JR1987: 1
	Vaporization Enthalpy	Vaporization Enthalpy < 50 kJ/mol The value is based on the enthalpy of vaporization of water	VE > 50 Val=0.8, Tol=2.0	Riedel1954: 1 CG1994: 1
Transport property	Density	$0.65 \text{ g/cm}^3 < \rho < 1.3 \text{ g/cm}^3$ Based on the values of the test solvents that solubilized the oil	$\rho < 0.65$ Val=0.1, Tol=0.1 $\rho > 1.3$ Val'=0.1, Tol'=0.075	MB2010: 1
	Viscosity	Viscosity < 5 mPa.s Based real data of the test solvents that solubilized the oil	$\mu > 5$ Val=0.5, Tol=0.5	JR1987: 0.5 CMMG2008: 0.5
	Surface tension	$15 \text{ mN/m} < ST < 40 \text{ mN/m}$ Based on the values of the test solvents that solubilized the oil	ST < 15 Val=0.8, Tol=1 ST > 40 Val'=0.8, Tol'=1	Pitzer1995: 0.5 ZuoStenby1997: 0.5
Water miscibility	$K_{OW}$	Log ( $K_{OW}$ ) < 4.5 Bioconcentration in living organism is a concern for high values of $K_{OW}$	Log ( $K_{OW}$ ) > 4.5 Val=0.4, Tol=0.6	MG2002: 0.5 HSKASG2012: 0.5
	$W_s$	Log ( $W_s$ ) < 4 log(mg/L) Low aqueous solubility is sought as the seeds contain some moisture	Log ( $W_s$ ) > 4 Val=0.4, Tol=0.5	MG2002: 1



### 2.3. Alternative solvents design methodology

The aforementioned three phases of Reverse Engineering methodology based on CAMD can be further refined in practice into five steps. The first step is the definition of the design problem by determining all the functionalities of the alternative solvents e.g. transport properties. The second step consists of converting the functionalities into physio-chemical properties that can be predicted e.g. volatility is converted to boiling point. For each property, a target value is determined based on the value of the existing solvents, or the process requirements (e.g. high flash point). The tolerance for each property performance function *Perf* is then defined, along with a weight for each property to compute the global performance function *GloPerf*. The third step is the identification and choice of the required property models (group contribution methods selected as to provide the smallest relative error with experimental data collected). Table 1 summarizes these three steps. The fourth step can be divided into 2 parts: (1) evaluation of solvents and (2) molecular design and structural optimization of compounds based on pre-defined chemical moieties implemented in the tool. This step provides a list of the best candidates along with their corresponding *GloPerf* values. The last step is the choice step. During this step, the best candidates are screened based on various criteria. Herein, the Chem 21 Solvent selection guide was employed for the screening of candidates based on safety, health and environment criteria. Each criterion is given a score between 1 and 10, 1 being the least impacting in the corresponding category. The safety criterion is mainly based on the flash point of the solvent, along with other properties that were ignored herein due to the lack of data (electrical conductivity, etc.). The health criterion takes into consideration the hazard statement given by REACH as well as the boiling point (as an assessment of volatility). The last criterion is also based on the environment hazard statement (H4xx) as well as the boiling point, as an estimation of VOC emissions but also the ease of recycling and separation from the product. A solvent is then classified as “recommended”, “problematic” or “hazardous” based on the three scores.

### 2.4. COSMO-RS computational method

COSMO-RS is a powerful theory based on a quantum-chemical approach, combined with statistical thermodynamics, for molecular description and thermodynamic properties prediction based on the chemical potential  $\mu$  of the molecule (Klamt, 1995; Eckert and Klamt, 2002). For molecules not found in the COSMObase, COSMO files are generated using a BP-TZVPD-FINE basis set, after geometry optimization using Turbomole V7.4 (GmbH). Statistical thermodynamics assure the transition from the microscopic molecular interactions, described by sigma ( $\sigma$ )-profile, to macroscopic thermodynamic properties, by means of the ( $\sigma$ )-potential. The latter is a measure of the system's affinity to a surface of polarity  $\sigma$ . The ( $\sigma$ )-potential similarity (SPS) is calculated as the sum of the differences between the 2 molecules using COSMOtherm (BIOVIA). It corresponds to the similarities of the compounds with respect to their properties as solvent. The smaller the compound's chemical potential profiles overlapping, the smaller the SPS. Rapeseed's bio-chemical composition varies widely depending on the seed's variety and other factors such as environmental conditions (Nehmeh et al., 2022). Triolein (OOO), is one of the main triglycerides present in the rapeseed oil with concentrations up to over 70% (Guan et al., 2016) and will be used as the base for the SPS calculations.

### 3. Results and Discussion

#### 3.1. Oil Sphere

After a series of solubility tests (25 wt% of oil) using a list of solvents carefully selected to fill and represent the Hansen space, the HSP of rapeseed oil were determined using HSPiP software. As the sphere did not have a perfect fit (around 0.9), the calculations were repeated 40 times in order to get a representative mean and standard deviation of the sphere centre coordinates as well as the radius ( $\delta D: 15.7 \pm 0.24$ ,  $\delta P: 2.8 \pm 0.44$ ,  $\delta H: 1.6 \pm 0.62$ ,  $R: 10.5 \pm 0.64$ ), as illustrated in figure 1. As can be predicted, the oil shows very low polarity and hydrogen bonding ability, mainly due to the presence of the triple ester groups ( $-\text{CH}_2\text{OOC}-$ ) on the glycerol backbone.

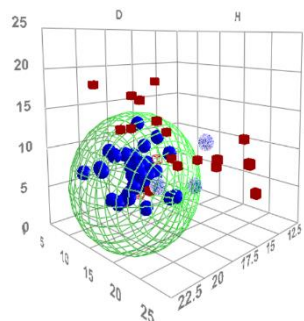


Figure 1: Hansen sphere for rapeseed oil.

#### 3.2. IBSS Evaluation

A list of over 3900 molecules (real and theoretical) were evaluated using IBSS. No molecular structure optimization was carried out. Over 2000 molecules showed a global performance higher than 0.8. This result is somewhat expected, as the oil sphere radius, 10.5, is quite large, and many molecules are within the solubility sphere ( $\text{RED} < 1$ : the property with the most weight). Therefore, the focus was put on the molecules that are closer to the sphere centre ( $\text{RED} < 0.5$ ), with acceptable boiling point ( $< 160^\circ\text{C}$ ), decreasing the selection down to 738 molecules. Over half of the remaining molecules were low polar hydrocarbons, like hexane, and are known to be usually toxic; they were then discarded. Special attention was given to oxygenated hydrocarbons, as their biosynthesis is generally possible, along with molecules containing heteroatoms e.g. nitrogen and fluorine. The Chem21 solvent selection guide was then applied to the 299 molecules left using all the hazard statement available to date, resulting in a list of 72 so-called “recommended” molecules, as per the same guide.

Name	Similarity	GloPerf	RED	BP (°C)	FP (°C)	Safety criteria	Health Criteria	Environment criteria	Final Criteria
Hexane	0.33	0.88	0.22	69	-24	7	7	7	Hazardous
Mol. 1	0.76	0.98	0.24	142	42	3	2	5	Recommended
Mol. 2	0.76	0.94	0.21	118	14	4	2	3	Recommended
Mol. 3	0.75	0.92	0.15	166	48	3	2	5	Recommended
Mol. 4	0.7	0.88	0.30	123	5	5	1	3	Recommended
Mol. 5	0.64	0.94	0.31	103	5	4	2	3	Recommended
Mol. 6	0.53	0.99	0.12	142	25	3	2	5	Recommended
Mol. 7	0.53	0.94	0.47	116	20	4	2	3	Recommended
Mol. 8	0.52	0.99	0.37	139	32	3	4	3	Recommended
Mol. 9	0.51	0.94	0.44		32	3	1	5	Recommended
Mol. 10	0.51	0.97	0.45	110	24	3	4	3	Recommended

Figure 2: Molecules with the highest SPS compared to hexane (BP & FP are experimental data).

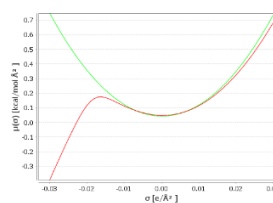


Figure 3: Triolein (red) vs hexane (green) sigma-potential.

#### 3.3. Sigma Profile and solubility calculation

The sigma potential similarity of triolein with the 72 recommended molecules ( $\text{GloPerf} > 0.8$ ,  $\text{RED} < 0.5$ ) from IBSS evaluation was then calculated. Figure 2 represents the most interesting molecules along with hexane. Based on the RED value, hexane seems like a good solvent for the solubilization of oil. Nevertheless, its SPS is very low, as its  $\sigma$ -potential profile shows a quite different shape than that of the oil (figure 3). Furthermore, hexane’s global performance is low (0.88), mainly due to its very low flash point, but also its relatively high vapour pressure ( $> 0.1$  bar), and it is classified as hazardous. Meanwhile, 10 of the studied molecules had a SPS higher than 0.5, and 40 other got a SPS higher than that of hexane, meaning that they potentially have better interactions with the oil.

#### 4. Conclusion

Alternative solvents for the extraction of rapeseed oil have been evaluated and compared with the conventional hexane, using Reverse Engineering approach. The Hansen Solubility Parameters of the oil were determined experimentally. Then, an evaluation and screening of thousands of molecules was achieved using IBSS, resulting in 299 molecules with a global performance higher than 0.8, and a Relative Energy Difference lower than 0.5. Among those, 72 are thought to have no or very little impact on safety, health, and environment. The  $\sigma$ -potential of these solvents as well as that of hexane were then compared to that of triolein. It can be clearly and safely said that hexane is theoretically not the most suitable solvent for oil extraction. Further experimental validation of the performance of the evaluated molecules is essential. Moreover, further research on the availability, price, and route of production of each alternative solvent is needed. Coupling Reverse Engineering with CAMD has allowed the efficient selection of new potential candidates for the substitution of hexane.

#### Acknowledgments

This research was funded by the 3BCAR Carnot Institute. This work has benefited from state aid managed by the Agence Nationale de la Recherche under the Investissements d'Avenir program bearing the reference ANR-18-EURE-0021.

#### References

- BIOVIA COSMOtherm, Release 2022; Dassault Systèmes. <http://www.3ds.com>.
- F. Eckert, A. Klamt, 2002. Fast solvent screening via quantum chemistry: COSMO-RS approach. *AIChE Journal* 48, 369–385.
- European Parliament, Council of 11 March 1999. Directive 1999/13/EC on the limitation of emissions of volatile organic compounds due to the use of organic solvents in certain activities and installations. URL <http://data.europa.eu/eli/dir/1999/13/oj/eng> (accessed 11.28.22).
- European Parliament, Council of 23 April 2009. Directive 2009/32/EC on the approximation of the laws of the Member States on extraction solvents used in the production of foodstuffs and food ingredients. URL <https://eur-lex.europa.eu/eli/dir/2009/32/oj> (accessed 28.11.22).
- M. Guan, H. Chen, X. Xiong, X. Lu, X. Li, F. Huang, C. Guan, 2016. A Study on Triacylglycerol Composition and the Structure of High-Oleic Rapeseed Oil. *Engineering* 2, 258–262.
- C. Hansen, 2007. Hansen Solubility Parameters: A User's Handbook, 2nd ed. CRC Press, Boca Raton.
- J. Heintz, J.-P. Belaud, N. Pandya, M. Teles Dos Santos, V. Gerbaud, 2014. Computer aided product design tool for sustainable product development. *Computers & Chemical Engineering* 71, 362–376.
- A. Klamt, 1995. Conductor-like screening model for real solvents: a new approach to the quantitative calculation of solvation phenomena. *The Journal of Physical Chemistry* 99, 2224–2235.
- M. Nehmeh, I. Rodriguez-Donis, A. Cavaco-Soares, P. Evon, V. Gerbaud, S. Thiebaud-Roux, 2022. Bio-Refinery of Oilseeds: Oil Extraction, Secondary Metabolites Separation towards Protein Meal Valorisation—A Review. *Processes* 10, 841.
- R.P. Pohanish, 2017. Sittig's Handbook of Toxic and Hazardous Chemicals and Carcinogens. William Andrew.
- D. Prat, A. Wells, J. Hayler, H. Sneddon, C. Robert McElroy, S. Abou-Shehada, P. Dunn, 2016. CHEM21 selection guide of classical-and less classical-solvents. *Green Chemistry* 18, 288–296.
- Regulation (EC) No 1272/2008 of the European parliament and of the council of 16 December 2008, 2008. Reach Classification, Labelling and Packaging of Substances and Mixtures.
- TURBOMOLE V7.4, TURBOMOLE GmbH, Karlsruhe, 2019, <http://www.turbomole.com>.

# Innovative process for manufacturing pharmaceutical mini-tablets using 3D printing

Varun Sundarkumar,<sup>a</sup> Wanning Wang,<sup>a</sup> Zoltan K. Nagy,<sup>a</sup> Gintaras Reklaitis<sup>a</sup>

<sup>a</sup>*Davidson School of Chemical Engineering, Purdue University, West Lafayette Indiana 47906 USA*

## Abstract

Providing drug products for pediatric patients is a challenging problem for the pharmaceutical industry. Children often require flexible low-dose medication with features like taste-masking and ease of swallowing. In recent years, mini-tablets have emerged as an attractive dosing solution that can meet these requirements. They are small form oral dosages around 2-4 mm in diameter that can be dispensed individually or in combination. Conventionally, they are made using methods like direct compression and hot melt extrusion. This study introduces a new technique to make mini-tablets: drop-on-demand 3D printing. Here the active ingredient is suspended in a liquid excipient, the formulation is printed as droplets and each drop is solidified to yield a mini-tablet. An optimal solvent bath that can uniformly capture mini-tablets is designed and dosages of Atorvastatin (active ingredient) are produced as a test case. Quality of these dosages is determined by measuring their content uniformity.

**Keywords:** pharmaceuticals, mini-tablets, 3D printing, drop-on-demand, low dose high precision drug products

## 1. Introduction

Pediatric medicines make up less than 10% of the overall drug market. Serving this patient group however poses a major challenge for the pharmaceutical industry. Factors like limited excipient compatibility, prevalence of rare diseases and smaller pediatric patient populations contribute to this (Milne & Bruss, 2008). Drug products made for children must also cater to their specific needs like having low dose strengths, taste masking and ease of swallowing etc., which in-turn make their production difficult.

Currently, two categories of drug products are popular for medicating children: liquid oral dosages and compounded dosages. Liquid oral drug products are attractive as they satisfy many of these requirements: they are flexible to dose, easy to consume and can be dispensed in low dose strengths. Compounded dosages are made in special pharmacies (called compounding pharmacies) where drug products are made at non-commercial specifications, by either altering the dose or changing the ingredients, to meet the needs of the patient. Both these products, however, face some limitations: 1) many active pharmaceutical ingredients (APIs) are unstable in aqueous media and prone to degradation 2) liquids are also susceptible to inaccurate dosing leading to adverse drug reactions. 3) In compounded dosages, achieving drug content uniformity and release profile matching commercial products is difficult (Zuccari et al., 2022).

In recent years, mini-tablets have emerged as an attractive alternative for pediatric medication. These are small-size solid oral drug products (2-4 mm in diameter) that are easy to swallow and can be dispensed individually or in combination in flexible doses. Two broad categories of techniques have been used to make them: powder based and 3D printing based. Powder based techniques include conventional methods like granulation

and direct compression (Stoltenberg & Breitzkreutz, 2011). Here the API is blended with excipient powders to yield a powder formulation. In granulation this formulation is processed into coarse aggregates. Each granule aggregate is then dispensed as a mini-tablet. In direct compression the powder formulation is compacted using small dies into mini-tablets. These methods rely on existing process technology and thus are easy to implement. They also can achieve high production rates. However, manufacturing low dose drug products is challenging for poorly flowing API powders. The second category of techniques is 3D printing, this is a manufacturing process in which the product is made in a layer-wise fashion. Hot melt extrusion is the most used 3D printing method for making mini-tablets (Krause et al., 2021). Here the API-excipient blend is extruded into a thin filament which is then melted and formed into the desired shape. Drug loading in the dosage can be easily customized by varying the number of printed layers. A key limitation of this method is the need for high operating temperatures which can cause degradation of the API or change its polymorphic form.

This study introduces a new method for making mini-tablets: using a drop-on-demand (DoD) inkjet printer. DoD falls under the umbrella of 3D printing methods, it processes a liquid formulation where the API is either dissolved or suspended in an excipient carrier fluid (İçten et al., 2017). The formulation is printed as droplets onto substrates like placebo tablets or capsules to make the drug product. This study aims to manufacture mini-tablets by solidifying individual drops generated by the DoD system. Printed droplets are discharged into an inert and viscous solvent bath where they settle slowly and solidify. Solid droplets are then washed and air dried to yield ready-to-use mini-tablets. The main advantage of this technique is producing dosages having low drug loadings with high precision. To develop this manufacturing route, first an optimal solvent bath is designed to uniformly capture and solidify the printed drops and next, mini-tablets for an API called atorvastatin are printed and tested for content uniformity.

## 2. Drop-on-demand printing of mini-tablets

As discussed before, DoD is a novel pharmaceutical 3D printing technique that builds dosages by printing multiple drops of an API containing formulation onto inert substrates. It has a versatile operation and can print a variety of formulation (melts, suspension etc.) and produce dosages with broad range of drug loadings. It can also incorporate emerging developments in pharmaceutical manufacturing like continuous processing and end-to-end operation (Sundarkumar et al., 2022).

Its working principle is as follows: the formulation ink for printing is held in the reservoir with constant agitation to ensure concentration homogeneity. Then a high-precision positive displacement pump dispenses a packet of fluid with constant volume through a nozzle in the form of a drop. This drop is then deposited on the desired substrate. The central idea in this study is to capture printed droplets and process them into individual mini-tablets. To achieve this a melt formulation is used, this formulation is then printed into an inert solvent bath (silicon oil) where the droplets self-solidify at room temperature (Figure 1). These solidified drops are then washed dried and collected as ready to use mini-tablets. Details of how the bath is designed is discussed in the results section.

Drug products have many critical quality attributes, such as, content uniformity, residual solvent content, dissolution behavior etc. For 3D printed drug products, content uniformity is an important metric as precise dosing is one of its key features. Thus, to test the quality of printed mini-tablets, uniformity in shape, weight and drug loading content across dosages are measured.

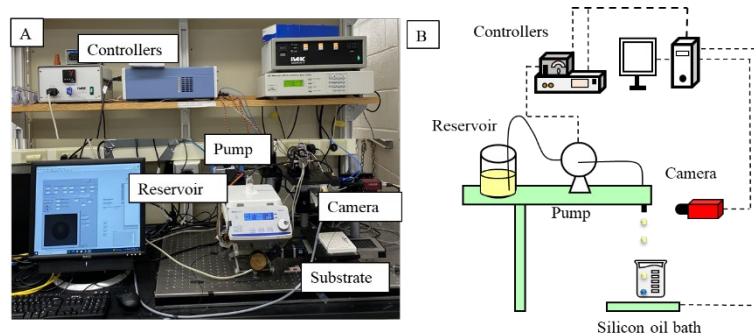


Figure 1. A) Apparatus and B) Schematic of the DoD printing system.

### 3. Materials

#### 3.1. Chemicals

Atorvastatin (Dr. Reddy's Laboratories), Polyethylene glycol 2000 (PEG) (Fisher chemicals), Silicon oil (polysiloxane, PMX 200 1000 cSt, Dow Chemical Company), Hexamethyldisiloxane (HMDSO) (Fisher chemicals). Before designing the process, the excipient material needs to be selected. PEG is an excipient widely used in drug products and acts as a dissolution enhancer. It melts at  $\sim 65$  °C and is a solid at room temperature. To solidify the printed droplets, silicon oil is used as a bath solvent as it is inert and does not interact with either the API or PEG. HMDSO is used as a washing fluid for the mini-tablets as it is a lighter inert silicon oil that can remove heavier oil adhering to the tablets.

#### 3.2. Instruments

Nikon Eclipse E600 microscope is used to image and size mini-tablets, Waters ultra-pressure liquid chromatography (UPLC) is used to measure drug loading in mini-tablets.

### 4. Results

#### 4.1. Designing the silicon oil bath

The first step in process development for this route is to design the inert solvent bath. Heat transfer (convective) due to the motion of the drop in the bath needs to be sufficient to solidify the mini-tablet. The critical process parameters here are bath viscosity and final drop temperature. Having low viscosity or high final drop temperature can lead to incomplete solidification or pooling of the droplets in the settling chamber respectively. On the other hand, high viscosities can lead to droplet aggregation (Figure 2). Optimal bath design is the one with the smallest settling chamber height (to reduce solvent amount required) and with the fastest rate of heat transfer.

To achieve this, a multi objective optimization problem can be posed with the goal of minimizing chamber height and solidification time (higher the heat transfer rate, shorter the solidification time). Decision variables here are bath viscosity  $\eta$  and final drop temperature  $T_f$ . Both variables are continuous: continuum in viscosity can be achieved by blending together base silicon oils (commercially available products) in required proportions. The bath is assumed to be at room temperature and thus temperature dependence of oil viscosity is ignored. To prevent pooling, an upper bound constraint is introduced on the final drop temperature; and to prevent drop aggregation, a minimum value is set for the distance between successive drops. Problem formulation is as follows:

$$\begin{aligned} & \min l_{\text{chamber}} \\ & \min t_{\text{solidification}} \end{aligned}$$

$$\text{st: } T_f < 35^\circ \text{C}, \quad \eta < \frac{5}{18} g d_{MT} (\rho_{MT} - \rho_{Si})$$

$$\eta \in [0.1, 10] \text{ Pas}, \quad T_f \in [23, 35]^\circ \text{C}$$

$l_{chamber}$  and  $t_{solidification}$  are calculated as follows:

1. Physical properties of the system:  $d_{MT} = 4 \text{ mm}$ ,  $\rho_{MT} = 1124 \text{ kgm}^{-3}$ ,  $C_{p_{MT}} = 2135 \text{ Jkg}^{-1}\text{K}^{-1}$ ,  $k_{MT} = 0.31 \text{ Wm}^{-1}\text{K}^{-1}$ ,  $T_i = 70^\circ \text{C}$ ,  $\rho_{Si} = 950 \text{ kgm}^{-3}$ ,  $k_{Si} = 0.15 \text{ Wm}^{-1}\text{K}^{-1}$ ,  $C_{p_{Si}} = 1250 \text{ Jmol}^{-1}\text{K}^{-1}$ ,  $T_{bath} = 23 \text{ C}$ ,  $\Delta H_{melt} = 100 \text{ kJkg}^{-1}$ ,  $T_{melting} = 45^\circ \text{C}$
2. First, calculate terminal settling velocity of the tablets.  $v_t = \frac{g d_{MT}^2}{18\eta} (\rho_{MT} - \rho_f)$ .
3. Next, compute heat transfer rate (assume convection dominated).  $h = Nu \frac{k}{d}$ ,  $Nu = 2 + \left(0.4Re^{0.5} + 0.06Re^{\frac{2}{3}}\right) Pr^{0.4}$ ,  $Re = \frac{\rho_f v_t d}{\eta}$ .
4. Finally, calculate solidification time  $t_s = \frac{\rho_{MT} V C_P}{hA} \ln\left(\frac{T_i - T_b}{T_f - T_b}\right) + \frac{\Delta H \rho V}{hA(T_m - T_b)}$  and  $l_{chamber} = v_t * t_s$ .
5. Verify if convection driven heat transfer assumption is valid, i.e., if  $\frac{hd}{k} \gg 1$ .

The first constraint acts as a variable bound,  $35^\circ \text{C}$  is chosen because it is half the initial temperature and the formulation is fully solid at this temperature. For the second constraint, we assume that for aggregation to be arrested a gap of at least one drop diameter between successive drops is necessary  $l_{inter\ drop} > d_{MT}$ . As drop generation rate is constant (one every 5 sec), the constraint can be framed in terms of settling velocity:  $v_t > \frac{d_{MT}}{5}$  which is equivalent to  $\eta < \frac{5}{18} g d_{MT} (\rho_{MT} - \rho_{Si})$ .

The problem is solved by building a Pareto front (figure 2) using the Non-dominated Sorting Genetic Algorithm II (NSGA2) available in the Python library 'pymoo'. Although this algorithm does not guarantee global optimality, it provides a quick and efficient way to generate diverse non-dominated solutions that can represent the best trade-offs between objectives. From this Pareto set, the point closest to having a bath viscosity of 0.95 Pas is chosen for operation as it is closest to the material at hand. The operating point chosen is as follows:

$$\eta = 0.95 \text{ Pas}, T_f = 34.9, l_{chamber} = 6.5 \text{ cm and } t_{solidification} = 40.8 \text{ sec.}$$

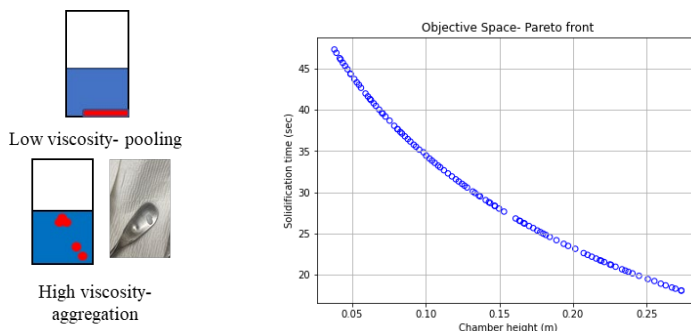


Figure 2. Design of solvent bath and settling chamber- optimization solution.

#### 4.2. Consistency measurements for the printed mini-tablets

After determining bath viscosity, mini-tablet manufacturing is carried out. Drug products with two dose levels of atorvastatin are targeted: 0.1 and 1 mg. In each case formulation drops are printed into the bath, solidified, washed with HMDSO and air dried (figure 3).

To test whether this operation yielded consistent mini-tablets, uniformity of three parameters is measured across dosages: shape, weight and drug loading. To measure variation in shape, the diameter of each mini-tablet is determined under a microscope. Drug loading in each mini-tablet is measured by dissolved it in methanol and analyzing on an offline UPLC. 25 dosages of each dose strength are used for this analysis and the results are summarized in tables 1 and 2.

The results show that for both dose strengths highly precise mini-tablets are obtained. The variation in all three parameters is very low and is well under the regulatory threshold (5% relative standard deviation for drug loading as specified by the United States Food and Drug Administration). Drop size can be altered by changing the DoD printing conditions like ejection rate, volume per drop etc. Based on concentration of the active ingredient in the formulation, drop size can be adjusted to deliver the desired dose strengths in mini-tablets. High consistency seen in the process can be explained by two reasons: 1) it is easier to disperse API in liquid excipients compared to powder excipients leading to more uniform drug loadings; and 2) each mini-tablet is made of exactly one droplet; thus, uniformity of drop formation translates into uniformity in shape and weight of mini tablets. In this study, polymorphic transformation of the API crystals is not considered. Dissolving the API in a molten polymer like PEG and solidifying it yields amorphous solid dispersions. If polymorph control is critical then the API crystals can be suspended in an inert excipient that does not dissolve it (like triglyceride oils). This method also faces some limitations: the mini-tablets produced lack hardness which would make their transportation difficult. Manufacturing high drug loading mini-tablets is also challenging, as increasing particle loading beyond a certain concentration yields a paste-like formulation that does not generate uniform droplets.

To realize the full potential of this DoD approach (flexible dosage production rate, quick product changeovers etc.), the mini-tablet manufacturing process needs to be made entirely continuous. Currently, a part of the process (up to drop solidification) is continuous while the subsequent washing and drying steps are carried out batch-wise. Continuous operation can be achieved by integrating the DoD system with a continuous filter dryer unit. To support emerging initiatives in pharmaceutical manufacturing like digitalization and quality-by-design, sensors can be installed into the platform to track drug loading in the manufactured mini-tablets in real time. This will allow for building quality into the product during the manufacturing stage itself, and thus will mitigate the need for post manufacturing quality checks.

Table 1. Consistency results for 0.1 mg mini-tablets.

Parameter	Mean	Relative stdev
Diameter	3.14 mm	1.04 %
Weight	21.27 mg	1.54 %
Drug loading	0.107 mg	2.31 %

Table 2. Consistency results for 1 mg mini-tablets.

Parameter	Mean	Relative stdev
Diameter	2.84 mm	1.22 %
Weight	16.26 mg	2.18 %

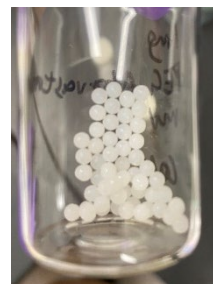


Figure 3. Printed atorvastatin mini-tablets.



Drug loading	1.01 mg	2.44 %
--------------	---------	--------

## 5. Conclusions

This study introduces a novel technique to manufacture pharmaceutical mini-tablets. API is dissolved in a molten PEG formulation and its droplets are printed and solidified in an inert silicon oil bath. These mini-tablets are seen to show a high degree of consistency in shape, weight and drug loading across dosages. To assess if these dosages are equivalent to commercial products, their dissolution behavior and long-term stability need to be determined. This technique fills an important gap in the pediatric drug market- producing high-precision low dose mini-tablets. In conjunction with the other manufacturing techniques, this method can accelerate the adoption of pharmaceutical mini-tablets.

## 6. Acknowledgements

The authors would like to thank the Food & Drug Administration for partially funding this work, in part, through grant (U01FD006738). Views expressed in written materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does any mention of trade names, commercial practices, or organization imply endorsement by the United States Government. This material is based upon work partially supported by the National Science Foundation under Grant No. 2132142.

## References

- İçten, E., Purohit, H. S., Wallace, C., Giridhar, A., Taylor, L. S., Nagy, Z. K., & Reklaitis, G. V. (2017). Dropwise additive manufacturing of pharmaceutical products for amorphous and self emulsifying drug delivery systems. *International Journal of Pharmaceutics*, 524(1–2). <https://doi.org/10.1016/j.ijpharm.2017.04.003>
- Krause, J., Müller, L., Sarwinska, D., Seidlitz, A., Sznitowska, M., & Weitschies, W. (2021). 3D printing of mini tablets for pediatric use. *Pharmaceutics*, 14(2). <https://doi.org/10.3390/ph14020143>
- Milne, C. P., & Bruss, J. B. (2008). The economics of pediatric formulation development for off-patent drugs. *Clinical Therapeutics*, 30(11). <https://doi.org/10.1016/j.clinthera.2008.11.019>
- Stoltenberg, I., & Breitkreutz, J. (2011). Orally disintegrating mini-tablets (ODMTs) - A novel solid oral dosage form for paediatric use. *European Journal of Pharmaceutics and Biopharmaceutics*, 78(3). <https://doi.org/10.1016/j.ejpb.2011.02.005>
- Sundarkumar, V., Nagy, Z. K., & Reklaitis, G. V. (2022). Small-Scale Continuous Drug Product Manufacturing using Dropwise Additive Manufacturing and Three Phase Settling for Integration with Upstream Drug Substance Production. *Journal of Pharmaceutical Sciences*. <https://doi.org/10.1016/j.xphs.2022.03.009>
- Zuccari, G., Alfei, S., Marimpietri, D., Iurilli, V., Barabino, P., & Marchitto, L. (2022). Mini-Tablets: A Valid Strategy to Combine Efficacy and Safety in Pediatrics. In *Pharmaceutics* (Vol. 15, Issue 1). <https://doi.org/10.3390/ph15010108>

# Process simulation coupled with life cycle assessment of an alternative polyamide precursor

Andrea Mio,<sup>a\*</sup> Maurizio Fermeglia,<sup>a</sup>

<sup>a</sup> *MolNBL@UniTS, Dipartimento di Ingegneria e Architettura, Università degli studi di Trieste, Piazzale Europa 1, 34127 Trieste, Italy*

## Abstract

A detailed process simulation for the production process of a valuable alternative reactant to produce a polyamide, *i.e.*, m-xylylenediamine (MXDA) using Aspen Plus has been performed. A comparative life cycle assessment of MXDA with an equivalent polymer precursor, *i.e.*, hexamethylenediamine (HMDA), has been carried out using the material and energy balances obtained through process simulation. Since neither of the two compounds outperformed the other in the impact assessment's environmental analysis, a general trade-off must be determined, even though slightly better performances are gained by m-xylylenediamine. The uncertainty related to the energy demand, which is one of the most influencing parameters, has been assessed to identify the implications of energy losses. The outcomes are actually reversed when the energy requirement for MXDA is increased by 25%, suggesting a general preference for HMDA. Therefore, particular care should be taken on energy optimization when implementing MXDA manufacturing process industrially.

**Keywords:** LCA, Process Simulation, Multiscale Modeling, polyamide, xylylenediamine, hexamethylenediamine

## 1. Introduction

After the pandemic, the manufacturing of polymers is facing significant challenges as a result of the enormous rise in demand, which is hardly being met by the supply. In this panorama, a major role is played by polyamide production, whose global market size for Nylon 6 and Nylon66 was estimated at 32.6 B\$ in 2022 and is expected to reach 53.04 B\$ in 2030 (Grand View Research, 2020). The tremendous growth in Nylon global market can mainly be attributed to an increasing demand of China market, along with a consistent growth of the automotive industry. The principal application of Nylon6 is in the textile industry (Deopura and Padaki, 2015), *e.g.*, carpets, fabrics and clothes, while Nylon 66, which is characterized by greater wear, thermal and structural resistances, is commonly employed in the transport sector (Ishikawa et al., 2018) and for metal replacements in advanced applications, *e.g.*, marine engine covers (Mio et al., 2021).

Due to the disruption of the supply chains, the sharp rise in the price of raw materials and the growing market demand for sustainable products, producers are being compelled to find substitute precursors. To prevent wasting time, effort, and money on a promising new synthetic route that might become more burdensome when applied at an industrial scale, the introduction of novel synthetic pathways for polyamides needs to be evaluated from an environmental standpoint already at an early-stage design. The environmental implications of a product or a manufacturing process are accurately evaluated by LCA methodology (International Standards Organisation, 2006), which allows to address the environmental impacts of the entire life cycle of goods, from raw materials extraction (cradle) to disposal at their end of life (grave). However, it is frequently challenging to

gather data on the materials and energy consumptions, the gaseous and liquid emissions, the wastes, and the by-products connected to the whole lifecycle of the substance under examination during product development in the chemical sector (Mio et al., 2018). Besides the usage of well-established databases, it is common practice to resort to process simulation for filling the gaps of life cycle inventories. Focusing on the polymer production sector, Gracida-Alvarez et al. (2019), coupled a carbon footprint analysis with a process simulation of fast pyrolysis for converting polyethylene plastic waste to monomers and aromatics, while Bello et al. (2020) conducted a life cycle assessment of a precursor of poly(ethylene furanoate), *i.e.*, 2,5-furandicarboxylic acid, using process simulation to perform sensitivity analyses and identify the major environmental hotspots. In this paper, a detailed process simulation for the production process of a valuable alternative reactant to produce a polyamide, *i.e.*, m-xylylenediamine (MXDA), has been assessed from an environmental standpoint.

## 2. Materials and Methods

Hexamethylenediamine (HMDA) and adipic acid are combined to form Nylon 66, following a polycondensation mechanism. While adipic acid is created via a multitude of syntheses that all start with benzene, industrial production of HMDA uses either butadiene or propylene via acrylonitrile. By comparing them from an economic and environmental perspective, Dros et al. (2015) have outlined and evaluated either the conventional production pathway and novel bio-based ones.

With the aim of providing a new reactant for producing a polyamide that can replace traditional Nylon 66, the production pathway for m-xylylenediamine (MXDA) has been modelled. In fact, being an aliphatic diamine with two reactive amine groups at the opposite side of the molecule, MXDA can easily replace HMDA in the polycondensation mechanism (Hugo et al., 2008a), or serve as an epoxy resin curing agent (Kutsuna and Kihara, 2008).

The production of MXDA has been described by several patented processes owned by different companies, *e.g.*, BASF (Ernst et al., 2008b, 2008a; Hugo et al., 2008a, 2008b, 2004) and Mitsubishi Gas Chemical Company (Kumano et al., 2013, 2011; Nakamura et al., 2003). As a result of a thorough investigation of the patents declared performance, the most recent process design suggested by Kumano et al. (2013) has been selected as the best available technique (BAT), due to a greater yield and lower energy consumption, as a result of the limited use of distillation columns. Moreover, this patent provides detailed information regarding the reaction duration, conditions, and, sometimes, waste by-products, which can be assumed as emissions. The published production process of MXDA follows five main steps, which have been modelled in the Aspen Plus flowsheet shown in Figure 1.

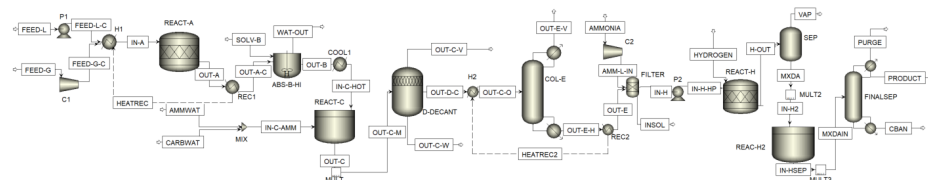


Figure 1: Flowsheet of the production of m-xylylenediamine

(i) The chain of reactions starts with an ammoxidation stage in REACT-A, where m-xylene interacts with ammonia and oxygen to produce isophthalonitrile (IPN) at 420 °C

and 2 bar. In the first separation step, m-tolunitrile (TN) is used to fill unit ABS-B to conduct a liquid-vapor absorption process that directly injects the liquid-vapor stream into the solvent at 140 °C and 0.57 bar, aiming to separate IPN from by-products and contaminants. (ii) The subsequent addition of an aqueous solution of ammonia and ammonium carbonate in REACT-C at 110 °C and 2 bar neutralizes the resultant solution. (iii) In decanter D-DECANT, the mixed solution of organic and aqueous phases is separated at atmospheric pressure. (iv) The organic phase is then further separated under vacuum (0.35 bar) in COL-E to isolate the IPN from the solvent TN and the major part of volatile components. (v) After an ammonia addition that precipitates the insoluble components, IPN is hydrogenated using a nickel catalyst under medium temperature (80 °C) and high pressure (80 bar) in REAC-H2 to produce MXDA, which is then purified in the FINAL-SEP rectifier under vacuum. Due to the formation of undesired chemicals from a partial hydrogenation, a polymerization, or cleavage of the product, ideal reaction conditions (Nakamura et al., 2003) and proper solvent (Row et al., 2007) are essential in this hydrogenation reaction to yield a high grade of MXDA.

Due to the nature of the chemicals involved, the NRTL thermodynamic model has been set, using UNIFAC group contribution method to generate missing binary interaction parameters. High-pressure hydrogenation required an Equation of State (EOS)-based model, *i.e.*, Peng-Robinson, instead. The experimental process follows a semi-batch regime and has been scaled-up considering an initial charge of 350 kg of xylene to meet the needs of an industrial scale production. Heat integration has been implemented in order to transfer the heat from outlet hot streams to inlet cold streams. Since there were no known kinetic models for the reactions involved, conversion reactors were employed in an effort to achieve the yield specified in the patent, while reactions enthalpies were calculated by means of the enthalpies of formation of the components. When the yields were not specified, a conversion of 95% of the reactants were assumed, as reported within ecoinvent guidelines for conservative scenarios (Weidema et al., 2013).

Aiming at determining the most environmentally friendly polymer precursor, openLCA was used to compare the life cycle assessments of the two alternatives, after obtaining the material and energy balances through process simulation. The functional unit was defined as *1 kg of polymer precursor in the European market*, using a cradle-to-gate approach, *i.e.*, including raw materials extraction and refining, transportations and production of the substances. For the manufacture of reactants (such as xylene, ammonia, oxygen, etc.) secondary data have been retrieved using ecoinvent v3.8, with allocation cut-off method. Moreover, the unidentified byproducts and the unreacted chemicals have been modelled as aqueous emissions of analogous substances with a characterization factor specified within ReCiPe 2016 Midpoint (H), which has also been used during the life cycle impact assessment phase.

### 3. Results and Discussion

The outcomes of the comparative LCA between the alternative precursors are shown in Figure 2 and Table 1, where the contributions from several sources (raw materials, utilities and production process) have been highlighted.

According to the results of the impact category analysis, MXDA is less burdensome than HMDA in 10 of the 18 impact categories that were analyzed, with MXDA scores ranging from 50 to 80% of HMDA's ratings. The environmental performances of MXDA are worse when considering freshwater, marine and terrestrial ecotoxicity, marine and freshwater eutrophication, water consumption, land use, ionizing radiation and human toxicity (non-carcinogenic).

Table 1: Life Cycle Impact Assessment results for HMDA and MXDA

Impact category	Unit	HMDA				MXDA			
		Raw Materials	Utilities	Production	Total	Raw Materials	Utilities	Production	Total
Fine particulate matter formation	kg PM2.5 eq	3.27E-03	2.55E-03		5.82E-03	3.34E-03	2.25E-03		4.17E-03
Fossil resource scarcity	kg oil eq	1.49E+00	1.72E+00		3.20E+00	2.21E+00	7.34E-01		2.69E+00
Freshwater ecotoxicity	kg 1,4-DCB	4.34E-02	3.50E-02	1.31E-02	9.16E-02	6.61E-02	7.29E-02	7.23E-02	1.71E-01
Freshwater eutrophication	kg P eq	2.01E-04	4.54E-04		6.55E-04	1.03E-03	7.85E-04		8.87E-04
Global warming	kg CO2 eq	2.16E+00	3.25E+00	6.16E-01	6.02E+00	3.86E+00	2.32E+00	5.52E-01	5.77E+00
Human carcinogenic toxicity	kg 1,4-DCB	6.79E-02	7.92E-02		1.47E-01	1.32E-01	7.85E-02		1.40E-01
Human non-carcinogenic toxicity	kg 1,4-DCB	8.12E-01	7.72E-01		1.58E+00	1.64E+00	1.21E+00		1.65E+00
Ionizing radiation	kBq Co-60 eq	5.45E-02	1.61E-01		2.15E-01	5.15E-01	3.57E-01		3.90E-01
Land use	m2a crop eq	1.04E-02	2.41E-02		3.45E-02	2.71E-02	2.64E-02		6.13E-02
Marine ecotoxicity	kg 1,4-DCB	5.78E-02	5.30E-02	1.86E-04	1.11E-01	8.97E-02	9.54E-02	6.31E-03	1.38E-01
Marine eutrophication	kg N eq	3.20E-05	3.27E-05		6.47E-05	2.99E-04	5.60E-05		2.87E-04
Mineral resource scarcity	kg Cu eq	1.20E-02	1.77E-03		1.38E-02	7.26E-03	2.13E-03		8.00E-03
Ozone formation, Human health	kg NOx eq	3.32E-03	4.22E-03		7.54E-03	5.84E-03	3.20E-03		7.28E-03
Ozone formation, Terrestrial ecosystems	kg NOx eq	3.52E-03	4.37E-03		7.89E-03	6.13E-03	3.26E-03		7.62E-03
Stratospheric ozone depletion	kg CFC11 eq	5.97E-07	1.18E-06		1.77E-06	1.05E-06	7.06E-07		1.31E-06
Terrestrial acidification	kg SO2 eq	9.91E-03	7.42E-03		1.73E-02	9.36E-03	6.18E-03		1.20E-02
Terrestrial ecotoxicity	kg 1,4-DCB	1.88E+00	3.98E+00	2.31E-05	5.86E+00	4.86E+00	3.09E+00	4.20E-01	7.59E+00
Water consumption	m3	3.33E-02	6.49E-03		3.98E-02	9.92E-02	1.25E-02		6.19E-02

The ionizing radiation, human non-carcinogenic toxicity, marine ecotoxicity, and freshwater eutrophication are primarily influenced by the electricity consumption, which is higher for MXDA than HMDA. This is counterbalanced by a greater heat demand for HMDA, resulting in lower impacts for MXDA in terms of global warming potential, fossil resource scarcity, fine particulate matter formation, human carcinogenic toxicity and ozone-related impact categories. Freshwater ecotoxicity is affected by the m-xylylenediamine production due to the presence of a significant quantity of contaminants, primarily byproducts and impurities in wastewater, that must be further purified before being discharged into the environment. Raw materials for MXDA mainly affects marine eutrophication, due to the usage of ammonia and ammonium carbonate, as well as terrestrial ecotoxicity due to the use of m-xylene, while water consumption has been negatively impacted by both xylene and ammonia production. Raw materials for HMDA are more burdensome in terms mineral resource scarcity and terrestrial acidification, due to sulfur-based reactants. While the impacts of raw materials among the different impact categories are equally distributed between the two precursors, the impacts related to the energy consumptions are generally higher for MXDA.

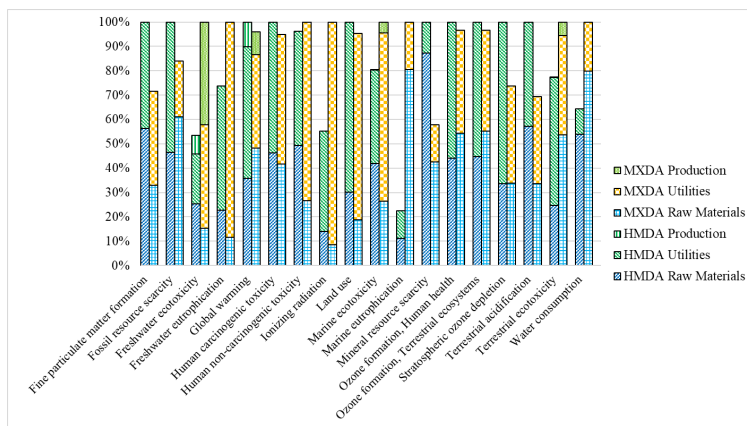


Figure 2: Comparison between the Life Cycle Impact Assessment results for hexamethylenediamine (HMDA) and m-xylylenediamine (MXDA)

Even though there is a limited use of the distillation column in the MXDA process, resulting in lower heat required, the electricity consumptions of compressors and pumps is still demanding from an environmental viewpoint. Since process simulation provided

limited information due to a lack of available data for MXDA, the assumptions on energy consumptions have been subjected to a sensitivity analysis. Therefore, in order to emphasize the effects of a different energy demand from simulated data, both the heat and electricity demands have been adjusted by a factor of  $\pm 25\%$ . The results of the sensitivity analysis are reported in Figure 3.

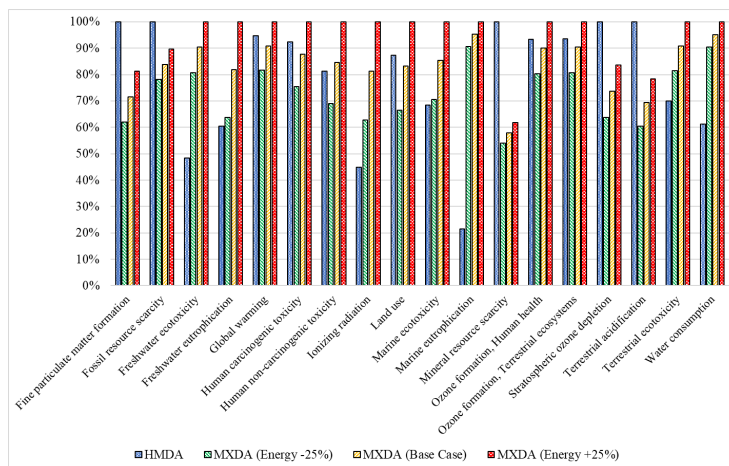


Figure 3: Sensitivity analysis results for different MXDA energy demand scenarios ( $\pm 25\%$ )

As expected, an energy demand increment of 25% worsens the environmental performance of MXDA for a multitude of impact categories, being the most burdensome precursor for 13 out of 18 impact categories. In particular, Global warming, Human carcinogenic toxicity, Land use, and Ozone formations gained worse scores than HMDA, degrading the MXDA base case that achieved superior performances in comparison with the other monomer. The precursors' relative positions do not change when the MXDA energy demand is reduced by 25% (apart from their non-carcinogenic toxicity to humans), but their impacts do get closer in situations where the MXDA results were higher than HMDA in the base case scenario, while their difference increases in impact categories where the MXDA results were lower than HMDA for base case scenario.

#### 4. Conclusions

This paper investigated the environmental performances of a polyamide precursor, *i.e.*, *m*-xylylenediamine, generating its life cycle inventory using a publish patent coupled with process simulation. The outcomes of the study have been compared with an analogous assessment performed on a traditional polyamide precursor, *i.e.*, hexamethylenediamine. The outcomes showed how the environmental performances of the two precursors are mostly equivalent, with slightly better performances for *m*-xylylenediamine. However, the production energy demand (both heat and electricity) played an essential role to determine the most environmentally friendly precursor. Moving to an industrial scale to validate the process simulations outcomes is therefore necessary in order to determine the correct energy usage and modify the life cycle inventory accordingly.

This paper used process simulation along with life cycle assessment to fill in the existing gaps in life cycle inventories of novel materials and substances, which is a well-known example of the application of multiscale modeling for the early-design assessment of products and processes.

## References

- Bello, S., Méndez-Trelles, P., Rodil, E., Feijoo, G., Moreira, M.T., 2020. Towards improving the sustainability of bioplastics: Process modelling and life cycle assessment of two separation routes for 2,5-furandicarboxylic acid. *Sep. Purif. Technol.* 233, 116056.
- Deopura, B.L., Padaki, N. V., 2015. Synthetic Textile Fibres: Polyamide, Polyester and Aramid Fibres. *Text. Fash. Mater. Des. Technol.* 97–114.
- Dros, A.B., Larue, O., Reimond, A., De Campo, F., Pera-Titus, M., 2015. Hexamethylenediamine (HMDA) from fossil- vs. bio-based routes: an economic and life cycle assessment comparative study. *Green Chem.* 17, 4760–4772. <https://doi.org/10.1039/c5gc01549a>
- Ernst, M., Hahn, T., Wenz, K., Hugo, R., Melder, J., 2008a. Method for producing a Xylylenediamine. US 2008/0214871 A1.
- Ernst, M., Wenz, K., Hugo, R., Melder, J., Hahn, T., 2008b. Method for producing a Xylylenediamine. US 2008/0154061 A1.
- Gracida-Alvarez, U.R., Winjobi, O., Sacramento-Rivero, J.C., Shonnard, D.R., 2019. System Analyses of High-Value Chemicals and Fuels from a Waste High-Density Polyethylene Refinery. Part 2: Carbon Footprint Analysis and Regional Electricity Effects. *ACS Sustain. Chem. Eng.* 7, 18267–18278. <https://doi.org/https://doi.org/10.1021/acssuschemeng.9b04763>
- Grand View Research, 2020. Nylon Market Size, Share & Trends Analysis Report.
- Hugo, R., Jourdan, S., Wenz, K., Preis, T., Weck, A., 2008a. Preparation of Xylylenediamine (XDA). US 7,323,598 B2.
- Hugo, R., Jourdan, S., Wenz, K., Preiss, T., Weck, A., 2008b. Preparation of Xylylenediamine (XDA). US7368610.
- Hugo, R., Wenz, K., Wambsganss, R., Jourdan, S., Preiss, T., 2004. Method for the Production of Diaminoxylene By Continuous Hydrogenation of Liquid Phthalonitrile. WO2005026099A1.
- International Standards Organisation, 2006. ISO 14044 Environmental management - Life cycle assessment - Requirements and guidelines.
- Ishikawa, T., Amaoka, K., Masubuchi, Y., Yamamoto, T., Yamanaka, A., Arai, M., Takahashi, J., 2018. Overview of automotive structural composites technology developments in Japan. *Compos. Sci. Technol.* 155, 221–246.
- Kumano, T., Ibi, Y., Arasuna, K., Nagao, S., 2013. Method for Producing Xylylenediamine. EP2671864 A1.
- Kumano, T., Nakaya, K., Kato, K., Shigematsu, R., 2011. Production of Xylylenediamines. EP 1873137 B1.
- Kutsuna, T., Kihara, S., 2008. Adhesive of epoxy resin and curing agent with xylylenediamine structure. US 7425598 B2.
- Mio, A., Bertagna, S., Cozzarini, L., Laurini, E., Bucci, V., Marinò, A., Fermeglia, M., 2021. Multiscale modelling techniques in life cycle assessment: Application to nanostructured polymer systems in the maritime industry. *Sustain. Mater. Technol.* 29, e00327. <https://doi.org/10.1016/J.SUSMAT.2021.E00327>
- Mio, A., Limleamthong, P., Guillén-Gosaíbez, G., Fermeglia, M., 2018. Sustainability Evaluation of Alternative Routes for Fine Chemicals Production in an Early Stage of Process Design Adopting Process Simulation along with Data Envelopment Analysis. *Ind. Eng. Chem. Res.* 7946–7960. <https://doi.org/10.1021/acs.iecr.7b05126>
- Nakamura, K., Amakawa, K., Shitara, T., 2003. Method for producing high purity Xylylenediamine. US 2003/0013917 A1.
- Row, S.W., Chae, T.Y., Yoo, K.S., Lee, S.D., Lee, D.W., Shul, Y., 2007. Effect of reaction solvent on the hydrogenation of isophthalonitrile for meta-xylylenediamine preparation. *Can. J. Chem. Eng.* 85, 925–928. <https://doi.org/10.1002/cjce.5450850614>
- Weidema, B.P., Bauer, C., Hischier, R., Mutel, C., Nemeček, T., Reinhard, J., Vadenbo, C.O., Wenet, G., 2013. Data quality guideline for the ecoinvent database version 3. *Ecoinvent Report 1 (v3)*. Swiss Cent. Life Cycle Invent. 3, 169.

# Connectivity Matrix-based Descriptors with Deep Learning for Estimation of Pure Component Properties

Qiong Pan,<sup>a</sup> Xiaolei Fan,<sup>b</sup> Jie Li,<sup>a</sup>

<sup>a</sup>*Centre for Process Integration, Department of Chemical Engineering, School of Engineering, The University of Manchester, Manchester M13 9PL, UK*

<sup>b</sup>*Department of Chemical Engineering, School of Engineering, The University of Manchester, Manchester M13 9PL, UK*

## Abstract

Physicochemical properties are fundamental for the design of chemical products and processes. Various approaches including cheminformatic and graph-based methods have been applied for the estimation of pure component properties. In this work, we coupled our connectivity matrix-based molecular structure representation method with a deep learning neural network to develop molecular structure-property relationship models. Molecular structure information is represented by the connectivity matrix, while substructures information is represented by the extracted submatrix. This extraction does not cause any loss of structural information, substructural information can be recovered using the stored matrix. Matrices are transferred into features and characterised based on their eigenvalues. Statistical and mathematical approaches, followed by deep learning neural network models, are developed to correlate the structure-property relationship. Results on the case studies of normal boiling point show that the accuracy of the deep neural network model improved by 9.48% over the previous neural network model.

**Keywords:** molecular descriptor, deep neural network, property estimation, connectivity matrix

## 1. Introduction

Molecular properties such as the physical, chemical, and structural properties are fundamental characteristics of molecules. Subtle structural differences may lead to significant property differences. Therefore, accurate property determination is essential for effective chemical compound design and utilisation (Katritzky et al., 1997). However, measuring molecular properties by experimental approaches is laborious and expensive. As a result, estimating molecular properties by developing mathematical models based on existing experimental data is a feasible approach (Wen et al., 2021). Generally, current approaches for property prediction include correlation-based approaches, the group contribution (GC) method, quantitative structure-activity/property relationship (QSAR/QSPR), and graph convolution methods. Correlation-based approaches are the easiest because of the concise format of correlation equations (Korsten, 1997). One of the shortcomings is that prediction accuracy is usually not guaranteed, especially for compounds with heteroatoms. The GC method is under the assumption that a molecule's properties are determined by solving a set of equations that maps the number of each group to experimental property value, a property of a molecule is then calculated through



the summation of all group contributions (Hukkerikar et al., 2012). One significant challenge of the GC method is that no method reported in the literature can be used for the systematic generation of a molecule's structural groups (or molecular fragments). The definition of groups usually has a significant effect on the performance of the prediction accuracy. It can lead to an inaccurate estimation if any of the predefined groups cannot describe one part of its molecular structure (Gani et al., 2005). Although reported GC methods predefined hundreds of groups that can represent most of the common structural characteristics of a compound, an automatic group generation approach tailored to the unique molecular pool can also improve the accuracy of structure-property correlation. Let alone the population of chemical compounds increases substantially because of the rapid progress of chemical synthesis techniques (Yoshida et al., 2011). Chemical graph theory is a branch of mathematical chemistry that models chemical topology phenomena mathematically. The basic idea of chemical graph theory is that atoms and bonds of a molecule can be thought of as nodes and edges in a graph (Randić, 2008). Therefore, topological indices representing molecular structure properties can be generated from the graph. Different to GC groups, which assume that each group contributes independently of other groups in the structure, many topological indices are a function of the entire graph and thus reflect the fundamental nature of the molecular structure. One important class of topological indices called connectivity indices considers how a graph is connected (Randic, 1975). The QSPR method correlates properties with molecular descriptors derived from molecular characteristics. Technically, molecular descriptors for QSPR models can be derived from GC groups, topological indices, and other defined descriptors (Austin et al., 2016). The major problem is that interpretation of the developed descriptors could be weak, provisional, or utterly lacking because some of the descriptors are chemically meaningless (Todeschini and Consonni, 2000). Recently, machine learning methods, especially graph neural networks (GNN), have been used to predict the structure-property relationship of molecules (Coley et al., 2019). GNN utilises a graph representation of molecules where atoms correspond to nodes and bonds to edges containing information about the molecular structure. However, the convolution process of GNN may result in the loss of information, making it difficult to trace the substructure's contribution to the property.

In this work, we applied the novel connectivity matrix-based molecular substructural descriptor generation method to normal boiling point property prediction. Different to connectivity indices which are usually derived from the adjacency matrix of a molecule, the connectivity matrix is a variant of the adjacency matrix. It includes the atomic number and actual bond type information other than the adjacency information. An extraction strategy is proposed to extract a plethora of submatrices (also called molecular structural fragments), each representing the environment of an atom/bond automatically and systematically from this connectivity matrix, rather than using predefined groups like GC methods. This extraction does not cause a loss of molecular information compared to methods with convolution steps. In addition, we increased the number of hidden layers of the neural network to increase model accuracy (Pan et al., 2023).

## 2. Methodology

### 2.1. Datasets

The normal boiling point (NBP) dataset used in this work is from the literature (Alshehri et al., 2021). As far as we know, it is the largest NBP database with molecular structural information represented in SMILES format and property values.

## 2.2. Generation of features

Our recent article describes the detailed method (Pan et al., 2023). The general idea is that the environment of an atom/bond can be represented by extracting entries in the matrix considering the connectivity information. All molecules stored in SMILES format are canonicalised before structural information extraction. Rdkit package and in-house Python scripts were implemented for the feature generation. Figure 1 shows the structure, and reference number of each atom in acetic acid, each atom in the molecule is labelled a number, which is used as row/column indexes when generating the connectivity matrix. In Table 1, a square matrix in which the number of rows equal to the number of atoms is generated. Diagonal entries store the atomic number of each corresponding atom, while off-diagonal entries store the connectivity information.

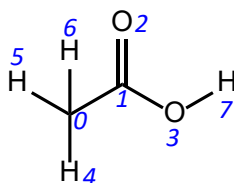


Figure 1 Structure of acetic acid and reference number of each atom

The environment of each substructure is represented by the extracted submatrix. For example, atom 1 can be represented by extracting intersections of rows 0, 1, 2, 3 and columns 0, 1, 2, 3 from the connectivity matrix. Eigenvalues and determinants of such extracted submatrices are calculated, which were used as feature names to classify the structural fragments and count the number of each feature (number of same fragments). Features were then characterised using statistical methods such as Pearson's correlation method and principal component analysis, as shown in Figure 2d and 2e. Although the determinants and eigenvalues of submatrices are chemically meaningless, substructural information can be recovered using the stored matrix, because atoms and connecting information were fully stored.

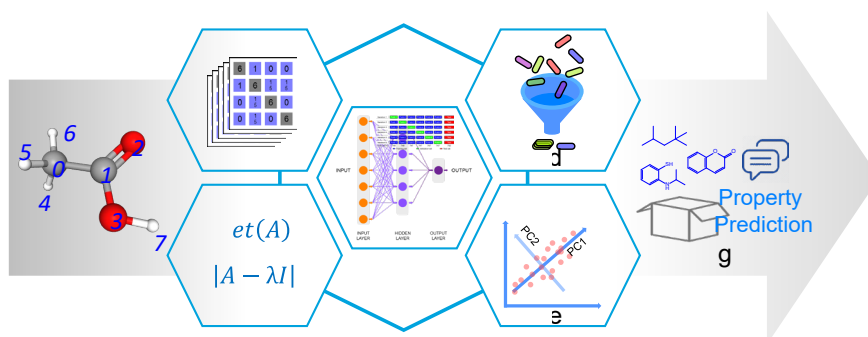


Figure 2 Flow chart of deep learning structure-property model development based on connectivity matrix

## 2.3. Development of structure-property models using DNN

Based on the previous work, we upgraded our neural network to a deep neural network (DNN) with multiple hidden layers and multiple nodes in each hidden layer. The dropout rate in each hidden layer is set as 25%. The number of neurons in the first hidden layer is

optimised, while the number of neurons in the second layer is fixed to 80. Other features of the model remain the same in as our previous work, such as the early stopping and checkpoints step of callback methods are adopted to avoid overfitting, cross-validation to evaluate model performance over different subsets of the training dataset.

Table 1 Connectivity matrix representation of acetic acid

Index	0	1	2	3	4	5	6	7
0	6	1	0	0	1	1	1	0
1	1	6	2	1	0	0	0	0
2	0	2	8	0	0	0	0	0
3	0	1	0	8	0	0	0	1
4	1	0	0	0	1	0	0	0
5	1	0	0	0	0	1	0	0
6	1	0	0	0	0	0	1	0
7	0	0	0	1	0	0	0	1

### 3. Results and discussion

Root mean squared error (RMSE) and mean absolute percentage error (MAPE) are used to evaluate the performance of the constructed model. The investigated number of the first hidden layer units varies from 10% to 130% of the number of input features with a 20 % step. The effects of hidden layer units of the DNN model on RMSE and MAPE are presented in Table 2. The DNN model with 209 neurons in the first hidden layer performs best with a minimum RMSE of 31.36 and MAPE of 4.20%. However, there is no noticeable trend regarding the number of neurons with the model performance.

Table 2 Effects of hidden layer units of the DNN model on RMSE and MAPE

Neurons	RMSE			MAPE (%)		
	training set	validation set	test set	training set	validation set	test set
41	29.66	32.56	33.73	3.63	4.45	4.83
125	24.47	27.53	31.83	2.99	3.94	4.43
209	22.69	26.29	31.36	2.76	3.80	4.20
293	23.07	26.71	32.55	2.86	3.82	4.29
377	24.12	27.49	31.88	2.94	3.92	4.25
460	23.48	26.99	31.49	2.88	3.86	4.26
544	24.98	27.83	31.93	3.12	4.00	4.43

The results of the DNN model over the previous artificial neural network (ANN) model are presented in Table 3. The average RMSE of the 6-fold cross-validation DNN model on the test set is 31.36, which is reduced by 6.50% compared to 33.54 of the ANN model. The MAPE (%) of the DNN model is reduced by 9.48%, from 4.64% to 4.20%.

Table 3 Comparison of the DNN model with the ANN model

Model	RMSE			MAPE (%)		
	training set	validation set	test set	training set	validation set	test set
ANN	30.02	39.76	33.54	3.28	4.23	4.64
DNN	22.69	36.33	31.36	2.76	3.80	4.20

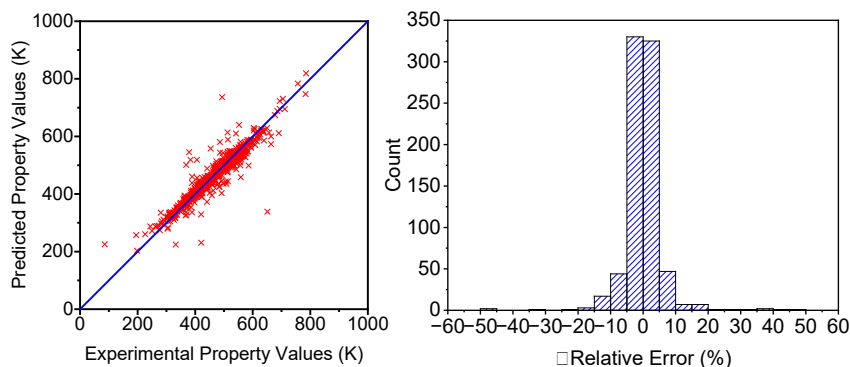
The detailed RMSE and MAPE values for the DNN model generated using the 6-fold cross-validation method are presented in Table 4. The RMSEs on the test set stayed within a relatively narrow band, ranging from 28.52 to 36.71.

The parity plot (i.e., predicted values vs. experimental values) is provided in Figure 3a, where red dots represent the predicted boiling point over the practical property value in the test set (at iteration 6). It can be observed that most of the 792 data points in the test set distribute along the diagonal line, indicating that the predicted values are close to the experimental values. Figure 3b shows the distribution of relative errors (REs) between the predicted and the experimental property values at iteration 6. The number of data points with a predicted RE between (-5 %, 0) and (0, 5 %) are 330 and 325, which are higher than the ANN model (265, 217). Our previous ANN model outperformed the popular GC method, and the results suggest that the DNN model has a better prediction accuracy than the ANN model.

Table 4 RMSE and MAPE values of the DNN model using the k -fold method

Iteration	RMSE			MAPE (%)		
	training set	validation set	test set	training set	validation set	test set
1	16.63	24.28	28.56	2.17	3.29	3.75
2	18.42	70.34	31.67	2.63	3.35	4.01
3	41.07	35.00	36.71	4.39	5.17	5.35
4	14.90	28.19	30.33	2.37	3.67	4.14
5	31.29	32.68	32.35	2.98	3.72	4.31
6	13.80	27.50	28.52	2.03	3.62	3.68

Figure 3 Parity plot from the test set (iteration 5)



#### 4. Conclusion

Structure information representation and mathematical modelling of structure-property relationships are critical factors in pure component properties prediction. In this work, based on the molecular structural features of molecules created using the novel connectivity matrix-based framework, we developed deep neural network (DNN) property prediction models. The effect of the number of hidden layer neurons was investigated. Compared to artificial neural networks, the increase in the number of hidden layers improved the accuracy of results. However, there was no apparent trend between the number of neurons and the model performance. The average RMSE of the DNN model is reduced by 6.50% compared to the ANN model (31.36 vs 33.54), while the MAPE (%) of the DNN model is reduced by 9.48% (4.20% vs 4.64%).

## Acknowledgements

The authors gratefully acknowledge the financial support from China Scholarship Council (CSC) (No. 202006440006).

## References

- A.S. Alshehri, A.K. Tula, F. You, R. Gani, 2021. Next generation pure component property estimation models: With and without machine learning techniques. *AIChE Journal*, e17469.
- N.D. Austin, N.V. Sahinidis, D.W. Trahan, 2016. Computer-aided molecular design: An introduction and review of tools, applications, and solution techniques. *Chemical Engineering Research and Design* 116, 2-26.
- Connor W. Coley, W. Jin, L. Rogers, T.F. Jamison, T.S. Jaakkola, W.H. Green, R. Barzilay, K.F. Jensen, 2019. A graph-convolutional neural network model for the prediction of chemical reactivity. *Chemical Science* 10, 370-377.
- R. Gani, P.M. Harper, M. Hostrup, 2005. Automatic creation of missing groups through connectivity index for pure-component property prediction. *Industrial & Engineering Chemistry Research* 44, 7262-7269.
- A.S. Hukkerikar, B. Sarup, A. Ten Kate, J. Abildskov, G. Sin, R. Gani, 2012. Group-contribution+ (GC+) based estimation of properties of pure components: Improved property estimation and uncertainty analysis. *Fluid Phase Equilibria* 321, 25-43.
- A.R. Katritzky, M. Karelson, V.S. Lobanov, 1997. QSPR as a means of predicting and understanding chemical and physical properties in terms of structure. *Pure and Applied Chemistry* 69, 245-248.
- H. Korsten, 1997. Characterization of hydrocarbon systems by DBE concept. *AIChE Journal* 43, 1559-1568.
- Q. Pan, X. Fan, J. Li, 2023. Automatic creation of molecular substructures for accurate estimation of pure component properties using connectivity matrices. *Chemical Engineering Science* 265, 118214.
- M. Randić, 1975. Characterization of molecular branching. *Journal of the American Chemical Society* 97, 6609-6615.
- M. Randić, 2008. On history of the Randić index and emerging hostility toward chemical graph theory. *MATCH Commun. Math. Comput. Chem* 59.
- R. Todeschini, V. Consonni, 2000. *Handbook of molecular descriptors*. Wiley-VCH, Weinheim ;
- H. Wen, Y. Su, Z. Wang, S. Jin, J. Ren, W. Shen, M. Eden, 2021. A systematic modeling methodology of deep neural network - based structure - property relationship for rapid and reliable prediction on flashpoints. *AIChE Journal*, e17402.
- J. Yoshida, H. Kim, A. Nagaki, 2011. Green and sustainable chemical synthesis using flow microreactors. *Chemsuschem* 4, 331-340.

# Novel biorefinery ideas for conversion of biomass to biofuel

Niloufar Ghavami,<sup>a</sup> Karhan Özdenkçi,<sup>a</sup> Simeone Chianese,<sup>b</sup> Dino Musmarra,<sup>b</sup>  
Cataldo De Blasio<sup>a</sup>

<sup>a</sup> Faculty of Science and Engineering, Åbo Akademi University, Rantakatu 2, 65100  
Vaasa, Finland

<sup>b</sup> Department of Engineering, University of Campania “Luigi Vanvitelli”, Via Roma 29,  
81031 Aversa, Italy

## Abstract

The current economic model produces large amounts of waste causing environmental problems. In contrast, circular economy focuses on the efficient utilization of resources, increasing materials' lifecycle, recycling, and decreasing environmental damage. Biorefineries are the foundation for accomplishing sustainable development goals and transitioning to the circular economy. Biomass wastes are sustainable feedstock for biorefineries and are increasing due to population growth and urbanization.

Thermochemical conversion technologies specifically supercritical water gasification, hydrothermal liquefaction, and hydrothermal carbonization are viable technologies for wet biomass feedstock. Therefore, the objective of this article is to investigate a possible biorefinery integration scenario with the mentioned processes and the feasibility of the integrated processes. A study on biorefinery processes could facilitate the transition to a circular economy and sustainability and provide a holistic view for decision-makers.

**Keywords:** Circular economy, Biomass waste, Biorefinery, Thermochemical conversion methods.

## 1. Introduction

Growing waste production due to urbanization and population growth results in several environmental risks. For instance, sewage sludge (SS) is one of the major waste streams. The generation of SS in the 27 European Union member countries, (EU-27), varies between 0.1 to 30.8 kg per population equivalent and year (Havukainen et al., 2022). SS contains organics content, inorganic compounds, various microorganisms and pathogens, nitrogen and phosphorous and toxic pollutants (Tiwari et al., 2022). Therefore, energy recovery and the disposal process of SS are quite challenging.

Different methods are applied for processing SS including supercritical water gasification (SCWG) at temperatures higher than 400 °C, hydrothermal liquefaction (HTL) at 250–400 °C, and hydrothermal carbonization (HTC) at 180–280 °C (Park et al., 2019). The superiority of these processes is the removal of costly drying pre-treatment (Özdenkçi et al., 2019). Water's physical characteristics have an important role in hydrothermal processes. The physical properties change significantly in sub-critical and critical regions: water behaves like a non-polar solvent and a catalyst in hydrothermal processes (Ghavami et al., 2021). In HTL, the organic matter is liquefied to biocrude oil and other by-products as aqueous phase (AP), bio-char, and gaseous products (Nallasivam et al., 2022). HTL process consists of three main steps: depolymerization, decomposition and recombination (Gollakota et al., 2018). HTC is a proper approach to converting wet biomass to a

carbonized solid fuel (hydrochar) with a moderately high heating value and nutrient-rich aqueous phase (Son Le et al., 2022; Xu et al., 2022). The HTC reactions include hydrolysis, dehydration, decarboxylation, aromatization, condensation, and polymerization (Ghavami et al., 2022). It is worth mentioning that the aqueous phase from HTL and HTC contains a huge amount of organic and inorganic compounds that need further processing (Merzari et al., 2019; Watson et al., 2020). SCWG is one of the effective technologies for converting biomass to hydrogen-rich gas products. Modell used it for the first time in the 1970s (Reddy et al., 2014; Wei et al., 2021). SCWG resulted in char formation being inevitable, especially with feedstocks involving unhydrolyzed large molecules, causing reactor plugging (Ghavami et al., 2021). Therefore, in order to reduce the char formation, the aqueous phase from HTL and HTC is processed in SCWG because of involving lighter and dissolved organics (Ong et al., 2020).

Integration of mentioned processes as a biorefinery concept is considered a sustainable approach to move towards a circular bioeconomy, carbon neutrality, waste management, and biofuel production (Culaba et al., 2022). Biorefineries are integrations of various processes producing multiple value-added products, in contrast with the traditional approach to a single product. Moreover, multiple products increase the rate of return on investment and protect against market uncertainties and fluctuations (Giwa et al., 2023). Designing a biorefinery requires a comprehensive understanding of integrated processes. In our previous article, we studied the integration of the HTC process into a biogas plant (Ghavami et al., 2022). Therefore, the objective of this study is to investigate the integration of hydrothermal processes with a biogas plant.

## 2. Process description and feedstock selection

Figure 1 depicts a complete view of the proposed biorefinery. SS feedstock is divided between anaerobic digestion (AD) and HTL unit operations, and the digestate from the AD process is directed to the HTC process. The aqueous phase of HTC and HTL processes enters the SCWG process for syngas production. The HTC process of SS digestate is at 200 °C and autogenous pressure with 30% solid load. The HTL process is simulated under the condition of 350 °C and autogenous pressure with 19% solid load. The aqueous phase (AP) from HTL and HTC is gasified in 600 °C and 250 bar conditions.

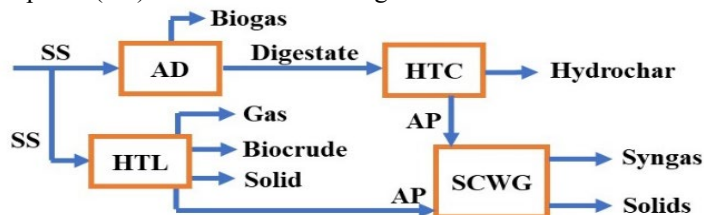


Figure 1. Flow diagram of the proposed biorefinery

Table 1 reports the dry basis compositions and the heating values (HHV) of the feedstocks and products based on the experiments (Parmar and Ross, 2019; Rahman et al., 2021).

Table 1. Analysis of feedstocks and products (Parmar and Ross, 2019; Rahman et al., 2021)

Sample	HHV (MJ/kg)	C	H	N	S	O	Ash	Yield
SS	14.1	33.1	5.5	5	0.7	25.9	29.8	-
SS digestate	14.9	28.7	3.1	3.4	1.5	16.4	46.9	-
Biocrude	28.3	53.3	6.01	3.9	1.1	29.9	4.7	37.1
Hydrochar	15.4	34.5	4.3	2.9	1.2	12.8	44.3	78

HTL, HTC, and SCWG processes are simulated in Aspen Plus V11.1 according to the following assumptions. The physical and chemical properties of non-conventional solids (NC solids) are estimated by 'HCOALGEN' and 'DCOALIGT'. Heating values for non-

conventional components are calculated with a unified correlation valid for a wide range of substances (Channiwala and Parikh, 2002). Gas phase in HTL contains 85.3 mol% CO<sub>2</sub>, 0.9 mol% H<sub>2</sub>, 6.4 mol% CH<sub>4</sub>, 4.4 mol% CO, and 2.9 mol% N<sub>2</sub>. The gas phase in HTC contains 90 mol% CO<sub>2</sub>, 1 mol% H<sub>2</sub>, 3 mol% CH<sub>4</sub>, and 6 mol% CO.

Prior to the simulation, it is needed to define components, thermodynamic method, and capacity. The component list is a combination of conventional and non-conventional substances including water, hydrogen, oxygen, carbon dioxide, carbon monoxide, methane, and nitrogen as conventional besides SS, digestate, hydrochar, biocrude oil, dissolved organics, and ash as non-conventional. The thermodynamic method is selected as PSRK for HTC and HTL sections because of accuracy at close-to-critical temperatures. Meanwhile, Peng-Robinson is used for the reactor and separation units in the SCWG section when converting non-conventional components to conventional ones and simulating the phase separation. In addition, mass balance is required to define yields on a non-inert basis for the RYIELD reactors in HTC and HTL processes. Yield reactors are used due to the non-conventional substances in the feedstocks. The HTC yields are calculated as 0.15, 0.81, 0.04, 0.005, 2.64e-6, 4.22e-5, and 0.00015 for hydrochar, water, organics, CO<sub>2</sub>, H<sub>2</sub>, CH<sub>4</sub>, and CO, respectively. Similarly, the HTL yields are 0.05, 0.02, 0.83, 0.02, 0.06, 3.14e-5, 0.0016, 0.0019, and 0.0013 for biocrude oil, solid, water, organics, CO<sub>2</sub>, H<sub>2</sub>, CH<sub>4</sub>, CO, and N<sub>2</sub>, respectively. The SCWG process is simulated as a yield reactor decomposing the non-conventional components into stable elemental components, separating ash, and a RGIBBS as an equilibrium reactor minimizing Gibbs free energy. Figure 2 depicts the process simulation of the proposed biorefinery.

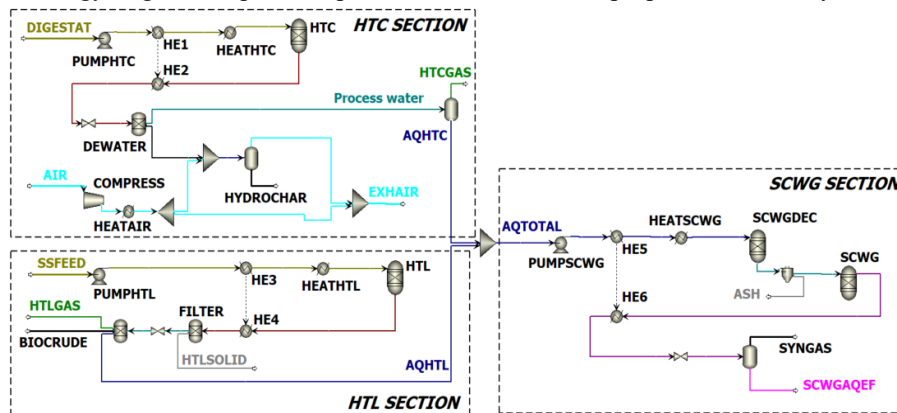


Figure 2. Process simulation of the proposed biorefinery

In the HTC section, the digestate feedstock (DIGESTAT) is pumped to 45 bars and is heated to 200 °C to reach the conditions in the HTC reactor. The stream exchanges heat with the reactor downstream by units 'HE1' and 'HE2' up to 180 °C and then reach the temperature of 196 °C by the external source in the HEATHTC heater. The product mixture goes through expansion and separation after heat exchange with the feedstock. Hydrochar is separated from the process water in the dewatering section and is dried down to 20% moisture by compressed air at 110 °C and 1.2 bars as described more in detail by Ghavami et al. (Ghavami et al., 2022). The HTC process water is conducted to a flash separation for removing gas products and then mixed with the HTL process water. In the HTL section, SS feedstock (SSFEED) is pumped to 170 bars and is heated to 350 °C to be prepared for the reaction condition, and same as the HTC part the reactor product exchange heat with the feedstock by units 'HE3' and 'HE4', then it reaches reaction condition by HEATHTL. The product mixture is passed through a filter to remove the



HTLSOLID after heat exchange with the SS feedstock. HTL products (gas, biocrude oil, and aqueous phase) are separated via a drum after expansion. The mixture of aqueous effluents (AQTOTAL) is pumped and heated with the same procedure as HTC and HTL to 250 bars and 720 °C. After the heat exchange with the aqueous mixture (units ‘HE5’ and ‘HE6’) and expansion, the syngas is separated in a flash unit.

### 3. Result and discussion

#### 3.1. Mass balance

Different methodologies are applied for mass balance calculations due to using various feedstock and data availability for each process. For the HTC and HTL sections, mass balance calculations were conducted by hydrochar, biocrude oil, liquid, and gas yields as well as proximate and ultimate analysis. Moreover, oxygen and hydrogen in organics are calculated by experimental chemical oxygen demand (COD) and biochemical methane potential (BMP). Ash content in biocrude oil is assumed as 4.72% based on error bars of the dry-ash-free elemental composition. In the SCWG section, mass balance calculations in the decomposition reactor were based on elemental compositions. Figure 3 depicts mass balances for HTC, HTL, and SCWG reactors.

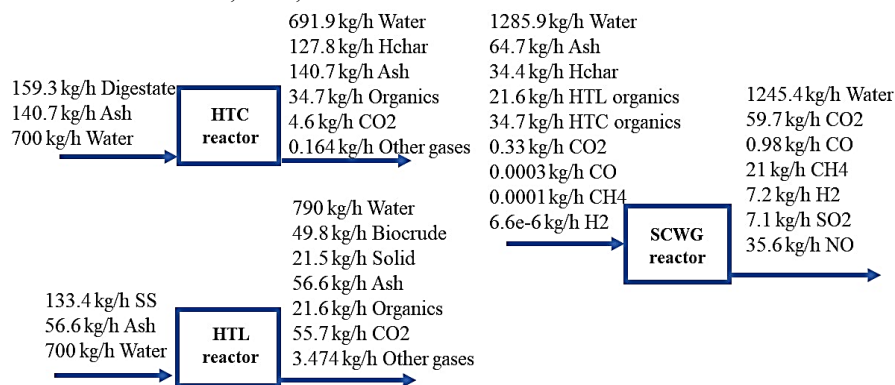


Figure 3. Mass balance in HTC, HTL, and SCWG reactors

#### 3.2. Energy balance

Table 2 reports the power and heat requirements for the proposed biorefinery as well as energy obtained from HTL solid and syngas.

Table 2. Power and heat requirement for the biorefinery

Electricity requirement (kW)		Heat requirement (kW)	
PUMPHTC	1.7	HEATHTC	16.1
PUMPHTL	7.8	HEATHTL	50.2
PUMPSCWG	15.1	HEATAIR	162.1
COMPRESS	56.4	DEWATER	5.9
0.25 x HEATSCWG	41.9	0.75 x HEATSCWG	125.6
<b>Total</b>	<b>122.9</b>	<b>Total</b>	<b>359.9</b>
Electricity outcome (kW)		Heat outcome (kW)	
Syngas	152.8	Syngas	458.3
HTL solid	38.5	HTL solid	115.5
<b>Total</b>	<b>191.3</b>	<b>Total</b>	<b>573.8</b>

The electricity requirement results from the pumps and the compressor while heat requirement results from the external heat supplied to air and HTC and HTL reactor inlets. Since the SCWG inlet is at a high temperature, it is compromising the combined heat and power (CHP) production. Therefore, the duty of the “HEATSCWG” unit is considered as

25 % electricity and 75 % heat requirement. Similarly, the CHP obtained from the combustion of syngas and HTL solid is also assumed as 25 % electricity and 75 % heat. Biocrude oil can also be used as a fuel in the process or be upgraded to a transportation fuel. Meanwhile, the hydrochar can be used as a fertilizer, benefiting from nutrients and minerals. In addition, the remaining fertilizer elements can be recovered from the SCWG process as the solid outlet.

### 3.3. Feasibility and future aspects

The proposed biorefinery process is applicable to biogas plants with various feedstocks. Hydrothermal processes are suitable for converting wet biomass. In addition, the nutrients and minerals can be recovered on the solid downstream regarding the usage as a fertilizer. The proposed process leads to the generation of 68.4 kW of electricity and 213.9 kW of heat as the net amounts.

The proposed process has high potential in terms of feasibility since the combination improves economic performance and overcomes the operational issues of each conversion method. HTC is industrially applied for sewage sludge (Buttmann, 2017). Meanwhile, HTC alone requires energy, e.g. 58 kW electricity and 61 kW heat consumption to process 1000 kg/h of agricultural residue digestate with 30 % moisture (Ghavami et al., 2022). Similarly, HTL alone also requires energy, despite being stated as a feasible process for large-scale (Seiple et al., 2020). Moreover, SCWG was stated to be feasible for a more challenging feedstock (black liquor) provided that the operational issues are addressed well (Özdenkçi et al., 2019). The proposed process provides net CHP production via SCWG of aqueous effluents while producing hydrochar and bio-oil as well as addressing the aqueous discharge issues and char formation in SCWG.

## 4. Conclusion

The scheme of a biorefinery is proposed in this article and integrated processes including HTC, HTL, and SCWG are simulated in Aspen Plus. Moreover, the process is able to produce CHP, especially via syngas after covering the energy requirements as well as the biocrude oil product and solids to be used as fertilizers.

## 5. Acknowledgement

We gratefully acknowledge the financial support of the Kaute Foundation, the Finnish Foundation for Technology Promotion (Tekniikan edistämissäätiö), The University Foundation in Ostrobothnia (Höskolestiftelsen i Österbotten), and the Regional Council of Ostrobothnia for their support on the EU-ERUF project named “Implementation of a hydrothermal carbonization process for the disposal sludge and wet organic streams”.

## 6. References

- Buttmann, M., 2017. Industrial scale plant for sewage sludge treatment by hydrothermal carbonization in Jinin/China and phosphate recovery by Terranova ultra HTC process. Presented at the European Biosolids and Organic Resources Conference.
- Channiwala, S.A., Parikh, P.P., 2002. A unified correlation for estimating HHV of solid, liquid and gaseous fuels. *Fuel* 81, 1051–1063. [https://doi.org/10.1016/S0016-2361\(01\)00131-4](https://doi.org/10.1016/S0016-2361(01)00131-4)
- Culaba, A.B., Mayol, A.P., San Juan, J.L.G., Ubando, A.T., Bandala, A.A., Concepcion II, R.S., Alipio, M., Chen, W.-H., Show, P.L., Chang, J.-S., 2022. Design of biorefineries towards carbon neutrality: A critical review. *Bioresour Technol* 128256. <https://doi.org/10.1016/j.biortech.2022.128256>
- Ghavami, N., Özdenkçi, K., Chianese, S., Musmarra, D., De Blasio, C., 2022. Process simulation of hydrothermal carbonization of digestate from energetic perspectives in Aspen Plus. *Energy Conversion and Management* 270, 116215. <https://doi.org/10.1016/j.enconman.2022.116215>
- Ghavami, N., Özdenkçi, K., Salierno, G., Björklund-Sänkiäho, M., De Blasio, C., 2021. Analysis of operational issues in hydrothermal liquefaction and supercritical water gasification processes: a review. *Biomass Conv. Bioref.* <https://doi.org/10.1007/s13399-021-02176-4>

- Giwa, T., Akbari, M., Kumar, A., 2023. Techno-economic assessment of an integrated biorefinery producing bio-oil, ethanol, and hydrogen. *Fuel* 332, 126022. <https://doi.org/10.1016/j.fuel.2022.126022>
- Gollakota, A.R.K., Kishore, N., Gu, S., 2018. A review on hydrothermal liquefaction of biomass. *Renewable and Sustainable Energy Reviews* 81, 1378–1392. <https://doi.org/10.1016/j.rser.2017.05.178>
- Havukainen, J., Saud, A., Astrup, T.F., Peltola, P., Horttanainen, M., 2022. Environmental performance of dewatered sewage sludge digestate utilization based on life cycle assessment. *Waste Management* 137, 210–221. <https://doi.org/10.1016/j.wasman.2021.11.005>
- Merzari, F., Langone, M., Andreottola, G., Fiori, L., 2019. Methane production from process water of sewage sludge hydrothermal carbonization. A review. *Valorising sludge through hydrothermal carbonization. Critical Reviews in Environmental Science and Technology* 49, 947–988. <https://doi.org/10.1080/10643389.2018.1561104>
- Nallasivam, J., Prashanth, P.F., Vinu, R., 2022. Chapter 4 - Hydrothermal liquefaction of biomass for the generation of value-added products, in: Varjani, S., Pandey, A., Taherzadeh, M.J., Ngo, H.H., Tyagi, R.D. (Eds.), *Biomass, Biofuels, Biochemicals*. Elsevier, pp. 65–107. <https://doi.org/10.1016/B978-0-323-88511-9.00018-5>
- Ong, B.H.Y., Walmsley, T.G., Atkins, M.J., Walmsley, M.R.W., 2020. A Kraft Mill-Integrated Hydrothermal Liquefaction Process for Liquid Fuel Co-Production. *Processes* 8, 1216. <https://doi.org/10.3390/pr8101216>
- Özdenkçi, K., De Blasio, C., Sarwar, G., Melin, K., Koskinen, J., Alopaeus, V., 2019. Techno-economic feasibility of supercritical water gasification of black liquor. *Energy* 189, 116284. <https://doi.org/10.1016/j.energy.2019.116284>
- Park, M., Kim, N., Lee, S., Yeon, S., Seo, J.H., Park, D., 2019. A study of solubilization of sewage sludge by hydrothermal treatment. *Journal of Environmental Management* 250, 109490. <https://doi.org/10.1016/j.jenvman.2019.109490>
- Parmar, K.R., Ross, A.B., 2019. Integration of Hydrothermal Carbonisation with Anaerobic Digestion; Opportunities for Valorisation of Digestate. *Energies* 12, 1586. <https://doi.org/10.3390/en12091586>
- Rahman, T., Jahromi, H., Roy, P., Adhikari, S., Hassani, E., Oh, T.-S., 2021. Hydrothermal liquefaction of municipal sewage sludge: Effect of red mud catalyst in ethylene and inert ambiances. *Energy Conversion and Management* 245, 114615. <https://doi.org/10.1016/j.enconman.2021.114615>
- Reddy, S.N., Nanda, S., Dalai, A.K., Kozinski, J.A., 2014. Supercritical water gasification of biomass for hydrogen production. *International Journal of Hydrogen Energy* 39, 6912–6926. <https://doi.org/10.1016/j.ijhydene.2014.02.125>
- Seiple, T.E., Skaggs, R.L., Fillmore, L., Coleman, A.M., 2020. Municipal wastewater sludge as a renewable, cost-effective feedstock for transportation biofuels using hydrothermal liquefaction. *Journal of Environmental Management* 270, 110852. <https://doi.org/10.1016/j.jenvman.2020.110852>
- Son Le, H., Chen, W.-H., Forruque Ahmed, S., Said, Z., Rafa, N., Tuan Le, A., Ağbulut, Ü., Veza, I., Phuong Nguyen, X., Quang Duong, X., Huang, Z., Hoang, A.T., 2022. Hydrothermal carbonization of food waste as sustainable energy conversion path. *Bioresource Technology* 363, 127958. <https://doi.org/10.1016/j.biortech.2022.127958>
- Tiwari, N., Garua, B., Bansal, M., Sharma, J.G., 2022. 20 - Biorefinery and waste management by co-digestion of sewage sludge with organic wastes, in: Shah, M.P., Rodriguez-Couto, S., Shah, N., Banerjee, R. (Eds.), *Development in Waste Water Treatment Research and Processes*. Elsevier, pp. 365–386. <https://doi.org/10.1016/B978-0-323-85584-6.00021-2>
- Watson, J., Wang, T., Si, B., Chen, W.-T., Aierzhati, A., Zhang, Y., 2020. Valorization of hydrothermal liquefaction aqueous phase: pathways towards commercial viability. *Progress in Energy and Combustion Science* 77, 100819. <https://doi.org/10.1016/j.pecs.2019.100819>
- Wei, N., Xu, D., Hao, B., Guo, S., Guo, Y., Wang, S., 2021. Chemical reactions of organic compounds in supercritical water gasification and oxidation. *Water Research* 190, 116634. <https://doi.org/10.1016/j.watres.2020.116634>
- Xu, Z.-X., Ma, X.-Q., Zhou, J., Duan, P.-G., Zhou, W.-Y., Ahmad, A., Luque, R., 2022. The influence of key reactions during hydrothermal carbonization of sewage sludge on aqueous phase properties: A review. *Journal of Analytical and Applied Pyrolysis* 167, 105678. <https://doi.org/10.1016/j.jaap.2022.105678>

# Waste integration for sustainable operation of WWTP's

Elena C. Blanco,<sup>a</sup> Mariano Martín,<sup>a</sup> Pastora Vega,<sup>b</sup>

<sup>a</sup> *Department of Chemical Engineering, University of Salamanca, 37008 Salamanca, Spain*

<sup>b</sup> *Department of Computer Science, University of Salamanca, 37008 Salamanca, Spain*  
*mariano.m3@usal.es*

## Abstract

The water-energy nexus is an important topic nowadays. Within it, the wastewater treatment plants (WWTPs) energy consumption is becoming an issue. WWTPs typically consume more energy than the one that can be obtained from the biogas produced from sludge anaerobic digestion. In this work, a process-level analysis is presented to study the feasibility of integrating both wastewater and municipal solid waste (MSW) treatment to achieve an energetically self-sustainable operation of a WWTP. The influence of the climate of different regions across the Iberian Peninsula on the energy requirements has also been evaluated. Mesophilic and thermophilic digestion is compared in Salamanca as a case base. Next, the optimal digestion temperature of 30 °C is found. Moreover, in all cities it is necessary for MSW to provide between 37-40% of the facility energy consumption. From an economic point of view, an investment between 0.09-0.16 €/kg of sewage sludge is obtained for the integrated process. Therefore, this techno-economic assessment demonstrates the feasibility of integrating these two treatments for a fully self-sufficient and sustainable process.

**Keywords:** energy, wastewater, municipal solid waste.

## 1. Introduction

Wastewater treatment plants are crucial facilities for reducing the potential environmental impact of wastewater discharges into receiving water bodies. However, they consume large amounts of energy. Until recently, energy was relatively affordable internationally, and many wastewater treatment facilities were not designed and operated with the goal of minimizing energy consumption, only to satisfy certain effluent quality requirements. Nowadays, rapid population growth and urbanization have led to an increase in wastewater production, so the number of WWTPs has also increased and effluent quality requirements have also become more demanding. This fact has resulted in a large energy consumption devoted to the operation of WWTPs, and the water-energy nexus has started to attract increasing attention from both environmental and economic points of view. On the one hand, WWTPs can represent 15–20% of the total energy consumption of municipal public structures and facilities, with energy also constituting anywhere 25-40% of the total operating costs in a conventional WWTP (Yang, et al., 2020). On the other hand, the operation of WWTPs involves greenhouse emissions due to the need for non-renewable energy sources; moreover, they have grown in recent years due to the increase in the volume of waste treated and the implementation of new processes aimed at achieving higher effluent quality. This water-energy nexus is promoting a series of studies on the relationship between energy and water for sustainable development, in particular,

the potential for energy self-sufficient WWTPs has become an area of increasing research and innovation.

Anaerobic digestion (AD) is a common technology for sludge treatment at WWTPs and allows the production of biogas, which is the main energy source in a WWTP. The use of biogas for digester heating and electricity generation is a sustainable way to recover energy from WWTPs, as it could replace fossil fuels, with consequent sludge reduction. However, while the use of AD for sludge processing and biogas production is widely reported in the literature (Bin Khawer, et al., 2022), in most cases, the biogas cannot cover the energy requirements, thermal and electrical, of the plant. Even in developed countries, there is still a large gap for the energy self-sufficient WWTPs, so it is necessary to investigate and apply new trends in energy production (renewables) and new technologies to achieve it. An alternative to generate clean energy that has gained importance in recent years is its production from waste, known as waste-to-energy (WTE). MSW can be used to provide such energy helping manage this residue at the same time. Several waste-to-energy technologies, both conventional and non-conventional, have been described in recent literature (Munir, et al., 2021). Conventional waste treatment or disposal techniques include composting, anaerobic digestion, and landfilling, while incineration, pyrolysis, gasification, and hydrothermal processing are considered non-conventional waste-to-energy technologies. Among all of them, incineration is regarded as the most mature.

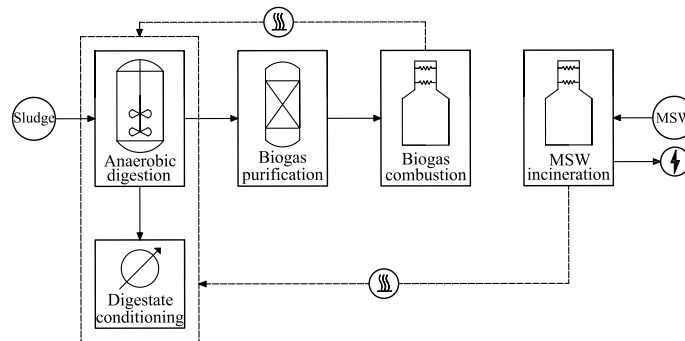
This work proposes the integration of sludge and MSW treatment for self-sufficient operation. The incineration of MSW produces steam to provide the thermal energy requirements in the AD process of the sewage sludge that cannot be covered by the biogas generated, and, also, to produce the total electricity requirements of the WWTP.

## 2. Process description and modeling

The integration of wastewater and MSW treatment combines two sections: sewage sludge treatment, that is the main thermal energy consumer within wastewater treatment, and the incineration of MSW, that is the provider of additional thermal energy and supplies power to the complete wastewater treatment, aiming at an energetically self-sufficient operation. The sewage sludge treatment is divided into four sections: biogas production, biogas purification, digestate conditioning and energy production (biogas combustion) (Figure 1). First, the sewage sludge is heated up and fed into a bioreactor where it is anaerobically digested to produce biogas, which contains methane, carbon dioxide, nitrogen, hydrogen sulfide, ammonia and moisture, and a decomposed substrate (digestate). The two main anaerobic digestion conditions are considered, mesophilic and thermophilic, which operate between 30-37°C and 50-55°C, respectively. Then, the digestate produced is conditioned in a reboiler, removing the excess of water and ammonia. This step is included so that the digestate can later be used as a fertilizer. The third step consists of the biogas purification by removing the CO<sub>2</sub> (and traces of NH<sub>3</sub>) and the H<sub>2</sub>S in a Pressure Swing Adsorption (PSA) system and in a fixed-bed reactor, respectively. Once the biogas is mainly methane, it is burned with air in a furnace to provide the thermal energy requirements of the process. They include heating up the combustion air and the sludge feed, and to produce steam to heat up the bioreactor and the heat exchanger of the sludge conditioning step.

The incineration of MSW is used to produce the additional thermal energy needed in the anaerobic digestion and the power requirements of the WWTP. A fraction of the MSW is burned in a furnace to produce the steam to heat the bioreactor and the heat exchanger of sludge conditioning, and the rest is used to produce steam to feed a Rankine cycle to

generate power. In the Rankine cycle, three different sections of the steam turbine are considered (high, medium and low-pressure turbines). The treatment of the combustion



gases generated in the incineration is not considered in this work.

Figure 1. Process diagram of the integrated process of sewage sludge and MSW treatment.

Both processes, stand-alone digestion in Salamanca and the integrated process in all cities, have been modelled using an equation-based approach. These models are formulated using mass and energy balances, thermodynamics, chemical and vapor-liquid equilibria, etc., to evaluate their performance and are based on the work of León and Martín (2016). The entire superstructure is formulated as a non-linear programming problem and is implemented in GAMS using CONOPT 3.0 as preferred solver in a multistart optimization approach. The objective function is the minimization of the external energy requirements, given by (Eq.(1)):

$$Z = (E_{\text{dig}} - E_{\text{biogas}}) - E_{\text{MSW}} \quad \text{Eq.1}$$

where  $E_{\text{dig}}$  is the total thermal energy required in the digestion process,  $E_{\text{biogas}}$  is the thermal energy supply by the generated biogas and  $E_{\text{MSW}}$  is the external thermal energy that needs to be provided by MSW.

This study is divided into two stages:

1) The evaluation of the operating conditions and the optimization of the operating temperature of the AD, minimizing in both cases the external supply of thermal energy. Firstly, mesophilic (35°C) and thermophilic (55°C) operation are compared. The variation of the water and air temperature that exists along the year in the city of Salamanca is considered, as a case base, because the thermal energy requirements depend on these temperatures. And, secondly, the operating temperature of the sludge digestion over a year has also been optimized, due to its effect on the amount of biogas produced and the energy consumption to heat up the sludge and dewatering step.

2) The study of the influence of the location on the thermal energy requirements of the optimized process. Different regions of the Iberian Peninsula have been evaluated considering the different climates that influence the air and water temperature each month, particularly: Oceanic in La Coruña (population: 244850, hab.eq.: 600000), Coastal Mediterranean in Barcelona (population: 1637000, hab.eq.: 1706250), Continental Mediterranean in Salamanca (population: 144000, hab.eq.: 260042) and Madrid (population: 3266000, hab.eq.: 1335000), and Oceanic Mediterranean in Sevilla (population: 689000, hab.eq.: 950000). The size of each WWTP is related to the amount of organic matter that is treated, considering that a person produces 26.9 kg of dry matter per year (Bianchini, et al., 2016). Also, it conditions its electricity consumption, considering that the average consumption of a WWTP in Spain is 5.6 kW/hab.eq. (IDEA, 2010).

### 3. Results

For the base case, the city of Salamanca, first, the operating conditions of the two different main types of AD have been studied. For the mesophilic and thermophilic digestion modes, the most energy intensive processes are the digestion itself and the digestate dewatering. The digester consumes 914.5 kW in both cases because it is considered as an isothermal process, so it only depends on the enthalpy of reaction. Nevertheless, the digestate dewatering consumes more energy in the mesophilic case, 467 kW vs. 358 kW, because the outlet temperature of the heat exchanger has always the maximum value (150°C), so the gradient temperature of the thermophilic case is lower. The other two duties are the sludge heating before digestion and the air heating before combustion and these values change over the months. The sum of these energy requirements goes from 103.5 kW in warm months to 175.4 kW in the coldest ones, in the case of mesophilic digestion, and, from 215.9 kW to 288.0 kW for thermophilic digestion. Obviously, they are higher as the digestion temperature increases. The amount of additional thermal energy is higher in the case of mesophilic digestion, although the digestion temperature is lower, an average of 39% vs. 37%. This fact is due to the significant contribution of digestate dewatering to the total thermal energy requirements of the process, that is higher in the mesophilic digestion (30% vs. 23%). In fact, the extra thermal energy without considering digestate conditioning is higher in the thermophilic case.

Then, digestion temperature has been optimized within the range of 30-55°C. The optimal operating temperature is 30°C, although it is shown that the additional thermal energy requirements is reduced with the increasing of digestion temperature, the problem tends to choose the lowest digestion temperature of the range to reduce the thermal energy involved in the process since the additional yield of biogas is not enough to mitigate the difference. This optimal temperature is the value used to study the influence of the climate, because it determines the value of the energy that heats up the sludge before digestion and the air before combustion. The change over the months and the different contribution on the total energy are shown in next section, where they are compared to other climates.

Once optimized the digestion temperature, the integrated process of wastewater a MSW treatment is modeled to make the process self-sufficient in terms of energy. The influence of the climate on the energy requirements has been studied for different cities (Salamanca, Madrid, Barcelona, La Coruña and Sevilla). The value of energy required in the digester and the one required to dehydrate the digestate per kilogram of sludge is the same in all cities, 680 kJ/kg, and 364 kJ/kg, respectively. In Figure 2, the energy requirements to heat sludge and air for each city over the months are shown. For all cases, June, July, and August present the lowest values because are the warmest months. Salamanca presents the maximum values, around 110 kJ/kg, because the winter in the Continental Mediterranean climate is the coldest one; nevertheless, in the warmest months, La Coruña exceeds the energy requirements of Salamanca, because the Oceanic climate tends to present mild summers, so water and air temperatures are lower, around 61-70 kJ/kg vs. 56-66 kJ/kg. The minimum values are obtained in Sevilla, from 37 kJ/kg to 86 kJ/kg, because the city is located in the most southern part of the Iberian Peninsula and the Continental Mediterranean climate presents the warmest temperatures, both air and water. The largest difference between the maximum and minimum value is shown in Madrid, 58%, because it presents the highest gradient temperature over the months and the minimum in La Coruña, 32%, where the thermal amplitude is lower due to ambient humidity; while, in Salamanca, this difference is 49%, in Sevilla, 56%, and, in Barcelona, 50%. This difference can also be seen in the shape of the curve presenting the data for the

extra energy (Figure 2), which is more pronounced in the case of Madrid. The amount of additional energy to be provided by the MSW slightly varies between the different cities and is between 37% and 40% of the total energy required for the digestion process for all of them. This percentage is similar because the main contributor to the total energy required in the process is the energy for the digestion and for the digestate conditioning, which represent around 91% of the total thermal energy. Its value increases as thermal energy requirements increase, so the same trend is observed as in the case of energy needs throughout the year between cities. Therefore, the extremes are Seville, with the lowest percentage of additional energy, and Salamanca, with the highest, except in summer, when it is surpassed by La Coruña. From January to April, between Madrid, Barcelona and La Coruña, the first one is the major consumer, and the same occurs from October to December; however, this trend changes between April and October, where the energy requirements of La Coruña increase, becoming even higher than those of Salamanca. The average amount of MSW that is fed to the integrated process is higher as the size of the WWTP increases because the energy and ranges from 0.26 kg/s in Salamanca to 1.67 kg/s in Barcelona. However, the amount of MSW per kilogram of sludge is almost the same for all cities, around 0.19 kg/kg, with a very slight difference between cities, being lower as the average of extra thermal energy decreases. Around 12% of the total MSW fed is devoted to the production of thermal energy, and the rest goes to the production of electricity. The databases show that there is enough MSW in each city to cover the energy requirements of the WWTP, considering that, a person produces 455 kg of MSW per year (Eurostat, 2022).

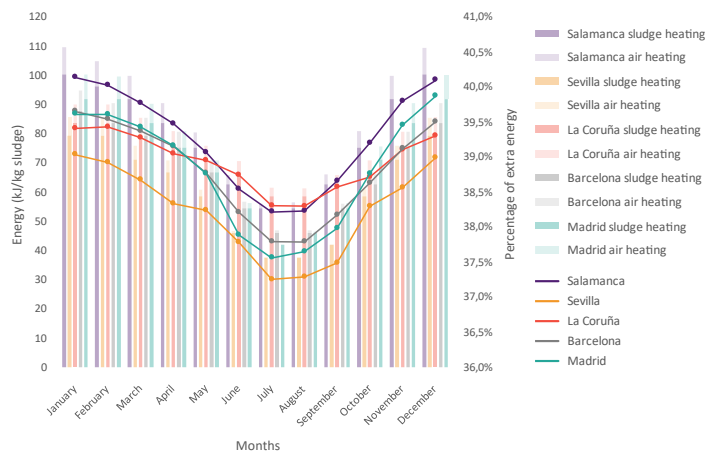


Figure 2. Energy contribution of sludge and air heating to digestion and percentage of extra energy needed in Salamanca, La Coruña, Sevilla, Madrid, and Barcelona per month.

An economic analysis of the integrated process has also been carried out for the different cities. Capital and production cost are obtained. Figure 3 shows the breakdown of the capital and the production cost for the different cities. The digester is the main item in the distribution of the capital cost in all cities, with around 51% of the total investment, followed by the steam turbines, with 25%. Regarding production costs, the main item is capital charges, with about 60% in all cities, followed by maintenance, with 20%. There is a slight difference in these percentages between cities, so they are higher as the size of the WWTP increases. Although, both capital and operating cost are higher in the larger WWTP, they benefit from economies of scale. The investment for the integrated process goes from 0.16 €/kg of sewage sludge in Salamanca, the smallest WWTP, to 0.09 €/kg in



Barcelona, which is the largest. The same is true for the investment per kilogram of MSW fed, which ranges from 0.86 €/kg in Salamanca to 0.46 €/kg in Barcelona.

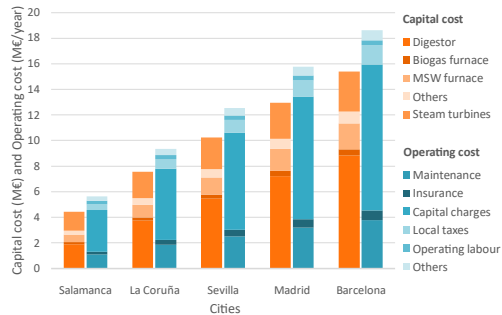


Figure 3. Capital and production cost for the different cities

#### 4. Conclusions

In this work, the energy requirements of the WWTP have been studied. The thermal energy requirements in sewage sludge treatment are a challenging problem because the anaerobic digestion of sludge does not produce enough biogas to cover them. Therefore, an integrated wastewater and municipal solid waste treatment process has been evaluated to make the WWT self-sufficient in terms of energy. Moreover, the influence of the climate in different region of the Iberian Peninsula has been studied. The results show that it is feasible to integrate the treatment processes because there is enough MSW to provide energy to the WWTP in all cities. An amount of MSW per kilogram of sludge is around 0.19 kg/kg, with a very slight difference between cities, being lower as the average of extra thermal energy decreases. The investment cost ranges from 0.09 €/kg to 0.16 €/kg of sewage sludge, depending on the size of the WWTP. This analysis allows the successful integration of these treatment processes with the aim of a fully self-sufficient and sustainable system.

#### Acknowledgments

The authors acknowledge MICINN Spain grant PID2019-105434RB-C31 and Programa Investigo JCYL to Elena C. Blanco.

#### References

- Bianchini, A., Bonfiglioli, L., Pellegrini, M. & Sacconi, C., 2016. Sewage sludge management in Europe: a critical analysis of data quality. *Int. J. of Environment and Waste Management*, 18(226).
- Bin Khawer, M. U., Raza Naqvi, S., Ali, I., Arshad, M., Juchelková, D., Waqas Anjum, M., Naqvi, M., 2022. Anaerobic digestion of sewage sludge for biogas & biohydrogen production: State-of-the-art trends and prospects. *Fuel*, 329(125416).
- Eurostat, 2022. Municipal solid waste generated 2005 and 2020.
- IDAE, 2010. Estudio de prospectiva. Consumo energético en el sector del agua.
- León, E. & Martín, M., 2016. Optimal production of power in a combined cycle from manure based biogas. *Energy Conversion and Management*, Volume 114, pp. 89-99.
- Munir, M., Mohaddespour, A., Nasr, A. & Carter, S., 2021. Municipal solid waste-to-energy processing for a circular economy in New Zealand. *Renewable and Sustainable Energy Reviews*, 145(111080).
- Yang, X., Wei, J., Ye, G., Zhao, Y., Li, Z., Qiu, G., Li, F., Wei, C., 2020. The correlations among wastewater internal energy, energy consumption and energy recovery/production potentials in wastewater treatment plant: An assessment of the energy balance. *Science of The Total Environment*, 714(136655).

# Multiscale design methodology for flow reactors supporting rapid prototyping

Tomoyui Taguchi,<sup>a</sup> Shigeru Kado,<sup>b</sup> Toshiyuki Watanabe,<sup>b</sup> Yoshiyuki Yamashita,<sup>c</sup>

<sup>a</sup>*Chiyoda Corporation, Pharmaceutical Engineering Department, 4-6-2, Minatomirai, Nishi-ku, Yokohama-shi, Kanagawa 220-8765, Japan*

<sup>b</sup>*Chiyoda Corporation, Research and Development Center, 3-13, Moriya-cho, Kanagawa-ku, Yokohama-shi, Kanagawa 221-0022, Japan*

<sup>c</sup>*Department of Chemical Engineering, Tokyo University of Agriculture and Technology, Tokyo 184-8588, JAPAN*

*Corresponding Author: taguchi.tomoyuki@chiyodacorp.com*

## Abstract

Continuous manufacturing has been incorporated in the pharmaceutical industry to achieve a beneficial production system for scale-up or scale-out, flexibility of demand response, and sustainability of material utilization (Yu, 2008). This paper proposes a rapid prototyping approach based on a mathematical model that integrates experiment and simulation. It is important to segregate the parameter identified for new manufacturing items from the parameter applied to common items based on the accumulation of experimental data. As an example case of rapid prototyping, a trickle bed reactor (TBR) as a multiphase heterogeneous catalytic reactor with complex interdependencies is presented. This design approach enables a rapid prototyping of production system supporting multiphase catalytic reaction. Both the hardware design, such as catalysts and filling systems in the equipment in multi-stage reactors, and the design of operating conditions can be decided considering the constraints of the equipment, materials and reagents. The cycle of data acquisition and model evaluation in laboratory scale can accelerate the actual manufacturing system establishment.

**Keywords:** Continuous pharmaceutical manufacturing, Modeling and simulation, Optimization, Process design and Trickle bed reactor

## 1. Introduction

TBR is used for applications ranging from large-scale manufacturing centered on petrochemicals to small-scale manufacturing, including pharmaceutical manufacturing. When compared to other types of reactors such as bubble column, fluidized bed, slurry type reactor, and so on, TBR is preferable as a multiphase catalytic reactor because of the simple separation process and an easy replacement of catalyst (Duduković et. al., 1999). In conventional large-scale production, a robust design of the wetting efficiency of the catalyst is essential for the prevention of catalyst and equipment burnout, owing to localized severe temperature fluctuations. Process modeling is often applied to TBR design in various industrial fields, as shown in Fig.1 (Al-Dahhan et. al., 1997).

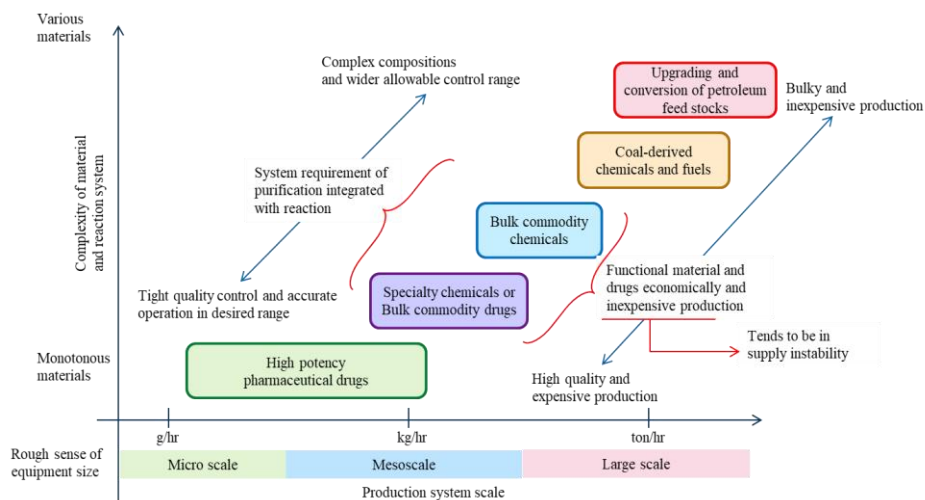


Fig.1 Conceptual map of a TBR application in various industrial fields

In the mesoscale region, engineering technology is required to establish a rapid and high-quality manufacturing system. It is necessary to determine the characteristics of the target product and establish a new database about the thermodynamic and transport properties (Masson et al., 2022). The mesoscale production system tends to be in supply instability of the functional material and drugs because of the inexpensive production considering the complicated manufacturing process. Frequently consumed mesoscale products need to be manufactured economically and inexpensively in robust facilities designed with process integration. In a mesoscale TBR design, it is necessary to select a good catalyst size to compensate for its deterioration during transport ability in the laminar flow area, to suppress the pressure drop, and to enhance the reaction rate at a high solid-liquid ratio during transport of the reactants. Furthermore, the target material must be recovered at high purity when compared to the conventional production of petroleum products, where the amount of impurities should be small. The other motivation is to increase the synthetic route of organic synthesis through applying continuous manufacturing. The reductive amination and removal of benzyl-type protecting groups are two major examples of reduction reactions that frequently use catalytic hydrogenation over precious metal catalysts (Carey et al., 2006). The basic concept of the proposed method is to utilize the pre-determined information about the process and equipment and to incorporate the reaction kinetics. It can reduce the experimental work for the preparation of the TBR system, concerning the characteristics of the reaction and catalysts.

## 2. Concept of rapid prototyping

The rapid prototyping approach consists of three steps: 1) parameter adjustment of the reaction kinetics through laboratory experiments; 2) verification of the integrated model parameters through bench-scale experiments under suppressed material consumption; and 3) determination of production scale and system specifications.

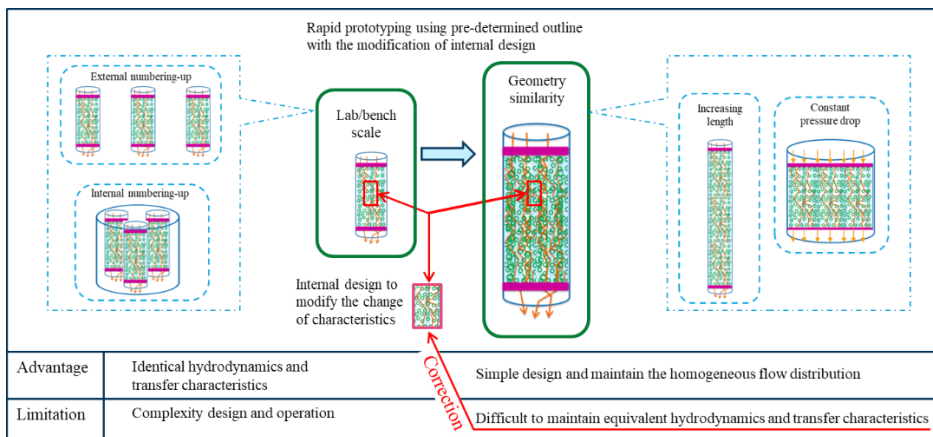


Fig.2 Concept of rapid prototyping of fixed bed multiphase catalytic reactor

With the pre-determined models in Sections 2 and 3, we can focus on the exploration of reactions and material properties characteristic, and quickly determine the required design prototypes for different manufactured products. Figure 2 shows the morphology of the scale-up policy (Dong, et. al., 2021). Numbering-up has the advantage of retaining the production conditions established from the lab/bench scale. However, there are disadvantages that it is difficult to establish design, operation and control philosophy for multiple reactors. On contrary, scaling up has the advantage concerning simple configuration and more flexible design space of internal packed features and operating condition, although the homogeneity of flow and transport phenomena are required to modify. However, it can adjust the reaction performance, heat and mass transport characteristics to the desired condition that is similar to the condition of the lab/bench scale. The navigation system consisted of segregating data base both material features and transport phenomenon can help to determine the appropriate process design. Fig.3 shows the concept of navigation system. New material features need to be created for the new target production repeatedly, although a data base of transport phenomenon is usually to provide the best route to achieve a high-quality production. The experience of development cycle will enhance this navigation system according to the accumulated properties.

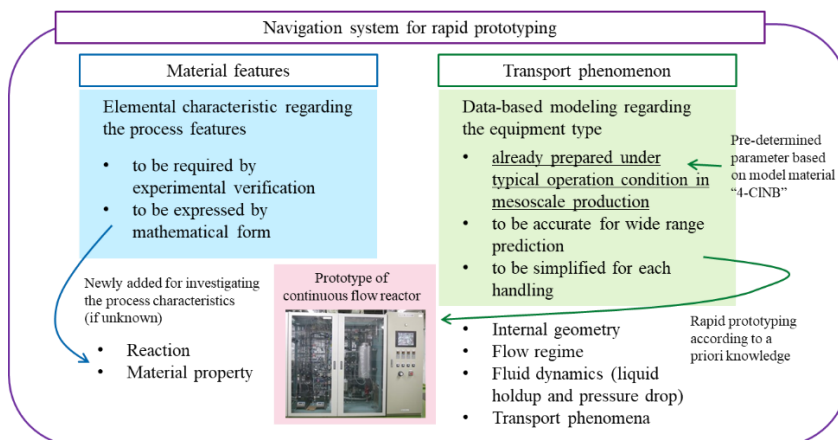


Fig.3 The concept of navigation system for rapid prototyping

### 3. Integrated modeling of trickle bed reactor

The modeling of TBR can be separated into three sections: 1) reaction kinetics; 2) the size of the equipment and catalyst; and 3) transport phenomena. If the model parameters in sections 2 and 3 are pre-determined, the reaction model with different fixed bed catalysts can be applied in a simple and rapid way. The parameters of the process model are adjusted considering the reaction kinetics, catalyst properties (size and metal loading), flow regime, axial dispersion, pressure drop, transport phenomena, and liquid holdup. Establishment of the TBR system requires the comprehension of complex independent phenomena by conducting careful experimentation and deconvolution of various parameters (Azarpour et al., 2021). The hydrogenation reaction of 4-chloronitrobenzene using a Pd/TiO<sub>2</sub> catalyst is adopted as a mockup substance and reaction system. The intermediate 4-chloronitroaniline is set as the product for the evaluation of reaction pathways by overreaction, by-products (nitrobenzene), and impurities (nitrosamines). To build this integrated model, 300 experimental data points were used under the following conditions:

Table 1 The reference data for providing mockup integrated model of TBR

Process conditions	Filling conditions	Equipment size
Substrate concentration: 50 mol/m <sup>3</sup> , 100 mol/m <sup>3</sup>	Catalyst amount: 1~50 g (based on Pd/TiO <sub>2</sub> )	Catalyst filling length: 10~320 mm
Liquid flow rate: 0.2~80 mL/min.	Metal content: 0.2~5 wt%	Diameter (I.D.): 12.2mm, 30mm
G/L ratio: 16.7~100 (volume ratio)		
Temperature: 35~60 °C		
Pressure: 0~0.5 MPaG		

Fig. 4 shows an overview of the integrated simulation model.

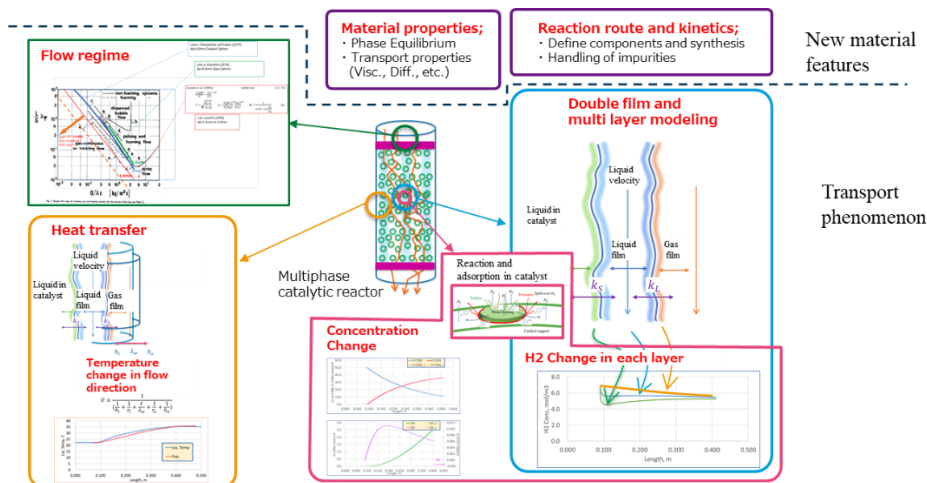


Fig.4 Integrated modeling of interdependency phenomenon

Based on the formula determined for the flow regime in a TBR (Gianetto et al., 1978), the formula is modified to adapt the process conditions to the packing size applied in the mesoscale region. As a result, the liquid holdup and pressure drop change, which affects the heat and mass transport phenomena and reaction conditions in the integrated simulation model.

#### **4. Model-based design approach**

To provide an accurate integrated model, the acquisition of high-quality experimental data is important to provide a map constituting the accuracy of process characteristics. On the other hand, scale design experiences will update a compass to predict the right reactor scale with the interior design. The integrated model can aim to design the following aspects;

##### *4.1. The internal design*

The size of the catalyst is set at a few hundred  $\mu\text{m}$ , which does not change significantly from the viewpoint of reactor pressure loss and liquid holdup capacity. However, the catalyst amount and metal loading are deemed the design parameters that affect the balance between the reaction heat generation and heat removal by the jacket coolant. The different combinations of catalyst amount and metal loading are evaluated for reaction performance. Fig. 5 shows the conversion of starting material (SM) and Fig.6 shows the selectivity of target material (TM). It can be confirmed that both Fig.5 and Fig.6 give a good estimation. In other words, it is possible to predict the desired conditions using the integrated model and to determine design conditions rapidly.

##### *4.2. The reactor scale design*

The reactor scale is an important design parameter. Higher velocity makes the transport ability better, but it makes the pressure drop worse. The liquid flow and Gas/Liquid ratio also have a significant impact on liquid holdup. Therefore, different sets of catalyst packing heights are verified. Fig.7 shows the conversion of SM, and Fig.8 shows the selectivity of TM, respectively. It can be confirmed that both Fig.7 and Fig.8 give good estimations. The integrated model can predict the TBR performance for different diameters and catalyst packing heights; therefore, the desired operating condition can be determined by the selection of the appropriate size of reactor. By combining equipment size selection and internal design, it is possible to rapidly prepare a multiphase catalytic reactor system that satisfies the required manufacturing scale and specifications.

#### **5. Conclusion**

This paper introduces a rapid prototyping method, which is a conceptual model based on a comparison between mesoscale experimental data and simulation. In the design of multiphase catalytic reactors in the mesoscale region, the basic data on transport phenomena is still insufficient. Therefore, the fluid parameter and transport phenomena are incorporated into the integrated TBR mathematical model. Improvement of the integrated TBR database can be achieved through repeated experiments and simulation.

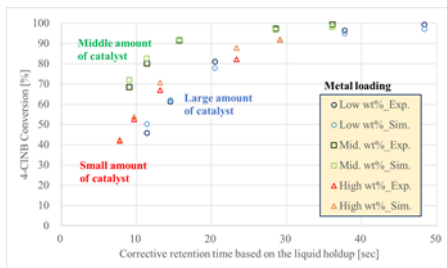


Fig.5 Prediction of conversion change based on the internal design

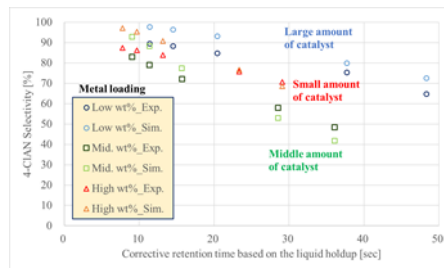


Fig.6 Prediction of yield change based on the internal design

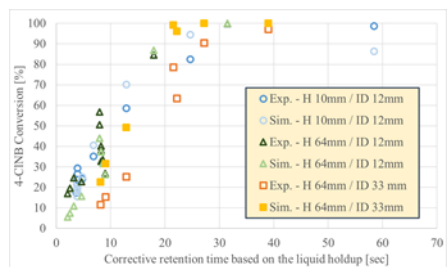


Fig.7 Prediction of conversion change based on reactor scale design

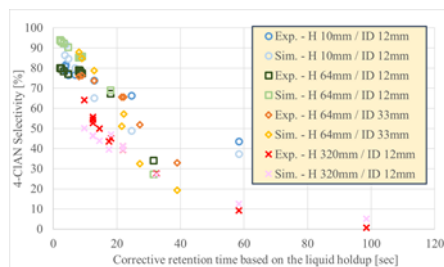


Fig.8 Prediction of yield change based on reactor scale design

This study has made it possible to evaluate the catalytic reaction performance and to quickly identify the parameters for the reaction characteristics. According to the reaction parameter incorporation into the integrated model, which can help the determination of reactor internal and scale design. This method can be used as a common platform to combine chemical and catalytic reactions, which can help researchers with the chemist and process engineer. This design approach can also make the frequent change of production items possible. Following that, while establishing the mesoscale continuous pharmaceutical manufacturing system, this approach can be used to study other processes such as synthesis and separation.

## References

- A. Azarpour et. al., 2021, Performance analysis and modeling of catalytic trickle-bed reactors: a comprehensive review, *Journal of Industrial and Engineering Chemistry*, 103, 1-41.
- J. Carey et. al., 2006, Analysis of the reactions used for the preparation of drug candidate molecules, *Organic & Biomolecular Chemistry*, 4, 2337-2347.
- M. H. Al-Dahhan, et. al., 1997, High-pressure trickle-bed reactors: A review, *Ind. Eng. Chem. Res.*, 36, 3292-3314.
- Z. Dong, Z. Wen, F. Zhao, S. Kuhn, and T. Noël, 2021, Scale-up of micro-and milli-reactors: An overview of strategies, design principles and applications, *Chemical Engineering Science: X*, 10, 100097.
- M. P. Duduković, F. Larachi, P. L. Mills, 1999, Multiphase reactors—revisited, *Chemical Engineering Science*, 5413-14, 1975-1995.
- A. Gianetto, G. Baldi, V. Specchia, S. Sicardi, 1978, Hydrodynamics and solid - liquid contacting effectiveness in trickle - bed reactors, *AIChE Journal*, 24, 6, 1087-1104.
- E. Masson et. al., 2022, Fixed bed continuous hydrogenations in trickle flow mode: A pharmaceutical industry perspective, *Org. Process Res. Dev.* 2022, 26, 2190–2223
- L. X. Yu, 2008, Pharmaceutical quality by design: product and process development, understanding, and control, *Pharmaceutical research*, 25(4), 781-791.

# Computer-aided multi-scale simulation of MOF-based membrane separation for CO<sub>2</sub> removal via integration of molecular and process simulation through machine learning

Xi Cheng,<sup>a</sup> Yangyanbing Liao,<sup>b</sup> Xiaolei Fan,<sup>b</sup> Jie Li,<sup>a</sup>

<sup>a</sup>*Centre for Process Integration, Department of Chemical Engineering, School of Engineering, the University of Manchester M13 9PL, UK*

<sup>b</sup>*Department of Chemical Engineering, School of Engineering, the University of Manchester M13 9PL, UK*  
*jie.li-2@manchester.ac.uk*

## Abstract

MOF membranes have shown high productivity in CO<sub>2</sub> removal from gas mixtures. In this work, we present a new integrated framework for reliable prediction of the process performance of MOF-based membranes (e.g. IRMOF-1). Firstly, molecular simulations are conducted to investigate the adsorption isotherms, self-diffusivity, activation energy, permeability and selectivity of IRMOF-1 for CO<sub>2</sub>/CH<sub>4</sub> separation under different operating conditions. The simulated isotherms at 298 K are in good agreement with the experimental results. As the operating conditions vary, the permeability of CO<sub>2</sub> can scale from 40898 to 381776 barrer. A similar observation is made regarding selectivity, highlighting the dramatic effect of operating conditions on membrane performance. Further, predictive models are developed with machine learning, allowing the calculation of permeability and selectivity. As the mean squared error is 0.0086 and the R-Square is 0.9822, the predictive models are shown high reliability. The predictive models integrated with the tanks-in-series model of a hollow fiber membrane separation process are simulated by the finite-volume method. The feasibility and competence of the proposed framework are illustrated in three case studies.

**Keywords:** Metal-organic framework (MOF), CO<sub>2</sub>/CH<sub>4</sub> separation, Molecular simulation, Machine learning, Membrane separation.

## 1. Introduction

As energy consumption increases due to large-scale industrial production, more and more greenhouse gases (mainly CO<sub>2</sub> and CH<sub>4</sub>) are released, leading to serious greenhouse effects such as global warming. In the meantime, the extraction of raw natural gas fields is increasing where the separation of CO<sub>2</sub> and CH<sub>4</sub> is becoming particularly crucial. For industrial applications of CO<sub>2</sub>/CH<sub>4</sub> gas mixture separation, the feed composition of CO<sub>2</sub> is widely distributed, covering a range of 30 to 90%. This makes it urgent to develop technologies for the efficient CO<sub>2</sub> removal. In order to reduce energy costs, membrane separation techniques are employed for the enrichment of lean feed gas. Membrane separation takes advantage of the selective separation function of materials. Therefore, operating conditions are typically moderate, leading to much lower energy consumption.



Furthermore, the membrane separation is generally easy to operate, control and scale up with high separation efficiency.

There are two parameters that characterize the efficiency of membrane separations, including permeability and selectivity. Conventional polymer membranes typically have a trade-off relationship between permeability and selectivity, failing to achieve high permeability and high selectivity simultaneously. Porous membranes are considered to be the alternative to polymer membranes due to their adsorption behavior and pore size sieving effect, which allow them to break through the limitation of Robeson trade-off upper bound of the polymer membrane (Robeson 2008). MOF stands for a new type of porous materials which have shown high productivity in CO<sub>2</sub> removal from gas mixtures. Various MOF materials have been synthesized experimentally, however, unfortunately their properties are generally evaluated at ambient temperature and 1 bar pressure (Erucar and Keskin 2013). Due to the fact that these properties can vary significantly depending on the operating conditions, this evaluation is no longer sufficient for industrial applications under varying operating conditions. On one hand, experimental studies are usually complex and expensive because of the uncertainty of uncontrollable errors and the variability of samples with impurities and defects. On the other hand, some MOF membrane performance have been evaluated by molecular simulations, most of these are carried out at ambient temperature and a fixed pressure (Li, Kresse et al. 2001). Therefore, existing work is insufficient to support a process optimization approach to identify the optimal design for CO<sub>2</sub>/CH<sub>4</sub> membrane separation.

Although molecular simulation is extensively applied to study gas separation in MOFs, it is computationally expensive in terms of resources and time costs (Pronk, Pall et al. 2013). For this reason, it is a challenge to integrate molecular simulations with process simulations. Through molecular simulation it is possible to generate the big dataset on the performance of MOFs under different operating conditions. A prediction model can be developed based on this big dataset by machine learning. This incorporation shows great potential as it can efficiently realize integration design. To the best of our knowledge, there is currently no framework developed for the integration of molecular simulation into membrane process simulation via machine learning. Consequently, the main objective of this study is to develop a novel multi-scale design framework via the integration of molecular and process simulation through machine learning, which is evaluated by exploring the potential of a MOF-based membrane for the separation of CO<sub>2</sub>/CH<sub>4</sub> mixture.

## 2. Methodology

### 2.1. Molecular Simulation

In this work, we firstly apply grand canonical Monte Carlo (GCMC) and molecular dynamics (MD) simulation to investigate the adsorption, diffusion, and permeation of a MOF-based membrane (e.g. IRMOF-1) for CO<sub>2</sub>/CH<sub>4</sub> mixtures separation under different operating conditions. We choose IRMOF-1 as the MOF-based membrane material because it is one of the earliest MOFs that can be stably synthesized and industrialized and has been successfully fabricated as film. It is also the parent of the classical IRMOF series with a large number of available literature and experimental data. The CO<sub>2</sub>/CH<sub>4</sub> binary mixtures with varied feed CO<sub>2</sub> compositions from natural gas is used in this work. The molecular simulation is carried out at temperatures of 273, 298, 323, 348, 373 and 423 K under pressure range of 0–20 bar. The adsorption isotherms of pure CO<sub>2</sub>, CH<sub>4</sub> and their mixtures in IRMOF-1 are evaluated by GCMC simulation. MD simulation is then performed in NVT ensemble to investigate the self-diffusivity and diffusion activation

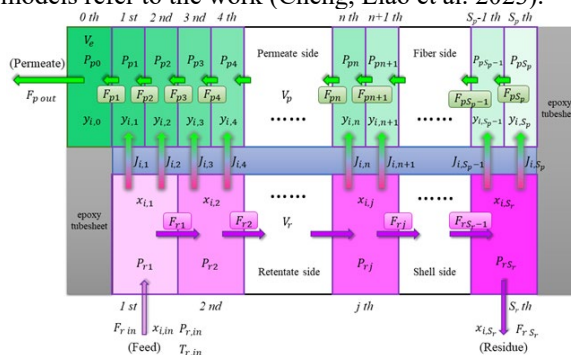
energy. The permeability and membrane selectivity of IRMOF-1 are also calculated based on the solution-diffusion model. For detailed parameters and validation refer to the work (Cheng, Liao et al. 2023).

### 2.2. Machine Learning (e.g. Artificial Neural Network)

Based on the GCMC & MD simulation for the CO<sub>2</sub>/CH<sub>4</sub> mixture separated in IRMOF-1 membrane under different operating conditions, two ANN models are developed and trained to describe the adsorption capacity and self-diffusivity, respectively. There are approximately 70% of the data selected randomly from dataset (1340 points) for model training, 15% of the data for validation and 15% of the data for testing. The ANN models use the feed CO<sub>2</sub> composition, temperature, and partial pressure as the input layer to create the input vector. The output layer of the network is a four-dimensional output vector containing adsorption capacity and self-diffusivity of each component. For detailed configuration of ANN refer to the work (Cheng, Liao et al. 2023).

### 2.3. Mathematical Modeling of Hollow Fiber Membrane

This work adopts a tanks-in-series model (Katoh, Tokumura et al. 2011) of the hollow fiber membrane as shown in Fig. 1, feeding from the shell side, capturing permeate CO<sub>2</sub> from the inside of fibers and yielding a product from the residue side. The tanks-in-series model consists of a series of equal sized and well-mixed tanks which are numbered in both sides from the feed end to the residue end. The number of tanks in the permeate side is set to double of that in the residue side. The mathematical model includes the relationship governing transport across the membrane, mass balance equations, pressure drop equation for permeate/feed sides and boundary and initial conditions that reflect the configuration and operation of the hollow fiber membrane module. For detailed expressions and models refer to the work (Cheng, Liao et al. 2023).



**Figure 1** Schematic of tanks-in-series model for a counter-current hollow-fiber

### 2.4. Implementation

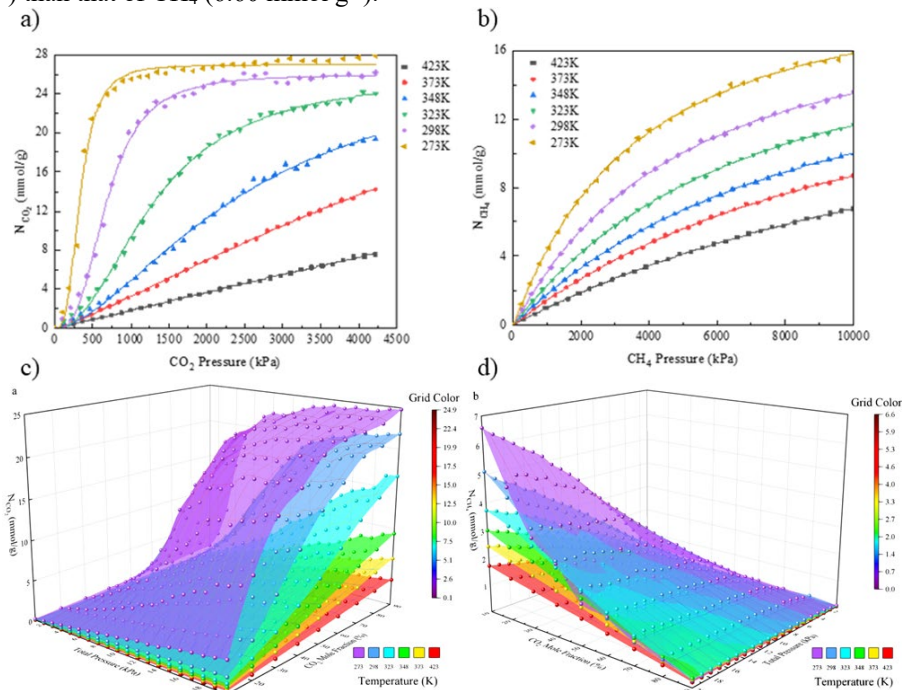
The GCMC and MD simulations are conducted in Materials Studio software. The ANN training, validation and testing are conducted in MATLAB providing early stopping applied in this work. The mathematical model of hollow fiber membrane separation process is simulated using the finite volume method implemented in MATLAB.

## 3. Results

### 3.1. GCMC & MD Simulation

Fig. 2a-b show the adsorption isotherms of pure CO<sub>2</sub> and CH<sub>4</sub> while Fig. 2c-d show adsorption isothermal surfaces of mixed CO<sub>2</sub> and mixed CH<sub>4</sub> in varied composition binary mixtures in IRMOF-1 from the GCMC simulation. On one hand, the pure gas GCMC results show that the maximum adsorption capacity of CO<sub>2</sub> is as high as 27.91

mmol g<sup>-1</sup>, while the maximum adsorption capacity of CH<sub>4</sub> reaches 15.86 mmol g<sup>-1</sup>. On the other hand, the mixture GCMC results demonstrate that CO<sub>2</sub> plays a dominant role in the mixed gas adsorption behavior, with a much larger adsorption capacity (24.80 mmol g<sup>-1</sup>) than that of CH<sub>4</sub> (6.60 mmol g<sup>-1</sup>).



**Figure 2** Adsorption isotherms of pure CO<sub>2</sub>, CH<sub>4</sub> and CO<sub>2</sub>/CH<sub>4</sub> binary mixture at different temperatures: a) pure CO<sub>2</sub> b) pure CH<sub>4</sub> c) mixed CO<sub>2</sub> d) mixed CH<sub>4</sub>

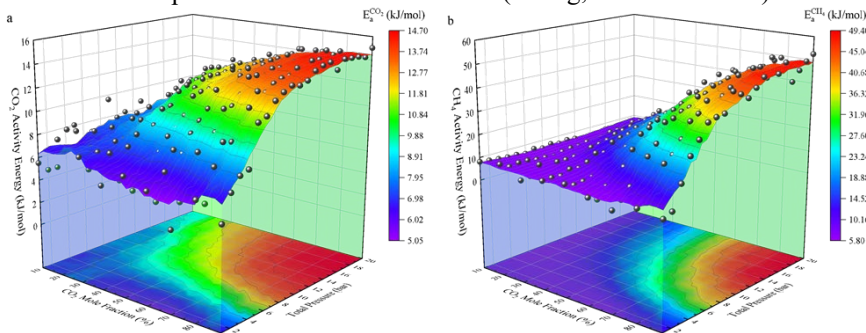
**Table 1** Diffusion activation energy of pure CO<sub>2</sub> and CH<sub>4</sub> at 1, 5, 10 bar

p (bar)	Pure CO <sub>2</sub> E <sub>a</sub> (kJ mol <sup>-1</sup> )	Pure CH <sub>4</sub> E <sub>a</sub> (kJ mol <sup>-1</sup> )
1	9.2	6.6
5	11.6	5.6
10	15.0	5.4

As observed from Table 1, the MD results for pure gases suggests that the activation energy of CO<sub>2</sub> diffusion in IRMOF-1 is greater than that of CH<sub>4</sub> under the same operating conditions, indicating that the diffusion of CO<sub>2</sub> is more difficult. This conclusion is consistent with the fact that the self-diffusivity of CO<sub>2</sub> is smaller than that of CH<sub>4</sub>. On the other hand, the mixture MD results are completely opposite to the pure gas results, which can be seen from Fig. 3. The activation energy of CO<sub>2</sub> diffusion in IRMOF-1 is much lower than that of CH<sub>4</sub> in the mixture, which is also confirmed by experimental phenomena of CO<sub>2</sub>/CH<sub>4</sub> binary mixture diffusion in other porous materials (Zhao, Feng et al. 2016). The combined effect of competitive adsorption and synergistic diffusion leads to a significant reduction in activation energy of mixed CO<sub>2</sub> diffusion. CH<sub>4</sub> carries CO<sub>2</sub> along with it through the membrane, while large amounts of CO<sub>2</sub> get stuck in the pores of membrane, blocking the pathway and making the diffusion of CH<sub>4</sub> difficult.

Based on the solution-diffusion model, permeability and selectivity could be calculated. The mixture results show that CO<sub>2</sub> permeability in IRMOF-1 is ranging from

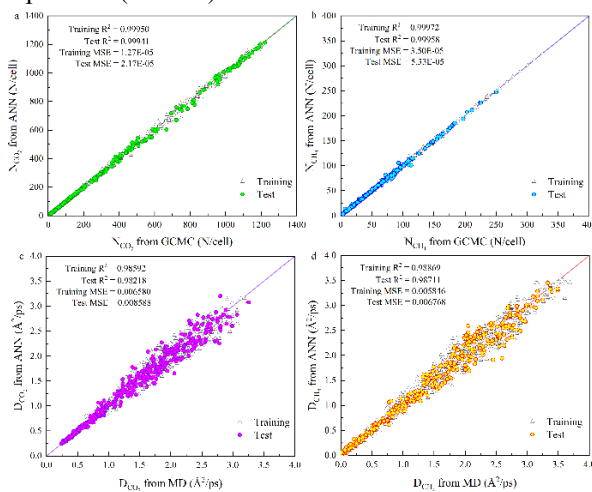
40898 to 381776 barrer with the variation in operating conditions, the same phenomenon is also reflected in the selectivity. These indicate the great effect of operating conditions on the performance of MOF membrane. The rest of available data and analysis of GCMC and MD results are presented in detail in the work (Cheng, Liao et al. 2023).



**Figure 3** Diffusion activation energy of CO<sub>2</sub>/CH<sub>4</sub> mixtures: a)  $E_a$  of mixed CO<sub>2</sub>, b)  $E_a$  of mixed CH<sub>4</sub>

### 3.2. ANN Prediction

Prediction models for adsorption and diffusion are developed through ANN. The MSE of prediction test is 0.0086, and R<sup>2</sup> is 0.9822, which indicate the high prediction accuracy of the developed ANN models. Fig. 4 reveals the parity plot for adsorption capacity and self-diffusivity of CO<sub>2</sub> and CH<sub>4</sub> in mixtures from ANN prediction and GCMC & MD results under various operating conditions such as temperature (273–423 K), pressure (0–20 bar), and feed CO<sub>2</sub> composition (0.1–0.9).



**Figure 4** Parity plot for prediction from ANN vs. GCMC & MD

### 3.3. Integrated Process Simulation

Three case studies by process integration and simulation of membrane separation of CO<sub>2</sub>/CH<sub>4</sub> mixture from raw natural gas, landfill gas and shale gas indicate that IRMOF-1 membrane can produce greater recovery with significantly smaller membrane area than the polymer PTMSP membrane from Ref. (Robeson 2008) to achieve the desired target purity. The computational results using IRMOF-1 membrane with the developed ANN model and using PTMSP membrane are provided in Table 2. All these demonstrate the proposed multi-scale design framework is feasible, practical, and superior. It integrates

the variation of permeability with operating conditions into the dynamic simulation of membrane separation process. The rest of available data and analysis of integration results are presented in detail in the work (Cheng, Liao et al. 2023).

**Table 2** Computational results of case study

Criterion	Case 1		Case 2		Case 3	
	This work	Reference work [9]	This work	Reference work [9]	This work	Reference work [9]
$F_{permeate}$ (Nm <sup>3</sup> /h)	652	716	686	800	1186	1482
$F_{residue-product}$ (Nm <sup>3</sup> /h)	1348	1284	1314	1200	814	518
<b>Fiber number</b>	6062	45600	7408	40500	17964	50480
<b>Membrane area</b> $A_m$ (m <sup>2</sup> )	<b>14.47</b>	108.9	<b>17.69</b>	96.70	<b>42.89</b>	120.5
<b>Purity<sub>CH<sub>4</sub></sub></b> (mole %)	90	90	80	80	65	65
<b>Recovery<sub>CH<sub>4</sub></sub></b> (%)	<b>76</b>	72	<b>81</b>	74	<b>88</b>	56

#### 4. Conclusion

In this paper, the integrated design framework we initially proposed is well demonstrated in the gas separation hollow fiber membrane process with the example of IRMOF-1, which provides an integration idea for future optimization of process conditions and innovative design of optimal MOF membrane material.

#### References

- Cheng, X., Y. Liao, Z. Lei, J. Li, X. Fan and X. Xiao (2023). "Multi-scale design of MOF-based membrane separation for CO<sub>2</sub>/CH<sub>4</sub> mixture via integration of molecular simulation, machine learning and process modeling and simulation." *Journal of Membrane Science*: 121430.
- Erucar, I. and S. Keskin (2013). "High CO<sub>2</sub> Selectivity of an Amine-Functionalized Metal Organic Framework in Adsorption-Based and Membrane-Based Gas Separations." *Industrial & Engineering Chemistry Research* **52**(9): 3462-3472.
- Katoh, T., M. Tokumura, H. Yoshikawa and Y. Kawase (2011). "Dynamic simulation of multicomponent gas separation by hollow-fiber membrane module: Nonideal mixing flows in permeate and residue sides using the tanks-in-series model." *Separation and Purification Technology* **76**(3): 362-372.
- Li, X.-G., I. Kresse, Z.-K. Xu and J. Springer (2001). "Effect of temperature and pressure on gas transport in ethyl cellulose membrane." *Polymer* **42**(16): 6801-6810.
- Pronk, S., S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl (2013). "GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit." *Bioinformatics* **29**(7): 845-854.
- Robeson, L. M. (2008). "The upper bound revisited." *Journal of Membrane Science* **320**(1-2): 390-400.
- Zhao, Y., Y. Feng and X. Zhang (2016). "Selective Adsorption and Selective Transport Diffusion of CO<sub>2</sub>-CH<sub>4</sub> Binary Mixture in Coal Ultramicropores." *Environmental Science & Technology* **50**(17): 9380-9389.

# Decision making software for cosmetic product design based on an ontology

Alex Gabriel<sup>a</sup>, Juliana Serna<sup>a,b\*</sup>, Valentin Plantard–Wahl<sup>c</sup>, Antoine Le Jemtel<sup>c</sup>,  
Vincent Boly<sup>a</sup>, Véronique Falk<sup>b</sup>

<sup>a</sup>*Université de Lorraine, Equipe de Recherche des Processus Innovatif, ERPI-ENSGSI, 8 Rue Bastien-Lepage, 54000 Nancy, France*

<sup>b</sup>*Université de Lorraine, Laboratoire Réactions et Génie des Procédés, CNRS-LRGP-ENSIC, 1 Rue Grandville, 54000 Nancy, France*

<sup>c</sup>*Université de Lorraine, TELECOM Nancy, 193 Avenue Paul Mulle, 54600 Villers-lès-Nancy, France*

\*[jsernaro@unal.edu.co](mailto:jsernaro@unal.edu.co)

## Abstract

The design of formulated chemical products such as cosmetic emulsions is normally carried out by interdisciplinary teams that need versatile and reliable tools to support them in the decision-making process. In view of this, this article presents a tool to support product designers during the conceptual design of emulsified cosmetic products. The tool is based on a previously developed ontology for cosmetic product design. It integrates concepts from emulsion science, cosmetic formulation, expert knowledge, and design heuristics in a systematic and flexible way. Additionally, it has a modern architecture based on mobile applications which makes it easy to use. The tool support three types of decisions: screening of ingredients, selection of ingredients considering multiple criteria (ingredient performance, origin, price, etc.), and evaluation of a formulation according to design heuristic. As a result, the tool allows the visualization of ingredient properties and a better selection of ingredients, as well as an early evaluation of formulations prior to experimentation.

**Keywords:** Product design, Cosmetic emulsions, Software application, Ontology

## 1. Introduction

The design of cosmetic products is a complex task that involves the management of multiple and heterogeneous information: raw material data, mathematical models, design heuristics, emulsions science principles, experimental data, among others (Zhang et al., 2020) (Kontogeorgis et al., 2022). Thereby, tools capable of handling varied information and comprehensively representing the solution space of design options are needed. They would have the potential to assist product designers in making decisions to seek original and context-specific solutions. Those tools should be flexible to be adapted to different design problems, easy to use and understand to help rather than hinder the design process, and reliable to be effective in supporting decision making.

There have been extensive contributions from the academy to produce tools for chemical product design. Some of the most notable are ProCAPD with an emphasis on the prediction of physicochemical properties for product optimization (Kalakul et al., 2018) and the virtual process-product design laboratory (VPPD-Lab) which contains a wide

database of property and molecular design models to support design decisions (Kontogeorgis et al., 2022). Previously mentioned tools are generic and not specific for cosmetic design. They are mostly applicable for the definition and resolution of optimization design problems, i.e., they can be used when the design problem is or can be well and explicitly defined. Additional examples of practical tools to support formulators in the design of cosmetics are the app-based resources from Abbott (2022), which shows how scientific principles of colloidal science can be used to explore design solutions for different formulated products, mainly paints, and the commercial software Ansys Granta MI for material design and selection (Ansys, 2022).

However, the drawbacks of the contributions are their complexity, their lack of specificity for cosmetics and their low accessibility and ease of use. In view of the above, this article presents a cross platform decision-making tool for the design of emulsion-based cosmetics at early stages. It is called Formultools. Its architecture was inspired by the OntoCosmetic ontology (Serna et al., 2022), whose scope is the formulation of emulsion-based cosmetic products.

## 2. Design methodology

Considering the premise that the tool must be accessible, a great importance was given to the interaction with the user/formulator. The design process was inspired by the 5-plans user-centered design approach (Garrett, 2011). It involved cosmetics experts and computer scientists in co-design. The development was done iteratively with intermediate usability tests with non-expert users. Each experience was assessed through a questionnaire using the method AttrakDiff (Lallemand et al., 2015) to follow usability improvement at each iteration. The problems encountered were progressively corrected until the current version was released.

### 2.1. Tool scope and functionality

The decision-making tool targets cosmetic product designers (the user). It is to be implemented at the early design stages. Three purposes were defined: the exploration of ingredients according to their properties; the selection of ingredients considering specific criteria, the evaluation of a possible formulation according to design heuristics.

## 3. Formultools: an application for the design of cosmetic products

### 3.1. Ontology as a foundation

The ontology OntoCosmetic is a knowledge representation containing all the concepts used for the development of the application. The ontology scope is the formulation of emulsion based cosmetic products. It is based on four main concepts: ingredients, formulations (list of ingredients with composition), properties, design heuristics, and their interrelations. Figure 1 presents an overview of the ontology:

- **Ingredients and formulation (dosage):** These concepts correspond to the variables that designers can select to achieve expected product properties, i.e., they are the independent variables because designers can select ingredients and their composition (dosage). Ingredients can be emollients, surfactants, thickeners, actives, other ingredients (stabilizers, preservatives, etc.).
- **Properties:** There are product properties and ingredients properties, and they characterize the behavior of the product and of the different ingredients respectively. Some properties are specific to a category of ingredient such as the HLB for surfactants. Some properties are quantitative (as the HLB) or qualitative (as the origin of ingredients, which can be natural or synthetic). As an example, Table 1 presents a list of the properties of emollients and their characteristics.

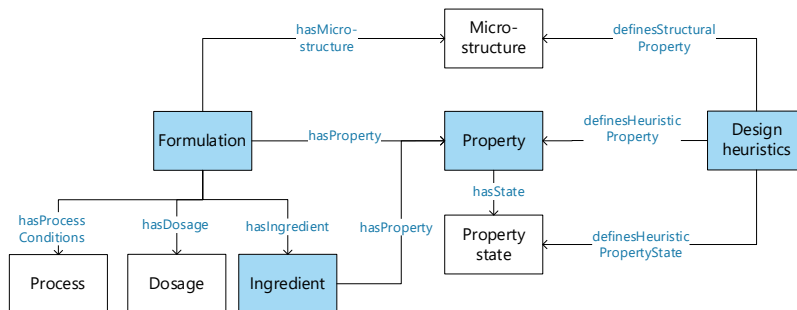


Figure 1. Overview of OntoCosmetic

Table 1. Main emollient properties in OntoCosmetic

CAS	Meaning	Range
INCI name	Abbreviation for International Nomenclature of Cosmetic Ingredients	String
Chemical class	chemical nature of an emollients: ester, fatty alcohol, hydrocarbons, silicones, triglyceride, fatty acid	String
Polarity (qualitative)	It determines the types of inter-molecular forces of the substance.	
Viscosity (qualitative)	Resistance of a fluid to the deformation	
Spreading (qualitative)	It identifies the easy of a fluid to spread on a surface	
Emollience (qualitative)	It indicates if an emollient produces a light skin feel or a rich skin feel	
After feeling (qualitative)	It indicates if an emollient leaves an oily residue after application	
State	indicates if the substance is liquid or solid.	
Biodegradability	It is the capacity for biological degradation of organic materials by living organisms	
Origin	It makes refer to the source of the substance: if it is of petrochemical origin or natural origin.	
Price	Ingredient cost	Float
HLBr	It represents the HLB needed by an oil or a mixture of oils to be more likely to be incorporated in a stable O/W emulsion.	Float
MW	Mass of a compound divided by the number of moles	Float

- **Design heuristics:** They are design rules based on expert knowledge and they relate ingredients, ingredients properties and the formulation with product properties and their values. They enable to predict the effects of design decisions in product properties. As an example, Table 2 presents a list containing some of the design heuristics.

Table 2. Example of design heuristics in OntoCosmetic

Product property	Heuristic
Stability	An emulsion tends to be more stable using a suitable combination of surfactants with a HLB equivalent to the HLB required by the emollients
Stability	It is recommended to use more than one surfactant
Stability	The quantity of surfactant should be between 14 to 20% of the total oily phase
Viscosity	Viscosity is low (lotion texture) if a low proportion of thickener is used
Viscosity	Viscosity is high (cream to gel texture) if a high proportion of thickener is used
Viscosity	Viscosity increases if the content of one of the following fatty alcohol increases: cetearyl alcohol, stearyl alcohol or cetyl alcohol
Viscosity	Viscosity is decreasing if a branched-chain higher alcohol is added

### 3.2. Formultools: a cross platform tool based on OntoCosmetic

Ontologies are great for formalizing knowledge but they are not intended for end users. Formultools aims to provide user-friendly interface to product designers who wants to



explore the design space at early design stages looking for product concepts to be tested. It was developed to be simple, accessible, useful, and versatile. Its data architecture is directly inspired by the OntoCosmetic ontology as the knowledge graph used is an instantiated version of OntoCosmetic with ingredients, heuristics, and properties. The structure of the application is split into two main sections as represented in **Error! Reference source not found.**. An exploration section, which is focused on ingredients searching within a collection, visualization, and ranking. And a formulation section calculating product properties based on heuristics for their verification. Its interface is inspired on mobile applications.

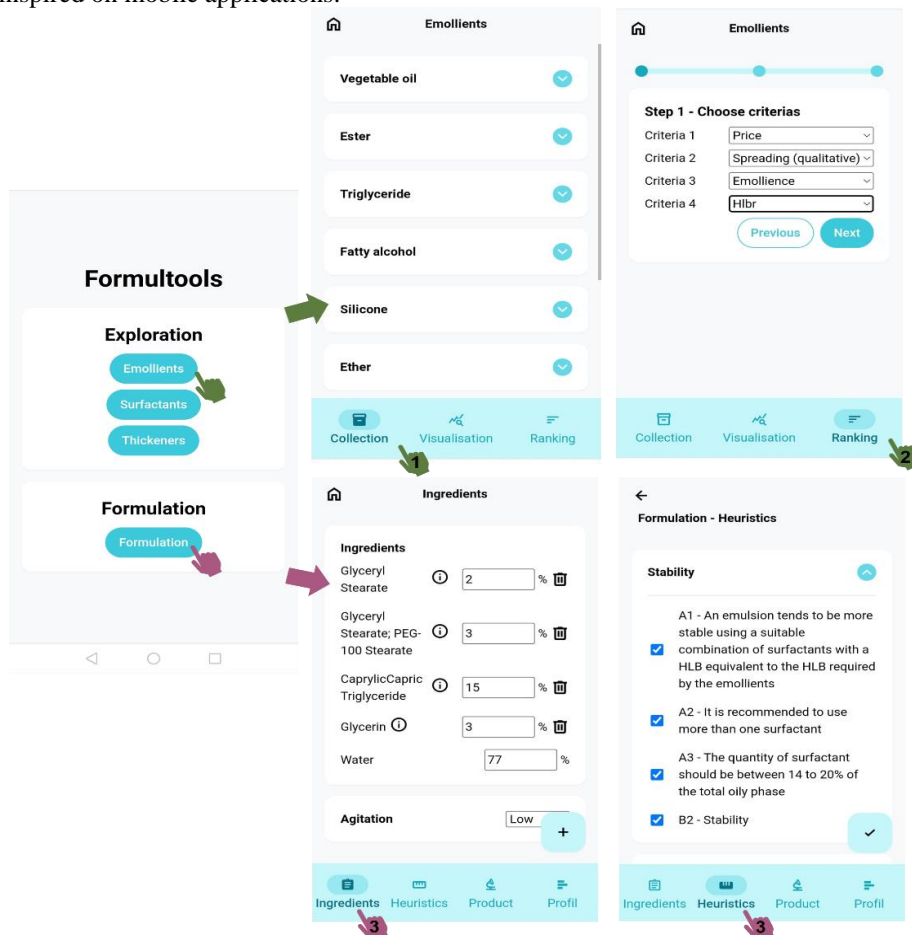


Figure 2. General interface of Formultools

The main functions of Formultools are explained below:

- **Ingredient searching and visualization:** The application has three categories of ingredients according to their functions in cosmetic products: emollients, surfactants, and thickeners, which can be accessed in the home page. Subsequently, the collection of ingredients of each category are listed and presented according to their chemical class. Ingredients can be explored in the following two ways: they can be viewed with their individual properties on the collection tap, they can be explored based on the

values of their properties and compared graphically on the visualization tap. The latter also presents the filter option, where ingredients can be screened according to their properties.

- Figure 3 presents the exploration options.

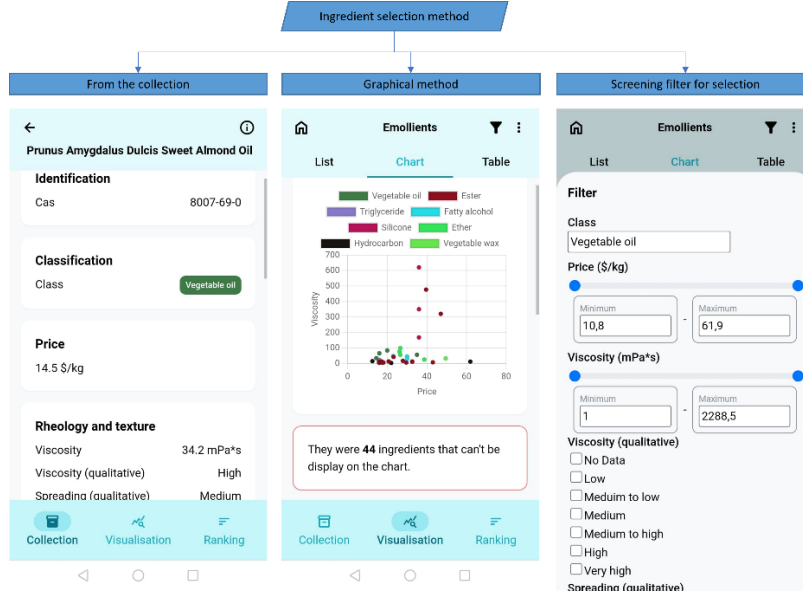


Figure 3. Exploring emollients using Formultools (collection, visualization, filter)

- Ingredient selection: The application enables the ranking of ingredients considering the properties that the designer wants to optimize. For this, the method Analytic Hierarchy Process (AHP) (Saaty, 1987) is used. With the aid of the tool, the formulator first defines the proprieties considered as criteria of selection in the first tap. Then, he defines which proprieties must be minimized and maximized in the second tap. Thirdly, he makes a pairwise comparison of proprieties to calculate their weights. The AHP method implies the calculation of a consistency ratio to evaluate the pairwise comparison, which can be also visualized. The last step is the visualization of the ranking. The four steps of ingredient selection are illustrated in Figure 4.

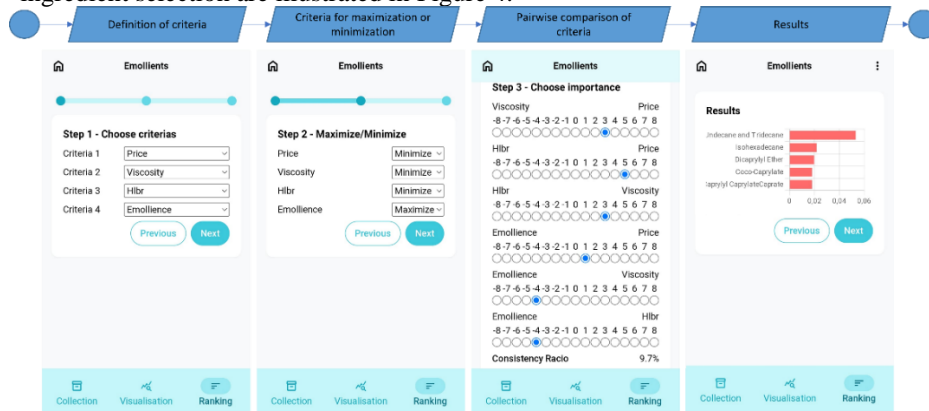


Figure 4. Ranking process of cosmetic ingredients using Formultools

- Formulation evaluation: FormulTools enables the evaluation of formulations suggested by designers according to design heuristic and recommendations contained in the data base. He defines a formulation (ingredients and composition), select design heuristics from the data base and based on the heuristics, the tool calculate product properties. In this way, the designer can propose solutions to design specifications and check if those solutions would correspond to specifications according to heuristics before testing them experimentally.

#### 4. Conclusions and perspectives

The article presents Formultools, a tool to support decision-making process for the design of cosmetic emulsions at early stages. Currently, the tool can perform three functions: exploration and selection of ingredients and verification of cosmetic formulations. It was designed to make it as practical as possible which is why it has the form of a mobile application. In a near future, it will be tested with experts and improved for its subsequently application in real design cases.

Formultools was developed in co-design between experts on cosmetic formulation and computer science. Additionally, it was continuously tested with users and iteratively improved. As a result, the tool was developed in a very agile process.

In the long term, the tool will be enhanced to assist the design process in a more holistic way. This will require the implementation of new features to enable information management, learning from each successful design, use by various design agents, among others, while maintaining its simplicity and usability.

#### References

- Abbott, S., 2022. Science for the Real World [WWW Document]. URL <https://www.stevenabbott.co.uk/index.php>
- Ansys, 2022. Materials: Leading Material Intelligence [WWW Document]. URL <https://www.ansys.com/products/materials#tab1-2>
- Garrett, J. J. (2011). The elements of user experience: User-centered design for the Web. In *Interactions*. New Riders. <https://doi.org/10.1145/889692.889709>
- Kalakul, S., Zhang, L., Choudhury, H.A., Elbashir, N.O., Eden, M.R., Gani, R., 2018. ProCAPD – A Computer-Aided Model-Based Tool for Chemical Product Design and Analysis, in: Eden, M.R., Ierapetritou, M.G., Towler, G.P.B.T.-C.A.C.E. (Eds.), *13 International Symposium on Process Systems Engineering (PSE 2018)*. Elsevier, pp. 469–474. <https://doi.org/https://doi.org/10.1016/B978-0-444-64241-7.50073-2>
- Kontogeorgis, G.M., Jhamb, S., Liang, X., Dam-Johansen, K., 2022. Computer-aided design of formulated products. *Curr. Opin. Colloid Interface Sci.* 57, 101536. <https://doi.org/10.1016/j.cocis.2021.101536>
- Lallemand, C., Koenig, V., Gronier, G., & Martin, R. (2015). Création et validation d'une version française du questionnaire AttrakDiff pour l'évaluation de l'expérience utilisateur des systèmes interactifs. *European Review of Applied Psychology*, 65(5), 239-252. <https://doi.org/10.1016/j.erap.2015.08.002>
- Saaty, R.W., 1987. The analytic hierarchy process—what it is and how it is used. *Math. Model.* 9, 161–176. [https://doi.org/10.1016/0270-0255\(87\)90473-8](https://doi.org/10.1016/0270-0255(87)90473-8)
- Serna, J., Rivera-gil, J.L., Gabriel, A., Arrieta-Escobar, J.A., Boly, V., Falk, V., Narvaez Rincon, P.C., 2022. Application of an ontology-based decision support system for the design of emulsion-based cosmetic products, in: *ESCAPE 32*. Toulouse.
- Zhang, L., Mao, H., Liu, Q., Gani, R., 2020. Chemical product design – recent advances and perspectives. *Curr. Opin. Chem. Eng.* 27, 22–34. <https://doi.org/10.1016/j.coche.2019.10.005>

# Evaluation of amines for the removal of acid gases from the liquid hydrocarbons

Siddig S. Khalafalla<sup>a</sup>, Umer Zahid<sup>a, b</sup>

<sup>a</sup> *Chemical Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia.*

<sup>b</sup> *Interdisciplinary Research Center for Membranes & Water Security, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia.*  
[uzahid@kfupm.edu.sa](mailto:uzahid@kfupm.edu.sa)

## Abstract

Liquefied Petroleum Gas (LPG) is a product made from crude oil and has several domestic and industrial applications. Continuous increases in energy demand around the world and strict environmental regulations require clean fuel for various applications. This requires the cleaning of liquid hydrocarbons from acid gases which are toxic to human health and disastrous to industrial equipment. Currently, most of the liquid hydrocarbon treatment is done using caustic solutions to remove acid gases. However, depending on the concentration of impurities in the feed and the amount of feed to be processed, a large amount of caustic solution is required which poses another disposal issue. This study proposes to replace the conventional caustic cleaning process with the amine process for liquid hydrocarbon treatment. The process consists of a liquid-liquid contactor where impurities are removed using the amine and then rich amine is regenerated in the stripper column. This study's goal is to assess various amine types for removing acid gases from liquid hydrocarbon feeds. The study is conducted using commercial simulation software ASPEN HYSYS<sup>®</sup> for the design and analysis of process performance. The studied variables include the concentration of the amine, the amine circulation rate, the effect of LPG feed temperature, and the lean amine temperature. The outcomes demonstrate that the removal of acid contaminants from the hydrocarbons can be a competitive process to achieve the required product purities.

**Keywords:** LPG sweetening; acid gases; process design; amines.

## 1. Introduction

LPG is typically a mixture of propane and butane in various proportions derived from the processing of crude oil or natural gas. LPG has a wide range of domestic and industrial usage. The global demand for LPG has been raised over the past decades due to the need for clean energy. Additionally, a compound annual growth rate (CAGR) of 2.2% is anticipated from 2021 to 2026 (*Global Five-Year LPG Demand Fuelled by New Petchems Plants, Res-Com; Supply to Overwhelm*, n.d.). Moreover, at the same time, strict environmental policies are in place to curtail environmental and health issues, therefore the demand for liquid hydrocarbon sweetening facilities has increased. The most common impurities present in LPG are carbon dioxide (CO<sub>2</sub>), hydrogen sulfide (H<sub>2</sub>S), carbon sulfides (COS), and mercaptans (RSH). These impurities, if not removed, can cause corrosion, operational inefficiencies, and other personnel health issues. Currently, many industries are using caustic solutions to treat the contaminants. However, usually, a large amount of caustic solution is required for the hydrocarbon sweetening which poses another environmental issue. Another method to remove impurities from the liquid hydrocarbon is to treat it with the amine solution. The typical amine process consists of

contactor-stripper columns in series where the acid impurities are removed with amine solvent in a liquid-liquid extractor (LLE) (Jayakumar et al., 2017).

Amines have been used for many years in the natural gas industry for the sweetening of natural gas (Zahid et al., 2020). However, the application of amines for the treatment of liquid hydrocarbon is rarely studied. Most of the existing studies have focused on hydrocarbon treatment using caustic solutions and various ways how the process efficiency can be improved. However, in the literature, there are only a few studies that have investigated the use of amines for liquid hydrocarbon treatment. Qeshta et al. (Qeshta et al., 2015) studied the LPG sweetening process using MDEA for the Takreer refinery in Abu Dhabi, UAE where the current LPG sweetening process is the MEROX process. They concluded that amine can be a good alternative process for the treatment of LPG which can reduce significant losses in the caustic process. Zoghi et al. (Zoghi et al., 2022) investigated the effect of 30 wt.% DEA solution for the treatment of liquid propane containing both  $H_2S$  and  $CO_2$ .

The primary purpose of this research work is to focus on the design and simulation of the liquid hydrocarbon sweetening process using different amines in a liquid-liquid extractor (LLE). Three amine solvents namely methyl diethanolamine (MDEA), diglycolamine (DGA), and diethanolamine (DEA) are evaluated to analyze the sweetening performance. The goal is to keep  $H_2S$  and  $CO_2$  in the sweet hydrocarbon stream leaving at the top of the contactor below 5 ppm. Finally, a parametric analysis has been performed to study the effect of various operational parameters.

## 2. Process Details

### 2.1. Thermodynamic

Aspen HYSYS was used to build the steady-state simulation model. For the calculation of the properties of the gas phase, the Peng-Robinson (PR) equation of state has been employed while the E-NRTL model for electrolytic thermodynamics with binary and electrolyte pair parameters optimized for liquid-liquid applications has been selected. Aspen HYSYS has a built-in library for all the equilibrium and kinetic reactions taking place in the acid gas removal (AGR) system which has been utilized from the acid gas-liquid treatment package. The available acid gas-liquid treatment thermodynamic package has been validated over a wide range of operating conditions with good accuracy, it can be used with confidence (Zahid et al., 2020).

### 2.2. Process Description

Figure 1 shows a typical LPG sweetening unit. Initially, LPG is fed at the bottom of the liquid-liquid extractor. As the amine solution flows countercurrent to the sour LPG, the components of the acid gas react with the amine to produce regenerable salts. The rich amine leaves at the bottom of the LLE, while the sweet LPG leaves at the top. Following this, any absorbed hydrocarbons are removed in a flash drum. The rich amine is then fed to the top of the stripper column after being heated in the rich/lean amine heat exchanger.  $H_2S$  and  $CO_2$  are stripped from the rich amine as it flows down the column, leaving the lean amine at the bottom. It is then recycled to the top of the absorber where it passes through a lean/rich exchanger and repeats the cycle.

### 2.3. Process Simulation

This study explored three simulation cases with different impurities in the sour LPG feed as shown in Table 1. The first one has both  $H_2S$  and  $CO_2$ , while the second and the third have only either  $CO_2$  or  $H_2S$  respectively. For each sour feed, three different amine solutions have been tested MDEA (35 wt.%), DGA (50 wt.%), and DEA (30 wt.%). The input variables to the simulation models are listed in Table 2. Trayed columns have been

used to model the absorber and stripper due to high solvent circulation flow. For towers with a low-pressure drop requirement and a diameter of fewer than 2.5 feet, packed columns are typically advised. For single amines, the circulation rate is adjusted according to the required product purity, which is not more than 5 ppm of H<sub>2</sub>S in the sweet gas. The stripper column's vent rate and reflux ratio have been specified to achieve column convergence, whereas the absorber column has a degree of freedom equal to zero.

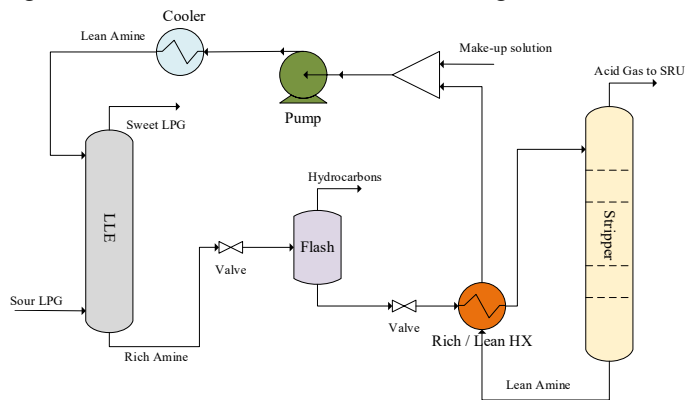


Figure 1: Process flow diagram of LPG sweetening unit.

Table 1: Sour LPG stream composition.

Component	<i>Mole fraction</i>		
	Case 1	Case 2	Case 3
Methane	0.009	0.006	0.011
Ethane	0.028	0.443	0.132
Propane	0.839	0.218	0.267
i-Butane	0.009	0.000	0.414
n-Butane	0.001	0.129	0.168
n-Pentane	0.001	0.068	0.005
n-hexane	0.000	0.060	0.000
CO <sub>2</sub>	0.019	0.077	0.000
H <sub>2</sub> S	0.093	0.000	0.002

Table 2: Input specifications used in the simulation for each LPG feed.

<i>Parameter</i>		<i>Case 1</i>	<i>Case 2</i>	<i>Case 3</i>
<b>LPG Temperature</b>	° C	38	15.6	51
<b>LPG Pressure</b>	bar	34.5	55.2	40.2
<b>LPG flow rate</b>	m <sup>3</sup> /h	20.9	159	24.4
<b>LLE Number of stages</b>		20	20	10
<b>Amine Temperature</b>	° C	54.44	45	51
<b>Amine Pressure</b>	bar	28.3	55.2	40.2
<b>Rich amine Reg. Feed Temp.</b>	° C	110	100	100
<b>Regeneration column number of stages</b>		10	10	20

### 3. Results

The simulation results such as rich loading, lean loading, and circulation rate have been summarized in table 3. Since each scenario has different lean loading and rich loading that reflect to have different circulation rates for the amine. For case 1 and case 3, DGA shows promising results while for case 2 MDEA was the best as illustrated in the next section.

Table 3: Main simulation results for the three cases.

	Case 1	Case 2	Case 3
<b>MDEA 35 wt. %</b>	Rich Loading: 0.45	Rich Loading: 0.249	Rich Loading: 0.273
	Lean loading: 0.01	Lean loading: 0.01	Lean loading: 0.002
	Circulation rate: 825 kmol/h	Circulation rate: 543 kmol/h	Circulation rate: 25 kmol/h
<b>DGA 50 wt. %</b>	Rich Loading: 0.35	Rich Loading: 0.35,	Rich Loading: 0.35
	Lean loading: 0.035	Lean loading: 0.01	Lean loading: 0.02
	Circulation rate: 600 kmol/h	Circulation rate: 263 kmol/h	Circulation rate: 12.3 kmol/h
<b>DEA 30 wt. %</b>	Rich Loading: 0.35	Rich Loading: 0.43	Rich Loading: 0.35
	Lean loading: 0.015	Lean loading: 0.03	Lean loading: 0.015
	Circulation rate: 1210 kmol/h	Circulation rate: 355 kmol/h	Circulation rate: 26 kmol/h

#### 3.1. Regeneration energy requirements

Energy analysis for the regeneration process has been studied, the results show that the developed simulation models are in good agreement with the available literature data. The result shows that the DGA is preferred in the presence of  $H_2S$  and  $CO_2$ , with 62% and 38% less energy compared to MDEA and DEA, respectively. However, for case 2 in the presence of  $CO_2$  only, MDEA has lower energy requirements with a slight improvement of approximately 3% and 11% compared to DEA and DGA. While in the case of  $H_2S$  only (Case 3), DGA has a lower energy requirement similar to case 1. Figure 2 illustrates the energy requirement of three cases for amine regeneration.

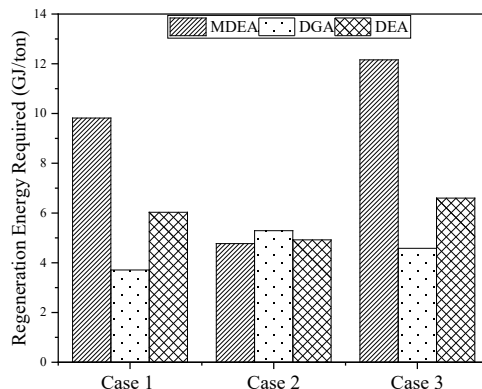


Figure 2: Energy required for the regeneration process.

### 3.2. Effect amine concentration

Acid gas removal in the LLE depends directly on amine concentrations. Figure 3 shows the effect of different amine concentrations on the acid gas in the sweet LPG and the energy required for the regeneration. It was observed that as the amine concentration increased, the acid gas in the sweet LPG decreased while there is an increase in the energy required for the regeneration. Similar trends were observed for Cases 2 and 3.

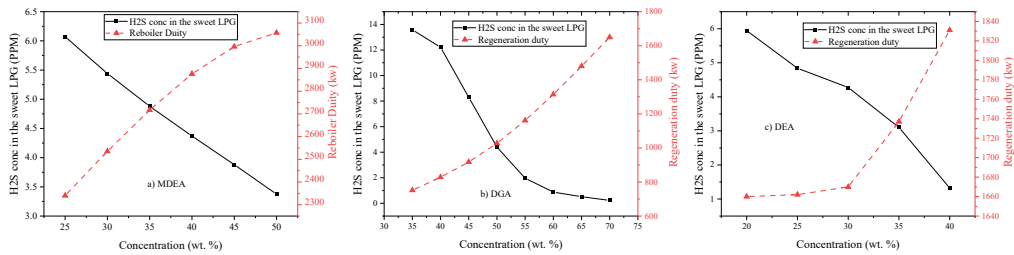


Figure 3: Effect of amines concentration

### 3.3. Effect of lean amine temperature

As shown in figure 4, the lean amine temperature influences both acid gas concentration in the sweet LPG and regeneration duty. As the lean amine temperature increased, the acid gas in the sweet LPG gradually increased due to the exothermic nature of the extraction of the acid gas with amines. However, the regeneration duty slightly decreases with the increase in the lean amine temperature. Similar trends were observed in other cases.

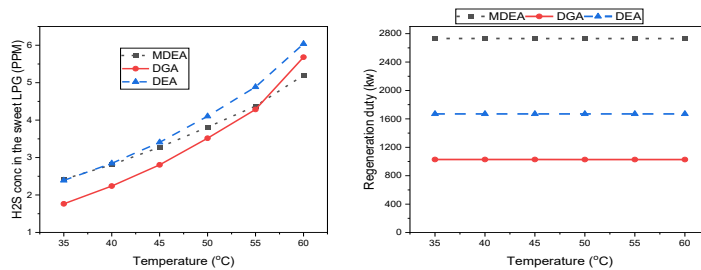


Figure 4: Effect of lean amine temperature

### 3.4. Effect of Lean amine circulation rate

One of the major parameters that directly affect the extraction of the acid gas from the LPG is the lean amine circulation rate. Figure 5 shows the effect of circulation rate on acid gas concentration and regeneration energy requirement. As the circulation rate increased, the acid gas concentration increased because of the lower residence time. On the other hand, the duty of the regeneration is also increased due to the high flow rate.

Moreover, the effect of the LPG feed temperature has been studied for Case 3. LPG feed temperature has been changed in temperature between 10 °C to 50 °C while keeping the difference between the lean amine temperature and the sour LPG feed fixed at 5 °C. The results show that the H<sub>2</sub>S concentration in the sweet LPG increases as the temperature of the feed increases while the regeneration duty slightly decreases.



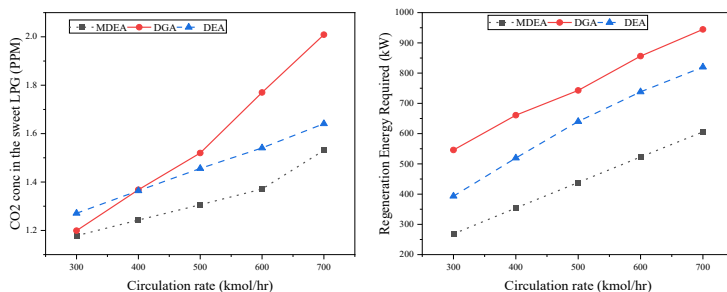


Figure 5: Effect of lean amine circulation rate

#### 4. Conclusion

In conclusion, the design specification of the sweetening of the sour LPG should be maintained in such a way that the concentration of the acid gas in the sweet LPG achieved the product specification. Sensitivity analyses have been conducted for a few of the significant process variables that could have an impact on the process performance. The results show that the acid contaminants removal from the hydrocarbons can be a competitive process to achieve the required product purities

**Acknowledgment:** The author(s) would like to acknowledge the support provided by the Deanship of Research Oversight and Coordination (DROC) at King Fahd University of Petroleum & Minerals (KFUPM).

#### References

- Global five-year LPG demand fuelled by new petchems plants, res-com; supply to overwhelm. (n.d.). Retrieved November 29, 2022, from <https://www.energyaspects.com/lpg-demand-insight>
- Jayakumar, K., Panda, R. C., & Panday, A. (2017). *A Review: State-of-the-Art LPG Sweetening Process*. 9(2), 175–206.
- Qeshta, H. J., Abuyahya, S., Pal, P., & Banat, F. (2015). Sweetening liquefied petroleum gas (LPG): Parametric sensitivity analysis using Aspen HYSYS. *Journal of Natural Gas Science and Engineering*, 26, 1011–1017. <https://doi.org/10.1016/j.jngse.2015.08.004>
- Zahid, U., Sakheta, A., & Lee, C. J. (2020). Techno-economic analysis of acid gas removal from associated and non-associated sour gas using amine blend. *International Journal of Greenhouse Gas Control*, 98(November 2019), 103078. <https://doi.org/10.1016/j.ijggc.2020.103078>
- Zoghi, A. T., Shokouhi, M., Naderi, F., Abbasghorbani, M., Fatehi, A., Pouladi, B., & Adhami, M. A. (2022). Investigation of Aqueous Diethanolamine Performance in Prediction of Hydrogen Sulfide and Carbonyl Sulfide Removal from Liquefied Propane. *Journal of Solution Chemistry*, 51(1), 84–96. <https://doi.org/10.1007/s10953-021-01131-1>

# Embedding Operating Flexibility in Process Design

Steven Sachio, Cleo Kontoravdi, Maria M. Papathanasiou

*Sargent Centre for Process Systems Engineering, Department of Chemical Engineering,  
Imperial College London, London, United Kingdom SW72AZ*

## Abstract

Inherently flexible processes are able to handle disturbances better than point-specific designed processes. However, there is a lack of systematic framework to embed operating flexibility in process design. In this work, we propose a model-based framework for integrated analysis of process flexibility and performance based on the identification and assessment of process design spaces. The proposed framework enables the identification of operating regions, quantification, within which the product and process meet the specifications. The operational flexibility and sensitivity of the process to design and operating parameters is also quantified. We demonstrate the capabilities of the framework on a Protein A chromatographic separation, used in biopharmaceutical manufacturing. We identify the feed flow rate as the most influential process parameter, while we quantify an acceptable range for the feed stream variability (concentration: 0.37 – 0.43 mg ml<sup>-1</sup> and flowrate: 0.72 – 0.88 ml min<sup>-1</sup>) for a fixed process design.

**Keywords:** computing and systems engineering, process design, chromatography

## 1. Introduction

Embedding flexibility into process design can enable more agile and sustainable processes. This is particularly true in the case of (bio-) pharmaceutical production processes, where disturbances resulting from variability in processing streams are dominant. For example, in the case of monoclonal antibodies (mAbs), an increase in the impurity content (e.g., concentration of host cell proteins, antibody aggregates or fragments) may lead to additional separation columns and/or purification cycles required to reach the desired product purity. Subsequently, this can increase the overall downstream cost and processing times, leading to less economically and environmentally efficient processes (Papathanasiou et al, 2020; Narayanan et al, 2020). In addition to this, manufacturers need to demonstrate that the processes can meet consistently product quality criteria with respect to functionality and purity for regulatory authorities to grant market authorization. For this, there is an imminent need to design processes that are adaptive to mixture variabilities and can guarantee eco-efficient, sustainable production within specifications.

Traditionally, product quality is confirmed through a Quality by Testing (QbT) approach before batch release, which incurs high experimentation costs and waiting times, delaying market release. Furthermore, manufacturers are not able to take in-process corrective actions with this approach leading to ‘off-spec’ products and hence lost batches. To this end, the Quality by Design (QbD) initiative aims to shift the focus in process development to revolve around product quality (ICH, 2011) moving away from QbT. A key step in QbD-supported initiatives is the identification of a design space defined by good candidate condition sets where the process is guaranteed to satisfy the target specifications. This approach offers greater operational flexibility, in contrast to point-

specific operation (Ding et al, 2021; Laky et al, 2022). Nonetheless, identification of a design space can lead to time- and cost- intensive exhaustive wet-lab experimentation. For this reason, computer-aided modelling tools are exploited for the pre-identification of good candidate operating condition sets (Kotidis et al, 2019; Diab et al, 2020; Kusumo et al, 2020). Here, we present a machine learning-aided design space identification framework that harnesses the potential of Sobol sequences for single-pass parallelizable sampling, allowing for rapid determination of acceptable operating conditions and space-specific sensitivity analysis.

## 2. Framework

The framework comprises of three main steps: 1) model development and problem formulation, 2) design space identification and 3) design space analysis.

### 2.1. Model Development and Problem Formulation

The process of interest is characterized using mathematical models, which can be mechanistic, hybrid, or fully data driven and are validated through experimental data. Following that, the manipulated variables (MVs), the KPIs and their associated bounds are identified. Finally, the KPIs of interest are selected and constraints with respect to both MVs and KPIs are defined.

### 2.2. Design Space Identification

First, virtual experimentation using Sobol sequence based on the formulated problem is performed to collect the dataset. Then, the constraints defined are applied to screen the dataset separating the satisfied and violated points. A collection of satisfied points, also known as a point cloud, is then set as the basis for the design space identification. The identification of the design space relies on geometrical hulls, known as alpha shapes which are a generalization of convex hulls with a parameter (alpha radius,  $\alpha_r$ ), determining the non-convexity of the resulting hull. As  $\alpha_r \rightarrow \infty$ , the formed hull is the convex hull, while as  $\alpha_r \rightarrow 0$ , the hull becomes increasingly non-convex until it reduces to the set of points at alpha radius equal to zero. The value of the alpha radius is critical for the identification of a continuous design space and is proportional to the number of points available in the space. In the case of non-linear problems, the design space shape is often non-convex and therefore at large alpha values, violated points may be found lying inside of the design space. However, at small alpha radius values, disjointed spaces start to be formed. To mitigate this, we propose two methods: namely the tolerance and the resolution support method.

#### 2.2.1. Tolerance Method

Here, we allow for a percentage of violated points to be included in the design space to allow for the identification of a design space at low resolution with less disjointed spaces. To do this, we set a tolerance value for the allowed number of violations inside the design space is set. The alpha radius should not be too small so that the hull does not reduce to thin disjointed spaces, but not too large either as this would result in many points inside the design space violating the constraints. Here, a bisection search is used to find the alpha radius. In each iteration, a check for how many violations lie inside of the design space is performed and the iteration is terminated once the number of violations tolerance is met. This method allows for the rapid identification of the design space which can be useful provided that the resulting violations in the space are analyzed further to ensure that the violation of the KPIs meets the required accuracy.

### 2.2.2. Resolution Support Method

An alternative to the tolerance method is the resolution support method. For this, we train artificial neural networks (ANNs) using the manipulated variables as inputs and the KPIs as outputs. The trained ANN is used as an interpolator to increase the resolution of the dataset. Provided that it is accurate enough, this allows for small alpha radius values to be used, resulting in very fine design spaces which capture non-convexity well.

### 2.3. Design Space Analysis

The space defined in 2.2 is used for process flexibility analysis typically based on a nominal operating point (NOP) of interest. Starting from the NOP, a uniform cube with respect to the parameter axes is formed and expanded until one of its vertices intersects the design space boundary. The resulting region is defined as the uniform acceptable operating region (AOR). This allows for the extraction of multivariate proven acceptable ranges (MPARs) of the operating parameters with respect to the NOP used, which is an underpinning element for analyzing process flexibility and controllability. This analysis is highly dependent on the number of data points used to construct the design space. Therefore, design spaces formed using low-resolution datasets are coarser. Comparing the tolerance based and resolution support design spaces, the latter results in a larger AOR, owing to the richer dataset.

### 2.4. Sobol Sensitivity Indices

In addition to the systematic flexibility analysis, the framework also offers information on the sensitivity of the monitored KPIs with respect to the manipulated variables. For this, the obtained dataset within the framework can be directly used to calculate Sobol sensitivity indices (Sobol, 2001). With the resolution support method, where ANNs are trained for interpolation, the sensitivity analysis can also be performed in any acceptable operating region without further simulation of the original model. This allows for the identification of the most impactful manipulated variables and their second order interactions.

## 3. Case Study: Protein A Chromatography

### 3.1. The System

The framework presented in section 2 is applied to the semi-continuous multicolumn capture process presented by Steinebach et al (2016). The process considers two chromatography columns operated in a semi-continuous cyclic manner. Figure 1 shows the process schematic. In step A, both columns are interconnected, and fresh feed is introduced into column I. While the breakthrough from column I is fed into column 2. Next (step B), column I is washed while column

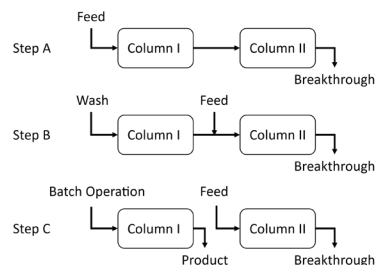


Figure 1. Process schematic adapted from Steinebach et al (2016).

II is fed the breakthrough from column I and the fresh feed simultaneously. In step C, the columns operate in batch mode with column I eluting the products and regenerating while column II is fed with the fresh feed. The process is repeated with the columns position switched. The process is modelled with Partial Differential and Algebraic Equations (PDAEs) with 50 discretization points over the column length. This results in a total of 808 ordinary differential equations (ODEs) and 812 algebraic equations. The process was validated by Steinebach et al (2016) for feed mAb concentration of 0.2 – 0.77 mg/ml and

feed flowrate of 0.5 – 1.5 ml/min. Both the process modelling and post-modelling analysis are carried out in Python.

### 3.2. Problem Formulation & Data Sampling

In this study, the objective is to investigate the process performance and flexibility of the process under the influence of feed concentration variations. The variables considered in this study are feed mAb concentration ( $c_{\text{feed}}$ ) [disturbance], feed volumetric flowrate ( $Q_{\text{feed}}$ ) [control variable], and the column switching time ( $T_{\text{switch}}$ ) [design variable]. The mAb concentration is considered as a disturbance due to variability in the performance of the upstream bioreactor. On the other hand, the volumetric flowrate is a control variable while the column switching time is a design variable. The manipulated variable bounds for the study are based on the model validation range (Table 1). The KPIs considered are yield and productivity. Following standard practice, minimum constraints of 99% and 4 mg ml<sup>-1</sup> h<sup>-1</sup> are assumed for yield and productivity, respectively. The formulated problem is used to collect a total of 4096 points via quasi-random Sobol sampling.

Table 1. Manipulated variable bounds used in the study based on the validated model range (Steinebach et al., 2019).

Manipulated Variable	Lower Bound	Upper Bound
$c_{\text{feed}}$ (mg ml <sup>-1</sup> )	0.21	0.63
$Q_{\text{feed}}$ (ml min <sup>-1</sup> )	0.5	1.5
$T_{\text{switch}}$ (min)	40	120

## 4. Results and Discussion

### 4.1. Design Space Identification

The collected dataset is screened using the constraints (> 99% yield and > 4 mg ml<sup>-1</sup> h<sup>-1</sup> productivity) specified to classify satisfied and violated points. Figures 2 (A), (B) and (C) illustrate the knowledge space (unconstrained), and design spaces (constrained) using the tolerance and the resolution support method, respectively. The proposed framework enables quantification of the volume of the design space including good candidate operating points. As discussed previously, in nonlinear problems as the one presented here, the design space may include points that violate one or more of the predefined constraints. A tolerance of 1% resulted in 12 out of 1281 points (0.94%) (large blue crosses) which violated the constraints inside of the design space. On the other hand, the resolution support method arrives at a zero-violation design space with a smoother surface. The framework also allows for the quantification of the size of the design space for exact comparison. In this case, the size is the volume of the space defined by the manipulated variables, hence the units are in mg ( $\frac{\text{mg}}{\text{ml}} \times \frac{\text{ml}}{\text{min}} \times \text{min} = \text{mg}$ ). The design space size of the tolerance-based method is 5.98 mg, while that of the resolution support method is 8.43 mg (41% larger). This is due to

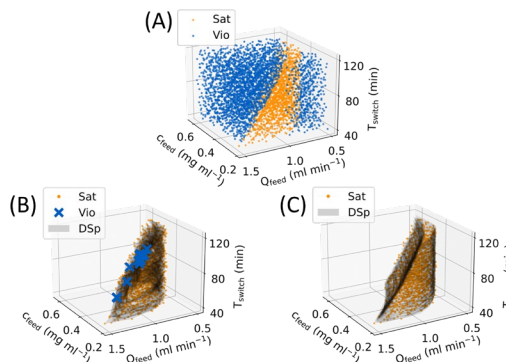


Figure 2. (A) Point cloud classification by applying constraints, (B) tolerance-based method, (C) resolution support method.

the higher resolution allowing for the use of a finer alpha radius which can capture the non-convexity of the problem better.

4.2. Design Space Analysis

The operational flexibility is then assessed by analyzing the identified design space. For this, we use the resolution support method (Figure 2 A(C)) as it provides the design space without any violated points. For the analysis, we choose two arbitrary nominal points of interest, as shown in Figure 3. The first nominal point lies in the center of the design space, while the second point is in the higher productivity region of the design space. The comparison between the two designs is detailed in Table 2.

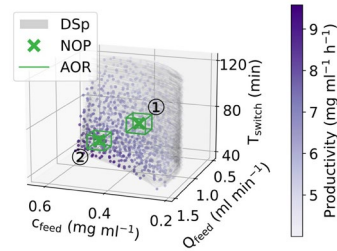


Figure 3. Good candidate operating point analysis with resolution support.

Table 2. Detailed comparison of NOP1 & NOP2

	NOP1	NOP2
AOR size (mg)	0.134	0.116
MPAR $c_{feed}$ (mg ml <sup>-1</sup> )	0.40 ± 0.033	0.55 ± 0.032
MPAR $Q_{feed}$ (ml min <sup>-1</sup> )	0.80 ± 0.080	0.75 ± 0.076
MPAR $T_{switch}$ (min)	70.0 ± 6.34	50.0 ± 6.05
Average Productivity (mg ml <sup>-1</sup> h <sup>-1</sup> )	5.91	7.54

The size of NOP1 is 15% larger than that of NOP2 indicating greater flexibility. This is further expressed in the respective MPARs of the manipulated variables which reduce by 5% going from NOP1 to NOP2. The average expected productivity when operating at NOP2 is 27% larger than that of NOP1. Using this framework, such trade-offs between flexibility and performance gain can be quantified systematically. Quantification of such metrics is essential for the design of robust and controllable processes that are crucial in biopharmaceutical manufacturing, bounded by stringent regulations.

4.3. Sobol Sensitivity Indices

We further investigate and quantify the impact of the studied variables and disturbances on the chosen KPIs. This is achieved through calculation of the Sobol sensitivity indices. For this, we use the dataset as generated in section 2. The Sobol sensitivity indices highly depends on the chosen bounds of the manipulated variables and the non-linearity of input output relationship within the chosen bounds. Hence, it is vital to be able to identify the sensitivity indices at different spaces especially for integration process units and process control. Figure 4 illustrates the Sobol sensitivity indices calculated for the acceptable

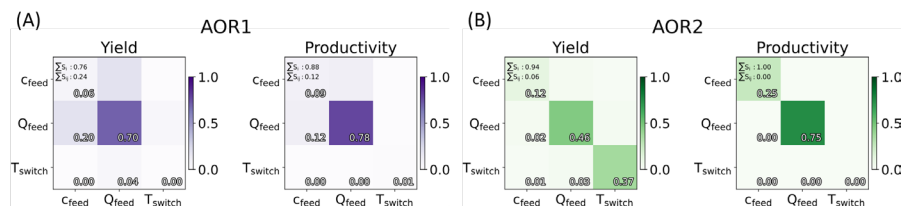


Figure 4. Sobol indices in (A): AOR1 & (B): AOR2.

operating region of NOP1 (AOR1) and NOP2 (AOR2). Whereby first order interactions are observed by looking at the main diagonals and second order interactions on the off-diagonal elements. First order indices refer to how the fractional contribution of a variable

on the variance of the KPI while the second order indices show the contribution of variable-variable interactions to the variance of the KPI. We assume that any interaction with an index below 10% is non-significant. There is a significant difference in the indices within AOR1 and AOR2 (Figure 4 (A) & (B)). The results indicate that for AOR1, the feed flowrate has the largest impact on both KPIs with indices larger than 0.70. AOR2 illustrates a different trend, where the switching time plays a bigger role in determining the yield of the product (from  $S_i = 0.00$  in AOR1 to  $S_i = 0.37$  in AOR2). On the other hand, the second order interactions observed between the feed concentration and flowrate in productivity of AOR1 ( $S_{ij} = 0.20$ ) are not present in AOR2 ( $S_{ij} = 0.02$ ). This is due to the nonlinear relationship between the KPIs and the manipulated variables which is a challenge for identifying the design space with zero violations. This framework allows for the unravelling of such complex relationships in a computationally efficient manner.

## 5. Conclusion

We presented a novel framework for the identification and analysis of process design spaces. The capabilities of the presented framework were illustrated through its application on a downstream purification unit used in biopharmaceutical production. We considered design (switching time) and operating decisions (flow rate), as well as disturbances (feed composition) simultaneously. Based on the process feasibility constraints and the target KPIs, we identified and quantified a design space that includes good candidate condition sets, predicted to meet the KPIs of interest. We further demonstrated how this space can be used for the evaluation and quantitative comparison of different operating conditions and their flexibility. This can assist manufacturers by offering a measurable estimate of the variability that a process at hand can handle without any further mediation and remaining within spec. In addition to this, we demonstrated how the framework enables space-specific sensitivity analyses without additional high-fidelity model simulation which is crucial for process control and integration. Importantly, the presented framework is generic, owing itself to the assistance of simultaneous design and flexibility assessment of different processes. In the future, the advantages of the framework will be further enhanced to incorporate formal optimization and control methodologies that can benefit from the process knowledge and good candidate initial guesses generated by the presented methodology.

## References

- Papathanasiou & Kontoravdi (2020): "Engineering challenges in therapeutic protein product and process design". *Current Opinion in Chemical Engineering* (27).
- ICH (2011): "ICH Harmonised Tripartite Guideline: Pharmaceutical Development Q8, Step 4".
- Ding and Ierapetritou (2021): "A novel framework of surrogate-based feasibility analysis for establishing design space of twin-column continuous chromatography". *Int. J. Pharm.* (609).
- Laky et al (2022): "Determination of probabilistic design spaces in the hybrid manufacture of an active pharmaceutical ingredient using PharmaPy". *Computer Aided Chem. Eng.* (49).
- Kotidis et al. (2019): "Constrained global sensitivity analysis for bioprocess design space identification". *Computers and Chem. Eng.* (125).
- Diab and Gerogiorgis (2020): "Design Space Identification and Visualization for Continuous Pharmaceutical Manufacturing". *Journal of Pharmaceutics* (235).
- Kusumo et al. (2020): "Bayesian approach to probabilistic design space characterization: A nested sampling strategy". *Ind. & Eng. Chem. Research* (59).
- Narayanan et al. (2020): "Bioprocessing in the Digital Age: The Role of Process models". *Biotechnol J* (15).
- Steinebach et al. (2016) "Model based adaptive control of a continuous capture process for monoclonal antibodies production". *J Chromatogr A.* (1444).

# Transfer learning for process design with reinforcement learning

Qinghe Gao<sup>a</sup>, Haoyu Yang<sup>a</sup>, Shachi M. Shanbhag<sup>a</sup>, Artur M. Schweidtmann<sup>a,\*</sup>

<sup>a</sup>*Process Intelligence Research Team, Department of Chemical Engineering, Delft University of Technology, Van der Maasweg 9, Delft 2629 HZ, The NETHERLANDS*

<sup>\*</sup>*Corresponding author. Email: [a.schweidtmann@tudelft.nl](mailto:a.schweidtmann@tudelft.nl)*

## Abstract

Process design is a creative task that is currently performed manually by engineers. Artificial intelligence provides new potential to facilitate process design. Specifically, reinforcement learning (RL) has shown some success in automating process design by integrating data-driven models that learn to build process flowsheets with process simulation in an iterative design process. However, one major challenge in the learning process is that the RL agent demands numerous process simulations in rigorous process simulators, thereby requiring long simulation times and expensive computational power. Therefore, typically short-cut simulation methods are employed to accelerate the learning process. Short-cut methods can, however, lead to inaccurate results. We thus propose to utilize transfer learning for process design with RL in combination with rigorous simulation methods. Transfer learning is an established approach from machine learning that stores knowledge gained while solving one problem and reuses this information on a different target domain. We integrate transfer learning in our RL framework for process design and apply it to an illustrative case study comprising equilibrium reactions, azeotropic separation, and recycles, our method can design economically feasible flowsheets with stable interaction with DWSIM. Our results show that transfer learning enables RL to economically design feasible flowsheets with DWSIM, resulting in a flowsheet with an 8% higher revenue. And the learning time can be reduced by a factor of 2.

**Keywords:** Reinforcement learning, process design, transfer learning

## 1. Introduction

The transition of chemical engineering to a sustainable and circular future requires new methods of process design (Fantke et al., 2021). Currently, methods for process design are mainly manual work with long development times and superstructure methods are also limited to predefined process configurations (Chen et al. 2017, Mitsos et al. 2018). Recently, reinforcement learning (RL), a branch of machine learning (ML), has shown promising results in process design (Midgley et al., 2020, Khan et al., 2020, Göttl et al., 2021, Stops et al., 2022, Kalmthout et al., 2022). One major challenge in RL for process design is the training process as it is trial-and-error based. Thereby, the learning process typically requires a large number of process simulations, which demands expensive computational power. Previous work (Khan et al., 2020, Göttl et al., 2021, Stops et al., 2022) mostly leverages short-cut process simulation methods to simulate the processes efficiently, which can lead to inaccurate results. Recent works employ rigorous process simulators such as COCO and Aspen Plus with rigorous models (Midgley et al., 2020, Kalmthout et al., 2022). Nevertheless, the problem of long simulation times hinders further developments (Kalmthout et al., 2022).



We propose to utilize transfer learning in process design with RL to facilitate the effectiveness and efficiency of the learning process. Transfer learning is a technique to improve learning performance by transferring knowledge from different but relevant domains to the target domain (Zhu et al., 2020). Adapting the concept of transfer learning for process design with RL, we first pre-train our recently proposed RL agent on a short-cut process simulation from our previous work (Stops et al., 2022). Then, we transfer the pre-trained agent to a rigorous process simulator DWSIM for further training. Finally, we illustrate the impact of transfer learning through one case study.

## 2. Methods

The RL problem can be formulated as a Markov decision process (MDP):  $M = \{S, A, T, R\}$ , which includes states  $s \in S$ , actions  $a \in A$ , transitions  $T: S \times A$ , and reward functions  $R$ . The agent aims to maximize the reward in the environment by iteratively taking action, evaluating the current reward, and updating the states. Specifically, in process synthesis tasks, states correspond to flowsheets. Actions are composed of two parts: Discrete and continuous actions. The discrete actions include the selections of a unit operation and its location in the flowsheet. The continuous actions define the design and operation variables of the corresponding unit operation. After the agent has performed actions, the states are updated. Based on the current state, a reward is calculated by the environment, e.g., the process simulation software, and returned to the agent as feedback on its actions. For process design, this reward is typically a design goal such as the process revenue. By repeating the process of performing actions and receiving rewards multiple times, the agent is trained to perform actions that result in a higher reward, corresponding to flowsheets with higher revenue.

### 2.1. Agent and environment

We adapt the RL framework from our previous work (Stops et al., 2022), where the states are presented as directed flowsheet graphs. Within the directed flowsheet graphs, nodes represent the unit operations and edges correspond to the process streams. Each node and edge is assigned a feature vector, respectively, storing information about the unit operation, e.g., type or size, and stream, e.g., thermodynamic data or flow rate of the stream. Furthermore, the agent architecture consists of three major parts: Graph encoder, actor networks, and critic networks. The graph encoder takes current flowsheet graphs as input and utilizes graph convolution in GNNs to learn information about the flowsheet graphs, in form of a vector representation, also referred to as flowsheet fingerprint. Actor networks are responsible for taking actions during the training process. There are three action levels: Selecting an open stream, selecting a unit operation, and selecting a corresponding design variable. Moreover, taking the flowsheet fingerprint as input, critic networks are used to estimate the reward of actions, and then actor networks will learn to take actions with the highest estimated reward. Specifically, the reward is calculated by using the DWSIM process simulator (Medeiros et al., 2018). The RL framework is implemented in Python including an interface for the agent to actively interact with DWSIM.

### 2.2. Transfer learning

We extend our RL framework by transfer learning. Specifically, we add a pre-training phase to the training of the agent. In the pre-training phase, we use a short-cut process simulation environment (Stops et al., 2022) and train the agent over 10000 episodes. Then, we transfer the pre-trained agent to a fine-tuning phase in which the agent is trained against a rigorous process simulator DWSIM for further 15000 training episodes. For the comparison, we directly train another agent with DWSIM environment over 15000

episodes. Note that in each episode, the agent generates a complete flowsheet. Both pre-training and fine-tuning processes are adapted from Proximal Policy Optimization (PPO) by OpenAI (Schulman et al., 2017). Then, the agent architecture is updated by gradient descent for the total loss function derived from summing up losses of actor networks, loss of critic networks, and corresponding entropies.

The agents are trained on a Windows server with a 3.5 GHz 24 cores Intel(R) Xeon(R) W-2265 CPU, NVIDIA GeForce RTX 3090 GPU and 64 GB memory.

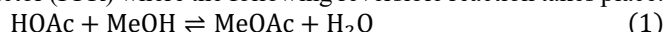
### 3. Illustrative case study

The production of methyl acetate (MeOAc) is chosen as an illustrative case study.

#### 3.1. Process simulation

The short-cut process methods for pre-training are illustrated in our previous work (Stops et al., 2022). Here, the process simulation with DWSIM is introduced. In this case study, the agent can choose reactors, distillation columns, and heat exchangers as unit operations. Besides, the agent can also decide to add recycles or claim open streams as products. The types of unit operations and corresponding design variables are defined as follows.

**Reactor** is deployed to convert reactants to the desired product (MeOAc). The reactor is modeled as a plug flow reactor (PFR) where the following reversible reaction takes place:



For operational simplicity, the reactor is simulated isothermally, in which the temperature is kept constant regarding the inlet stream temperature. Besides, catalyst loading is not considered. The reaction kinetics are obtained from (Xu et al, 1996), with the equilibrium being related to temperature. The reactor cross-sectional area is determined by the relation:  $N/10$ , where  $N$  is the inlet molar flow (Stops et al., 2022). The design variable is the reactor length, which will be determined by the agent in the third-level continuous decision process. The range is from 3 to 10 m.

**Distillation column** is applied to separate MeOAc from the quaternary system. Rigorous distillation columns are used instead of shortcut columns in the previous work (Stops et al., 2022) to account for more realistic factors such as intermolecular interactions. The rigorous column models provide multiple possible choices of design parameters, from which the distillate to feed ratio (D/F) is selected as the third-level decision. Other adjustable parameters such as the number of stages and reflux ratio are set as fixed values (35 and 1.5, respectively). The D/F ratio can range from 0.4 to 0.6.

**Heat exchanger** is a DWSIM heater model. In the proposed framework, the heat exchanger is simulated based on the outlet temperature which is determined by the third-level decision. A temperature range from 278.15 K to 330.05 K is applied, where the upper limit refers to the lowest boiling point of the components which is MeOAc.

**Recycle** action consists of additional units including a splitter and a mixer. Firstly, the process stream is split into a recycle stream and a purge. Secondly, the recycled stream is merged with the selected feed with a mixer. Thereby, the split ratio is the third-level decision of the agent, which lies in the range of 0.1 to 0.9 for pragmatic consideration.

#### 3.2. Reward

The reward determines the economic viability of generated flowsheets and teaches the agent to take feasible actions. First, a reward of 0 € is given when the incomplete flowsheets can converge after every single action, because the economic value is difficult to assess for an incomplete flowsheet. Second, whenever the agent fails the simulation by taking infeasible actions, the episode will be terminated immediately and a negative

reward -10M€ is given. Finally, when a flowsheet is completed, we calculate the reward according to Equation 2:

$$r = \sum P_{product} - \sum C_{feed} - \sum (C_{operation} + 0.15 \cdot C_{invest})_{units} \quad (2)$$

where  $P_{product}$  is the revenue of the sold product (Seider et al., 2008),  $C_{feed}$  is the costs of feeds,  $C_{operation}$  is the operation costs (Smith, 2016) and  $C_{invest}$  is the total capital investment which is multiplied by factor 0.15 (Seider et al., 2008). In the case of negative rewards, a reduction factor 10 is applied to encourage the exploration of the agent.

#### 4. Results and discussion

Figure 1 shows the learning curves for the agents with and without transfer learning. The scores represent the moving average rewards, i.e., the economic viability of the flowsheets, over 100 episodes. The training for the agent without transfer learning took 72 hours over 10000 episodes. For the agent with transfer learning, the pre-training took 2 hours over 10000 episodes and the further training took 72 hours over 10000 episodes. During the first 3500 episodes, the agent without transfer learning generates predominantly infeasible lengthy flowsheets, resulting in the learning curves rising slowly. In fact, due to the complexity of the design space, the agent has difficulty in learning from the previous failed flowsheets, leading to negative scores. After 3500 episodes, the agent mainly produces flowsheets with positive scores, which indicates that the designed processes are economically viable. Besides, within the training episodes, the learning curve slowly converges and reaches maximally to about 42. In comparison, the agent with transfer learning shows a quicker learning process. At the beginning of the learning process, the agent is able to mostly produce positive scores, and then the learning curve rises steeply. This demonstrates that the agents successfully leverage the pre-trained information from a short-cut process environment to make favorable decisions even in the early training stages. After about 4500 episodes, the score begins to fluctuate between 30 and 40 and maximumly reaches 46 which is higher than the agent without transfer learning.

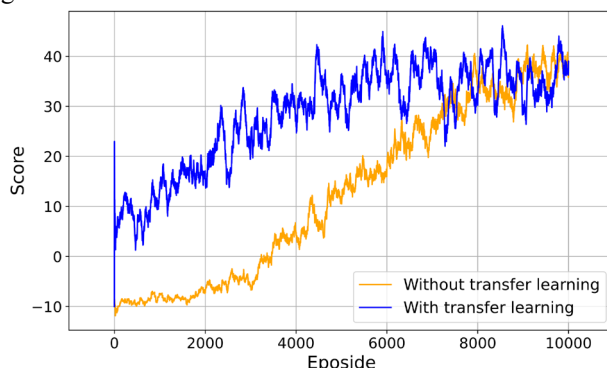


Figure 1: Learning curves of the agent with and without transfer learning. The blue line depicts the learning curve with transfer learning and the orange line indicates the learning curves without transfer learning. The scores are moving average rewards over 100 episodes and each learning curve runs over 10000 episodes.

Figure 2 displays the best flowsheet generated by the transfer learning agent and the continuous design variables are shown in Table 1. First, the feed (F1) is fed directly into three consecutive reactors (R1, R2, and R3) where the MeOAC and H<sub>2</sub>O are produced by

the esterification of HOAC and MeOH. Then, the resulting quaternary mixture is split up in a column (C1). The mixture is distilled from the top part of the column and sent to one heat exchanger (Hex1) to get the product (P1), containing enriched MeOAc and residues of HOAC and H<sub>2</sub>O. The bottom product of the first column is further split up into the second column (C2) to produce pure H<sub>2</sub>O in the distillate (P2) and a mixture of MeOH and H<sub>2</sub>O in the third product stream (P3). And 90 % of the bottom product is recycled and mixed back into the feed.

While the resulting flowsheet has a positive reward, it is still far from a realistic engineering solution and future research is required. In particular, there are three major issues observed in the optimal flowsheet solution: (1) In industrial applications, MeOAc is primarily produced in reactive distillation (Huss et al., 2003). As our agent does not include reactive distillation as a unit operation, this cannot be identified. In future work, intensified unit operations can be added to our framework or the agent could operate on a phenomena level rather than a unit operation level. (2) The best flowsheet generated by the agent in this work contains three consecutive PFRs. The reason is that the length of a single PFR is limited to 10 m, which is not sufficient to finish the reaction. Therefore, the agent learns to choose multiple PFRs to fulfill the reaction and maximize the product. (3) One unnecessary heat exchanger is added after the distillation column C1 and before the product P1. We believe this is due to the small operations cost of the heat exchangers and the minor impact on the overall reward. Future research should further investigate possible mitigation strategies such as longer training or further hyperparameter tuning.

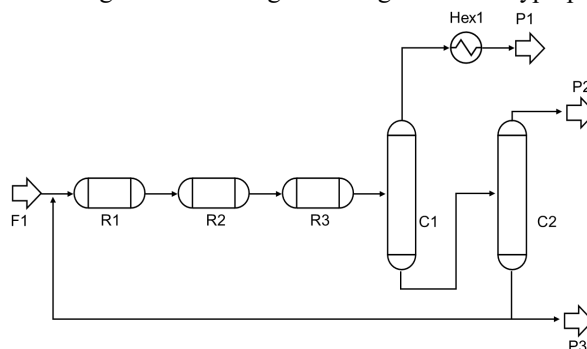


Figure 2: Best flowsheet generated by the transfer learning agent. First, MeOAc and H<sub>2</sub>O are produced from the feed (F1) in three consecutive reactors (R1, R2, and R3). Then the mixture is separated in the first column (C1). The first product (P1) is enriched with MeOAc but also contains residues of HOAC and H<sub>2</sub>O after one heat exchanger (Hex1). Then the bottom mixture of MeOH and H<sub>2</sub>O is further separated in the second column (C2). Pure H<sub>2</sub>O is in the distillate (P2) and the third product (P3) is the mixture of MeOH and H<sub>2</sub>O. And 90% of P3 is recycled and mixed with the feed stream.

Table 1. Design variables selected by the agent with transfer learning in the best flowsheet.

Unit operation	Design variable	Unit	Best run
Reactor 1 (R1)	Length	m	9.42
Reactor 2 (R2)	Length	m	9.25
Reactor 3 (R3)	Length	m	8.38
Column 1 (C1)	Distillate to feed ratio	-	0.58
Column 2 (C2)	Distillate to feed ratio	-	0.4
Heat exchanger 1 (Hex1)	Outlet temperature	K	315
Recycle	Recycled ratio	-	0.9

## 5. Conclusion

We propose to deploy the transfer learning for process design in RL to accelerate the learning process of the agent. The GNNs-based agent is first pre-trained with a short-cut simulation environment and then transferred to the rigorous process simulator environment for further training. In the illustrative case study, the agent is able to design economically feasible flowsheets in the process simulator DWSIM environment. Furthermore, the learning curves demonstrate that transfer learning indeed improves the efficiency of the learning process significantly and thus can be used to reduce the overall training time significantly. This work thus demonstrates that transfer learning can accelerate the learning process of graph-based RL with rigorous process simulator environments.

## References

- Chen, Q., & Grossmann, I. E. (2017). Recent developments and challenges in optimization-based process synthesis. *Annu. Rev. Chem. Biomol. Eng.* 8(1), 249-283.
- Fantke, P., Cinquemani, C., Yaseneva, P., De Mello, J., Schwabe, H., Ebeling, B., & Lapkin, A. A. (2021). Transition to sustainable chemistry through digitalization. *Chem.* 7(11), 2866-2882.
- Göttl, Q., Grimm, D. G., & Burger, J. (2022). Automated synthesis of steady-state continuous processes using reinforcement learning. *Frontiers of Chemical Science and Engineering*, 16(2), 288-302.
- Huss, R. S., Chen, F., Malone, M. F., & Doherty, M. F. (2003). Reactive distillation for methyl acetate production. *Computers & chemical engineering*, 27(12), 1855-1866
- Khan, A., & Lapkin, A. (2020). Searching for optimal process routes: a reinforcement learning approach. *Computers & Chemical Engineering*, 141, 107027.
- Midgley, L. I. (2020). Deep Reinforcement Learning for Process Synthesis. arXiv preprint arXiv:2009.13265.
- Mitsos, A., Asprion, N., Floudas, C. A., Bortz, M., Baldea, M., Bonvin, D., ... & Schäfer, P. (2018). Challenges in process optimization for new feedstocks and energy sources. *Computers & Chemical Engineering*, 113, 209-221.
- Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347.
- Seider WD, Lewin DR, Seader JD, Widagdo S, Gani R, Ng KM. *Product and process design 16 principles: Synthesis, analysis, and evaluation* (4th edition). Hoboken, NJ: Wiley . 2017
- Smith R. *Chemical process: Design and integration* (2nd edition). Chichester, West Sussex, United Kingdom: Wiley . 2016.
- Stops, L., Leenhouts, R., Gao, Q., & Schweidtmann, A. M. (2022). Flowsheet generation through hierarchical reinforcement learning and graph neural networks. *AIChE Journal*, e17938.
- van Kalmthout, S. C., Midgley, L. I., & Franke, M. B. (2022). Synthesis of separation processes with reinforcement learning. arXiv preprint arXiv:2211.04327.
- Xu, Z. P., & Chuang, K. T. (1996). Kinetics of acetic acid esterification over ion exchange catalysts. *The Canadian Journal of chemical engineering*, 74(4), 493-500.
- Zhu, Z., Lin, K., & Zhou, J. (2020). Transfer learning in deep reinforcement learning: A survey. arXiv preprint arXiv:2009.07888.

# Data augmentation for machine learning of chemical process flowsheets

Lukas Schulze Balhorn<sup>a</sup>, Edwin Hirtreiter<sup>a</sup>, Lynn Luderer<sup>a</sup>, Artur M. Schweidtmann<sup>a,\*</sup>

<sup>a</sup> *Process Intelligence Research, Department of Chemical Engineering, Delft University of Technology, Van der Maasweg 9, Delft 2629 HZ, The Netherlands*

*\*Corresponding author. Email: [a.schweidtmann@tudelft.nl](mailto:a.schweidtmann@tudelft.nl)*

## Abstract

Artificial intelligence has great potential for accelerating the design and engineering of chemical processes. Recently, we have shown that transformer-based language models can learn to auto-complete chemical process flowsheets using the SFILES 2.0 string notation. Also, we showed that language translation models can be used to translate Process Flow Diagrams (PFDs) into Process and Instrumentation Diagrams (P&IDs). However, artificial intelligence methods require big data and flowsheet data is currently limited. To mitigate this challenge of limited data, we propose a new data augmentation methodology for flowsheet data that is represented in the SFILES 2.0 notation. We show that the proposed data augmentation improves the performance of artificial intelligence-based process design models. In our case study flowsheet data augmentation improved the prediction uncertainty of the flowsheet autocompletion model by 14.7%. In the future, our flowsheet data augmentation can be used for other machine learning algorithms on chemical process flowsheets that are based on SFILES notation.

**Keywords:** Data Augmentation, Flowsheet Autocompletion, SFILES, Transformers

## 1. Introduction

The design of a flowsheet topology is an important step in early process synthesis. This step consists of selecting and arranging unit operations for a chemical process. Artificial intelligence (AI) methods have the potential to learn from previous flowsheets and support engineers in process development (Hirtreiter et al., 2022; Oeing et al., 2022; Schweidtmann, 2022; Vogel et al., 2023). For instance, Vogel et al. (2023) proposed an algorithm for the autocompletion of flowsheets. This autocompletion algorithm is inspired by text-autocompletion from natural language processing (NLP) that is based on generative transformer models (Radford et al., 2019). In addition, Hirtreiter et al. (2022) showed that the prediction of control structure elements from Process Flow Diagrams (PFDs) can be interpreted as a translation task between PFDs and Process and Instrumentation Diagrams (P&IDs). Hence, they deployed a sequence-to-sequence transformer architecture which is commonly used for translation of text between different languages. These flowsheet transformers rely on machine-readable flowsheet representations.

To represent flowsheets in a machine-readable format, we depict them as graphs or as text using unique, i.e., canonical, SFILES 2.0 strings (Vogel et al., 2022b). In general, flowsheets are drawings of chemical processes. Chemical engineers use flowsheets for the communication, planning, operation, simulation, and construction of these processes. An example flowsheet is given in Figure 1.

An intuitive way to represent flowsheets is via graphs with unit operations as nodes and stream connections as directed edges. Besides the graph representation, flowsheets can also be represented as strings. D’Anterrosches (2005) introduced the Simplified Flowsheet Input-Line Entry-System (SFILES) notation, which we recently extended to include control structures and other features in (Vogel et al., 2022b). When creating the SFILES, we traverse the graph by starting at an input node and following the stream direction until we reach a product node or a recycle. In case the stream branches at a node, i.e., a splitter, we need to decide which stream to follow first. To determine the order of the branches in the linear string, the SFILES algorithm assigns each node a unique rank. The SFILES string for the flowsheet from Figure 1 is given by:

$$(\text{raw})(\text{hex})\{1\}(\text{r})<\&|(\text{raw})(\text{pp})\&|(\text{mix})<1(\text{v})(\text{dist})[\{\text{tout}\}(\text{prod})]\{\text{bout}\}(\text{splt})1(\text{prod})\text{n}|(\text{raw})(\text{hex})\{1\}(\text{prod}).$$

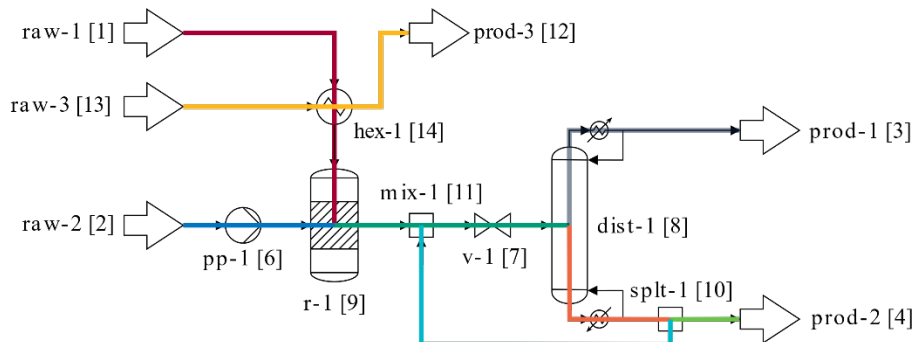
While AI models require big training data, machine-readable flowsheet data is typically limited. The reason for limited data is that flowsheets are mainly depicted as images and therefore not machine-readable (Schweidtmann, 2022). Recently, we propose to automatically find flowsheet images in literature (Schulze Balhorn et al., 2022) and to make them machine-readable via computer vision (Theisen et al., 2023). However, these methods need to be implemented at a large scale. In addition, the majority of flowsheets are not publicly available due to company’s intellectual property protection. Chemical process datasets with machine-readable flowsheets are rare and often contain only dozens of flowsheets (Hirtreiter et al., 2022; Oeing et al., 2022; Vogel et al., 2023; Zhang et al., 2018; Zheng et al., 2022). However, AI methods like transformers usually require big training data.

One promising approach to overcome limited training data for artificial intelligence is data augmentation. Data augmentation builds on the idea to generate additional artificial training data by masking, modifying, perturbing the available original training data at hand. For example, in computer vision, data augmentation is an established method to improve the model performance by adding modified copies of already existing images (Shorten & Khoshgoftaar, 2019). However, no data augmentation method for flowsheets exists yet. Thus, a data augmentation method is needed which adds modified copies of already existing flowsheets to mitigate the issue of limited available flowsheet data.

We propose a novel data augmentation method for process flowsheets. Specifically, we use a text-based augmentation method for SFILES which is inspired by the augmentation of SMILES strings for molecules (Bjerrum, 2017). Our approach is to randomize the branching decisions in the SFILES generation. We demonstrate the proposed flowsheet augmentation method in the context of a flowsheet autocompletion model from (Vogel et al., 2023).

## 2. Data augmentation methodology

To augment the flowsheet data sets, we modify the branching decision in the SFILES generation algorithm to create multiple SFILES strings representing the same flowsheet. In the case of determining canonical SFILES the branching decisions are made by assigning every node of the graph to a unique rank. Hence, canonical SFILES are a unique mapping of a flowsheet graph to a string. When generating augmented SFILES, the branching decisions are made randomly, resulting in a non-canonical form. The difference lies in the order of branches in the linear string. The resulting augmented SFILES contain the identical information as the canonical SFILES, thus describing the same process flowsheet. Hence, all augmentations can be translated back to the original canonical SFILES. In Figure 1 we show the augmentation of the example flowsheet.



### canonical SFILES

```
(raw)(hex){1};(r)-&(raw)(pp)&(mix)<1(v)(dist)[{tout};(prod)]{bout}(spl)1(prod)n(raw)(hex){1};(prod)
```

### augmentation SFILES

```
(raw)(hex){1};(r)-&(raw)(pp)&(mix)<1(v)(dist)[{bout}(spl)1(prod)]{tout}(prod)n(raw)(hex){1};(prod)
```

```
(raw)(hex){1};(r)-&(raw)(pp)&(mix)<1(v)(dist)[{bout}(spl)1[(prod)]]{tout}(prod)n(raw)(hex){1};(prod)
```

```
(raw)(hex){1};(r)-&(raw)(pp)&(mix)<1(v)(dist)[{tout}(prod)]{bout}(spl)1[(prod)n(raw)(hex){1};(prod)
```

Figure 1: Augmentation example for the flowsheet from Figure 1. The node rank is given in square brackets after the node name.

During augmentation, only the uniqueness of the SFILES representation is lost while the full flowsheet topology information is preserved. Notably, we do not change the order of the input branches during data augmentation. Otherwise, flowsheets with disconnected sub-graphs cannot be translated back to the original canonical form. This is for example important for independent processes with heat integration, such as  $(raw)(hex)\{1\}(prod)$  in the example flowsheet from Figure 1.

The number of augmented SFILES that can be derived from a single flowsheet graph is limited. Specifically, the number of potential augmented SFILES depends on the number of available branching points for a given flowsheet, with more branching points leading to an exponential increase of augmentation possibilities. For example, the flowsheet in Figure 1 contains two branching points. This results in three augmented SFILES representations and four SFILES representations in total.

The two branching points in this case are the distillation column  $dist-1$  and the splitter after the distillation column  $spl-1$ . At the first branching point, we switch the order of the top and bottom outlet streams. The second augmentation affects only the bottom product stream  $prod-2$ . Recycles without unit operations always appear directly after the splitting unit, here that leads to  $(spl-1)1$  in all cases. In case the product branch is visited before the recycle branch, additional squared brackets around  $prod-2$  are used. To make the methodology for flowsheet augmentation openly accessible, we include it in our public SFILES 2.0 Github repository (Vogel et al., 2022a).

## 3. Case study and Results

### 3.1. Data and Data augmentation

We use two SFILES datasets which were created by (Vogel et al., 2023). Firstly, we use their proposed flowsheet generator to generate a large-scale dataset of about 8,000



artificial flowsheets. Here, we can flexibly scale the dataset size. Secondly, we use a dataset made from 223 Aspen and DWSIM chemical process simulations. We call this the real dataset. Before starting the training runs, we split each dataset into a training dataset, a validation dataset, and a test dataset. For the two training datasets we created a maximum of five augmentations for each SFILES, which roughly increases the dataset five-fold. By limiting the number of augmentations to five, we ensure that larger flowsheets are not over-represented.

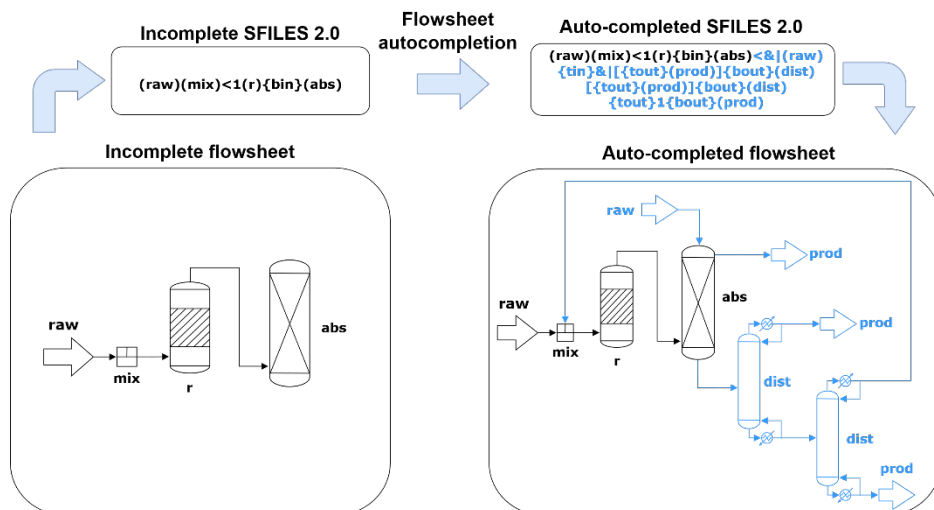


Figure 2: Example prediction of the flowsheet autocompletion model. Figure adapted from (Vogel et al., 2023).

### 3.2. Flowsheet autocompletion

We consider a process design case study and use the flowsheet autocompletion model from (Vogel et al., 2023). The objective of the model is to support chemical engineers in the design of a new process topology. It can suggest the following unit operation for an incomplete flowsheet, similar to sentence completion in messenger apps. Specifically, the start of a SFILES is given and the autocompletion model predicts how the sequence ends by iteratively predicting the following building block, token, of the SFILES (Figure 2). It should be highlighted that the prediction is only based on the process topology and does not consider operating points, components, and material flows. The autocompletion model is built on the transformer architecture in a decoder-only approach (Radford et al., 2019). Because transformer models are very data-intensive, the generated dataset is used to pretrain the flowsheet autocompletion. The real dataset is then used to fine tune the flowsheet autocompletion. For a more detailed description of the model we refer to (Vogel et al., 2023).

To study the effectiveness of data augmentation we retrain the model for flowsheet autocompletion from (Vogel et al., 2023). Specifically, we train the model both with augmented and non-augmented flowsheet data. Overall, we consider three different training scenarios. First, we train the model with the non-augmented generated dataset and fine-tune this model with the real dataset (i). This model is used to reproduce the results from (Vogel et al., 2023). Secondly, we augment the real data and use them to fine-tune the pretraining model, resulting in model (ii). We thus only alter the fine-tuning. Finally, we train the flowsheet autocompletion with augmented data for both pretraining

and fine-tuning, yielding model (iii). For a fair comparison of training runs with and without data augmentation, we use the same hyperparameters.

### 3.3. Results

We use perplexity to measure the model performance. Perplexity describes the uncertainty of a model in its predictions. Therefore, a low perplexity is desirable. Here, a lower perplexity means that the model is more confident that the suggested unit operation is reasonable. Perplexity is the exponential of the negative average log-likelihood of the next token prediction. It is also equivalent to the exponential of the cross-entropy loss obtained during model training. The perplexity is computed as

$$PP(T) = \exp\left(-\frac{1}{n} \sum_i^n \log P(t_i | t_{1:i-1})\right), \quad (1)$$

where  $T = (t_1, \dots, t_n)$  is a sequence of  $n$  tokens and  $P(t_i | t_{1:i-1})$  describes the predicted probability for the next token  $t_i$  given the sequence  $t_{1:i-1}$ .

Table 1: Perplexity  $PP$  results of data augmentation. Best test result in bold font. The column "Augm." shows whether the training data were augmented or not. Pretraining and fine-tuning test perplexity are both evaluated after the model training is completed with the fine-tuning dataset.

Model	$PP$ pretraining		$PP$ fine-tuning			Training Time	
	Augm.	Test	Augm.	Train	Val		Test
(i)	No	5.38	No	3.13	4.23	5.02	51 min
(ii)	No	6.33	Yes	3.32	4.12	4.69	57 min
(iii)	Yes	<b>5.16</b>	Yes	3.07	3.80	<b>4.28</b>	1 h 31 min

The results of the different training runs are shown in Table 1. For the fine-tuning, model (ii) performs slightly better than model (i) on the test set. We explain this improvement by the fact that the fine-tuning data are limited in size (i) and with the data augmentation we can make the model more robust (ii). Even though the test perplexity is lower with data augmentation, the training perplexity is higher. This shows that the model trained with data augmentation is less prone to overfitting and generalizes better to unseen data. With data augmentation also applied to the pretraining (model (iii)), we see the best performance in all categories, improving the fine-tuning perplexity by 14.7% compared to model (i). We conclude that data augmentation can also improve the pretraining with a relatively large, generated dataset, resulting in a better fine-tuning performance.

In general, we see that the augmented SFILES are valid flowsheet representations and that they help the flowsheet autocompletion to learn the SFILES grammar and chemical process structure. For future work it would be interesting to investigate if it is more favorable for pretraining to generate more artificial data, to increase the dataset with data augmentation, or to combine both methods. It is worth noticing, that the training and test perplexity are in every case more similar for the models trained with augmented data. Due to the higher variance in the training data, these models are less prone to overfitting.

## 4. Conclusions

We propose a data augmentation of chemical flowsheet data by randomizing the branching decisions in the graph traversal for producing SFILES. We thereby demonstrate a way to increase the available flowsheet data for subsequent training of AI models. We apply augmented SFILES to the problem of flowsheet autocompletion and show that the augmentation improves the model performance in low data regimes. In

future research, we aim to apply the data augmentation methodology to further NLP models to improve their performance. The long-term goal should be to increase the flowsheet dataset size by mining additional flowsheets from literature (Schweidtmann, 2022; Schulze Balhorn et al., 2022) and companies and, if necessary, digitizing them (Theisen et al., 2023). A combination of both, data augmentation and increased dataset, size will be necessary and most beneficial.

## References

- Bjerrum, E. J. (2017). Smiles enumeration as data augmentation for neural network modeling of molecules. arXiv preprint arXiv:1703.07076.
- d'Anterrockes, L. (2005). Process flowsheet generation & design through a group contribution approach. [CAPEC], Department of Chemical Engineering, Technical University of Denmark.
- Hirtreiter, E., Schulze Balhorn, L., & Schweidtmann, A. M. (2022). Towards automatic generation of Piping and Instrumentation Diagrams (P&IDs) with Artificial Intelligence. arXiv preprint arXiv:2211.05583.
- Oeing, J., Welscher, W., Krink, N., Jansen, L., Henke, F., & Kockmann, N. (2022). Using artificial intelligence to support the drawing of piping and instrumentation diagrams using dexpi standard. *Digital Chemical Engineering*, 4, 100038.
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, 1(8), 9.
- Schulze Balhorn, L., Gao, Q., Goldstein, D., & Schweidtmann, A. M. (2022). Flowsheet Recognition using Deep Convolutional Neural Networks. In *Computer Aided Chemical Engineering* (Vol. 49, pp. 1567-1572). Elsevier.
- Schweidtmann, A. M. (2022). Flowsheet mining. Manuscript, In preparation. TU Delft.
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*, 6(1), 1-48.
- Theisen, M. F., Flores, K. N., Schulze Balhorn, L., & Schweidtmann, A. M. (2023). Digitization of chemical process flow diagrams using deep convolutional neural networks. *Digital Chemical Engineering*, 6, 100072.
- Vogel, G., Schulze Balhorn, L., Hirtreiter, E., & Schweidtmann, A. M. (2022a). Process-intelligence-research/sfiles2: V1.0.0 (Version Release). Github. <https://github.com/process-intelligence-research/SFILES2>
- Vogel, G., Schulze Balhorn, L., Hirtreiter, E., & Schweidtmann, A. M. (2022b). SFILES 2.0: An extended text-based flowsheet representation. arXiv preprint arXiv:2208.00778.
- Vogel, G., Schulze Balhorn, L., & Schweidtmann, A. M. (2023). Learning from flowsheets: A generative transformer model for auto-completion of flowsheets. *Computers & Chemical Engineering*, 171, 108162.
- Zhang, T., Sahinidis, N. V., & Sirola, J. J. (2019). Pattern recognition in chemical process flowsheets. *AIChE Journal*, 65(2), 592-603.
- Zheng, C., Chen, X., Zhang, T., Sahinidis, N. V., & Sirola, J. J. (2022). Learning process patterns via multiple sequence alignment. *Computers & Chemical Engineering*, 159, 107676.

# Systematic synthesis of pathways for hydrogenation and hydrogenolysis of acetylene over catalytic surfaces

Manjusha C. Padole, Swayam P. Misra, V. Sai Phani Kumar, Parag A.

Deshpande

*Quantum and Molecular Engineering Laboratory, Department of Chemical Engineering, Indian Institute of Technology Kharagpur, Kharagpur 721302, India*

## Abstract

Search for a good catalyst to catalyse surface reactions with high product selectivity has always been an important aspect of heterogeneous catalyst design. To this end, detailed insights into the mechanism of a reaction need to be developed, which is at times difficult for heterogeneous reactions. Efficient techniques are required which accurately provide all possible pathways over a heterogeneous catalytic surface given a set of chemically plausible species present. Linear algebra along with graph theoretical techniques can be used for identifying different possible reaction pathways from a set of elementary reactions. This study demonstrates the use of one such technique resulting in a reaction route diagram for acetylene hydrogenation-hydrogenolysis reaction system. A systematic protocol for the identification of all possible pathways is presented. Multitude of reaction pathways influencing the product selectivity of the reaction was observed with the possibility of existence of pathways with the desired selectivity.

**Keywords:** elementary reactions; surface species; reaction graph; incidence matrix; reaction route diagram.

## 1. Introduction

Rational design of a heterogeneous catalytic reactor requires a detailed understanding of the surface chemistry of the associated reaction. An important issue in the engineering of heterogeneous catalytic reactions is the identification of all possible reaction pathways leading to a specified product given an initial set of reactants. This problem can be quite challenging even for macroscopically simple reactions like acetylene hydrogenation, as has been highlighted in this study. Due to the possibility of existence of a large number of surface species on catalytic surfaces and their simultaneous interactions, on-surface synthesis has been reported by several investigators to yield a range of products based on the reaction and surface conditions. All these observations motivate the necessity of knowledge of all possible reaction pathways which can be observed over a heterogeneous surface. Recent developments in theoretical and computational frameworks for the analysis of reactive systems have proved to be helpful in gaining insights into on-surface systems. Such methods for the systematic reaction pathway search for the automated generation of reaction mechanisms work efficiently for small molecules only, and are time intensive for large molecules. Theoretical techniques reported till date require an efficient technique for the synthesis of reaction pathways. This work provides an efficient approach towards the systematic

search of reaction pathways considering as the test case acetylene hydrogenation and hydrogenolysis. Since the reactions are carried out heterogeneously, as mentioned earlier, the possibility of existence of a large number of surface intermediates exists leading to varied macroscopically observed products. Apart from the desired reaction, namely the selective hydrogenation, complete hydrogenation to  $C_2H_6$  is a possibility. Further, conversion of all three forms of  $C_2$  gases to  $CH_4$  by hydrogenolysis cannot be ruled out. Therefore, it is desirable to obtain all possible pathways during the aforementioned conversions. This remains the central issue addressed in this study.

With an aim of addressing the aforementioned issues, we propose a combination of mathematical techniques involving linear algebra and graph theory to get the correct picture. For acetylene- $H_2$  surface reactions, attempts have been made in the past to implement graph theory to complex reaction systems. Christiansen [1] and Temkin et al. [2] introduced the use of reaction graph for catalytic and noncatalytic reactions. Fishtik et al. [3] confirmed the use of this graph theory approach to be suitable for linear as well nonlinear kinetic mechanisms for single and multiple overall reactions. The success of such methods, however, depends on the identification of reaction pathways. In this study, we demonstrate a novel and generic approach for the analysis of surface catalytic systems which is capable of handling systems with a large number of surface species and macroscopic parallel reactions. Our approach considers the presence and influence of every possible surface species contributing towards the overall reaction. We demonstrate its applicability using hydrogenation and hydrogenolysis of  $C_2H_2$  as the test case.

## 2. Problem formulation and methodology

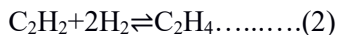
Macroscopically observable chemical species involved in the system under consideration are  $C_2H_2$ ,  $C_2H_4$ ,  $C_2H_6$ ,  $CH_4$  and  $H_2$ . These five chemical species in different adsorbed states over a catalytic surface can give rise to three categories of macroscopically observable chemical reactions given below in the scheme of Figure 1. Application of linear algebra tells us that in the above system of six series-parallel reactions with five chemical species, only three reactions are independent. This can be arrived at by finding the rank of the stoichiometric matrix with reactions as rows and chemical species as columns. However, chemical insights invoke some fundamental questions to consider all six reactions for further analysis. Analysis of the sub-set consisting of reactions (1-3) will show one of equations (2) and (3) to be redundant. This means that  $C_2H_6$  formation can take place by sequential hydrogenation of  $C_2H_2$  via  $C_2H_4$ . However, reaction (2) chemically means that  $C_2H_6$  formation takes place from  $C_2H_2$  and we ask this question: is it possible to obtain  $C_2H_6$  from  $C_2H_2$  without observing  $C_2H_4$  in any of the intermediate steps. For this, it is required to identify a pathway from  $C_2H_2$  to  $C_2H_6$  such that no intermediate species is  $C_2H_4$ . Similar analysis is required for hydrogenolysis reactions also.

The proposed algorithm for synthesis of pathways given an initial reactant and a final product is shown in Fig. 1.

Selective hydrogenation:



Complete hydrogenation:



Hydrogenolysis:

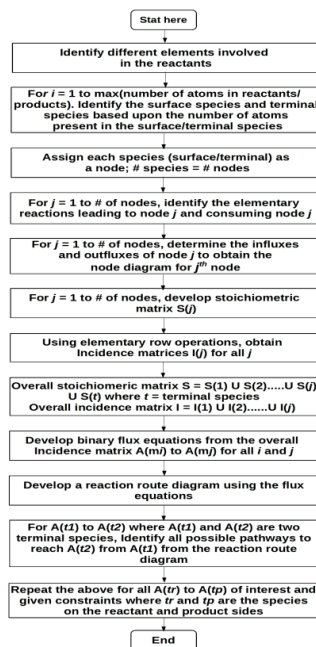
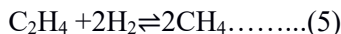
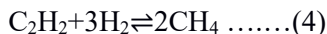


Fig. 1: Reactions involving the hydrogenation and hydrogenolysis are shown in left and the flowchart depicting the algorithm developed for systematic synthesis of pathways for a heterogeneously catalysed surface reactions is shown in right.

### 3. Results and discussion

The incidence matrix is a vital tool for the development of reaction route diagram as it gives a precise way of converting a set of reactants into a set of products. To obtain the stoichiometric sub-matrices,  $S(i)$  ( $i= 1$  to 33), the elementary reactions are written in the mathematical form as

$$\sum_{n=1}^{33} [a_n n_j] = 0 \forall r; r = 1 \text{ to } 119 \dots \dots \dots (7)$$

where  $n_j$  in Eq. 7 represents the chemical species. We identified a total of thirty-three surface species and five terminal species for the system under consideration. Correspondingly, thirty-three node diagrams were developed. We have prepared predefined pathways before analysing them and this process of synthesis of pathways can be automated. A sample node diagram for one of the surface species is shown in Figure 2a. Based on the surface species and node diagrams, a hundred and nineteen elementary reactions were identified which would explain hydrogenation and hydrogenolysis of acetylene over a catalytic surface. We have verified that no other possible surface species and elementary reaction exists.

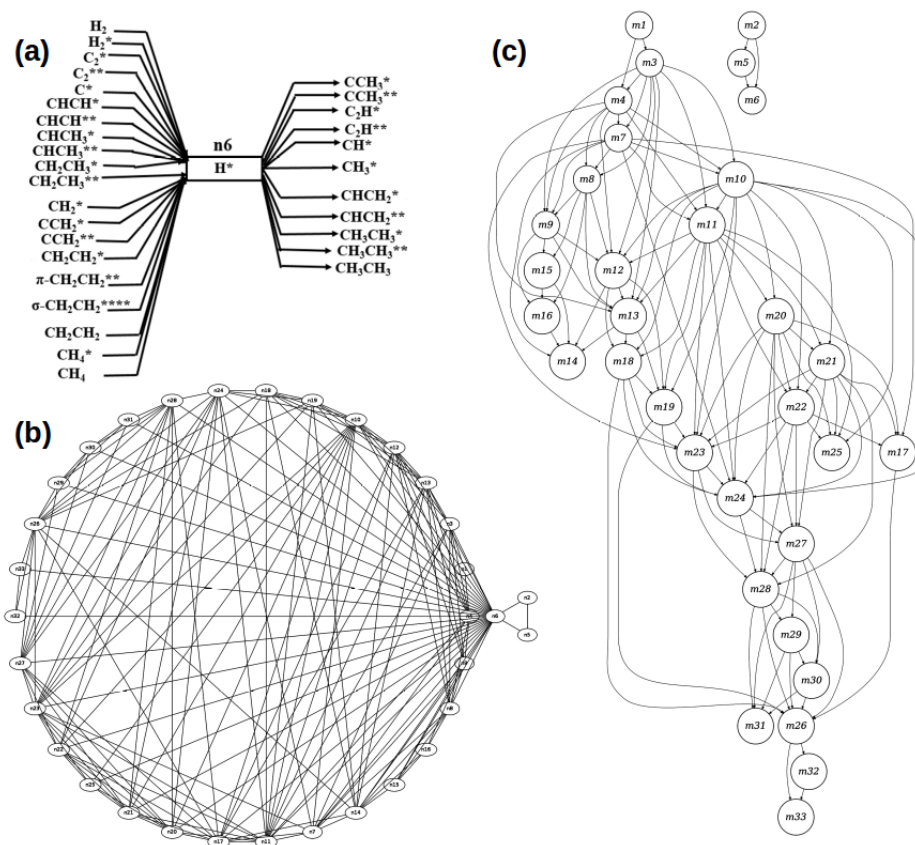


Fig. 2: Catalytic hydrogenation and hydrogenolysis of  $C_2H_2$ . Fig.2a: Node diagrams for the different chemical species identified. Fig. 2b: Reaction network diagram. Nodes  $n_i$  represent species connected by an undirected arrow, Fig. 2c: Reaction route graph.

A node appearing in the reaction scheme as a result of a surface reaction can be characterised distinctly by the appearance of both influxes and outfluxes in its node diagram. This can be seen in the node diagram of ( $H^*$ ), for example. The conversion of the given species to other species and its formation from other species can be seen from the influxes and outfluxes of this node. It is to be noted that it is possible to obtain a given species from a combination of other species resulting in treatment of such interactions as influxes while the same species can as well be treated to be resulting into a combination of other species resulting in the appearance of such species as outfluxes. This happens because of the inherent reversibility of surface reactions. This, however, does not affect our analysis in any manner as we define the interconnections in the reaction network diagram as direction-less and the appearance of a given species as influx or outflux in a given node results in the same network. The set of all node diagrams indicating individual influxes and outfluxes can be collated as a reaction network diagram which has no directionality (Figure 2b) and further as a reaction route graph which has directionality (Figure 2c). The reaction route diagram can be used to obtain the reaction pathway diagram, as shown in Fig. 3. The two pathways for  $C_2H_2$  hydrogenation to  $C_2H_4$ , for example, can be observed to involve two different modes of

adsorption of  $C_2H_2$ , one being the single site adsorption ( $C_2H_2^*$ ) while the other being the two site adsorption mode ( $C_2H_2^{**}$ ). The intermediates following the adsorption steps can also be seen to be distinct. Both of these pathways have been reported previously for  $C_2H_2$  hydrogenation to  $C_2H_4$  [4, 5]. This shows the success of our method in capturing different pathways. Pathways for further hydrogenation of  $C_2H_4$  to  $C_2H_6$  were also identified, the conversion of  $C_2H_2$  to  $C_2H_6$  via  $C_2H_4$  is shown in Fig. 3.

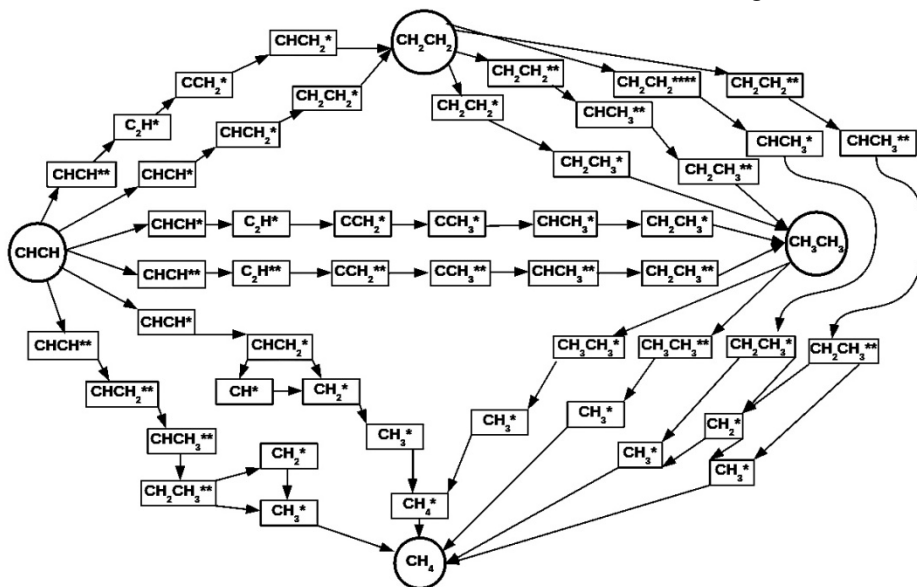


Fig. 3: Two sample pathways between any two terminal specified

Another important observation from the figure is the existence of distinct pathways for  $C_2H_2$  hydrogenation to  $C_2H_6$  without observing  $C_2H_4$ . These pathways have never been discussed in the literature and the existence of such pathways shown by our study signifies the importance of our systematic synthesis protocol in identification of pathways. Identification of hydrogenolysis of different  $C_2$  gases with the aforementioned constraints of observing or not observing a particular species can similarly be done, three instances for example are given in Fig. 3.

## Conclusions

A novel and efficient technique utilising linear algebra and graph theory was developed for the identification of all possible pathways for heterogeneously catalysed surface reactions. The applicability of the method was demonstrated successfully for  $C_2H_2$  hydrogenation-hydrogenolysis reaction. The macroscopically simple reactions involving  $C_2H_2$ ,  $C_2H_4$ ,  $C_2H_6$ ,  $CH_4$  and  $H_2$  were found to have complex surface interactions resulting in thirty-three chemical species and a hundred and nineteen elementary surface reactions. A reaction route graph, developed from elementary reactions via the stoichiometric matrix and the incidence matrix, demonstrated complex surface interactions among the surface species and a multitude of reaction pathways for partial hydrogenation, complete hydrogenation and hydrogenolysis with existence of direct reaction routes from  $C_2H_2$  to rest all terminal species. This method is capable of determining all possible pathways in short time frame compared to other expensive methods, for example techniques like quantum chemical calculations. Although the



presented model has demonstrated for hydrogenation of ethylene, the same can be explored for other surface reactions as well.

## References

- [1] J. A. Christiansen, 1953, Elucidation of reaction mechanisms by the method of intermediates in quasi-stationary concentrations, *Adv. Catal.* 5, 311-353.
- [2] O. N. Temkin, A. V. Zeigarnik, D. G. Bonchev, 1996, *Chemical reaction networks: A graph theoretical approach*, CRC Press: New York.
- [3] I. Fishtik, C. A. Callaghan, R. Datta, Reaction route graphs. II. Examples of enzyme- and surface-catalyzed single overall reactions, *J. Phys. Chem. B.* 108 (2004) 5683-5697.
- [4] D. Mei, M. Neurock, C. M. Smith, 2009, Hydrogenation of acetylene-ethylene mixtures over Pd and Pd-Ag alloys: First-principles-based kinetic monte carlo simulations, *J. Catal.* 268, 181-195.
- [5] D. Mei, P. A. Sheth, M. Neurock, C. M. Smith, 2006, First-principles-based kinetic monte carlo simulation of the selective hydrogenation of acetylene over Pd(111), *J. Catal.* 242, 1-15.

# New Methodology Integrating Tray Efficiency Predictions with Internals Design for Valve-Tray Distillation Columns

Chenguang Zhu,<sup>a</sup> Megan Jobson,<sup>a</sup> Nan Zhang<sup>a</sup>

<sup>a</sup>*Centre for Process Integration, Department of Chemical Engineering, The University of Manchester, Manchester, M13 9PL, United Kingdom*

## Abstract

Column internals design is a complex problem with many degrees of freedom and hydraulic constraints. This work proposes a new automatic design methodology for the internal design of valve-tray columns, iteratively applying results from rigorous distillation simulations, hydraulic correlations and tray efficiency predictions for valve trays. The feasibility of the proposed internals is checked against hydraulic constraints, including jet flooding, downcomer flooding and weir loading. The methodology is applied to an existing industrial valve-tray column separating C4 hydrocarbons; the results show the impact of the internals design approach on column hydraulics and tray efficiency. The proposed design can reduce the column diameter while avoiding hydraulic bottlenecks. Future work aims to implement the methodology for column design optimisation.

**Keywords:** Hydraulic analysis; Hydraulic Design; Tray efficiency;

## 1. Introduction

Distillation is a widely used, effective fluid separation process in industrial practice but requires substantial energy input and capital investment. Several published papers have addressed the optimal design of distillation columns, i.e. the column diameter, and number of trays in each section, to meet given separation specifications with minimum total annualised cost (TAC). However, these optimal column designs are still conceptual designs; details of column internals are not accounted for. The traditional approach to dimensioning column internals is to use trial and error: Kister et al., 1992; Chuang and Nandakumar, 2000; Towler and Sinnott, 2021. No definitive procedure to guide the design of column internals is available in the open literature. Typically, assigning hardware to a conceptual design depends on the designer's experience using the trial-and-error method, without a comprehensive search for internals that meets hydraulic limits and operating constraints and targets with the lowest capital cost. As a result, the design is time-consuming and a new column is excessively conservative and the effect of the column's internal design on the column efficiency and sizing has not been considered.

Valve tray is used in this study as valve tray columns are the most popular choice for trayed distillation column internals in the industry since lower pressure drop than bubble cap trays and have a wide range of applicability (Kister et al., 1992). The capital cost of a valve-tray column depends on column or section diameter, the number of trays, and the column internals – namely, type of tray; type, dimensions and layout of valves; tray spacing, pass configuration, downcomer dimensions, etc. At the same time, the column

internal dimensions affect the stage efficiency, and hence the theoretical number of stages, and thus the reflux ratio. The reflux ratio will in turn affect the required dimensions of the internals, which implies that design optimisation should consider column internal dimensions and stage efficiency iteratively (or simultaneously).

Due to the development of computer-aided process engineering, limited research in the open literature has developed optimisation approaches for internals design for sieve-tray columns (Ogboja and Kuye (1990); Lahiri, 2014; Souza et al., 2022). However, these papers optimised the sieve tray column internals dimension by giving fluid information. Interaction between tray efficiency and column internals dimensions are ignored.

Lahiri (2014) used particle swarm optimisation and the Aspen Plus simulator to optimise the design of sieve-tray distillation columns, including tray specifications. The continuous design variables considered are: downcomer dimensions, tray dimensions and tray spacing, while the discrete variables are the number of stages, feed location and pass configuration. The objective function is the total annualised cost; the constraints are hydraulic limits and pressure drop. Lahiri (2014) only considers the continuous variables during the optimisation, while the discrete variables are specified manually. Significantly, tray efficiency is not evaluated within the optimisation. In practice, neglecting the tray efficiency could lead to different internal flows, and violation of hydraulic constraints.

Souza et al. (2022) formulated a mixed-integer nonlinear problem for optimising the internals dimensions (diameter, tray spacing, hole diameter, weir length, etc.) of sieve-trayed columns using column capital cost or the mass of shell and trays as an objective function, applying hydraulic (including jet flooding, downcomer flooding) and pressure drop constraints. Souza et al. (2022) first obtain the flow rate and physical properties from the Aspen Plus simulator with operating conditions and overall efficiency specified manually without any updating.

Ogboja and Kuye (1990) developed a formulation considering tray efficiency for sieve tray optimisation. Chan and Fair's (1984) correlation was used to calculate tray efficiency; the objective is then to maximise the tray efficiency with rigorous consideration of hydraulic limits and pressure drop. However, Ogboja and Kuye (1990) used a set of flow rates and fluid properties presented by literature to conduct the design, which means flow rate, fluid properties and the total number of stages are fixed in the optimisation. Although this work uses an exhaustive method to analyse all sets of variables (internal dimensions) to determine tray efficiencies, the effect of tray efficiency on the flow rate and fluid properties was not considered during the iteration of the tray efficiency evaluation.

Meanwhile, published research has not explored details of the hydraulic evaluation of valve-tray columns with cost as a design criterion, and the impact of the tray efficiency of the key tray in each section on the flow rate, and fluid properties should be assessed.

This work develops a systematic and automatic column design methodology for valve-tray columns considering column internals dimensions design, hydraulic performance of the trays, pressure drop and tray efficiency, with a focus on minimising column internals dimensions with the consideration of tray efficiency, where the operating limits (jet flooding, downcomer flooding and downcomer backup) in the column should be avoided.

## 2. Valve-trayed column design methodology

Figure 1 represents the proposed methodology which employed converged simulation of a simple distillation column (using Aspen Plus V11) and Python code which can exchange inputs and outputs with the process simulator automatically.

This work proposes a design methodology including an equilibrium-based 'rigorous' simulation of a column considering section efficiency (Murphree efficiency) to achieve a

specified separation of a given feed; calculation of column internal dimensions (diameter, deck thickness, downcomer sizing, tray spacing, valve dimensions, tray spacing) under specified flooding limits and fractional hole area; and prediction of section efficiency of mean stage efficiencies within the section. The column is simulated using an established process simulation software assuming the stage efficiencies (as input variables), tray spacing, and the number of passes. The simulation results (stage-flow rates and fluid properties) are used as inputs for column hydraulic analysis and tray and downcomer calculations for every stage. The feasibility of the design is checked in terms of hydraulic limits and pressure drop, and the internal dimensions are then modified to avoid hydraulic limits. Next, the corresponding internal dimensions (active area, deck thickness, fractional hole area) and fluid properties are input to the stage efficiency prediction model of Chen and Chuang (1993). The calculated section efficiencies are re-input to the process simulator to recalculate the material and the energy flows and thus update the internal dimensions and section efficiencies, which means the column simulation and internal dimensions calculation are repeated until converged. The three steps are repeated until the calculated and assumed efficiencies are equal.

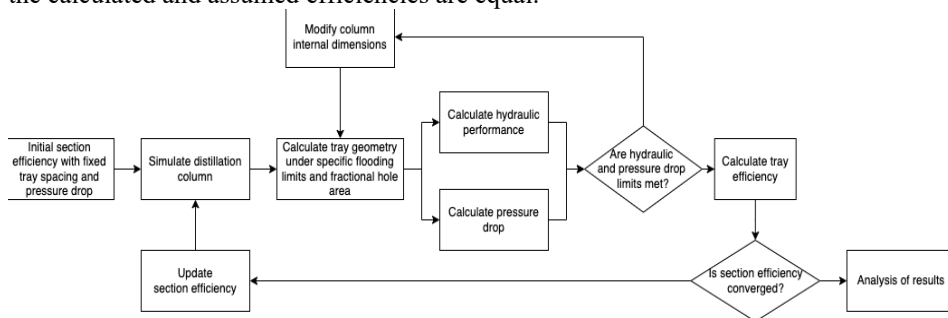


Figure 1. Proposed design methodology for valve-tray distillation column

### 2.1. Hydraulic analysis: methodology

Distillation columns can only operate efficiently with containing trays within certain hydraulic limits. A smaller diameter means a smaller active area or downcomer area for vapour-liquid pass-through, which will cause jet flooding or downcomer flooding. Thus, the selection of jet flooding and downcomer flooding upper limits have a significant effect on the diameter.

A maximum value of 80% is assumed since this figure is widely used in design textbooks. Jet flooding is predicted using the Glitsch correlation, also known as the ‘Equation 13’ (Branan, 2005) for conventional floating valve trays. Downcomer choke flooding is evaluated in terms of the volumetric flow rate at the downcomer entrance following the Glitsch correlation (Kister et al., 1992). Another typical constraint weir loading, the ratio of liquid volume flow rate to weir length, is assumed to limit 107 m<sup>3</sup>m<sup>-1</sup>h<sup>-1</sup>, as the value is typically applied in the practice (Kister et al., 1992). The clear liquid height, downcomer backup and pressure drop analysis are evaluated following correlations of Koch-Glitsch (2013); the vapour capacity factor and approximate tower diameter are calculated using the correlations of Branan (2005).

Downcomer backup is evaluated knowing the aerated liquid height in the downcomer, which is calculated by an aerated factor, the ratio of the clear liquid height to the aerated liquid height. A correlation is proposed, regressed from the aerated factor chart for valve trays of Klein (cited by Kister et al., 1992). The regressed model is shown in Eq. (1), where  $F_{va}$  is the superficial F-factor based on the active area.

$$\beta = -1.31F_{va} + 2.07F_{va}^2 - 1.71F_{va}^3 + 0.72F_{va}^4 - 0.15F_{va}^5 + 0.01F_{va}^6 + 0.97 \quad (1)$$

Eq. (1) is only valid between the  $F_{va}$  from 0 to  $3.05 \text{ kg}^{0.5}\text{s}^{-1}\text{m}^{-0.5}$ .

### 2.2. Tray efficiency evaluation: methodology

The correlation of Chen and Chuang (1993) was developed based on a sieve tray in the hydrocarbon systems. However, Vennavelli et al. (2014) validated this correlation using FRI valve tray efficiency data (hydrocarbon test data) and showed that Chen and Chuang's (1993) model predicted the FRI valve tray efficiency better than the four existing valve tray efficiency models presented in published literature. Vennavelli et al. (2014) suggested that the sieve tray models present reasonable alternatives if they are developed from a wide range of data and operating conditions. Thus, there is good evidence that the Chen and Chuang (1993) model is valid for valve-tray columns in hydrocarbon systems.

## 3. Case study: separation of i-butane/n-butane mixture

The distillation system to demonstrate the proposed methodology is an industrial-scale valve-tray column separating the i-butane and n-butane mixture. However, the composition of distillate and bottom stream, and the section efficiency presented by Ilme et al. (2001) are not self-consistent. Thus, a self-consistent distillate and bottom stream composition, and section efficiency simulated and calculated following the methodology above are used instead of the results presented by Ilme et al. (2001), while the rest of the operating conditions and column internals dimensions are the same as that presented by Ilme et al. (2001). Thus, the i-butane/n-butane column based on the modified distillate and bottom stream composition, and section efficiency is used as a base case. The detailed operating conditions, stage distribution and product specifications are shown in Figure 2. Standard valve trays (Koch-Glitsch Ballast V-1 trays) are assumed for both sections. Section efficiency is applied instead of the tray efficiency of each tray, which is the average value of the tray efficiency of each tray in rectifying or stripping section predicted using Chen and Chuang (1993) model. The design methodology was applied automatically via the Python code to meet different flooding limits (jet flooding and downcomer flooding). The fractional hole area, which is the ratio between the hole area and the active area, was also varied. The impact of different flooding conditions and fractional hole area on the column diameter, section efficiency, pressure drop, and heat duty of the condenser and reboiler were analysed. The feed conditions, light key component purity in the distillate stream and distillate rate were maintained throughout the study. Table 1 presents the details of column internal dimensions and operating conditions under different flooding conditions (from 70% to 80%) and fractional hole area (from 16.6% to the lower limit).

The percentage of jet and downcomer flooding calculated based on the base case internal dimensions given by Ilme et al. (2001) are 69% and 52%. It is found that the diameter can reduce by 10% when the column operates under 80% of both jet and downcomer flooding and the fractional hole area is 16.6%. However, higher flooding condition means higher froth liquid height on the tray, leading to a slightly higher pressure drop than the base case, while section efficiencies are almost the same as the base case.

It is also found that the section efficiency is mainly dominated by the fractional hole area. The section efficiency can increase by around 7 to 10% in both rectifying and stripping sections under the same flooding condition. However, it shows that the lower limits of fractional hole areas are different under different flooding conditions due to downcomer backup constraints. Thanks to the increasing of section efficiency, the reflux ratio is smaller than the base case, leading to a low heat duty demand in both the condenser and reboiler.

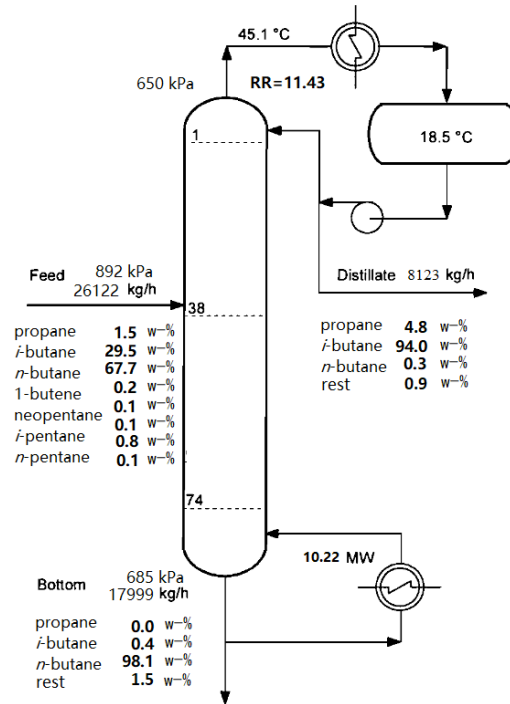


Figure 2. Flowsheet and operating data for an i-butane/n-butane column

Table 1. Column internal dimensions, hydraulic performance and pressure drop performance under different flooding limits and fractional hole area.

Flooding condition	base case	70%			80%		
		16.6%	13.0%	8.7%	16.6%	13.0%	10.1%
% hole area	16.6%	16.6%	13.0%	8.7%	16.6%	13.0%	10.1%
Diameter	2,900 mm	2,796 mm	2,765 mm	2,721 mm	2,593 mm	2,564 mm	2,539 mm
Pressure drop	0.48 kPa	0.51 kPa	0.55 kPa	0.73 kPa	0.54 kPa	0.64 kPa	0.89 kPa
Heat duty	10.60MW; 10.22 MW	10.59MW; 10.21 MW	10.38MW; 10.00 MW	10.09MW; 9.71 MW	10.52MW; 10.15 MW	10.31MW; 9.94 MW	10.13MW; 9.76 MW
Section efficiency	110%, 115%	110%, 115%	114%, 119%	121%, 124%	112%, 116%	115%, 120%	119%, 123%

Since the section efficiency is integrated into the column internals design, the column diameter can reduce an extra 2% than only changing the flooding condition. However, the reduction of fractional hole area allows the vapour velocity through the valve tray to be increased, which results in a significant increase in pressure drop. As shown in the table, the pressure drop can be increased by up to 60%, from 0.54 kPa to 0.89 kPa, under the same flooding conditions.

#### 4. Discussion

In this work, the complex tray internals design and tray efficiency prediction can be integrated automatically by computer-aided design. Thus, the size of the distillation column no longer needs to be too conservative based on the designers' experience, and the separation efficiency is also improved, resulting in substantial savings in capital investment. At the same time, it helps engineers clearly understand the current production capacity and the potential to increase production, which is also important for safe production.

#### 5. Conclusion

The case study illustrates the benefit of the proposed automatic design methodology to design and evaluate design results to reduce the column diameter while satisfying the hydraulic limits (jet flooding, downcomer flooding, downcomer backup, and weir loading).

The methodology can be applied to trayed distillation columns and accounts for the impact of the change of hydraulic limits and fractional hole area on the pressure drop, downcomer backup and section efficiencies. In this case study, the distillation column diameter could be reduced by up to 12%, compared to the base case, by increasing the flooding conditions, without violating hydraulic constraints. Since the tray efficiency is integrated into the internal design, additional benefits can be obtained while the fractional hole area is changed. For the tray column, the fractional hole area is contributed by many variables, including valve leg length or hole diameter, deck thickness, and pitch of the hole, which will also affect the tray efficiency according to Chen and Chuang (1993) correlation. Thus, there is still a wide range to explore an optimal design.

However, the feed location and the number of actual stages are fixed in this work; these may not be optimal for the total annual cost. Cost-benefit analysis would be needed to further evaluate the design solutions.

#### References

- Branan, C. R. (2005). *Rules of Thumb for Chemical Engineers: A manual of Quick, Accurate Solutions to Everyday Process Engineering Problems*. Elsevier
- Chen, G.X. & Chuang, K.T. (1993), Prediction of point efficiency for sieve trays in distillation. *Industrial & Engineering Chemistry Research*, 32(4), 701–708.
- Chuang, K. T., & Nandakumar, K. (2000). Tray Columns: Design. In I. D. Wilson (Ed.), *Encyclopedia of Separation Science*, pp. 1135–1140. Academic Press.
- Ilme, J., Klemola, K., Aittamaa, J., & Nystrom, L. (2001). Calculating distillation efficiencies of multicomponent i-butane/n-butane column. *Chemical Engineering Communications*, 184(1), 1–21.
- Kister, H. Z., Haas, J. R. ., Hart, D. R. ., & Gill, D. R. . (1992). *Distillation Design (Vol. 1)*. McGraw-Hill New York.
- Koch-Glitsch. (2013). Ballast tray design manual bulletin 4900. [Online] <https://www.koch-glitsch.com/getattachment/ebcafcad-ed20-419a-994b-b05ef37cf682/attachment.aspx> [Accessed October 2022]
- Lahiri, S. K. (2014). Particle swarm optimization technique for the optimal design of plate-type distillation column. In J. Valadi & P. Siarry (Eds.), *Applications of Metaheuristics in Process Engineering* (pp. 153–182). Springer International Publishing.
- Ogboja, O., & Kuye, A. (1990). A procedure for the design and optimisation of sieve trays., *Trans. I. Chem. E.*, 68(5), 445–452.

- Souza, A. R. C., Bagajewicz, M. J., & Costa, A. L. H. (2022). Globally optimal distillation tray design using a mathematical programming approach. *Chemical Engineering Research and Design*, 180, 1–12.
- Stichlmair, J. G., Klein, H., & Rehfeldt, S. (2021). *Distillation: Principles and Practice*. John Wiley & Sons. (pp. 487-533)
- Towler, G., & Sinnott, R. (2021). *Chemical Engineering Design: Principles, Practice and Economics of Plant and Process Design*, 3<sup>rd</sup> ed. Butterworth-Heinemann.
- Vennavelli, A. N., Whiteley, J. R., & Resetarits, M. R. (2014). Predicting valve tray efficiency. *Chemical Engineering Research and Design*, 92(11), 2148–2152.





# Process Design and Intensification of Circulating Catalytic Fluidized Bed Membrane Reactor for Oxidative Coupling of Methane

Moustafa Ali<sup>a</sup>, Yuhe Tian<sup>b</sup>, Dustin Kenefake<sup>a,c</sup>, Efstratios N. Pistikopoulos<sup>a,c\*</sup>

<sup>a</sup>Texas A&M Energy Institute, Texas A&M University, College Station, Texas, USA, 77845.

<sup>b</sup>Department of Chemical and Biomedical Engineering, West Virginia University, Morgantown, West Virginia, United States, 26505.

<sup>c</sup>Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX, USA, 77840

\*Corresponding Author: Prof. Efstratios N. Pistikopoulos - [stratos@tamu.edu](mailto:stratos@tamu.edu)

## Abstract

The catalytic oxidative coupling of methane (OCM) process has received intense interest in the literature due to the potential to directly convert natural gas or methane to value-added chemicals at a reduced cost, energy consumption and carbon emissions when compared to conventional processes. However, major challenges such as low yield, catalyst deactivation, and reactor scale-up still challenge the commercialization of this process. A promising solution to address these challenges is to develop innovative OCM reactor designs leveraging the recent advances in modular process intensification, e.g. membrane reactors. The aim of this work is to design an optimal OCM process at commercial scale leveraging the concept of modular process intensification. We investigate an intensified fluidized bed membrane reactor (FBMR) catalyzed by  $\text{La}_2\text{O}_3/\text{CaO}$  and compare its performance to a conventional fluidized bed reactor (FBR). The use of membrane for oxygen feed distribution has been reported to result in better  $\text{C}_2+$  yield and selectivity by selectively enhancing the desired reactions.

**Keywords:** OCM, Process Intensification, Fluidized Bed Membrane Reactors

## 1. Introduction

The oxidative coupling of methane has been the target of intense scientific and commercial interest for more than forty years due to the tremendous potential of the technology to reduce costs, energy, and environmental emissions in the production of ethylene. In OCM, methane ( $\text{CH}_4$ ) and oxygen ( $\text{O}_2$ ) react over a catalyst exothermically to form ethylene ( $\text{C}_2\text{H}_4$ ), ethane ( $\text{C}_2\text{H}_6$ ), water ( $\text{H}_2\text{O}$ ) and heat, according to the following reactions (1 to 4) (Penteado et al., 2016). While the benefits of OCM have been known since the early 1980s, past efforts did not result in a viable process design with acceptable performance needed for commercialization. Despite the pioneering work in the catalysis of OCM that has managed to bring the commercialization realization of OCM (Galadima & Muraza, 2016; Onoja et al., 2019), there is need for a novel reactor design to achieve highly economically and environmentally relevant optimal OCM process design with optimal  $\text{C}_2+$  yield, selectivity and  $\text{CH}_4$  conversion (Bartreau, 2022). These catalysts, while at times achieving promising yield and selectivity, were hampered by very high operating temperatures, low activities, and short lifetimes on the order of hours to days.



The paper by (Cruellas et al., 2017) outlined that there are more than 20 types of reactors that were designed for experiments or only modeled for the optimized OCM production. In the aforementioned paper, it outlined that the most popular and widely applied reactors that, produce the highest  $C_2+$  selectivity%,  $CH_4$  conversion%, and the highest  $C_2+$  yield%. The most common types of conventional reactors used were the packed bed reactors and fluidized bed reactors. The advantages of fluidized bed reactors (FBR) are: (i) Perform almost isothermally and (ii) Stable and efficient performance especially concerning the reactor operability and the safety aspects (Chen et al., 2020; Schwarz et al., 2014; Yaghobi & Ghoreishy, 2008). Concerning the packed bed reactors (PBR), issues associated with the performance of the reactor were identified. In the application of OCM, the reactor suffers from the following: (i) Hot spot formations on the catalyst causing the reduction of  $C_2+$  formation drastically due to high operation temperatures. (ii) Fast degradation of catalyst. This in fact leads to low yield of  $C_2+$ . While fluidized bed reactors overcame the catalyst deactivation and high heat management issues, still FBR showed that only low yield of 19% could be achieved (Jašo et al., 2011). Several works have tried to overcome some of the bottlenecks in both conventional reactors such as reducing the operating temperatures while improving the yield and for better conversion, however, the highest yield% reached was approximately 20% (Chen et al., 2020; Yaghobi & Ghoreishy, 2008). Membrane reactors offer a great advantage to overcome those issues, however, the overall conversion, yield and selectivity were relatively low. The maximum methane conversion  $C_2+$  yield and selectivity achieved from membrane reactors were 24%, 20% and 47% respectively (Onoja et al., 2019). Although membrane reactors demonstrated an improved selectivity of  $C_2+$ , those systems suffered from low methane conversion due to the low flux of oxygen throughout the membrane which hindered the overall performance. While this issue could be avoided by raising the operating temperature, it negatively affected the selectivity of  $C_2+$  (Farrell & Linic, 2016). Cruellas et al., 2020 introduced the concept of bubbling fluidized bed reactor with a membrane installed and loaded with catalyst of interest. The advantages behind this system are the following (Cruellas et al., 2017): (i) Distributed oxygen feeding results in a local lower  $O_2$  concentration, which allows to obtain higher yields, (ii) Stability of the catalyst particles. Typically, the advantages of traditional fluidized beds and use of membranes are combined. Thus, the objective of this paper is to design a novel intensified circulating Fluidized Bed Membrane Reactor (FMBR) for the application of OCM to overcome all the above-mentioned issues leveraging the use of membrane for better oxygen distribution and the internal catalyst circulation to achieve higher  $C_2+$  yield. This paper summarizes the main findings behind the oxidative coupling of methane (OCM) system, main key elements to model an intensified OCM Circulating Fluidized Bed Membrane Reactor system. This rest of the paper is organized as follows: section 2 covers the kinetics of oxidative coupling of methane. Section 3 presents the design and validation of the conventional Fluidized Bed Reactor (FBR). Section 4 showcases the design and intensification of circulating Fluidized Bed Membrane Reactor followed by results and technical insights.

## 2. Kinetics of Oxidative Coupling of Methane

The reported kinetic model is considered the most used kinetic OCM model for the development of an OCM reactor (Ching et al., 2002; Cruellas et al., 2020). A comprehensive 10-step kinetic model of the oxidative coupling of methane to  $C_2+$  hydrocarbons over a 10 %  $La_2O_3/CaO$  catalyst was developed on the basis of kinetic measurements in a micro catalytic fixed bed reactor covering a wide range of reaction conditions by Stansch et al., 1997. The reaction scheme shown in Figure 1 contains three primary and seven consecutive steps that involves many species. The conversion of hydrocarbons and of carbon monoxide with oxygen were described by applying Hougen-Watson type rate equations. For the remaining reaction, power law equations were applied. The kinetics model's parameters were reported by (Ching et al., 2002; Stansch et al., 1997)

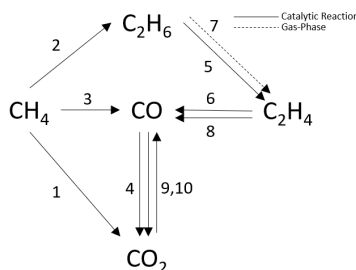


Figure 1: Reaction scheme of the kinetic model proposed by (Stansch et al., 1997)

## 3. Modeling of Conventional Fluidized Bed Reactor (FBR)

To systematically identify the optimal design solution, 2D high fidelity conventional Fluidized Bed Reactor (FBR) was adapted from (Cruellas et al., 2020a; Cruellas et al., 2020b) and modeled utilizing gPROMS ModelBuilder (PSE, 2020), which comprise partial differential algebraic equations accounting for mass balances, hydrodynamics, catalyst solid distribution and radial velocities across the reactor. The model is a core-annulus system which describes the two reactor regions that are characterized by different fluid dynamic conditions and catalyst concentrations. Figure 2 illustrates the hydrodynamic mechanistic of the 7 meters length and 0.6 meters diameter adapted reactor design. The model assumes that gas is transported upward in the core section, and it is stagnant in the outer region and accounts for radial gas mass transfer between the two sections. It was assumed that undiluted catalyst particles are fed to the reactor. The reactor is operated at 800 °C and 2 bar with methane to oxygen feed of 4:1. The reactor was simulated with a terminal velocity of 2 m/s and the value of mass transfer coefficient was fixed at 0.05 m/s. Results shows that  $C_2+$  yield was found to be 15% with a selectivity of 50% and a methane conversion of 26% which matches the adapted model from (Cruellas et al., 2020a). The validation of FBR hydrodynamics is shown in Figure 3.

## 4. Design and Intensification of Circulating Fluidized Bed Membrane Reactor (FBMR)

Oxygen permeation only membranes are integrated in the circulating FBR design to investigate the performance the proposed intensified FBMR design. A configuration had been selected where the membrane tube is placed in middle of the reactor along the axial direction and the catalyst bed is positioned around the membrane tubes. Due to the high

operating temperature of the process,  $\text{Ba}_{1-x}\text{Sr}_x\text{Co}_{1-y}\text{Fe}_y\text{O}_{3-\delta}$  (BSCF) perovskite membrane is selected to be integrated with the conventional fluidized bed reactor with a diameter of 0.01 meters, whereas the membrane length is equal to the reactor length to maximize the oxygen distribution (Spallina et al., 2015).

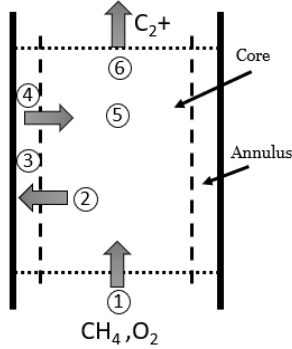


Figure 2: FBR Hydrodynamics Mechanism

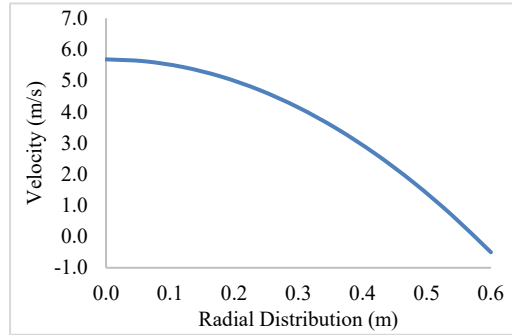


Figure 3: Hydrodynamics Validation - FBR Design

This membrane can withstand high temperature ranges up to 1000 °C. In addition, the selected membrane is an  $\text{O}_2$  permeation only meaning that air could be used as inlet along with methane for better economics (CAPEX and OPEX). The utilization of membrane along the axial distribution in the reaction offers a uniform distribution of  $\text{O}_2$  thus, it helps to better control  $\text{O}_2$  concentration along the reactor, and eventually improves  $\text{C}_2+$  selectivity%, yield% and an increase in the  $\text{CH}_4$  conversion%. The equations were modified to accommodate the placement of the membranes in the FBR design. The model includes the membrane characteristics, governing equations of both fuel and air sides, hydrodynamics, catalyst solid distribution and radial velocities across the reactor. The membrane characteristics including permeabilities, mass and heat coefficients were adapted from (Spallina et al., 2015; Xu & Thomson, 1999). The Fluidized Bed Membrane Reactor (FBMR) governing equations are presented in table 1.

Table 1: Fluidized Bed Membrane Reactor (FBMR) Governing Equations	
Mass Balance in the core	$\frac{dc_i^c v^c}{dx} = r_i(c_j^c, \gamma^c) - K_i^{ca}(c_i^c - c_i^a) + h_i^f a^f (c_i^c - c_i^f)$
Mass Balance in the annulus	$r_i(c_j^a, \gamma^a) + \frac{V^c}{V^a} K_i^{ca}(c_i^c - c_i^a) = 0$
Air Side Equations	$\frac{dc_i^{air} v^{air}}{dx} = -h_i^{air} a^{air} (c_i^{air} - c_i^{m,air})$
Total Mass Balance	$c^c = \sum_{i=1}^N c_i^c$
Average Solid Fraction	$\gamma^c = 1 - \varepsilon^c = \frac{1}{r_c^2} \int_0^{r_c} 2r\varepsilon(r) dr$
	$\gamma^a = 1 - \varepsilon^a = \frac{1}{r_a^2 - r_c^2} \int_{r_c}^{r_a} 2r\varepsilon(r) dr$
	$\varepsilon(r) = 1 - \frac{G(r)}{\rho_s V_p(r)}$
Radial Velocity Profile – Solid reflux	$\frac{G(r)}{G_s} = a \left( 1 - \left( \frac{r}{r_a} \right)^5 \right) + 1 - \frac{5a}{7}$
Radial Velocity Profile – Particle velocity	$V_p(r) = \left( 1.5 \frac{U_{g,c}}{\varepsilon^c} - V_t \right) \left( 1 - \left( \frac{r}{r_c} \right)^2 \right)$

### 5. FBMR Results and Technical Insights

The fluidized bed membrane reactor was simulated with the same conditions selected for the conventional fluidized bed reactor with catalyst particle size of  $70\ \mu\text{m}$  and net solids flux of  $100\ \text{kg m}^{-2}\text{s}^{-1}$ . The amount of oxygen has been varied from 10% to 30% whereas the catalyst concentration was fixed at 10%. The results showed that the optimal feed ratio  $\text{CH}_4/\text{O}_2$  of 4 with a bed temperature of 800 and pressure of 2 bar. The computed axial mole fraction profiles of methane and  $\text{C}_2^+$  are given in Figure 4. The overall process performance resulted in a  $\text{C}_2^+$  yield of 35%,  $\text{C}_2^+$  selectivity of 67% and methane conversion of 67%. The axial  $\text{C}_2^+$  selectivity,  $\text{C}_2^+$  yield and conversion profiles are presented in Figure 5. The integration of membrane in the FBR design along with the catalyst distribution offered feeding the oxygen along the axial direction. While this integration helps in keeping the oxygen partial pressure low, it does consequently help in enhancing the reactor's performance towards the yield of more  $\text{C}_2^+$ . At the entrance of the reactor, the kinetics dominates over mass transfer allowing all the reaction to take place in the core only. It is important to note that both the kinetics resistance and mass transfer resistance are comparable as the reactions progress along the axial direction in both the core and annulus. This in fact could help protecting the  $\text{C}_2^+$  formed inside the intensified FBMR. Finally, one of the drawbacks of the OCM kinetics is the enormous production of carbon dioxide. It is one of most critical issues faced as it has a significant impact on the energy consumption, the environment and the product's purity (Galadima & Muraza, 2016). As the implementation of the membrane in the intensified FBMR design shifted the kinetics towards the production of more  $\text{C}_2^+$ , a  $\text{CO}_2$  reduction was achieved from 6.8% in the FBR configuration to 1.3% in the intensified FBMR design.

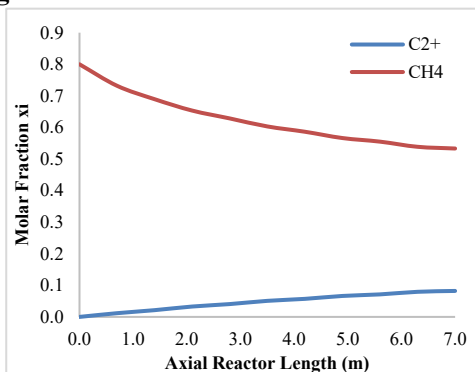


Figure 4: FBMR molar fraction profile of  $\text{CH}_4$  and  $\text{C}_2^+$

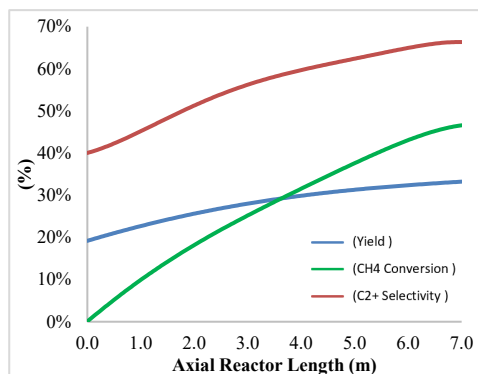


Figure 5:  $\text{C}_2^+$  Selectivity, Yield and  $\text{CH}_4$  Conversion of FBMR

### Conclusion

In this paper, we present the new intensified reactor design system for the simulation of oxidative coupling of methane. The conventional fluidized bed reactor was simulated and validated with a  $\text{C}_2^+$  yield, selectivity, and methane conversion of 15%, 54% and 26% respectively. A membrane was integrated with the conventional reactor coupling a perovskite oxygen selective membrane along the axial direction to improve on the three parameters mentioned above. The reactor's design shows a noticeable improvement with a  $\text{C}_2^+$  yield of 35%,  $\text{C}_2^+$  selectivity of 67% and improved conversion of methane up to 50% with a reduced  $\text{CO}_2$  production by almost 81%. Finally, the intensified fluidized bed membrane reactors could be a solution to improve the  $\text{C}_2^+$  yield%,  $\text{C}_2^+$  selectivity% and

CH<sub>4</sub> conversion%. Through this paper, it has been demonstrated that the concept of the intensified FMBR design can outperform the conventional FBR design.

## 6. Future Work

Although the new proposed intensified FMBR design solves some of the limitation for the OCM reaction to take place, it opened few research questions to uncover the role of membrane in achieving high C<sub>2</sub><sup>+</sup> selectivity%, C<sub>2</sub><sup>+</sup> yield% and CH<sub>4</sub> conversion%. The effect of these parameters (e.g. oxygen vacancy dilution coefficient (D<sub>v</sub>), species diffusion coefficient (D<sub>i</sub>) and perovskite membrane mass transfer coefficients (k<sub>f</sub> & k<sub>a</sub>)) will be investigated to enhance the economics of the OCM reactions and performance of FMBR. While the investigation of OCM reaction systems has mostly been focused on steady-state conceptual designs, ignoring the process dynamics limits the applicability of evaluating the implementation for commercial use.

## 7. Acknowledgement

We acknowledge the financial support from the Texas A&M Energy Institute, Artie McFerrin Department of Chemical Engineering at Texas A&M University, Shell USA and RAPID SYNOPSIS Project.

## References

- Barteau, M. A. (2022). Is it time to stop searching for better catalysts for oxidative coupling of methane? *Journal of Catalysis*, 408, 173–178. <https://doi.org/10.1016/j.jcat.2022.03.006>
- Chen, L., Pannala, S., Broekhuis, R., Gautam, P., Gu, T., West, D., & Balakotaiah, V. (2020). Three-dimensional CFD simulation of pattern formation in a shallow packed-bed reactor for oxidative coupling of methane. *Chemical Engineering Journal*, 400(March), 125979. <https://doi.org/10.1016/j.cej.2020.125979>
- Ching, T. T., Mohamed, A. R., & Bhatia, S. (2002). Modeling of catalytic reactor for oxidative coupling of methane using La<sub>2</sub>O<sub>3</sub>/CaO catalyst. *Chemical Engineering Journal*, 87(1), 49–59. [https://doi.org/10.1016/S1385-8947\(01\)00191-7](https://doi.org/10.1016/S1385-8947(01)00191-7)
- Cruellas, A., Melchiori, T., Gallucci, F., & van Sint Annaland, M. (2017). Advanced reactor concepts for oxidative coupling of methane. *Catalysis Reviews - Science and Engineering*, 59(3), 234–294. <https://doi.org/10.1080/01614940.2017.1348085>
- Cruellas, A., Melchiori, T., Gallucci, F., & van Sint Annaland, M. (2020). Oxidative Coupling of Methane: A Comparison of Different Reactor Configurations. *Energy Technology*, 8(8), 1–15. <https://doi.org/10.1002/ente.201900148>
- Cruellas, Aitor, Heezius, J., Spallina, V., van Sint Annaland, M., Medrano, J. A., & Gallucci, F. (2020). Oxidative coupling of methane in membrane reactors: A techno-economic assessment. *Processes*, 8(3), 1–25. <https://doi.org/10.3390/pr8030274>
- Farrell, B. L., & Linic, S. (2016). Oxidative coupling of methane over mixed oxide catalysts designed for solid oxide membrane reactors. *Catalysis Science and Technology*, 6(12), 4370–4376. <https://doi.org/10.1039/c5cy01622c>
- Galadima, A., & Muraza, O. (2016). Revisiting the oxidative coupling of methane to ethylene in the golden period of shale gas: A review. *Journal of Industrial and Engineering Chemistry*, 37, 1–13. <https://doi.org/10.1016/j.jiec.2016.03.027>
- Jašo, S., Arellano-Garcia, H., & Wozny, G. (2011). Oxidative coupling of methane in a fluidized bed reactor: Influence of feeding policy, hydrodynamics, and reactor geometry. *Chemical Engineering Journal*, 171(1), 255–271. <https://doi.org/10.1016/j.cej.2011.03.077>
- Oono, O. P., Wang, X., & Keehagiopoulos, P. N. (2019). Influencing selectivity in the oxidative coupling of methane by modulating oxygen permeation in a variable thickness membrane reactor. *Chemical Engineering and Processing - Process Intensification*, 135(August 2018), 156–167. <https://doi.org/10.1016/j.ccep.2018.11.016>
- PSE. (2020). *Process System Enterprise.gPROMS*. [www.pcenterprise.com/products/gproms](http://www.pcenterprise.com/products/gproms)
- Schwarz, H., Geske, M., Franklin Goldsmith, C., Schlögl, R., & Horn, R. (2014). Fuel-rich methane oxidation in a high-pressure flow reactor studied by optical-fiber laser-induced fluorescence, multi-species sampling profile measurements and detailed kinetic simulations. *Combustion and Flame*, 161(7), 1688–1700. <https://doi.org/10.1016/j.combustflame.2014.01.007>
- Spallina, V., Melchiori, T., Gallucci, F., & Van Sint Annaland, M. (2015). Auto-thermal reforming using mixed ion-electronic conducting ceramic membranes for a small-scale H<sub>2</sub> production plant. *Molecules*, 20(3), 4998–5023. <https://doi.org/10.3390/molecules20034998>
- Stansch, Z., Mleczko M. L., & Baerns. (1997). Comprehensive Kinetics of Oxidative Coupling of Methane over the La<sub>2</sub>O<sub>3</sub>/CaO Catalyst. *Industrial and Engineering Chemistry Research*, 36(7), 2568–2579. <https://doi.org/10.1021/ie960562k>
- Xu, S. J., & Thomson, W. J. (1999). Oxygen permeation rates through ion-conducting perovskite membranes. *Chemical Engineering Science*, 54(17), 3839–3850. [https://doi.org/10.1016/S0009-2509\(99\)00015-9](https://doi.org/10.1016/S0009-2509(99)00015-9)
- Yaghobi, N., & Ghoreishy, M. H. R. (2008). Oxidative coupling of methane in a fixed bed reactor over perovskite catalyst: A simulation study using experimental kinetic model. *Journal of Natural Gas Chemistry*, 17(1), 8–16. [https://doi.org/10.1016/S1003-9953\(08\)60019-5](https://doi.org/10.1016/S1003-9953(08)60019-5)

# Solvent pre-selection for extractive distillation using Gibbs-Helmholtz Graph Neural Networks

Edgar Ivan Sanchez Medina,<sup>a\*</sup> Kai Sundmacher,<sup>a,b</sup>

<sup>a</sup> *Chair for Process Systems Engineering, Otto-von-Guericke University, Universitätsplatz. 2, Magdeburg, 39106, Germany*

<sup>b</sup> *Process Systems Engineering, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, Magdeburg, 39106, Germany*

\* *sanchez@mpi-magdeburg.mpg.de*

## Abstract

When selecting candidate solvents for extractive distillation, activity coefficients at infinite dilution are commonly employed. With them, the selectivity of different solvents for a given mixture can be estimated, and the solvents can be ranked accordingly. However, given the large chemical space of potential solvents and the limited experimental data for activity coefficients at infinite dilution, predictive methods become necessary to perform a solvent pre-selection process across a broad solvent space. In this work, a method for selecting solvents using the Gibbs-Helmholtz Graph Neural Network as predictive model for infinite dilution activity coefficients is presented. Different case-studies are given to illustrate the efficacy of this methodology ranging from aliphatic-aromatic separations to the separation of olefin-paraffin mixtures. The extended Margules equation is employed to estimate the vapor-liquid equilibria of the system of interest allowing the minimum solvent-to-feed ratio to be estimated. The results show that industrially relevant solvents are selected for different type of mixtures while exploring a much larger solvent space compared to the one delimited by the available experimental data. Also, it is shown that the pre-selection of solvents based on infinite dilution conditions differs from the one performed using the minimum solvent-to-feed ratio criterion. The latter is recommended because it approximates the conditions usually encountered in a real extractive distillation column.

**Keywords:** Graph Neural Network, Activity coefficient, Extractive distillation, Solvent selection.

## 1. Introduction

Separation processes in general, and distillation in particular, account for a significant part of the whole world's energy consumption (Sholl et al., 2016). Despite the efforts in alternative separation technologies, distillation is envisioned to remain as the predominant separation technology in practice during the coming decades. When dealing with close-boiling or azeotropic mixtures, extractive distillation is of particular interest. The reason is that by using an external agent, the entrainer, the thermodynamic equilibrium of the original mixture is conveniently modified, and the separation becomes possible or much more energy-efficient. However, the process of selecting the most appropriate solvent to enhance the separation is far from being straightforward. The



solvent not only has to be thermodynamically powerful, but also economically viable and environmentally responsible.

The process of selecting the best solvent for a given mixture is particularly difficult because, typically, a large number of solvents need to be screened or novel solvents need to be designed from the vast chemical space. Therefore, an efficient pre-selection methodology that reduces the number of candidate solvents which need to be considered for more detailed evaluation is valuable. Brouwer et al. (2021) recently proposed a simple methodology for solvent pre-selection that does not rely on specialized software (e.g., COSMO-RS). Instead, it only uses experimental infinite dilution activity coefficients and the extended Margules equation. This makes it accessible to a broader range of engineers and researchers around the globe that might not have access to specialized software. However, its application is limited by the availability of experimental values for activity coefficients which in turn limits the number of solvents that are considered during the screening. Here, we extend this methodology by using our recently proposed Gibbs-Helmholtz Graph Neural Network approach (GH-GNN) (Sanchez Medina et al., 2022b) as a predictive method for enlarging the solvent search space. This removes the limitation of experimental activity coefficients availability and increases the number of solvents that can be considered during early-stage process design. Moreover, the original method's assumption of partly symmetric Margules parameters is here removed given that all binary mixtures can be potentially calculated by the GH-GNN.

## 2. Background

The relative volatility of the key components  $i$ ,  $j$  in a mixture indicates its ease of separation. At low or moderate pressures, it is defined as the ratio of the pure component vapor pressures  $P^\circ$  multiplied by their activity coefficients  $\gamma$ . Since the ratio of the vapor pressures ( $P_i^\circ/P_j^\circ$ ) within a small temperature range is constant, the modification of the relative volatility, at the same operating pressure, can only be accomplished by the introduction of a solvent that effectively modifies the ratio of the activity coefficients ( $\gamma_i/\gamma_j$ ). This ratio in the presence of a solvent is called selectivity. The maximum impact that a solvent has on the mixture's relative volatility occurs at infinite dilution conditions where the original mixture is infinitely diluted in the solvent. As a result, it is a common practice to rank the efficiency of different solvents according to the value of the selectivity at infinite dilution achieved by such solvents.

Recently, graph neural networks (GNNs) have shown promising performance on predicting activity coefficients at infinite dilution (Sanchez Medina et al. 2022a, Sanchez Medina et al., 2022b, Rittig et al. 2022). In this context, molecular graphs are constructed that contain atomic information (such as the type and hybridization of the atom) in their nodes and chemical bond information (such as the type of bond) in their edges. This information is then processed via a message-passing scheme. Moreover, in the GH-GNN model, global-level features (molecular descriptors describing the polarity and polarizability) are used as part of the learning process of activity coefficients at infinite dilution (Sanchez Medina et al. 2022b). First, the GH-GNN model uses two molecular graphs (corresponding to the solute and solvent in the mixture) that are passed through a molecular-level GNN. Then, a two-node mixture graph is constructed in which each node is defined by the solute and solvent embeddings learned by the first GNN. Intermolecular interaction information is gathered in the mixture graph using hydrogen-bonding information and the global-level embeddings of the solute and solvent graphs. This mixture graph is passed through a second GNN which learns the molecular interactions. Finally, an integrated Gibbs-Helmholtz expression is used to model the temperature-

dependent activity coefficients at infinite dilution. This GH-GNN model is here used to estimate the necessary activity coefficients to perform the solvent pre-selection.

### 3. Methodology

#### 3.1. Pseudo-binary relative volatility at infinite dilution (*inf-RV*)

The relative volatility between the key components  $i$  and  $j$  in the presence of the solvent is called pseudo-binary relative volatility and, at infinite dilution conditions, it is defined as

$$\alpha_{ij}^{\infty} = \frac{P_i^{\circ}}{P_j^{\circ}} \left( \frac{\gamma_i^{\infty}}{\gamma_j^{\infty}} \right)_s \quad (2)$$

where the activity coefficients at infinite dilution  $\gamma^{\infty}$  refer to the solute (i.e., key component  $i$  or  $j$ ) infinitely diluted in the solvent. These coefficients were calculated using the GH-GNN model. The pure component vapor pressures were calculated using the Antoine equation with parameters available from the NIST Chemistry WebBook (2022). The screening was performed over 911 molecular solvents obtained from the original GH-GNN work (Sanchez Medina et al. 2022b) which contrasts with the screening performed by Brouwer et al. (2021) where only 69 molecular solvents were screened due to the restrictions of limited experimental data availability. The pseudo-binary relative volatility at infinite dilution (*inf-RV*) with respect to all solvents was calculated at different temperatures of industrial relevance, and the solvents were ranked according to the mean value across all utilized temperatures. The applicability domain prediction of the GH-GNN model based on chemical classes and Tanimoto similarity was used to only consider predictions within the applicability domain of the model.

#### 3.2. Minimum solvent to feed ratio (*min-SF*)

As pointed out by Brouwer et al. (2021), the selection of solvents based on the pseudo-binary relative volatility at infinite dilution corresponds to situations where the solvent-to-feed ratio is very large. However, often much smaller solvent-to-feed ratios are used in practice. Therefore, the pre-selection of solvents was also carried out according to the minimum solvent-to-feed ratio (*min-SF*) that is necessary to have a pseudo-binary relative volatility equal to 3. This relative volatility value has been recommended as a threshold for which the energy savings due to employing extractive distillation are still considered attractive (Blahušiak, et al. 2018). Typically, if the relative volatility is increased beyond 3 the energy-savings are not as significant and extractive distillation becomes less attractive from an energetic perspective.

The extended Margules model (Eqs. 3-7) (Mukhopadhyay et al., 1993) is here used to estimate the activity coefficients at finite dilution. If one assumes that the ternary interaction parameter  $A_{123}$  is equal to zero, this equation has as its only parameters the activity coefficients at infinite dilution of the species involved which are here predicted by the GH-GNN model.

$$\ln(\gamma_1) = 2(x_1x_2A_{21} + x_1x_3A_{31}) + x_2^2A_{12} + x_3^2A_{13} + x_2x_3B_{123} - 2G^E \quad (3)$$

$$\ln(\gamma_2) = 2(x_2x_3A_{32} + x_2x_1A_{12}) + x_3^2A_{23} + x_1^2A_{21} + x_3x_1B_{123} - 2G^E \quad (4)$$

$$\ln(\gamma_3) = 2(x_3x_1A_{13} + x_3x_2A_{23}) + x_1^2A_{31} + x_2^2A_{32} + x_1x_2B_{123} - 2G^E \quad (5)$$

$$G^E = x_1x_2(x_2A_{12} + x_1A_{21}) + x_1x_3(x_3A_{13} + x_1A_{31}) + x_2x_3(x_3A_{23} + x_2A_{32}) + x_1x_2x_3B_{123} \quad (6)$$

$$B_{123} = 0.5(A_{12} + A_{21} + A_{13} + A_{31} + A_{23} + A_{32}) - A_{123} \text{ with} \quad (7)$$

$$A_{ik} = \ln(\gamma_{ik}^{\infty})$$

Therefore, by employing only the infinite dilution activity coefficients together with the extended Margules model, the vapor-liquid equilibrium behavior of a given mixture, in the presence of a solvent, can be estimated. And the minimum solvent-to-feed ratio (*min-SF*) to reach a relative volatility of 3 can be used as the ranking criteria for pre-selecting solvents.

## 4. Results

### 4.1. Aliphatic-aromatic separations

We studied a collection of binary systems composed of benzene and one of the following alkanes: n-hexane, n-heptane, n-octane, n-nonane and n-decane. The above mentioned two metrics (i.e., *inf-RV* and *min-SF*) were used to rank the set of solvents considered during the screening. It is possible that, these two metrics lead to different solvents during the selection process depending on how the solvent's impact varies at the finite dilution. For instance, Figure 1 shows the top 10 solvents selected by each metric for the n-heptane-benzene system. While some solvents were selected by both metrics (indicated with "(Both)"), most solvents differ depending on the metric used for the selection.

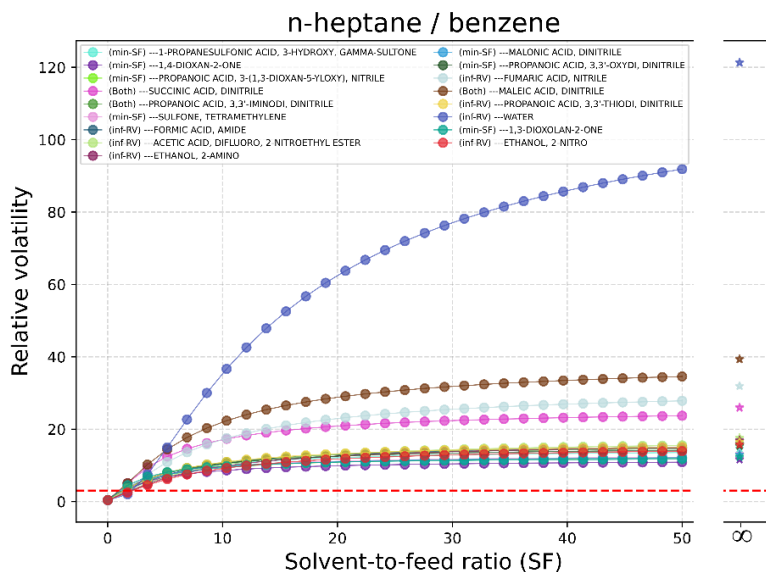


Figure 1. Pseudo-binary relative volatility of the top 10 solvents selected for the mixture n-heptane-benzene using *min-SF* and *inf-RV*. The horizontal red dotted line indicates a pseudo-relative volatility of 3. The star markers denote the corresponding pseudo-binary relative volatilities at infinite dilution.

Table 1. Top 5 solvents selected for the binary mixture benzene and the indicated alkane based on the minimum solvent-to-feed ratio (*min-SF*) to reach a pseudo-binary relative volatility of 3. The range of *min-SF* for the selected solvents is also indicated for each mixture.

Rank	n-hexane <i>min-SF</i> : 0.3	n-heptane <i>min-SF</i> : 1.0-1.2	n-octane <i>min-SF</i> : 2.3-3.7	n-nonane <i>min-SF</i> : 3.3-8.4
1	1,2,3-Tris(2-cyanoethoxy)propane	(2Z)-2-Butenedinitrile	(2Z)-2-Butenedinitrile	Water
2	1,3-Propane sultone	Succinonitrile	Succinonitrile	(2Z)-2-Butenedinitrile
3	(Methylsulfonyl)ethane	Ethylene carbonate	Water	Succinonitrile

4	3,3'-(oxybis(ethyleneoxy))dipropionitrile	Sulfolane	3,3'-Iminodipropanenitrile	Carbamoyl
5	Sulfolane	1,3-Propane sultone	Carbamoyl	(2E)-2-Butenedinitrile

The above occurs whenever the intersection of the pseudo-binary relative volatility trajectory of one solvent intersects the one of another solvent above the threshold of relative volatility equal to 3. For this reason, even though water was selected as the best solvent for this system based on *inf-RV*, it is actually the worst among the 17 solvents depicted on Figure 1 based on *min-SF* (i.e., its impact decays more rapidly as the solvent-to-feed ratio decreases compared to the other solvents). Given that in practice very large solvent-to-feed ratios are not plausible, a better pre-selection criterion should be *min-SF*. Table 1 shows the top 5 solvents selected for each of the considered aromatic-aliphatic binary systems.

As can be seen in Table 1, industrially relevant solvents are often selected which strongly indicates that the proposed methodology is able to pre-select promising candidates. For example, sulfolane (Tilstam, 2012), ethylene carbonate (Mohsen-Nia et al., 2010) and succinonitrile (Tan et al., 2022) are often reported as promising solvents for aliphatic-aromatic separations. For the mixture n-decane/benzene the top 5 selected solvents are similar to the ones for the system with n-nonane with a *min-SF* ranging from 5.6 for the top solvent to 30.5 for 1,4-Butynediol.

#### 4.2. Olefin-paraffin separations

To exemplify olefin-paraffin separations the separation of the following binary mixtures was studied: n-hexane/1-hexene, n-butane/2-butene, n-heptane/1-heptene and n-propane/propene. Table 2 contains the top 5 performing solvents based on the *min-SF* metric along with the ranges needed to achieve a pseudo-binary relative volatility of 3. In the case of the n-heptane/1-heptene mixture only three solvents were found to increase the relative volatility up to 3 (corresponding to the top 3 solvents for the mixture n-hexane/1-hexene). This shows the difficulty of separating olefin-paraffin mixtures. Dimethyl sulfoxide stands out from the selected solvents for the n-propane/propene mixture agreeing with an extractive process already patented (Carter, 1981).

Table 2. Top 5 solvents selected for the shown olefin/paraffin mixtures based on the minimum solvent-to-feed ratio (*min-SF*) to reach a pseudo-binary relative volatility of 3. The range of *min-SF* for the selected solvents is also indicated for each mixture.

Rank	n-hexane/1-hexene <i>min-SF</i> : 5.2-439	n-butane/2-butene <i>min-SF</i> : 1.0-1.3	n-propane/propene <i>min-SF</i> : 1.0-2.5
1	Fumaronitrile	Carbamoyl	Glycerin
2	(2Z)-2-Butenedinitrile	Cyanogen	Dimethyl sulfoxide
3	Succinonitrile	Fumaronitrile	Carbamoyl
4	m-Phenylenediamine	2,4-Dinitrobenzaldehyde	Cyanogen
5	Morpholine	Succinonitrile	1,3-Benzenediamine

## 5. Conclusions and future work

The pre-selection of good candidate solvents for extractive distillation is an important step during chemical process development. Obtaining experimental data for all potential solvents in a large chemical space is practically impossible given the limitations of time and resources to which the exploration is bounded. Therefore, accurate and efficient predictive methods that assist experts in selecting suitable solvents are necessary. Here, we extend the methodology proposed by Brouwer et al. (2021) by using the Gibbs-

Helmholtz Graph Neural Network (GH-GNN) (Sanchez Medina et al., 2022b) together with the extended Margules equation to estimate the activity coefficients of all chemical species in the mixture of interest. With this, the minimum solvent-to-feed ratio needed to increase the relative volatility of the mixture to a certain point can be estimated. The impact of assuming an infinite solvent-to-feed ratio is also discussed by comparing the solvent selection obtained by the relative volatility at infinite dilution to the one obtained by using the minimum solvent-to-feed ratio metric. The latter is recommended because it approximates the real conditions of an extractive distillation column better. Industrially relevant solvents were selected using the proposed methodology for both aliphatic/aromatic and olefin/paraffin separations showing the validity of the proposed framework. The proposed methodology has been made open-source at [https://github.com/edgarsmdn/SolvSelect\\_GHGNN](https://github.com/edgarsmdn/SolvSelect_GHGNN).

**Acknowledgments:** Edgar Ivan Sanchez Medina is also affiliated with the International Max Planck Research School for Advanced Methods in Process and Systems Engineering - IMPRS ProEng at the Max Planck Institute for Dynamics of Complex Technical Systems Magdeburg.

## References

- Blahušiak, M., Kiss, A. A., Babic, K., Kersten, S. R., Bargeman, G., Schuur, B., 2018. Insights into the selection and design of fluid separation processes. *Separation and purification technology*, 194, pp. 301-318.
- Brouwer, T., Kersten, S.R., Bargeman, G. and Schuur, B., 2021. Solvent pre-selection for extractive distillation using infinite dilution activity coefficients and the three-component Margules equation. *Separation and purification technology*, 276, pp. 119230.
- Carter, C. O., 1981. U.S. Patent No. 4,267,034. US4267034A. Washington, DC: U.S. Patent and Trademark Office.
- Mohsen-Nia, M., and Doulabi, F. M., 2010. Separation of aromatic hydrocarbons (toluene or benzene) from aliphatic hydrocarbon (n-heptane) by extraction with ethylene carbonate. *The Journal of Chemical Thermodynamics*, 42(10), pp. 1281-1285.
- Mukhopadhyay, B., Basu, S., Holdaway, M. J., 1993. A discussion of Margules-type formulations for multicomponent solutions with a generalized approach. *Geochimica et Cosmochimica Acta*, 57(2), pp. 277-283.
- NIST Chemistry WebBook, 2022. NIST Standard Reference Database Number 69. Accessed on 02 October 2022 at <https://doi.org/10.18434/T4D303>.
- Rittig, J.G., Hicham, K.B., Schweidtmann, A.M., Dahmen, M. and Mitsos, A., 2022. Graph neural networks for temperature-dependent activity coefficient prediction of solutes in ionic liquids. *arXiv preprint arXiv:2206.11776*.
- Sanchez Medina, E. I., Linke, S., Stoll, M., Sundmacher, K., 2022a. Graph neural networks for the prediction of infinite dilution activity coefficients. *Digital Discovery*, 1(3), pp. 216-225.
- Sanchez Medina, E. I., Linke, S., Stoll, M., Sundmacher, K., 2022b. Gibbs-Helmholtz Graph Neural Network: capturing the temperature dependency of infinite dilution activity coefficients. *arXiv preprint arXiv:2212.01199*.
- Sholl, D.S. and Lively, R.P., 2016. Seven chemical separations to change the world. *Nature*, 532(7600), pp. 435-437.
- Tan, T., Cheng, H., Chen, G., Song, Z. and Qi, Z., 2022. Prediction of infinite-dilution activity coefficients with neural collaborative filtering. *AIChE Journal*, 68(9), pp.e17789.
- Tilstam, U., 2012. Sulfolane: A versatile dipolar aprotic solvent. *Organic Process Research & Development*, 16(7), pp. 1273-1278.
- Yang, D., Zhang, S., & Jiang, D. E., 2019. Efficient absorption of SO<sub>2</sub> by deep eutectic solvents formed by biobased aprotic organic compound succinonitrile and 1-ethyl-3-methylimidazolium chloride. *ACS Sustainable Chemistry & Engineering*, 7(10), pp. 9086-9091.

# **Dhrushti-AI: A multi-screen multi-user eye-tracking system to understand the cognitive behavior of humans in process industries**

Thasnimol Valuthottiyil Shajahan,<sup>a</sup> Rahul Madbhavi,<sup>a</sup> Mohammed Aatif Shahab,<sup>a</sup> Babji Srinivasan,<sup>a,c</sup> Rajagopalan Srinivasan<sup>b,c</sup>

<sup>a</sup>*Department of Applied Mechanics, Indian Institute of Technology Madras, Chennai, 600036, India*

<sup>b</sup>*Department of Chemical Engineering, Indian Institute of Technology Madras, Chennai, 600036, India*

<sup>c</sup>*American Express Lab for Data Analytics, Risk and Technology, Indian Institute of Technology Madras, Chennai, 600036, India*

[babji.srinivasan@iitm.ac.in](mailto:babji.srinivasan@iitm.ac.in), [raj@iitm.ac.in](mailto:raj@iitm.ac.in)

## **Abstract**

Operator performance is key to ensuring safety in process industries. Therefore, a comprehensive assessment of their performance is critical for smooth and efficient plant operation. Traditional performance assessments are not comprehensive as these ignore cognitive aspects of performance. On the other hand, while eye-tracking-based approaches do provide a cognitive assessment during operating training, their applicability in a real-time real setting is limited. Existing eye-tracking systems come with many constraints affecting users' mobility (restricted movement in all directions) in their environment. In addition, it is beyond the scope of current eye trackers to track multiple users working in an environment, such as in control rooms. Satisfying the above requirements makes these eye-trackers expensive. In this work, we demonstrate the capabilities of an in-house developed cost-effective eye-tracking system to track users' eye movement in an unconstrained environment while giving freedom of head movement. Human subject studies are conducted to compare operators' gaze patterns and the quality of data with that obtained using commercial eye trackers. The results generally agree with data quality obtained using commercial eye trackers. Hence, the performance of the developed eye-tracking system is comparable to existing commercial eye-trackers while overcoming their limitations, such as restricted user movement, single-user tracking, and high cost.

**Keywords:** Operator performance, Eye-tracking, Cognitive behavior, Human error, Safety

## **1. Introduction**

The use of innovative automation methods and safety management systems has helped the process industries become safer. Nonetheless, accidents of varying severity continue to occur. Studies indicate that human error is the primary cause of 75% of industrial accidents (Jung et al., 2020). The safe running of the plant is guaranteed by cooperation between the human and automation system. Operators must constantly keep track of the functioning of several control mechanisms and step in when the automation fails to provide as expected. However, the rise of digitization and the complexity of automation

have changed the operator's role and added new difficulties (Nazir et al., 2014). To ensure optimum operating conditions, operators must collect, prioritize, and utilize data from several information sources. This adds to their cognitive workload and worsens if the operators are not well trained. Therefore, keeping track of the operator's performance is crucial to ensure safe plant operation.

Addressing operators' performance issues involves improving their competency and continuously observing their behavior in the field (Dai et al., 2016). There are hardly any studies focusing on monitoring operator performance in real time. For operators to become more competent, research has largely focused on providing them with the proper feedback based on their performance. Most performance assessment methods depend on subjective expert opinion that disregards the operators' cognitive behavior (Nazir et al., 2015). There is widespread agreement that cognitive behavior plays a significant role in unusual circumstances and accident causation.

In recent years, researchers have used eye-tracking to get insights into human cognitive behavior. It has been extensively used in understanding human cognitive processes such as situation awareness, expert level, learning ability, and mental state of the operator in several safety-critical domains such as healthcare, aviation, automobile driving, and nuclear power plant (Srinivasan et al., 2019). This work shows the potential of an in-house developed eye-tracker (Dhrushti-AI) capable of monitoring human performance in an unconstrained environment. Human subject studies were conducted to compare the quality of data obtained from Dhrushti-AI and commercial eye trackers.

Dhrushti-AI can enhance safety by addressing the challenges associated with operator performance during training and in real-time plant operation. Firstly, it can be used along with operator training simulators to identify flaws in the operator's understanding of the process during training. Secondly, it can be deployed in real control room settings to monitor the operator's performance in real-time. In addition, we can enhance operators' situational awareness based on the insights from eye-tracking data. For instance, if operators miss any critical information from HMI (Human Machine Interface) appropriate feedback can be provided. Adequate situational awareness and skills are crucial, especially when dealing with abnormal situations (Bhavsar et al., 2017).

## 2. Literature Review

Eye tracking is a technique to record a person's eye movement. It provides eye gaze coordinates (where the person is looking) and pupil size. There are two primary categories of eye trackers in the market: remote and wearable. The latter must be worn, whilst the former may be positioned close to the test screen. Wearable eye trackers are appropriate when the scenario is not constrained to a narrow space and the test participants are allowed to walk about the test setting. However, it may create discomfort for the user to wear a device for a long time, such as in a control room setting where operators work for long hours. On the other hand, a remote eye-tracker is set up in the user environment and it doesn't hinder the subject. Current key market players are Tobii, Smart Eye, SMI (SensoMotoric Instruments), and SR Research (Hermens et al., 2013).

Eye-tracking measures have been found to correlate with various aspects of human performance in safety-critical domains. An eye-tracking study using SMI Eye-tracking Glasses was conducted with air traffic controllers to capture their mental workload (Rodríguez et al., 2015). The study reveals that increased pupil diameter is associated with an increased mental workload. Another study found that changes in mental workload

can be measured using frequency domain analysis of pupil size measurements. An increase in workload is connected with an increase in power level at low frequencies (Peysakhovich et al., 2015). The experiments are conducted by using a single-screen setup. During the whole experiment, participants' gaze position and pupil diameter were recorded with a remote SMI RED eye-tracker (SensoMotoric Instruments GmbH, Germany) at a sampling rate of 120 Hz. Wu et al. (2020) used eye-tracking to analyze the difference in cognitive behavior of 32 novices and 7 expert operators when doing activities involving standard operating procedures of accident scenarios in nuclear power plants (NPPs). The study uses iView X head-mounted eye tracking device (SMI, Germany) at a sampling rate of 50 Hz to capture the eye movements of operators. It was found that novice operators had a greater fixation rate, illustrating their intense search in various places and inexperience in tackling process disturbances.

Our previous studies with control room operators reveal distinct mental models of expert and novice operators (Shahab et al., 2021). During an abnormal situation, expert operators have higher dwell time on important regions of HMI than novices (Sharma et al., 2016). Further, operators who failed to complete a disturbance task have higher gaze entropy indicating that their eye gaze is distributed in several regions on the HMI unrelated to the disturbance (Bhavsar et al., 2017). These approaches provide interesting insights into operators' cognition and help identify the reason behind human failure. These eye-tracking studies were conducted using Tobii TX300 eye tracker to record participant eye movements at sampling rates of 120 Hz.

In summary, existing commercial eye-tracking systems support single-user and single-screen HMI with limited user movement. However, this is not the case in real control rooms where the operator needs to gather information displayed on multiple screens and involves a lot of head movements. In addition, satisfying the above requirements makes these eye-trackers expensive. In this work, we demonstrate the potential of an in-house developed cost-effective eye-tracker system (Dhrushti-AI) to record multiple users' eye movements from an unconstrained environment.

### 3. Methodology

In this work, we demonstrate the capabilities of an eye-tracker (Dhrushti-AI) developed to track users' gazes in the presence of head movements in unconstrained environments. The eye-tracker comprises an array of cameras fitted with lenses and filters, illuminators, an optics controller, power sources, and a computing unit. The camera array observes a region of the environment to detect users' faces and track their gazes, referred to as the headbox. The eye-tracking system illuminates the headbox with light in the near-infrared region using the illuminators. The optics controller generates the trigger signals to control the camera exposure and the illuminators. The compute unit detects the faces in the images obtained from the camera array using Histogram of Oriented Gradients (HOGs) based face detectors. Face recognition algorithms are utilized to identify faces and enable

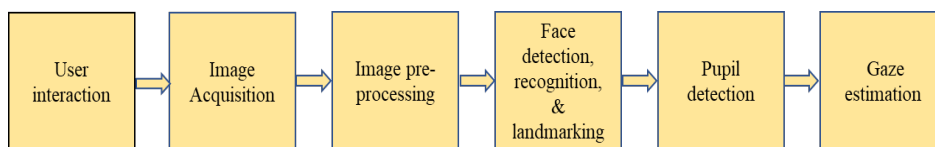


Fig. 1. Steps involved in the proposed eye-tracking system



the application of user-specific calibrated gaze models. Key landmarks in the faces (such as corners of the eyes, nose, etc.) are extracted and are used to estimate the users' head orientation. The pupils from these images are detected using several image processing techniques, such as thresholding, closing, and contour extraction. The eye-tracker utilizes the head orientation, pupil center locations, and the user-specific calibrated gaze model to estimate the gaze direction. The output of the developed eye-tracker is the user's gaze location (coordinates on screen) and pupil size variation. The overall methodology to get the output is shown in Fig. 1. The raw gaze data from Dhrushti-AI is processed further to obtain fixation and saccades using the Velocity-Threshold Identification (I-VT) filter.

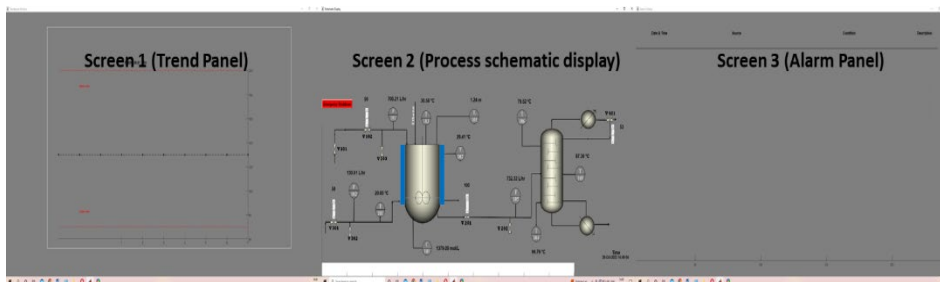


Fig. 2. HMI with the schematic and display units for three screens.

To demonstrate the applicability of the developed multi-screen eye-tracker (Dhrushti-AI), we used a simulated chemical process plant with which operators interacted via HMI. The HMI is presented to the operator on a display comprising three screens. Eight participants volunteered to play the role of the control room operator. They were asked to monitor and intervene during abnormalities. The participants first interacted with the process via a single-screen HMI of the chemical process displayed on the center screen. During this period, the other two screens were blank. In the next study, we displayed the HMI on all three screens. A typical task involves bringing a process variable(s) to within the acceptable limits as specified by the region between the upper limit and lower limit in Fig. 2. Eye-tracker calibration for each participant was performed at the beginning of the experiment. Next, we compare the results obtained from Dhrushti-AI and a commercial eye-tracker.

#### 4. Result & Discussion

In this section, we illustrate the potential of Dhrushti-AI in capturing user eye movements in an unconstrained environment (typical control room setup) without restricting their head movements. The eye-tracker was set up to monitor user head movements at a distance of 45-60 cm from the center camera while allowing 25 cm horizontal movement and 35 cm vertical movement. Initial results have been obtained with image capture rates of 30 frames per second (FPS), while the hardware capabilities allow image capture at rates higher than 120 FPS. Gaze angle errors of less than  $1.5^\circ$  were obtained.

As discussed earlier, we conducted experiments with human subjects using Dhrushti-AI and commercial eye-tracker Smart-eye Aurora. We compared the quality of data obtained from the eye-tracker using data loss (proportion of missing gaze samples including blinks; (Holmqvist, 2017) and proportion of all fixation durations (dwell time) outside of various information sources, i.e. AOI's (Areas of interest) on the HMI. When dealing with abnormal situations operators primarily focus on various information sources which are relevant to the situation at hand (Sharma et al., 2016). Therefore, it is expected that the

proportion of fixation on non-information sources obtained from Dhrushti-AI should be minimum and comparable to commercial eye trackers.

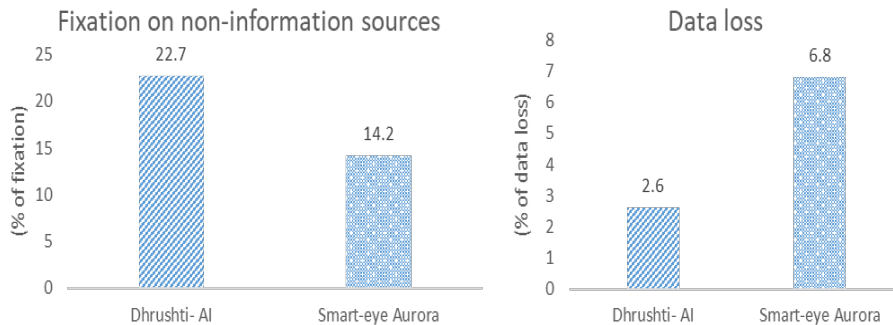


Fig. 3. Comparison of results obtained from Eye-trackers (Dhrushti-AI and Smart Eye Aurora): based on (a) Fixation on non-information sources, (b) Data loss

The results for both Dhrushti-AI and Smart Eye Aurora obtained from eight participants are shown in Fig. 3a and 3b. As observed from Fig. 3, the data loss for Dhrushti-AI is lesser compared to the smart-eye aurora (Fig.3b). This signifies that even with a significant head movement of participants, Dhrushti-AI is able to capture the operator's gaze position from all three screens. Similarly, the proportion of time on non-information sources obtained from Dhrushti-AI is also similar to that of Smart Eye Aurora (Fig.3b). It is important to note that few of the AOIs on the HMI were smaller and even with such lower size, Dhrushti-AI indicates that operators were looking at the AOIs most of the time.

In addition to the data quality, we also compare the distribution of dwell time on AOIs for operators who successfully completed the task and those who failed. The results indicate that experts have higher dwell time on key variables and exhibit lower gaze entropy than novices. These results are in general agreement with earlier works (Sharma et al., 2016; Bhavsar et al., 2017; Shahab et al., 2021) conducted using commercial eye trackers. In light of these results, the performance of the developed eye-tracking system (Dhrushti-AI) is comparable to existing commercial eye-trackers while overcoming their limitations such as restricted head movement, single screen setup, and high cost. Thus, the developed eye tracker has the potential to be deployed in control rooms for assessing operator performance.

## 5. Conclusion

In this work, we demonstrated the potential of a multi-screen- multi-user eye tracking system (Dhrushti-AI) in capturing operator cognitive behavior in a process plant setting. There is widespread agreement that cognitive behavior plays a substantial role in the causes of abnormal situations and accidents. The capabilities of Dhrushti-AI are demonstrated with a 3- screen workstation as the HMI. We conducted experimental studies to validate the quality of the developed eye tracker with that of the commercially available eye-trackers. Results indicate that the performance of the Dhrushti-AI is lesser than the existing commercial eye trackers, in terms of data loss. The operator's cognitive

behavior obtained from Dhrushti-AI is also in general agreement with our earlier works (Sharma et al., 2016) with research-grade eye trackers (smart eye Aurora). Thus, the developed multi-screen eye-tracker can be deployed in actual control rooms to infer the cognitive behavior of operators previously confined to controlled environments. Additionally, it can identify flaws in the thought process of novice operators during training and help them in acquiring knowledge about the process dynamics. In the future, we can increase the system's potential to meet the requirements of large-scale industrial studies.

## References

- Bhavsar, P., Srinivasan, B., & Srinivasan, R. (2017). Quantifying situation awareness of control room operators using eye-gaze behavior. *Computers & chemical engineering*, 106, 191-201.
- Dai, Y., Wang, H., Khan, F., & Zhao, J. (2016). Abnormal situation management for smart chemical process operation. *Current opinion in chemical engineering*, 14, 49-55.
- Holmqvist, K. (2017). Common predictors of accuracy, precision, and data loss in 12 eye-trackers. In *The 7th Scandinavian Workshop on Eye Tracking*.
- Jung, S., Woo, J., & Kang, C. (2020). Analysis of severe industrial accidents caused by hazardous chemicals in South Korea from January 2008 to June 2018. *Safety Science*, 124, 104580.
- Nazir, S., Kluge, A., & Manca, D. (2014). Automation in the process industry: Cure or curse? How can training improve operator's performance. In *Computer Aided Chemical Engineering (Vol. 33, pp. 889-894)*. Elsevier. doi: 10.1016/B978-0-444-63456-6.50149-6
- Nazir, S., Øvergård, K. I., & Yang, Z. (2015). Towards effective training for process and maritime industries. *Procedia Manufacturing*, 3, 1519-1526. doi: 10.1016/j.promfg.2015.07.409
- Peysakhovich, V., Causse, M., Scannella, S. and Dehais, F., 2015. Frequency analysis of a task evoked pupillary response: Luminance-independent measure of mental effort. *International Journal of Psychophysiology*, 97(1), pp.30-37.
- Rodríguez, S., Sánchez, L., López, P., & Cañas, J. J. (2015, September). Pupillometry to assess Air Traffic Controller workload through the Mental Workload Model. In *Proceedings of the 5th International Conference on Application and Theory of Automation in Command and Control Systems (pp. 95-104)*. ACM.
- Shahab, M. A., Iqbal, M. U., Srinivasan, B., & Srinivasan, R. (2022). HMM-based models of control room operator's cognition during process abnormalities. 1. Formalism and model identification. *Journal of Loss Prevention in the Process Industries*, 76, 104748.
- Shahab, M. A., Srinivasan, B., & Srinivasan, R. (2021, December). Analysis of Control Room Operators' Competence using Cognitive Engineering Approaches to Improve Process Safety. In *2021 International Conference on Maintenance and Intelligent Asset Management (ICMIAM) (pp. 1-6)*. IEEE.
- Sharma, C., Bhavsar, P., Srinivasan, B., & Srinivasan, R. (2016). Eye gaze movement studies of control room operators: A novel approach to improve process safety. *Computers & Chemical Engineering*, 85, 43-57.
- Srinivasan, R., Srinivasan, B., Iqbal, M. U., Nemet, A., & Kravanja, Z. (2019). Recent developments towards enhancing process safety: Inherent safety and cognitive engineering. *Computers & Chemical Engineering*, 128, 364-383.
- Wu, Y., Liu, Z., Jia, M., Tran, C. C., & Yan, S. (2020). Using artificial neural networks for predicting mental workload in nuclear power plants based on eye tracking. *Nuclear Technology*, 206(1), 94-106.

# Digital Twin- A System for Testing and Training

Michael Schueler<sup>a</sup>, Tanja Mehling<sup>b</sup>

<sup>a</sup> *Siemens AG, DI PA SW ID MBS OA , Industriepark Hoechst, B598, 65926 Frankfurt*  
[michael.schueler@siemens.com](mailto:michael.schueler@siemens.com)

<sup>b</sup> *Siemens AG, DI PA SW ID MBS DS, Industriepark Hoechst, B598, 65926 Frankfurt*  
[tanja.mehling@siemens.com](mailto:tanja.mehling@siemens.com)

## Abstract

An Operator Training System (OTS) is a digital twin of a process plant, that enables to study the plant before it really exists. That does not just give an opportunity to train the operators of the plant, before plant startup, but it also helps to get a better understanding of the process.

In this work, an OTS of a novel process to produce biochemicals from wood was developed. Aiming to train the staff while the plant is still under construction and to gain a deep knowledge of this plant, a digital twin of the plant was built. According to the requirements for the training, a mid-fidelity process model was combined with a high-fidelity model.

To finalize the OTS, the process model was connected to the original Distributed Control System (DCS) environment.

It can be shown that the development of a process plant digital twin is highly effective in accelerating the commissioning of a new plant. On the one hand, the operators are enabled to gain a deep knowledge of the process and the operation of the plant even before the plant is finalized. On the other hand, the detailed simulation gives valuable insights to the process, which help to understand the process, the conditions, and the regulation concept.

**Keywords:** Digital Twin, Operator Training System, dynamic Simulation.

## 1. Introduction

In the recent years, Digital Twins became omnipresent in a huge variety of applications. They are of major contribution to the digitalization of the industries, still one of the actual megatrends. Of course, also the process industry applies Digital Twins over the life cycle of a production plant like described by Labisch (2019). An Operator Training System (OTS) is a digital twin of a process plant, that enables to study the plant even before it really exists. That does not just give an opportunity to train the designated operators of the plant, before plant startup, but it also helps to get a better understanding of the process and the plants functionalities as investigated by NAMUR (2006).

The aim of an OTS as a Digital Twin is, to reproduce the physical plant as well as the real drive and control mechanisms, the automation and operation. Both, the real plant and the virtual plant, i.e. the Digital Twin are compared in Figure 1.

The overall plant including the automation can be separated in five layers. For the real plant the first layer is the real process plant. The process plant contains all units like columns, reactors, vessels pumps, valves, pipes and so on.

The second layer of the real system contains all actuators and sensors. Actuators are typically pump motors and valve drives. The sensors measure all relevant process data like temperature, pressure, flow, and composition.

The third layer with the remote IOs (Input/Output) describes the operations or assets, transferring the data from the second layer (sensors and actuators) to the automation system (AS), the fourth layer.

Finally, the automation system is connected to the human machine interface (HMI). The fifth layer. By means of the HMI the state of the process and the plant is visualized, and the operator can control the plant.

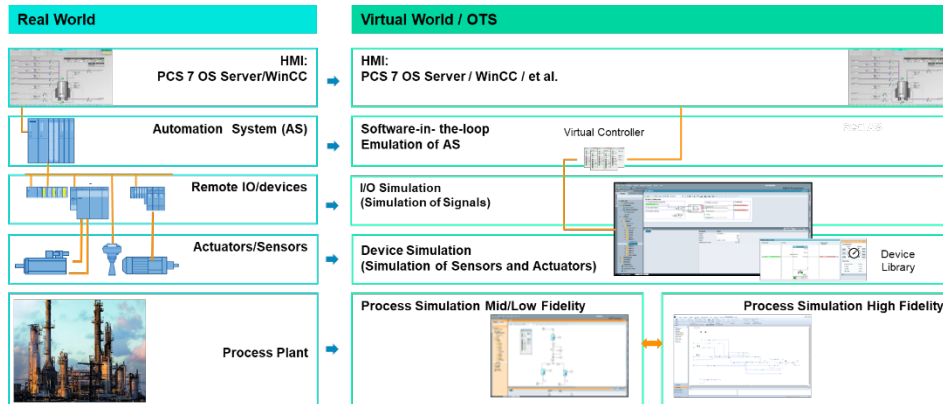


Figure 1: structure of a real plant and an OTS

Analogue to the overall real plant, an operator training system as a digital twin of a real plant has the same five layers. But in contrast to the real plant, most of the five layers are virtualized in an operator training system.

The basis for an operator training system is the process simulation. This virtual plant represents the real process plant (layer one). By means of the process model the behavior of the plant is simulated. For an operator training system mainly dynamic process models are used. With a dynamic process model the transient behavior of the process can be described. Dynamic simulations can be quite computationally intensive. Thus, it is aimed, to model as accurate as possible but only as accurate as needed, to save time and resources. Based on the needs for the training and considering the computational effort the fidelity of the process model can be selected. Even the combination of low or mid fidelity models with high fidelity models is possible. Within the process model all necessary process measurements are calculated and transferred to the next layer.

In this layer, the device simulation, the behavior of the measurement devices and actuators is simulated. The time dependent behavior of valve drives and motors as well as all necessary feedback signals for the automation system are calculated. Based on a library containing the necessary device modules, the relevant device simulation model can be composed.

The device Simulation is connected to the IO simulation (third level). In this level the signals from the device simulation are transferred to the fourth level, the emulation of the automation system and vice versa. The IO-level can also be used without the first two layers (process simulation and device simulation) to generate signals for testing the automation system.

The automation system itself is emulated in the fourth layer. With virtual controllers running on standard personal computers (PCs) the behavior of the automation system is simulated. The configuration of the automation system of the real plant can be loaded to the virtual controller without any change. The automation is executed and behaves like the real automation system.

Finally, as in the real plant control, the automation system of the Digital Twin is connected to the original human machine interface (HMI) of the distributed control system (DCS). In this fifth layer the process can be controlled with the very same operating screens like in the control room. Thus, with the OTS, the operator will not realize the differences between the Digital Twin and the real plant. Neither the operating system, i.e. the HMI, nor the effect of operating actions differ from the real plant. The only difference might be the actual surrounding of the operator, which might be anywhere but is not limited to the real control room.

In the current work, an example of an OTS is given as develop and implemented for a real process. The process itself is described, the proceeding of the development and the benefits are shown.

We think the correct split of model fidelity is an essential decision in the design of an OTS. Furthermore the use of tools that can direct reuse the original data of the automation system reduces the effort of design.

## 2. Development of the Operator Training System.

In this work, an OTS was realized, representing a novel process to produce biochemicals from wood. Aiming to train the staff while the plant is still under construction and to gain a deep knowledge of this plant and the corresponding process, a digital twin of the plant was built. According to the requirements for the training, to reproduce the real process and to consider the complexity of the process, a mid-fidelity process model was combined with a high-fidelity model. To finalize the OTS, the process model was connected to the original DCS environment. Thus, including the sensor and actor level, virtual controllers are applied to train the effect of interferences while the process is running, but also to train the startup, shutdown, disturbances, and an emergency case.

In the next section the process itself is briefly presented. Furthermore, the modeling of the process and it challenges are described.

### 2.1. Biochemical Process

The biochemical process is divided in several main steps (see Figure 2). In a first step the tree trunks are grinded into chips. These chips are washed and prepared for the next process steps. The wood is then enzymatically converted into intermediate products. These intermediate products are separated from the slurry and purified. Further, these intermediate products are chemically converted to the final product. In a last step the biochemicals are purified in a distillation train to achieve the required purities. The final purified products are stored in several tanks.

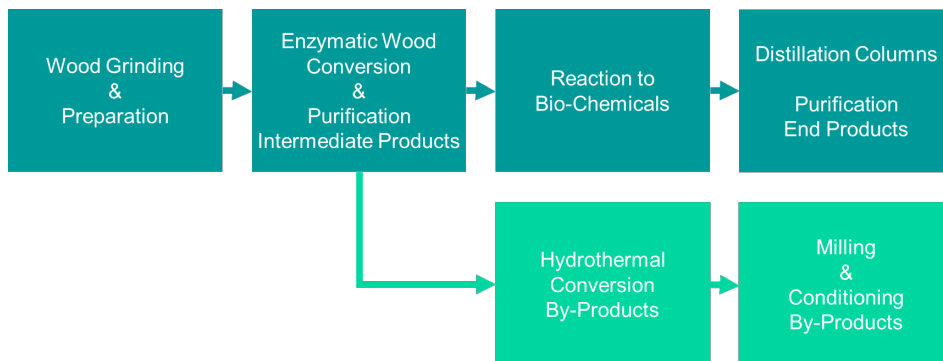


Figure 2: Block flow diagram of the process

The residual components separated from the slurry after the enzymatic conversion are further processed. This complex mixture is converted hydrothermally. These side products are further conditioned by drying and milling and the final byproducts are stored in tanks for further distribution.

## 2.2. Process modelling

To create a detailed but economic OTS, the process was divided in many individual process steps. Some of the process steps require an intensive, high-fidelity simulation, while others can be represented by less intensive mid-fidelity models. Some process steps were not considered at all, since their complexity is very low, the operation is well known and the benefit of a Digital Twin is comparatively low.

Therefore, the wood unloading, storage and coarse chopping are not represented in the Digital Twin. Furthermore, since the process model is focused on the main operation of the plant, cleaning facilities (cleaning in place (CIP) systems) and safety relief valves, sampling points and drain valves were neglected.

Simple models, thus, a mid-fidelity simulation was chosen for plant components such as tank farms, truck unloading stations and storage. Furthermore, utility systems, if relevant for the process simulation, are also simulated with simple models. Complex chemical processes such as reactions or phase changes typically do not take place in these parts of the plant, so that it is reasonable, to apply the simpler models in these cases.

In contrast, plant sections with reactions and apparatuses in which the phase equilibrium plays a decisive role (e.g. distillation columns) are simulated with high fidelity models. But not only the necessary complexity was decisive for the division into mid fidelity or high-fidelity models, but also that there are not too many interfaces between the two simulators.

### 2.2.1. High fidelity process model

Decisive for each process model is the application of reliable and evaluated physical properties of the participating components. Within this work, property packages adjusted to the requirements of the single process steps were defined for the high-fidelity models. On the one hand, all the necessary components for the various process steps had to be available. On the other hand, the number of components should be kept low, to keep the simulation simple and therefore not too slow. Some of the physical property packages contain up to about 50 different components. This is necessary to consider all the relevant educts and products of a reaction as well as solvents and components effecting the equilibrium or the reaction conditions like pH value and so on. Especially the many different materials, which are part of the natural raw material wood behave differently throughout the single process steps and must be accounted for individually. Nevertheless, the component list was simplified, wherever possible, to keep the simulation as simple as possible. In any case, the property packages were optimized to reduce the simulation effort, but precisely describe the phase behavior and the heat and mass balance at the same time. This was achieved by using the non-random-two-liquid model (NRTL model). Another challenge beside the large number of components was to get all the necessary parameters in the NRTL model. For well-known components like water these parameters are known. For other components like the components of wood these parameter, not only the pure component parameters but also the binary interaction parameters, must be adapted to the existing measurements.

An accurate simulation also requires a detailed description of the reaction kinetics in dependence of time and conditions (e.g. temperature, pressure or pH value). In the described process, several conversions (enzymatical, chemical, and hydrothermal) need

to be considered. All of the reactions are complex and depend on the detailed composition of the raw material, which cannot be assumed to be constant. Accordingly, the detailed mechanism for these very complex reactions is not yet investigated in such a detail, as the input for the simulation requires. Thus, based on lab trials and data from pilot plans, the key influencing parameters for each conversion step was identified. In simplified descriptions, the effect of those parameters on the conversion rate was defined and the respective kinetics was implemented into the simulation.

Beside a reliable physical property package and adequate reaction mechanisms, models for the simulation of the single unit operations are required. For many parts of the plant, the model could be created with the standard elements from the simulation library. The library contains, for example, devices such as pumps, valves, tanks, heat exchangers and distillation columns.

However, there are also devices/plant parts that could not be created directly with devices from the library. For example, a model for a filter press had to be developed in-house. The filter press runs continuously, operating in different steps. Thus, the filling with suspension, dehumidifying by pressing, washing with fresh water, repeated pressing, and eventually further drying with compressed air need to be covered by the developed model. All these steps were implemented with a simplified substitute model, using the available library models like component splitters, pumps, and tanks, linked, controlled, and coordinated with an adjusted logic.

A tailor-made model also had to be implemented in case of the reactors. Although different reactors are available in the library, the kinetics derived for the reactions could not simply be integrated into them. In most cases the reactor was represented by a simple tank. The reaction is then calculated in a spreadsheet, taking the current state of the influencing reaction conditions into account. The actual material conversion and heat of reaction is then implemented by a virtual recirculated material stream.

In total, the entire high fidelity dynamic process model includes about 200 process and instrumentation diagrams (PIDs) with the corresponding devices. The total process model has been divided into around 15 individual models, so that the complete model can also be calculated in real time during the transitions from one state to the other with sufficient accuracy and with sufficiently small simulation steps. These individual models are simulated in parallel. The individual models are connected to each other so that the streams are transmitted from one model to the other at the interfaces.

### *2.2.2. Mid fidelity process model*

Compared to the simulation with the high-fidelity model, the mid-fidelity model uses simplified approaches. A maximum of 20 different components are taken into account in the mid fidelity model, only. In addition, ideal gases or ideal liquids are assumed for the physical properties. Phase changes are not taken into account. These simplified assumptions are sufficient for the plant parts to be modeled like storage tanks, truck unloading and loading stations. All of the used models are part of the available libraries.

### *2.2.3. Model integration*

The individual sub models of the high-fidelity process model are connected to the mid fidelity simulation. The corresponding information for the streams is exchanged at the transitions. The entire plant is thus simulated with the two types of process models. In addition, the process models are linked to the device level (layer two). In the device level the manipulated variable for the actuators in the process model is calculated, such as the valve position or the engine speed. In the same way, the corresponding signals for the control technology are calculated from the measured values of the process model.



### 3. Benefits of the OTS

The aim of developing the OTS for the described process, was to be able to train the future operators already while the plant is still under construction. Nevertheless, additional benefits were generated also on the way to develop the OTS. Important insights into the process could already be gained in the first phases of the project. On the one hand, the process design was improved for the distillation columns. The column design was based on the heat and mass balance generated from the steady state simulation, fitted to lab experiments and pilot plant data. Since manual adjustments, applied in the steady state simulation, cannot be used for dynamic simulation, a more general and robust approach was chosen. With improved physical properties, the simulation could be better adapted to the existing measurement results. However, it turns out that the required purities could not be achieved in one separation step with the chosen design of the column. Based on this knowledge, the design of the column could be adjusted just before ordering and a potentially expensive redesign during commissioning could be avoided. Normally such a redesign is not done, and the equipment is specified based on the heat and mass balance. In addition, the control concept could be adapted and improved based on the dynamic simulation, especially in the area of the distillation columns. Based on the dynamic behavior of the system, the control concepts could be optimized. Above all, the interactions between several columns in a row could be considered. Furthermore, it was possible to define the necessary number and position of complex online measurement methods to determine the purity.

### 4. Conclusion

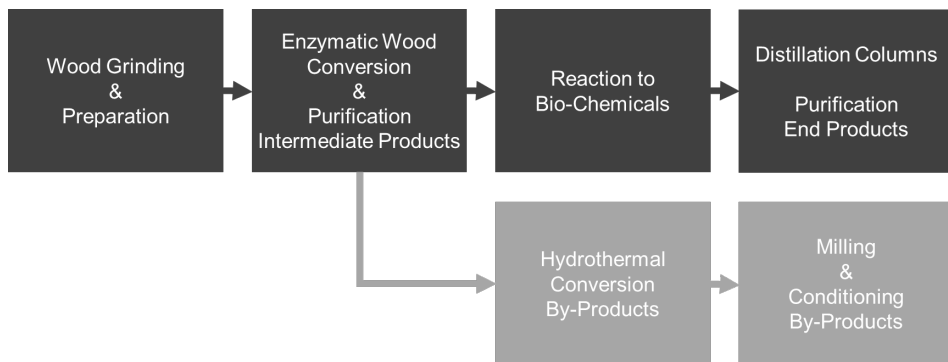
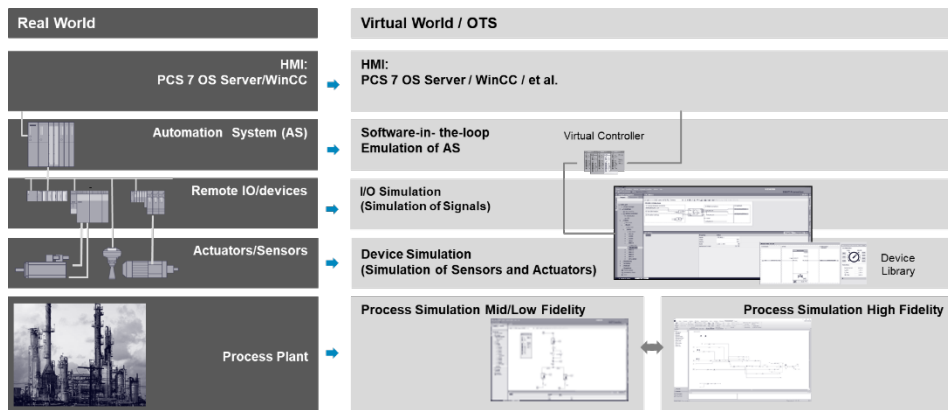
In this work, an example for a successful development of an OTS is given. By the combination of mid- and high-fidelity models, an accurate, fast simulation of a novel, high complex and comprehensive process was realized. The OTS was successfully applied, to train the operators prior to the commissioning.

Furthermore, it was shown, that a digital twin in the form of an operator training system (OTS) can not only be used to train the plant operator. The normal operation of the plant as well as the start-up and shutdown of the plant can be practiced without risk. In addition, a wide variety of malfunctions and emergency situations can be simulated and the optimal behavior in these cases can be practiced again and again. All of this can also be done before the system is actually put into operation.

In addition, however, the dynamic simulation for an operator training system also provides valuable information in relation to system design and control concepts. In this way, knowledge about the later operation of the system can be gained in the early phases and can be incorporated into the construction of the system.

### References

- D. Labisch, Ch. Leingang, O. Lorenz, M. Oppelt, B-M. Pfeiffer, F. Pohmer, 2019, Evolution eines Digital Twin am Beispiel einer Ethylen-Anlage, atp magazin, 06/07 2019
- Schulze, K., 2014, Trainingssimulation in der Prozessindustrie. Atp magazin, 56(01-02),
- Cox, R. K., Smith, J. F., und Dimitratos, Y., 2006, Can simulation technology enable a paradigm shift in process control?: Modeling for the rest of us. Computers & chemical engineering, 30(10-12), (pp.1542-1552).
- NAMUR, 2006, NA 60: Management von Trainingssimulatorprojekten. NAMUR: [www.namur.net](http://www.namur.net)
- Siemens AG, Operator Training System (OTS) – Für effizientere Analogennutzung, <https://assets.new.siemens.com/siemens/assets/api/uuid:4a5f99a1-8f40-4614-b468-0ca9d9a2f645/285-operator-training-systems-de.pdf>





# Process Validation in Modular Plants considering the Scale-Up and Prior Knowledge

J. Mädler<sup>a\*</sup>, A. Koch<sup>a</sup>, A. Bamberg<sup>b</sup>, L. Urbas<sup>a</sup>

<sup>a</sup>*Chair of Process Control Systems & Process Systems Engineering Group,  
TU Dresden, Dresden 01062, Germany*

<sup>b</sup>*Merck KGaA, Frankfurter Str. 250, Darmstadt 64293, Germany  
[jonathan.maedler@tu-dresden.de](mailto:jonathan.maedler@tu-dresden.de)*

## Abstract

Modular plants built from process equipment assemblies (PEAs) provide a promising solution to challenges like increasing requirements of flexibility and faster time to market. To leverage the potential in processes under good manufacturing practice (GMP) conditions new methods are required to support the process validation in modular plants. In this paper, we present an integrated approach to Quality by Design and Process Performance Qualification between lab and production scale in modular plants applying the Product-Process-Resource model by Bamberg et al. (2021). Furthermore, we reformulate the Extended Flexibility Analysis by Ochoa et al. (2021) adapting the terminology to the Quality by Design context and adding the control inputs in the form of the CPPs to the optimization problem. The problem is solved by means of a grid search algorithm. The approach is studied and illustrated in a simulation-based case study with a simple reaction system.

**Keywords:** Modular Plant, Quality by Design, Design Space, Process Performance Qualification, Scale-Up

## 1. Introduction

Modular Plants, in accordance with VDI 2776 (VDI, 2020) and VDI/VDE/NAMUR 2658 (VDI/VDE/NAMUR, 2019) consist of Process Equipment Assemblies (PEAs) enabling increased flexibility and faster time to market. These properties make this concept attractive to industry sectors which often operate under the good manufacturing practice (GMP) conditions. Although aspects like PEA selection (Schindel et al., 2021) and orchestration (Klose et al., 2022) have been considered in literature before, process validation and quality by design in modular plants remain rather underdeveloped. However, Mädler et al. (2022) did present a digital twin-concept for smart process equipment assemblies supporting process validation in modular plants. This concept aims to make prior knowledge Findable, Accessible, Interoperable, and Reusable (FAIR) and relates different kinds of information to a set of methods for process validation in modular plants. In this paper we build upon these results and propose a method for Quality by Design (QbD) and Process Performance Qualification (PPQ) in the production scale based on knowledge gained in process development with lab scale PEAs. The remainder of this paper is structured as follows: section 2 provides background information about Process Validation, Quality by Design, and the Product-Process-Resource model. In section 3, a method for the Design Space Identification and Design of Experiments in the

PPQ is introduced. This method is researched in a simulation-based case study with a reactor PEA in section 4. Section 5 provides a conclusion and an outlook.

## 2. Background

### 2.1. Process Validation

According to the FDA Guidance (FDA, 2011), Process Validation consists of the (1) Process Design, the (2) Process Qualification (PQ), and the (3) Continued Process Verification stage. In the Process Design stage, the process development takes place which is frequently carried out according to the Quality by Design (QbD) paradigm. Furthermore, the scale-up of the process from lab scale to production scale must be achieved. The Process Qualification stage consists of two parts: the “Design of Facilities & Qualification of Equipment and Utilities” and the Process Performance Qualification (PPQ) (Katz and Campbell, 2012). The PPQ aims to confirm the capability of the process and plant for reproducible production of sufficient quality (Katz and Campbell, 2012). The Continued Process Verification aims to assure that the process is kept under control and to facilitate continued process optimization (Katz and Campbell, 2012).

### 2.2. Quality by Design and Process Performance Qualification

The Quality by Design paradigm proclaims the application of scientific methodologies and risk management in order to assure product quality. An important sub-concept is the Design Space (DS), which can be interpreted as a mathematical model relating the Critical Process Parameters (CPPs), the Critical Material Attributes (CMAs) of the feed streams, and sufficient values for the Critical Quality Attributes (CQAs) of the product. The Design Space is enclosed by the Knowledge Space (KS), which “contains all the information about all regions of the process that have been investigated” (Boukouvala et al., 2017) in the process development. To assure product quality under disturbances, usually a Normal Operating Range (NOR) is chosen within - but smaller than - the Design Space (cf. Boukouvala et al., 2017). The identification of the Design Space frequently requires extensive experimentation which is often very expensive at the production scale. In such cases, the identification of the Design Space is done using lab scale experiments. Afterwards, the results are scaled to production scale and tested in the PPQ.

### 2.3. Product-Process-Resource Model

The Product-Process-Resource model outlines the distribution of knowledge (Bamberg et al., 2021). The Capability Model (CM) is a product-independent description of the PEA capabilities like e.g., cooling capacity. The Transformation Model (TM) is a plant-independent description of the process. Combining these two models yields the Operation Model which can be used to characterize, identify, or predict the Design Space.

## 3. Methodology

The method presented in this paper is based on a scale-down concept employing behavioral models of PEAs in lab scale and production scale. Experiments for the identification of the Design Space are carried out utilizing a lab scale PEA which mimics the capabilities of the production scale PEA (e.g., Services with related Process Parameters). The gained knowledge is then used to derive an Operation Model for the production scale which is applied to estimate the Design Space in the production scale PEA and plan experiments for the PPQ.

### 3.1. A behavioral model for the Knowledge Space

In this paper, we consider a system of nonlinear algebraic equations (NAEs) for behavioral models of Knowledge Space

$$\mathbf{0} = \mathbf{h}(\mathbf{x}, \mathbf{u}_{\text{CMA}}, \mathbf{u}_{\text{CPP}}, \boldsymbol{\theta}_{\text{TM}}, \boldsymbol{\theta}_{\text{CM}}) \quad (1)$$

where  $\mathbf{x}$  denotes the states of the system,  $\mathbf{u}_{\text{CMA}}$  the independent inputs related to the CMAs,  $\mathbf{u}_{\text{CPP}}$  the independent inputs related to the CPPs,  $\boldsymbol{\theta}_{\text{TM}}$  the parameter values related to the Transformation Model of the material system, and  $\boldsymbol{\theta}_{\text{CM}}$  the parameters related to the Capability Model of a PEA in lab scale or in production scale.

### 3.2. Identification of a Knowledge Space in Lab Scale

In the lab scale, a Knowledge Space can be identified applying prior knowledge from Product Development and the Capability Model of the lab scale PEA. This Capability Model provides one or multiple potential behavioral models according to equation (1). Furthermore, it does provide the parameters  $\boldsymbol{\theta}_{\text{CM}}$  including their values which have been captured during a separate characterization of the PEA and/or through other previous experimental data. In addition, it provides information about  $\mathbf{u}_{\text{CPP}}$  of the PEA, which relate to the services of the PEA and their service parameters. Furthermore, the characteristics of the educts ( $\mathbf{u}_{\text{CMA}}$ ) are assumed to be controllable during lab experiments. The Knowledge Space is identified by applying statistical DoE to obtain initial parameter estimates and afterwards model-based design of experiment (MBDoe) for parameter estimation and for model discrimination. The results of these experiments are used to identify the parameter values  $\boldsymbol{\theta}_{\text{TM}}$  including uncertainty estimates, applying e.g., the maximum likelihood method and confidence intervals.

### 3.3. Identification of the Design Space and Process Performance Qualification in the Production Scale

The Transformation Model and its parameters  $\boldsymbol{\theta}_{\text{TM}}$  are extracted from the Operation Model of the lab scale PEA and combined with the Capability Model of the characterized production scale PEA. The resulting Operation Model is used to identify the Design Space in the production scale and optimally design experiments for the PPQ.

For the identification of the Design Space from a Knowledge Space described based on a behavioral model according to equation (1), the Extended Flexibility Analysis approach by Ochoa et al. (2021) is adopted. In contrast to this publication, we consider the control inputs in our approach, which relate to the Critical Process Parameters  $\mathbf{u}_{\text{CPP}}$  (or  $z$  in Ochoa et al. (2021)). In the Flexibility Analysis, the permitted ranges of the CQAs are described by the feasibility functions

$$g_j(\mathbf{x}, \mathbf{u}_{\text{CMA}}, \mathbf{u}_{\text{CPP}}, \boldsymbol{\theta}_{\text{TM}}, \boldsymbol{\theta}_{\text{CM}}) \leq 0 \quad (2)$$

These functions are combined with the behavioral model from equation (1) and constraints to the independent inputs and parameters. The following formulation the flexibility index can be calculated

$$\begin{aligned} F &= \max_{\delta_{\text{CMA}} \in \mathbb{R}^+} \delta_{\text{CMA}} \\ \text{s.t. } \chi(\boldsymbol{\theta}_{\text{CM}}) &= \max_{\mathbf{u}_{\text{CMA}} \in \mathcal{I}_{\text{CMA}}} \min_{\mathbf{u}_{\text{CPP}} \in \mathcal{I}_{\text{CPP}}} \max_{\boldsymbol{\theta}_{\text{TM}} \in \mathcal{I}_{\text{TM}}} \max_{j \in \mathcal{J}} g_j(\mathbf{x}, \mathbf{u}_{\text{CMA}}, \mathbf{u}_{\text{CPP}}, \boldsymbol{\theta}_{\text{TM}}, \boldsymbol{\theta}_{\text{CM}}) \leq 0 \\ \mathbf{h}(\mathbf{x}, \mathbf{u}_{\text{CMA}}, \mathbf{u}_{\text{CPP}}, \boldsymbol{\theta}_{\text{TM}}, \boldsymbol{\theta}_{\text{CM}}) &= \mathbf{0} \\ T_{\text{CMA}}(\mathbf{u}_{\text{CMA}}) &= \left\{ \mathbf{u}_{\text{CMA}} : \mathbf{u}_{\text{CMA}}^{\text{N}} - \delta_{\text{CMA}} \Delta \mathbf{u}_{\text{CMA}}^- \leq \mathbf{u}_{\text{CMA}} \leq \mathbf{u}_{\text{CMA}}^{\text{N}} + \delta_{\text{CMA}} \Delta \mathbf{u}_{\text{CMA}}^+ \right\} \\ T_{\text{CPP}}(\mathbf{u}_{\text{CPP}}) &= \left\{ \mathbf{u}_{\text{CPP}} : \mathbf{u}_{\text{CPP}}^{\text{LB}} \leq \mathbf{u}_{\text{CPP}} \leq \mathbf{u}_{\text{CPP}}^{\text{UB}} \right\} \\ T_{\text{TM}}(\boldsymbol{\theta}_{\text{TM}}) &= \left\{ \boldsymbol{\theta}_{\text{TM}} : \boldsymbol{\theta}_{\text{TM}}^{\text{LB}} \leq \boldsymbol{\theta}_{\text{TM}} \leq \boldsymbol{\theta}_{\text{TM}}^{\text{UB}} \right\} \end{aligned} \quad (3)$$

where  $\delta_{\text{CMA}}$  denotes the scalar flexibility index of a CMA and the superscripts LB and UB denote lower and upper bounds respectively. The optimization problem reads as follows: the maximum among the feasibility functions  $g_j$  is maximized based on the range of the uncertain transformation variables  $\boldsymbol{\theta}_{\text{TM}}$ , which have been identified within the lab

scale. Hence, the multistage optimization problem generates the worst-case scenario in terms of the CQA values. The CPPs are then used to derive minimized values for the feasibility functions in the worst-case. Afterwards, the feasibility function values are again maximized based on the allowed range of CMA values. If the feasibility function values remain below 0, the range of CMAs is feasible, and the flexibility test is passed. Finally, the range of CMAs is maximized by maximizing the flexibility index  $\delta_{\text{CMA}}$ . This multistage optimization problem is known to be very computationally expensive (cf. Ochoa and Grossmann, 2020). Thus, Ochoa and Grossmann (2020) list two reformulations for special cases like e.g., the vertex enumeration approach for cases with 1-D quasi-convex feasibility functions, and two general reformulations.

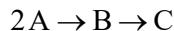
For this publication, we use a similar algorithm to the vertex enumeration approach according to Ochoa and Grossmann (2020), but instead of only considering the vertices, a grid of points is used which is spanned by the CMAs and Transformation Model parameters. This is done to get an approximation to the global solution in cases of non-convex feasibility functions. This approach is computationally very expensive, but easy to implement and therefore suits the demonstration of the approach.

The result of this Extended Flexibility Analysis is a hyperrectangle formed by the CMAs. In PPQ, the vertexes of this hyperrectangle with some safety margin should be tested in combination with the corresponding CPPs to span a Normal Operating Range. Furthermore, the Operational Model can be used to search for an optimal operation point.

## 4. Case Study

### 4.1. Example case

Let us assume that an owner/operator aims to scale-up a synthesis of a product B in an exothermic, homogeneous second order reaction from educt A, but B degrades in another reaction to C:



Two stirred tank reactor (STR) PEAs (one in lab scale and one in production scale) are available to the owner/operator. Both PEAs can be operated in a continuous STR (CSTR) operation mode, but their minimal residence times and cooling capacities may differ. The behavioral models under sufficient homogenization can be described as follows

$$0 = \frac{1}{\tau} (c_{i,\text{in}} - c_i) + \sum_{j=1}^{\text{NR}} \left[ \nu_{i,j} k_{0,j} \exp\left(-\frac{E_{A,j}}{RT}\right) \prod_{i=1}^{\text{NC}} c_i^{\gamma_{i,j}} \right] \quad (4)$$

where the subscripts  $i$  and  $j$  denote the species and the reactions with the number of components NC and the number of reactions NR respectively,  $c$  are the molar concentrations,  $\nu$  the stoichiometric coefficients,  $k_{0,j}$  the Arrhenius constants,  $E_{A,j}$  the activation energies,  $R$  the universal gas constant,  $T$  and  $\gamma$  the reaction orders regarding a species. Assuming a dosing and tempering service of ideal behavior, the residence time  $\tau$  and the temperature  $T$  can usually be classified as CPPs. The concentration of species A in the feed  $c_{A,\text{in}}$  is a CMA. The owner operator aims to achieve the following CQAs:

$$\text{Profit} = \frac{100c_B - 25c_{A,\text{in}}}{\tau} \geq 2 \quad (5)$$

$$\text{Purity} = \frac{c_B}{c_A + c_B + c_C} \geq 0.8 \quad (6)$$

### 4.2. Identification of a Knowledge Space in Lab Scale

For the sake of simplicity, we assume that the Transformation Model parameter values including uncertainties in Table 1 have been identified in lab scale applying MBDoE.

Table 1: Assumed parameter values including uncertainty measures identified in lab scale

Parameter	Unit	Nominal value	Minimum	Maximum
$k_{0,1}$	L/(mol×min)	3.2e+06	3.19e+06	3.21e+06
$k_{0,2}$	1/min	5.2e+06	5.19e+06	5.21e+06
$E_{A,1}$	J/mol	4.0e+04	3.99e+04	4.01e+04
$E_{A,2}$	J/mol	5.5e+04	5.49e+04	5.51e+04

### 4.3. Identification of the Design Space in Production Scale

Figure 1 visualizes the CQAs (Profit, Purity) in dependence of the CMA ( $c_{A,in}$ ) and the CPPs ( $\tau, T$ ) for the nominal parameter values  $\theta_{TM}$  according to Table 1. Applying the Flexibility Analysis formulation in equation (3), the relation for the flexibility index according to Figure 2 is calculated. A conservative estimate of the flexibility index can be derived to be 0.175 assuming a  $\Delta u_{CMA}$  value of 1 mol/L for  $c_{A,in}$ . This can also be seen in the Figure 2 on the right side, where it can be recognized that the Profit and Purity values barely reach the sufficient level for combination of CPPs ( $\tau = 44$  min,  $T = 300$  K).

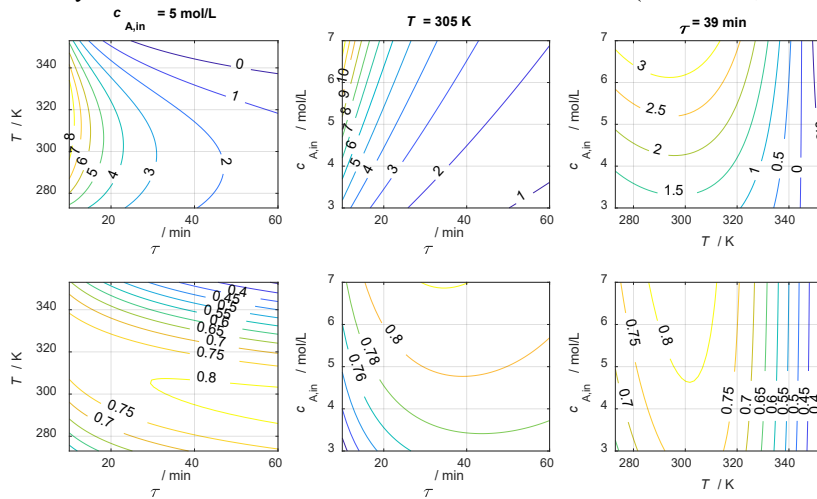


Figure 1: Profit (top row) and Purity (bottom row) in dependence of the CMA (molar concentration of species A in the feed  $c_{A,in}$ ) and the CPPs (residence time  $\tau$ , temperature  $T$ )

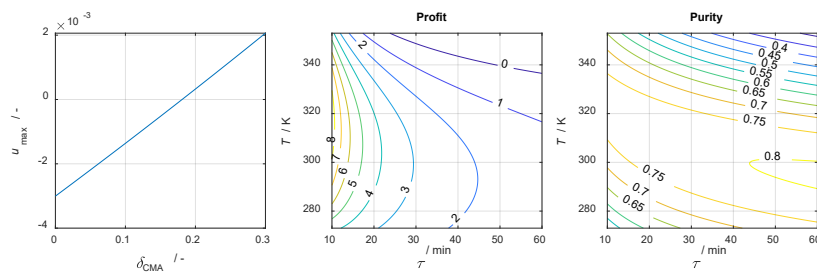


Figure 2: Maximum value of the feasibility functions  $u_{max}$  in dependence of the flexibility index  $\delta_{CMA}$  (left); and the corresponding the Profit (central) and Purity (right) at  $c_{A,in} = 4.825$  mol/L



## 5. Conclusions

In this paper, we introduced a methodology for Quality by Design and Process Performance Qualification in modular plants. This approach is based on similar PEAs in lab scale and production scale, which are integrated with Capability Models including behavioral models and Capability Model parameters. To avoid expensive experimentation in the production scale, experiments are carried out in the lab scale PEA initially based on statistical design of experiments and afterwards based on MBDoE. The identified Transformation Model including the Transformation Model parameter values with uncertainty measures are then integrated with the Capability Model of the production scale PEA. The resulting Operation Model is used to identify the Design Space and plan experiments for the PPQ. Furthermore, we suggested a reformulation of the Extended Flexibility Analysis problem proposed by Ochoa et al. (2021) by adding the CPPs (or control inputs) to the optimization problem. This allows the consideration of control actions based on measured deviation of the CMAs in the operation phase.

In future work, the approach should be applied to more complex and real-world problems. In addition, the design of experiments, experimentation and identification in lab scale should be further elaborated. Furthermore, the solution algorithm for the Extended Flexibility Analysis should be reworked adopting the general formulations by Ochoa and Grossmann (2020), or surrogate-based formulations.

## References

- Bamberg, A., Urbas, L., Bröcker, S., Bortz, M., Kockmann, N., 2021. The Digital Twin – Your Ingenious Companion for Process Engineering and Smart Production. *Chem Eng Technol* 44, 954–961. <https://doi.org/10.1002/ceat.202000562>
- Boukouvala, F., Muzzio, F.J., Ierapetritou, M.G., 2017. Methods and Tools for Design Space Identification in Pharmaceutical Development, in: *Comprehensive Quality by Design for Pharmaceutical Product Development and Manufacture*. pp. 95–123. <https://doi.org/10.1002/9781119356189.ch6>
- Katz, P., Campbell, C., 2012. FDA 2011 process validation guidance: Process validation revisited. *Journal of GXP Compliance* 16, 18.
- Klose, A., Lorenz, J., Bittorf, L., Stark, K., Hoernicke, M., Stutz, A., Weinhold, H., Krink, N., Welscher, W., Eckert, M., Unland, S., Menschner, A., Da Silva Santos, P., Kockmann, N., Urbas, L., 2022. Orchestration of modular plants: Procedure and application for orchestration engineering. *atp* 63, 68–77. <https://doi.org/10.17560/atp.v63i9.2599>
- Mädler, J., Rahm, J., Viedt, I., Urbas, L., 2022. A digital twin-concept for smart process equipment assemblies supporting process validation in modular plants, in: *Montastruc, L., Negny, S. (Eds.), Computer Aided Chemical Engineering*. Elsevier, pp. 1435–1440. <https://doi.org/10.1016/B978-0-323-95879-0.50240-X>
- Ochoa, M.P., García-Muñoz, S., Stamatis, S., Grossmann, I.E., 2021. Novel flexibility index formulations for the selection of the operating range within a design space. *Computers & Chemical Engineering* 149, 107284. <https://doi.org/10.1016/j.compchemeng.2021.107284>
- Ochoa, M.P., Grossmann, I.E., 2020. Novel MINLP formulations for flexibility analysis for measured and unmeasured uncertain parameters. *Computers & Chemical Engineering* 135, 106727. <https://doi.org/10.1016/j.compchemeng.2020.106727>
- Schindel, Polyakova, Harding, Weinhold, Stenger, Grünewald, Bramsiepe, 2021. General approach for technology and Process Equipment Assembly (PEA) selection in process design. *CEP:PI* 159, 108223. <https://doi.org/10.1016/j.ccp.2020.108223>
- VDI, 2020. *Modulare Anlagen - Grundlagen und Planung modularer Anlagen - Blatt 1 (VDI 2776:2020-11)*.
- VDI/VDE/NAMUR, 2019. *Automatisierungstechnisches Engineering modularer Anlagen in der Prozessindustrie - Allgemeines Konzept und Schnittstellen - Blatt 1 (VDI/VDE/NAMUR 2658-1:2019-10)*.

# Digital Twins for Scale-Up in Modular Plants: Requirements, Concept, and Roadmap

A. Koch<sup>a\*</sup>, J. Mädler<sup>a</sup>, A. Bamberg<sup>b</sup>, L. Urbas<sup>a</sup>

<sup>a</sup>*Chair of Process Control Systems & Process Systems Engineering Group,  
TU Dresden, Dresden 01062, Germany;*

<sup>b</sup>*Merck KGaA, Frankfurter Str. 250, Darmstadt 64293, Germany,*

## Abstract

The application of Digital Twins for scale-up in Modular Plants is envisioned to significantly reduce the necessary time-to-process for new-to-market products and address scale-up challenges. To identify application specific requirements of a Digital Twin for scale-up in Modular Plants, expert interviews were conducted. The spectrum of identified requirements as well as plant life cycle aspects are incorporated into a workflow concept which illustrates the partial models of Digital Twins, as well as the interfaces for the exchange of information. Also presented is a roadmap illustrating the cluster-specific development pathways for Digital Twins for scale-up in Modular Plants.

**Keywords:** Digital Twin, Modular Plants, process development, scale-up

## 1. Introduction

In order to maintain a competitive edge within dynamic markets, such as pharmaceuticals and specialty chemicals, Modular Plants (MP) must address the following requirements: (1) reduce the time-to-market, (2) provide increased flexibility, and (3) have increased efficiency (Bernshausen et al., 2019). An additional hurdle is the scale-up from laboratory to production. The integration of Digital Twins (DT) into process development workflows could potentially significantly speed up the equipment selection of a manufacturing process. For this purpose, the combination of standardized, predesigned equipment such as Process Equipment Assemblies (PEA) with DTs is poised to be an integral tool to facilitate rapid time-to-process (Mädler et al., 2022a; Mädler et al., 2022b).

## 2. Background

### 2.1. Digital Twins

According to Rosen et al. (2019), a Digital Twin is a semantically linked collection of all digital artefacts, including design and engineering data as well as operational data and behavior descriptions. Moreover, a Digital Twin is intended for a specific purpose (Rosen et al, 2019). Bamberg et al. (2021) view the Digital Twin as consisting of different partial models which can be structured in the **Product-Process-Resource-Model** (Bamberg et al., 2021). The model of Bamberg et al. (2021) was applied to Modular Plants by Mädler et al. (2022b) to illustrate how the knowledge distribution in a Digital Twin is divided between the PEA manufacturers and the owners/operators (O/O). The **capability model** is a product independent description of a PEA containing structural, operational, and behavioral information about the PEA, independent of a concrete process. The

**transformation model** is a plant independent and product specific description of the process used to derive information for a concrete PEA such as residence times. Combining the capability model and the transformation model yields the **operation model**, which is used for recipes and optimization of process parameters (Mädler et al., 2022b).

### 2.2. Modular Plants

The structure of Modular Plants is outlined in the VDI 2776 Part 1. Modular Plants (MPs) consist of functional equipment assemblies (FEAs), which are combined into process equipment assemblies (PEAs) (VDI, 2020). PEAs can be further characterized as intelligent modules which are computerized and (static) self-describing (Mädler et al., 2022a). Automation in Modular Plants is described in VDI/VDE/NAMUR 2658 (VDI/VDE/NAMUR, 2019). The operation of multiple PEAs in a modular plant is carried out via the process orchestration layer (POL). Here, the Module Type Package (MTP) serves to integrate PEA into the POL, yielding a standardized and non-proprietary means to describe the automation in a PEA (Klose et al., 2019).

### 2.3. Process Design in Modular Plants

Applying modular-based plant design methodology results in decreased planning and construction times by incorporating predefined modules and devoting time to the selection process instead of to individual design of each piece of equipment (Bramsiepe et al., 2012). Methodologies for matching PEAs in process design can be found in literature. One example is Schindel et al. (2021), who utilize Matching Matrices to match the technical attributes of a PEA with the process requirements (Schindel et al., 2021).

## 3. Methodology: Expert Interviews

In order to delineate the task-specific requirements of a Digital Twin for Modular Plants for process development and scale-up, semi structured interview questionnaires based on a set of working hypotheses were conducted with experts from industry and academia with experience ranging from 3 to 45 years (cf. Koch, 2022). To elicit requirements, the interview transcripts were then analyzed using the coding agenda methodology of Mayring (Mayring, 2021). The coded requirements were grouped into five clusters, which were adapted from Oppelt (2016). These clusters, shown in Figure 1, are defined as follows:

- (1) **Acceptance:** demonstrate DT and gain acceptance by Owners/Operators (O/O)
- (2) **Efficiency:** address bottlenecks in process development workflows with DTs
- (3) **Integration:** exchange of information using interfaces
- (4) **Reusability:** reusability of models and information for new products
- (5) **Usability:** accessibility and ease of use of DT by O/O

To illustrate how these requirements are applied to the DT workflow concept and roadmap, four selected requirements from the integration cluster are discussed: interfaces, semantically linked models, CAE System, and DEXPI. These requirements are well suited to describe the DT workflow in Section 5. A detailed explanation of all requirements can be found in Koch (2022). The requirement of a CAE System serves as a tool to store the digital artefacts including design and engineering data (cf. Rosen et al. (2019)), which are part of the essential information contained in a DT. The requirement of interfaces refers to the interfaces, which are needed as part of a DT in order to transfer information between engineering tools, such as a simulation tool and a CAE system. Interface also refers to interfaces between a simulation tool and operating data contained in a POL. One specific identified bottleneck in the process development workflow is the difficulty in transferring information between engineering tools.

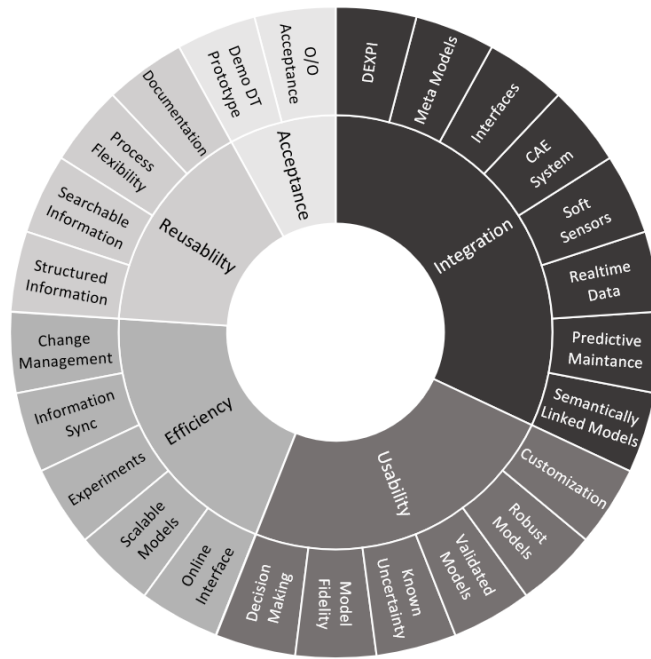


Figure 1: Clustered, expert identified requirements of Digital Twin for scale-up (Koch, 2022).

The integration of interfaces to support the obligatory semantic linkage between CAE tools could address this specific bottleneck, resulting in a DT concept that is in line with Rosen et al. (2019). The last requirement explained here, is DEXPI (Wiedau et al., 2019), which refers to a specific, openly available information model which could be used to transfer data between engineering tools and should be integrated into a DT concept.

## 4. RoadMap: Vision for Digital Twins in Scale-Up

### 4.1. Background and Prerequisites

In order to create a DT concept, the requirement catalog (see Figure 1) was workshopped and refined with software experts from industry. The final requirements catalog was used to create a Roadmap, shown in Figure 2, which outlines the necessary steps to enable this Digital Twin Vision for scale-up in Modular Plants. As a starting point for this specific Digital Twin, a minimum baseline was defined with two prerequisites: (1) MTP is integrated into both the lab and production PEAs and (2) the lab and production PEA setup makes both scale-up and scale-down possible.

### 4.2. Explanation

The roadmap depicts five different pathway lines with a series of different symbols. Each pathway line aligns with a requirement cluster explained in Section 3, and each symbol refers to a different requirement identified from the expert interviews shown in Figure 1. Additionally, each requirement was allocated to a specific tool or interface, which are categorized as follows: simulation tool (circle), plant engineering tool (triangle), interface (star), both simulation tool and plant engineering tool (diamond), and undefined/other (square). Moreover, each requirement is allocated to a specific time frame represented by the background color: short term (dark grey), medium term (medium grey), and long term (light grey).

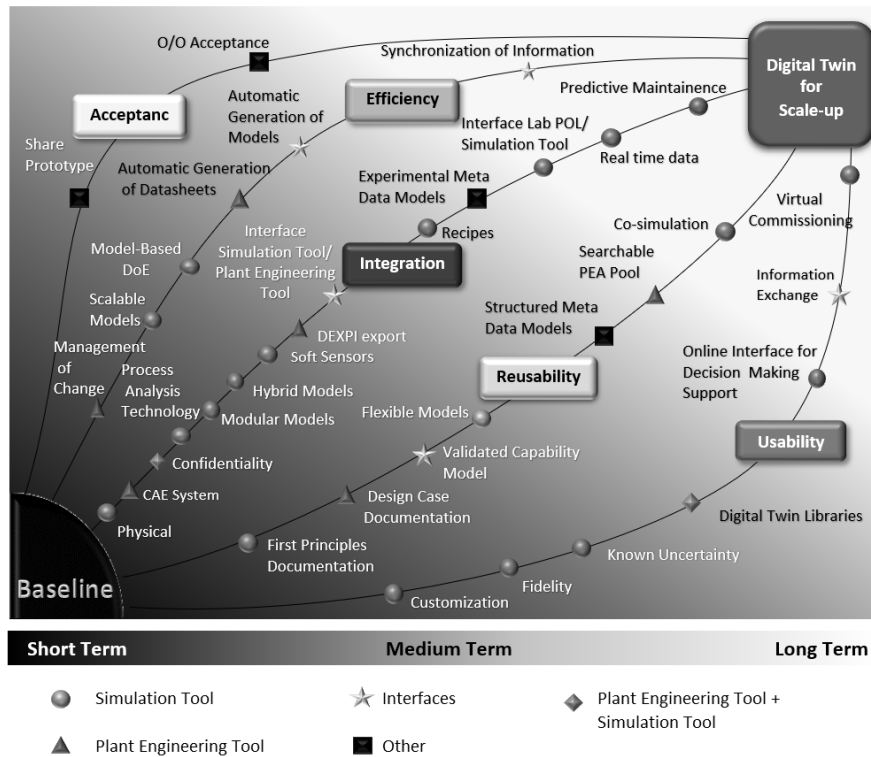


Figure 2: Roadmap for a Digital Twin for Scale-up using Modular Plants. Adapted from Koch (2022).

The requirements are ordered along each cluster pathway and allocated to a time frame based on both the importance to experts as well as the feasibility of the required technology needed to implement a requirement. Addressing each of these requirements yields the vision of a Digital Twin for scale-up potentially facilitating faster time-to-process.

## 5. Digital Twin Concept

### 5.1. Workflow Concept

Short- and medium-term requirements illustrated in the roadmap in Figure 2 are addressed with a concept workflow for a Digital Twin for scale-up, which is shown in Figure 3. This presented DT concept workflow assumes that information such as design and engineering data is contained in a Plant Engineering Tool, behavior descriptions in a simulation tool, operational data from the production PEA in the Production POL, and experimental data in the Lab POL. The partial models within the simulation tool are aligned with Bamberg et al. (2021). Here, the operation model for the Lab PEA consists of a capability model, representing the characterized capability of the Lab PEA resource and a transformation model for a specific product. This transformation model is shared with the production PEA and also contains relevant information from product development. This product specific transformation model is combined with the capability model from the production PEA into an operation model. Operational data from the Production PEA is used to improve the Operation Model for the Production PEA, illustrating the semantic linkage between models. The result of this workflow is an in-spec product for the supply chain.

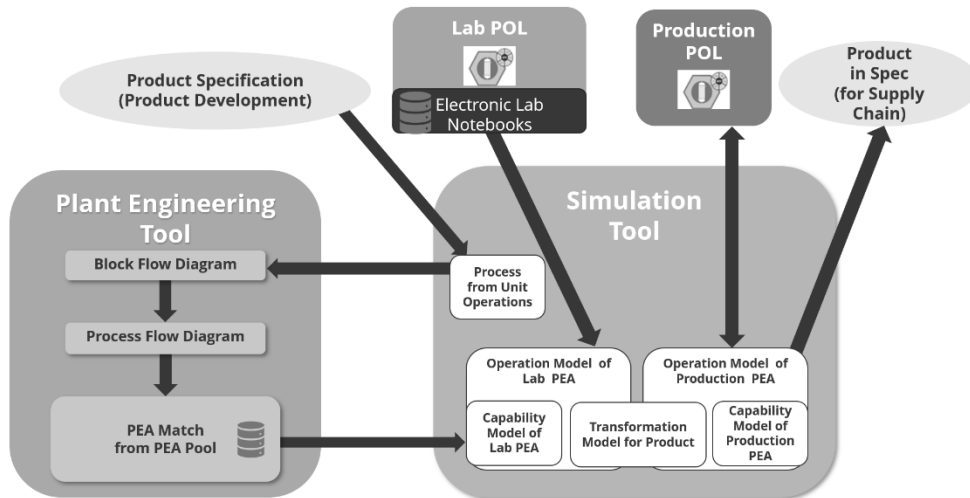


Figure 3: Workflow for a Digital Twin for Scale-up. Adapted from Koch (2022).

### 5.2. Requirements of an Interface to Support PEA Matching

To demonstrate how the expert requirements were integrated into the workflow shown in Figure 2, the interface between the simulation tool and plant engineering tool is explained. This Digital Twin concept centers around having a repository of resource models, or PEA Pool, stored in a searchable databank of a plant engineering tool containing design and engineering data. This addresses the requirement of structured and searchable information as well as CAE system. Here, the PEA matching methodology of Schindel et al. (2021) is referenced to identify the requirements of an interface that supports matching process models (from a simulation with unit operations) to resource models in a PEA Pool. This workflow begins with a resource independent simulation of a process for a new product based on a simulation created from unit operations (reaction, distillation, etc.). Schindel et al. (2021) assume that the information about a process provided by a simulation for matching PEAs to processes is low. Moreover, technical criteria (e.g., temperature, pressure, and viscosity) are identified as the most important information for successful matching. These technical criteria are found to be unit operation and material flow specific; thus, important technical criteria for the selection of a heat exchanger are not the same as the criteria for the selection of a distillation column. Hence, for a match to be successful it is essential that the simulation tool propagates these necessary technical criteria through an interface to a plant engineering tool. The exchange of this information across this interface could be supported by an information model such as DEXPI (Wiedau et al., 2019). A detailed explanation of this workflow concept, interfaces, and partial models as well as how the requirements were used to create this concept can be found in Koch (2022).

## 6. Discussion and Conclusion

In this paper, expert interview derived requirements of Digital Twins for scale-up in Modular Plants are presented and discussed. The expert interviews yielded purpose specific requirements that must be integrated into a Digital Twin for scale-up, thereby supporting process development and potentially facilitating faster time-to-process. Considering these requirements and challenges along with plant life cycle aspects, a concept DT workflow was presented. This DT concept was used to illustrate the linkage

of the capability models and transformation models of a Digital Twin as well as the integration of information contained in the plant engineering tool, Lab POL, and Production POLs. This semantic linkage enables efficient engineering workflows, resulting in faster time-to-process and ultimately model reuse when PEAs are integrated into new processes. These interfaces support PEA matching, which is critical in order to support faster time-to-process. Moreover, the presented roadmap illustrates clear pathways to enable this vision of a Digital Twin and the necessary steps which must be taken to achieve this goal. Using this DT concept and roadmap, the next step is to develop prototypical implementations demonstrating the added benefit of a Digital Twin, resulting in owner/operator acceptance of a Digital Twin and integration into existing engineering workflows within process development. Moreover, prototypes should address the uncertainty and challenges in scale-up, specifically focusing on the linkage between scalable simulation models and information models. Here, alliance with existing standards is necessary to support the exchange of essential and potentially proprietary information between the partial models of the Digital Twin.

## References

- A. Bamberg, L. Urbas, S. Bröcker, M. Bortz, and N. Kockmann, 2021, The Digital Twin—Your Ingenious Companion for Process Engineering and Smart Production, *Chem. Eng. Technol.*, 44, 6, 954–961.
- J. Bernshausen, A. Haller, H. Bloch, M. Hoernicke, S. Hensel, A. Menschner, M. Stutz, M. Maurmaier, T. Holm, C. Schäfer, L. Urbas, U. Christmann, C. Fleischer-Trebes, & F. Stenger, 2019. Plug & Produce auf dem Sprung in den Markt: Neuerungen in Spezifikation und Implementierung des MTP. *Atp Magazin*, 61(1–2), 56–69.
- B. Bramsiepe, S. Sievers, T. Seifert, G. D. Stefanidis, D.G. Vlachos, H. Schnitzer, B. Muster, C. Brunner, J.P.M. Sanders, M.E. Bruins, G. Schembecker, G., 2012. Low-cost small scale processing technologies for production applications in various environments—Mass produced factories. *Chem. Eng. Process*, 51, 32–52.
- A. Klose, S. Merkelbach, A. Menschner, S. Hensel, S. Heinze, L. Urbas, C. Schäfer, S. Szmais, M. Eckert, T. Ruede, T. Scherwies, P. Santos, F. Stenger, T. Holm, W. Welscher, N. Krink, T. Schenk, A. Stutz, M. Maurmaier, & F. Apitz, 2019. Anforderungen an die Orchestrierung für modulare Prozesseinheiten, *Chem. Eng. Technol.*, 42, 11, 2282–2291.
- A. Koch, 2022. Digital Twins for Lab2Plant in Modular Process Plants. Diploma Thesis. Unpublished
- J. Mädler, J. Rahm, I. Viedt, and L. Urbas, 2022a, A digital twin-concept for smart process equipment assemblies supporting process validation in Modular Plants, In *Computer Aided Chemical Engineering*, 51, 1435–1440, Elsevier.
- J. Mädler, I. Viedt, J. Lorenz, and L. Urbas, 2022b, Requirements to a digital twin-centered concept for smart manufacturing in Modular Plants considering distributed knowledge. In Y. Yamashita & M. Kano (Hrsg.), *Comput. Aided Chem. Eng.* 49, 1507–1512.
- P. Mayring, P. (2021). *Qualitative content analysis: A step-by-step guide*. SAGE Publications.
- M. Oppelt, 2016. Towards an integrated use of simulation within the life-cycle of a process plant.
- R. Rosen, J. Fischer, & S. Boschert, 2019, Next generation digital twin: An ecosystem for mechatronic systems?, *IFAC-Paper*, 52, 15, 265–270.
- A.-L. Schindel, M. Polyakova, D. Harding, H. Weinhold, F. Stenger, M. Grünewald, and C. Bramsiepe, 2021, General approach for technology and Process Equipment Assembly (PEA) selection in process design, *CEP:PI*, 159, 108223.
- VDI/VDE/NAMUR, 2019, VDI/VDE/NAMUR 2658-Automation engineering of modular systems in the process industry-General concept and interfaces – Part 1, Beuth Verlag GmbH.
- VDI 2776-1:2020-11, 2020. *Modular Plants - Fundamentals and planning Modular Plants*.
- M. Wiedau, L. von Wedel, H. Temmen, R. Welke and N. Papkonstantinuous, 2019, ENPRO Data Integration: Extending DEXPI Towards the Asset Lifecycle, *Chem. Eng. Tech.*, 91, 249–255.

# Modeling and Simulation of a PSA Process for CO<sub>2</sub>/Syngas Separation

Magno Fonseca Santos,<sup>a</sup> Stefano Ferrari Interlenghi,<sup>b</sup> Antonio Esio Bresciani,<sup>a</sup>  
Newton Libanio Ferreira,<sup>a</sup> Gabriel Soares Bassani,<sup>c</sup> Rita Maria de Brito Alves<sup>a</sup>

<sup>a</sup>*Universidade de São Paulo, Av. Prof. Luciano Gualberto, n. 380, trav. 3, São Paulo 05508-010, Brazil*

<sup>b</sup>*Unidade SENAI CETIQT Parque, Rua Fernando de Souza, n. 120, Rio de Janeiro 21941-857, Brazil*

<sup>c</sup>*Repsol Sinopec Brazil, Praia de Botafogo, 300, Rio de Janeiro 22250-040, Brazil*

## Abstract

This work aims to model and simulate a pressure swing adsorption (PSA) unit for CO<sub>2</sub>/syngas separation using rigorous multi-bed and single-bed approaches. The PSA unit was designed to purify syngas from reverse water-gas shift (RWGS) reaction to application in the Fischer-Tropsch synthesis. Process simulations were performed in Aspen Adsorption® v10 software, using the Peng-Robinson equation of state. The association of 7 adsorption beds was considered, allowing the continuous production of purified streams, in a concept of cyclic steady-state operation. The single-bed approach reduces the system complexity and improves the processing power/time without losing precision. A sensitivity analysis showed that integrated RWGS and PSA processes can produce syngas with purity above 94.65% with an H<sub>2</sub>/CO molar ratio of 1.50. The total recovery was around 99.98% and 99.57% for H<sub>2</sub> and CO, respectively, for an adsorption pressure of 20.00 bar, and a desorption pressure of 1.50 bar.

**Keywords:** Pressure swing adsorption, Syngas purification, Adsorption of CO<sub>2</sub>, Process simulation.

## 1. Introduction

Syngas is a mixture composed mainly of carbon monoxide (CO) and hydrogen (H<sub>2</sub>), whose importance as a building block for oxygenates production and Fischer-Tropsch (FT) synthesis to hydrocarbons (Theampetch et al., 2021) has increased. Syngas composition depends on its production technology, which includes methane reforming processes and reverse water-gas shift (RWGS) reaction (Dasireddy et al., 2021), which converts CO<sub>2</sub> and H<sub>2</sub> into CO. For the syngas to achieve the specification of the desired derivative, for instance, FT synthesis, major impurities such as CO<sub>2</sub> must be removed from the gas stream. Methods such as membranes, absorption, and adsorption have been used to separate CO<sub>2</sub> from syngas. Among these, adsorption is a well-established purification technology due to its efficiency, operational flexibility, low costs, and absence of by-products when combined with an effective regeneration process (Wibowo et al., 2021). In parallel adsorbent beds, alternated adsorption/desorption steps are used to maintain continuous feeding and high productivity. Adsorption performances depend on the operation parameters, such as gas type, mixture properties, temperature, and pressure (Kacem et al., 2015). The CO<sub>2</sub>/syngas separation can potentially be carried out,



at moderate temperatures and pressures, using adsorbent materials, such as zeolites and activated carbon (Wilson *et al.*, 2020), which affects the unit process design. These processes may be categorized according to the desorption technique employed, such as temperature swing adsorption (TSA), vacuum swing adsorption (VSA), pressure swing adsorption (PSA), and combined options. Only a few applications of TSA to CO<sub>2</sub> capture have been found and large-scale VSA is highly electricity-demanding, which seriously impacts the electrical system. PSA is widely used in gas separation applications (Golmakani *et al.*, 2017). This work aims to model and simulate a PSA unit for CO<sub>2</sub>/syngas separation using rigorous multi-bed and single-bed approaches. The PSA unit was designed to purify syngas from the RWGS reaction to application in the FT synthesis. The novelty of the proposal lies in the application of PSA technology with a new arrangement of adsorption beds and operational stages to separate large amounts of impurities (~64% wt of CO<sub>2</sub>).

## 2. Methods and Modeling

PSA is a separation process whereby the gaseous components with weak physical interaction with the considered fixed phase, a solid material of high porosity, are separated from other substances. Highly interactive substances are adsorbed at higher pressures, and, after an adequate interval, the adsorption bed is intermittently depressurized to recover the product and regenerate the fixed phase at lower pressures (Durán *et al.*, 2022). In this proposal, CO<sub>2</sub> must be adsorbed, allowing syngas to pass through the adsorbent bed. The desorbed gas mixture, rich in CO<sub>2</sub>, is recycled to the RWGS reactor previously evaluated by Santos *et al.* (2022), considering a purge of 1% (Figure 1).

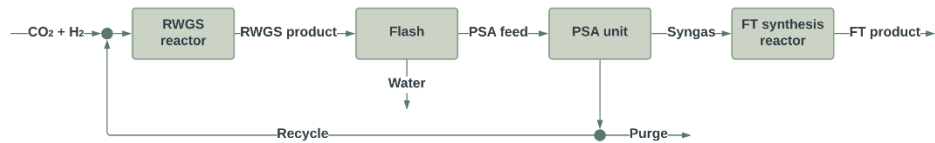


Figure 1. Process Flow Diagram of the general topology of the process.

### 2.1. Model implementation and parameters

Aspen Adsorption® v10 software was used to simulate the PSA unit. The Peng-Robinson equation of state was used for simulation. The extended Langmuir adsorption isotherms (Equation 1) and adsorbent parameters by Liu *et al.* (2020), and linear driving force adsorption kinetics (Equation 2) by Xiao *et al.* (2016) were used to predict the adsorption equilibrium behavior on a packed bed of activated carbon as the fixed phase. The same H<sub>2</sub> adsorption isotherms were used for H<sub>2</sub>O. This assumption is justified by the fact that H<sub>2</sub>O does not significantly affect the CO<sub>2</sub>/syngas separation, similar to H<sub>2</sub>, due to the hydrophobic character of activated carbon (Kumita *et al.*, 1994).

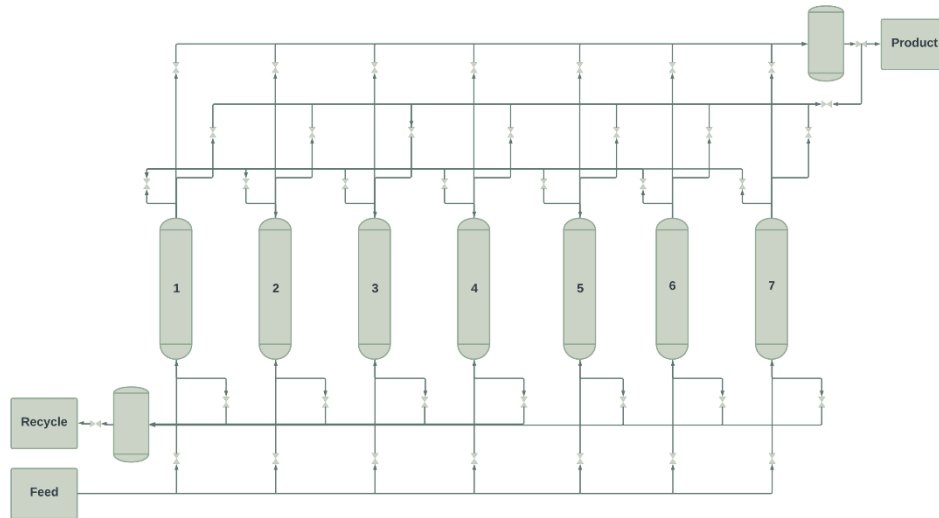
$$n_i^* = \frac{IP_{1,i} e^{\frac{IP_{2,i}}{T}} p_i}{1 + (\sum_{i=1}^N IP_{3,i} e^{\frac{IP_{4,i}}{T}} p_i)}, i = 1, \dots, N \quad (1)$$

$$\frac{\partial n_i}{\partial t} = k_i(n_i^* - n_i), i = 1, \dots, N \quad (2)$$

where  $n_i^*$  is the equilibrium adsorption amount of species  $i$  (mol/kg);  $IP_{k,i}$  are parameter  $k$  ( $k=1,4$ ) of the Extended Langmuir model of species  $i$ , (mol/kg bar), (K), (1/bar), (K);  $n_i$  is the dynamic adsorption amount coefficient of species  $i$  (mol/kg),  $k_i$  is the mass transfer coefficient of species  $i$  (1/s).

2.2. PSA unit simulation

The association of 7 adsorption beds was considered for the PSA unit, allowing the continuous production of purified streams, in a concept of cyclic steady-state operation. The operational phases considered were: 1- Adsorption (A): Retention of higher adsorbent affinity species; 2- Equalizations (depressurizations) (EnD): Sequential bed depressurization to recovery product; 3- Proving purge (PP): Product utilization to clean adsorption bed interstitial spaces; 4- Blowdown (B): Depressurization for regeneration; 5- Purge (P): Bed interstitial clean-up for regeneration; 6- Equalizations (pressurizations) (EnP): Product recovery through sequential bed pressurization; and 7- Final pressurization (FP) with the clean product. Figure 2 brings the time-dependent diagram, showing the integration of each adsorption bed during a complete operating cycle.



Bed	Steps																													
1	A											E1D	E2D	E3D	PP	B	P	E3P	E2P	E1P	FP									
2	E1P	FP	A										E1D	E2D	E3D	PP	B	P	E3P	E2P	E1P	FP								
3	E3P	E2P	E1P	FP	A										E1D	E2D	E3D	PP	B	P	E3P	E2P	E1P	FP						
4	B	P	E3P	E2P	E1P	FP	A										E1D	E2D	E3D	PP	B	P	E3P	E2P	E1P	FP				
5	E3D	PP	B	P	E3P	E2P	E1P	FP	A										E1D	E2D	E3D	PP	B	P	E3P	E2P	E1P	FP		
6	E1D	E2D	E3D	PP	B	P	E3P	E2P	E1P	FP	A										E1D	E2D	E3D	PP	B	P	E3P	E2P	E1P	FP
7	A		E1D	E2D	E3D	PP	B	P	E3P	E2P	E1P	FP	A																	

Figure 2. Integration of adsorption beds during a complete operating cycle.

To simulate the multi-bed column adsorption, rigorous multi-bed and single-bed approaches were used. In the first approach, all existing beds were added to the simulation and connected via material streams by interaction valves and holdup tanks. In the single-bed approach, only one bed was modeled, and an interaction block was used. This block records the material profile from the rigorously modeled column and reruns it at user-defined points to simulate returning material. To make this approach possible, each adsorbent bed must be identical and undergoes the same steps in a given cycle.

Since the adsorption process is exothermic, the enthalpy variation is small, and the simulations were considered isothermal. The temperature of the adsorbent bed was kept constant at 298.15 K during the complete operating cycle. The pressure of the adsorption step was set at 20.00 bar for integration with the FT synthesis. The pressure of the purge step was set at 1.50 bar to avoid vacuum operation. Identical dead volumes before and after the adsorbent bed were considered. Table 1 summarizes the pressure and duration of the steps of an operating cycle.

Table 1. Conditions of the PSA process operation cycle.

Step	Pressure (bar)	Time (s)
A	20.00	360
E1D	20.00 → 15.42	60
E2D	15.42 → 8.57	120
E3D	8.57 → 5.58	60
PP	5.58 → 3.82	120
B	3.82 → 1.50	60
P	1.50	120
E3P	1.50 → 3.60	60
E2P	3.60 → 7.71	120
E1P	7.71 → 10.26	60
FP	10.26 → 20.00	120

A sensitivity analysis was performed to find the appropriate residence time for the syngas production respecting the length/diameter ratio of 4 for the adsorbent bed. To assess the performance of the operating cycle when the system reaches the steady state, the process performance indicators were estimated: syngas (CO + H<sub>2</sub>) purity (Equation 3), H<sub>2</sub>/CO molar ratio in the product, and CO, H<sub>2</sub>, CO<sub>2</sub>, CH<sub>4</sub> and H<sub>2</sub>O recovery per pass (Equation 4) and global recovery (Equation 5).

$$Pur_s = \frac{M_{p,s}}{M_p} \quad (3)$$

$$Rec_{i,pp} = \frac{M_{p,i,pp}}{M_{p,pp}} \quad (4)$$

$$Rec_{i,o} = \frac{M_{p,i}}{M_p} \quad (5)$$

where  $Pur_s$  is the syngas purity (%),  $M_{p,s}$  is the syngas in the product (kmol),  $M_p$  is the total product (kmol),  $Rec_{i,pp}$  is the recovery per pass of species  $i$  (%),  $M_{p,i,pp}$  is species  $i$  in the product per pass (kmol),  $M_{p,pp}$  is the total product per pass (kmol),  $Rec_o$  is the overall recovery (%),  $M_{p,i}$  is species  $i$  in the product (kmol).

### 3. Results and Discussion

#### 3.1. Rigorous multi-bed approach

The rigorous multi-bed approach allows the accurate calculation of column interaction and the more precise inclusion of advanced control systems. However, systems with over two beds are extremely complex. The system initialization is extremely detailed, and convergence is slow and hard. Each element of each bed must be independently included in the cycle manager, generating huge data tables with exact specifications. In this approach, when two different pressure streams are separated by a block, the simulator

automatically and instantly equalizes the pressures, preventing the step-by-step pressure reduction. For instance, if the top of two columns, at 1.50 bar and 20.00 bar, respectively, are separated by a closed valve, the calculation proceeds as normal. The moment the valve is slightly opened, both columns tops instantly go to around 15.00 bar. This creates unrealistic pressure profiles with huge increases in pressure. Thus, the modeling of each interaction is not recommended for multi-bed/column adsorption, since it is very time-consuming, a hard-to-converge system, and a power demanding processing.

### 3.2. Single-bed approach

The single-bed approach reduces the system complexity and improves the processing power/time without losing precision. Table 2 shows the effect of the residence time on the outcome of the PSA cycles with the single-bed approach. As expected, an opposite trend is observed in the pattern of syngas purity and CO and H<sub>2</sub> recovery per pass. Therefore, it is usually necessary to reach a compromise when selecting the optimum cyclic configuration. CO<sub>2</sub>/syngas separation also depends on the requirements of the intended application.

Table 2. Simulation results in terms of process performance indicators.

Residence time (s)	27	37	47	57
Adsorbent bed diameter (m)	1.83	2.03	2.20	2.34
Adsorbent bed length (m)	7.31	8.12	8.79	9.38
Syngas purity (%)	74.33	84.66	95.50	97.86
H <sub>2</sub> /CO molar ratio	1.51	1.67	2.14	8.91
CO recovery (%)	81.45	73.12	56.74	13.48
H <sub>2</sub> recovery (%)	99.02	98.61	97.97	96.65
CO <sub>2</sub> recovery (%)	40.14	20.13	5.96	1.66
CH <sub>4</sub> recovery (%)	60.10	43.93	18.47	2.10
H <sub>2</sub> O recovery (%)	99.02	98.61	97.97	96.65

The results of the sensitivity analysis showed that in the residence time of 47 s the purified syngas presented the molar composition of 0.03% CH<sub>4</sub>, 30.04% CO<sub>2</sub>, 5.53% CO, 64.39% H<sub>2</sub>, and 0.32% H<sub>2</sub>O. This composition is similar to that presented by Graciano et al. (2018) for FT synthesis in a slurry bubble column reactor. After the recycle convergence of the CO<sub>2</sub>-rich stream to the RWGS reactor, the proposed system to treat 50 tons/h of CO<sub>2</sub> was able to produce syngas with 94.65% purity and an H<sub>2</sub>/CO molar ratio of 1.50. The overall recovery was 99.98% and 99.57% for H<sub>2</sub> and CO, respectively. Table 3 shows the main streams of the proposed system shown in Figure 1. Each adsorbent bed in the PSA unit is 13.08 and 3.27 m in length and diameter, respectively.

Table 3. Component mass flow rate (ton/h).

Parameter	CO <sub>2</sub> + H <sub>2</sub>	RWGS product	PSA feed	Syngas	Recycle
Total	54.99	163.05	145.20	36.04	108.06
CO <sub>2</sub>	50.00	93.04	93.03	5.45	86.70
H <sub>2</sub>	4.99	3.00	3.00	2.95	0.05
CO	-	48.46	48.46	27.39	20.86
H <sub>2</sub> O	-	18.00	0.15	0.15	0.00
CH <sub>4</sub>	-	0.55	0.55	0.10	0.45

## 4. Conclusion

A PSA unit for CO<sub>2</sub>/syngas separation integrated into the RWGS process was studied. The single-bed approach was the best option for a multi-bed/column system, capable of

reducing the system complexity and improving processing power/time without losing precision. Considering syngas purity and CO recovery, a residence time of 47s was in good agreement with the literature data. After integrating RWGS and PSA units, the proposed system can properly treat and separate the RWGS reaction products for a CO<sub>2</sub> load of 50 tons/h to produce 94.65% pure syngas. The total recovery was 99.98% and 99.57% for H<sub>2</sub> and CO, respectively, at adsorption and desorption pressures of 20 bar and 1.5 bar. Thus, PSA units using activated carbon as adsorbent material for CO<sub>2</sub>/syngas separation is a technically feasible process.

## 5. Acknowledgment

The authors gratefully acknowledge Repsol Sinopec Brazil for its financial and technical support and ANP (Brazilian National Oil, Natural Gas, and Biofuels Agency) for the strategic importance of its support through the R&D levy regulation. The authors also thank the National Council for Scientific and Technological Development - CNPq (#314598/2021-9). This study was financed in part by the Coordination for the Improvement of Higher Level Personnel - CAPES - Finance Code 001.

## References

- A. Golmakani, S. Fatemi, and J. Tamnanloo, 2017, Investigating PSA, VSA, and TSA methods in SMR unit of refineries for hydrogen production with fuel cell specification, *Separation and Purification Technology*, 176, 73–91.
- A. Theampetch, C. Prapainainar, S. Tungkamani, P. Narataruksa, T. Sornchamni, L. Árnadóttir, G. N. Jovanovic, 2021, Detailed microkinetic modelling of syngas to hydrocarbons via Fischer Tropsch synthesis over cobalt catalyst, *International Journal of Hydrogen Energy*, 46, 48, 24721–24741.
- B. Liu, X. Yu, W. Shi, Y. Shen, D. Zhang, and Z. Tang, 2020, Two-stage VSA/PSA for capturing carbon dioxide (CO<sub>2</sub>) and producing hydrogen (H<sub>2</sub>) from steam-methane reforming gas, *International Journal of Hydrogen Energy*, 45, 46, 24870–24882.
- H. Wibowo, H. Susanto, N. Grisdanurak, D. Hantoko, K. Yoshikawa, H. Qun, M. Yan, 2021, Recent developments of deep eutectic solvent as absorbent for CO<sub>2</sub> removal from syngas produced from gasification: Current status, challenges, and further research, *Journal of Environmental Chemical Engineering*, 9, 4, 105439.
- I. Durán, F. Rubiera, and C. Pevida, 2022, Modeling a biogas upgrading PSA unit with a sustainable activated carbon derived from pine sawdust. Sensitivity analysis on the adsorption of CO<sub>2</sub> and CH<sub>4</sub> mixtures, *Chemical Engineering Journal*, 428, 132564.
- J. E. A. Graciano, B. Chachuat, and R. M. B. Alves, 2018, Conversion of CO<sub>2</sub>-Rich Natural Gas to Liquid Transportation Fuels via Trireforming and Fischer–Tropsch Synthesis: Model-Based Assessment, *Industrial & Engineering Chemistry Research*, 57, 30, 9964–9976.
- J. Xiao, Y. Peng, P. Bénard, and R. Chahine, 2016, Thermal effects on breakthrough curves of pressure swing adsorption for hydrogen purification, *International Journal of Hydrogen Energy*, 41, 19, 8236–8245.
- M. Kacem, M. Pellerano, and A. Delebarre, 2015, Pressure swing adsorption for CO<sub>2</sub>/N<sub>2</sub> and CO<sub>2</sub>/CH<sub>4</sub> separation: Comparison between activated carbons and zeolites performances, *Fuel Processing Technology*, 138, 271–283.
- M. Kumita, F. Watanabe, and M. Hasatani, 1994, Effect of water vapor on CO<sub>2</sub>-PSA separation performance, in *Ecomaterials*, Elsevier, 329–332.
- S. M. W. Wilson, D. A. Kennedy, and F. H. Tezel, 2020, Adsorbent Screening for CO<sub>2</sub>/CO Separation for Applications in Syngas Production, *Separation and Purification Technology*, 236, 116268.
- V. D. B. C. Dasireddy, D. Vengust, B. Likozar, J. Kovač, and A. Mrzel, 2021, Production of syngas by CO<sub>2</sub> reduction through Reverse Water–Gas Shift (RWGS) over catalytically-active molybdenum-based carbide, nitride and composite nanowires, *Renewable Energy*, 176, 251–261.

# Simultaneous design of integrated cyanobacteria-based biorefinery and its heat exchanger network

Matías Ramos<sup>a,b</sup>, Romina Lasry Testa<sup>a,b</sup>, Fernando Ramos<sup>a,b</sup>, Vanina Estrada<sup>a,b</sup>,  
Maria Soledad Diaz<sup>a,b</sup>

<sup>a</sup>*Planta Piloto de Ingeniería Química (PLAPIQUI CONICET-UNS), Camino La Carrindanga km. 7, Bahía Blanca, Argentina*

<sup>b</sup>*Departamento de Ingeniería Química, Universidad Nacional del Sur (UNS), Bahía Blanca, Argentina*

## Abstract

In this work we propose a mixed integer nonlinear programming (MINLP) model for the optimal design of an integrated cyanobacteria-based biorefinery and its heat exchanger network (HEN), considering alternative *in-vivo* and *in-silico* *Synechocystis* sp. PCC 6803 strains, designed in our previous work. The objective is to potentially produce pigments, biopolymers, bioethanol, biogas and biofertilizers. The production process includes different alternatives embedded within a superstructure. To assess the feasibility of the biorefinery, we considered the Sustainability Profit (SP), a monetary-based sustainability metric as the objective function. The Net Present Value (NPV) was also calculated. Both indexes give positive values and numerical results provides useful insights on advanced biofuels production with metabolic engineered cyanobacteria. Comparison with suboptimal solutions using *in-vivo* strains is also provided.

**Keywords:** MINLP, *Synechocystis* sp. PCC 6803, Cyanobacteria-based biorefinery, *In-silico* cyanobacteria

## 1. Introduction

Cyanobacteria are promising candidates to become cell factories, as these photoautotrophic microorganisms are capable of growing on atmospheric and/or industrial carbon dioxide, inorganic phosphorus (P) and nitrogen (N) as macronutrients, and solar and/or artificial light as energy source. One interesting high added-value product for food and nutraceutical industries produced by cyanobacteria is the photosynthetic pigment phycocyanin, which possesses an intense blue color. Also, some cyanobacteria are able to store carbon as poly- $\beta$ -hydroxybutyrate (PHB), a biopolymer similar to polypropylene with a wide range of applications. On the other hand, photoautotrophic fourth generation bioethanol production by cyanobacteria has been widely studied *in-vivo* (Dienst *et al.*, 2014) and *in-silico* (Lasry Testa *et al.* 2019, 2022) with mutant strains of *Synechocystis* sp. PCC6803 (*Synechocystis*).

In this work, we aim to assess the efficiency of *in-silico* strains of *Synechocystis* sp. PCC6803, developed in previous work (Lasry Testa *et al.* 2019, 2022), in a large scale biorefinery. We propose a mixed-integer nonlinear programming (MINLP) model for the simultaneous optimal plant design and heat exchanger network synthesis (HEN) of a cyanobacteria-based integrated biorefinery for pigments (phycocyanin and zeaxanthin), PHB, biofuels (bioethanol and biogas) and biofertilizers. To enhance cyanobacterial biomass valorization and biofuels production, and to close the nutrient (N and P) cycle in

the integrated biorefinery, an anaerobic digester (AD) is included in the superstructure. Four *Synechocystis* strains are considered as production options: three *in-silico* strains previously developed by our group (Lasry Testa *et al.* 2019, 2022) and one *in-vivo* strain (Vidal Vidal, 2009; Delpino *et al.* 2014). The idea behind evaluating the performance of the developed *in-silico* strains is to determine if it is worth the effort to design them, based on metabolic engineering in the laboratory. The objective function is the sustainability profit (Zore *et al.*, 2017) for the assessment of economic, environmental and social benefits

## 2. Process Description

### 2.1 *Synechocystis* strains

In this work, four different *Synechocystis* strains constitute the main “cell factories” for this biorefinery superstructure, as we aim at assessing *in-silico* developed strains efficiency within the frame of an integrated biorefinery. S1 (Vidal Vidal, 2009; Delpino *et al.*, 2014) is an *in-vivo* genetically engineered strain that harvests the insertion of *pdh* and *adh* genes from *Zymomonas mobilis*, to create an ethanol production pathway (as *Synechocystis* do not naturally produce ethanol). The remaining strains are *in-silico* ones, we have designed in previous work. S2 is a wild type strain (no ethanol production) represented by its genome scale model. It comprises 784 reactions and 535 metabolites, with 80 exchange reactions that include cytoplasm, carboxisome, thylacoidal lumen, thylacoidal membrane, cytoplasmic membrane, periplasm, extracellular space (Lasry Testa *et al.*, 2019). S3 includes the reactions codified by the genes *pdh* and *adh* and produces ethanol coupled to growth (Lasry Testa *et al.*, 2019), and S4 produces PHB coupled to growth (Lasry Testa *et al.*, 2022). The coupled strains were designed by formulating a bilevel optimization problem that identifies gene deletions to achieve the desired coupling. Coupling production to growth has the objective of turning the desired product production into a subproduct of growth, so that it becomes necessary to the microorganism’s metabolic function. In the bilevel optimization problem, the outer objective function is to maximize product production rate, while setting an upper bound to the number of gene deletions (represented through binary variables), and the inner optimization problem minimizes product production rate subject to the metabolic network model (LP). The problem has been reformulated as a single level optimization problem by applying dual theory; i.e., replacing the inner LP by a set of equations comprising the primal LP constraints, its dual problem constraints and the strong duality condition (primal LP objective function equal to dual problem objective function). In this case, the dual problem objective function has bilinear continuous-discrete terms, which have been replaced by exact linearizations (Glover, 1975). The resulting Mixed Integer Linear Programming (MILP) model renders a genetically engineered strain that couples ethanol production to cell growth through fourteen genetic intervention in the case of S3 (Lasry Testa *et al.*, 2019) and with sixteen genetic interventions for PHB production in the case of S4.

### 2.2 Process Superstructure

The proposed superstructure of a cyanobacteria-based biorefinery for the potential production of pigments (phycocyanin and zeaxanthin), PHB, bioethanol, biogas and

biofertilizers is shown in Fig. 1. It includes three main processing stages: production, separation and purification.

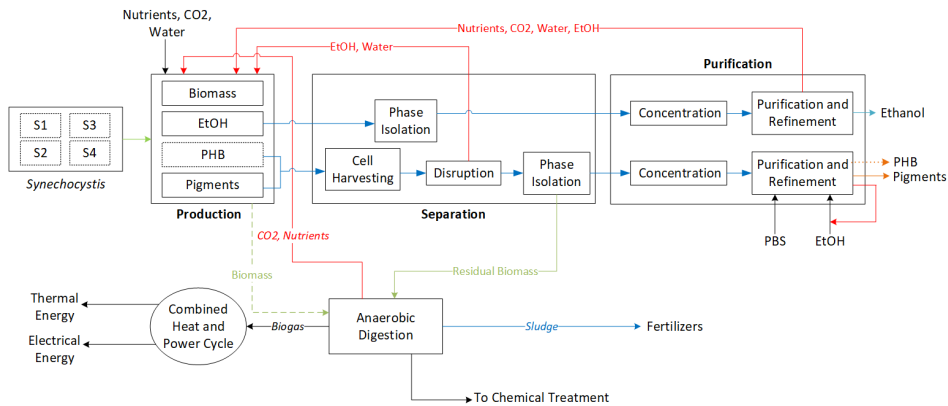


Figure 1. Cyanobacteria-based biorefinery simplified superstructure

The process begins with the production stage, where the four *Synechocystis* strains are considered for cultivation in open ponds at 30 °C, using CO<sub>2</sub> as carbon source, water, and N and P as nutrients. If S1 or S3 are chosen for cultivation, the harvesting process is carried out by microfiltration membranes to separate ethanol and culture media from the biomass. On the other hand, if S2 or S4 are selected, a centrifugation step is required to remove and recycle the culture media from the biomass.

A fraction of the harvested ethanol and water mixture (containing CO<sub>2</sub> and nutrients) from the microfiltration membranes (S1 or S3), is recycled to the open ponds in order to reach a minimal ethanol concentration threshold of 5 g/L, to attain an efficient performance of pervaporation membrane units in the subsequent ethanol separation and purification stages. The stream fed to the ethanol separation section is pre-heated to 50 °C in a heat exchanger before entering the polydimethylsiloxane membrane (PDMS) for pervaporation (where a concentration of 55 g ethanol/L is achieved). Next, ethanol is fed to a vapor compression distillation system where it reaches a concentration of 76%. The compressed vapor (ethanol/water mixture) is sent to a ceramic membrane followed by a cooling exchanger (25 °C) in order to obtain bioethanol fuel-grade concentration of 99.5% (Lopes *et al.*, 2019). Removed culture media streams from PDMS, ceramic membranes and vapor compression distillation system are also recycled to the cultivation system in order to enhance the sustainability feature of the proposed cyanobacterial biorefinery. For the four strains, the main stream of harvested biomass can be directly fed to an anaerobic digester for the production of fertilizers and biogas, and to recycle nutrient streams to the cultivation stage (García Prieto *et al.*, 2017). Also, it can be led to the separation section for pigments or PHB (S4 strain only) extraction. At the pigment separation stage, the wet biomass is dried in a spray drier with hot air (110 °C) in a 5.03 air-to-water ratio. In this step, most of the water and ethanol remnants are removed and recycled to the production stage after a cooling (40 °C) and flash operation. The dried biomass is disrupted in a bed mill, and then, a phosphate buffer solution (PBS) in a 17.25 PBS-to-phycoerythrin ratio is used in a diafiltration system to separate phycoerythrin from the residual biomass and zeaxanthin. In order to obtain commercial phycoerythrin, the purification of this pigment is conducted by freeze drying, removing mostly all of the water containing phosphates. For the case of the hydrophobic pigment, zeaxanthin, it is



extracted in a solid mixer-settler extractor, using ethanol as solvent in a 7.43 ethanol-to-zeaxanthin ratio. Finally, the residual biomass is sent to the anaerobic digester and the zeaxanthin is purified by chromatographic columns where the removed ethanol is reused in the solvent-extractor. The separation and purification stage for PHB, when strain S4 is cultivated, is described in detail in Ramos *et al.* (2019).

### 3. Mathematical Model

The proposed superstructure is formulated as a mixed-integer nonlinear programming (MINLP) problem and implemented in GAMS 35.2.0 (McCarl *et al.*, 2021) in order to determine the optimal design of an integrated cyanobacteria-based biorefinery and its HEN. The objective function is to maximize the monetary-based sustainability metric, Sustainability Profit (SP) (Zore *et al.*, 2017). The proposed superstructure includes mass and energy balances for the integrated biorefinery process, as well as its heat exchanger network (HEN) design and connection equations to link process design variables with HEN variables. Binary variables are associated to potential units and to heat exchanger matches.

### 4. Numerical Results

The MINLP model formulated for sustainability optimization and simultaneous process and HEN design includes 4,522 discrete variables, 24,840 continuous variables and 34,905 constraints. It was solved to an objective function value of SP=50.77 MMUS\$/y using DICOPT, with CONOPT and CPLEX as nonlinear and linear subsolvers, respectively (Grossmann *et al.*, 2003). Table 1 shows the alternative strain yields, used as input for the model.

Table 1. *In-vivo* (S1) and *in-silico* (S2, S3, S4) strain yields

Strain	Biomass yield (g/(L.d))	Ethanol yield (g/(L.d))	PHB yield (g/(L.d))
S1	0,339	0,288	0
S2	1,375	0	0
S3	0,669	0,875	0
S4	0,229	0,00	0,119

A required production of 180 t/y of phycocyanin/h (industrial level) and 1700 t/y of ethanol is fixed. The optimal scheme corresponding to SP maximization includes the use of *Synechocystis* strains S2 and S3 for the production of 20.55 kg phycocyanin/h, 0.53 kg zeaxanthin/h, 200 kg /h ethanol and 10.86 kg biofertilizer/h. The flowsheet, with the optimal mass flowrates of the main biorefinery streams is presented in Fig. 2. It must be noted that the production of PHB is not selected in the optimal biorefinery scheme.

Regarding heat integration, additional runs were carried out without heat integration and numerical results show that the optimal integrated scheme increased the number of heat exchangers required, as compared to the optimal base case without HEN (10 vs 7) and it also reduced the cost of external utilities by 60 %. Moreover, the reduction of utilities matches from 7 to 5 accounts for significant utility savings that improve the sustainability objective function. As it occurs with other sustainability analysis papers from the literature, the most important contribution to the objective function is given by the economic pillar (Pedrozo *et al.*, 2022). According to the results shown in Table 2, a negligible difference of 0.53 % is found in this pillar between the two proposed cases.<sup>a</sup>

This is due to the increase in the number of heat exchangers for the heat integration case (10 HXs) with respect to the one without heat integration (7 HXs). However, there is a meaningful improvement of nearly 22 % in the environmental pillar, while only an enhancement of 2.70 % can be reached in the social pillar. Regarding the economic pillar, a further analysis demonstrates that utilities reduction does not impact this pillar as other variables, such as maintenance and repairs cost. The social pillar difference is associated with the increase in the number of heat exchangers after heat integration, representing more workers' contracts. A decrease of fossil fuels requirements (external utilities), positively benefits the environmental pillar in the objective function.

In order to gain more insight into the study of the genetically engineered *in-silico* strains, we run an additional optimization case, imposing that S3 (*in-silico* ethanol producer) cannot be selected. *In-vivo* (S1) and *in-silico* (S2) strains are selected for ethanol and pigment production. A comparison between the latter case with the optimal one, can be found in Table 2 (fourth column). The sustainability objective function has a 60% decrease (from 50.78 to 20.55 MMUS\$/y) by the use of the *in-vivo* strain. The main difference is due to higher capital and operating costs associated to the suboptimal solution (S1 and S2) that negatively impact the economic pillar. Moreover, a decrease in the environmental pillar was also found. The obtained results are associated with the *in-vivo* lower biomass and bioethanol yields, which result in higher process flowrates that demand larger equipment size and energy requirements.

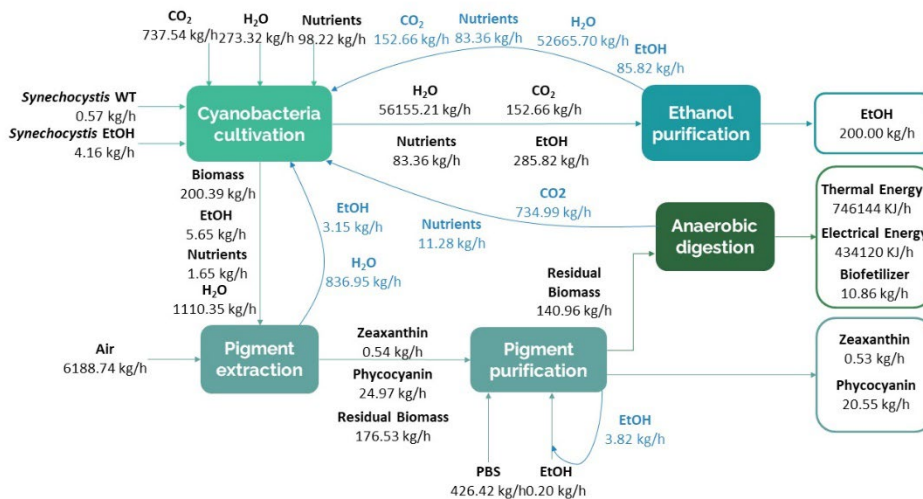


Figure 2. Integrated biorefinery optimal scheme for SP maximization

Table 2. Sustainability Profit pillars comparison

	Without HEN	With HEN	
	S2 and S3 (optimal)	S2 and S3 (optimal)	S1 and S2
SP (MMUS\$/y)	50.44	50.78	20.55
$P_{\text{economic}}$ (MMUS\$/y)	47.43	47.18	19.14
$P_{\text{environmental}}$ (MMUS\$/y)	2.64	3.22	1.02
$P_{\text{social}}$ (MMUS\$/y)	0.37	0.38	0.39

## 5. Conclusions

In this work, we have addressed the production of pigments, biopolymers, bioethanol, biogas and biofertilizers with alternative strains of *Synechocystis* sp. PCC 6803 by a mixed integer nonlinear programming (MINLP) model for the optimal design of an integrated biorefinery and its heat exchanger network (HEN). A superstructure with different process alternatives and different *Synechocystis in-vivo* and *in-silico* strains was formulated. *In-silico* strains S2 and S3 are the ones selected in the optimal solution. Both the SP and the NPV for the biorefinery are positive. A production cost of 0.38 USD/kg for fourth generation bioethanol is achieved due to incomes from pigments (phycocyanin and zeaxanthin) sales. Overall, the numerical results obtained are encouraging into further improving and testing the superstructure model, as well as they show that the use of *in-silico* S2 and S3 strains increases up to 60% the SP, as compared to the *in-vivo* ethanol producer strain S1. Strains S2 and S3 would be possible candidates to be designed by metabolic engineering through gene editing tools such as CRISPR-Cas systems.

## References

- I. E. Grossmann, J. Viswanathan, A. Vecchiotti, R. Raman, E. Kalvelagen, 2003. GAMS/DICOPT: A discrete continuous optimization package. Washington, DC, USA.
- C. Delpino, V. Estrada, J. Laglecia, R. Vidal Vidal, F. Florencio, M. García Guerrero, M. S. Diaz, 2014. Dynamic flux balance analysis in cyanobacteria for ethanol production with simultaneous optimization approaches. *Comp. Aided Chem. Eng.* 33, 1165-1170.
- D. Dienst, J. Georg, T. Abts, L. Jakowec, E. Kuchmina, T. Börner, A. Wilde, U. Dühring, H. Enke, W.R. Hess, 2014. Transcriptomic response to prolonged ethanol production in the cyanobacterium *Synechocystis* sp. PCC6803. *Biotechnol. Biof.* 7, 21.
- C. V. García Prieto, F. D. Ramos, V. Estrada, M. A. Villar, M. S. Diaz, 2017. Optimization of an integrated algae-based biorefinery for the production of biodiesel astaxanthin and PHB. *Energy*. 139, 1159-1172.
- F. Glover, 1975. Improved linear integer programming formulations of nonlinear integer problems. *Manag. Sci.* 22, 455-460.
- R. Lasry Testa, C. Delpino, V. Estrada, M. S. Diaz, 2019. Bioethanol Production with Cyanobacteria by a Two-Stage Fermentation Strategy. *Comp. Aided Chem. Eng.* 46, 499-504.
- R. Lasry Testa, C. Delpino, V. Estrada, M. S. Diaz, 2022. Development of *in silico* strategies to photoautotrophically produce poly- $\beta$ -hydroxybutyrate (PHB) by cyanobacteria. *Algal Research*. 62, 102621.
- T. F. Lopes, C. Cabanas, A. Silva, D. Fonseca, E. Santos, T. L. Guerra, C. Sheahan, A. Reis, F. Girio, 2019. Process simulation and techno-economic assessment for direct production of advanced bioethanol using a genetically modified *Synechocystis* sp. *Bioresource Technology Reports*. 6, 113-122.
- B. A. McCarl, A. Meeraus, P. van der Eijk, M. Bussieck, S. Dirkse, P. Steacy, F. Nelissen, 2017. McCarl GAMS user guide, Washington, DC, USA.
- H. A. Pedrozo, A. I. Casoni, F. D. Ramos, V. Estrada, M. S. Diaz, 2022. Simultaneous design of macroalgae-based integrated biorefineries and their heat exchanger network. *Comp. Chem. Eng.* 164, 107885.
- H. Shokravi, M. Heidarrezaeri, Z. Shokravi, H. C. Ong, W. J. Lau, M. F. Md Din, A. F. Ismail, 2022. Fourth generation biofuel from genetically modified algal biomass for bioeconomic development. *J. Biotechnol.* 360, 23-36.
- R. Vidal Vidal, 2009. Producción fotosintética de etanol por la cianobacteria *Synechocystis* sp. PCC6803. PhD Dissertation, University of Sevilla, Spain.
- Z. Zore, L. Čuček, Z. Kravanja, 2017. Syntheses of sustainable supply networks with a new composite criterion – Sustainability profit. *Comp. Chem. Eng.* 102, 139-155.

# Sustainable Conceptual Design of a Hydrogen Production via Natural Gas pyrolysis

Patience B. Shamaki,<sup>a</sup> Galo A.C Le Roux<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, Politechnic school, universidade de Sao Paulo, Av. Prof. Luciano Gualberto, trav.3,n380, Sao Paulo, Brazil*  
[galoroux@usp.br](mailto:galoroux@usp.br)

## Abstract

Methane pyrolysis involves a simple 1-step endothermic reaction with solid carbon (C) and hydrogen (H<sub>2</sub>) as the main products. The aim of this study is to perform a sustainable conceptual process design including heat integration and sustainability analysis of hydrogen production through natural gas (NG) thermal decomposition. Natural gas was used as process feed instead of pure methane and five component products: H<sub>2</sub>, C, C<sub>2</sub>H<sub>4</sub>, C<sub>2</sub>H<sub>2</sub>, and C<sub>6</sub>H<sub>6</sub>. The process design, thermodynamic analysis, heat integration and sustainability analysis considering the economics and emissions using different energy sources was performed. As expected, results show conversion at lower temperatures for NG as compared to pure methane feed. It also showed the potential of utilizing heat from the process to improve energy efficiency in an integrated process. The cost of NG, C sales and C credit incentive significantly influence the process profitability. The economics also showed operating cost reduction by 21.6% after heat integration and about 78kgCO<sub>2</sub>/kgH<sub>2</sub> was avoided using this process to produce H<sub>2</sub>.

**Keywords:** methane pyrolysis, process design, hydrogen.

## 1. Introduction

The global pressure for sustainable processes and fuels to mitigate carbon footprint has garnered considerable attention to hydrogen. As a versatile and valuable raw material for industrial processes, hydrogen can also serve as clean fuel, being an energy carrier. Moreover, it can easily be mixed with one of its main raw materials CH<sub>4</sub> for cleaner and better combustion (Ishaq & Dincer, 2020). Even though hydrogen itself is a clean and abundant element present in most compounds, its production is usually not as clean. Hydrogen is primarily produced through the reforming of hydrocarbons, which produce high CO<sub>2</sub> content as a process by-product besides the emission from the process energy source, about 8-12 kgCO<sub>2</sub>/kgH<sub>2</sub> is produced from the steam methane reforming process (SMR). A cleaner alternative is the use of water splitting, but the process is thermodynamically limited given that only the hydrogen in the water compound contains energy, the oxygen byproduct is essentially not so useful. This is why thermal decomposition of methane (TDM) also known as methane pyrolysis (MP) has received considerable attention in recent years as an economical means of producing low to zero emission hydrogen, in comparison to other hydrogen alternatives. It involves a simple 1-step endothermic reaction primary process to produce hydrogen and solid carbon as the main products. Due to the stability of methane, the decomposition process involves some secondary and auxiliary reactions producing other higher hydrocarbons at increasing

temperature and consequent considerable heat consumption before it fully decomposes to the primary products (Sánchez-Bastardo et al., 2021). However, most literature only report the main products  $H_2$ , C without considering the impact of the intermediate products. Furthermore, the two main challenges limiting its commercialization involves the complexity of dealing with high temperature production, as well as the production of excess carbon beyond market demand. This is why several researches are ongoing to exploit this potential low to zero emission hydrogen production route in the current climate faced with increasing energy demand and declining climatic conditions due to increased emissions. (Naikoo et al., 2021) provided a detailed analysis on the different types of MP such as thermo-catalytic decomposition to reduce reaction temperature however, limited by catalyst degeneration due to high temperature, the use of molten metals or salt to promote heat integration (Leal Pérez et al., 2021) or the use of plasma torch and regenerative heat exchanger reactor (RHER) (Keipi Tiina, 2017; Kerschler et al., 2021) at very high temperature. Additionally, more studies of heat integration to reduce energy waste from the process are necessary. In this study, the thermodynamical feasibility of a sustainable process design is investigated and analyzed, even though the kinetics could be also limiting it is important to investigate the boundaries of the process to enable further research and innovative process design. Therefore, they are not going to be considered in this contribution.

## 2. Methodology and Process Description

### 2.1. Methodology

The process of hydrogen production using natural gas thermal decomposition was simulated using the Peng-Robinson- Boston-Mathias (PR-BM) as the property model method. The Gibb's energy minimization as the model for the reactor in ASPEN Plus. The base process as presented by Cheon, et al. (2021) was simulated, although in this work a RHER was not used. One of our objectives is to study the thermodynamically feasible extent of process heat integration, so as to investigate the important modalities required for a sustainable process design. The process data from ASPEN Plus simulation was used to perform heat integration via the PINCH analysis in the ASPEN Energy Analyzer, given that the reaction is slightly endothermic, but is favored by high temperatures. The analysis enabled minimum utility requirement/maximum exchanger requirement and the cheapest design was selected as seen in figure 1. The thermodynamic and sustainability analysis of the heat integrated process design considering the economics and the carbon footprint from energy sources was performed since the process itself does not produce  $CO_2$  directly. The main process components include a pyrolysis reactor, a cyclone for solid separation and a PSA separation unit as shown in the process flow diagram from fig 1. The analysis based on the process thermodynamic properties performed enabled the choice of initial process flow design operating conditions.

### 2.2. Process simulation description and assumptions

As shown in the base process flowsheet diagram in figure 1, the natural gas feed is pre-heated to reactor temperature and fed into the reactor at  $1195^\circ C$  and 1 bar, this provides a global conversion rate of 95% and the products consisting of  $H_2$ , solid carbon, ethylene, acetylene and benzene, are sent to a cyclone (dust collector) with high efficiency to remove the solid carbon before it is cooled for hydrogen purification in the PSA unit (compressor, cooler and SEP component). For the separation 90% of the product gas is considered hydrogen, 8% unconverted methane and other hydrocarbons are recycled and 2% purged as flue gas. 5% of the recovered hydrogen is fed back to the process to improve the feed quality and the 95wt % pure hydrogen contains some gases from the product

mixture. The base heat supply for the process is natural gas, other sources of heat supply (grid electricity (coal), renewable sourced energy) was also investigated for economic and emission footprint, given that process itself does not produce CO<sub>2</sub> directly but through the heating sources. The simulation was performed for a H<sub>2</sub> capacity of 414 kg/hr this equals a medium scale hydrogen production process.

Given the peculiarity of the high temperature favored pyrolysis process, the economic feasibility analysis was performed using the order of magnitude costs curve methods of capital and operating cost estimation techniques proposed in Towler and Sinnott (2012). Here the total invested cost (TIC) is divided by the total hydrogen produced to obtain the unit cost of hydrogen, and an optimization was performed to minimize the unit cost of hydrogen produced subject to the plant capacity. The itemized cost data was obtained from literature (Chen et al., 2021; Cheon et al., 2021; Towler & Sinnott, 2013) and ASPEN Plus software, and scaled to the predefined plant capacity; with preliminary capital cost estimation accuracy of +/- 30%. Table 1 shows the simulated process flow information.

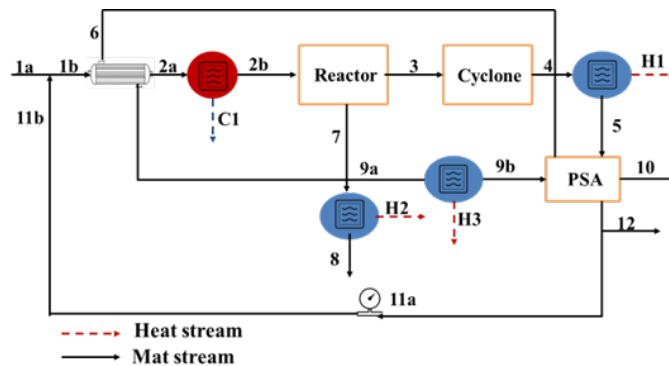


Figure 1: Heat Integrated methane pyrolysis process

### 3. Results and Discussions

#### 3.1. Thermodynamic Analysis

The influence of the presence of other hydrocarbons in natural gas can be seen in the decomposition at lower temperatures when compared to utilization of pure methane feed from figure 2. The secondary products ethylene, benzene and acetylene were produced in very little quantities, with acetylene produced at higher quantity than the ethylene and benzene at very high temperature. Given that the quantity produced is very small, the secondary products considered were separated from the H<sub>2</sub> in the PSA unit and sent together with the unconverted methane to recycle, this significantly improved conversion rate to primary products. Figure 2 also showed that conversion rate is favored by low pressure. The consequence of utilizing methane as opposed to natural gas for process feed is evident in the emission rate as shown in table 1. This is due to the heat duty required to raise the temperature high enough to decompose pure methane.

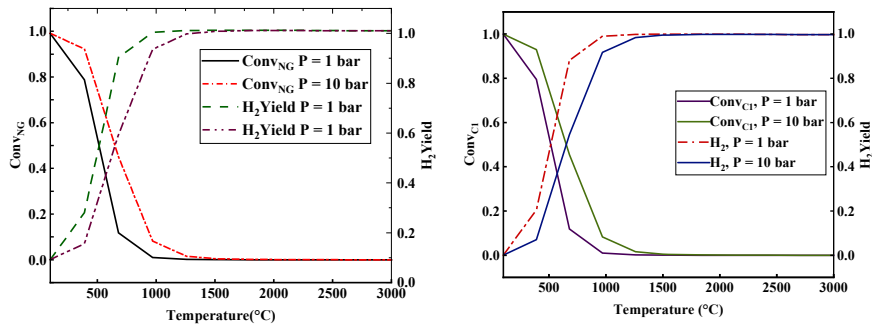


Figure 2: NG decomposition vs Methane conversion rate and product yield

Table 1: Material balance of thermal decomposition process

Param/Stream No	1a	1b	2a	2b	3	4	5	6
Temperature (°C)	25	23	440	1195	1195	1195	50	50
Pressure (bar)	1	1	1	1	1.01	1.01	1	11
<b>Mol flow (kmol/h)</b>								
NG	99	118	118	118	342	230	230	205
C1	112	132	132	132	357	244	244	219
Param/Stream No	7	8	9a	9b	10	11a	11b	12
Temperature (°C)	1195	30	487	139	50	50	26.3	50
Pressure (bar)	1.01	1	11	11	11	11	1.11	11
<b>Mol flow(kmol/h)</b>								
NG	112	112	230	230	230	19	19	6
C1	112	112	244	244	244	20	20	5

### 3.2. Sustainability analysis

#### 3.2.1. Economics

For the economic analysis, the heat integration performed reduced the utility cost by 52% and consequently the operating cost by 21.6%. The heat integration for the single process of thermal decomposition was constrained by the process thermodynamics, more heat could be recovered when this process is integrated with other processes depending on the end use of the produced hydrogen and carbon. Furthermore, there is added technological complexity due to the high temperature required for the process as well as solid carbon recovery from the reactor, one challenge is the ability to recover heat from the solid carbon produced. This process requires a special heat exchange design as it cannot be simulated with the common shell and tube heat exchanger. (Pérez, et al. 2021) utilized a plate heat exchanger designed for solid-fluid (liquid or gas) heat exchange. These challenges reflect on the total cost of production and consequently on the price of hydrogen from MP process. Sensitivity analysis was performed on the designed process plant to ascertain parameters that can facilitate the point of ‘break even’, that is the point where the net profit value (NPV) equals zero. The selling price of hydrogen, carbon, discounted cash flow rate of return (DCFRROR) from 0-35%, as well as carbon credit for

CO<sub>2</sub> emission avoidance was evaluated varying hydrogen price from 1 – 10\$/kg, NG feed cost 0.1-1\$/kg, carbon price from 0.00 – 2 \$/kg, and carbon credit from 0-0.01\$/kgCO<sub>2</sub> avoidance. The analysis showed that hydrogen selling price, considering NG price at 0.4651\$/kgNG (being the highest spot price with NG cost hike), without any carbon sales and carbon credit must be 2.5\$/kg to break even. This hydrogen selling price is a slightly above the price of hydrogen from SMR but less than and competitive with hydrogen produced via SMR with carbon capture sequestration (CCS). Carbon sales at 0.1618\$/kg could enable hydrogen from MP to be competitive or at par with SMR with selling price of 2\$/kg and feed cost of 0.4651\$/kg. Moreover, carbon credit at 0.0051\$/kgCO<sub>2</sub>avoided could also lead to break even resulting in the 2\$/kgH<sub>2</sub> and feed cost of 0.4651\$/kgNG. However, for the price 0.3\$/kgNG (the average spot price) used for this work, H<sub>2</sub> selling price of 1.8\$/kg breaks even, the DCFROR analysis also showed that at 29.48% interest rate, the NPV equals zero as shown in figure 4.

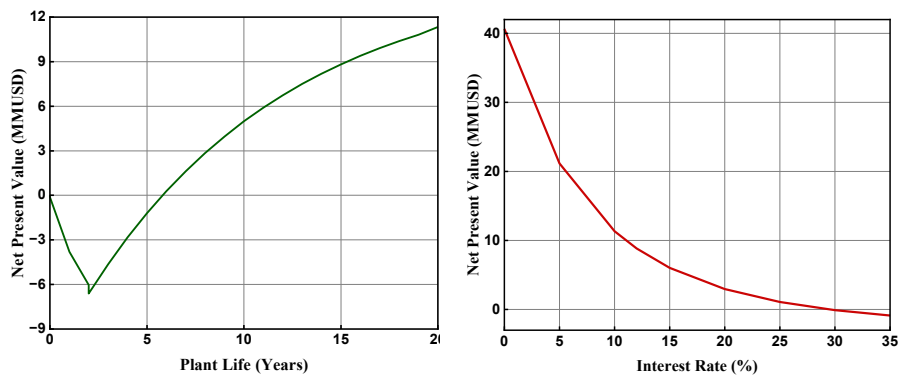


Figure 3: DCFROR and Hydrogen selling price sensitivity analysis

### 3.2.2. Carbon footprint

For the carbon footprint, we analyzed of the CO<sub>2</sub> emitted and CO<sub>2</sub> emission avoided when using NG as process feed against methane as process feed with three different energy sources (Natural gas, Coal electricity from the grid (US) and renewable energy sources). As expected, the process with methane as a feed emits slightly more than natural gas feed given that the main source of CO<sub>2</sub> emission is the energy source, and the heat required for the decomposition is higher with methane feed. On the average the process produces 3.3kg CO<sub>2</sub>/kgH<sub>2</sub> and 0.04 kgCO<sub>2</sub>/kgH<sub>2</sub> when renewable energy source is used, noting that the process with NG gas presents some CO<sub>2</sub> in the waste stream, accounting for the CO<sub>2</sub> limit in pipeline gas. Notwithstanding, this stream waste is insignificant as the process with pure methane still emitted more than that with NG feed. When comparing to other hydrogen production processes such as steam reforming with an average of 10kg CO<sub>2</sub>/kgH<sub>2</sub> without carbon capture, the thermal decomposition of NG/methane significantly reduced the carbon footprint.

The total heating duty required for the base case was 5,474kW and 5,874kW, for NG feed and CH<sub>4</sub> feed respectively, and the cooling duty was 4,167kW and 4,353kW respectively for each feed. And after heat integration, the heating duty was reduced to 4,699kW and 5,054kW for each feed, and cooling duty was also reduced to 3387kW and 3715kW for each.



Table 2: CO<sub>2</sub> emission and CO<sub>2</sub> avoidance per heat source

CO <sub>2e</sub> (kg CO <sub>2</sub> /kg H <sub>2</sub> )	NG source	Grid elect (coal)	Renewable source
NG feed	3.42	9.99	0.046
C1 feed	3.53	10,282	0.015
Tot.CO <sub>2av</sub> (kg CO <sub>2</sub> /hr)	NG source	Grid elect source	Renewable source
NG feed	-33885	-31160	-35284
C1 feed	-43437	-40448	-44993

#### 4. Conclusion

A sustainable design of the thermal decomposition of natural gas was considered in this work, an optimal process heat integration design was made and the process analyzed considering thermodynamics, economics and carbon footprint mitigation. The results showed that selling carbon and carbon tax/credit is essential for plant profitability as enable an economical low emitting hydrogen production process. Furthermore, results show that the methane pyrolysis process enables better energy efficiency when integrated with other processes than as a stand-alone based on process complexity from thermodynamics and technological limitations. Given the economic potential, it will be beneficial to develop novel designs to overcome these challenges. Further works will focus on integrating the process with the end use process, conceptual heat exchanger network design that could enable recovery of heat from the hot carbon

#### Acknowledgement

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001 and Petroleum Technology development fund – Nigeria (PTDF).

#### Reference

- Chen, Z., Zhang, R., Xia, G., Wu, Y., Li, H., Sun, Z., & Sun, Z. (2021). Vacuum promoted methane decomposition for hydrogen production with carbon separation: Parameter optimization and economic assessment. *Energy*, 222.
- Cheon, S., Byun, M., Lim, D., Lee, H., & Lim, H. (2021). Parametric study for thermal and catalytic methane pyrolysis for hydrogen production: Techno-economic and scenario analysis. *Energies*, 14(19). <https://doi.org/10.3390/en14196102>
- Ishaq, H., & Dincer, I. (2020). Performance investigation of adding clean hydrogen to natural gas for better sustainability. *Journal of Natural Gas Science and Engineering*, 78.
- Keipi Tiina. (2017). Technology Development and Techno-Economic Analysis of Hydrogen Production by Thermal Decomposition of Methane.
- Kerscher, F., Stary, A., Gleis, S., Ulrich, A., Klein, H., & Spliethoff, H. (2021). Low-carbon hydrogen production via electron beam plasma methane pyrolysis: Techno-economic analysis and carbon footprint assessment. *International Journal of Hydrogen Energy*, 46(38), 19897–19912. <https://doi.org/10.1016/j.ijhydene.2021.03.114>
- Leal Pérez, B. J., Medrano Jiménez, J. A., Bhardwaj, R., Goetheer, E., van Sint Annaland, M., & Gallucci, F. (2021). Methane pyrolysis in a molten gallium bubble column reactor for sustainable hydrogen production: Proof of concept & techno-economic assessment. *International Journal of Hydrogen Energy*, 46(7), 4917–4935.
- Naikoo, G. A., Arshad, F., Hassan, I. U., Tabook, M. A., Pedram, M. Z., Mustaqeem, M., Tabassum, H., Ahmed, W., & Rezakazemi, M. (2021). Thermocatalytic Hydrogen Production Through Decomposition of Methane-A Review. *Frontiers in Chemistry*, 9(October), 1–24.

- Sánchez-Bastardo, N., Schlögl, R., & Ruland, H. (2021). Methane Pyrolysis for Zero-Emission Hydrogen Production: A Potential Bridge Technology from Fossil Fuels to a Renewable and Sustainable Hydrogen Economy. *Industrial and Engineering Chemistry Research*, 60(32), 11855–11881. <https://doi.org/10.1021/acs.iecr.1c01679>
- Towler, G. P., & Sinnott, R. K. (2013). *Chemical engineering design : principles, practice, and economics of plant and process design*. Butterworth-Heinemann.



# The influence of biomass characteristics and their uncertainties on the production of sustainable aviation fuel

Moaaz Shehab,<sup>a,b,c</sup> Diego Freire Ordóñez,<sup>c</sup> Mai Bui,<sup>c</sup> Kai Moshhammer,<sup>a</sup> Edwin Zondervan<sup>b</sup>

<sup>a</sup>*Physikalisch Technische Bundesanstalt (PTB), Bundesallee 100, Braunschweig 38116, Germany*

<sup>b</sup>*Twente University, Drienerlolaan 5, Enschede 7522 NB, Netherlands*

<sup>c</sup>*Imperial College London, Exhibition Rd, South Kensington, London SW7 2BX, UK.*

## Abstract

Sustainable aviation fuel (SAF) plays an important role in decarbonizing the aviation sector. The ASTM D7566 dictates several pathways to produce sustainable fuels that share the same characteristics as the conventional jet fuel. One of the common pathways to produce SAF is via the gasification of biomass to generate syngas. In this work a model was developed for the process to evaluate the use of different biomass sources on the SAF yield. Moreover, the influence of the measurement uncertainty in the biomass characteristics on the process performance was investigated. The study has shown that the hydrogen content is the most crucial element in the biomass to obtain a higher SAF yield. On the other hand, the uncertainty of the biomass characteristics such as the moisture content causes a 1.5% variation in the final SAF yield. Such analysis shows the urge for an accurate and reliable measurement of the biomass characteristics, this allows for a structural embedding of the uncertainty while making decisions during the process design and evaluation stage.

**Keywords:** Biomass gasification, sustainable aviation fuel, biomass characteristics.

## 1. Introduction

Aviation fuel accounts for 3% of the EU's total greenhouse gas emissions. To tackle this problem, the EU proposed a mandate to include a sustainable aviation fuel (SAF) uptake as a drop-in fuel, with 2% SAF starting from 2025 and increase up to 63% by 2050. SAF can be produced via different pathways which are regulated by the ASTM standard D7566. These pathways vary for different feedstocks. The most common feedstocks are waste oil and fats, cover crops, agriculture and forest residuals and the organic fraction in municipal solid waste (MSW). The Hydro processed Esters Fatty Acids (HEFA) process uses waste oil and fats to produce SAF, while for the agriculture and forest residual, gasification coupled with fischer tropesch (GFT) or alcohol to jet (ATJ) could be used. As gasification technology is widely and commercially established on coal gasification, the focus of this study is on utilizing different types of biomass in the gasification process to produce syngas. This syngas will be upgraded by the Fischer-Tropsch process (FT) to produce liquid hydrocarbons. To describe and understand which potential yields of SAF can be obtained from biomass, data on the biomass characteristics is required. However, such attributes are subject to enormous variations from source to source and from season to season. Therefore, the aim of this study is to analyze the effect of the different biomass

types, their characteristics such as moisture content, elemental analysis and the energy content as well as their respective uncertainties on the production process of SAF. The biomass data used in the analysis was provided by the BIOFMET project where different biomass types were characterized and their respective uncertainties were determined.

## 2. Methodology

A block diagram was developed for the process of biomass conversion into sustainable aviation fuel based on the available literature data for biomass gasification and the Fischer-Tropsch process as shown in figure 1. (Almena et al. 2022; König et al. 2015; Larson et al. 2020).

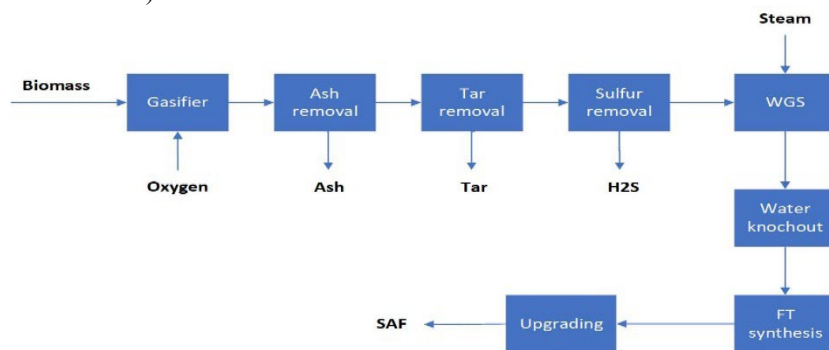
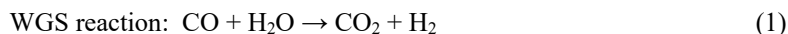


Figure 1. Block diagram for biomass to SAF conversion process.

The biomass is fed to a gasifier to produce syngas using oxygen. The syngas produced from the gasifier contains ash, tar and H<sub>2</sub>S which need to be removed before being supplied to the water gas shift reactor (WGS) as those impurities can deactivate the catalyst in both the WGS and FT. In the WGS, the clean gas reacts with a super-heated steam to increase the H<sub>2</sub>-CO ratio to 2:1 according to equation 1.



The steam will react with the carbon monoxide to produce more hydrogen and carbon dioxide. The syngas will be directed to the FT to produce long chain hydrocarbons according to equation 2.



After the formation of hydrocarbons, the syngas enters a series of flash drums where the temperature and pressure are adjusted accordingly to distil the SAF fraction. ASPEN Plus V10 was used to simulate the process and calculate the mass- and energy balances of the process streams. As a property method, the Peng-Robinson equation of state in combination with the Boston-Mathias equation was used. Oxygen gasification was selected as the air separation unit produces a nitrogen-free syngas which results in a syngas that is better suitable for the Fischer-Tropsch process. FT can be performed at lower as well as at higher temperatures. For the synthesis of long chain hydrocarbons low temperature operation of FT is recommended. In the FT reactor a closed loop design was opted to maximize the CO conversion which resulting in a higher SAF yield. The scale of the model was set around 1 GW which is equivalent to a feed rate of 300 ton/hour biomass. Three different types of biomass were analysed, wood chips of high quality (WC-HQ), wood chips of industrial quality (WC-IQ) and wood pellets (WP). Within the

framework of the BIOFMET project, the characteristics of those samples were measured, and their respective uncertainties were determined and minimized.

Table 1. Biomass properties.

Property	Wood chips - IQ	Wood chips – HQ	Wood pellet
C	47.13	47.08	50.22
H	5.69	5.61	5.84
N	0.198	0.074	0.093
S	0.01245	0.0058	0.0069
O	44.99	46.99	43.56
Ash	1.97	0.24	0.28
Moisture	10.76	11.24	5.93
Volatile Matter (VM)*	<u>17–25</u>	<u>17–25</u>	<u>17–25</u>
Fixed Carbon (FC)*	<u>70–90</u>	<u>70–90</u>	<u>70–90</u>

\* The values of the volatile matter and fixed carbon are assumed based on the literature data for woody biomass. An average value of for VM and FC were used (Dai et al. 2015).

The model was used to test how the different biomass sources and their characteristics could influence the production of SAF. Operating conditions such as the temperature, pressure and the flow rate need to be optimized and fixed to have a fair comparison.

### 2.1 Model optimization

The gasification process decomposes the biomass and leads to a syngas that contains CO, CO<sub>2</sub>, H<sub>2</sub>, CH<sub>4</sub> along with undesired impurities such ash, tar and H<sub>2</sub>S. Therefore, optimizing the gasifier is crucial to maximize the CO and H<sub>2</sub> yields, while lowering the concentration of the impurities. An effective gasification process operates at an equivalence ratio (ER) of 0.2 to 0.4. A value less than 0.2 lead to a dominant pyrolysis reaction, while for an ER of more than 0.4, the combustion reaction will dominate and more CO<sub>2</sub> and H<sub>2</sub>O produced. In figure 2, a temperature of 840 °C was selected for the gasifier and different flowrates of oxygen were tested to determine the optimum ER ratio.

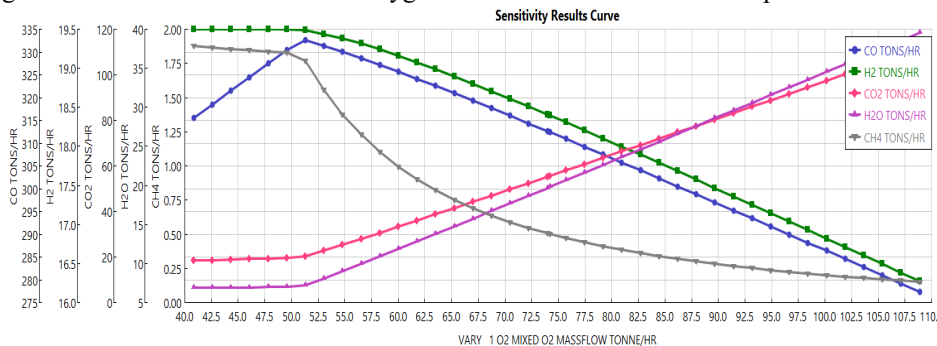


Figure 2. Sensitivity analysis for the oxygen flowrate against the syngas composition.

An oxygen flow rate of 78 ton/hr was opted which corresponds to an ER ratio of 0.2. This flow rate guarantees that more CO and H<sub>2</sub> are produced. For FT, different catalysts such as Ru, Fe, Ni and Co can be used to produce hydrocarbons. Ni has a high selectivity towards methane, while Ru is too expensive for use at a commercial scale. An iron catalyst is relatively tolerant to sulfur as compared to cobalt and can be used to adjust the H<sub>2</sub>/CO ratio thanks to its water gas shift reaction capability. However, it has a low selectivity towards long chain hydrocarbons. Therefore, iron is not suitable for the production of sustainable aviation fuel. Cobalt based catalysts produce less unsaturated hydrocarbons and alcohols, they have a long lifetime and are suitable to produce long

chain alkanes. The Anderson–Schulz–Flory distribution, with a propagation probability  $\alpha$  (growth factor = 0.95) as the only parameter was used to predict the fractional distribution of the different alkanes. The temperature and pressure of the FT reactor was adapted from the experimental results performed by König et al. (König et al. 2015). The reactor operates at 225 °C and a pressure of 2.5 MPa.

### 3. Results and discussion

The Aspen model was used to analyse the differences between the wood chips and wood pellet in terms of their influence on SAF production as shown in figure 3.

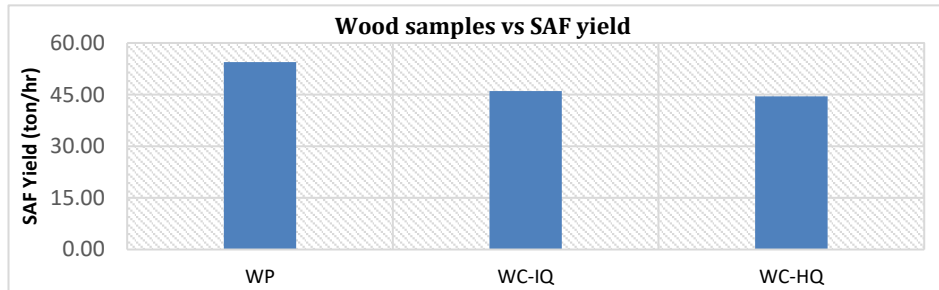


Figure 3. SAF yield of different biomass types.

WP provides the higher SAF yield compared to WC because of having a higher hydrogen and carbon content, lower moisture and ash content.

#### 3.1. The influence of moisture uncertainty on the SAF yield

After drying the biomass and before feeding it to the gasifier, a sample is analysed to determine its final moisture content. Determining the moisture content and its respective uncertainty is scientifically challenging as a lot of factors influence the final results, starting from the experimental procedure, the surrounding conditions and the biomass composition. Typically, the uncertainty of the moisture content of the biomass can vary from  $\pm 1\text{--}5\%$ . In the model, the operating parameters were maintained fixed to analyse the influence of the moisture change on the final SAF yield as shown in figure 4.

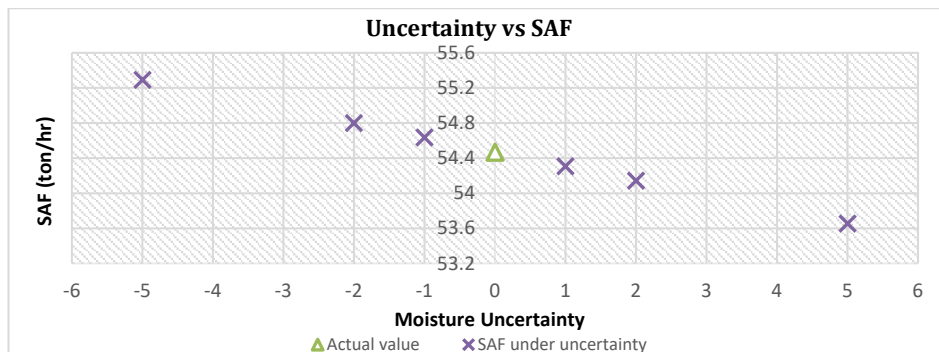


Figure 4. SAF yield of WP vs moisture uncertainty.

Wood pellet was used to perform the above-mentioned analysis. The actual value represents the obtained yield by using the measured moisture of 5.93% as given in table 1. It can be concluded that the final yield varies depending on the uncertainty. At an uncertainty of  $\pm 5\%$ , the SAF production will increase or decrease by 1.5% (around 0.83 ton/hr). Theoretically, 1.5% could be accepted, in reality a decrease of 0.83 ton/hr is significant as this corresponds to thousands of tons per year. Moreover, unforeseen higher/lower level of the moisture content due to the uncertainty leads to problems in the

gasifier. In case of an increase in the moisture content, the temperature of the gasifier will decrease which influences the syngas quality and lowers its heating value. Furthermore, the rate of tar cracking would decrease due to lower temperatures. In that case, the operator has to adjust the gasifier temperature. Therefore, accurately measured data with specified uncertainty is important to help the operator in deciding how to minimize the impact of any unpredicted variation in the process.

### 3.2. The influence of C,H,N,S,O uncertainty on the SAF yield

As syngas contains primary CO and H<sub>2</sub>, the percentage of carbon, hydrogen and oxygen in the biomass would largely determine the SAF yield. Generally, biomass has an average ratio of 45–55% carbon, hydrogen of 5–7% and oxygen of 35 – 50% (Dai et al. 2015). During the process design phase it is important to consider the uncertainty of the different parameters to correctly size the instruments and have an accurate average for the final yield. For example, hydrogen has a very low density which leads to a large specific volume. For the given values above, the average uncertainties as provided from the BIOFMET data is for H around ± 0.15 %, while for both C and O is ± 1%. Small changes in composition can influence the yield as shown in figure 5.

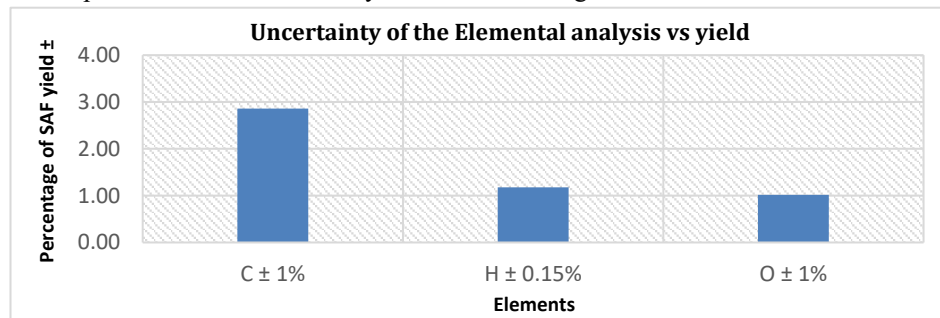


Figure 5. SAF yield vs the uncertainty of the elemental analysis.

A change of ±1% in C and O would influence the SAF yield by up to ± 3 and ± 1% respectively. In case of hydrogen a change of ± 0.15 % influences the yield by ± 1.3%. Hydrogen is the most influential factor in the process. Therefore, the choice for a biomass source should be based on its hydrogen content and not the carbon content. The average range of carbon in the different biomass types is enough to provide the process with a sufficient CO.

### 3.3 The influence of the calorific value on the gasification performance

The energy content of the biomass could be determined experimentally by a bomb calorimeter or by using the elemental analysis. According to a survey performed within the scope of the BIOFMET project, the uncertainty of the energy content supplied to industry reaches up to 5 % (Shehab et al. 2022). The energy content influences the calculations of the cold gas efficiency (CGS) of the gasifier and the overall process performance. Therefore, a small error in the measurement raises the overall uncertainty of the process, as the uncertainty distribution might not be concentrated around the average. To calculate the cold gas efficiency, the lower heating value (LHV) of biomass should be calculated as  $LHV = HHV - 10.55(W + 9H)$ .  $W$  is the moisture weight fraction and  $H$  is the hydrogen percentage in the biomass. An example is given below for the wood pellet. Shehab *et al.* 2022, provided an improved technical protocol to determine the energy content of the biomass with low uncertainty of ± 1%. Based on those experimental data, the LHV of wood pellet is 19.92 MJ/kg, while the resulting LHV of the syngas provided by the ASPEN model is 13.56 MJ/kg. Equations 3



was used to calculate the CGS of the gasifier in case of 5% uncertainty and 1% uncertainty.

$$\text{Cold gas efficiency: } \frac{\text{LHV of syngas}}{\text{LHV of Biomass}} * 100 = 68.08 \% \quad (3)$$

An uncertainty of  $\pm 5\%$  causes inaccuracy in the efficiency of the gasifier by over  $\pm 3.2\%$ . While an uncertainty of  $\pm 1\%$  in the energy content will lead to a substantially lower uncertainty of  $\pm 0.6\%$ . Such an approach shows the importance of having a reliable testing mechanism for the biomass as it directly influences the process performance and efficiency. This information helps the process designer to provide a realistic range of uncertainty to the unit operation and the process efficiency.

#### 4. Conclusions

An Aspen model was developed to simulate the conversion of biomass into sustainable aviation fuel. Optimizing the model was crucial to provide an accurate prediction of the syngas composition when using oxygen as a gasifying agent. Using different biomass types on the model proved that a slight variation in the biomass characteristics could substantially influence the SAF yield. Hydrogen is the most sensitive element in the biomass, therefore, selecting the biomass should be based on hydrogen content. Knowledge about the uncertainty of the biomass characteristics is crucial, so the process designer could provide a better estimation of the unit operation uncertainty and the final SAF yield

#### Acknowledgments

This BIOFMET research (19ENG09) was funded from the EMPIR programme (European Metrology Programme for Innovation and Research) co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme.

#### 5. References:

- Almena, Alberto; Thornley, Patricia; Chong, Katie; Röder, Mirjam (2022): Carbon dioxide removal potential from decentralised bioenergy with carbon capture and storage (BECCS) and the relevance of operational choices. In *Biomass and Bioenergy* 159, p. 106406. DOI: 10.1016/j.biombioe.2022.106406.
- Dai, Jianjun; Saayman, Jean; Grace, John R.; Ellis, Naoko (2015): Gasification of Woody Biomass. In *Annual review of chemical and biomolecular engineering* 6, pp. 77–99. DOI: 10.1146/annurev-chembioeng-061114-123310.
- Klüh Daniel; Gaderer Matthias (2022): Integrating a Fischer-Tropsch Fuel Production Process into Pulp Mills. In *Chemical Engineering Transactions* 94, pp. 13–18. DOI: 10.3303/CET2294002.
- König, Daniel H.; Baucks, Nadine; Dietrich, Ralph-Uwe; Wörner, Antje (2015): Simulation and evaluation of a process concept for the generation of synthetic fuel from CO<sub>2</sub> and H<sub>2</sub>. In *Energy* 91, pp. 833–841. DOI: 10.1016/j.energy.2015.08.099.
- Larson, Eric D.; Kreutz, Thomas G.; Greig, Chris; Williams, Robert H.; Rooney, Tim; Gray, Edward et al. (2020): Design and analysis of a low-carbon lignite/biomass-to-jet fuel demonstration project. In *Applied Energy* 260, p. 114209. DOI: 10.1016/j.apenergy.2019.114209.
- Shehab, Moaaz; Stratulat, Camelia; Ozcan, Kemal; Boztepe, Aylin; Isleyen, Alper; Zondervan, Edwin; Moshhammer, Kai (2022): A Comprehensive Analysis of the Risks Associated with the Determination of Biofuels' Calorific Value by Bomb Calorimetry. In *Energies* 15 (8), p. 2771. DOI: 10.3390/en15082771.

# Translation Ontology of optimal Decision Making along a Distributed Production Chain by Example

Peter Klein<sup>a</sup>, Heinz A. Preisig<sup>b</sup>, Natalia Konchakova<sup>c</sup>

<sup>a</sup>*Fraunhofer Institute for Industrial Mathematics, Kaiserslautern, Germany*

<sup>b</sup>*Norwegian University of Science and Technology, Trondheim, Norway*

<sup>c</sup>*Helmholtz-Zentrum Hereon, Geesthacht, Germany*

## Abstract

Traditionally, product design is done from within a company. From a purely economic point of view, it will focus on optimising product performance while minimising costs and waste. The optimisation uses key performance indicators, which often result from modelling the production processes and the product's properties. Increasingly, though, product design and product manufacturing involve groups of companies, each of which has its own objectives and strategies. The central idea is to couple the single company-based silo optimisation to a collaborative optimisation and decision-making to unlock hereto hidden product innovation potentials on a digital platform. This platform supports the modelling process from describing the distributed manufacturing and the involved materials to product design, which we call *Translation*. In this contribution, we discuss a joint translation ontology, capturing the translation process and the recursive business-to-business process implemented in an H2020-funded innovation platform VIPCOAT.

**Keywords:** translation ontology, optimally distributed production, innovation environment, product design.

## 1. Introduction

The European Materials Modelling Council (EMMC) has developed a Translation process performed by so-called Translators to support the industry in utilising materials modelling and facilitating innovation. The delivered Translation Guide (Hristova-Bogaerds et al. 2019) describes in detail a six-step sequential translation process: business case identification → industrial case understanding → analysis of available data space → identification (translation) to modelling workflow → model execution and validation → back-translation of modelling results to business case. An extension of this guide, in particular of its last step, establish Business Decision Support Systems (BDSS), which introduce business-related key performance indicators (KPI) into the picture (Dykeman et al. 2020) and systematise the original back-translation process. Recent developments in this area are summarised in a report on the progress of implementing the translation and BDSS concept in various running European projects, see (Klein, Konchakova, et al. 2021).

In parallel to the standardisation of translation processes, first steps in materials modelling standardisation and ontologisation have been made by the EMMC community. These attempts are documented in a CEN workshop agreement on Modelling Data, MODA CWA (CEN workshop agreement, CWA 17284:2018 (E) 2018), and Elemental

Multiperspective Material Ontology (EMMO) (Ghedini et al. 2021). An ontologisation of the MODA CWA, compatible with the EMMO, has been described for one particular industrial process in (Klein, Preisig, et al. 2021), where the development of corrosion-inhibiting coatings served as an example of a translation process initially documented in a MODA and a set of roles to automatically map this MODA description to an EMMO compatible format of materials modelling workflows. The construction principle used is now systematised in a further CEN workshop agreement, see (CEN workshop agreement, CWA 17960:2022 (E) 2022).

In this paper, we describe the basic steps for the ontologisation of the corresponding translation process suitable for coating development and take into account that the development and production processes of coatings are distributed (chained) among different actors in the industry.

## 2. Industrial challenges and requirements

In general, new innovative product designs within a company aim to optimise product performance and customer acceptance while minimising costs and waste. On top, most companies implement long-term strategies like maintaining good supplier and customer relations, technological leadership, and, nowadays, sustainable, resilient, and green production and supply. The optimisation uses key performance indicators (KPIs), often resulting from modelling the production processes and the product's properties.

Increasingly, though, product design and product manufacturing involve groups of companies, each of which has its own objectives and strategies. The central idea of this work is to couple the single company-based silo optimisation to a collaborative optimisation and decision-making process to unlock hereto hidden product innovation potentials on a digital platform. This approach necessarily calls for interoperability along the value-added chain, which is greatly supported by ontologies. In this contribution, we discuss a joint translation ontology, capturing the translation process and the recursive business-to-business (B2B) process implemented in an H2020-funded innovation platform VIPCOAT (“Virtual Open Innovation Platform for Active Protective Coatings Guided by Modelling and Optimization”) for the development of new coating systems. The best-known inhibiting pigments in metal coatings contain hexavalent chromium-based compounds, or Cr(VI) compounds for short, as pigments. Acknowledging human health and environmental issues has led these Cr(VI) compounds to be heavily regulated, so a replacement in corrosion-inhibiting coatings, to be developed by coating suppliers, is necessary.

The project involves three major production companies: a nanoscale corrosion inhibitor producer, a coating formulation producer and a large consumer with very tight and demanding specifications related to the performance of the applied coating formulations on metal surfaces. In this B2B2B innovation environment, collaborative decision support is deployed on a cloud platform with endpoints in the companies, representing our ontology's first use case. An ontologisation of the materials modelling support to be used within each company is provided in ref. (Klein, Preisig, et al. 2021).

## 3. Ontology construction principles

The first step towards ontologisation of the VIPCOAT platform was done for the modelling layer in ref (Klein, Preisig, et al. 2021), compatible with the EMMO (Ghedini et al. 2021). As an intermediate step in this works, we first construct an ontologisation of

the traditional translation process in a Translator to Business (T2B) relation, reflecting the EMMC Translation Guide (Hristova-Bogaerds et al. 2019). This ontology is consistent with a developed crosswalk from an initial VIPCOAT MODA workflow of the three industrial applications to the EMMO.

Translation as an interactive process is conceptually described in the ontology as subclasses of the translation class: partners and performance. While the Partners' subclass model basically the interaction between humans involved in a translation process, the Performance subclass models the realisation of business goals.

#### 4. T2B and T2B2B Translation ontology

This initial Translation ontology models a one-to-one relationship between the two triples (Translator, Modelling, Simulation) and (Client, Business Case, Industrial Case) - T2B relationship. In the next step, this ontology will be extended to handle a one-to-many relationship (T2B2B) between the triple (Translator, Modelling, Simulation) and a chain of (Client, Business Case, Industrial Case) triples, thus ontologically modelling a B2B2B environment.

The approach used in this works implements a somewhat different perspective on ontologies. Usually, ontologies are constructed to systematise existing knowledge. Consecutively, we use the ontology-defined syntax and semantics to construct correspondingly structured knowledge, like ontology-driven tools and knowledge graphs. This approach is our paradigm for implementing ontology-based interoperable processes and tools.

##### 4.1. Ontology for T2B Translation Processes

Translation as an interactive process using the original T2B pattern, see (Klein, Konchakova, et al. 2021), consist of two important concepts which are described as subclasses of an ontological Translation class: Partners and Performance. While the Partners model the interaction basically between humans involved in a Translation process, the Performance class models the realisation of business goals in an industrial Case or process. RDF-type data file for further processing of this ontology can be made available upon request.

One Translator- one client relationship structure has been added to the ontology and mapped into an RDF-type data file (see Figure 1).

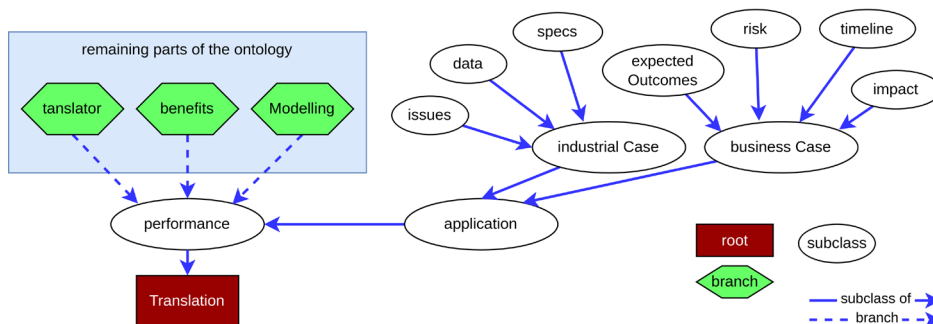


Figure 1: Section of the T2B ontology

The BDSS system deployed at the industrial partner sites may be used to visualise the results of the tasks and/or of experimental data individually at each site.

Decision Support is implemented on the VIPCOAT platform by multi-criteria optimisation and interactive explorers of the solutions of corresponding optimisation models. While multi-criteria optimisation is a well established mathematical technology used to find best balanced compromises between conflicting objectives (in the technological area usually called Key Performance Indicators), decision makers may interactively explore what-if-scenarios and add context information not modelled in the multi-criteria optimisation in order to take informed final decisions on products and processes.

#### 4.2. Ontology for one-to-many Translation Processes

Applied to the example, VIPCOAT project picks this challenge up as depicted in Figure 2.

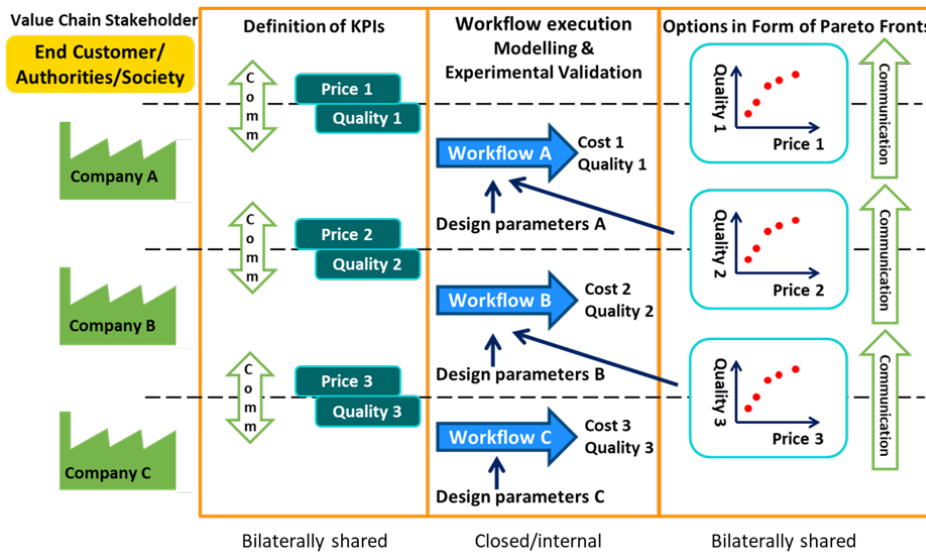


Figure 2: B2B2B environment used in VIPCOAT.

Partner C develops new pigments containing regulation-compliant inhibiting molecules. Partner B adds these pigments to formulate coatings, and Partner C applies this product on surfaces and tests their inhibiting performance to meet the End Customer's expectations. This VIPCOAT platform demonstration use case is thus structured along a value/production chain in a B2B2B environment. Typical for such environments is that the requirements flow is upstream of the production chain, starting at the End Customer. In VIPCOAT, the requirements specification on coatings is formulated to be applicable for aircraft manufacturers. The KPIs are very detailed and contain many more technical details to be fulfilled by a paint producer's coating formulation besides the corrosion-inhibiting properties. These requirements generally persist while the Cr(VI) compounds are to be replaced, a typical B2B supplier - OEM (original equipment manufacture) relation: The paint producer buys the pigments from his supplier. However, many of the requirements are not that hard specified. Also, the interaction between the different

involved materials is not fully predictable. Together this provides a certain level of freedom of choice, allowing the complete chain to choose an optimal combination satisfying the coating-applying partner's requirements.

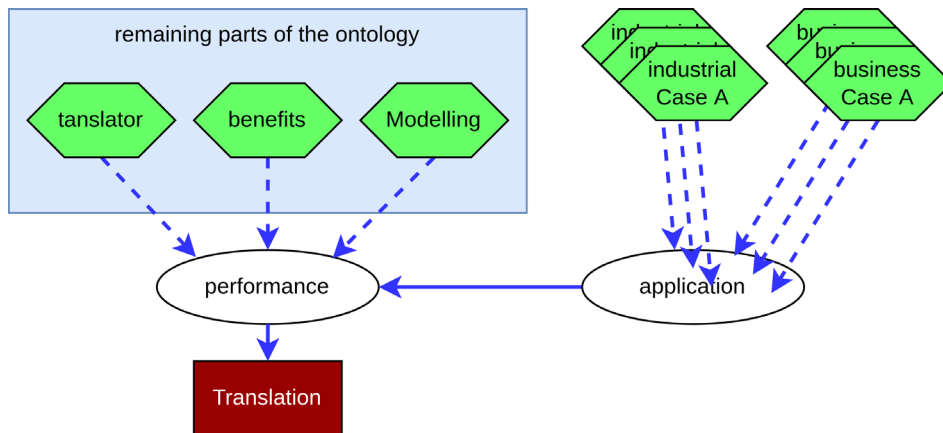


Figure 3: Sketch of the T2B2B ontology: relationship of one Translator and three industrial and business Cases (A, B, C)

VIPCOAT addresses these development challenges directly through its modelling tasks serving the most important steps in coating development by utilisation of integrated simulation and optimisation tools.

Further developments towards B2B2B collaborative decision-making, one of the VIPCOAT platform goals, call for a good understanding of B2B2B relations to implement such a new product development style. An ontology segment for the One-Translator and a B2B2B environment applied to active protective coating creation example is presented in Figure 3.

## 5. Conclusion

This contribution presents an ontological approach along a distributed production chain designed for protective coating developments. In this example, the materials modelling approach has already been ontologised, so a complete set of ontologies, including business-related ontological classes, are created. In its current implementation stage, the VIPCOAT platform takes advantage of these ontologies to support inter-company collaboration for fast innovations of corrosion-inhibiting coatings used in the aeronautic industry. The next step will be the introduction of domain ontologies to capture more general coating system formulation. These ontologies will get constructed in parallel to the VIPCOAT platform developments to support cross-fertilisation at least in the coating industry.

**Acknowledgements:** VIPCOAT H2020-NMBP-TO-IND-2020, Grant Agreement No. 952903; MARKETPLACE H2020-NMBP-25-2017, Grant Agreement No. 760173.

## References

- CEN workshop agreement, CWA 17284:2018 (E). 2018. “Materials Modelling: Terminology, Classification and Metadata.”
- CEN workshop agreement, CWA 17960:2022 (E). 2022. “ModGra - Graphical representation of physical process models.”
- D. Dykeman, A. Hashibon, P. Klein, and S. Belouettar, 2020, Guideline Business Decision Support Systems (BDSS) for Materials Modelling, Technical report. EMMC. <https://doi.org/10.5281/zenodo.4054009>.
- E. Ghedini, G. Goldbeck, J. Friis, et al., 2021, Elementary Multiperspective Material Ontology (EMMO), GitHub. <https://github.com/emmorepo/410EMMO>.
- D. Hristova-Bogaerds, P. Asinari, N. A. Konchakova, L. Bergamasco, A. M. Ramos, G. Goldbeck, D. Höche, O. Swang, and G. J. Schmitz, 2019, EMMC Translators’ Guide, Technical report. EMMC. <https://doi.org/10.5281/zenodo.3552260>.
- P. Klein, N. Konchakova, D. G. Hristova-Bogaerds, M. Noeske, A. Simperler, G. Goldbeck, and D. Höche, 2021, Translation in Materials Modelling: Process and Progress, Technical report. OntoTRANS – FORCE. <https://doi.org/10.5281/zenodo.4729917>.
- P. Klein, H.A. Preisig, M. T. Horsch, and N. Konchakova, 2021, Application of an Ontology Based Process Model Construction Tool for Active Protective Coatings: Corrosion Inhibitor Release, In Proceedings of JOWO 2021, 26. CEUR-WS.
- Virtual Open Innovation Platform for Active Protective Coatings Guided by Modelling and Optimization (VIPCOAT, Grant Agreement No 952903). 2021. <https://cordis.europa.eu/project/id/952903/de>, <https://ms.hereon.de/vipcoat/>.

# The Integration of Drug Substance and Drug Product Manufacturing Models: The Missing Link for Model-based End-to-End Process Development

Charalampos Christodoulou<sup>a,\*</sup>, Samir Diab<sup>b</sup>, Gabriele Bano<sup>c</sup>, Magdalini Aroniada<sup>b</sup>, Neil Hodnett<sup>a</sup>, Simeone Zomer<sup>b</sup>

<sup>a</sup> GlaxoSmithKline, Gunnels Wood Road, Stevenage SG1 2NY, United Kingdom

<sup>b</sup> GlaxoSmithKline, Park Road, Ware SG12 0DP, United Kingdom

<sup>c</sup> GlaxoSmithKline, 1250 S Collegeville Rd, Collegeville (PA) 19426, United States

\*Corresponding author email: [charalampos.x.christodoulou@gsk.com](mailto:charalampos.x.christodoulou@gsk.com)

## Abstract

Demonstrations of pharmaceutical process modelling have been made in the literature in the pursuit of accelerated development. Modelling activities for drug substance (DS) and product (DP) manufacturing are often conducted separately, as project time constraints often require that the DS and DP processes be developed simultaneously, each with fixed decision criteria and requirements. We illustrate how connecting DS and DP process models is essential for a full process understanding and how key attributes of an isolated DS can propagate through a DP tableting line to impact the DP dissolution. We do so by performing an uncertainty analysis on a flowsheet model of a crystallization and direct compression line. The dissolution sensitivity highlights the need for end-to-end flowsheet modelling to realize the full benefits of modelling. We also discuss the key technical and regulatory obstacles that must be overcome to achieve this.

**Keywords:** Flowsheet model; Pharmaceutical manufacturing; Drug Substance; Drug Product

## 1. Introduction

The development of robust pharmaceutical manufacturing processes that ensure product quality, manufacturability and material efficiency requires a sound understanding of the impact of the process' inputs (e.g., raw material properties, process operating conditions) on product quality (e.g., purity, content uniformity, dissolution) across the operating space. Consideration of the end-to-end process in a rigorous manner can be challenging due to the large number of inputs to account for in a given process, as well as their multivariate interactions that potentially impact product quality and manufacturability.

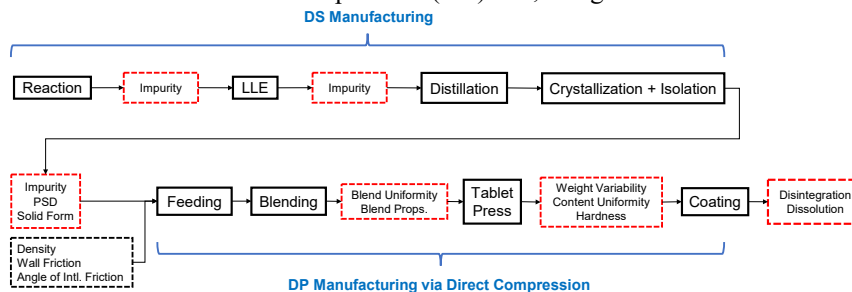
Mathematical modelling can assist this by describing the relevant physical and chemical phenomena occurring in each unit operation and connecting these sub-models into an integrated flowsheet model to assess the criticality of multiple input factors *in-silico* (Diab et al., 2022a; 2022b). In industrial practice, modelling activities for drug substance (DS) and drug product (DP) manufacturing can occur independently, as the short lifespan of a patent means that DS and DP process developments usually must occur simultaneously and risk not being perfectly harmonized, e.g., the selected DS solid form may change under certain DP process conditions. Additionally, most DP models, besides those employing physics-based methods such as Discrete Element Method (DEM) are semi-empirical and thus the benefit of connecting them to other first-principles models more commonly available in DS processes are then not as significant. This does not allow for a full consideration of the impact of DS input variability on the DP process and quality. This “disjointed” consideration of DS and DP models misses the full potential of end-to-end modelling to accelerate development.



## 2. Pharmaceutical Manufacturing Process Modelling

Various efforts have been made towards the modelling of DS system models. Examples include the use of models to predict input factor impacts on related impurity levels (Diab et al., 2022a; 2022b), avoidance of undesirable solid state forms and control of particle size distribution, PSD (Montes et al., 2018). Examples of models for DP manufacturing include wet granulation (Boukouvala et al., 2013), dry granulation (Bano et al., 2022a) and direct compression (García-Muñoz et al., 2018).

There are few literature cases of end-to-end process modelling that span the interfaces between DS, DP manufacture and *in-vitro* product performance. Benyahia et al. (2012) were the first to model an entire continuous pilot plant but did not consider product performance. Nagy et al. (2021) modelled a DS synthesis and crystallization process with *in vitro* dissolution, but without explicit modelling of the DS isolation and DP manufacturing unit operations. Most examples omit isolation, even though this can drastically impact the DS physical properties (Guo, 2009). Fig. 1 shows an example end-to-end process and how each unit operation potentially impacts a quality attribute. While many DS process models predict purity and PSD, they often do not output responses that are key inputs to DP models. For example, changes in crystallizer operation could impact the DS PSD and shape, which may then impact the powder (blend) flowability, tablet content uniformity and weight variability, etc. We now illustrate how variation in the DS properties that are DP model outputs can propagate through to the final DP quality for a tablet manufactured in a direct compression (DC) line, using a flowsheet model.



**Figure 1.** Example of an end-to-end pharmaceutical process. Black solid = Unit operation. Red = quality attribute. Black dashed = Material property not predicted from DS process model.

## 3. Case Study

The case study in this paper considers a seeded antisolvent crystallization from the final stage of a DS manufacturing process, a DC process model for DP manufacturing and subsequent *in vitro* product performance. The considered sub-models are all parameterized vs. industrial process data.

### 3.1 Drug Substance Crystallization

We consider a seeded, antisolvent DS crystallization. The supersaturated solution of DS in organic solvent is seeded at a fixed temperature and the resulting mixture is then aged. Following isothermal de-supersaturation, antisolvent is added and the suspension is further aged. The API mass balance is given by Eq. 1, where  $C_b$  = bulk solution concentration,  $\rho_c$  = crystal density,  $k_v$  = crystal volume-shape factor,  $G$  = overall growth rate,  $L$  = DS particle size,  $n$  = number size distribution. Growth contributions come from both pure API ( $G_{API}$ ) and impurity ( $G_{Imp}$ ) in Eqs. 2–4, where  $C_{sat}$  equilibrium solubility (a function of temperature,  $T$ , and antisolvent content,  $x_{AS}$ ),  $\rho$  = solvent density,  $C_{imp}$  = impurity concentration in the feed to the crystallization, and model parameters,  $k_g$ ,  $g$ ,  $s$ ,  $k_i$  and  $l$  are fit to de-supersaturation and impurity content (Taylor et al., 2021).

$$dC_b^{API}/dt = 3k_v\rho_c \int GL^2n(L)dL \quad (1) \quad dP_d/dt = \Lambda d_h^2 \varepsilon p_c / S_p \mathcal{T}_{or}^2 \mu P_d \quad (11)$$

$$G_{API} = k_g \left( \frac{C_b^{API} - C_{sat}}{\rho} \right)^g (x_{AS} + 1)^s \quad (2) \quad \frac{dM_t}{dt} = [2\rho_s(1 - \varepsilon)\Omega(t)] \frac{dP_d}{dt} \quad (12)$$

$$G_{imp} = G_{API} k_i (x + 1)^l C_{imp} \quad (3) \quad p_c = (\sigma \cos \theta_e) / d_h ; \quad d_h = [\varepsilon / (1 - \varepsilon)] (d_p / 6) \quad (13)$$

$$\frac{dL}{dt} = G = G_{API} + G_{imp} \quad (4) \quad \frac{\partial N_i}{dt} = B_i \delta(l - l_{0,i}) + R_{i,l} \frac{\partial N_i}{dl} \quad (14)$$

$$K = 1.5V_b^{1/3} F_h \omega_b t_{blend} \quad (5) \quad B_i = \frac{1}{\rho_p} \left( \frac{x_i}{S_p l_{0,i}^3} \right) \frac{dM_t}{dt} \quad (15)$$

$$sf = a_{sf} (1 + b_{sf} P) / (1 + a_{sf} b_{sf} P) \quad (6) \quad R_{i,l} = \dot{M}_l / \rho_p \quad (16)$$

$$TS = TS_0 ((1 - \beta) + \beta \exp(-\gamma K)) \quad (7) \quad \dot{M}_l = K_D (C_{sat} - C_b)^{n_{si}} \quad (17)$$

$$TS_0 = a_1 \exp(b_1 (1 - sf)) \quad (8) \quad K_D \approx 1/K_r + 1/K_d = (K_d K_r) / (K_d + K_r) \quad (18)$$

$$\beta = a_2 (1 - sf) + b_2 \quad (9) \quad C_b^{API} = x_{API} \rho_t \left( \frac{V_{t,0} - V_t}{V_t} \right) - \left( \frac{m_{API,3} \rho_p S_p}{V_m} \right) \quad (19)$$

$$V_c = (H_c - \dot{\varepsilon} t) A_t \quad (10) \quad \%LC = 100 C_b^{API} V_m / x_{API} \rho_t V_{t,0} \quad (20)$$

### 3.2 Blending and Direct Compression of Drug Product

The DC model is comprised of a blender and a tablet press. The Extended Kushner model (Nassar et al., 2021) is used to describe the tablet tensile strength ( $TS$ ) as a function of the extent of blend lubrication,  $K$ , as per Eqs. 5–9, where  $V_b$  = blender volume,  $F_h$  = fractional headspace in blender,  $\omega_b$  = blend speed,  $t_{blend}$  = blending time,  $sf$  = attained tablet solid fraction,  $P$  = pressure applied in tablet press,  $TS_0$  = tablet tensile strength at zero-porosity. Parameters  $a_{sf}$ ,  $b_{sf}$ ,  $a_1$ ,  $b_1$ ,  $a_2$ ,  $b_2$  and  $\gamma$  are estimated from experimental data.

### 3.3 In-Vitro Product Performance: Disintegration and Dissolution

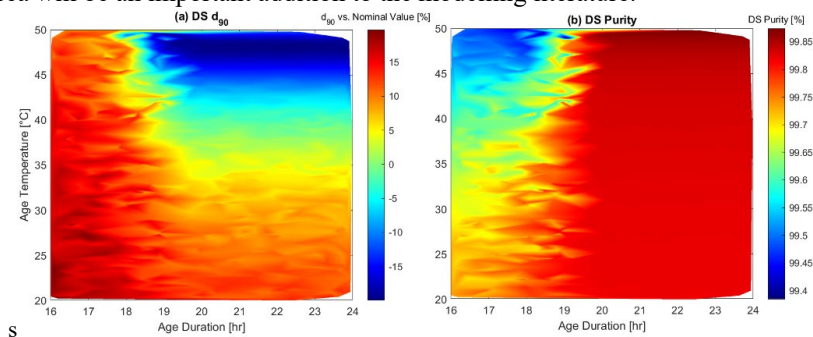
Both erosion and swelling are considered in the tablet disintegration (Markl et al., 2017). Erosion is described by Eq. 10, where  $V_c$  and  $H_c$  = coating layer volume and thickness, respectively,  $A_t$  = tablet surface area,  $\dot{\varepsilon}$  = erosion rate. The penetration depth due to swelling is given by Eq. 11, where  $P_d$  = penetration depth,  $d_h$  = pore hydraulic diameter,  $\mathcal{T}_{or}$  = tablet tortuosity,  $\mu$  = liquid viscosity,  $p_c$  = capillary pressure. Model parameters  $S_p$  and  $\Lambda$  are to be estimated from experimental data (Peppas and Colombo, 1989). The change in tablet mass,  $M_t$ , is given by Eqs. 12 and 13, where  $\varepsilon$  = average porosity of the swollen product,  $\rho_s$  = solid true density,  $\sigma$  = liquid surface tension,  $\theta_e$  = solid-liquid contact angle,  $d_p$  = tablet equivalent spherical volume diameter (Kuentz and Leuenberger, 1998) from disintegration time (DT). The DT is defined when  $dP_d/dt = 0$ .

Dissolution is described by Eqs. 14–20, where  $N_i$  = number size distribution,  $B_i$  = rate of particle release,  $l$  and  $l_{0,i}$  = particle size at  $t$  and  $t = 0$ , respectively,  $\rho_p$  = particle density,  $x_i$  = mass fraction,  $M_t$  = tablet mass,  $R_{i,l}$  = particle dissolution rate,  $\dot{M}_l$  = overall mass flux,  $K_D$  = overall dissolution coefficient, which is calculated from the surface integration constant,  $K_r$ , and the diffusion mass transfer coefficient,  $k_d$  (estimated from diffusion coefficient,  $D$ , and unstirred layer thickness,  $h_l$ ),  $n_{si}$  = order of dissolution. Model parameters  $K_r$  and  $n_{si}$  are to be estimated from dynamic dissolution data for tablets with different solid fractions. The dissolution extent is defined as the percentage label content, %LC, where  $V_m$  = liquid volume in the USP <711> (2011) test vessel.

#### 4. Uncertainty Analysis

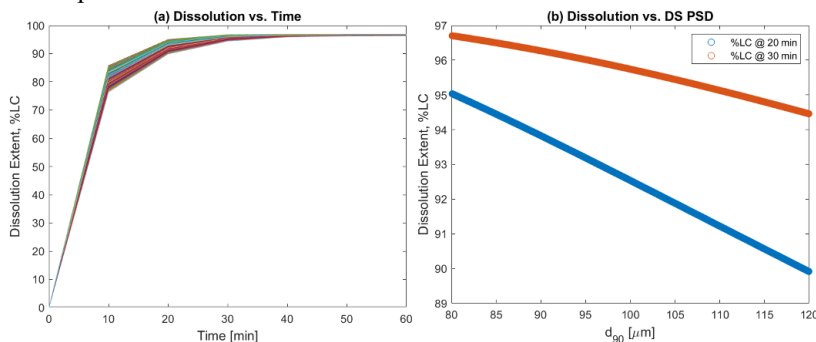
We quantify the impact of DS property variability propagation through to DP dissolution in two steps: (1) uncertainty analysis on the crystallization process model to understand the expected variance in DS properties, (2) uncertainty analysis to quantify the impacts of DS property variability propagating through the DC process to the DP dissolution using Sobol sampling in gPROMS FormulatedProducts. The input factors for the crystallization uncertainty analysis were the seeding and age temperature and age duration, seed loading, antisolvent addition duration and quantity. The responses were DS  $d_{90}$  and purity.

Fig. 2 shows the variation in DS  $d_{90}$  and purity as a function of age temperature and age duration, both of which impact the selected responses when varied over the ranges investigated. The total range of the simulated DS  $d_{90}$  was approximately  $\pm 20\%$  vs. the nominal value. The range of attained DS purity = 99.38–99.87 mol%. Even low levels of key DS impurities can be severe with respect to product quality, e.g., if they are genotoxic or mutagenic, and can also impact the crystallization kinetics and the resulting PSD, which is a direct input to the DP process. Modelling the impacts of impurities on crystallization kinetics is certainly not trivial (Capellades et al., 2022); further research in this area will be an important addition to the modelling literature.



**Figure 2.** Age temperature and duration vs. DS properties: (a)  $d_{90}$ , (b) purity.

For the DC analysis (step 2), the input was the DS  $d_{90}$  (whose range is defined from the observed responses in step 1). The considered response was the DP dissolution, whose variation is shown in Fig. 3. The extent of dissolution at 20 min varies significantly ( $\sim 90$ –95 %LC), while at 30 min the variation is slightly less ( $\sim 94.4$ –96.7 %LC) vs. DS  $d_{90}$  variation. Of course, the selection of the timepoint(s) at which dissolution extent is examined impacts how sensitive the dissolution behavior is.



**Figure 3.** (a) Variation in dissolution extent vs. time. (b) Dissolution extent at 20 and 30 min vs. DS  $d_{90}$ .

## **5. Conclusions and Discussion – Industrial and Regulatory Considerations**

The presented case study exemplifies how variation in DS properties can propagate to the final DP product performance and why modelling an end-to-end process is valuable. The modelling approach we have shown could also be used to directly inform the questions posed in the International Council for Harmonization (ICH) quality guideline Q6A (2000) Decision Tree #3, to assess whether an acceptance criterion is required for DS PSD. However, one must acknowledge that there remain certain obstacles to realizing end-to-end process modelling in general.

**Material Science:** Advancements in proven methods for DS bulk property prediction from the DS particle properties need to be made. A first principles description for this needs computational approaches to interpret and predict physical properties of crystalline solid forms incorporated into an integrated process model, including DS isolation steps – such studies are few in the literature (Destro et al., 2021).

**Blend Property Prediction:** Predicting blend properties from its composition and individual component properties is essential to describe the formulation-dependence of a process. Modelling the behavior of granular solid blends without DEM is not arbitrary. Many efforts have been made in the literature, only some of which are industrially pragmatic. Statistical approaches are limited by the range of data collated and cannot be reliably extrapolated outside of the considered calibration data space – such models that are currently available consider only a small range of flow characteristics for OSD forms. Approaches using the granular Bond number, although based more on fundamental particle behavior and interactions, require measurements that are not routinely collected in an industrial environment (e.g., particle surface energy). Bano et al. (2022b) proposed mixing rules for blend flow properties based on the van der Waals rules for fluid mixtures, while Jolliffe et al. (2022) proposed a protocol for modelling powder blend compaction properties. More efforts such as these for a wider set of blend properties are needed.

**Standardized, Systematic Model Evaluation:** Process modelling is still in its relative infancy in pharmaceutical manufacturing, and thus there can be a lack of widespread confidence and uptake of models by the pharmaceutical industry. Specifically, there may be a lack of scientists with the appropriate expertise for model development in a company, as well as resource constraints and a lack of management and stakeholder support. One component of the latter point is that there is no standardized workflow for systematic model development and evaluation that also considers uncertainty propagation. A recent publication by Braakman et al. (2021) describes a workflow for the systematic evaluation of quantitative systems pharmacology models – a similar workflow for integrated process models that is supported by both industry and regulators alike would be invaluable.

**Regulatory Guidance:** While there have been an increasing number of literature contributions from industrial research groups on modelling applications, there is not a clear consensus or guidance to what extent these approaches will be accepted by global regulatory agencies. While the ICH quality guidelines provide some guidance on model development, the level of detail that should be included in a marketing application on model usage to define a control strategy is less clear. Consequently, this leaves industrial practitioners with uncertainty on the likelihood of success; and therefore, resorting to an empirical approach based on Design of Experiments (DoE) and statistical analysis, which is a more widely applied aspect of the Quality by Design (QbD) framework, is thus the safer strategy. This highlights the need for better engagement between regulators and industry on model development and usage.

The increasing number of modelling applications by pharmaceutical companies is certainly promising. We believe that overcoming these barriers are critical to realizing the full benefits of end-to-end modelling for accelerated pharmaceutical development. Continued demonstration of successful case studies of model applications on active projects are key to move towards the systematic application of modelling by first intent for pharmaceutical process development.

### Acknowledgments

This study was funded by a Digital Design capability project at GlaxoSmithKine (GSK).

### References

- G. Bano, R.M. Dhege, S. Diab, D.J. Goodwin, L. Gorringer, M. Ahmed, R. Elkes, S. Zomer, 2022a. Streamlining the development of an industrial dry granulation process for an immediate release tablet with systems modelling. *Chem. Eng. Res. Des.*, 178, 421–437.
- G. Bano, M. Aroniada, Y. Vueva, 2022b. A model-based approach to predict the flowability of directly compressed pharmaceutical blends. *Comput.-Aided Chem. Eng.*, 51, 31–6.
- F. Boukouvala, A. Chaudhury, S. Sen, R. Zhou, L. Mioduszewski, M.G. Ierapetritou, R. Ramachandran, 2013. Computer-aided flowsheet simulation of a pharmaceutical tablet manufacturing process incorporating wet granulation. *J. Pharm. Innov.*, 8, 11–27.
- S. Braakman, P. Pathamanathan, H. Moore, 2022. Evaluation framework for systems models. *CPT Pharmacometrics Syst. Pharmacol.*, 11, 264–289.
- G. Capellades, J.O. Bonsu, A.S. Myerson, 2022. Impurity incorporation in solution crystallization: diagnosis, prevention, and control. *CrystEngComm*, 24, 1989–2001.
- F. Destro, I. Hur, V. Wang, M. Abdi, X. Feng, E. Wood, S. Coleman, P. Firth, A. Barton, M. Barolo, Z.K. Nagy, 2021. Mathematical modeling and digital design of an intensified filtration-washing-drying unit for pharmaceutical continuous manufacturing. *Chem. Eng. Sci.*, 244, 116803.
- S. Diab, C. Christodoulou, G. Taylor, P. Rushworth, 2022a. Mathematical modeling and optimization to inform impurity control in an industrial active pharmaceutical ingredient manufacturing process. *Org. Process Res. Dev.*, 26, 10, 2864–2881.
- S. Diab, G. Bano, C. Christodoulou, N. Hodnett, A. Benedetti, M. Andersson, S. Zomer, 2022b. Application of a system model for continuous manufacturing of an active pharmaceutical ingredient in an industrial environment. *J. Pharm. Innov.* DOI: [10.1007/s12247-021-09609-7](https://doi.org/10.1007/s12247-021-09609-7)
- S. García-Muñoz, A. Butterbaugh, I. Leavesley, L.F. Manley, D. Slade, S. Bermingham, 2018. A flowsheet model for development of a continuous process for tablets. *AIChE J.*, 64, 511–25.
- Y. Guo, 2009. Impact of solid-state characteristics to the physical stability of drug substance and drug product. *Handbook of Stability Testing in Pharmaceutical Development*, 241–246.
- International Council on Harmonization, 2000. ICH Topic Q6A Specifications: Test procedures and acceptance criteria for new drug substances and new drug products: chemical substances.
- H.G. Jolliffe, E. Ojo, C. Mendez, I. Houson, R. Elkes, G. Reynolds, A. Kong, E. Meehan, F.A. Becker, P.M. Piccione, S. Verma, A. Singarju, G. Halbert, J. Robertson, 2022. Linked experimental and modelling approaches for tablet property predictions. *Int. J. Pharm.*, 626, 122116.
- M. Kuentz, H. Leuenberger, 1998. Modified Young's modulus of microcrystalline cellulose tablets and the directed continuum percolation model. *Pharm. Dev. Technol.*, 3, 13–19.
- D. Markl, S. Yassin, D.I. Wilson, D.J. Goodwin, A. Anderson, J.A., Zeitler, 2017. Mathematical modelling of liquid transport in swelling immediate release tablets. *Int. J. Pharm.*, 526, 1–10.
- F.C.C. Montes, K. Gernaey, G. Sin, 2018. Dynamic plantwide modeling, uncertainty, and sensitivity analysis of a pharmaceutical synthesis: ibuprofen case study. *Ind. Eng. Chem. Res.* 57, 30, 10026–37.
- B. Nagy, B. Szilágyi, A. Domokos, B. Vészi, K. Tacsí, Z. Rapi, H. Pataki, G. Marosi, Z.K. Nagy, Z.K. Nagy, 2021. Dynamic flowsheet model development and digital design of continuous pharmaceutical manufacturing with dissolution modeling. *Chem. Eng. J.*, 419, 129947.
- J. Nassar, B. Williams, C. Davies, K. Lief, R. Elkes, 2021. Lubrication empirical model to predict tensile strength of directly compressed powder blends. *Int. J. Pharm.*, 592, 119980.
- N.A. Peppas, P. Colombo, 1989. Development of disintegration forces during water penetration in porous pharmaceutical systems, *J. Control. Release*, 10, 245–250.
- G. Taylor, H. Yao, B. Williams, R.E. Yule, N. Mitchell, 2021. Kinetic modelling and design space exploration of impurity inclusion in an API crystallization. *International Symposium on Industrial Crystallisation*.
- USP <711> Dissolution. The United States Pharmacopeial Convention. 2011.

# Direct Compression Flowsheet Modelling to Assess Tablet Manufacturability and *In Vitro* Dissolution

Magdalini Aroniada<sup>a,\*</sup>, Charalampos Christodoulou<sup>b</sup>, Houda Khaled<sup>a</sup>

<sup>a</sup>GlaxoSmithKline, Park Road, Ware SG12 0DP, United Kingdom

<sup>b</sup>GlaxoSmithKline, Gunnels Wood Road, Stevenage SG1 2NY, United Kingdom

\*Corresponding author email: [magdalini.x.aroniada@gsk.com](mailto:magdalini.x.aroniada@gsk.com)

## Abstract

Direct Compression (DC) is a desirable manufacturing route for pharmaceutical tablets, having fewer unit operations than granulation processes. Assessment of the impact of process operating set points and lot-to-lot raw material property variation on manufacturability and product quality attributes is essential for the development of a robust process. In this study, we describe the development of a DC process model to assess the influence of process operating parameters and material properties on *in vitro* dissolution behavior. We summarize the model equations, followed by its calibration and validation vs. industrial data. We then use Global Sensitivity Analysis to investigate the effects of process operating parameters and material properties on *in vitro* dissolution.

**Keywords:** Direct compression; Flowsheet model; Pharmaceutical manufacturing, Global Sensitivity Analysis.

## 1. Introduction

Detailed understanding of the impact of critical material attributes (CMAs) and critical process parameters (CPPs) that belong to different unit operations of the manufacturing process on tablet intermediate and/or critical quality attributes (iCQAs and CQAs, respectively) requires quantifying how their combinations and interactions of CMAs and CPPs affect product performance. This detailed understanding can be obtained by developing an integrated (flowsheet) model that links together the different unit operations of the process, often termed a “system” model in the pharmaceutical industry. Such a model is based on mechanistic and/or empirical sub-models, each simulating different unit operations of the process and can be used to enhance process understanding.

In this paper, we have developed a system model for an industrial direct compression (DC) line. We analyze the system by performing a variance-based global sensitivity analysis to identify and rank the critical process parameters, gain process understanding and support process development. The paper is structured as follows: (a) materials and methods used, (b) models used to develop the system model and validation, (c) Global Sensitivity Analysis (GSA) and input to the control strategy definition.

## 2. Materials and methods

All data presented and used in this study were performed with formulations containing a GlaxoSmithKline (GSK) proprietary active pharmaceutical ingredient (API), microcrystalline cellulose (MCC), croscarmellose sodium and lubricant. For the formulations used, another solid form, “API Form X” can be present. The levels of this

form depend on the incoming drug substance material and the tablet press operation. The software used for all modelling activities was gPROMS®.

### 3. Models

This section describes the approach used to develop the system model for the DC process, how unit operations and product performance were modelled individually and how the sub-models were connected in the validated system model.

#### 3.1. Blender and Tablet press

The purpose of the semi-empirical blender model is to calculate the extent of lubrication in the blender ( $K_b$ ). The extent of lubrication, introduced by Kushner et al. (2012), is a metric to establish the direct correlation between mixing history and the amount of lubrication observed in the system. We considered that the extent of lubrication in the feed frame ( $K_f$ ) is also significant (Ketterhagen, 2017).

The tablet press model is an empirical model which combines the Kawakita and Lüdde (1971) and the Extended-Kushner (Nassar et al., 2021) models for the calculation of the tablet porosity (i.e., solid fraction) and tensile strength, respectively, as a function of extent of lubrication. These models require calibration with data for the relevant process equipment. The tablet tensile strength can be calculated from Eq. 1:

$$TS/TS_0 \equiv (1 - \beta) + \beta \exp(-\gamma K) \quad (1)$$

where  $K$  = total extent of lubrication,  $\gamma$  = tensile strength decay rate,  $\beta$  = material lubrication extent coefficient,  $TS_0$  = tensile strength of the unlubricated blend. The total extent of lubrication,  $K = K_b + K_f$  and  $K_f = v_{tip}\tau$  where  $v_{tip}$  is the tip speed of the feed frame paddle wheel, and  $\tau$  is the mean particle residence time in the feed frame. The tablet tensile strength is then related to the hardness (the practically measured quantity) as per Pitt and Heasley (2013). The generation of API Form X in the tablet press is captured by a data-driven model whose inputs are API  $d_{90}$ , level of API form X and tablet hardness.

#### 3.2. Tablet disintegration

The tablet disintegration model includes two mechanisms: (1) liquid penetration kinetics and (2) swelling of the tablet (Markl et al., 2017). According to Desai et al. (2016), the break-up of inter-particle bonds can also be caused by strain recovery. Strain recovery is the forcing/transitioning of macromolecules (polymers/disintegrants) into a metastable configuration, e.g., due to spontaneous crystallisation during compaction. In contrast to the omni-directional enlargement of the swollen particles, strain recovery causes a unidirectional increase in the size of the particles. All the mentioned mechanisms of disintegration are interlinked and influenced strongly by the microstructure of the tablet. In this study, since water absorption leads to the particle expansion (swelling) which in turn causes the breakup of the tablet structure, we assume that the penetration rate is the

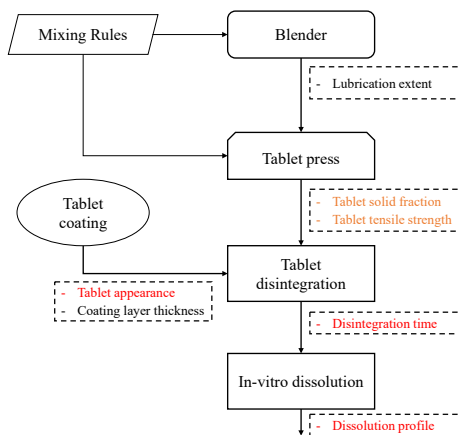


Figure 1. System model flow diagram (highlighted with orange: iCQAs, and red: CQAs).

disintegration rate-limiting step (Markl et al., 2017). The fluid flow in a tablet is described by the volume averaged velocity vector  $\mathbf{u}_p$ . This can be viewed as the volumetric flow rate per unit area. Following Bear (1988), to calculate  $\mathbf{u}_p$ , one can solve the continuity and momentum (Darcy's equation) balance equations:

$$\partial_x \mathbf{u}_p = 0 ; \mathbf{u}_p = -(\mathcal{K}/\mu_l) \partial_x p_p - \rho_l \mathbf{g} \quad (2)$$

where  $p_p$  = pressure inside the pores,  $\mathcal{K}$  = porous medium permeability,  $\mathbf{g}$  = gravitational acceleration,  $\rho_l$  and  $\mu_l$  = density and viscosity of the penetrating liquid, respectively. In Eq. 2,  $\mathcal{K}$  can be found from the Carman-Kozeny permeability equation,  $\mathcal{K} = d_h^2 \epsilon / (S_p \mathcal{T}_{or}^2)$  where  $d_h$  = pore hydraulic diameter. To account for the fact that  $\mathcal{K}$  varies within the same tablet, we introduced an empirical correction factor,  $\Lambda$ , which is fitted to tablet disintegration data. The corrected permeability is given by  $\mathcal{K}^* \equiv \Lambda \mathcal{K}$ .

To calculate average volume velocity vector, one needs to estimate the average porosity of the tablet. According to Markl et al. (2017), the latter is given by  $\epsilon = \epsilon_0 (R_c/R_{c0})^2$  where  $R_c$  and  $R_{c0}$  = wet and dry tablet region pore radius. As the tablet swells, it increases in size and this is causing the capillary radius to decrease, approximated as follows:

$$R_c = R_{c0} - \left(\frac{d_p}{2}\right) \left[ \left(\frac{\delta_0 + \lambda t}{\delta_0}\right)^{\frac{1}{3}} - 1 \right] \quad (3)$$

where  $\delta_0$  = initial dry pore length, and  $\lambda$  = swelling rate which can be calculated following Peppas and Colombo (1989). To calculate the wetting front inside the pores,  $P_d$ , the wetting front profile can be defined as  $F_p(r, z, t) \equiv P_d(r, t) - z$  (Leal, 2007). The substantial derivative of  $F_p$  is zero and this allows us to write:

$$\frac{dF_p}{dt} = 0 \Leftrightarrow \partial_t F_p + \mathbf{u}_p \partial_x F_p = 0 \Rightarrow \partial_t P_d + u_r|_{h_p} \partial_r P_d - u_z|_{h_p} = 0 \quad (4)$$

As stated above, the penetration rate is the disintegration rate-limiting step, thus, we calculate the rate of disintegration using the following expression:

$$\frac{dM_t}{dt} = [2\rho_s(1 - \epsilon)\Omega(t)] \frac{dP_d}{dt} \quad (5)$$

where  $M_t$  = mass of the tablet at time  $t$ ,  $\Omega$  = tablet cross-sectional surface area, and  $\rho_s$  = solid true density.

### 3.3. In-vitro dissolution

For the *in-vitro* dissolution model, we have used the population balance model by Wilson et al. (2012) to simulate the change in the particle size distribution during dissolution. We calculate the rate of particle release from the mass concentration and the released particles' number size distribution. The extent of dissolution is defined as the percentage Label Content (%LC) in solution, calculated as:

$$\%LC = 100 C_b^{API} V_m / x_a \rho_t V_{t0} \quad (6)$$

where  $C_b$  = mass concentration of the API in the bulk liquid,  $V_m$  = USP <711> (2011) dissolution test vessel liquid volume,  $V_t$  = volume of the tablet at  $t = 0$  and  $\rho_t$  = tablet density which is estimated from  $\rho_t \approx \rho_s(1 - \epsilon)$ . When %LC = 100% the dissolution process is complete.



### 3.4. Coating

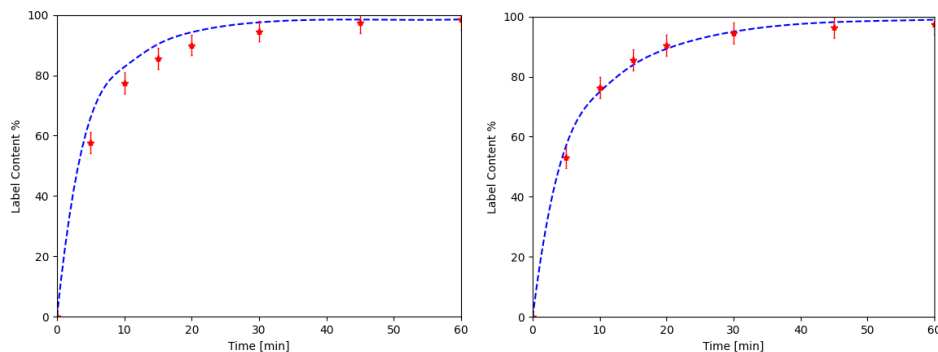
To capture the influence of the coating layer, we used the “erosion rate” approach of Wilson et al. (2012) and Bano et al. (2022a). The main assumption of the model is that the liquid does not penetrate the coating layer, but the latter erodes with a constant erosion rate  $\hat{\epsilon}$ . The coating volume  $V_c$  on the surface of the tablet with surface area  $= A_t$  at any time is given by  $V_c(t) = [H_c(t) - \hat{\epsilon}t]A_t$  where  $H_c$  = coating thickness.

### 3.5. Mixing rules

We have also used a model-based approach to predict the fill bulk density of the blends from individual components. The model is that by Bano et al. (2022b) and used in the blender and tablet press models.

### 3.6. Calibration and validation of the models

All the above models were calibrated using experimental data from industrial DC campaigns. The system model was then validated vs. experimental data are shown in Fig. 2. The predictions and experimental data are in good agreement.



**Figure 2.** Model validation (left = nominal blend; right = different blender fill level and compaction force). Red points + error bars = experimental data; Blue curve = model predictions.

## 4. Results

Using the validated model, we performed a Global Sensitivity Analysis (GSA) to assess the impacts of different process parameters and material properties on tablet *in-vitro* dissolution. The input factors are: API particle size, form X level in API, moisture content, compaction force, MCC bulk density, lubrication time, feed frame rotational speed, tablet production rate, dissolution medium volume, blender rotational speed, paddle speed, MCC particle size, batch size and medium temperature. The output response is %LC at 30 min based on the USP <711> (2011) dissolution test.

### 4.1. Global Sensitivity Analysis

The first step on GSA is to perform the Elementary Effects (EE) method (Morris, 1991). This identified that the factors that have the biggest effect on the response of interest are the API particle size, level of API form X, the moisture content, the compaction force and the MCC bulk density. We then used the Sobol (1993) variance-based method.

There are two types of Sobol sensitivity indices: the first-order effect index ( $S_i$ ), that represents the main effect contribution of each input to the variance of the output, and the total effect index ( $S_{Ti}$ ), which measures the effect of both first- and higher-order effects

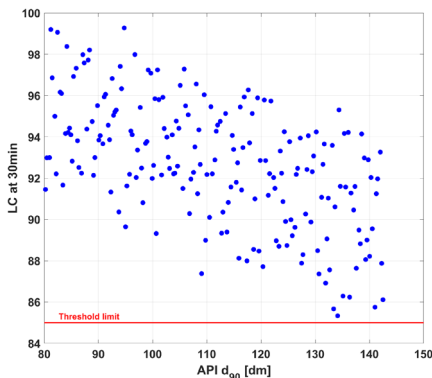
(interactions). For this study, the total effect indices for the response that the most important factors affecting the Label Content (LC)% are the API Form X and the API particle size ( $d_{90}$ ). Moreover, the interactions between MCC density, compaction force and water content (or moisture content) are expected to be important (Table 1).

**4.2. Assessment of Process Operating Space**

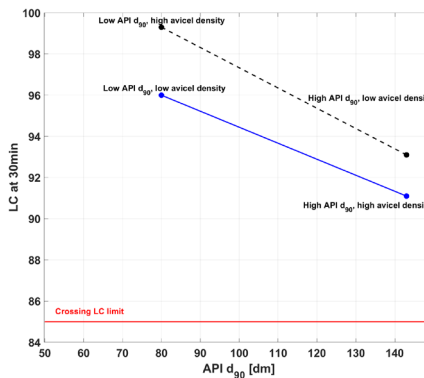
In this study we have used the validated model to run simulated experiments to assess the factor criticality (influential factors found in section 4.1). With this analysis we can capture scenarios in the considered process operating space where the tablet dissolution approaches or passes a threshold limit. Using the model, we can observe overall trends across the considered operating space by visualising simulated experiments as a function of different input factors. For example, in Fig. 3 we show a scatter plot that allows us to see the trend when increasing the API particle size. We also observe that some simulations are very close to the threshold limit ( $\%LC < 85\%$  at 30 min). For these cases close to the threshold, the level of form X in API and water content are high. This highlights how the model could help to highlight points in the operating space which may pose a product quality issue, which may have been challenging from experiments alone.

**Table 1.** Variance-based sensitivity analysis. Factor ranking; response: LC% @30min.

Ranking	Input Factors	$S_i$	$S_{Ti}$
1	Form X level [in API]	0.510	0.560
2	API particles X90	0.265	0.414
3	Moisture content	0.034	0.061
4	Compaction force	0.000	0.051
5	MCC density	0.000	0.026



**Figure 3.** Scatter plots of *in-silico* experiments API  $d_{90}$  effect on LC at 30min



**Figure 4.** Factor-factor interaction plots, %LC at 30min

In addition, we performed simulations to visualize factor-factors interactions. In Fig. 4, we performed simulations with high and low values of API particle size vs. high and low MCC densities. In this case, the lines are almost parallel (highlighting that factor-factor interactions are almost negligible) and the limit is not crossed. This factor-factor interaction (API particle size and MCC density) is negligible.

**5. Conclusions**

In this study, we developed a system model for a DC tableting line. The system model was calibrated and validated against data collected from industrial campaigns. At this stage, the validation showed that the model is predictive vs. *in vitro* dissolution data and has a very good prediction fidelity. Later, the model was used to perform *in-silico*

calculations of dissolution profiles under USP <711> conditions. Based on this, we identified the critical process parameters which affect tablet dissolution (main CQA). In this work, it is highlighted that the modelling approach can be used to provide actionable insight into control strategy definition discussions and to help industrial practitioners develop robust pharmaceutical manufacturing processes.

### Acknowledgments

This study was funded by a Digital Design capability project at GlaxoSmithKline (GSK).

### References

- G. Bano, R.M. Dhenge, S. Diab, D.J. Goodwin, L. Gorringer, M. Ahmed, R. Elkes, S. Zomer, 2022a. Streamlining the development of an industrial dry granulation process for an immediate release tablet with systems modelling. *Chem. Eng. Res. Des.*, 178, 421–437.
- G. Bano, M. Aroniada, Y. Vueva, 2022b. A model-based approach to predict the flowability of directly compressed pharmaceutical blends. *Comput.-Aided Chem. Eng.*, 51, 31–6.
- J. Bear, 1988. *Dynamics of fluids in porous media*. Courier Corporation.
- P.M. Desai, C.V. Liew, P.W.S. Heng, 2016. Review of disintegrants and the disintegration phenomena. *J. Pharm. Sci.*, 105, 9, 2545–2555.
- K. Kawakita, K. Lüdde, 1971, Some considerations on powder compression equations, *Powder Technology*, 4(2), pp.61-68
- W. Ketterhagen, A. Aliseda, M. am Ende, A. Berchielli, P. Doshi, B. Freireich, A. Prpich, 2017. Modeling tablet filmcoating processes. In *Predictive Modeling of Pharmaceutical Unit Operations* (pp. 273-316). Woodhead Publishing.
- J. Kushner, 2012. Incorporating Turbula mixers into a blending scale-up model for evaluating the effect of magnesium stearate on tablet tensile strength and bulk specific volume. *Int. J. Pharm.*, 399, 19–30.
- L.G. Leal, 2007. *Advanced transport phenomena: fluid mechanics and convective transport processes*. Cambridge University Press.
- D. Markl, S. Yassin, D.I. Wilson, D.J. Goodwin, A. Anderson, J.A., Zeitler, 2017. Mathematical modelling of liquid transport in swelling immediate release tablets. *Int. J. Pharm.*, 526, 1–10.
- M.D. Morris, 1991. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33, 2, 161–174.
- J. Nassar, B. Williams, C. Davies, K. Lief, R. Elkes, 2021. Lubrication empirical model to predict tensile strength of directly compressed powder blends. *Int. J. Pharm.*, 592, 119980.
- N.A. Peppas, P. Colombo, 1989. Development of disintegration forces during water penetration in porous pharmaceutical systems, *J. Control. Release*, 10, 245–250.
- K.G. Pitt, M.G. Heasley, 2013. Determination of the tensile strength of elongated tablets. *Powder Technology*, 238, pp.169-175.
- I. Sobol', 1993, Sensitivity analysis for non-linear mathematical models. *Mathematical Modeling & Computational Experiment (Engl. Transl.)*, 1, pp. 407–414.
- USP <711> Dissolution. The United States Pharmacopeial Convention. 2011.
- D. Wilson, S. Wren, G. Reynolds, 2012. Linking dissolution to disintegration in immediate release tablets using image analysis and a population balance modelling approach. *Pharm. Res.*, 29, 198–208.

## Enhanced hot-liquid water pretreatment of biomass with recovery and valorization of side products

Tamara Janković,<sup>a</sup> Adrie J. J. Straathof,<sup>a</sup> Anton A. Kiss<sup>a,b</sup>

<sup>a</sup> *Department of Biotechnology, Delft University of Technology, van der Maasweg 9, 2629 HZ Delft, The Netherlands*

<sup>b</sup> *Department of Chemical Engineering, Delft University of Technology, van der Maasweg 9, 2629 HZ Delft, The Netherlands*  
*A.A.Kiss@tudelft.nl*

### Abstract

Lignocellulosic biomass potentially represents a great feedstock for biofuel production, but its' pretreatment needs to be enhanced in order to make biorefineries competitive with fossil fuel based alternatives. One way to make biorefineries more economically viable is to recover and valorize all generated by-products during the biomass pretreatment step. The main goal of this original research is to design an optimal process for recovering valuable by-products after hot-liquid water pretreatment of poplar biomass. Rigorous models for all operations included in the recovery process are developed using Aspen Plus as a CAPE tool. An optimal downstream processing sequence, consisting of multiple distillation steps, is designed to recover several valuable components, such as acetic acid, formic acid, furfural and 5-hydroxymethyl-furfural (HMF).

**Keywords:** biorefineries, lignocellulosic biomass, valorization, distillation

### 1. Introduction

The increasing concerns on environmental pollution and reduction of available fossil carbon sources lead to a need for a transition towards biofuels and bio-based chemicals from renewable sources. Technologies for the production of the first generation biofuels are already well established. On the other hand, their implementation directly competes with global food production as the first generation biomass can also be used for food supply. The potential solution is usage of lignocellulosic biomass as a source for biorefineries. As the second generation feedstock, it is a non-food crop that can be easily grown on land not suitable for food production. Therefore, the production of biofuel and biochemicals from lignocellulosic biomass does not threaten the food production process (Chiaramonti et al., 2013).

However, technologies and markets for conversion of the second generation feedstocks are still not well established. In order to become competitive with fossil-based production, biorefineries need to be more economically viable (Kiss & Infante Ferreira, 2016). Due to the complexity of the lignocellulosic biomass structure, the pretreatment step is needed to facilitate further conversion of lignocellulosic components to fermentable sugars. Since this process step significantly contributes to the total costs (up to 30-40%) (Ramaswamy et al., 2013), its' enhancement can drastically increase the viability of biorefineries. One way to improve the competitiveness of biorefineries is to recover and valorize valuable by-products generated during the biomass pretreatment step, which has not been the subject of many studies so far. This research gap is covered in this study.

## 2. Problem statement

### 2.1. Enhanced biomass pretreatment step

A common problem with different pretreatment methods is usage of additional chemicals. These chemicals might facilitate the pretreatment step but complicate downstream processing, as well as valorization of generated by-products. In that context, hot-liquid water pretreatment potentially presents an eco-friendly solution due to relatively mild process conditions, lack of additional chemicals, low process costs and a possibility to recover and valorize formed by-products (Chen et al., 2022).

### 2.2. Recovery of valuable by-products from the pretreatment step

After the biomass pretreatment step by hot-liquid water, the solid fraction is sent to enzymatic hydrolysis while the potential value of the liquid fraction is usually neglected. This research focuses on downstream processing of the liquid fraction after hot-liquid water pretreatment of poplar biomass. The main goal is to fully recover valuable by-products (e.g. acetic acid, formic acid, furfural, HMF). These components are not only economically valuable, but they can also inhibit fermentation reactions. Therefore, their removal will result in higher bioethanol yields. A crucial problem regarding by-product recovery is a very dilute solution after the pretreatment step (Kiss et al., 2016). Consequently, improved fluid separation processes are needed in order to recover by-products in a high-purity form that can be valorized on the market.

## 3. Methodology/Approach

The main objective of this research is to develop an optimal way to recover by-products from the liquid obtained after hot-liquid water pretreatment. The composition of the liquid, which is the feed stream for the designed downstream processing sequence, is obtained from literature (Jiminez-Gutierrez, 2023) and is presented in Table 1. This stream is diluted solution of different products generated by pretreatment of poplar biomass. Acetic acid is formed by hydrolysis of acetyl groups in the hemicellulose chain. C-5 sugars (mainly xylose and arabinose) from hemicellulose can degrade to furfural, while C-6 sugars (mainly glucose, mannose and galactose) from both cellulose and hemicellulose can degrade to HMF. Formic acid can be generated from both furfural and HMF, while levulinic acid can be formed by degradation of HMF. Rigorous Aspen Plus simulations are developed for every process operation. Components are defined in Aspen Plus according to the literature (Aspen Technology, 2021). Due to presence of different carboxylic acids in the initial mixture, NRTL property method coupled with Hayden-O'Connell model with correction for vapor phase association of carboxylic acids is used (Aspen Technology, 2020). Properties of the components and possible formation of azeotropes (Gmehling et al., 2004) are given in Table 2.

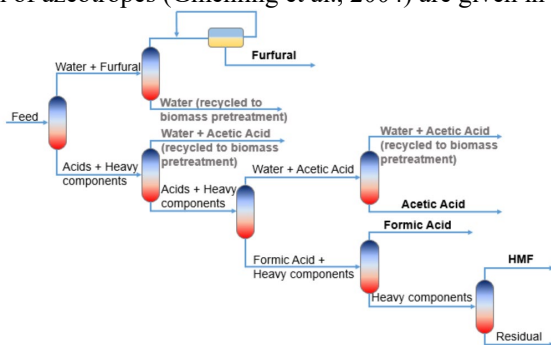


Figure 1. Simplified process flow scheme of the by-products recovery process

Table 1. Feed stream

Component	%wt
Water	96.568
Acetic acid	1.747
Formic acid	0.427
Furfural	0.297
HMF	0.128
Levulinic acid	0.005
Glucose and xylose	0.097
Other heavy components	0.730

Table 2. Components and azeotropes (1 atm)

Pure components		Azeotropes		
Component	T <sub>b</sub> (°C)	Component	Molar fraction	T (°C) / type
Water	100.02	Water	0.9066	97.79
Formic acid	100.55	Furfural	0.0934	/heterogeneous
Acetic acid	118.01	Water	0.3607	106.50
Furfural	161.35	Formic acid	0.4998	/homogenous
Levulinic acid	256.98	Acetic acid	0.1395	
HMF	281.85	Water	0.4566	106.81
Glucose	343.85	Formic acid	0.5434	/homogenous

Different methods for recovery of present components from diluted solutions are proposed in the literature (Li et al., 2016), but recovery yields are usually insufficient and the usage of additional chemicals complicates downstream processing. Research is often done on lab-scale and scale-up is not considered. To avoid adding extra chemicals, distillation is chosen and component separation is proposed in a defined order (Figure 1). Since the feed is very diluted, the first step is to concentrate it by separating as much water as possible. Due to the formation of light heterogeneous azeotrope (Figure 2), furfural is separated together with water. It can further be recovered by simple phase separation in a decanter. Recovery of acetic and formic acids from an aqueous solution is complex due to existence of four distillation areas and two azeotropes (Figure 3). As acetic acid concentration is the highest of all by-products in the feed, its' recovery is crucial. Therefore, most of acetic acid and water are separated from formic acid and heavier components (HMF, levulinic acid, dissolved sugars and solubilized lignin components) at temperature lower than needed for ternary azeotrope formation. Further recovery of acetic acid is complex due to the tangent pinch at high water concentrations. Formic acid, plus the remaining water and acetic acid, can be separated from HMF and other high boiler components by distillation. The obtained mixture can be further treated in order to obtain a high-purity formic acid product. HMF can be recovered in a high-purity form, while other components can be sent to the fermentation step.

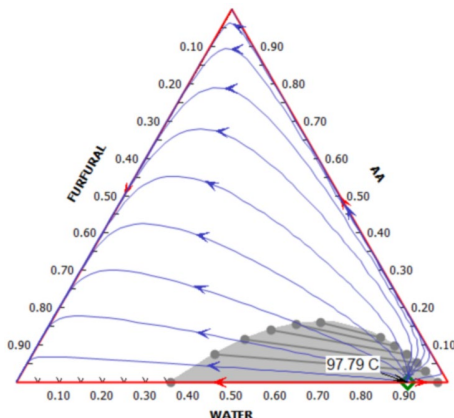


Figure 2. Ternary diagram for system water (W) - acetic acid (AA) - furfural (FURFURAL), mole basis

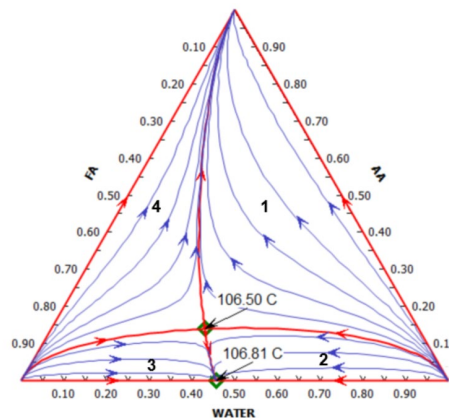


Figure 3. Residue Curve Map for system water (W) - formic acid (FA) - acetic acid (AA), mole basis

## 4. Results and discussion

### 4.1. By products recovery process design

The proposed optimal process design for recovery of valuable by-products after biomass pretreatment by hot-liquid water recovers acetic acid, furfural, HMF and formic acid. This process (Figure 4) consists of seven distillation columns and one decanter. Light azeotrope water-furfural is separated in column C-1, after which furfural is purified by additional distillation in column C-2 and phase separation in decanter DEC-1. This results in 100% furfural recovery, while purity of product stream is 97.9%wt. After separation of the water – furfural azeotrope, the solution is still very dilute and is further concentrated in column C-3. Some acetic acid is separated together with water, resulting in top product from this column with 0.5%wt of acetic acid (13.8% from acetic acid present in the feed stream). Since addition of acetic acid in pretreatment step improves digestibility of biomass by decreasing pH to mild-low values (Jiminez-Gutierrez, 2023), this stream can be recycled back to the biomass pretreatment step. Bottom product from column C-3 is further distilled in column C-4. Acetic acid, together with most of the water, is separated as a top product, while formic acid, levulinic acid, HMF and other heavy components are obtained in the bottom stream. Acetic acid is purified in column C-5 and obtained in a high purity form (99.8%wt), with total recovery of 78.8 %. HMF and heavier components are separated from formic acid, levulinic acid and remains of acetic acid and water in column C-6. HMF is further purified in column C-7 and obtained as high-purity (100%wt) top product with total recovery of 99.6 %, while heavy components that are obtained at the bottom can be recycled back to the fermentation step. Top product from column C-6 is stream with 73.6%wt of formic acid. Since composition of this stream is in distillation region 3 in Figure 3, high recovery of pure formic acid cannot be obtained by simple distillation due to the thermodynamic limitations of the water – formic acid azeotrope formation. Therefore, reactive distillation with methanol could be applied to completely recover formic acid (Painer et al., 2015). Moreover, additional purification of formic acid might not be necessary as there are applications of 70% formic acid solution (Singhal et al., 2015). Condition and composition of feed stream and product streams are given in Table 3.

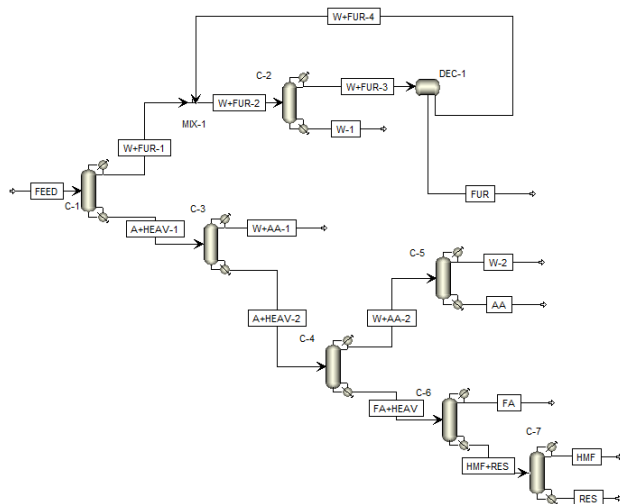


Figure 4. Flowsheet of by-products recovery process

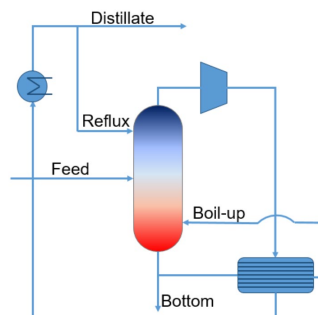


Figure 5. MVR heat pump design

Table 3. Condition and composition of feed stream and product streams for by-products recovery process

Stream	FEED	FUR	AA	FA	HMF
Temperature [°C]	99.70	27.36	26.00	34.82	26.00
Pressure [bar]	1	1	1	0.05	0.01
Mass Flow [kg/h]	110000	333.89	1517.00	635.11	140.70
Mass Fraction					
Water	0.9657	0.0206	0.0003	0.2187	0
Acetic Acid	0.0175	0	0.9982	0.0360	0
Formic Acid	0.0043	0	0.0015	0.7358	0
Furfural	0.0030	0.9794	0	0	0
HMF	0.0013	0	0	0.0003	1
Levulinic acid	0.0001	0	0	0.0091	0
Glucose, xylose and other heavy components	0.0083	0	0	0	0

Table 4. Design parameters of distillation columns in by-products recovery process

Column	C-1	C-2	C-3	C-4	C-5	C-6	C-7
No. Stages	52	30	60	48	66	20	20
Feed Stage	35	12	40	36	33	10	13
Pressure [bar]	1	1	1	1	1	0.05	0.01

#### 4.2. Distillation columns design

Design parameters (number of stages, feed stage and column pressure) for each of the columns are given in Table 4. As the feed stream is very dilute, large amounts of water need to be evaporated in C-1, C-3, C-4 and C-5. Hence these columns have larger number of stages and higher reboiler duties than the other columns. However, the top vapor from these columns can be used to evaporate bottom liquid by implementing mechanical vapor recompression as heat pump. Columns C-6 and C-7 operate at lower pressure to reduce temperatures and avoid thermal degradation.

#### 4.3. Energy analysis

Mechanical vapor recompression (MVR) and heat integration (HI) are implemented to enhance base process design. MVR is a heat pump design that can be applied to decrease energy use when separating close boiling components (Kiss & Infante Ferreira, 2016). This implies compressing top vapor and using it to evaporate bottom liquid (Figure 5). In the designed optimal recovery process, MVR is applied to the columns C-1, C-3, C-4 and C-5. Comparison of key performance indicators for processes before and after enhancing process design is given in Table 5. The inefficiencies in power generation are taken into account for calculation of the total energy requirements through an electric to thermal conversion factor of 2.5. Total investment cost (TIC), total operating cost (TOC) and total annual cost (TAC) are calculated according to a published procedure (Kiss, 2013), with a payback time of 10 years. A Marshall and Swift cost index of 1638.2 (end of 2018) is used (Mantingh & Kiss, 2021). Emissions of CO<sub>2</sub> are calculated using the equations from the literature (Mantingh & Kiss, 2021).

Table 5. Key performance indicators for processes with and without MVR

Performance indicator	Base case	Enhanced case	Savings (%)
Thermal energy requirements (kW <sub>th</sub> h/kg <sub>product</sub> )	75.13	0.18	99.8
Electrical energy requirements (kW <sub>e</sub> h/kg <sub>product</sub> )	n/a	3.53	n/a
Primary energy requirements (kW <sub>th</sub> h/kg <sub>product</sub> )	75.13	8.99	88.0
Total Investment Cost (k\$)	24923	48284	n/a
Total Operating Cost (\$/kg <sub>product</sub> )	2.29	0.33	85.5
Total Annual Cost (\$/kg <sub>product</sub> )	2.41	0.56	76.6
CO <sub>2</sub> emissions (kg <sub>CO2</sub> /kg <sub>product</sub> )	10.84	1.88/ 0.03*	82.7/ 99.7*

\*electricity from non-renewable sources / green electricity



## 5. Conclusions

This original work proposes an optimal process design for recovering by-products after biomass pretreatment by hot-liquid water. Improved fluid separation processes are used to obtain valuable products from very dilute feed stream (>96%wt water). Several biomass pretreatment products are effectively recovered, including: acetic acid, formic acid, furfural and HMF, with recoveries of around 78.8%, 99.5%, 100% and 99.6% respectively. Acetic acid, furfural and HMF are obtained in high-purity forms (99.8%wt, 97.9%wt and 100%wt respectively), while formic acid is recovered as a 73.6wt% solution. MVR and HI are efficiently implemented to further decrease the energy usage. Due to the use of additional compressors, TIC after implementation of MVR and HI is about 93.7% higher, but TOC and TAC are decreased for 85.5% and 76.6% (from 2.29 to 0.33 \$/kg<sub>product</sub> and from 2.41 to 0.56 \$/kg<sub>product</sub> respectively). Moreover, usage of MVR resulted in 88.0% energy savings and significantly lower CO<sub>2</sub> emission leading to a more environmentally friendly process.

## References

- Aspen Technology, *Aspen Physical Property System* (2020)
- Aspen Technology, *Characterization of Lignocellulosic Biomass with Specification in Aspen Properties*. (2021). <https://knowledgecenter.aspentech.com>
- Chen, W. H., Nižetić, S., Sirohi, R., Huang, Z., Luque, R., M.Papadopoulos, A., Sakthivel, R., Phuong Nguyen, X., & Tuan Hoang, A. (2022). Liquid hot water as sustainable biomass pretreatment technique for bioenergy production: A review. *Bioresource Technology*, 344(10). <https://doi.org/10.1016/j.biortech.2021.126207>
- Chiaromonti, D., Giovannini, A., Janssen, R., & Mergner, R. (2013). *Lignocellulosic ethanol production plant by Biochemtex in Italy Lignocellulosic Ethanol Process and Demonstration A Handbook Part I*. WIP Renewable Energies. [www.biolyfe.eu](http://www.biolyfe.eu)
- Gmehling, J., Menke, J., Krafczyk, J., & Fischer, K. (2004). *Azeotropic Data* (Second Add). Wiley-VCH.
- Jimenez-Gutierrez, J. M. (2023). *Pretreatment of lignocellulosic biomass for acetic acid co-valorization*. Delft University of Technology.
- Kiss, A. A. (2013). *Advanced Distillation Technologies: Design, Control and Applications*. Wiley. <https://doi.org/10.1002/9781118543702>
- Kiss, A. A., & Infante Ferreira, C. A. (2016). Heat Pumps in Chemical Process Industry. In *Heat Pumps in Chemical Process Industry*. CRS Press. <https://doi.org/10.1201/9781315371030>
- Kiss, A. A., Lange, J. P., Schuur, B., Brillman, D. W. F., van der Ham, A. G. J., & Kersten, S. R. A. (2016). Separation technology—Making a difference in biorefineries. *Biomass and Bioenergy*, 95(5), 296–309. <https://doi.org/10.1016/j.biombioe.2016.05.021>
- Li, Q.-Z., Feng, X.-J., Zhang, H.-B., Liu, H.-Z., Xian, M., Sun, C., Wang, J.-M., & Jiang, X.-L. (2016). Recovery Processes of Organic Acids from Fermentation Broths in the Biomass-Based Industry. *Journal of Microbiology and Biotechnology*, 26(1), 1–8. <https://doi.org/10.4014/jmb.1505.05049>
- Mantingh, J., & Kiss, A. A. (2021). Enhanced process for energy efficient extraction of 1,3-butadiene from a crude C4 cut. *Separation and Purification Technology*, 267(3). <https://doi.org/10.1016/j.seppur.2021.118656>
- Painer, D., Lux, S., & Siebenhofer, M. (2015). Recovery of Formic Acid and Acetic Acid from Waste Water Using Reactive Distillation. *Separation Science and Technology*, 50(18), 2930–2936. <https://doi.org/10.1080/01496395.2015.1085407>
- Ramaswamy, S., Huang, H.-J., & Ramao, B. V. (2013). *Separation and Purification Technologies in Biorefineries*. Wiley. <https://doi.org/10.1002/9781118493441>
- Singhal, N., Kumar, M., Kanaujia, P. K., & Virdi, J. S. (2015). MALDI-TOF mass spectrometry: An emerging technology for microbial identification and diagnosis. *Frontiers in Microbiology*, 6(8), 1–16. <https://doi.org/10.3389/fmicb.2015.00791>

# Development of deep learning framework to predict physicochemical properties for Ionic liquids

Sadah Mohammed<sup>a</sup>, Fadwa Eljack<sup>a</sup>, Saad Al-Sobhi<sup>a</sup>, Monzure-Khoda Kazi<sup>b</sup>

<sup>a</sup>*Qatar University, Department of Chemical Engineering, College of Engineering, P.O. Box 2713, Doha, Qatar*

<sup>b</sup>*Texas A&M Energy Institute, Texas A&M University, College Station, TX 77843-3372, United States*

*Fadwa.Eljack@qu.edu.qa*

## Abstract

In this paper, a deep learning-based group contribution approach has been developed to identify the optimum structure for ionic liquids (ILs) and to maximize the CO<sub>2</sub> absorption capacity. The suggested methodology demonstrates the steps required to build two deep-learning-based group contribution models, DNN-GC and ANN-GC, separately for IL viscosity and CO<sub>2</sub> solubility in the programming language ‘Python’ by using the two widely used modules, ‘Scikit-Learn’ and ‘TensorFlow’. The CO<sub>2</sub> solubility of IL candidates may not result in the best CO<sub>2</sub> absorption process performance because many other properties of IL, such as viscosity, may affect process performance. To resolve this issue, the two developed models were merged to increase the model accuracy and to predict CO<sub>2</sub> solubility and IL viscosity at the same operating pressure and temperature. Based on the latter steps, an illustrative post-combustion CO<sub>2</sub> capture case study was performed to assess the applicability of the model to select the best IL candidate. With this method, a trade-off between CO<sub>2</sub> solubility and viscosity of ILs can be explicitly studied.

**Keywords:** Deep learning, Group contribution, Ionic liquids, Physiochemical properties

## 1. Introduction

Recently, ionic liquids (ILs) have been the subject of intensive research due to their tunability to fit the needs of a certain application. Also, these solvents possess several advantages including low energy required for energy regeneration, low vapor pressure, and non-corrosive (Nath & Henni, 2020). Despite these merits, their tunability also makes it difficult to identify the optimum candidate for CO<sub>2</sub> capture from the thousands of potential candidates. The number of novel ILs is on the rise and yet the experimental data on these novel ILs are scarce or not readily available in the open literature. Therefore, a robust and less time-consuming computational method such as predictive models, especially deep learning (DL) algorithms is needed for the optimal design of ILs and to accurately understand the dynamics of their physiochemical properties (Jian et al., 2022). Deep neural networks (DNNs) have a greater depth in terms of network layers and a larger number of network parameters compared to

conventional shallow neural networks (Deng et al., 2020). Deep learning (DL) is the primary machine-learning model for predicting structure properties and is composed of deep neural networks (DNN) and artificial neural networks (ANN) (Schmidhuber, 2015). DL which is a subset of Machine learning (ML) are empirical models that apply multiple structures of a neural network to significantly enhance the computing efficiency leading to an increase in neuronal complexity (Pezhman et al., 2022). Artificial neural networks (ANNs), employ most DL algorithms, consisting of an input layer, an output layer, and one hidden layer (Pezhman et al., 2022). Both the DNN and ANN model receives inputs as artificial features from raw data, where they get correlated between layers until they reach the property prediction at the final layer. However, in DNN the hidden layer comprises a multi-layer to allow the neural network to reach deeper levels and to have more network parameters. Moreover, they are computationally advanced in terms of finding non-linear correlations (Deng et al., 2020). The chief objective of this work is to develop two deep learning-based group contribution (GC) models in python to predict ILs viscosity and their CO<sub>2</sub> absorption capacity at different pressure and temperature based on their cation-anion skeletons and functional groups. Furthermore, the two models are merged to optimize new IL molecule structure based on viscosity, cation-anion skeleton, and cation-substituents to achieve maximum CO<sub>2</sub> solubility in a post-combustion carbon capture case study.

## 2. Methodology

### 2.1 Data collection

In this study, a total of 3466 data points on viscosity and CO<sub>2</sub> solubility for 48 ILs, 1733 data points each, as well as their critical properties ( $T_c$ ,  $P_c$ , and  $\omega$ ) were collected from Ionic liquids Database-ILThermo (v2.0) (*Ionic Liquids Database - ILThermo*, n.d.) at different temperature and pressure. The classifications of the cations in the selected ILs are Imidazolium (41%), Ammonium (15%), Phosphonium (15%), and Pryridinium (29%).

### 2.2 DL-GC model development

GC model is widely used in IL solvent design and predicting their thermophysical properties, and for many various decomposition approaches of IL molecules (Chen et al., 2022). A total of 28 different functional groups were obtained from IL molecular structure consisting of 4 cations, 13 anions, and 11 cation substituents (e.g. organic and inorganic groups) as shown in Table 1. So far, many DL-based non-linear GC models have been suggested for property estimation, and the results show that these non-linear models have higher prediction accuracy compared to linear models. Therefore, here the aim is to incorporate DNN and ANN with the GC model to construct a predictive model for the IL viscosity and CO<sub>2</sub> solubility. DNN was selected to predict IL viscosity due to better performance over ANN during model development. The optimum DL-GC architecture and hyperparameters will be identified through the grid search method in TensorFlow for DNN and ANN to determine the optimum number of neurons in each hidden layer.

Table 1: Functional group considered for ionic liquid design

IL functional group	GC parameters
<b>Cation core</b>	$[\text{Im}_{13}]^+$ , $[\text{Im}_{123}]^+$ , $[\text{Pyr}_{11}]^+$ , $[\text{Pyr}_{13}]^+$ , $[\text{N}]^+$ , $[\text{P}]^+$
<b>Anion core</b>	$[\text{BF}_4]^-$ , $[\text{Cl}]^-$ , $[\text{DCA}]^-$ , $[\text{PF}_6]^-$ , $[\text{SCN}]^-$ , $[\text{C}(\text{CN})_3]^-$ , $[\text{Tf}_2\text{N}]^-$ , $[\text{MeSO}_4]^-$ , $[\text{CH}_2\text{SO}_4]^-$ , $[\text{TfO}]^-$ , $[\text{CF}_3\text{COO}]^-$ , $[\text{HCOO}]^-$ , $[\text{NO}_3]^-$ , $[\text{DCN}]^-$ , $[\text{AC}]^-$ , $[(\text{CH}_2)_2\text{PO}_2]^-$
<b>Cation substituents</b>	$[\text{CH}_3]$ , $[\text{CH}_2]$ , $[\text{OH}]$ , $[\text{aC-CH}_3]$ , $[\text{aN-CH}_3]$ , $[\text{aN-CH}_2]$ , $[\text{cN-CH}_3]$ , $[\text{cN-CH}_2]$ , $[\text{N-CH}_2]$ , $[\text{N-CH}_3]$ , $[\text{P-CH}_2]$

### 2.3 Data pre-processing

As a first step, the collected data need to go through the pre-processing step to remove anomalies within the data such as outliers or missing data values. After that, feature selection was applied to select the independent variables for model inputs. In the DNN-GC model, certain features were dropped to improve the performance of the model. In total, 39 and 46 independent variables including the group contribution parameters (see Table 1) for DNN-GC and ANN-GC models, respectively, were identified to predict both dependent variables, the viscosity and the  $\text{CO}_2$  solubility at different temperatures and pressure, resulting in 39 and 46 input nodes in the input layer and one output node in the output layers in both models. The following step is splitting the acquired dataset into two sets, the training set (90%) and the test set (10%), by using an in-built function in the Scikit-learn python library called `train_test_split` function.

### 2.4 Building and training the DL model

To continue with DL model development, both DNN and ANN were constructed and trained separately. A general structure for DNN-GC consists of multiple hidden layers and each layer comprises several neural elements. ANN-GC model consists of only one hidden layer that contains many neural elements. Both models were built using the Keras-Dense layer available in the TensorFlow package. A commonly used activation function named Rectified linear unit (ReLU) was selected for this study because this function acts as an excellent predictor that activates all neurons at the same time (Zarra et al., 2019). To speed up the optimization process during the training stage, the adaptive moment estimation (Adam) optimizer of TensorFlow was used. The loss function 'mean\_absolute\_error' (MAE) was used during training to identify the best regression based on the value of MAE. To assess the predictive performance of both models, three standard metrics were chosen namely Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and coefficient of correlation ( $R^2$ ).

### 3. Results and discussion

#### 3.1 Identifying the optimum structure for DNN and ANN

To increase the accuracy of the two models, the grid search hyperparameters tuning method was applied to search for the optimum number of hidden layers, neurons, and epochs. When the grid search was used for the DNN-GC model, it was observed that 2 hidden layers with 30 and 256 neurons each, with a number of epochs 10,000, and activation function ReLU and sigmoid for the hidden layer and the output layer, respectively, were required to increase the model accuracy. Also, MAE is used for training and validation loss function as stated earlier to validate the accuracy of the model. Furthermore, regularization and dropout methods were also used to avoid overfitting or underfitting during the training. A similar approach was followed for the ANN-GC model, only one hidden layer with 40 neurons was enough to increase the accuracy of the model and neither regularization nor dropout was required for the ANN-GC model. After training both models for nearly an hour, and based on the optimum structures, both models were validated graphically and statistically. Figure 1 depicts the correlated viscosity and CO<sub>2</sub> solubility values for both models versus the corresponding experimental data. Both predicted values for both models are accumulated around the 45° line which indicates a good fit. As for the statistical evaluation, the results have shown high R<sup>2</sup> of 95% and 96%, RMSE of 0.052 and 0.065, and MAE of 0.042 and 0.021 for DNN-GC and ANN-GC, respectively.

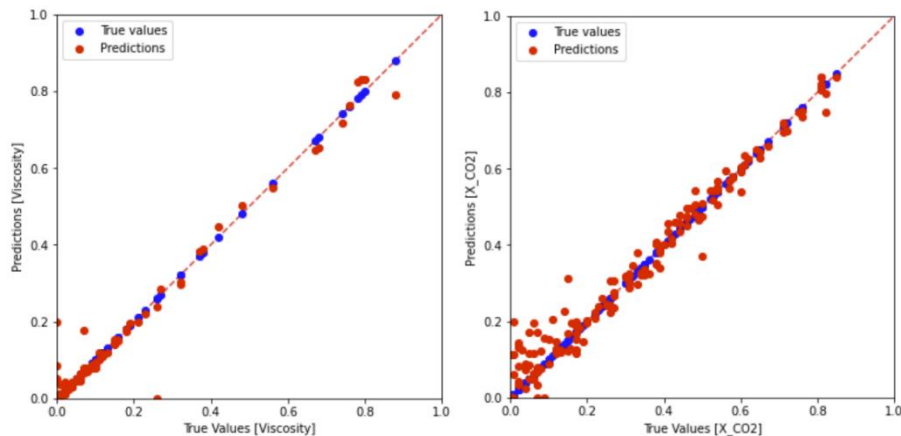


Figure 1: Cross-plot of true values versus predictions for Viscosity using DNN (Left) and CO<sub>2</sub> solubility using ANN (Right)

#### 3.2 Illustrative case study

To explicitly study the trade-off between the CO<sub>2</sub> solubility and viscosity of ILs, both trained models were merged by making the output of the DNN-GC model (viscosity) as an input feature to the ANN-GC model to simultaneously understand the behavior of the viscosity of the ILs and their CO<sub>2</sub> solubility capacity by varying temperature and pressure. The ‘Keras.Model’ option available in TensorFlow was used to merge both models. To find the optimum parameters (activation function and epochs), a grid search was applied.

Optimizers ReLU and sigmoid for hidden and output layers, respectively, gave better performance, and an epoch of 8000 improved model accuracy (training time = 25 minutes). The previously used metrics were also applied to the merged model. As for graphical validation, a good fit is depicted and almost all predicted data are accumulated around the 45° line as shown in Figure 2. For statistical validation, results showed that the merged model has better accuracy with an R<sup>2</sup> of 97%. RMSE and MAE are 0.041 and 0.022, respectively. This indicates that by merging models, the performance of the model is enhanced.

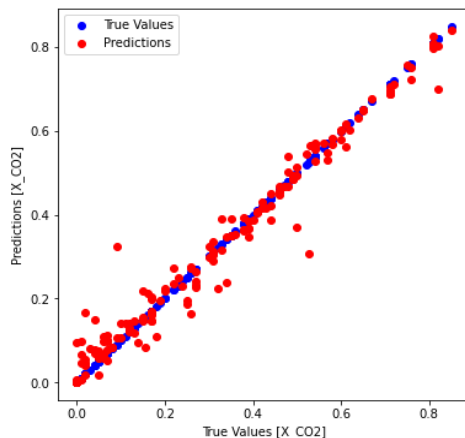


Figure 2: Cross-plot of true values versus predictions for CO<sub>2</sub> solubility using the merged model – DNN-ANN

In this illustrative study, the objective is to demonstrate the capability of the developed models to find the optimum structure of IL with maximal CO<sub>2</sub> absorption capacity using the given functional group in Table 1. Two constraints were set for the feasibility of ILs: (1) the maximum number of cation substituents [CH<sub>2</sub>] was set to be 8 and (2) viscosity less than 0.1 Pa.s. Also, only cation core [Im<sub>13</sub>]<sup>+</sup> and [Im<sub>123</sub>]<sup>+</sup> were considered in this example. To illustrate a post-combustion CO<sub>2</sub> capture, absorption conditions were pre-specified as 293.15 K and 1.5 MPa. A python code was developed to predict the optimum combination of a functional group to obtain an IL with high CO<sub>2</sub> capacity and low viscosity. Table 2 presents the best IL candidates, with the following IL structure cation [Im<sub>13</sub>]<sup>+</sup> anion [Tf<sub>2</sub>N]<sup>-</sup>, 1OH, 8CH<sub>2</sub>, 1 aN-CH<sub>3</sub>, 1aN-CH<sub>2</sub> having the highest CO<sub>2</sub> removal and low value of viscosity (< 0.1 Pa.s). However, IL with anion [PF<sub>6</sub>]<sup>-</sup> might not be a viable option for the CO<sub>2</sub> capture process due to its high viscosity.

Table 2: Merged model results of the optimal ILs

Optimal IL group combination	IL Viscosity (Pa.s)	CO <sub>2</sub> Solubility
1OH, 8CH <sub>2</sub> , 1[Im13] <sup>+</sup> , 1 aN-CH <sub>3</sub> , 1aN-CH <sub>2</sub> , [Tf <sub>2</sub> N] <sup>-</sup>	0.0111	0.642
1OH, 8CH <sub>2</sub> , 1[Im13] <sup>+</sup> , 1 aN-CH <sub>3</sub> , 1aN-CH <sub>2</sub> , [PF <sub>6</sub> ] <sup>-</sup>	0.553	0.543
1CH <sub>3</sub> , 7CH <sub>2</sub> , 1[Im123] <sup>+</sup> , 1 aN-CH <sub>3</sub> , 1aN-CH <sub>2</sub> , [AC] <sup>-</sup>	0.080	0.647

#### 4. Conclusion

This paper presents a new modeling framework for the design of the optimal molecular structure of IL for CO<sub>2</sub> solubility, from a given set of functional groups. One important target property considered in this study is viscosity since it affects the pumping power required and ultimately the operating cost. To achieve this, two predictive deep learning models, DNN and ANN, were built and trained separately for IL viscosity and CO<sub>2</sub> solubility, respectively. The grid search method for hyperparameters tuning was employed to find the optimum DL-GC structure with regularization and dropout applied to avoid overfitting and underfitting. Results of validation metrics R<sup>2</sup> were found to be 95% and 96% for DNN and ANN, respectively. Then, both models were merged, with the output of the DNN model (viscosity) used as an input feature to the ANN model (CO<sub>2</sub> solubility). This improved the performance of the model (R<sup>2</sup> = 97%). Furthermore, to investigate the applicability of the merged model, a case study was carried out to design optimum ILs based on a set of structural and viscosity constraints. Results showed that IL with [Im<sub>13</sub>]<sup>+</sup>, [Tf<sub>2</sub>N]<sup>-</sup>, and cation substituents (1OH) is a good candidate in terms of viscosity and CO<sub>2</sub> solubility.

#### 5. Acknowledgment

The authors acknowledge the paper was made possible by grant QUHI-CENG-22-23-465 from Qatar University. The statements made herein are solely the responsibility of the author[s].

#### 6. Reference

- Chen, Y., Peng, B., Kontogeorgis, G. M., & Liang, X. (2022). Machine learning for the prediction of viscosity of ionic liquid–water mixtures. *Journal of Molecular Liquids*, 350, 118546. <https://doi.org/10.1016/j.molliq.2022.118546>
- Deng, T., Liu, F. hai, & Jia, G. zhu. (2020). Prediction carbon dioxide solubility in ionic liquids based on deep learning. *Molecular Physics*, 118(6), 1–8. <https://doi.org/10.1080/00268976.2019.1652367>
- Ionic Liquids Database - ILThermo. (n.d.). Retrieved September 11, 2022, from <https://ilthermo.boulder.nist.gov/>
- Nath, D., & Henni, A. (2020). Fluid Phase Equilibria Solubility of carbon dioxide ( CO<sub>2</sub> ) in four bis ( tri fl uoromethyl-sulfonyl ) imide ( [ Tf<sub>2</sub>N ] ) based ionic liquids. *Fluid Phase Equilibria*, 524, 112757. <https://doi.org/10.1016/j.fluid.2020.112757>
- Pezhman, S., Atashrouz, S., Nakhaei-kohani, R., & Hadavimoghaddam, F. (2022). Modeling of H<sub>2</sub>S solubility in ionic liquids using deep learning : A chemical structure-based approach. *Journal of Molecular Liquids*, 351, 118418. <https://doi.org/10.1016/j.molliq.2021.118418>
- Schmidhuber, J. (2015). Deep Learning in neural networks: An overview. *Neural Networks*, 61, 85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>
- Zarra, T., Galang, M. G., Jr, F. B., & Belgiorno, V. (2019). Environmental odour management by arti fi cial neural network – A review. *Environment International*, 133(May), 105189. <https://doi.org/10.1016/j.envint.2019.105189>

# Economic optimization of the Northern Italian supply chain for residual plastic packaging waste treatment

Fabio Cieno<sup>a</sup>, Daniel Crîstiu<sup>a</sup>, Federico d'Amore<sup>a,b</sup>, Fabrizio Bezzo<sup>a</sup>

<sup>a</sup>*CAPE-Lab – Computer-Aided Process Engineering Laboratory, Department of Industrial Engineering, University of Padova, via Marzolo 9, Padova IT-35131, Italy*

<sup>b</sup>*Politecnico di Milano, Department of Energy, via Lambruschini 4, IT-20156 Milano (Italy)*

*fabrizio.bezzo@unipd.it*

## Abstract

Plastics are present in many products and nearly every commercial sector. However, the management of plastic waste is still an open challenge. This work proposes a mixed integer linear programming optimization model of the Northern Italian supply chain for mixed plastic waste treatment. The model aims at maximizing the economic performance of the overall supply chain, taking into account both the selection of the treatment technologies (i.e. incineration, gasification and pyrolysis) and finding the optimal distribution of the material flows. The best economic scenario is achieved by a supply chain where only incineration plants are selected, yielding a gross profit of 31.8 M€/y for the overall supply chain. However, by varying oil price and specific treatment costs both pyrolysis and incineration plants are selected, showing that pyrolysis could be an economically competitive technology for chemical recycle of the mixed plastic waste, leading to a gross profit of 21.6 M€/y (- 32 % w.r.t. base case scenario).

**Keywords:** mixed plastic waste, plastic recycling, optimization, supply chain.

## 1. Introduction

Plastics are key and ubiquitous materials whose increasing consumption poses a challenge with regard to the management of plastic waste, as significant fractions of plastics still end up in landfills or leak into the ocean (Li et al., 2022). The largest fraction of plastic waste is made of plastic packaging; unfortunately, only a limited fraction can be mechanically recycled. The residual fraction, called mixed plastic waste (MPW), is either used as an ingredient for secondary solid fuel production, or directly incinerated or landfilled. The challenges in managing MPW are due to the variability in its composition, and incineration, which is an industrially well-established pathway, can help reducing the amount of waste that needs to be landfilled, and recover electric energy as well. Pyrolysis and gasification represent potentially attractive emerging technologies for MPW management; however, pyrolysis only preserves some molecular complexity and can be considered the first step of a chemical recycle path. Pyrolysis and gasification were indeed identified as the most mature and commercially viable chemical recycling technologies suitable for MPW treatment, according to the Threshold Readiness Level (TRL) assessment of Solis and Silveira (2020).

Supply Chain (SC) optimization raises as an important research topic in achieving an efficient plastic circular economy, and it allows assessing both the optimal selection of



MPW treatment technologies and an optimal distribution of the material flows among the different stages of the SC. The combinatorial complexity of the MPW SC is given by its multi-stage, spatially-explicit nature, and mixed integer linear programming (MILP) models are regarded as powerful mathematical optimization tools that can deal with such systems (Kallrath, 2000). This study aims at proposing a spatially explicit MILP framework for the economic optimization (in terms of maximizing the annual gross profit [€/y]) of MPW SCs, focusing on the Northern Italian region and considering multiple technological options.

## 2. Modelling framework

This section will describe the stages of the optimized MPW SC in terms of key model input parameters. The problem is formulated as an annualized spatially-explicit MILP modelling framework based on the location (in terms of latitude and longitude) of Italian provinces, sorting centres, incineration plants and oil refineries. The model optimization outcomes provide quantitative information on the SC in terms of economic results and technology selection under different scenarios. Both waste-to-energy (i.e. incineration) and chemical recycling technologies (i.e. pyrolysis and gasification) are considered as treatment options. The block diagram of the MPW SC is represented in Figure 1.

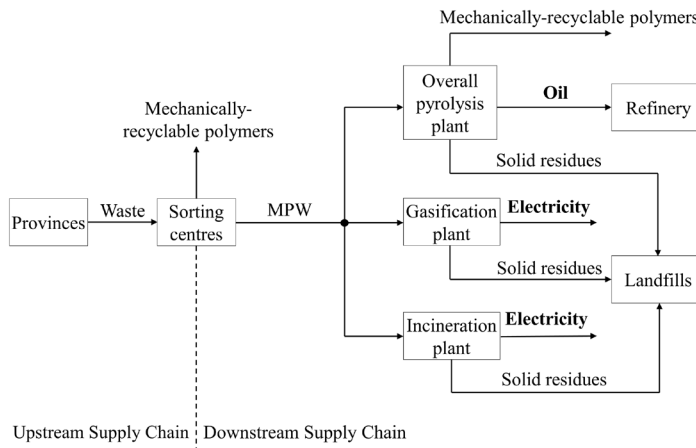


Figure 1. Block diagram of the MPW Supply Chain.

The stages of the MPW SC are divided into upstream SC and downstream SC. The upstream SC deals with the transport of plastic packaging waste (PPW) from provinces to sorting centres, separating the PPW share that can be mechanically recycled from the remaining part, i.e. the MPW. The downstream SC contains the MPW treatment options (incineration, gasification, pyrolysis), whose outputs are: electric power in the case of incineration and gasification, and oil in the case of pyrolysis plants, to be transported to a refinery for further processing (the system boundary is at the gate of the refinery in the case of pyrolysis oil). MPW is a heterogeneous mixture, and, to comply with the quality requirements for the pyrolysis oil produced, an additional sorting step is coupled with the pyrolysis plant; its costs and sorting performance are taken into account together with those of pyrolysis. Two options of pyrolysis technology are implemented: *Po. Pyrolysis* – the reactor feed consists of the polyolefin fraction, producing an oil similar to Brent oil. *PoPs. Pyrolysis* – using polyolefin and polystyrene fractions, producing an oil similar to

that of Western Canadian Select (WCS) (Mastellone, 2020). Road transport is considered as the only transport option since it reflects current practice for plastic waste.

### 2.1. Modelling assumptions and inputs

The spatially-explicit nodes are described in the model through a set  $n = \{n_{1-86}\}$  comprising 47 provinces, 12 sorting centres, 24 incinerators and 3 refineries. Each node is characterized by its exact geographic coordinates. The incineration plants are the plants currently available in Northern Italy (Utilitalia, 2019). It is assumed that gasification plants can be installed only in proximity of incineration plants; similarly, pyrolysis plants can be sited either by incineration plants or by refineries. The yearly PPW flowrate of each province are taken from COREPLA (2020) and are referred to 2019. Sorting centres are characterized in terms of plant capacity and sorting performances (their PPW-to-MPW conversion factor). All treatment plants generate solid residues that are sent to nearby landfills. For each technology (except incineration) three plant sizes have been considered (*Small, Medium and Large*). Each technology has been characterized in terms of relevant mass balances, energy balances, and each plant size and each incinerator has been characterized by its Specific Treatment Costs (*STC* [€/kt]) defined as:

$$STC = \frac{CAPEX + OPEX}{C} \quad (1)$$

where  $C$  is the annual plant capacity [kt/y],  $OPEX$  is the Operational Expenditure [M€/y] and  $CAPEX$  is the annualized Capital Expenditure [M€/y] calculated as:

$$CAPEX = TCI \cdot \frac{i \cdot (1 + i)^N}{(1 + i)^N - 1} \quad (2)$$

$TCI$  [M€] is the Total Capital Investment (updated using CEPCI, 2019),  $i$  is the interest rate, and  $N$  is the plant life. For all technologies and plant sizes, it is assumed that  $i = 15\%$  and  $N = 15$  years.

The decision variables in the optimization problem are:

- the PPW flowrates distribution from the provinces to the sorting centres;
- the MPW flowrates distribution from the sorting centres to the treatment plants;
- the selection of technology, plant size and location of the treatment plants.

### 2.2. Mathematical Formulation

The objective function of the MILP aims at maximizing the SC annual gross profit:

$$Objective = \max\{GP\} \quad (3)$$

Gross profit  $GP$  [€/y] is defined as the difference between the total annual revenues ( $TR$  [€/y]) and total annual costs ( $TC$  [€/y]):

$$GP = TR - TC \quad (4)$$

Total revenues  $TR$  [€/y] of Eq.(4) sums up revenues from gate tariffs cashed in  $Rev^{T,tot}$  [€/y], and from selling mechanically-recyclable polymers sorted out in the SC  $Rev^{Mech}$  [€/y], pyrolysis oil  $Rev^{Oil}$  [€/y], and electric power produced  $Rev^{El}$  [€/y].

$$TR = Rev^{T,tot} + Rev^{Mech} + Rev^{Oil} + Rev^{El} \quad (5)$$

A gate tariff value of 210 €/t of treated MPW is considered. Revenues from selling the mechanically-recyclable polymers from pyrolysis plants are computed based on 2019 market price 378 €/t. To calculate the revenues from the pyrolysis oil, the oil market price

is referring to 2019 average prices, and based on the choice of pyrolysis, either only from polyolefin fraction producing an oil similar to Brent oil (64.3 US\$/bbl) or from polyolefin and polystyrene fractions producing an oil similar to WCS (44.3 US\$/bbl). Revenues from electric power consider a market price of 52346 €/GWh, which is the average 2019 selling price in the Italian Power Exchange trading market.

The total cost  $TC$  [€/y] of Eq. (4) is defined as the sum of total MPW treatment cost ( $TPTC$  [€/y]), total landfilling cost ( $TLC$  [€/y]) and total transport cost ( $TTC$  [€/y]):

$$TC = TPTC + TLC + TTC \quad (6)$$

Total MPW treatment cost ( $TPTC$  [€/t]) is calculated as the sum between the overall plant specific treatment cost ( $STC$  [€/kt]) and incineration costs for remaining MPW  $PTC$  [€/y]:

$$TPTC = \sum m_{k,s}^{MPW} \cdot STC_{k,s} \cdot c_{k,s} + PTC \quad (7)$$

where  $m_{k,s}^{MPW}$  [kt/y] is the MPW flowrate treated in a plant  $k$  of size  $s$ , while  $c_{k,s}$  is a factor accounting for cost saving in the case a pyrolysis plant co-exists with a refinery.

The total landfilling cost ( $TLC$  [€/y]) is computed from the total residues produced in the SC ( $m^R$  [kt/y]) and landfill gate tariff  $T^l$  that equals 210 €/t of residues:

$$TLC = m^R \cdot T^l \quad (8)$$

Total transport cost ( $TTC$  [€/y]) is the sum of transporting waste, MPW, remaining MPW, oil and residues ( $TC^w$ ,  $TC^{MPW}$ ,  $TC^{RMPW}$ ,  $TC^{Oil}$ ,  $TC^R$ , respectively, all in [€/y]).

$$TTC = TC^w + TC^{MPW} + TC^{RMPW} + TC^{Oil} + TC^R \quad (9)$$

Each term is computed based on Unitary Transportation Cost ( $UTC$  [€/kt/km]) specific to the transported material (for solids:  $UTC^{MPW}$ , and for oil:  $UTC^{Oil}$ ), and the matrix of linear distances between nodes calculated with the spherical law of cosines (d'Amore and Bezzo, 2017).

### 3. Results

The model has been implemented in GAMS software and optimized by means of CPLEX solver. Two different scenarios are considered: a base case scenario containing all treatment options, and an alternative scenario excluding incineration.

The economic performance of the MPW SC for the different scenarios, and a cost breakdown showing are presented in Table 1.

Table 1. Total supply chain cost  $TC$  [M€/y] and the contribution of each component into the total cost: total MPW treatment cost  $TPTC$  [M€/y], total landfilling cost  $TLC$  [M€/y], total transport cost  $TTC$  [M€/y]. Annual gross profit  $GR$  [M€/y] are also presented.

Scenario	$TC$	$TPTC$		$TLC$		$TTC$		$GP$
	[M€/y]	[M€/y]	[%]	[M€/y]	[%]	[M€/y]	[%]	[M€/y]
Base Case	59.1	38.6	65.3	5.4	9.1	15.1	25.5	31.8
1-Pyro	68.2	45.5	66.7	7.1	10.4	15.6	22.9	
Alternative	99.4	66.0	66.4	11.7	11.8	21.7	21.8	21.6

### 3.1. Base Case Scenario Results

The base case scenario considers all MPW treatment technologies discussed for the economic optimization of the supply chain. It is characterized by a gross profit for the overall SC of 31.8 M€/y. Electric power contributes to about 21 % of total revenues. However, electricity revenues do not compensate for the total treatment costs. The largest contribution to total revenues (79 %) is given by the gate tariff received by treatment plants for treating MPW. The supply chain configurations are shown in Figure 2, for both the upstream and downstream sections.

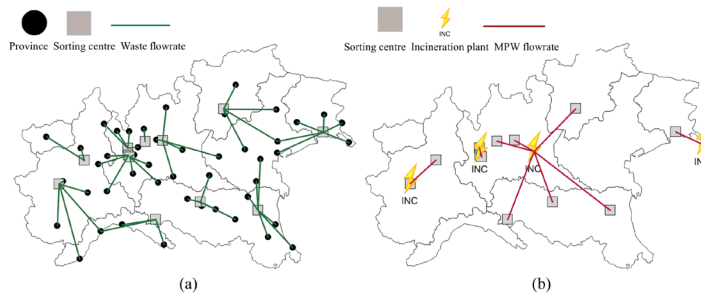


Figure 2. Base case supply chain configuration: (a) upstream SC, (b) downstream SC.

From Figure 2(a) it can be observed that the waste flowrate of each province is sent to the nearest sorting centre. This result suggests that whatever the advantages of sending the waste flowrates to sorting centres that are close to the final treatment plants, they are outmatched by the waste flowrates transport costs. From Figure 2(b), only the Northern Italian incineration plants are selected while no pyrolysis nor gasification plants are selected.

We verified that for the solver to select a different technology the oil price should be increased to a value of 72.7\$/bbl for Brent oil and to 50.1\$/bbl for WCS oil. The resulting configuration is shown in Figure 3, and it can be observed that both a pyrolysis plant and incineration plants are selected. The selected pyrolysis plant is of the largest plant size (benefiting from economies of scale) and it is located where a refinery exists, suggesting that the cost savings due to the coexistence of an oil refinery play a significant role.

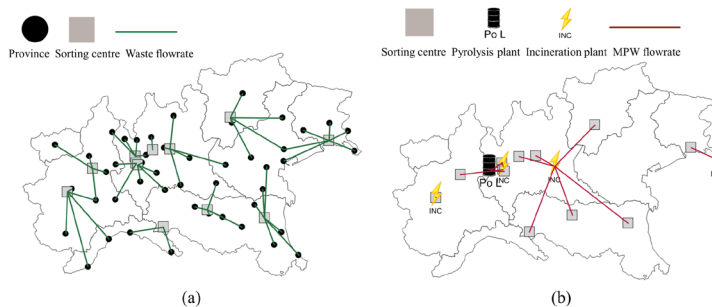


Figure 3. 1-Pyro Scenario MPW SC configuration: (a) upstream SC, (b) downstream SC.

### 3.2. Alternative Scenario Results

The alternative scenario excludes incineration (except for treating the remaining MPW from pyrolysis). The MPW SC configuration is shown in Figure 4, and it can be observed that only pyrolysis plants are selected, while gasification appears to not be economically

competitive w.r.t to the other technologies. From Table 1, the gross profit related to the alternative scenario is 21.6 M€/y, i.e. about 32 % lower w.r.t. the base case scenario.

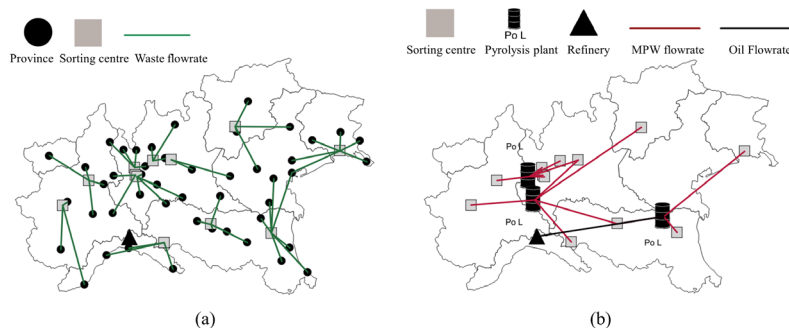


Figure 4. Alternative scenario MPW SC configuration: (a) upstream SC, (b) downstream SC

From Table 1 it can be observed that the total MPW treatment cost (*TPTC*) has the largest contribution representing more than 65 % of the to the total cost (*TC*) in all analysed scenarios. It can be observed that all cost components increase when pyrolysis plants are selected. In particular, the total landfilling cost (*TLC*) increases due to increasing flowrate of total solid residue generated by the pyrolysis in the SC.

#### 4. Conclusions

This study presented a MILP framework for the optimal design of an MPW supply chain in Northern Italy. Both waste-to-energy (i.e. direct incineration) and chemical recycling technologies (i.e. pyrolysis and gasification) were considered as possible MPW treatment options. The economic optimization of the base case scenario gives a configuration in which only incineration is selected, exhibiting an annual gross profit of 31.8 M€/y. However, it was proven for a higher oil price (comparable to 2022 oil prices), pyrolysis becomes economically competitive. Gasification plants are never a viable alternative. If incineration is excluded, only pyrolysis is selected, giving a gross profit of 21.6 M€/y (- 32 % w.r.t. base case scenario). Future work will address a multi-objective optimization accounting for both economic and environmental performances.

#### References

- F. d'Amore, F. Bezzo, 2017, Economic optimization of European supply chains for CO<sub>2</sub> capture, transport and sequestration. *International Journal of Greenhouse Gas Control*, 65, 99-116.
- CEPCI, 2019, Chemical engineering plant cost index [WWW Document, Accessed 21/11/2022], <https://www.chemengonline.com/pci-home>.
- COREPLA, 2020, Management Report 2019, [WWW Document, Accessed 21/11/2022] [https://www.corepla.it/sites/default/files/documenti/relazione\\_sulla\\_gestione\\_2019](https://www.corepla.it/sites/default/files/documenti/relazione_sulla_gestione_2019).
- J. Kallrath, 2000, Mixed Integer Optimization in the Chemical Process Industry: Experience, Potential and Future Perspectives, *Chemical Engineering Research and Design*, 78, 6.
- H. Li, et al., 2022, Expanding Plastics Recycling Technologies: Chemical Aspects, Technology Status and Challenges, *Green Chemistry Journal*, 1463-9262.
- M.L. Mastellone, 2020, Technical description and performance evaluation of different packaging plastic waste management's systems in a circular economy perspective, *The Science of the Total Environment*, 718, 137233.
- M. Solis, S. Silveira, 2020, Technologies for chemical recycling of household plastics – A technical review and TRL assessment, *Waste Management*, 105.
- Utilitalia, 2019, [WWW Document] <https://www.utilitalia.it/pdf/9ea3b0ab-6079-4975-88b0-814663c3fb15> [Accessed 21/11/2022].

# Optimisation modeling for decision support in the industrial circular economy activities

K. G. Stylianopoulou<sup>a</sup>, E. M. Kondili<sup>a</sup>, C. M. Papapostolou<sup>a</sup>, J. K. Kaldellis<sup>b</sup>

*<sup>a</sup>Optimisation of Production Systems Lab., Mechanical Engineering Department, University of West Attica, Greece*

*<sup>b</sup>Soft Energy Applications and Environmental Protection Lab., Mechanical Engineering Department, University of West Attica, Greece*

## Abstract

Recently circular economy has attracted a lot of attention in its ability to handle global challenges related to resource scarcity, sustainability and supply chain unpredictability. Designing circular products is probably one of the most critical factors that manufacturers need to consider when it comes to implementing a closed-loop production strategy. In continuation to our in-progress research related to circular economy applications in the manufacturing and process industry, the present work focuses in the development of optimisation models for the circular economy activities in the industry that will assist substantially in the decision making for implementing the most efficient and sustainable activities. The model development for such complex processes has very significant added value for the evaluation of the economic, environmental and technical circular economy options available. In this context, optimisation models are presented as well as the main problem variables and parameters are identified and the problem constraints are developed. Furthermore, the optimisation criteria for different industrial systems are also described. The results and conclusions of the work highlight the characteristics of the complex circular economy alternatives and develop the basis for the continuation and the expansion of the work in different industrial units.

**Keywords:** Circular economy modelling, Optimisation, industrial activities.

## 1. Introduction

The notion of circular economy (CE) the last few years has been widely accepted from manufacturing industries in order to transform their products processing and consumption to more sustainable throughout their supply chain. A transition from linear to a closed loop or circular system requires a shift away from the traditional model of producing products. Manufacturing industries have a significant impact on engineering process that must be adaptable, sustainable and efficient while maintaining high quality and low cost. Nonetheless, the big market demand for a long period of time is resulting in depleting natural resources and making products less sustainable. Over the past few decades efforts have been initiated to reduce our environmental impact through clean manufacturing and life cycle management initiatives (Stylianopoulou et al., 2021). Thus, this necessitates an innovative approach in researching the challenges and limitations of traditional linear economy to circular economy by making products and the supply chain of products more sustainable. One of the activities and focus is extending to a second life products by remanufacturing/recycling used products which are discarded by end-users when they are no longer needed. Advances towards CE mean efforts to improve the circulation of resources, materials and energies by closing loops. Thus, the objective of the present work is to highlight the relevance of CE with optimisation and make a concise review of the

current status in optimisation CE models, also, indicating the problem parameters and variables.

## 2. Tools for Circular Economy

Industries play an important role in implementing the CE notion. Product innovation, redesign, remanufacture, or value chain reconfiguration, can have a significant impact on how resources are managed and consumed. A remanufacturing industry must increase its quantity to effectively control variation in its supply chain and reduce the processing times required to intelligently incorporate remanufactured materials into new ones, such as the use of Industry 4.0 (I4.0). As a result, remanufacturing should be planned in such a way that allows high quantity content in the shortest amount of time when the products are produced (Baratsas et al., 2021, Nascimento et al., 2019). The use of harmonized CE standards and tools can be critical in the implementation of CE. This section presents selected CE tools which promote circularity in industries. Following, selective tools are presented which main goal is to promote circularity in industrial business.

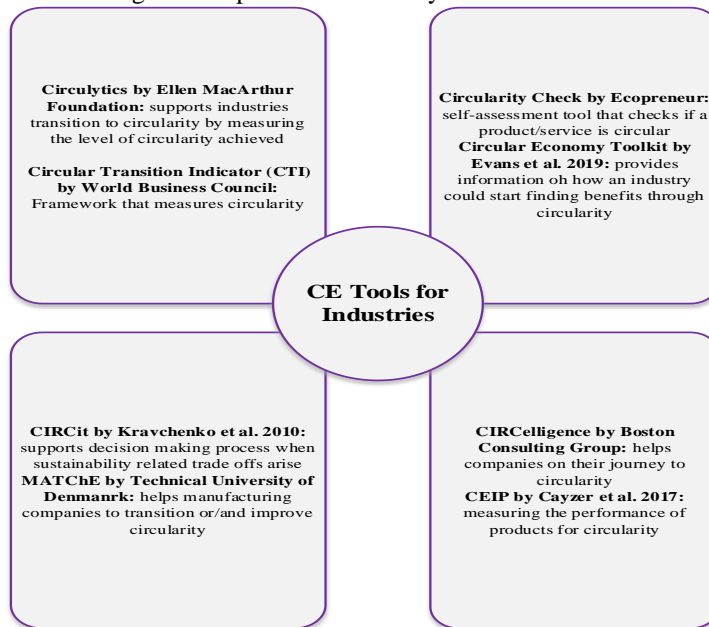


Figure 1: Existing CE Tools for Industries

However, these tools fail to give a quantitative solution in the CE approach by making them more of a framework than quantitative tool. Their approach is on a small amount of product life cycle without giving solution to the entire product lifecycle. Thus, there is a need for more accurate methodologies such as mathematical optimisation to begin implementing CE in industries.

## 3. Optimisation Models in CE

Circular economy in industries has been gaining interest from researchers and the industries themselves in view of transitioning from linear to circular economy. However, how an industry can approach circularity depends on the framework available to each country and how they can implement circularity.

Reference	Optimisation Model	Description	Field of Research
Zhao et al. 2012	Empirical study	Propose CE model for coal mine enterprise	Mining Industry
Li 2012	Quantitative evaluation model	Waste input-output analysis table	Circular Economy Enterprise
Wu et al. 2014	DEA window analysis	CE efficiency ranking	Country level
Ma et al. 2014	Evaluation system	Estimate circular Economy performance	Iron&steel industry
George et al. 2015	Theoretical model	CE model with two types of resources, polluting input & a recyclable input	CE activities
Igarashi et al. 2016	Multi-criteria optimization	Lower disassembly cost, higher recycling, & CO <sub>2</sub> saving rate	Recycling factories
Rada & Cioca 2017	MSW modelling	Planning waste management for CE	Municipal Solid Waste
Lin 2018	Empirical study	Define product design for smart manufacturing	Glass Industry
Jabbour et al. 2018	ReSOLVE model	Use of I4.0 to enhance CE principal application	Manufacturing Industries
Angouria-Tsorochidou 2018	Techno-economic assessment	Optimised collection of EoL electronic products	Electronic Industry
Rajput & Singh 2019	PCA-DEMATEL	Link between I4.0 & CE	Supply Chain networks
Nascimento et al. 2019	Qualitative research	Integrating I.4 & CE	Waste recovery to the supply chain
Velazquez Mar et al. 2019	Statistical entropy analysis	Optimizing a lithium-ion battery waste system	Battery waste sieving system
Garcia-Barragan et al. 2019	Mathematical model	Measuring the Circular Economy	Virgin & recycled materials
Bal & Badurdeen 2020	IoT & PSS model	Achieve CE through closed-loop material flow	EoL product recovery facilities
Atabaki et al. 2020	Multi-objective MILP model	Cost, CO <sub>2</sub> emissions & energy consumption for durable Products	Closed-loop supply chain
Yadav et al. 2020	Hybrid Best-Worst Method Elimination& Choice Expr. Reality	Framework for sustainable supply chain	Automotive sector
Baratsas et al. 2021	Framework & decision-making tool	CE PSE framework & decision-making tool for the modeling & optimisation of food SC	Food Supply Chains
Karayilan et al. 2021	Linear single-objective optimisation models	Environmental benefits & circularity in plastics wastes supply chains	Plastic Packaging

Table 1: Selected literature review of CE optimisation

For instance the EU supports recycling and recovery targets where in the US there is no relevant framework (Saidani et al., 2019, Bjørnbet et al., 2021). Performance indicators are also required in order to monitor targets in CE models in the manufacturing sector. Our research was conducted in over 100 articles relevant to CE implementation in manufacturing, food and electronic industries. However, only a few of them are related



to the optimisation models development in CE. In the context of our focus in developing optimisation models expressing CE strategies for the industry, we have completed a detailed review in the suggested CE relevant optimisation models. The analysis presented in Table 1 is mainly focused in the structure of the optimisation models, its criterion and constraints. Through this research it was identified that Industry 4.0 (I4.0) with CE enables a sustainable innovation model that improves end-of-life activities, reduces waste and tracks production and consumption operations. On the other hand, publications that are not using I4.0 are either generally referring to optimisation models or are not giving solutions across the entire product lifecycle from manufacturing to use/reuse to end of life. What has been discovered is that product design optimisation is a rather unexplored area. Some publications proposed frameworks so as to analyse trade-offs between profit, environmental impact and circular performances but without optimizing the physical parameters of a product.

#### 4. Mathematical Model to be developed

The problem considered here is to approach CE and cleaner production with a mathematical model. In the publications reviewed, there are plenty of quantitative models which are focusing in I4.0, smart manufacturing, Industrial Internet of Things (IIoT), but its integration with the optimisation models focusing on CE and cleaner production is attempted by only a few researchers.

<i>Parameters</i>	<i>Variables</i>
Demand for new product	Qty of product
Demand of new & refurbished product	Qty of product shipped from plant to warehouse to market
Demand for used products	Qty of returned product of least reusable
Demand for remanufacturing products	Qty of new products to be produced
Capacity for processing/production	Qty of remanufactured products to be produced
Energy consumption	Qty of products to be recycled
Production time	Qty of products to be refurbished
Number of returned products	Qty of products to be disposed
Cost of remanufactured products	<b>Constraints</b>
Cost of used products	Warehouse for returned product
Cost for new products	Qty of used products allocated from least reusable level
Setup Cost for production	Warehouse for products
Waste disposal cost	<b>Sets</b>
Warehouse cost	Set of products
Returned products to warehouse cost	Set of plant
Transportation Cost	Set of warehouses
Cost of disassembling & inspecting returned products	Set of market
Cost of products to be manufactured, cost of refurbished products	Set for time period
Collection and logistics costs	

Table 2: Description of proposed mathematical model parameters, variables and sets

Thus, there is a need of more focused and advanced optimisation models modelling circularity to fulfill demand, clean production and zero waste. Table 2 presents model assumptions, sets, parameters, and variables to formulate a mathematical optimisation model for industries to implement CE.

Currently there are three dimensions that the circularity in industry is approached in our work.

1. Feasibility models are developed to investigate in detail all the investment and operational costs of potential remanufacturing/ reuse investments. The feasibility model is generic; however specific case study is implemented for the chemical recycle of plastics.
2. Environmental performance software is being implemented to provide various performance indices that are exploited then as inputs in the optimisation models.
3. Optimisation models are currently developed for various CE alternatives in specific industrial cases.

## **5. Conclusion**

The industrial / manufacturing sector plays an important role in transitioning from linear to circular economy. Industries can influence the way resources are consumed and managed through implementing circular economy in their supply chains. Most researchers use the 4Rs strategies, reduce reuse recycle and recover, to approach this topic, however the 10Rs (refuse, rethink, reduce, reuse, repair, refurbish, remanufacture, repurpose, recycle and recover) of CE is more preferable solution in waste scale so as to prevent waste to landfill and achieve zero waste. In order to achieve this there is a must of implementing strategies for product and supply chain design and material selection. However, there are several barriers that backtrack such a transition to a more circular model. Those barriers are the lacking of a unifying definition of CE and a common understanding of the concept, lack of technological innovations in the supply chains of industries that are focusing in CE. In addition, optimisation models presented in the publications are more general and they are not giving solutions across the entire product lifecycle and supply chain of a product. Thus, more research is needed for better understanding of how CE industries can contribute to sustainable production and consumption. Furthermore, the optimisation models for CE product design needs to aim to more pertinent analysis for the whole product life cycle of an industry and comprehend the CE strategies. For this the use of harmonized CE standards and tools can play an important role. Further research needs to be done for optimisation models in material section from larger material databases which are going to benefit the CE product life cycle.

## **References**

- E. Angouria-Tsorochidou, C. Cimpan, K. Parajuly, 2018, Optimised collection of EoL electronic products for Circular economy: A techno-economic assessment, *Procedia CIRP*, 69, 986 – 991
- M.S. Atabaki, M. Mohammadi, B. Naderi, 2020, New robust optimisation models for closed-loop supply chain of durable products: Towards a circular economy, *Computers & Industrial Engineering*, 146
- A. Bal, F. Badurdeen, 2020, A Multi-Objective Facility Location Model to Implement Circular Economy, *Procedia Manufacturing*, 51, 1592-1599
- S.G. Baratsas, E.N. Pistikopoulos, S. Avraamidou, 2021, Circular Economy Systems Engineering: A case study on the coffee supply chain, *Computer Aided Chemical Engineering*, 50, 1541-1546
- M.M. Bjørnbet, C. Skaar, A.Ma. Fet, K. Ø. Schulte, 2021, Circular economy in manufacturing companies: A review of case study literature, *Journal of Cleaner Production*, 294
- Ecopreneur, 2022, Circularity Check. <https://ecopreneur.eu/circularity-check-landing-page/>.
- Ellen MacArthur Foundation, 2019, Circulytics: measuring circularity, <https://www.ellenmacarthurfoundation.org/resources/apply/circulytics-measuring-circularity>

- J. Evans, N. Bocken, University of Cambridge, 2022, Circular Economy Toolkit, <http://circularconomytoolkit.org/Toolkit.html>
- J.F. García-Barragán, J. Eyckmans, S. Rousseau, 2019, Defining and Measuring the Circular Economy: A Mathematical Approach, *Ecological Economics*, 157, 369–372
- D.A.R. George, B.C. Lin, Y. Chen, 2015, A circular economy model of economic growth, *Environmental Modelling & Software*, 73, 60–63
- K. Igarashi, T. Yamada, S.M. Gupta, M. Inoue, N. Itsubo, 2016, Disassembly system modeling and design with parts selection for cost, recycling and CO<sub>2</sub> saving rates using multi criteria optimisation, *J. Manuf. Syst*, 38, 151–164
- A.B.L. Jabbour, C.J.C. Jabbour, M.G. Filho, D. Roubaud, 2018, Industry 4.0 and the circular economy: a proposed research agenda and original roadmap for sustainable operations, *Ann. Oper. Res.*, 270, 273–286
- S. Karayılan, O. Yılmaz, C. Uysal, S. Naneci, 2021, Prospective evaluation of circular economy practices within plastic packaging value chain through optimisation of life cycle impacts and circularity, *Resources, Conservation & Recycling*, 173
- M. Kravchenko, D.C.A. Pigosso, T.C. McAloone, 2020, A Guidance for Navigating Trade-Offs to Support Sustainability-Related Decision-Making, <https://doi.org/10.11583/DTU.12933431>
- S. Li, 2012, The Research on Quantitative Evaluation of Circular Economy Based on Waste Input-Output Analysis, *Procedia Environmental Sciences*, 12, 65 – 71
- K.Y. Lin, 2018, User experience-based product design for smart production to empower industry 4.0 in the glass recycling circular economy, *Comput. Ind. Eng.*, 125, 729–738
- S. Ma, Z. Wen, J. Chen, Z. Wen, 2014, Mode of circular economy in China's iron and steel industry: a case study in Wu'an city, *Journal of Cleaner Production*, 64, 505–512
- D.L.M. Nascimento, V. Alencastro, O.L.G. Quelhas, G.G.C. Rodrigo, A.G.R. Jose, L.R. Lona, G. Tortorella, 2019, Exploring Industry 4.0 technologies to enable circular economy practices in a manufacturing context: a business model proposal, *J. Manuf. Technol. Manag.*, 30, 3, 607–627
- S. Rajput, S.P. Singh, 2019, Identifying Industry 4.0 IoT enablers by integrated PCA-ISM-DEMATEL approach, *Manag. Decis.*, 57, 8, 1784–1817
- E.C. Rada, L. Cioca, 2017, Optimizing the Methodology of Characterization of Municipal Solid Waste in EU Under a Circular Economy Perspective, *Energy Procedia*, 119, 72–85
- H. Rubel, A. Meyer zum Felde, J. Oltmanns, C. Lanfer, L. Bayer, 2020, CIRCelligence by BCG it's Time to close our future resource loops, <https://www.bcg.com/de-de/circelligence-by-bcg-close-future-loops>
- M. Saidani, B. Yannou, Y. Leroy, F. Cluzel, A. Kendall, 2019, A taxonomy of circular economy indicators, *Journal of Cleaner Production*, 207, 542–559
- K.G. Stylianopoulou, E.M. Kondili, J.K. Kaldellis, 2022, Process Systems Engineering prospects in Circular Economy implementation in industry, *Computer Aided Chemical Engineering*, 51, 1309–1314
- Technical University of Denmark, 2019, MATCHE - making the transition to circular economy, <https://ready2loop.org/en/about>
- O.V. Velazquez Martínez, K.G. Van Den Boogaart, M. Lundstrom, A. Santasalo-Aarnio, M. Reuter, R. Serna-Guerrero, 2019, Statistical entropy analysis as tool for circular economy: Proof of concept by optimizing a lithium-ion battery waste sieving system, *Journal of Cleaner Production*, 212, 1568–1579
- WBCSD, 2022, CTI tool, <https://ctitool.com/>
- H. Wu, Y. Shi, Q. Xia, W. Zhu, 2014, Effectiveness of the policy of circular economy in China: A DEA-based analysis for the period of 11<sup>th</sup> five-year-plan, *Resources, Conservation and Recycling*, 83, 163–175
- G. Yadav, S. Luthra, S.K. Jakhar, S.K. Mangla, D.P. Rai, 2020, A framework to overcome sustainable supply chain challenges through solution measures of industry 4.0 and circular economy: an automotive case, *J. Clean. Prod.*, 254
- Y. Zhao, L. Zang, Z. Li, J. Qin, 2012, Discussion on the Model of Mining Circular Economy, *Energy Procedia*, 16, 438 – 443

# Model-Based Design of Experiments for the identification of microalgae growth models with limiting nutrients

Alberto Saccardo, Beatriz Felices-Rando, Eleonora Sforza, Fabrizio Bezzo

*Department of Industrial Engineering, University of Padova, Via Marzolo 9, Padova  
35131, Italy*

*fabrizio.bezzo@unipd.it*

## Abstract

Nutrients play a major role in several microalgal industrial processes and their efficient exploitation has an impact on environmental sustainability. The Droop model is suitable for the description of nutrients effect on microalgal growth but presents challenges in parameter identification for the high experimental effort required. In this work, a Model-Based Design of Experiments (MBDoE) was employed to plan information-rich experiments for precise parameter estimation. An initial dynamic sensitivity analysis showed that parameters of the Droop model can be practically identified with measurements of biomass and nitrogen content of the medium only. Finally, MBDoE was employed to plan *in silico* experiments, proving that two optimal experiments are sufficient to achieve accurate parameters estimation.

**Keywords:** Model Based Design of experiments; Droop model; microalgae; continuous photobioreactors; parameter estimation.

## Introduction

Microalgae are a wide variety of photosynthetic microorganisms, both prokaryotic and eukaryotic, that can be found in all ecosystems. Thanks to their ability to exploit photosynthesis to fix CO<sub>2</sub> and consume nutrients as inorganic nitrogen and phosphorus from wastewater, they represent a sustainable alternative for multiple industrial applications (Kumar et al., 2020; Olaiwola, 2003). Modelling the effect of nutrients on microalgal growth is of particular interest due to their strong relevance in different microalgal industrial processes and their impact on environmental sustainability. Even though Monod equations are conventionally used to describe microalgal growth, the Droop model better describes the growth rate as a function of the internal quota of the limiting nutrient, since microalgae still exhibit growth when dissolved nutrients are exhausted (Bernard, 2011). The main hurdle when working with the Droop model is related to the high number of measurements and experiments needed to reliably determine the kinetic parameters, resulting in long and time-consuming experimental campaigns (Turetta et al., 2022). In addition, working in steady state continuous cultures, which are more suitable for microalgal production, is even more time intensive, up to 11 weeks for each nutrient.

Based on those reasons, in this work, we propose a Model-Based Design of Experiments (MBDoE) approach to optimize the quality of experimental data while minimizing time and resources consumption (Franceschini and Macchietto, 2008). Accordingly, the objective is to exploit MBDoE to design experiments for maximising the information

required for accurate parameter estimation of the Droop model for continuous microalgal cultivation from measurements of biomass and nitrogen concentrations only. After the introduction of the Droop model and the definition of the mathematical foundation of MBDoE, a dynamic sensitivity analysis will be presented to highlight possible issues in parameter estimation. To show the capability of MBDoE of designing experiments for accurate parameter estimation, iterative *in silico* experiments were performed, according to the MBDoE procedure.

## 1. Model and methodology

### 1.1. The Droop model

In the Droop model, the specific biomass growth rate  $\mu$  ( $\text{d}^{-1}$ ) is a function of both light intensity  $I$  ( $\mu\text{mol m}^{-2} \text{s}^{-1}$ ) and the internal quota  $q$  ( $\text{g}_\text{N} \text{g}_\text{x}^{-1}$ ) of the limiting nutrient: in this work, nitrogen is considered as the limiting nutrient. Thus, the specific biomass growth can be expressed as (Bernard 2011):

$$\mu = \mu_{max} \frac{I}{I + K_I \left( \frac{I}{I_{opt}} - 1 \right)^2} \left( 1 - \frac{q_{min}}{q} \right) - k_d \quad , \quad (1)$$

where  $\mu_{max}$  ( $\text{d}^{-1}$ ) is the maximum specific growth rate of the microorganism,  $K_I$  ( $\mu\text{mol m}^{-2} \text{s}^{-1}$ ) is the half saturation constant of light response curve,  $I_{opt}$  ( $\mu\text{mol m}^{-2} \text{s}^{-1}$ ) is the optimal light intensity for growth,  $k_d$  ( $\text{d}^{-1}$ ) is the specific decay rate and  $q_{min}$  ( $\text{g}_\text{N} \text{g}_\text{x}^{-1}$ ) is the minimum internal cell quota. Light intensity varies along the depth of the reactor and for a flat plate reactor it can be expressed through the Lambert-Beer law:

$$I = I_0 \exp(-k_a c_x z) \quad , \quad (2)$$

where  $I_0$  ( $\mu\text{mol m}^{-2} \text{s}^{-1}$ ) is the incident light intensity,  $k_a$  ( $\text{m}^{-2} \text{g}_\text{x}^{-1}$ ) is the biomass light absorption coefficient,  $c_x$  ( $\text{g}_\text{x} \text{m}^{-3}$ ) is biomass concentration, and  $z$  (m) is the spatial coordinate along the thickness of the reactor. The biomass growth rate  $r_x$  ( $\text{g}_\text{x} \text{m}^{-3} \text{day}^{-1}$ ) can be expressed as

$$r_x = \mu c_x \quad , \quad (3)$$

and the average biomass growth rate  $r_x^{avg}$  ( $\text{g}_\text{x} \text{m}^{-3} \text{day}^{-1}$ ) along the culture depth  $W$  (m) is calculated as:

$$r_x^{avg} = \frac{1}{W} \int_0^W r_x dz \quad . \quad (4)$$

The nutrient uptake rate  $\rho$  ( $\text{g}_\text{N} \text{g}_\text{x}^{-1} \text{d}^{-1}$ ) depends both on the external dissolved nutrient concentration  $c_N$  ( $\text{g}_\text{N} \text{m}^{-3}$ ) and on its internal quota (Turetta et al., 2022):

$$\rho = \rho_N \frac{c_N}{K_N + c_N} \left( 1 - \frac{q}{q_{max}} \right) \quad , \quad (5)$$

with  $\rho_N$  ( $\text{g}_\text{N} \text{g}_\text{x}^{-1} \text{d}^{-1}$ ) as the maximum nutrient uptake,  $K_N$  ( $\text{g}_\text{N} \text{m}^{-3}$ ) as the nutrient half saturation constant and  $q_{max}$  ( $\text{g}_\text{N} \text{g}_\text{x}^{-1}$ ) as the maximum nutrient quota in the biomass. For a continuous photobioreactor material balances for biomass are:

$$\frac{dc_x}{dt} = \mu(q)c_x - Dc_x \quad , \quad (6)$$

$$\frac{dc_N}{dt} = -\rho(c_N)c_x - Dc_N + Dc_{N,in} \quad , \quad (7)$$

where  $D$  ( $\text{d}^{-1}$ ) is the dilution rate and  $c_{N,in}$  ( $\text{g}_\text{N} \text{m}^{-3}$ ) the inlet nitrogen concentration. The internal quota is maintained by nutrient uptake from external medium and consumed by cells duplication and thus its material balance can be expressed as:

$$\frac{dq}{dt} = \rho(c_N) - \mu(q)q \quad . \quad (8)$$

In this work a MBDoE was implemented for planning of experiments to estimate Droop parameters  $K_N$ ,  $q_{max}$ ,  $q_{min}$ ,  $\rho_N$ , while values of light parameters  $\mu_{max}$ ,  $K_I$ ,  $I_{opt}$ ,  $k_d$ ,  $k_a$  were retrieved from Saccorda et al. (2022).

### 1.2. MBDoE

MBDoE is a mathematical technique that, building on *a priori* knowledge embedded in a mechanistic mathematical model, aims at obtaining the maximum information from an experimental apparatus that will yield the most informative data for modelling activity, in a statistical sense, avoiding waste of time and resources (Franceschini and Macchietto, 2008). This approach consists of exploiting the model equations and current parameter values to estimate the information content of the experiment by evaluating a certain objective function. An optimisation framework is then applied to the solution of the resulting numerical problem. The methodology was successfully applied in several fields, including model identification in microalgae systems (Bernardi et al., 2016).

For MBDoE application, Equations 1-8 can be rearranged in a compact form:

$$\begin{cases} \mathbf{f}(\dot{\mathbf{x}}(t), \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}, \mathbf{k}, t) = 0 \\ \hat{\mathbf{y}}(t) = \mathbf{h}(\mathbf{x}(t)) \end{cases} \quad (9)$$

where  $\mathbf{x}$  is the  $N_x$ -dimensional vector of all the model state variables and,  $\hat{\mathbf{y}} = [c_x, c_N]^T$  is the  $N_y$ -dimensional vector of the measured response variables,  $\mathbf{u} = [D, c_{N,in}]^T$  is the  $N_u$ -dimensional set of time-varying inputs to the process,  $\boldsymbol{\theta} = [K_N, q_{max}, q_{min}, \rho_N]^T$  is the  $N_\theta$ -dimensional vector of parameters to be determined, determined,  $\mathbf{k} = [\mu_{max}, K_I, k_d, k_a, I, W]^T$  the  $N_k$ -dimensional vector of constants of the model and  $0 < t < \tau$  is time, with  $\tau$  as the duration of the experiment. The symbol  $\hat{\cdot}$  is used to define the estimate of a variable (e.g.  $\hat{\boldsymbol{\theta}}$  contains the best available parameters estimation). Variable measurements are collected at time instant defined in the  $N_{sp}$ -dimensional vector  $\mathbf{t}_{sp}$ . Each experiment is defined by the experiment design vector  $\boldsymbol{\phi} = [\mathbf{u}(t), \mathbf{t}_{sp}, \tau]^T$ .

A key point of MBDoE is the formulation of a metric for the measurement of experimental information content. In mathematical terms, decreasing the size of the inference regions of the model parameters allows to improve parameter precision. This is equal to decrease the value of the elements of the parameter variance-covariance matrix  $\mathbf{V}$  defined as (Franceschini and Macchietto, 2008):

$$\mathbf{V}(\hat{\boldsymbol{\theta}}, \boldsymbol{\phi}) = \left[ \sum_r^{N_y} \sum_s^{N_y} \tilde{\sigma}_{rs} \mathbf{Q}_r^T \mathbf{Q}_s \right]^{-1}, \quad (10)$$

with

$$\mathbf{Q}_r = \left. \frac{\partial \hat{y}_r}{\partial \theta_m} \right|_{t_l} \quad l = 1, \dots, n_{sp} \quad m = 1, \dots, N_\theta \quad (11)$$

$\mathbf{Q}_r$  is the  $(n_{sp} \times N_\theta)$  dynamic sensitivity matrix of the  $r^{th}$  response. The term  $\tilde{\sigma}_{rs}$  in Equation 10 is the  $(r, s)$  element of the inverse of variance-covariance matrix of the experimental error. The final designed experiment  $\boldsymbol{\phi}^{opt}$  is the solution the following minimisation/maximisation problem:

$$\boldsymbol{\phi}^{opt} = \operatorname{argmin} \{ \det[\mathbf{V}(\hat{\boldsymbol{\theta}}, \boldsymbol{\phi})] \} = \operatorname{argmin} \{ \det[\mathbf{V}^{-1}(\hat{\boldsymbol{\theta}}, \boldsymbol{\phi})] \} \quad (12)$$

defined with the so-called D-criterion (Franceschini and Macchietto, 2008). In other words, the optimum experiment is the one that minimises the parametric uncertainty or maximises the experimental information content.

## 2. Results

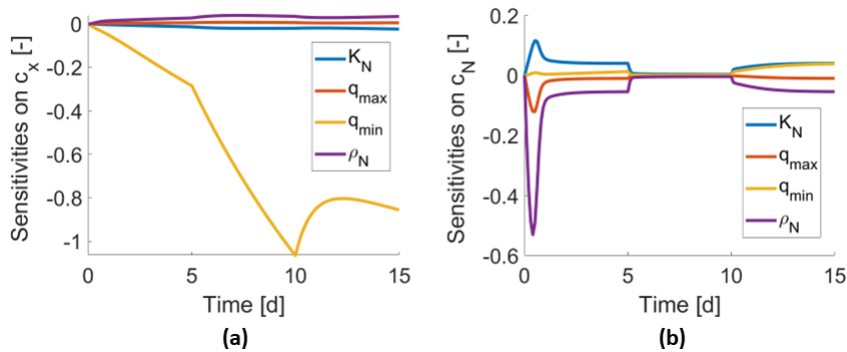
### 2.1. Sensitivity analysis

A local dynamic sensitivity analysis was performed to highlight possible parameter identifiability issues. The dynamic trajectories of measured variables were perturbed by one-at-a-time variations of the parameters to be estimated. Nominal values of the parameters to be estimated are shown in Table 1. Additional parameters and constants (fixed light parameters from Saccardo et al. (2022), light intensity and reactor thickness)

are shown in Table 2. The dynamic sensitivity  $s_{i,p}(t)$  for the  $i^{th}$  variable perturbed by the  $p^{th}$  parameter at time  $t$  was computed as:

$$s_{i,p}(t) = \frac{\hat{y}_i(t) - \hat{y}_{i,p}^{pert}(t)}{\Delta\% \hat{y}_{i,REF}} \quad (13)$$

where  $\hat{y}_i(t)$  is the value calculated with the nominal set of parameters,  $\hat{y}_{i,p}^{pert}(t)$  the perturbed value calculated with the variation in the  $p^{th}$  parameter,  $\Delta\%$  is the percentage perturbation of the parameter (equal to 1% for all parameters) and  $\hat{y}_{i,REF}$  is a reference value for the normalisation. Results of the sensitivity analysis are shown in Figure 1, where the effects of step wise changes in the inputs for all the perturbed parameters on the biomass and nitrogen concentrations are reported. Figure 1(a) highlights that  $q_{min}$  is the parameter with the highest impact on the biomass concentration and thus its estimation is expected to be linked to this measurement. Even though  $K_N$ ,  $q_{max}$ ,  $\rho_N$  have a minor impact on biomass concentration, Figure 1(b), shows that their main effect is on the nitrogen concentration and thus this variable is expected to be mainly responsible for their estimation. Figure 1(b) also evidences that some degree of correlation between  $q_{max}$  and  $\rho_N$  exists as well as anticorrelation between  $K_N$  and  $q_{max}$ , thus suggesting possible issues in the estimation of those parameters. However, Figure 1 also suggests that measurements of biomass and nitrogen concentration alone are sufficient for precise estimation of all Droop parameters if experiments are planned adequately.



**Figure 1:** Dynamic sensitivity analysis: sensitivities of model parameters on biomass (a) and nitrogen concentration (b).

**Table 1:** Reference values of parameters of Droop model.

Parameter	Value
$K_N$	14.23 g <sub>N</sub> m <sup>-3</sup>
$q_{max}$	0.90 g <sub>N</sub> g <sub>x</sub> <sup>-1</sup>
$q_{min}$	0.005 g <sub>N</sub> g <sub>x</sub> <sup>-1</sup>
$\rho_N$	0.62 g <sub>N</sub> g <sub>x</sub> <sup>-1</sup> d <sup>-1</sup>

## 2.2. In silico experiments

*In silico* iterative experiments have been performed to inspect the capability of MBDoE to plan experiments for precise parameter estimation.

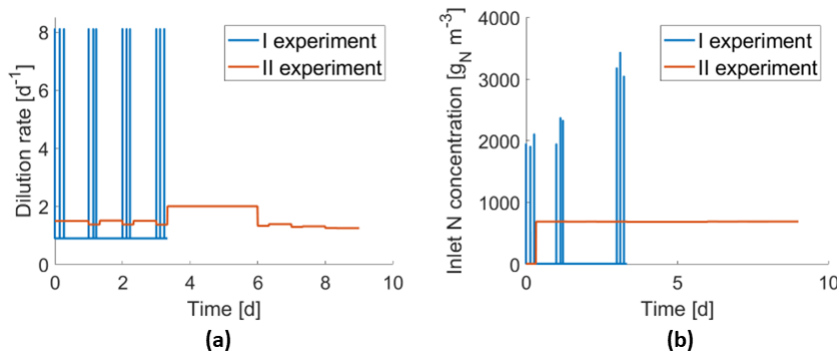
**Table 2:** Fixed light parameters (Saccardo et al., 2022). and constants for simulations

Name	Value
$\mu_{max}$	2 d <sup>-1</sup>
$K_I$	110 $\mu\text{mol m}^{-2} \text{s}^{-1}$
$I_{opt}$	405 $\mu\text{mol m}^{-2} \text{s}^{-1}$
$k_d$	0.45 d <sup>-1</sup>
$k_a$	0.09 m <sup>2</sup> g <sub>x</sub> <sup>-1</sup>
$I$	222 $\mu\text{mol m}^{-2} \text{s}^{-1}$
$W$	0.035 m

Reference parameter values of Table 1 have been perturbed by 50% and used to plan a first MBDoE. Then, a simulation using the “true” parameters of Table 1 has been performed with the input trajectories planned by the MBDoE. White noise has been then applied on these simulation results to obtain *in silico* data, employed for a subsequent parameter estimation. If the estimation was not precise the last parametric set was then used for a new MBDoE and the entire procedure was repeated until a precise estimation was achieved. It was found that two planned experiments of 5 and 9 days respectively are sufficient for a precise parameter estimation. Dynamic profiles of the first and second experiment are shown in Figure 2. Notice that in the first experiment input variables were allowed with impulse profiles. Moreover, all the input variables were allowed to vary from 9 AM to 5 PM to reflect real working conditions. Table 3 shows the initial perturbed parameters, parameters fitted after the first and second *in silico* experiment and the corresponding 95% *t*-values. All parameters from the second fit are close to those in Table 1 and have *t*-values above the reference one all above the reference *t*-value, showing high parametric accuracy. Results of the second fitting are shown in Figure 3 for both experiment I and II.

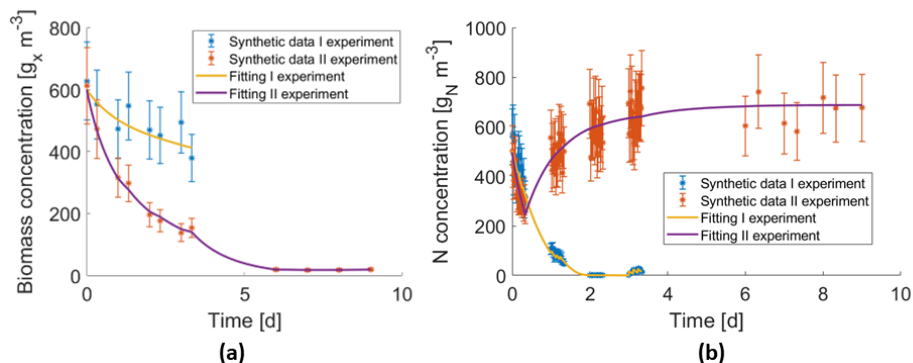
**Table 3:** Model parameters for *in silico* experiments. Reference *t*-value is equal to 1.67.

Parameter	Perturbed value	First fit	First fit 95% <i>t</i> -value	Second fit	Second fit 95% <i>t</i> -value
$K_N$ [g <sub>N</sub> m <sup>-3</sup> ]	7.115	14.74	4.9	14.73	5.01
$q_{max}$ [g <sub>N</sub> g <sub>x</sub> <sup>-1</sup> ]	0.45	0.8662	2.5	0.83	2.95
$q_{min}$ [g <sub>N</sub> g <sub>x</sub> <sup>-1</sup> ]	0.0075	0.00285	0.15	0.0055	1.72
$\rho_N$ [g <sub>N</sub> g <sub>x</sub> <sup>-1</sup> d <sup>-1</sup> ]	0.93	0.64	3.29	0.67	3.73



**Figure 2:** Dynamic profiles of input variables in the optimally designed experiments: dilution rate (a), inlet nitrogen concentration (b).





**Figure 3:** Second fitting results of synthetic data from the two optimal *in silico* experiment: biomass (a) and nitrogen concentration (b).

## Conclusions

In this work, the effectiveness of MBDoE techniques for precise parameter estimation of the Droop model was assessed. An initial local sensitivity analysis showed that the biomass concentration is mainly affected by  $q_{min}$  and the nitrogen concentration by  $K_N, q_{max}, \rho_N$ . Although some degree of correlation (between  $q_{max}$  and  $\rho_N$ ) and anticorrelation (between  $K_N$  and  $q_{max}$ ) exists, parameters are expected to be practically identifiable from measurements of biomass and nitrogen concentrations only. MBDoE was iteratively employed to plan *in silico* experiments, showing that two optimal experiments (of 5 and 9 days) can be sufficient to estimate all the parameters with accuracy. Moreover, MBDoE allowed to reduce the experimental time by over 80%, from about 77 days to 14 days for both experiments. Future work will aim at assessing the effectiveness of the proposed methodology experimentally.

## Acknowledgements

This project has received funding from the European Union's Horizon 2020 programme under the Marie Skłodowska-Curie grant agreement No. 955520 "DigitAlgaesation".

## References

- Bernard, O. (2011). Hurdles and challenges for modelling and control of microalgae for CO<sub>2</sub> mitigation and biofuel production. *Journal of Process Control*, 21(10), 1378–1389.
- Bernardi, A., Nikolaou, A., Meneghesso, A., Morosinotto, T., Chachuat, B., Bezzo, F. (2016). High-fidelity modelling methodology of light-limited photosynthetic production in microalgae. *PLoS One* 11, 1–20.
- Franceschini, G., Macchietto, S. (2008). Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science*, 63(19), 4846–4872.
- Kumar, M., Sun, Y., Rathour, R., Pandey, A., Thakur, I. S., Tsang, D. C. W. (2020). Algae as potential feedstock for the production of biofuels and value-added products: Opportunities and challenges. In *Science of the Total Environment* (Vol. 716). Elsevier B.V.
- Olaizola, M. (2003). Commercial development of microalgal biotechnology: from the test tube to the marketplace. *Biomolecular Engineering*, 20(4–6), 459–466.
- Saccorda, A., Bezzo, F., Sforza, E. (2022). Microalgae growth in ultra-thin steady-state continuous photobioreactors: assessing self-shading effects. *Front. Bioeng. Biotechnol.* 10, 1–9.
- Turetta, M., Barbera, E., Trentin, G., Bertucco, A., Sforza, E. (2022). Modeling the production of cyanophycin in *Synechocystis* sp. PCC 6803 cultivated in chemostat reactors. *Bioresource Technology Reports*, 19, 101132.

# Enabling circularity: Analysis of factors influencing MSW sorting behaviour in Central Norway

Tuva Grytli,<sup>a</sup> Cansu Birgen,<sup>b</sup>

<sup>a</sup> SINTEF Industry, P.O. Box 4760 Torgarden, 7465 Trondheim, Norway

<sup>b</sup> SINTEF Energy Research, P.O. Box 4760 Torgarden, 7465 Trondheim, Norway

## Abstract

A comprehensive waste composition analysis was undertaken in 2021, covering ten municipal waste companies in Central Norway. The results are used in a multiple linear regression model to explore factors influencing the municipal solid waste sorting behaviour. The aim of the study is to better understand what affects waste sorting, to aid in the progress of achieving circular economy targets. The share of food waste and mixed waste in the mixed waste bin is analysed considering three variables: Food waste collection system, glass and metal packaging waste collection system and type of settlement. The waste collection systems are found to have a significant effect on the sorting behaviour, where the *possibility* for sorting and the *user-friendliness* and *transparency* of the sorting system are important factors. Type of settlement is not found to have a significant effect on the sorting behaviour. We recommend further in-depth studies to confirm and expand on the findings in this study, and to better utilise the potential of waste composition analyses to enhance understanding of waste sorting behaviour. This understanding is crucial to reach a more circular waste system.

**Keywords:** Municipal solid waste, sorting behaviour, waste management, regression analysis, circular economy.

## 1. Introduction

To achieve a more circular future, municipal solid waste (MSW) management plays a key role, as large quantities of materials with great potential for reuse and recycling are handled. The 2021 Norwegian material recovery rate for MSW was 43 % (Statistics Norway 2022), far from the 2025 target in line with the EU of 55 % (Directive (EU) 2018/851 2018). Effective sorting at source is vital to increase material recovery, and there is room for improvement as more than 60 % of the waste in the mixed waste bin is recyclable materials (Fagerheim et al 2021). Effective waste strategy planning relies on good data (Thomas 2004), and this paper aims to contribute by studying the effects of various factors on the sorting behaviour.

This work was initiated to test an assumption that there is a difference in MSW sorting behaviour between different size settlements, i.e., cities versus more sparsely populated areas. To test this assumption, consistent and comparable data was needed. In 2021 the largest waste composition analysis in Norway to date was carried out, for the circular waste cluster CIVAC (Fagerheim et al 2021). The analysis was carried out for eight MSW companies in Central Norway, further including results from waste composition analyses performed in the same time period by two other MSW companies in the region. The resulting analysis covers an area of more than 750 000 inhabitants, and more than ten tonnes of waste was analysed. This data set made it possible to compare variations

between cities and villages/towns, as well as variations between different sorting systems, on a sample area basis.

## 2. Methodology

Using the results from the waste composition analysis, a multiple linear regression analysis was carried out, to study what significantly affects waste sorting.

The waste composition analysis investigated the contents of the mixed waste (MW) bin, identifying the share (weight percentage) of the 11 fractions shown in Table 1. The share of the fraction was used as dependent variable, and the regression was run for each fraction. Using a 5 % significance level, significant results were found for five fractions (in italic in Table 1). The focus of this study is the two largest fractions: Food waste including paper towel (FW) and MW which comprises the correctly sorted MW.

Table 1 Overview of dependent variables: waste fractions in the mixed waste bin.

Waste fractions
Paper and cardboard
<i>Food waste incl. paper towel</i>
Garden waste
Bags for waste
<i>Plastic packaging</i>
<i>Glass packaging</i>
<i>Metal packaging</i>
Other metal
Recyclable textiles
Hazardous waste and WEEE
<i>Mixed waste</i>

Twenty distinct sample areas were analysed, nine in cities and 11 in smaller settlements. 15 of the areas have separate collection of FW, and nine have kerbside collection of glass and metal packaging waste (G&M). One of the two additional MSW companies included reported only the totals, which cover eight sample areas, but in our analysis only counts as one sample area: a city without separate FW and kerbside G&M collection. See Table 2 for an overview of the MSW companies' solutions for sorting at source. This methodology made it possible to isolate three independent variables for the regression analysis, shown in Table 3. As the independent variables are categorical, they were all initially coded as dummy variables (Wooldridge 2009).

Table 3 Overview of MSW companies' source sorting solutions: basis for independent variables.

	MSW 1	MSW 2	MSW 3	MSW 4	MSW 5	MSW 6	MSW 7	MSW 8	MSW 9	MSW 10
Mixed waste	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Food waste	Green	Red X	Green	Green	Green	Green	Green	Green	Red X	Red X
Paper/cardboard	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Plastic	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Glass/metal	Green	Yellow	Green	Yellow	Green	Yellow	Yellow	Yellow	Green	Yellow
Kerbside collection	Central coll./recycling point									
Optibag	No separate collection: X									

Table 2 Overview of independent variables and their coding in the model.

Independent variables	Short name	Values		
		-1	0	1
Separate FW collection	sepFW	No separate collection	Optibag	Separate FW bin
Kerbside G&M collection	kerbG&M		Central recycling point or civic amenity site	Kerbside G&M collection
Settlement			Town/village	City

For separate FW collection, two systems were represented in the sample areas: A separate bin, and the optibag system. In the optibag system, several waste fractions are sorted in colour-coded bags and collected in the same bin, to be machine sorted centrally. Using a

scatter plot, in Figure 1, we found that the results from the optibag system lie between no separate collection and a separate FW bin. To corroborate this finding, the share of FW in the MW bin for a city which also uses the optibag system, is included in the figure as *Alt. optibag*. The coding of the dummy variable was changed to linear, increasing the explanatory power of the model ( $R^2$ ) by 10 percentage points.

This study has a small number of observations for a multiple regression analysis, which can increase uncertainty. This is because it was carried out at a sample area level rather than a household level. A sample area is however aggregated waste from typically 30-40 households, and we expect that the analysis represents around 600-800 households. Each bin collected is a snapshot in time, which may not be representative of the households' average waste composition. Aggregating into sample areas may thus make a more representative waste composition, but for an area rather than a household. As we do not have any information about each household, the additional resolution would give limited value, as variations cannot be explained. On the other hand, using data from only one analysis will ensure consistent data collection, reducing other sources of uncertainty. Differences between waste composition analyses can include analysis methods, such as sample size and location, or types and number of waste components (Dahlén & Lagerkvist 2008). In conclusion, the consistent data collection was considered to compensate for the small number of observations, and the study deemed a good contribution to advance the understanding of how external factors can affect MSW source-sorting results. To strengthen the findings and increase generalisability, similar studies should be carried out.

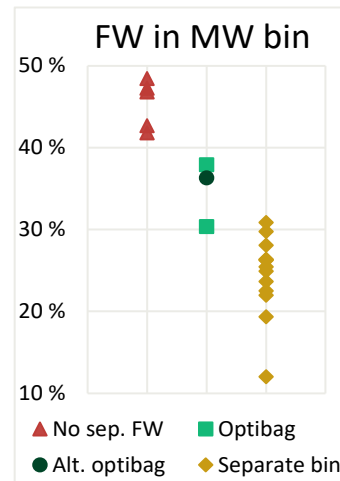


Figure 1 Actual shares of FW in the MW bin, by source sorting solution.

### 3. Results and discussion

The results from the linear regression model for FW are shown in Table 4. The model explains ( $R^2$ ) 85 % of the variation in the fraction, but only separate FW collection has a significant impact ( $p < 0.01$ ) on the amount of FW in the MW

Table 4 Results for food waste fraction regression model.

Dependent variable: Food waste					
$R^2$	85 %			95 % confidence intervals	
Obs.	20	Coeff.	P-value		
(Intercept)		35 %	< 0.01	32 %	39 %
sepFW		-9.7 %	< 0.01	-12.5 %	-7.0 %
kerbG&M		-3.3 %	0.13	-7.6 %	1.1 %
Settlement		1.7 %	0.45	-3.0 %	6.5 %

bin. With each increment in the separate FW collection variable, i.e. from no collection to optibag, or from optibag to dedicated FW bin, the share of FW in the MW bin is reduced by almost 10 percentage points. Due to how the sepFW variable is coded, this means the FW share goes from 45 % with no collection, to 35 % with optibag, to 25 % with dedicated bin, which amounts to an average 25 % reduction per step. The confidence intervals indicate a 95 % likelihood that the real reduction is between 7 and 12.5 percentage points. Kerbside collection of glass and metal packaging waste and the settlement type do not show any significant effects on the amount of FW in the MW bin.

The results from the linear regression model for correctly sorted MW are shown in Table 5. The model explains 82 % of the variation in the fraction, and both separate FW collection and kerbside G&M collection have significant effects ( $p < 0.01$ ) on the dependent variable. With each increment in the separate FW collection variable, the share of correctly sorted MW increases by almost 7 percentage points, going from 24 % with no collection, to 31 % with optibag, to 37 % with dedicated FW bin. This amounts to an increase of about 25 % per step on average. The confidence intervals indicate an increase between 4 and 9.5 percentage points. For kerbside collection of G&M, the share of correctly sorted MW increases by 7.5 percentage points, from 31 % to 38.5 %, which amounts to a 24 % increase. There is a 95 % probability that the increase is between 3 and 12 percentage points.

Table 5 Results for mixed waste fraction regression model.

Dependent variable: <b>Mixed waste</b>					
R <sup>2</sup>	82 %			95 % confidence intervals	
Obs.	20	Coeff.	P-value		
(Intercept)		31 %	< 0.01	27 %	34 %
sepFW		6.7 %	< 0.01	3.9 %	9.4 %
kerbG&M		7.5 %	< 0.01	3.1 %	11.8 %
Settlement		-4.4 %	0.07	-9.1 %	0.3 %

The increase in the share of correctly sorted MW from kerbside collection of G&M is larger than can be explained by a reduction in G&M alone, as G&M are small weight fractions. This suggests that other factors may be at play, and indeed we found a significant effect on the share of plastic packaging waste in the MW bin: When there is kerbside G&M collection there is less plastic packaging waste in the MW bin. This interrelation cannot be explained by our dataset, but Mikkelsen (2017) also found a connection between sorting of different waste fractions, where better sorting in one fraction coincides with better sorting in other fractions. He points at communication and information as important factors behind this, further discussed below for other variables.

Despite size of settlement not having a statistically significant effect on the share of correctly sorted MW, it is not far off, with a p-value of 0.07. Mikkelsen (2017) showed that typical city traits, such as limited indoor and outdoor space, more households likely to have small children (citizens in 20s and 30s age groups), more citizens with foreign background and more lower income households, are related to a lower degree of source sorting. The results from Mikkelsen (2017) were further analysed in Fagernæs (2018), and better communication differentiated by target group is presented as a key measure to improve poor source sorting. One could hypothesize that targeted communication is easier and social control stronger in a smaller community. The insignificant effect of settlement type on correctly sorted MW indicates better sorting in smaller communities.

Based on the results we see that two factors are important for improved MSW sorting at source: (1) The possibility to sort at source. For FW, we see that having separate FW collection reduces the amount of FW in the MW bin with both systems. This is in many ways obvious, as the only option if there is no separate FW sorting, is to dispose of FW in the MW bin (or home composting). (2) The convenience of the sorting system. We see that kerbside pickup of G&M reduces the amount of G&M in the MW bin, with significant reductions of the glass and metal packaging waste fractions, as well as a significant increase in correctly sorted MW. The easier it is for the consumer to sort at source, and the less transport required to use the correct waste bin, the more likely it is that they sort better at source.

However, while a more convenient sorting system means more of the fraction is collected separately, studies have shown that it also leads to more missorting in the case of glass and metal packaging. Syversen et al (2019) show that while the amount of sorted G&M waste collected with a kerbside system is more than 25 % higher, the share of waste that is missorted more than doubles, from 4 % to almost 10 %. This means that the quality of the sorted fraction is worse, and it will need more processing before recycling.

A third factor, which was outside the scope of this study, is the transparency of the waste system and the perceived fate of the waste. For FW, we see that the optibag system gives inferior results compared to a dedicated bin collected separately. Based on discussions with industry partners in the CircWtE project, we believe the main reason for this is that the inhabitants do not know enough about how the waste is further treated downstream. When several waste fractions are collected in one bin, they believe all the waste goes to the same place, most likely to incineration, hence perceiving it as less important to sort correctly. Industry partners reported that they see the same effect when several bins are collected with the same multi-compartment truck (CircWtE workshop, 31 May 2022, Trondheim, Norway). It is however worth noting that the optibag system is a space saving solution, requiring a single bin for four separate fractions. In densely populated areas where space is limited, this can enable source sorting of more fractions. This does not apply to indoor space limitations though, where multiple bins are still needed. The optibag system also enables collection of four fractions with one single-compartment truck, which can be beneficial in sparsely populated areas with long transport distances.

The type of settlement was not found to be a significant factor for the MSW sorting at source in this study. This may be rooted in how the type of settlement was determined. As we did not have detailed knowledge about where the waste was collected, sample areas were defined as city or town/village based on the name of the settlement. Most Norwegian cities are small and not densely populated, implying that the differences between a city and a town may be minor. Demographic traits also vary between neighbourhoods and different degrees of urbanization, meaning the sample areas from cities may not represent typical city traits. Without detailed knowledge about the locations, we were not able to analyse this further.

Despite interesting findings, a limited number of variables were available for analysis. Demographic variables are found to have significant effects on degrees of sorting at source (Mikkelsen 2017), but the dataset did not allow for such variables. With only two sample areas per MSW company, quantitative analysis of differences between companies was impossible, despite anecdotal indications that e.g. communication strategies can affect results significantly. In the presentation of the waste composition analysis, one MSW company was highlighted for outstanding FW sorting results, despite launching separate sorting only recently. This was explained by an active communication strategy, combined with a strict policy of not collecting wrongly sorted waste. Bad sorting is photodocumented, and the documentation including an explanation of why the waste is not collected is sent to the customer immediately. According to the MSW company, this active communication led to good sorting results as well as high customer satisfaction, which was validated in that they have the two highest shares of correctly sorted MW of all 20 sample areas (CIVAC: Presentation of results from waste composition analysis, 7 December 2021, Stjørdal, Norway). In summary, many variables other than sorting systems affect the results of MSW sorting, and these should be considered when

designing systems to enable effective sorting at source. As for the effect of the sorting systems, comparable results have been found in a similar study by the authors (Grytli & Birgen 2023), strengthening the findings of this study.

#### 4. Conclusions

The results from our regression analyses show that the possibility to sort at source, as well as the user-friendliness and transparency of the waste system, are significant factors impacting MSW sorting behaviour. We recommend further in-depth studies of waste composition analyses to confirm and expand on these findings.

Despite no significant effects from type of settlement in this study, demographic variables have been found to have an effect in other studies. Targeted communication and sanction systems also appear to impact sorting results. These types of variables should be studied further, to aid MSW companies in developing strategies for increased sorting.

This study intends to better utilise the potential of waste composition analyses and the results obtained contribute to a deeper understanding of sorting behaviour and its drivers. The study can help MSW companies better design and select effective collection systems and policymakers in implementing measures to improve sorting and increase circularity.

#### Acknowledgements

This work is part of the CircWtE project co-funded by industry and public partners and the Research Council of Norway under the SIRKULÆRØKONOMI program (CircWtE - 319795). The authors thank the personnel at CIVAC for providing data and valuable information on the waste management system.

#### References

- L. Dahlén, A. Lagerkvist, 2008, Methods for household waste composition studies, *Waste Management*, Vol. 28, 7, 1100-1112, <https://doi.org/10.1016/j.wasman.2007.08.014>.
- Directive (EU) 2018/851, 2018, Directive (EU) 2018/851 of the European Parliament and of the Council of 30 May 2018 amending Directive 2008/98/EC on waste (Text with EEA relevance)
- A.B. Fagerheim, E.L. Mikkelsen, S. Bjørnerud, 2021, *Avfallsanalyser restavfall Midt-Norge 2021*, <https://civac.no/dokumentasjon/hvilke-ressurser-finnes-i-restavfallet/>, acc. 2022-11-29
- C.C. Fagernæs, 2018, Recycling and material recovery of plastic packaging and food waste in Oslo municipality - How to change behaviour and attitude through facilitated measures?, Master thesis, Norwegian University of Life Sciences.
- T. Grytli, C. Birgen, 2023, A Critical Look at Waste Composition Analyses: Challenges and Opportunities, Proceedings of the 33rd European Symposium on Computer Aided Process Engineering (ESCAPE33), June 18-21, 2023, Athens, Greece. (Accepted)
- E.L. Mikkelsen, 2017, Increased material recycling from household waste in Oslo Municipality – the role of socio-demographics and socio-psychological factors, Master thesis, Norwegian University of Life Sciences.
- Statistics Norway, 2022, Table 12263: Selected key figures for household waste (M) 2015 - 2021
- F. Syversen, K. Kirkevaag, S. Bjørnerud, 2019, Renere råvarer. Case: Optimalisering av verdikjeden for glass- og metallemballasje, <https://avfallnorge.no/fagomraader-og-faggrupper/rapporter/renere-r%C3%A5varer>, accessed 2022-11-29
- C. Thomas, 2004, Effective use of data in waste strategy planning in the UK. In: ISWA World Environment Congress, 17-21 Oct 2004, Rome, Italy.
- J. Wooldridge, 2009, *Introductory Econometrics: A Modern Approach*, 4<sup>th</sup> Edition, Mason, Ohio: South Western Cengage Learning.

# A Critical Look at Waste Composition Analyses: Challenges and Opportunities

Tuva Grytli,<sup>a</sup> Cansu Birgen,<sup>b</sup>

<sup>a</sup> SINTEF Industry, P.O. Box 4760 Torgarden, Trondheim, 7465, Norway

<sup>b</sup> SINTEF Energy Research, P.O. Box 4760 Torgarden, 7465 Trondheim, Norway

## Abstract

This paper presents regression analysis results on time series data of waste composition analyses from a municipal waste company in Norway, with a discussion on methodological challenges. The aim is to investigate what affects municipal solid waste (MSW) source-sorting results, to facilitate achieving circular economy targets. The share of four MSW fractions in the mixed waste bin are studied considering five independent variables: Collection/sorting systems for food waste, glass and metal packaging waste, and mixed waste; year of analysis, and pre/post covid-19. We find that only the source sorting systems have significant impacts on the waste composition, and that user-friendliness matters. We recommend more in-depth studies to increase data quality to better utilise the potential of waste composition analyses for a more circular waste system.

**Keywords:** Municipal solid waste, sorting behaviour, waste management, regression analysis, circular economy.

## 1. Introduction

To improve circularity, recycling, and reuse, we need to improve municipal solid waste (MSW) management systems, where large quantities of materials have the potential to be recovered. The 2021 material recovery rate in Norway was 43 %, well below the 55 % target for 2025 (Statistics Norway 2020, Directive (EU) 2018/851 2018). More than 60 % of the waste in the mixed waste (MW) bin is recyclable, so understanding how households sort their waste is vital, and good data is key for effective waste strategy planning (Grytli & Birgen 2023). The aim of this study is to increase the insight into what affects sorting behaviour, to enable increased circularity in the waste treatment system.

This study is based on waste composition analysis data from a municipal waste company in Norway that serves more than 100 000 inhabitants. Waste composition analyses of the MW bin were carried out in 2016, 2017, 2019 and 2021, comprising a time series where in total almost 4 000 kg waste has been physically analysed.

## 2. Methodology

To better understand what affects waste sorting at source, the results from the waste composition analyses were investigated using multiple linear regression analysis. The number of sample areas analysed per year were two (2016-2017), three (2019) and four (2021), giving a total of 11 observations.

The waste composition analyses examine the contents of the MW bin, estimating the share, in weight percentage, of ten different waste fractions. Table 1 shows the sample area average share per fraction. A regression analysis was run for each fraction, with the



share of the fraction in the bin as the dependent variable. Significant effects were found for five fractions, and this study focuses on four of them: Food waste (FW), glass packaging waste (GW), metal packaging waste (MeW), and MW which is the correctly sorted MW.

To identify independent variables, the differences between the 11 observations were assessed. Two sample areas have separate FW collection. Kerbside glass and metal packaging waste (G&M) collection was rolled out between 2017 and 2019, and five observations took place after this roll-out.

Three sample areas have underground MW collection, where the waste is discarded in a large underground container via a waste inlet aboveground. The other eight have regular rolling bins for MW. Four observations were from late fall of 2021, i.e. after Covid-19. These differences gave rise to five potential independent variables, described in Table 2. As all independent variables (except year) are categorical, they were coded as dummy variables (Wooldridge 2009).

Table 2 Overview and coding of independent variables.

Independent variables	Short name	Values			
		0	1	2	3
Separate FW collection	<i>sepFW</i>	No separate collection	Separate FW bin		
Kerbside G&M collection	<i>kerbG&amp;M</i>	G&M to central recycling point or civic amenity site	Kerbside G&M collection		
Underground waste collection	<i>Undergr.</i>	Aboveground MW bin	Underground MW system		
Post Covid-19	<i>Covid19</i>	Before Covid-19	After Covid-19		
		1	2	4	6
Year	<i>Year</i>	2016	2017	2019	2021

There were some challenges with the identified independent variables. Most had some correlation with the time variable, as they either happened at a certain time (kerbside G&M collection and Covid-19) or were included only in later years (underground waste collection). This can cause a multicollinearity problem, as discussed later. The next challenge was a very small number of observations for a regression analysis. Each observation is however the aggregated waste from 30-40 bins. This means that the 11 sample areas represent more than 330 bins, and almost 400 households. For practical reasons this study was done on a sample area level. It is unlikely that additional resolution would have added much value, as we had no information about each household, thus could not have explained any additional variation. As each bin is a snapshot in time and may not be representative of the household's average waste composition, an average of a sample area may be considered more representative of an average waste composition. In general, a small number of observations increases uncertainty and means that the results should be interpreted carefully and may not be generalisable.

During analysis, unexpected results emerged. Close inspection of the data revealed large increases in one fraction, that could not be explained even after communication with the

Table 1 Overview and average share of waste fractions in the mixed waste bin.

Waste fractions with average shares	
Paper and cardboard	7.65 %
Food waste incl. paper towel	43.6 %
Garden waste	4.83 %
Plastic packaging	9.43 %
Glass packaging	2.94 %
Metal packaging	1.43 %
Other metal	0.98 %
Recyclable textiles	3.03 %
Hazardous waste and WEEE	1.35 %
Mixed waste	24.8 %

MSW company. Thorough debugging revealed a summation error for a subtraction of the fraction in a spreadsheet, causing a hard to spot error that had significant impacts on the results. Data from waste composition analyses is registered on paper and punched manually into a spreadsheet template. As a copy of the template was used, the mistake spread to five sample areas over two years, causing a systematic error. Left undiscovered, this systematic error would have impacted the accuracy of the model significantly. Several random errors were also discovered, where numbers were punched incorrectly. Although most random errors were too small to affect overall results, they would have reduced the model’s precision.

Due to Norway’s sparse and spread population and complex geography, MSW companies are often small, which induces several of the methodological complications discussed. Small companies have limited resources for carrying out waste composition analyses, resulting in fewer observations, and for quality assuring results. The uncovered systematic error gave a faulty picture of the company’s performance in collection, sorting and recycling, ultimately giving an incorrect base for their strategic development. This shows the benefit provided by research, as in addition to giving insight into effects of measures implemented, the MSW company’s data is improved and quality assured.

### 3. Results and discussion

The results from the linear regression model for the FW fraction are shown in Table 3. The model explains ( $R^2$ ) 91 % of the variation in the fraction, but the only variable with a significant effect ( $p < 0.01$ ) is separate FW collection. The model finds that the share of FW is reduced by 25 percentage points (pp) when there is separate FW collection, from 53 % to 28 % FW in the MW bin, amounting to a 48 % reduction. The confidence intervals indicate that there is a 95 % likelihood that the reduction is between 10 and 40 pp. None of the other variables were found to have a significant effect on the share of FW in the MW bin.

Table 3 Results for food waste fraction regression model.

Dependent variable: <b>Food waste</b>					
$R^2$	91 %	Adj. $R^2$	82 %	95 % confidence	
Obs.	11	<i>Coeff.</i>	<i>P-value</i>	<i>intervals</i>	
(Intercept)		53 %	< 0.01	43 %	63 %
sepFW		-25 %	< 0.01	-40 %	-10 %
kerbG&M		7.4 %	0.37	-12 %	27 %
Undergr.		0.3 %	0.95	-10 %	11 %
Covid19		-6.7 %	0.40	-25 %	12 %
Year		-1.5 %	0.40	-6 %	3 %

Multicollinearity (a correlation between several of the independent variables) may give unreliable results in a regression analysis. To assess whether multicollinearity was a problem in the model, the variance inflation factors (VIF) for the independent variables were analysed (Wooldridge 2009). There is no agreed-upon threshold for when multicollinearity becomes problematic, but threshold values of 5 or 10 are commonly used. The VIFs for the independent variables are never over 10, but for kerbside G&M collection, year and Covid-19, the VIF is over 5. This became a problem for the GW and MeW fractions, where results were significantly affected. The collection system for G&M has been shown to impact the sorting behaviour in other studies (Grytli & Birgen 2023, Syversen et al 2019), and it is therefore assumed to be more important than development over time, or the effect of a pandemic. To avoid the multicollinearity problem, the variables Post Covid-19 and Year were excluded from the models for the GW and MeW fractions, and a simplified model is presented.

The results for the simplified regression models are shown in Table 4 for GW and Table 5 for MeW. For GW we found that the adjusted  $R^2$  increased when running the model with only kerbside G&M collection as independent variable, indicating that the simplified model may be overfitted. Results from both runs are thus included in Table 4, and the discussion will focus on the *Only G&M* model for GW. The underlying reasons will be discussed later.

Table 4 Results for glass packaging waste fraction regression models.

Dep. variable:		Simplified model				Only G&M			
<b>Glass pack. waste</b>		R <sup>2</sup>	54 %	Adj. R <sup>2</sup>	34 %	R <sup>2</sup>	50 %	Adj. R <sup>2</sup>	45 %
Obs.	11	<i>Coeff.</i>	<i>P</i>	95 % <i>CI</i>		<i>Coeff.</i>	<i>P</i>	95 % <i>CI</i>	
(Intercept)		3.7 %	< 0.01	2.6 %	4.8 %	3.7 %	< 0.01	2.8 %	4.7 %
sepFW		0.8 %	0.51	-1.9 %	3.5 %				
kerbG&M		-2.2 %	0.04	-4.3 %	-0.1 %	-1.8 %	0.01	-3.1 %	-0.4 %
Undergr.		0.4 %	0.65	-1.6 %	2.5 %				

For GW the model explains ( $R^2$ ) 50 % of the variation in the fraction, while the MeW model explains more of the variation, with an  $R^2$  of 66 %. Only the kerbside G&M collection affects the share of G&M waste in the MW bin significantly in any of the models. For MeW the effect is more clearly significant ( $p < 0.01$ ) than for GW ( $p = 0.01$ ). The estimated reduction in MeW is 0.9 pp, from 1.6 % to 0.7 % MeW in the MW bin, amounting to a 55 % reduction. The confidence intervals indicate a true reduction between 0.3 and 1.5 pp. For GW the reduction is 1.8 pp, from 3.7 % to 1.9 % GW in the MW bin, or a 48 % reduction, with a true reduction estimated between 0.4 and 3.1 pp. The findings are in line with Syversen et al (2019), where the increase in collected metal is larger than the increase in collected glass when introducing a kerbside G&M collection. They found an increase in the amounts sorted at source of 11 % for glass and 157 % for metal. As metal packaging is lighter than glass packaging, we can assume that a smaller reduction in the share of metal in the MW can account for a larger increase in the amount collected from sorting at source.

For the MW, the results from the full regression model are shown in Table 6. The model explains ( $R^2$ ) 90 % of the variation in correctly sorted MW in the MW bin. The only significant effect is

Table 5 Results for simplified metal packaging waste fraction regression model.

Dependent variable: <b>Metal packaging waste</b>					
R <sup>2</sup>	66 %	Adj. R <sup>2</sup>	51 %	95 % <i>conf. intervals</i>	
Obs.	11	<i>Coeff.</i>	<i>P-value</i>		
(Intercept)		1.6 %	< 0.01	1.3 %	1.9 %
sepFW		0.7 %	0.07	-0.1 %	1.4 %
kerbG&M		-0.9 %	< 0.01	-1.5 %	-0.3 %
Undergr.		0.4 %	0.18	-0.2 %	0.9 %

Table 6 Results for mixed waste fraction regression model.

Dependent variable: <b>Mixed waste</b>					
R <sup>2</sup>	90 %	Adj. R <sup>2</sup>	80 %	95 % <i>conf. intervals</i>	
Obs.	11	<i>Coeff.</i>	<i>P-value</i>		
(Intercept)		21 %	< 0.01	13 %	29 %
sepFW		18 %	0.01	6 %	29 %
kerbG&M		-2.7 %	0.66	-18 %	12 %
Undergr.		-0.02 %	1	-8 %	8 %
Covid19		9.1 %	0.16	-5 %	23 %
Year		-0.5 %	0.71	-4 %	3 %

separate FW collection ( $p=0.01$ ), which leads to an increase in the share of correctly sorted waste in the MW bin of 18 pp, from 21 % to 39 %, or an 83 % increase. The confidence intervals indicate a real increase in correctly sorted MW with separate FW collection of 6 - 29 pp.

We see high explanatory powers ( $R^2$ ) and large effects for both the FW and MW fractions, when separate FW collection is the significant variable. Two factors help explain this: (1) FW is heavy, constituting by far the largest fraction in the MW bin for households without separate FW collection, as shown in Figure 1. With separate FW collection, the correctly sorted MW is the largest fraction. (2) If there is no separate FW collection, the only option is to dispose of FW in the MW bin (or home composting). This implies that

when we say *correctly sorted MW*, it does not mean that households without separate FW collection are doing anything wrong by discarding their FW in the MW bin, but rather that there should be separate FW collection, so this fraction can be recycled.

On the other hand, we saw how kerbside G&M collection did significantly affect the shares of GW and MeW in the MW bin, but not the share of correctly sorted MW. This is likely due to GW and MeW being small fractions in the MW bin. Effects may thus be obscured by variations in the larger fractions, such as food, plastic, and paper/cardboard. This shows the importance of looking at each fraction separately, to catch individual variations also in the smaller fractions.

Due to the weight of FW, we expected separate FW collection to impact the results of the shares of GW and MeW in the MW bin. As the heavy fraction is removed, the relative shares of all other fractions should increase. For MeW we saw this effect be close to significant, but we did not observe this for GW in the simplified model. The effect of the kerbside G&M collection was also less significant for GW in the simplified model. Through in-depth analysis of the background data, the reason for these results was found to be one outlier data point in a sample area after the rollout of kerbside G&M collection. The outlier lifts the sample area average above the lowest value before the rollout, and above the average of the areas with separate FW collection. This shows how the small number of sample areas can inflate the impact of noise from outliers. It is also a challenge when working with waste composition rather than absolute amounts. We found that by subtracting the FW from the waste composition, the problematic sample area's GW share falls below the lowest value before the rollout. This is because the earlier sample area had a very large share of FW, and the later (problematic) sample area a share below average, thus affecting the relative share of GW in the opposite direction. Both of these problems caused noise that made the simplified GW model seem overfitted, and as a consequence the choice to further simplify the model was made.

As discussed, the small number of observations caused some difficulties. However, using data from only one MSW company yields consistent data collection, which reduces other sources of uncertainty. Different methods for the waste composition analysis, in terms of for example sample size and location, or types and number of waste components, are good

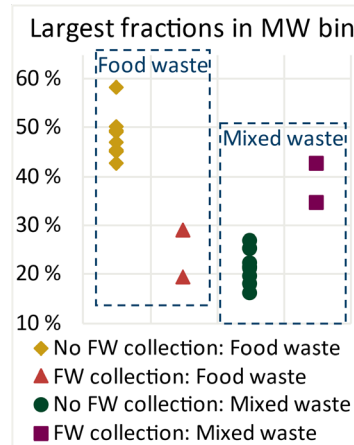


Figure 1 Shares of the two largest fractions in the mixed waste bin.

examples of such sources (Dahlén & Lagerkvist 2008). We assessed that the benefits from the data collection consistency outweigh the disadvantages of few observations and consider this study a good contribution to understanding the effects of external factors on the results of MSW sorting at source. To strengthen findings and make them more generalisable, similar analyses should be carried out for other MSW companies.

#### 4. Conclusions

The regressions show that it is mainly separate sorting and user-friendly systems that influence the sorting results in the MW bin. The more fractions it is possible and easy to sort, the better the sorting results. We found no significant impacts over time, post covid-19 or from a different type of MW bin.

There are many factors that can influence sorting at source that we were not able to analyse in this study, such as demographic variables, communication, and information, which is further discussed in Grytli & Birgen (2023). To optimize the design of MSW systems for a circular future, it is crucial to identify and understand all relevant factors, and more research should be carried out to this effect.

This study has revealed challenges related to waste composition analyses due to manual data registration. As most MSW companies in Norway are small, the capacity for quality assurance is limited, increasing the risk of errors. Through this work we were able to provide debugging and data quality improvements for the MSW company. Improving the waste composition statistics will help the company better understand how their implemented measures are working and which new ones to consider.

Besides helping the specific MSW company, this study can help similar companies better understand and design their collection systems to increase circularity. We also aim to contribute to the necessary high-quality information and data needed for effective waste strategy planning by policymakers.

#### Acknowledgements

This work is part of the CircWtE project co-funded by industry and public partners and the Research Council of Norway under the SIRKULÆRØKONOMI program (CircWtE - 319795). The authors thank the personnel at the municipal solid waste company for providing data and valuable information on the waste management system.

#### References

- L. Dahlén, A. Lagerkvist, 2008, Methods for household waste composition studies, *Waste Management*, Vol. 28, 7, 1100-1112, <https://doi.org/10.1016/j.wasman.2007.08.014>.
- Directive (EU) 2018/851, 2018, Directive (EU) 2018/851 of the European Parliament and of the Council of 30 May 2018 amending Directive 2008/98/EC on waste (Text with EEA relevance)
- T. Grytli, C. Birgen, 2023, Enabling circularity: Analysis of factors influencing MSW sorting behaviour in Central Norway, Proceedings of the 33rd European Symposium on Computer Aided Process Engineering (ESCAPE33), June 18-21, 2023, Athens, Greece. (Accepted)
- Statistics Norway, 2022, Table 12263: Selected key figures for household waste (M) 2015 - 2021
- F. Syversen, K. Kirkevaag, S. Bjørnerud, 2019, Renere råvarer. Case: Optimalisering av verdikjeden for glass- og metallemballasje, <https://avfallnorge.no/fagomraader-og-faggrupper/rapporter/renere-r%C3%A5varer> (accessed 2022-11-29)
- J. Wooldridge, 2009, *Introductory Econometrics: A Modern Approach*, 4<sup>th</sup> Edition, Mason, Ohio: South Western Cengage Learning.

# Waste valorisation within the Energy-Water-Food Nexus: A hybrid techno-geospatial optimisation approach

Mohammad Alherbawi, Sarah Namany, Maryam Haji, Gordon McKay, Tareq Al-Ansari \*

*College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar*

[\\*talansari@hbku.edu.qa](mailto:*talansari@hbku.edu.qa)

## Abstract

Waste of multiple categories such as agricultural, municipal, and industrial waste are the most abundant biomass sources worldwide. The mismanagement of such waste results in severe impacts on the environment, human health, and resources availability for future generations. However, waste can be utilised within energy, water and food sectors. As such, evaluating the optimal processing pathway, in addition to the optimal supply chain is crucial to fulfilling an adequate balance in the Energy-Water-Food (EWF) Nexus considering national priorities. In this study, 10 waste biomass categories were located over 35 locations across the Qatar using the geographic information systems (GIS), while their quantities, calorific values, and possible associated product yields were defined. Meanwhile, three biomass processing plants representing the EWF industries have been considered, including: biorefinery, desalination plant, and livestock feed developer. An optimisation model was then developed to select optimal biomass and minimise the overall transportation costs for all three industries. Three scenarios were introduced to ensure an EWF Nexus balance throughout the biomass allocation process. The scenarios included: a) equal biomass-based production amongst EWF industries, b) equal biomass distribution, and c) equal biomass-based sales. The model was able to allocate biomass resources from almost all sites at different utilisation percentages. Biomass transportation cost was minimised to 1.98 \$/t in scenario (b), where 0.48 M t/y of biomass were utilised, 49.8 M t/y of products were generated and sold at 82.5 M \$/y. However, scenario (c) achieved higher utilisation rate of biomass at 1.49 M t/y, transported at a minimised cost of 4.33 \$/t. The established model provides an insight on advanced decision-making approaches for optimal waste biomass valorisation to meet the growing demand on EWF resources.

**Keywords:** EWF nexus; Optimisation; GIS; Decision-making; Supply chain; Qatar.

## 1. Introduction

The Energy-Water-Food (EWF) Nexus is an emerging approach that supports the security of these three vital sectors and the adequate management of their resources (Namany et al. 2019). The importance of this approach lies in its capability to trace the interlinkages between the EWF systems and subsystems and alert any possible nexus imbalances (Al-Ansari et al. 2015). Biomass presents itself as a crucial aspect of EWF Nexus systems, where the resource is mainly utilised for clean energy production, while it may significantly occupy lands and consume water at the expense of food and water sectors. Several EWF models have been presented earlier for biomass and biofuel industry.

López-Díaz et al. (2018) developed a framework to enhance the EWF resources' management efficiency for biodiesel-bioethanol industry in Mexico. The model targeted the generation of biofuel to fulfill the market demand at a high profit, while minimising water consumption. Whereas Moioli et al. (2018) developed a "Nexus Index" to examine different EWF resources needed for the production of 1 unit of energy. Moreover, Alherbawi et al. (2021) evaluated 11 processing pathways of biomass to select an optimal route that maximises profit and energy, while minimising emissions, lands and water consumption via genetic algorithm.

In this study, a novel framework was developed to ensure an EWF Nexus balance. A hybrid techno-geospatial approach was developed to allocate local biomass resources for different EWF industries, maximise the production and minimise the gate-cost of biomass. ArcGIS was utilised to map the biomass supplying sites and EWF representative industries and to define the distances amongst them.

A mathematical optimisation model was then developed to optimally allocate biomass supplying sites for each industry, given their location, quantity, calorific values, and possible processing pathways. Three EWF balancing scenarios were studied: a) equal EWF production, b) equal biomass distribution, and c) equal biomass-driven sales.

## 2. Methodology

### 2.1 Data Collection and Mapping

The required data was adapted from authors' earlier study (Alherbawi et al. 2022), where geospatial, quantitative, and qualitative data (i.e., heating value) for 10 different waste categories in Qatar were re-mapped at 35 sites in ArcGIS (10.7.1) using Universal Transverse Mercator (UTM) coordinates system.

In addition, three industries representing energy, water and food sectors were selected and located on the map, including an earlier proposed biorefinery site (Alherbawi et al. 2022), a desalination plant, and a livestock feed developing industry.

The biorefinery was assumed to produce syngas via gasification of biomass at a yield of 65% (Alherbawi et al. 2020). Whereas the desalination plant was assumed to be equipped with a biomass-based combined heat and power plant (CHP) at 34% net efficiency (IEA 2007) and consumes 11.4 MJ/m<sup>3</sup> (Rayburn 2012). Nevertheless, the feed developer was assumed to utilise food waste and agro-waste only for the manufacturing of livestock feed with a 75% nutrients recovery. All three plants operate for 8000 hours a year with a biomass feed capacity of 30 tonnes per hour. Moreover, the distances between the 35 biomass supply sites and the 3 industries were calculated in ArcGIS along with the corresponding transportation costs.

### 2.2 Optimisation Model Development

The optimisation model aims at minimising the overall biomass transportation costs to the 3 industries, while achieving an EWF Nexus balance via optimal biomass allocation. A summary of the mathematical model is presented in Table 1. Three constraints scenarios were introduced to restrict the model for the achievement of EWF Nexus balance. Scenario (a) leads to the production of equal quantities (i.e., syngas, water, and livestock feed) in (tonnes/y). While scenario (b) distributes biomass equally between the EWF industries, within the given capacities. Finally, scenario (c) achieves equal gross product sales, given that the maximum possible biomass quantity is valorised. The problem was solved using Microsoft Excel solver.

**Table 1:** Mathematical optimisation model for optimal biomass allocation.

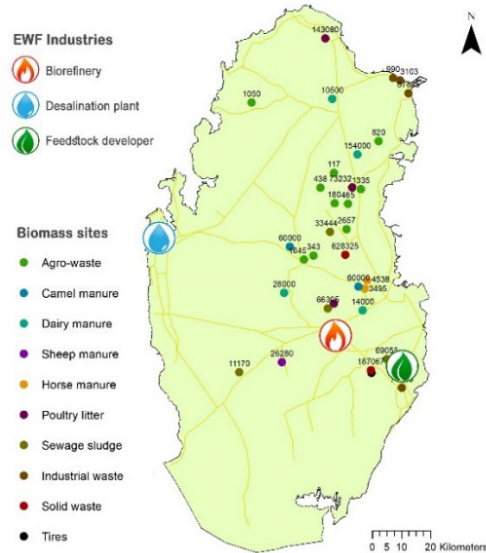
Function	Description
<b>Objective function:</b>	
$MIN TransCost = C * \sum_{i=1}^3 \sum_{j=1}^{35} D_{ij} * \frac{Q_j}{L} * A_{ij}$ <p>C: Cost of transportation (\$/km).                      D<sub>ij</sub>: Distance (km) between biomass site (j) to industry site (i).                      Q<sub>j</sub>: Total quantity (t) of biomass at site (j).                      L: Freight load (t/y).</p>	Minimising the total biomass transportation cost.
<b>Decision variable:</b>	
A <sub>ij</sub> : Allocated biomass (t/y) from supply site (j) to industry (i).	
$Distance (D_{ij}) = \sum_{i=1}^3 \sum_{j=1}^{35} \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2}$ <p>(X<sub>i</sub>, Y<sub>i</sub>): Coordinates of industry (i).                      (X<sub>j</sub>, Y<sub>j</sub>): Coordinates of biomass site (j).</p>	Distance between biomass supplying site (j) to industry (i) based on UTM coordinates.
$Gross Biomass (B_{ij}) = \sum_{i=1}^3 \sum_{j=1}^{35} Q_j * A_{ij}$	Total biomass utilised (t/y)
$Gross Generation (G_{ij}) = \sum_{i=1}^3 \sum_{j=1}^{35} B_{ij} * R_{ij}$ <p>(R<sub>ij</sub>): Rate of production (t/t) at industry (j) using biomass (i).</p>	Total products generated (t/y)
$Gross Sales (S_{ij}) = \sum_{i=1}^3 \sum_{j=1}^{35} G_{ij} * P_{ij}$ <p>(P<sub>ij</sub>): Price of product at industry (j) using biomass (i).</p>	Total products sales (\$/y)
<b>Constraints:</b>	
$\sum_{j=1}^{35} A_j \leq 1 \quad \forall i$	Biomass consumed at site (j) should not be more than its capacity.
$A_{2,13}, A_{2,14}, \dots, A_{2,35} = 0$	Biomass from sites (13-35) is not suitable for industry (2).
$A_{i,1}, A_{i,2} \leq 0.578$	Only non-recycable biomass can be utilised from sites (1,2).
$A_{2,1}, A_{2,2} \leq 0.427$	Only food waste can be utilised from these sites for industry (2).
$G_{ij} \leq CP_{ij}$	Gross production at industry (i) should not exceed its capacity (CP <sub>ij</sub> )
<b>EFW Balancing Constraints:</b>	
$\sum_{j=1}^{35} G_{1,j} = \sum_{j=1}^{35} G_{2,j} = \sum_{j=1}^{35} G_{3,j}$	Scenario (a): Equal products generation.
$\sum_{j=1}^{35} B_{1,j} = \sum_{j=1}^{35} B_{2,j} = \sum_{j=1}^{35} B_{3,j}$	Scenario (b): Equal biomass utilisation.
$\sum_{j=1}^{35} S_{1,j} = \sum_{j=1}^{35} S_{2,j} = \sum_{j=1}^{35} S_{3,j}$	Scenario (c): Equal products sales.



### 3. Results and Discussion

#### 3.1 Data Mapping

The mapped sites of biomass resources and the three targeted EWF industries are presented in Figure 1. Most of the sites exist along the eastern coast towards the middle of the map, while the industries are distributed in the southern and western sides.



**Figure 1:** Quantities (t/y) and geospatial distribution of waste in Qatar.

#### 3.2 Biomass Allocation

The Excel solver yielded the optimal biomass allocation shown in Table 2 and Figure 2. In scenario (a), given equal production for the three industries (at 0.24 M tonne/y), 1.15 M tonne of biomass were utilised at an average transportation cost of 2.45 \$/t. While in scenario (b), equal biomass quantities were allocated for all industries (total: 0.48 M tonne/y), with the lowest biomass transportation cost attained at 1.98 \$/t and highest gross production and sales at 49.8 M t/y and 82.5 M \$/y, respectively. Whereas in scenario (c), equal products sales were achieved at 24 M \$/y, where 1.49 M tonne of biomass were utilised at an average transportation cost of 4.33 \$/t.

**Table 2:** Optimal allocation of biomass over EWF industries.

Biomass Category	(j)	Allocated biomass ( $A_{ij}$ )								
		Scenario (a)			Scenario (b)			Scenario (c)		
		E	W	F	E	W	F	E	W	F
		i=1	i=2	i=3	i=1	i=2	i=3	i=1	i=2	i=3
Solid waste	1	12%	0.00%	88%	12%	18%	6%	0%	0%	100%
	2	4%	0.00%	96%	10%	23%	67%	0%	0%	100%
Agro-waste	3	0%	0.00%	25%	0%	88%	0%	31%	30%	31%
	4	0%	0.00%	100%	50%	50%	0%	31%	27%	32%
	5	0%	0.00%	100%	50%	50%	0%	31%	25%	32%
	6	0%	0.00%	100%	50%	50%	0%	35%	18%	37%
	7	0%	0.00%	47%	0%	88%	11%	31%	29%	31%

Biomass	Category	(j)	Allocated biomass ( $A_{ij}$ )								
			Scenario (a)			Scenario (b)			Scenario (c)		
			E	W	F	E	W	F	E	W	F
			i=1	i=2	i=3	i=1	i=2	i=3	i=1	i=2	i=3
		8	0%	0.00%	42%	0%	88%	8%	36%	22%	37%
		9	0%	0.00%	100%	50%	50%	0%	31%	25%	32%
		10	0%	0.00%	100%	50%	50%	0%	36%	25%	37%
		11	0%	0.00%	100%	50%	50%	0%	35%	14%	37%
		12	0%	0.00%	100%	50%	50%	0%	35%	0%	38%
	Tyers	13	1%	0.00%	0%	1%	0%	0%	68%	0%	0%
	Dairy	14	5%	0.00%	0%	4%	0%	0%	90%	0%	0%
	manure	15	54%	0.12%	0%	0%	0%	0%	64%	9%	0%
		16	10%	0.01%	0%	6%	0%	0%	80%	0%	0%
		17	4%	0.00%	0%	0%	0%	0%	93%	0%	0%
	Sheep	18	9%	0.01%	0%	6%	0%	0%	82%	0%	0%
	manure	19	2%	0.00%	0%	0%	0%	0%	97%	0%	0%
		20	18%	0.01%	0%	0%	0%	0%	63%	0%	0%
	Camel	21	21%	0.01%	0%	7%	0%	0%	58%	1%	0%
	manure	22	21%	0.01%	0%	14%	0%	0%	58%	1%	0%
	Horse	23	2%	0.00%	0%	1%	0%	0%	87%	0%	0%
	manure	24	1%	0.00%	0%	1%	0%	0%	67%	0%	0%
	Poultry	25	19%	0.01%	0%	16%	0%	0%	61%	0%	0%
	litter	26	26%	0.02%	0%	0%	0%	0%	49%	1%	0%
		27	50%	0.03%	0%	0%	0%	0%	60%	1%	0%
	Sewage	28	12%	0.00%	0%	4%	0%	0%	76%	0%	0%
	sludge	29	23%	0.01%	0%	19%	0%	0%	53%	0%	0%
		30	24%	0.01%	0%	16%	0%	0%	51%	0%	0%
		31	4%	0.00%	0%	1%	0%	0%	92%	0%	0%
	Industrial	32	22%	0.01%	0%	0%	0%	0%	56%	0%	0%
	waste	33	31%	0.01%	0%	0%	0%	0%	48%	0%	0%
		34	26%	0.01%	0%	12%	0%	0%	48%	0%	0%
		35	0%	0.00%	0%	0%	100%	0%	36%	25%	0%

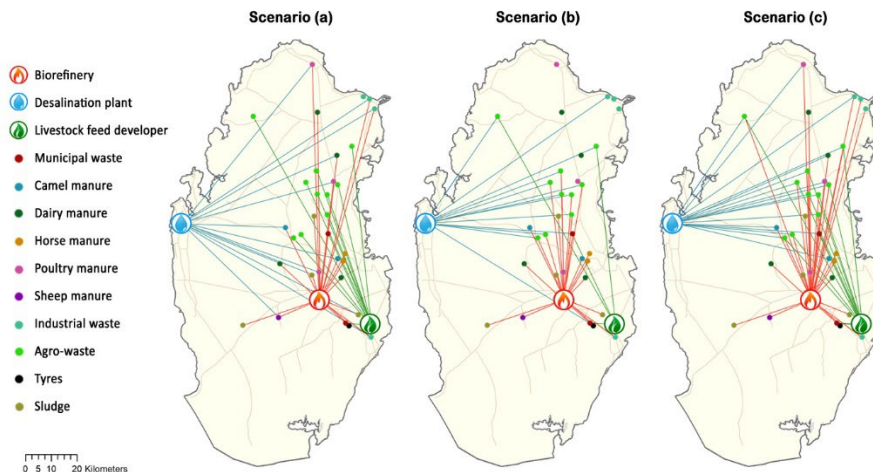


Figure 2: Representation of optimal biomass allocation over EWF industries.

#### 4. Conclusion

An EWF Nexus balancing framework was developed in this study to optimally allocate local waste biomass over EWF industries using ArcGIS and Excel tools. The framework has considered geospatial, quantitative, and qualitative aspects of biomass supplying sites and the receiving EWF industries. The model aimed at selecting optimal biomass resources and minimising their overall transportation costs. Moreover, three scenarios to achieve a Nexus balance were evaluated: a) equal biomass-based production, b) equal biomass distribution, and c) equal biomass-based sales. Scenario (b) achieved the lowest biomass transportation costs at 1.98 \$/t, however, lower biomass quantity has been utilised (0.48 M t/y). Whereas the maximum biomass utilisation was attained in scenario (c), with 1.49 M t/y of biomass transported at a minimised cost of 4.33 \$/t. This model provides insights on possible means to optimally utilise biomass resources at lower costs, while ensuring an EWF Nexus balance.

#### Acknowledgement

The authors would like to thank Qatar National Research Fund (QNRF) for their support of this research through NPRP-11S-0107-180216, Hamad Bin Khalifa University (HBKU) and Qatar Foundation (QF).

#### References

- C. Rayburn, 2012, *Water Reuse: Potential for Expanding the Nation's Water Supply through Reuse of Municipal Wastewater*.
- E. Moiola, F. Salvati, M. Chiesa, R.T. Siecha, F. Manenti, F. Laio, and M.C. Rulli, 2018, "Analysis of the Current World Biofuel Production under a Water–Food–Energy Nexus Perspective." *Advances in Water Resources* 121:22–31.
- IEA, 2007, *Biomass for Power Generation and CHP*.
- López-Díaz, D. Celeste, L.F. Lira-Barragán, E. Rubio-Castro, M. Serna-González, M.M. El-Halwagi, and J.M. Ponce-Ortega, 2018, "Optimization of Biofuels Production via a Water–Energy–Food Nexus Framework." *Clean Technologies and Environmental Policy* 20(7):1443–66.
- M. Alherbawi, A. AlNouss, G. McKay, and T. Al-Ansari, 2020, "Optimum Utilization of Jatropha Seedcake Considering the Energy, Water and Food Nexus." Pp. 229–34 in *30 European Symposium on Computer Aided Process Engineering*. Vol. 48, edited by S. Pierucci, F. Manenti, G. L. Bozzano, and D. B. T.-C. A. C. E. Manca. Elsevier.
- M. Alherbawi, A. AlNouss, G. McKay, and T. Al-Ansari, 2021, "Optimum Sustainable Utilisation of the Whole Fruit of Jatropha Curcas: An Energy, Water and Food Nexus Approach." *Renewable and Sustainable Energy Reviews* 137:110605.
- M. Alherbawi, G. McKay, H.R. Mackey, and T. Al-Ansari, 2022, "Multi-Biomass Refinery Siting: A Gis Geospatial Optimization Approach." *Chemical Engineering Transactions* 92:73–78.
- S. Namany, T. Al-Ansari, and R. Govindan, 2019, "Sustainable Energy, Water and Food Nexus Systems: A Focused Review of Decision-Making Tools for Efficient Resource Management and Governance." *Journal of Cleaner Production* 225:610–26.
- T. Al-Ansari, A. Korre, Z. Nie, and N. Shah, 2015, "Development of a Life Cycle Assessment Tool for the Assessment of Food Production Systems within the Energy, Water and Food Nexus." *Sustainable Production and Consumption* 2:52–66.

# Levelised Energy-Water-Food (LEWF) Indicator for Biomass Assessment and Optimisation

Mohammad Alherbawi, Ahmed AlNouss, Gordon McKay, Tareq Al-Ansari \*

*College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar*

\*[talansari@hbku.edu.qa](mailto:talansari@hbku.edu.qa)

## Abstract

The Energy-Water-Food (EWF) Nexus approach has recently become an important tool to ensure efficient and optimum management of the three vital resources by considering the inter-dependencies amongst them. However, EWF optimisation complexity increases at times when resources are prioritised during trade off-driven decision-making, coupled with differences between the three resources in terms of measurement units. Therefore, this study presents a novel ‘Levelised-EWF indicator’ (LEWF) which merges the net value of energy, water and food requirement for a system or a product into a single value in (Megajoule equivalent “MJ-e”) unit. This tool enables a convenient comparison between biomass sources and processing technologies in the light of the EWF Nexus by converting multi-objective problems into single objective. Moreover, the study proposed a 10-level EWF resilience chart, to evaluate the system resilience to the changes in EWF resources supplies based on the obtained LEWF value. As a case study, the proposed LEWF tool is used to evaluate the various processing routes for *Jatropha curcas*’ seedcake to produce various energy products. Required data are obtained via earlier Aspen Plus<sup>®</sup> simulations of potential processes including gasification, pyrolysis, hydrothermal liquefaction, incineration, fermentation, and anaerobic digestion. The results demonstrated that pyrolysis is the optimal processing route for *Jatropha* seedcake with the lowest LEWF value of 0.23 MJ-e/MJ (product), with an EWF Resilience Level of 9 out of 10. The importance of this tool lies in its ability to ensure an adequate management of EWF resources within biomass systems and ensure system resilience against possible changes in supply chains.

**Keywords:** LEWF; Resilience; Biomass; Optimisation; Simulation; *Jatropha*.

## 1. Introduction

Energy, water, and food are the three essential resources for human existence and nations development. The increasing demand on these resources has raised the need to efficiently manage their consumption in a balanced manner. However, efficient management becomes more pushing at times when complicated interlinkages between these resources exist. Such a relationship was officially introduced in the Rio+20 Summit as “EWF Nexus” (Galmés et al. 2014).

The EWF Nexus concept has emerged as a key tool to evaluate the EWF resources’ synergies and tradeoff to ensure an adequate consumption balance. Several EWF Nexus frameworks have then been developed based on existing quantitative and qualitative approaches.

Al-Ansari et al. (2015) presented a Nexus framework based on lifecycle assessment (LCA) concept to study the possible tradeoffs amongst EWF systems and sub-systems. Meanwhile, Daher and Mohtar (2015) developed a scenario-based framework (WEF Tool 2.0) to trace the flow of EWF resources for different case studies towards reaching a self-sufficiency scenario. In addition, Alherbawi et al. (2021) developed a technology-based optimisation framework to evaluate the key biomass processing technologies and their interventions with EWF resources. Whereas Flammini et al. (2014) introduced a “Nexus Rapid Appraisal” method as a decision-making tool, through which the systems’ associated EWF resources are plotted in spider charts against standard values. The chart’s polygon area is considered a rapid visual representation of the possible intervention between the three resources. Moreover, Moioli et al. (Moioli et al. 2018) developed a “Nexus Index” as reference to evaluate the system’s EWF requirement to produce one energy unit.

In this study, a new concept is introduced, named “The Levelised Energy, Water, Food Nexus (LEWF Indicator)” for biomass-biofuel systems. The indicator merges the system’s associated EWF resources input into a single value, with energy equivalent unit (E-eq) based on the concept of resource payback that was introduced earlier by the authors (Alherbawi et al. 2020, Alherbawi et al. 2021). Merging the EWF into a single value facilitates a convenient comparison between the various biomass resources and the different processing technologies.

## 2. Methodology

### 2.1. Data Collection

Required data for the establishment of the LEWF concept and the development of its formulation have been collected from literature as displayed in Table 1. The table includes information on the interlinked requirements amongst the EWF resources and sectors.

**Table 1:** Data collected to assist the formulation of the LEWF Indicator.

Scope	Detailed item	Value*	Unit
Energy to water	Water requirement for coal-power generation	0.0790	(m <sup>3</sup> /GJ)
	Water requirement for NG-power generation	0.0760	(m <sup>3</sup> /GJ)
	Water requirement for nuclear-power generation	0.0180	(m <sup>3</sup> /GJ)
	Water requirement for solar-power generation	0.0064	(m <sup>3</sup> /GJ)
	Water requirement for wind-power generation	0.0002	(m <sup>3</sup> /GJ)
	Water requirement for hydro-power generation	0.3000	(m <sup>3</sup> /GJ)
	Water requirement for diesel fuel production	5.3000	(m <sup>3</sup> /GJ)
Water to Energy	Energy requirement for rainfed irrigation	0.0000	(GJ/m <sup>3</sup> )
	Energy requirement for groundwater provision	0.0013	(GJ/m <sup>3</sup> )
	Energy requirement for FO desalination	0.0029	(GJ/m <sup>3</sup> )
	Energy requirement for RO desalination	0.0108	(GJ/m <sup>3</sup> )
	Energy requirement for MED desalination	0.0130	(GJ/m <sup>3</sup> )
	Energy requirement for MSF desalination	0.0114	(GJ/m <sup>3</sup> )
	Energy requirement for wastewater treatment	0.0022	(GJ/m <sup>3</sup> )
Food to Energy	Energy requirement for cereal production	3.5600	(GJ/t)
	Energy requirement for beef production	44.7500	(GJ/t)

Land to Energy	Land requirement for coal-power generation	37.6040	(m <sup>2</sup> /GJ)
	Land requirement for NG-power generation	38.2200	(m <sup>2</sup> /GJ)
	Land requirement for nuclear-power generation	39.1440	(m <sup>2</sup> /GJ)
	Land requirement for solar-power generation	133.971	(m <sup>2</sup> /GJ)
	Land requirement for wind-power generation	217.557	(m <sup>2</sup> /GJ)
	Land requirement for hydro-power generation	970.815	(m <sup>2</sup> /GJ)
	Land requirement for diesel fuel production	0.2400	(m <sup>2</sup> /GJ)
Land to Food	Land requirement for cereal production	0.00023	(m <sup>2</sup> /t)
Fertilisers to Food	N-fertiliser requirement for cereal production	0.014	(t/t)
LS Feed to Food	Protein-feed requirement for beef production	0.13	(t/t)

\* Resources:(Alherbawi et al. 2021; Dickinson 2018; FAO 2000; Jungbluth, Meili, and Wenzel 2018; Ladha et al. 2016; Stevens et al. 2017)

## 2.2. LEWF Indicator Development

The LEWF indicator is formulated as given in Equation 1. The mathematical formula translates the water consumed or produced in a system into an equivalent energy amount, that is the amount of energy required to generate an equivalent quantity of water. Similarly, the formula converts the amount of food consumed into energy equivalent value based on the amount of energy required to produce similar amount of food. Nevertheless, the formula considers the embodied water and land of 7 energy systems to comprehensively account for all EWF resources' consumption throughout the lifecycle of a biomass-based process. Meanwhile, the different conversion factors (CF) required for the conversion are calculated and reported in the results section based on the data collected in Table 1.

$$LEWF = 1 - \left[ \frac{\left( \sum_{e=1}^{e'} E_e \cdot CF_e + \sum_{w=1}^{w'} W_w \cdot CF_w + \sum_{f=1}^{f'} F_f \cdot CF_f \right)}{\left( \sum_{p=1}^{p'} Q_p \cdot HV_p \right)} \right] \quad (1)$$

Where:

**LEWF**: Levelised Energy-Water-Food indicator (GJ-e<sub>in</sub>/GJ<sub>out</sub>).

**E**: Total energy consumed throughout the process' lifecycle via different energy supplies (e).

**W**: Net water consumed throughout the process' lifecycle via different water supplies (w).

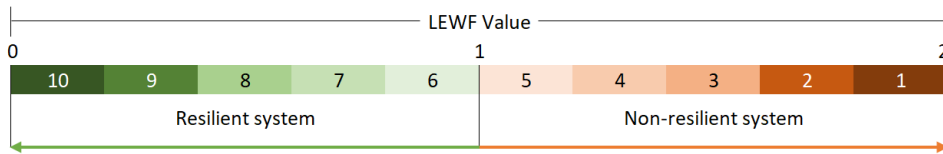
**F**: Net food consumed throughout the process' lifecycle via different food supplies (f).

**CF**: Conversion factor into energy equivalent value.

**Q**: Gross mass quantity of different energy products generated (p).

**HV**: Heating value of product (p).

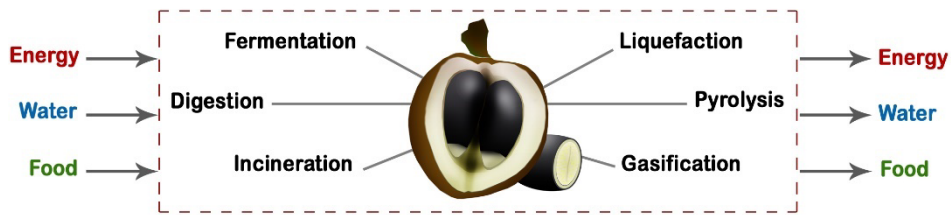
The LEWF obtained value may offer an indication of how resilient the system is. Where LEWF values of less than 1 (MJ-e/MJ product) indicate that the biomass-based system can generate sufficient energy to payback for the water and food consumed, while the resource payback is performed using an energy currency rather than money. Thus, the system hedges towards the worst-case scenarios, where in case of EWF supply chain disturbance, the system must be able to generate the required resources based on the extent of energy generated. Therefore, a 10-step EWF Resilience Chart is formulated based on the LEWF values, with a step size of 0.2, as illustrated in Figure 1.



**Figure 1:** Proposed EWF Resilience Chart.

### 2.3. Case Study

To evaluate the proposed LEWF tool and Resilience chart, a case study from author's previous work has been considered (Alherbawi et al. 2021), in which *Jatropha curcas* seedcake biomass was processed into multiple energy-producing technologies as illustrated in Figure 2. Energy, water and food-related data (Table 2) were adapted and plugged into the LEWF formula (Equation 1).



**Figure 2:** *Jatropha* seedcake's processing technologies within EWF boundaries.

**Table 2:** Case study's data on EWF systems consumption and generation.

Technology	Energy Generated (GJ/y)	Energy consumed (GJ/y)	Net water consumed ( $10^3 \text{ m}^3/\text{y}$ )	land used ( $10^3 \text{ m}^2/\text{y}$ )	By-products credit (Fertilisers) (tonne/y)
Gasification	1,039,967	291,924	38,557	25,000	0
Pyrolysis	1,645,898	291,735	38,431	25,000	0
Liquefaction	1,072,398	292,112	38,552	25,000	0
Fermentation	180,661	292,559	38,533	25,000	21,712 (N: 5%)
Digestion	664,308	291,782	38,398	25,000	2,778 (N: 5%)
Incineration	475,820	293,154	38,375	25,000	0

## 3. Results and Discussion

### 3.1. LEWF Indicator's Database

Utilising the data collected in Section 2, a set of conversion factors from different EWF resources into an energy equivalent value is provided in Table 3. The database comprises 7 energy supply resources and 7 water supply resources, in addition to food-related resources as energy crops or lands that are diverted from food cultivation towards energy production.

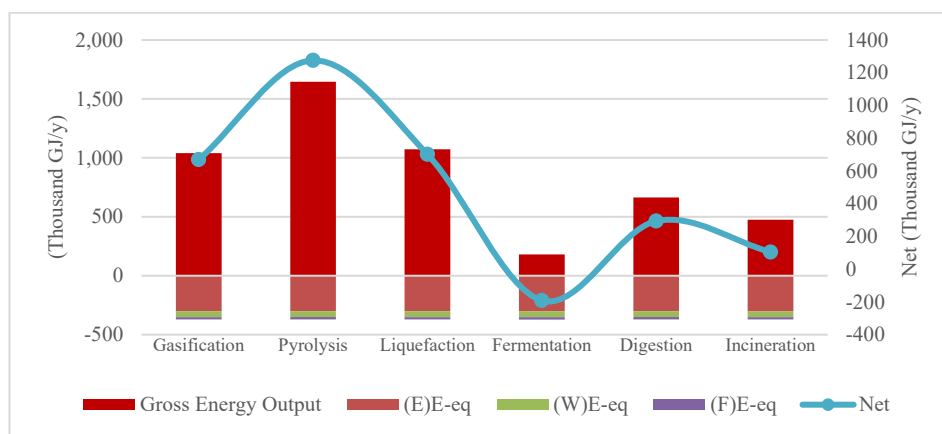
Nevertheless, food-related by-products (i.e., fertilisers, livestock feed) were taken into consideration, and were associated with a negative conversion factor as they represent a credit to the system, which cuts from its net resources consumption.

**Table 3:** LEWF Indicator’s conversion factor.

Source	Description	Energy equivalent conversion factor (GJ-e)
<b>Energy (GJ)</b>	Coal-Power	1.0313
	Natural Gas-Power	1.0317
	Nuclear- Power	1.0322
	Solar-Power	1.1097
	Wind- Power	1.1781
	Hydro- Power	1.7967
	Diesel fuel	1.0317
<b>Water (m<sup>3</sup>)</b>	Average	1.1731
	Rainfed	0.0000
	Groundwater	1.3E-03
	Forward osmosis	2.9E-03
	Reverse osmosis	1.1E-02
	Multi-effect distillation	1.3E-02
	Multi-stage flash	1.1E-02
	Wastewater treatment	2.2E-03
	Average	5.9E-03
<b>Food-related (tonne)</b>	Energy Crops	3.5600
	Land (m <sup>2</sup> )	8.2E-04
	By- products credit	-0.0498
	Fertilisers (N content) Animal feed (Protein content)	-5.8175

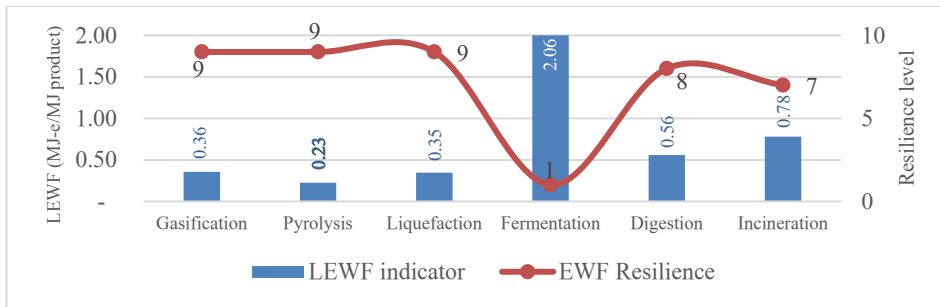
### 3.2. Model Application

The proposed LEWF Indicator has been applied to a *Jatropha* seedcake case study. The breakdown of total equivalent energy requirement for the different processing pathways is plotted in Figure 3. Whereas the final LEWF values and the corresponding EWF Resilience levels are illustrated in Figure 4.



**Figure 3:** Energy equivalent analysis of *Jatropha* seedcake processing technologies.





**Figure 4:** LEWFs and Resilience levels of *Jatropha* seedcake processing systems.

Pyrolysis achieved the lowest LEWF value (0.23 MJ-e/MJ) and highest EWF Resilience level (9/10), therefore, it is selected as an optimal processing technology. Similar conclusion has been reached through an EWF multi-objective optimisation conducted in MATLAB earlier (Alherbawi et al. 2021), which further validate the proposed LEWF Indicator.

#### 4. Conclusion

A novel LEWF Indicator is introduced for the evaluation of EWF resources requirement for energy systems. The indicator merges the three EWF input values into a single figure in energy equivalent unit (E-eq), which enables a convenient evaluation of EWF interlinkages in a system and facilitates the comparison with other systems. The indicator was implemented in a case study to select optimal processing pathway of *Jatropha* seedcake, where pyrolysis achieved the lowest LEWF at 23 (MJ-e/MJ product).

#### References

- A. Flammini, M. Puri, L. Pluschke, and O. Dubois, 2014, *Walking the Nexus Talk: Assessing the Water-Energy-Food Nexus in the Context of the Sustainable Energy for All Initiative*. Rome: Food and Agriculture Organization of the United Nations (FAO).
- B.T. Daher, and R.H. Mohtar, 2015, "Water-Energy-Food (WEF) Nexus Tool 2.0: Guiding Integrative Resource Planning and Decision-Making." *Water International* 40(5-6):748-71.
- E. Moiola, F. Salvati, M. Chiesa, R.T. Siecha, F. Manenti, F. Laio, and M.C. Rulli, 2018, "Analysis of the Current World Biofuel Production under a Water-Food-Energy Nexus Perspective." *Advances in Water Resources* 121:22-31.
- E.W. Dickinson, 2018, *Solar Energy Technology Handbook*. CRC Press.
- FAO, 2000, *The Energy and Agriculture Nexus*. Rome.
- J.K. Ladha, A. Tirol-Padre, C. K. Reddy, K. G. Cassman, Sudhir Verma, D. S. Powlson, C. van Kessel, Daniel de B Richter, Debashis Chakraborty, and Himanshu Pathak, 2016, "Global Nitrogen Budgets in Cereals: A 50-Year Assessment for Maize, Rice, and Wheat Production Systems." *Scientific Reports* 6:19355.
- L. Stevens, B. Anderson, C. Cowan, K. Colton, and D. Johnson, 2017, "The Footprint of Energy: Land Use of US Electricity Production." *STRATA: Logan, UT, USA*.
- M. Alherbawi, A. AlNouss, G. McKay, and T. Al-Ansari, 2021, "Optimum Sustainable Utilisation of the Whole Fruit of *Jatropha Curcas*: An Energy, Water and Food Nexus Approach." *Renewable and Sustainable Energy Reviews* 137:110605.
- M. Alherbawi, A. Ahmed AlNouss, G. McKay, and T. Al-Ansari, 2020, "Optimum Utilization of *Jatropha* Seedcake Considering the Energy, Water and Food Nexus." Pp. 229-34 in 30 European Symposium on Computer Aided Process Engineering. Vol. 48, Manca. Elsevier.
- N. Jungbluth, C. Meili, and P. Wenzel, 2018, "Life Cycle Inventories of Oil Refinery Processing and Products." *ESU-Services Ltd. Commissioned by BFE, BAFU, Erdöl-Vereinigung, Schaffhausen, Switzerland, Retrieved from: Www. Lc-Inventories. Ch*.
- T. Al-Ansari, A. Korre, Z. Nie, and N. Shah, 2015, "Development of a Life Cycle Assessment Tool for the Assessment of Food Production Systems within the Energy, Water and Food Nexus." *Sustainable Production and Consumption* 2:52-66.
- U.V. Galmés, R. Smith, I. Bokova, G.M. Candido de Castro, R. Toneto Jr., P. Brabeck-Letmathe, S. Cleasby, J. Korff, K. Beckett, S. Haye, and V. Somavilla, 2014, *The Road from RIO+20*.

# The controlled environment agriculture: a sustainable agrifood production paradigm empowered by systems engineering

Liang Wang<sup>a,d</sup>, Guoping Lian<sup>a,b</sup>, Zoe Harris<sup>a</sup>, Mark Horler<sup>c</sup>, Yang Wang<sup>d</sup>, Tao Chen<sup>a</sup>

<sup>a</sup>*Faculty of Engineering & Physical Sciences, University of Surrey, Guildford, GU2 7XH, UK*

<sup>b</sup>*Unilever R&D Colworth, Sharnbrook, Bedfordshire, MK40 1LQ, UK*

<sup>c</sup>*UK Urban AgriTech, 46 Jamaica Street, Liverpool, L1 0AF, UK*

<sup>d</sup>*National Innovation Center for Digital Fishery, China Agricultural University, Beijing, 100083, China*

## Abstract

Controlled environment agriculture (CEA) has some clear advantages over traditional farming, such as: reliable and consistent production capability; efficiency in water and space use; reducing the use and runoff of fertiliser and pesticides; etc. As such CEA can greatly benefit from the CAPE (computer-aided process engineering) approach – cross-fertilization of these two apparently distinct areas may result in new methods and applications to improve CEA and process engineering, with potentially significant contribution to circular economy. In this paper, we discuss several important aspects of CEA drawing from our own experiences in aquaculture and aeroponics, including product development, process design and process operation, and the potential contribution of CAPE. Finally, we postulate a systems platform for CEA, aiming to foster a long-lasting partnership between the two scientific communities.

**Keywords:** Controlled environment agriculture; Computer-aided process engineering; Digital twin; Optimisation; Process control; Process modelling.

## 1. Introduction

A recent breaking news is that the global population has exceeded 8 billion people, (United Nations, 2022). This means that more and more agricultural products and raw materials are required in future. The traditional production methods face a tremendous challenge to meet the ever-increasing demand whilst aiming to achieve net zero and minimal environmental impact (Hemathilake et al., 2022). Thus, advanced agricultural systems with high efficiency, high production capacity and low emission need to be developed.

Controlled environment agriculture (CEA) refers to the production of food (e.g., crops, aquatic and livestock) and non-food (e.g., biomass, fur) agricultural products under conditions that are more controllable than traditional farming in the open field/ponds. It uses adjustable production environment and facility conditions to optimal and correct the growth system over real-time (Benke et al., 2017), provides a high-efficiency, high-yield, potentially low carbon and sustainable production. At scale, CEA can spare the use of land and water bodies with significant ecological benefits. However, CEA faces the challenges of high capital investment and operating costs, the need for crop/animal to

adapt to the indoor environment, and complex cultural and political resistance. Some of these issues can be addressed by technological advancement including the use of systems engineering thinking and methods.

CEA encompasses a wide range of primary production methods, ranging from loosely controllable environment in polytunnel greenhouses, aquaculture pens, to totally controlled production systems in plant factories and recirculating aquaculture systems. Like any emerging and rapidly developing subject, this field has been given many names which often overlap in their connotation but have different focuses, such as plant factory, vertical farming, urban farming, facility agriculture, to name a few. We choose CEA to emphasise its key difference from traditional, uncontrolled farming.

In this paper, we discuss from the perspective of systems engineering the three tasks in CEA: product development, process design and process operation, with a focus on the latter. The narrative will mainly draw from our own experiences in aquaculture, aquaponics and aeroponics (Li et al., 2021; 2022; Guo et al., 2022; Zhou et al., 2021), supplemented by work reported in the open literature. Figure 1 illustrates two examples of CEA production systems. Other crop and livestock production systems, such as indoor poultry and insect farming, have their own special requirements in terms of product development, process design and operation. However, they share the same conceptual framework of using systems engineering approach to enable more sustainable production.

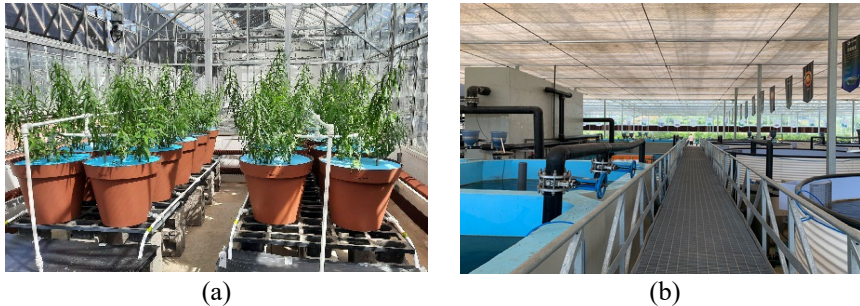


Figure 1. CEA examples: (a) Aeroponic cultivation of bioenergy willows in a greenhouse; (b) Industrialised aquaponics to produce fish and vegetables.

## 2. Systems engineering techniques for product and process design

### 2.1. Computer aided product design

Products play an important role on the production system. In CEA, product design / redesign is often needed to adapt to the spatiotemporal characteristics and the constraints of resources. For example, the willows growing in an indoor, aeroponic environment (Figure 1(a)) should ideally be dwarf, fast-growing, comfortable with aeroponic irrigation and nutrition, and resistant to disease. The same idea is also used in aquaponics factories, where the high-density production environment requires fish species with fast growth and disease resistance. Such controlled environment and the adapted crops or animals may also require new formulations of the feed.

Computer aided breeding is a well-established concept. Statistical design of experiments has been widely used in decades. Advanced machine learning / artificial intelligence methods, incorporating multi-faceted data such as imaging, genomics and phenomics, have shown promises (Harfouche et al., 2019; Marsh et al., 2021). More fundamental,

systems biology models, such as gene regulatory networks, have been applied to obtain the favourable traits with mechanistic understanding of why and how (Lavarenne et al., 2018). Notably, these computational methods have the maximum potential when they are integrated with high-throughput experiments – CEA can provide such high-throughput facilities with close control of the environmental parameters.

## *2.2. Computer aided process design*

CEA at industrial scale is analogous to biomanufacturing processes. The centrepiece is the “bioreactor” to produce the biomass (animal or plant). A wide range of facilities, equipment and devices are needed to support the microclimate environment, water quality, nutrient supply, disease control, downstream processing of products and/or wastes, etc. These decisions mirror process design in CAPE. For most facilities, the key design considerations are to do with HVAC (heating, ventilation and air conditioning), water supply, and power supply, to provide required temperature, humidity, ventilation, supplementary light, irrigation, etc.

Specific requirements are needed beyond the above general principle. For example, recirculating aquaculture systems (RAS) require design of not just the fish tanks, but also aeration, sterilisation, water quality adjustment, multi-stage filtration (physical and biological), etc. For soilless plant production, the design of nutrient supply and recycle units (Bittsánszky et al., 2016), supplementary lighting / shading facilities, etc., will be required in addition to HVAC.

It is necessary to fully consider the integration between product design (breeding) and process design, as opposed to treating them as two separate tasks. Computer aided process design in CEA is emerging (Guo et al., 2022; Holzworth et al., 2015), though its full potential is yet to materialise, especially in whole-process simulation and design.

Energy consumption has always been a major factor in agricultural production. The requirement for heating during winter has restricted the use of many agricultural technologies in high latitude regions. This is particularly an issue for CEA. A potential solution is the use of a multitude of energy sources, including solar, wind energy, biomass, geothermal and tidal energy. In high latitude regions, distributed energy systems for combined heat and power hold promise to reduce reliance on the grid, reduce cost and reduce emission. We conducted large-scale aquaponics production experiments by integrating solar, geothermal and biomass energy in Northern China (Le et al., 2021), supported by computational fluid dynamic simulations (Le et al., 2020).

## **3. Systems engineering methods for process operation**

Similar to chemical manufacturing, the operation of CEA processes can be organized in a hierarchy from regulatory control, optimisation to scheduling, planning and supply chain management. Agricultural production is inherently slow and in batch mode. Slow dynamics in biomass production suggests some level of tolerance to operation disturbances. For example, although the desired temperature for Salmon growth is around 12 °C, a short exposure to temperature moderately outside this range is usually tolerable, creating degrees of freedom for optimisation. On the other hand, the nutrient dynamics can be fast, such as nitrification in a fish tank, so is the dynamics of microclimate parameters. This disparity in dynamics may pose challenges to process monitoring, control, optimisation and scheduling.

### *3.1. Soft sensor for process monitoring*

Effective process operation requires reliable measurement. In traditional agriculture, a key consideration is the cost and maintainability because of the often harsh environment in the open field. These challenges are alleviated to some extent in CEA. Nevertheless, since agricultural products are relatively low value, cost and maintainability are still big drivers. For example, many commercial ammonia sensors need to be calibrated once a month – this might be an affordable, standard practice in a wastewater processing plant in the UK (private communication) but impractical for agriculture.

Therefore, software based sensor self-calibration has been proposed as a solution (Yuri et al., 2015), for temperature, dissolved oxygen (DO) and ammonia (Yu et al., 2020). Liu et al. (2013) developed a multi-source heterogeneous data preprocessing technologies to improve the measurement accuracy. Li et al. (2021) applied a moving horizon estimation approach, widely used in CAPE, with a mechanistic nitrogen cycle model as a state estimator to improve the resolution and reliability of total nitrogen monitoring from infrequent sensor readings.

### *3.2. Process control & optimisation*

Process control in CEA normally refers to the control of the microclimate parameters, i.e., temperature, humidity, water quality, nutrient supply, light intensity, CO<sub>2</sub> concentration (if CO<sub>2</sub> enrichment is possible in a totally controlled environment), etc. In the context of aquaculture DO control has been the primary focus, because fish health and viability is very sensitive to DO. Aeration is an effective manipulated variable to control DO, but it associates with significant energy cost. DO's importance has warranted a large number of reports of controller design, such as adaptive PID control enhanced by evolutionary algorithms (Zhou et al., 2021). In principle, model predictive control (MPC) should give improved results as shown in controlling DO in wastewater processing (O'Brien et al., 2011), though the application of MPC in agriculture is still under-explored. In addition, due to the difficulties in real-time measurement of other water quality indices, e.g. ammonia nitrogen, their feedback control is less common in practice.

Direct control and/or optimisation of the biomass growth has been less reported in the literature. In principle these are similar to batch process control/optimisation problems. The challenges are two-fold: the limitation on measurement and the incredibly long batch duration. Direct measurement of biomass growth could be achieved by manual sampling or using imaging based techniques, but the latter itself is an open research area. Long batch duration, from weeks to months, renders run-to-run control impractical. There have been reports of reference tracking using MPC (Chahid et al., 2021) and direct use of a mechanistic model in growth optimisation (Li et al., 2022). If computer models are developed to incorporate the prediction of both production and environmental parameters, they could be used to formulate economic indices and thus cast into an economic MPC framework for dynamic, direct optimisation of the economic measures. Currently such methods are limited to simulations and have a long way to be taken up in practice.

### *3.3. Scheduling, planning & supply chain*

Due to the long lead time of agriculture produces, scheduling, planning and supply chain integration is key to commercial viability and success. At the process level, scheduling models could be developed at steady state to mixed integer linear programs (MILP) or even nonlinear ones, similar to those in the vast CAPE literature, and solved accordingly. With distributed energy systems, the objective is typically to maximise capacity

utilisation with minimal energy costs, considering variable energy inputs/costs due to the use of renewables. Alternatively economic MPC, as mentioned above, might be useful to make dynamic decisions enclosing both the scheduling and optimisation levels. The key enabler is digitalisation throughout the production, and integrated with pre-production to post-production for optimising the fresh/cold supply chain. An especially interesting case is aquaponics, where the production of fish and crops should be scheduled in a staggered manner to meet variable market demand.

### *3.4. Big data, IoT and AI*

Agriculture is traditionally data poor but has seen rapid transition to big data, enabled by massive connectedness, industrial internet-of-things (IoT), and data processing platforms such as the Apache Spark and Storm. Big data has allowed the application of advanced AI for the intelligent perception, prediction, and decision making in agrifood production (Wang et al., 2021). A particularly important application of big data AI in CEA is disease control. AI is well suited to integrating data from low-cost camera with environment data, and by using machine learning algorithms to providing early warning of pathogens (Wen et al., 2012; Zhang et al., 2020). We recently demonstrated a proof-of-concept of using acoustic signal to determine fish feeding intensity (Cui et al., 2022).

## **4. Concluding remarks**

In this paper we provided an overview of CEA as an emerging method for efficient and sustainable agrifood production. Whether CEA can fulfill its perceived benefits is likely to depend on how the products are processes are designed, and how the processes are operated, all in a holistic approach encompassing commercial viability, energy consumption, environmental impact, and circularity. There is a strong case for adapting CAPE methods to the application in CEA, as shown in increasing research reported in the literature. However, challenges abound. Agriculture produces are of relatively low value per kg produced compared with many chemical products familiar to the CAPE community. Most processes are slow with a long time delay / time constant between input and output. Measurement of key process parameters, especially those to do with the biomass properties (e.g., growth and health of the animal/crop), is difficult. On the flip side, the live biomass being produced are often tolerant to minor environment changes, providing more space for optimisation. We foresee a fruitful and long-lasting partnership between CAPE and CEA to address these and emerging challenges in the near future.

## **References**

- Kurt Benke , Bruce Tomkins , 2017, Future food-production systems: vertical farming and controlled-environment agriculture, *Sustainability: Science, Practice and Policy*, 13:1, 13-26.
- A. Chahid , I. N'Doye, J.E. Majoris, M.L. Berumen, T.M. Laleg-Kirati, 2021, Model predictive control paradigms for fish growth reference tracking in precision aquaculture, *Journal of Process Control*, 105, 160-168.
- M. Cui, X Liu, J. Zhao, J. Sun, G. Lian, T. Chen, M.D. Plumbley, D. Li, W. Wang, 2022, Fish Feeding Intensity Assessment in Aquaculture: A New Audio Dataset AFFIA3K and a Deep Learning Algorithm," 2022 IEEE 32nd International Workshop on Machine Learning for Signal Processing (MLSP), 2022, pp. 1-6.
- Y. Guo, S. Zhang, S. Li, M. Zhang, H. Benli, Y. Wang, 2022, Numerical investigation for effects of natural light and ventilation on 3D tomato body heat distribution in a Venlo greenhouse, *Information Processing in Agriculture*, in press.

- A.L. Harfouche, D.A. Jacobson, D. Kainer, J.C. Romero, A.H. Harfouche, G.S. Mugnozza, M. Moshelion, G.A. Tuskan, J.J.B. Keurentjes, A. Altman, 2019, Accelerating Climate Resilient Plant Breeding by Applying Next-Generation Artificial Intelligence, *Trends in Biotechnology*, 37: 1217-1235.
- D.M.K.S. Hemathilake, D.M.C.C. Gunathilake, 2022, Chapter 32 - High-productive agricultural technologies to fulfill future food demands: Hydroponics, aquaponics, and precision/smart agriculture, *Future Foods*, Academic Press, Pages 555-567.
- D.P. Holzworth, V. Snow, S. Janssen, I.N. Athanasiadis, M. Donatelli, G. Hoogenboom, J.W. White, P. Thorburn, 2015, Agricultural production systems modelling and software: Current status and future prospects, *Environmental Modelling & Software*, 72: 276-286.
- J. Lavarenne, S. Guyomarc'h, C. Sallaud, P. Gantet, M. Lucas, 2018, The Spring of Systems Biology-Driven Breeding, *Trends in Plant Science*, 23, 706-720.
- A.T. Le, Wang, L.; Wang, Y.; Vu, N.T.; Li, D, 2020, Experimental Validation of a Low-Energy-Consumption Heating Model for Recirculating Aquaponic Systems, *Energies*, 13, 1958.
- A.T. Le, L. Wang, Y. Wang, D. Li, 2021, Measurement investigation on the feasibility of shallow geothermal energy for heating and cooling applied in agricultural greenhouses of Shouguang City: Ground temperature profiles and geothermal potential, *Information Processing in Agriculture*, 8: 251-269.
- H. Li, S. Chatzifotis, G. Lian, Y. Duan, D. Li, T. Chen, 2022, Mechanistic model based optimization of feeding practices in aquaculture, *Aquacultural Engineering*, 97: 102245.
- H. Li, W. Li, M. McEwan, D. Li, G. Lian, T. Chen, 2021, Adaptive filtering-based soft sensor method for estimating total nitrogen in aquaponic systems, *Computers and Electronics in Agriculture*, 186: 106175.
- S. Liu, L. Xu, D. Li, Q. Li, Y. Jiang, H. Tai, L. Zeng, 2013, Prediction of dissolved oxygen content in river crab culture based on least squares support vector regression optimized by improved particle swarm optimization, *Computers and Electronics in Agriculture*, 95: 82-91.
- J.I. Marsh, Hu, H., Gill, M. et al., 2021, Crop breeding for a changing climate: integrating phenomics and genomics with bioinformatics. *Theor Appl Genet* 134, 1677–1690.
- M. O'Brien, J. Mack, B. Lennox, D. Lovett, A. Wall, 2011, Model predictive control of an activated sludge process: A case study, *Control Engineering Practice*, 19, 54-61.
- United Nations, 2022, Data Portal Population Division, <https://population.un.org/dataportal/home>.
- Wang T, Xu X., Wang C, Li Z., Li D, 2021, From Smart Farming towards Unmanned Farms: A New Mode of Agricultural Production, *Agriculture*, 11, 145.
- H. Wen, L. Zhang, Z. Fu, X. Li, Y. Su, 2012, Agricultural disease-control scene determination based on audio-visual fusion, *Journal of Food, Agriculture & Environment*, 10: 867-870.
- H. Yu, Y. Chen, S. Gul Hassan, D. Li, 2016, Prediction of the temperature in a Chinese solar greenhouse based on LSSVM optimized by improved PSO, *Computers and Electronics in Agriculture*, 122: 94-102.
- H.H. Yu, L. Yang, D. Li, Y. Chen, 2020, A hybrid intelligent soft computing method for ammonia nitrogen prediction in aquaculture, *Information Processing in Agriculture*, (20)30034-2.
- P. Zhang, L. Yang, D. Li, 2020, EfficientNet-B4-Ranger: A novel method for greenhouse cucumber disease recognition under natural complex environment, *Computers and Electronics in Agriculture*, 176: 105652.
- X. Zhou, D. Li, L. Zhang, Q. Duan, 2021, Application of an adaptive PID controller enhanced by a differential evolution algorithm for precise control of dissolved oxygen in recirculating aquaculture systems, *Biosystems Engineering*, 208, 186-198.

## Acknowledgement

We thank funding support from the UK BBSRC (BB/S020896/1), the UK Biomass Feedstocks Innovation Programme, and the National Science Foundation of China (62339001).

# CO<sub>2</sub> Capture by Amine-Functionalized Magnesium Oxide: Experimental and Modelling Studies

Sagheer A. Onaizi

*Department of Chemical Engineering and Center of Excellence in Nanotechnology,  
King Fahd University of Petroleum and Minerals, Dhahran 31216, Saudi Arabia*

## Abstract

The concentration of carbon dioxide in the atmosphere has reached an alarming level. CO<sub>2</sub> can be captured using the traditional amine absorption process. However, this process is an energy-intensive and environmentally unfriendly. Thus, we have developed in this work a novel adsorbent using magnesium oxide (MgO) that has been functionalized with 3-aminopropyl-triethoxysilane (APTES). This CO<sub>2</sub> adsorbent (APTES-MgO) is reported for the first time in this work to the best of our knowledge. APTES-MgO can significantly adsorb CO<sub>2</sub> at ambient conditions. It is fully regenerable and the regeneration process can be carried out at mild conditions (120°C and atmospheric pressure) under a nitrogen environment. This low regeneration temperature indicates that the energy consumption of CO<sub>2</sub> capture process using APTES-MgO could be low. The adsorption capacity of APTES-MgO is more than 1.5 mmol/g at 30 °C and 1 atm. The regeneration process of APTES-MgO can fully restore its CO<sub>2</sub> adsorption capacity. The adsorption process has been modelled using a first order adsorption model. The mathematical model can track the experimental data quite well, allowing the estimation of the key adsorption parameters. The effectiveness of APTES-MgO in adsorbing CO<sub>2</sub> at ambient conditions and its regenerability and stability suggest its potential commercial application in CO<sub>2</sub> capture industry.

**Keywords:** Amine-functionalized magnesium oxide, Adsorption, Regeneration, Ambient conditions, Carbon dioxide (CO<sub>2</sub>) capture, Mathematical modelling

## 1. Introduction

Next paragraphs should start with heading as given in below example. The emission of CO<sub>2</sub> to the atmosphere is steadily increasing. One of the serious impacts of increasing CO<sub>2</sub>, and other greenhouse gases, level in the atmosphere is global warming, which has several negative consequences. One of such negative consequence is the significant increase in the water level of oceans and seas and, thus, the submerging of some islands and coastal cities. Thus, CO<sub>2</sub> has to be captured and managed in a proper way. The current main process for CO<sub>2</sub> capture is the absorption using aqueous solutions of some amines such as methyldiethanolamine (MDEA) and monoethanolamine (MEA). In addition to environmental concerns on the utilization of these organic solvents (i.e., amines), amine absorption process is corrosive and the regeneration of the spent amine is energy intensive (i.e., expensive) [1-3]. Thus, alternative processes have been (and are still) actively sought by academia and relevant industries.

One of the attractive alternatives is adsorption. Nonetheless, several adsorbents reported in the literature are either not regenerable or/and require relatively high regeneration energy. Furthermore, some adsorbents are only effective at high pressures or sub-zero temperatures. Accordingly, there is a pressing need for developing CO<sub>2</sub> adsorbents that



are effective at (or near) ambient conditions. Besides their effectiveness at ambient conditions, these adsorbents should be stable and fully regenerable under mild conditions. Thus, MgO has been synthesized in this work and used for CO<sub>2</sub> capture. To enhance the CO<sub>2</sub> adsorption capacity, MgO was functionalized with APTES. The kinetics of CO<sub>2</sub> adsorption on both MgO and the APTES-modified MgO was fitted using three different adsorption kinetics models (pseudo-first order, pseudo-second order and Avrami), enabling the extraction of the values of the adsorption parameters.

## 2. Methodology

### 2.1. Preparation of adsorbents

MgO was synthesized using the following protocol. First, an aqueous solution (270 mM) of magnesium nitrate hexahydrate was prepared. After obtaining a homogenous solution of this magnesium salt via stirring at room temperature for 10 min, a certain volume of 30% ammonium hydroxide (NH<sub>4</sub>OH) was added to make the molar ratio of Mg: NH<sub>4</sub>OH as 1:5. The new mixture was stirred for few minutes to thoroughly mix its ingredients. After that, the glass bottle containing the mixture was closed and incubated in a hot water bath (60 °C) for 6 h. The mixture was stirred while incubated in the hot water bath. Then, the mixture was removed from the hot water bath, the bottle cap was removed (to allow some liquid to evaporate) and the bottle content was stirred for additional 24 h at room temperature, leading to the formation of a thick paste (consists of Mg(OH)<sub>2</sub> and NH<sub>4</sub>NO<sub>3</sub>) [4]. To remove the formed NH<sub>4</sub>NO<sub>3</sub>, which is water-soluble, the thick paste was washed extensively with distilled water. Such an extensive wash is expected to also remove any unreacted magnesium precursor. After drying at 60 °C, the material was ground and calcinated (under an air environment) at 400 °C for 10 h to yield MgO. The synthesized MgO was characterized using BET and energy-dispersive X-ray spectroscopy (EDX).

After the synthesis of MgO, it was functionalized with APTES according to the following procedure. An aqueous dispersion of MgO containing 0.83 g/L was prepared by vigorously stirring the MgO solution for 5 min. After that a certain amount of APTES (MgO/APTES mass ratio = 1/2) was added to the MgO suspension. Then, the resultant solution was heated at 80 °C and refluxed for one day. Following this step, the obtained amine-functionalized MgO was rinsed thoroughly with distilled water. The produced APTES-MgO was collected via centrifugation and then it was dried at 60 °C overnight. The synthesized APTES-MgO was also characterized using BET and EDX.

### 2.2. Adsorption kinetics of CO<sub>2</sub>

The CO<sub>2</sub> adsorption kinetics on the bare and the APTES-modified MgO was monitored using a thermogravimetric analysis (TGA) technique. In each experiment, the TGA pan (made from alumina) was thoroughly cleaned and dried and then a certain amount (roughly 20 mg) of either MgO or APTES was placed in the TGA pan. To remove moisture, the adsorbent was degassed at 100 °C using a stream of nitrogen. The mass of the adsorbent was continuously monitored while degassing until it reached a steady value and remained constant for at least 30 min. After that, the temperature was dropped to the desired adsorption temperature (i.e., 30 °C). The temperature was continuously monitored until it stabilized at 30 °C and remained stable. Then, nitrogen gas valve was closed and the carbon dioxide (> 99.9% pure) valve was opened. CO<sub>2</sub> flowrate was kept constant at 100 mL/min. To ensure the attainment of adsorption equilibrium, the

adsorption of CO<sub>2</sub> was monitored in time until no significant change in the sample mass (adsorbent + adsorbate) was observed.

### 2.3. Regeneration of spent amine-functionalized adsorbent

The regeneration of APTES-MgO (given the effectiveness of this adsorbent relative to the unmodified MgO as will be presented and discussed in the subsequent section) was studied. The regeneration was carried out using nitrogen gas flowing at 100 mL/min. The protocol of APTES-MgO regeneration is as follows. APTES-MgO was contacted with CO<sub>2</sub> until adsorption equilibrium was approached. Then, CO<sub>2</sub> flow (100 mL/min) was halted and N<sub>2</sub> valve was opened. Once, the adsorbent was fully regenerated (i.e., the mass of APTES-MgO returned to its original value before CO<sub>2</sub> injection), the flow of N<sub>2</sub> was stopped and CO<sub>2</sub> valve was opened. Four cycles of adsorption-regeneration were conducted. In each of these cycles, CO<sub>2</sub> adsorption took place at 30 °C while the regeneration temperature was set to 120 °C. The pressure of adsorption and regeneration processes was fixed at 1 atm.

## 3. Results and discussion

### 3.1. Characterization of the synthesized MgO and APTES-MgO

After the synthesis of MgO and its modification with APTES, elemental analysis of the composition of these two adsorbents was carried out using EDX. The results obtained from the elemental analysis are summarized in Table 1. The MgO sample contains about 53.61 atomic% magnesium and 46.39 atomic% oxygen, indicating the formation of magnesium oxide. The hypothetical chemical formulas of the formed magnesium oxide is MgO<sub>0.87</sub>, which is reasonably close to the chemical formula of magnesium oxide (MgO). Upon the modification of MgO with APTES, other elements (N, C, and Si) appeared in the elemental analysis results. These elements originate from the APTES molecules (molecular formula: C<sub>9</sub>H<sub>23</sub>NO<sub>3</sub>Si). The atomic ratios of N, C and Si in the APTES-MgO adsorbent are close to their atomic ratios in the APTES molecules. The increase in the oxygen content of the amine-modified MgO relative to MgO stems from the presence of oxygen in the APTES molecules.

Table 1: Elemental compositions of the synthesized MgO-based adsorbents.

Sample	Mg atomic%	O atomic%	C atomic%	N atomic%	Si atomic%
MgO	53.61	46.39	–	–	–
APTES-MgO-A	15.85	44.75	31.83	3.01	4.56

Besides the elemental analysis, textural properties (i.e., BET surface area, pore volume and pore diameter) of MgO and APTES-MgO were studied. The textural properties of the synthesized MgO and the amine-modified MgO are shown in Table 2. The BET surface area and pore volume of MgO were about 246 m<sup>2</sup>/g and 0.086 cm<sup>3</sup>/g, respectively. Upon the functionalization of MgO with APTES, the BET surface area and the pore volume of the modified MgO dropped by a factor of ~5. Such a reduction in the BET surface area of MgO and its pore volume upon the functionalization with APTES is expected since the amine molecules are likely to fill some of MgO pores [5, 6]. Unlike the drop in the BET

surface area and pore volume of APTES-MgO relative to the unmodified MgO, the pore size has increased by about 20% upon the attachment of APTES to MgO. The attachment of APTES to MgO-A was carried out at 80 °C under continuous stirring and refluxing for 24 h. Such a thermal treatment, despite being mild, coupled with APTES attachment to MgO-A might have assisted in expanding the size of the pores of the APTES-MgO-A sample [7]. The expansion in the pore size upon the functionalization of MgO with APTES does not contradict the observed reduction in pore volume and surface area. However, despite having higher pore size, the number of pores in the APTES-MgO sample is likely much lower than that in the unmodified MgO due to the blockage of some pores by APTES molecules.

Table 2: BET surface area, average pore size and average pore volume of the synthesized MgO-based adsorbents.

Samples	BET Surface area (m <sup>2</sup> /g)	Pore volume (cm <sup>3</sup> /g)	Pore size (nm)
MgO	246.421	0.086	1.389
APTES-MgO	48.651	0.020	1.644

### 3.2. Adsorption kinetics of CO<sub>2</sub>

CO<sub>2</sub> adsorption on MgO and APTES-MgO was studied at ambient conditions (i.e., 30 °C and 1 atm). Before contacting CO<sub>2</sub> with these two adsorbents, they were degassed using nitrogen gas as described in the previous section. Figure 1 shows the kinetics of CO<sub>2</sub> adsorption on the unmodified MgO. The adsorption capacity approached about 0.68 mmol/g when the adsorption process approached equilibrium. The adsorption data were fitted to pseudo-first order, pseudo-second order and Avrami kinetics model as shown in Figure 1. Among these three models, the pseudo-second order kinetics model provides the best fit to the experimental data ( $R^2 = 0.9996$ ). The estimated parameter values of the CO<sub>2</sub> adsorption kinetics using these models are presented in Table 3.

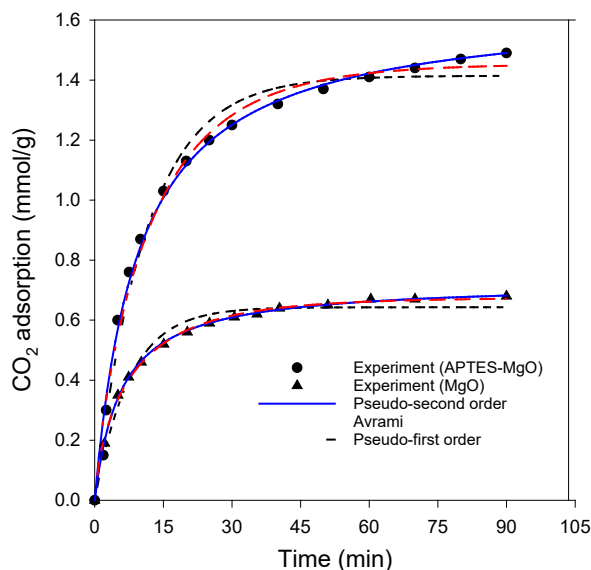


Figure 1: Adsorption kinetics of CO<sub>2</sub> on MgO and APTES-MgO. The solid curves represent the calculated values using the adsorption kinetics models.

The adsorption of CO<sub>2</sub> on APTES-MgO was also studied and modeled using the above adsorption kinetics models (see Figure 1). The estimated values of the adsorption parameters are shown in Table 3. The results displayed in Figure 1 and Table 3 (particularly the R<sup>2</sup> values) clearly demonstrate the superiority of the pseudo-second order in tracking the experimental data relative to the other two models. However, unlike the case of MgO where Avrami kinetics model provided a good data fitting, the fitting of CO<sub>2</sub> adsorption on APTES-MgO using this model is relatively poor. Nonetheless, Avrami model still provides a better data fitting for CO<sub>2</sub> adsorption on both MgO and APTES-MgO compare to the pseud-first order model.

Table 3: Values of the fitting parameters of CO<sub>2</sub> adsorption kinetics.

MgO					
Model	Model Equation	Adsorption rate constant (k)	Adsorption capacity at equilibrium (qe)	Avrami exponent (n)	R <sup>2</sup>
Pseudo-first order	$q_t = q_e(1 - e^{-kt})$	0.131 min <sup>-1</sup>	0.643 mmol.g <sup>-1</sup>	-----	0.9849
Pseudo-second order	$q_t = kq_e^2 t / (1 + kq_e t)$	0.241 g.mmol <sup>-1</sup> .min <sup>-1</sup>	0.725 mmol.g <sup>-1</sup>	-----	0.9996
Avrami	$q_t = q_e(1 - e^{-(kt)^n})$	0.117 min <sup>-0.706</sup>	0.675 mmol.g <sup>-1</sup>	0.706	0.9980
APTES-MgO					
Model	Model Equation	Adsorption rate constant (k)	Adsorption capacity at equilibrium (qe)	Avrami exponent (n)	R <sup>2</sup>
Pseudo-first order	$q_t = q_e(1 - e^{-kt})$	0.089 min <sup>-1</sup>	1.414 mmol.g <sup>-1</sup>	-----	0.9879
Pseudo-second order	$q_t = kq_e^2 t / (1 + kq_e t)$	0.064 g.mmol <sup>-1</sup> .min <sup>-1</sup>	1.647 mmol.g <sup>-1</sup>	-----	0.9951
Avrami	$q_t = q_e(1 - e^{-(kt)^n})$	0.082 min <sup>-0.846</sup>	1.454 mmol.g <sup>-1</sup>	0.846	0.9921

### 3.3. Regeneration of spent amine-functionalized adsorbent

In addition to the good CO<sub>2</sub> adsorption capacity on the APTES-functionalized MgO, it has also an excellent regenerability as shown in Figure 2. The CO<sub>2</sub> adsorption capability of APTES-MgO is preserved with the repetitive use. This means that the adsorbent is stable and the regeneration process of the spent adsorbent does not compromise its adsorption capacity. This is an attractive characteristic that is highly required in industrial applications. Another attractive characteristics of APTES-MgO is the mild regeneration conditions (120 °C and 1 atm). Exposing the spent APTES-MgO to N<sub>2</sub> at 120 °C and 1 atm can fully restore the original CO<sub>2</sub> adsorption capacity of this adsorbent within a very short time (< 20 min). This regeneration time can shortened with increasing the flow rate of the regeneration gas.

## 4. Conclusions

The modification of MgO with APTES greatly enhances CO<sub>2</sub> adsorption by a factor of more than 2 despite the significant drop (more than 5 times) in the adsorbent surface area

and pore volume. Kinetically, the adsorption of CO<sub>2</sub> on both adsorbents follows a pseudo-second order adsorption kinetics. The pseudo-first order model failed to reproduce the experimental data while the data fitting using the Avrami model was not as good as using the pseudo-second order. The full restoration of the adsorption capability of APTES-MgO upon the contact with N<sub>2</sub> at mild conditions in addition to the stability of the adsorbent make it a good candidate for commercial CO<sub>2</sub> capture.

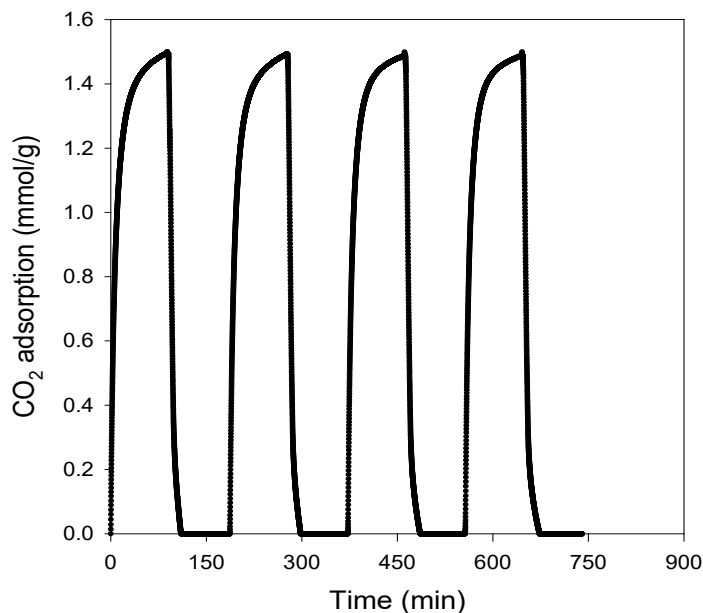


Figure 2: Regeneration of APTES-MgO using nitrogen gas. The regeneration temperature and pressure are 120 °C and 1 atm, respectively.

## References

- [1] P. Luis, 2016, Use of monoethanolamine (MEA) for CO<sub>2</sub> capture in a global scenario: Consequences and alternatives, *Desalination*, 380:93-99.
- [2] J. Liu, S. Wang, B. Zhao, H. Tong, C. Chen, 2009, Adsorption of carbon dioxide in aqueous ammonia, *Energy Procedia*, 1:933-940.
- [3] R. Strube, G. Pellegrini, G. Manfreda G, 2011, The environmental impact of post-combustion CO<sub>2</sub> capture with MEA, with aqueous ammonia, and with an aqueous ammonia-ethanol mixture for a coal-fired power plant, *Energy*, 36:3763-3770.
- [4] A. Hanif, S. Dasgupta, A. Nanoti A, 2016, Facile Synthesis of High-Surface-Area Mesoporous MgO with Excellent High-Temperature CO<sub>2</sub> Adsorption Potential, *Industrial & Engineering Chemistry Research*, 55:8070-8078.
- [5] X. Xu, C. Song, J. M. Andresen, B. G. Miller, 2002, Saroni AW. Novel Polyethylenimine-Modified Mesoporous Molecular Sieve of MCM-41 Type as High-Capacity Adsorbent for CO<sub>2</sub> Capture, *Energy & Fuels*. 16:1463-9.
- [6] S.-H. Liu, W.C. Hsiao, W.-H. Sie, 2007, Tetraethylenepentamine-modified mesoporous adsorbents for CO<sub>2</sub> capture: effects of preparation methods, *Adsorption*, 18:431-7.
- [7] Y. Matsuo, Y. Nishino, T. Fukutsuka, Y. Sugie, 2007, Introduction of amino groups into the interlayer space of graphite oxide using 3-aminopropylethoxysilanes, *Carbon*, 45:1384-1390.

## Development of 'Green' LNG through a CO<sub>2</sub> allocation procedure within a CO<sub>2</sub> utilisation network in Qatar

Razan Sawaly, Mohammad Alherbawi, Tareq Alansari\*

*College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar*  
*\*Corresponding email: talansari@hbku.edu.qa*

### Abstract

Natural gas has a crucial role in minimising carbon emissions and is considered a transition fuel. Liquefied natural gas (LNG), which is derived from natural gas is one of the most promising alternative fuels, and LNG producers are actively developing their methods to reduce or offset their carbon footprints. The allocation of carbon dioxide (CO<sub>2</sub>) emissions is crucial in identifying who is responsible for reducing emissions at the economy or enterprise level. Different strategies for allocating CO<sub>2</sub> emissions have been developed and used during the last few decades. As such, this study introduces options for low carbon LNG within a CO<sub>2</sub> utilisation network in the state Qatar; where a linear programming (LP) model is developed to address the optimal CO<sub>2</sub> allocation based on environmental and economic objectives. The model considers a single CO<sub>2</sub> source (Qatar Gas), where a post combustion carbon capture technology (PCC) is utilised to capture the CO<sub>2</sub> emissions, and then allocated to 7 sinks that use CO<sub>2</sub> as input to their production (including Enhanced Oil Recovery (EOR)). Three scenarios are considered to allocate CO<sub>2</sub> to different sinks to obtain the optimal solution for each scenario. Scenario 1 had the highest total profit of 55 billion USD. The lowest CO<sub>2</sub> emissions were produced by scenario 2, which had zero emissions. Scenario 3 had a total profit of 20 billion USD and a total of 2 million CO<sub>2</sub> emissions. For all scenarios, EOR sinks had the highest CO<sub>2</sub> allocation. The total global warming potential through the lifecycle of LNG is reduced from 0.46 to 0.14 tonne CO<sub>2</sub>e/tonne of LNG produced.

**Keywords:** LNG, CO<sub>2</sub> allocation, Sustainability, Carbon Dioxide, Carbon capture and utilisation (CCU)

### 1. Introduction

Due to the growing environmental concerns, there are indicators that expanding symbiotic linkages between industries within the same vicinity may significantly contribute to the development of sustainable industrial operations. Several energy corporations have recently revealed far-reaching carbon-neutrality goals, accompanied by more or less remote time horizons. Regardless of the fact that natural gas is the cleanest-burning hydrocarbon, regulators are paying closer attention to GHG emissions linked to the natural gas supply chain, including that of liquefied natural gas (LNG) (Roman-White et al., 2021). To guarantee natural gas's place in the energy transition in a decarbonizing world with a rising proportion of renewable energy sources, natural gas emissions must be reduced, and this can be made possible through a systems approach on an industrial level. For resource allocation in industrial park, a variety of tools has been utilised previously, including optimisation, game theory, agent-based modeling, network analysis, etc. Due to the established technologies in many industries, which are mostly focused on the use of fossil fuels, the direct reduction of the emissions is currently complex in the short term. Therefore, the concept of Carbon Capture and Storage (CCS) arises as a compelling proposal for reducing emissions, where absorption has been

demonstrated as an effective approach to produce a high capture yield with various solvents, such as piperazine (PZ) or monoethanolamine (MEA) (Li and Zhang, 2018). The requirement of large storage capacity of carbon dioxide gas is a challenge that affects the adoption of CCS technologies (IPCC, 2005). Furthermore, Carbon Capture and Utilisation (CCU) aims to gain an advantage by using CO<sub>2</sub> in various industrial processes in place of conventional raw materials (Aresta, 2010). Utilising CO<sub>2</sub> is a possibility for supplying a renewable energy source for the creation of numerous valued items. The optimum method of utilising CO<sub>2</sub> is thought to be its conversion into fuel. Some of the chemicals that can be generated by using captured CO<sub>2</sub> as an input or feedstock include methane, methanol, alkanes, and syngas (Marchese et al., 2021).

Research on energy conservation in an eco-industrial park (EIP) is mostly unexplored in the field of mathematical optimisation. Boix et al. (2015) stated that many objective functions, such as EIP energy efficacy, CO<sub>2</sub> emission reductions, waste heat recovery, and network payback period, have been proposed in the literature. However, the challenge of achieving multiple conflicting objectives at once is not efficiently examined yet as well as carbon focused eco-park resource allocation with an aim of producing low carbon and zero-carbon LNG. As such, this study is an extension of previous study conducted by Sawaly et. al, (2022), with the integration of a multi-objective model and Alshaheen field as a new EOR sink. A linear programming (LP) model is developed to address the optimal CO<sub>2</sub> allocation with multi-objective to maximise the overall profit of the system while minimising CO<sub>2</sub> emissions. The model considers a single CO<sub>2</sub> source (Qatar Gas), where a post-combustion carbon capture technology (PCC) is utilised to capture the CO<sub>2</sub> emissions, and then allocated to sinks that use CO<sub>2</sub> as input in their production. Thus, this study aims at introducing additional options for low-carbon LNG within a CO<sub>2</sub> capture and utilisation (CCU) network in the state Qatar to optimise and enhance eco-parks exchanges.

## 2. Model Design and Optimisation

### 2.1 Available data and assumptions

The objective is to facilitate a resource trade scheme for CO<sub>2</sub> emissions to produce low carbon LNG in Qatar Gas. With a total of roughly 21 LNG trains, Qatar Gas (QG) is the largest LNG producer in the world, operating 14 LNG trains with a 77 MT annual production capacity (Al-Yaeshi, 2020). According to Mohammed, (2016), the corresponding CO<sub>2</sub> emissions are 35.8 MT/year (Scope 1 – Direct total GHG emissions). Therefore, a post combustion carbon capture technology (PCC) is utilised to capture 70% of the total CO<sub>2</sub> emissions emitted by QG (25 MT year), and then allocate to sinks that use CO<sub>2</sub> as input to their production (QAFCO, QAFAC, Pearl GTL, Oryx GTL, Dukhan Field Well (EOR), Alshaheen Field (EOR)).

$$\text{Enhanced Lifecycle GWP of Qatar Gas} = 30\% \times \frac{\text{tonne of CO}_2 \text{ emitted/year}}{\text{tonne of LNG produced/year}} \quad (1)$$

With an OPEX of \$40 per tonne of CO<sub>2</sub>, chemical absorption using amine solvents was expected to capture CO<sub>2</sub> with 100% efficiency. In order to prevent pressure reductions throughout the pipeline and to preserve the operational parameters of the sink, CO<sub>2</sub> must be compressed. For every 100 km along the pipeline, the unit cost for compression and subsequent transportation is roughly \$1.5/tonne CO<sub>2</sub> (Al-Yaeshi, 2020; Sawaly et. al, 2022). It is assumed that the injection cost and emissions for EOR is neglected. Table 1 lists the sinks' products, production rate, CO<sub>2</sub> capacity, and emission rate. The suggested methodology uses a Linear Programming (LP) optimisation framework to address CO<sub>2</sub> allocation and find solutions for reducing CO<sub>2</sub> emissions at a realistic cost while

also maximising the profits of value-added products. The proposed methodology is used to implement the model's outcomes using the Excel solver and MATLAB. Carbon capture and compression, along with transportation are the only expenditures taken into account in this study that have an influence on profit, whilst the only emissions included are process related emissions. Three scenarios are implemented in this study: (1) A scenario for allocating CO<sub>2</sub> from the source to the sinks with a single objective of maximising the profit of the overall network (2) A scenario for allocating CO<sub>2</sub> from the source to the sinks with a single objective to minimise the CO<sub>2</sub> emission of the overall network, and; (3) A scenario for allocating CO<sub>2</sub> from the source to the sinks with a multi-objective for profit maximisation and emission minimisation of the overall network. The commodity prices of the products are collected over the period 2014-2018 (Intratec Solutions, 2019).

Table 1: Characteristics of the sinks (Sawaly et al., 2022; Ghiat et al., 2021; IEA, 2018; Mohammed, 2016).

Sink	Process/ product	Emission Rate	Production Rate	CO <sub>2</sub> Capacity	Distance from source
Unit	-	Tonne CO <sub>2</sub> out/tonne CO <sub>2</sub> feed	Tonne of product/tonne of CO <sub>2</sub>	Million tonne/year	Kilometre
QAFAC H <sub>2</sub>	Methanol from H <sub>2</sub>	0.02	5.9	1.83	100
QAFAC NG	Methanol from NG	2.09	3	0.60	100
QAFCO	Urea	0.23	1.35	1.01	100
Pearl GTL	ATR-GTL-Wax	2.33	0.884	2.19	5
Oryx GTL	ATR-GTL-Wax	9.75	0.884	0.36	3
Dukhan Field	EOR	0	0.2	15.6	111
AlShaheen Field	EOR	0	0.32	13.9	80

## 2.2 Model Formulation

$$\text{Maximise } \sum_{i=1}^7 \sum_{j=1}^5 x_{ij} [P_{ij} PR_{ij} Q - C_{operating,ij}] \quad (2)$$

$$\text{Minimise } \sum_{i=1}^7 \sum_{j=1}^5 x_{ij} Q ER_i \quad (3)$$

Subject to

$$\sum_{i=1}^7 \sum_{j=1}^5 x_{ij} Q PR_i \leq C_{Pij} \quad (4)$$

$$\sum_{i=1}^7 x_{ij} = 1 \quad (5)$$

$$\text{Where; } C_{operating,ij} = \sum_{i=1}^7 \sum_{j=1}^5 (C_{transport} + C_{pcc} + C_{OPEX}) \quad (6)$$

$$C_{transport} = Q x_{ij} \times \left( \frac{D_i}{100} \right) \times 1.5 \quad (7)$$

$C_{operating,ij}$  = Cost for pcc, transport, OPEX of sink i for year j

$C_{p,ij}$  = Capacity of sink i for year j

$Q$  = Carbon Captured from source (tonne/year)

$x_{ij}$  = Fractional amount allocated to sink i for year j

$PR_{ij}$  = Production Rate of sink i for year j as a function of carbon feed

$ER_{ij}$  = Emission Rate of sink i for year j as a function of carbon feed

$P_{ij}$  = Product price of industry/sink i for year j

$D_i$  = Distance from the source to sink i

## 3 Results and Discussion

The proposed methodology is used to implement the model's outcomes using the Excel solver for scenario 1 and scenario 2, and MATLAB for scenario 3. **Scenario 1** aims to



determine the optimal solution for allocating 25x10<sup>6</sup> tonne/year of CO<sub>2</sub> to all 7 sinks with an objective to maximise the profit of the overall system. According to the findings, the market prices of the products and production rate are the key drivers of the techno-economic optimisation. The total annual profits for all sinks ranged from 8.1 to 14.6 billion USD and the total profit of the overall system throughout the specified time period is 55 billion USD. The annual amount of CO<sub>2</sub> emitted by the sinks is 10.1 million tonnes CO<sub>2</sub>/year and the total emissions of the overall system throughout the specified time period is 50.7 million tonnes CO<sub>2</sub>. Figure 1 represents the allocation across the sinks, with QAFACO, QAFAC H<sub>2</sub>, QAFAC NG, Pearl GTL, and Oryx GTL being fully satisfied due to their higher commodity prices and smaller CO<sub>2</sub> capacity in comparison to Dukhan field well and Alshaheen field, which use CO<sub>2</sub> for enhanced oil recovery. Alshaheen field has a larger share of CO<sub>2</sub> allocation than Dukhan field, which is due to Alshaheen field's higher production rate and consequently generates more profit. The total CO<sub>2</sub> utilisation of the system is 25x10<sup>6</sup> tonne/year, which implies that the captured amount of CO<sub>2</sub> from Qatar Gas (25MT/year) is fully utilised by the sinks. The allocation ratio throughout the years is relatively in the same range for all years between 2014 and 2018.

This is due to the price linkage between products, which causes them to shift together. In **Scenario 2**, the goal is to find the best way to allocate 25x10<sup>6</sup> tonnes of CO<sub>2</sub> per year among the seven sinks with the least amount of emissions possible for the overall network. The allocation's primary factor, as indicated by the results, was each sink's emission rate. During enhanced oil recovery, the majority of the CO<sub>2</sub> that is injected stays there trapped in the geological formation, causing little to no emissions of CO<sub>2</sub>. Therefore, the allocation will shift toward the sink with the lowest emissions rate in the case of emission reduction. The captured CO<sub>2</sub> was allocated equally to Dukhan Field and Alshaheen field (figure 2) since their process generates zero emissions. The total annual profits for all sinks ranged from 1.43 to 5.09 billion USD and the total profit of the overall system throughout the specified time period is 14.6 billion USD.

**Scenario 3** was generated using MATLAB for finding a set of solutions that define the best trade-off between the competing objectives. The model generated 70 different Pareto optimal solutions (figure 4), however, most of the solutions favored EOR sinks (both, Dukhan Field and AlShaheen Field), which is because they generate zero emissions, and got higher CO<sub>2</sub> capacity allowing for more profit as well regardless of the fact that they both have the lowest production rate. Furthermore, following the EOR sinks, QAFAC H<sub>2</sub> obtained the highest allocation percentage due to its high production rate and minimal CO<sub>2</sub> emissions. Out of the 70 solutions that were developed, the one with the highest profit had a GWP of 4.5 MT CO<sub>2</sub>e and a profit of 22.6 billion USD. The solution that achieved the lowest GWP had a profit of 15 billion USD and zero GWP (zero emissions). The mean solution had a profit of 20 billion USD and a GWP of 2 MT CO<sub>2</sub>-eq. Figure 3 illustrates the CO<sub>2</sub> allocation for the mean solution.

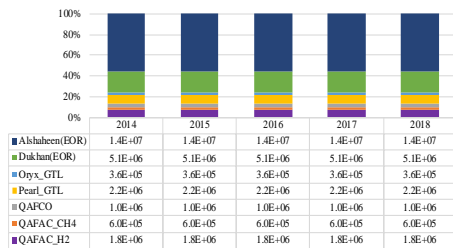


Figure 1: CO<sub>2</sub> Allocation of Scenario 1.

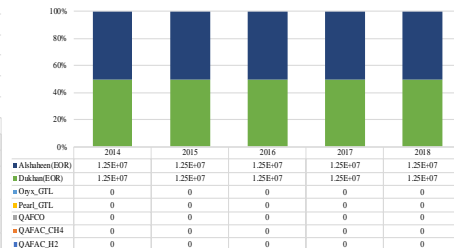


Figure 2: CO<sub>2</sub> Allocation of Scenario 2.

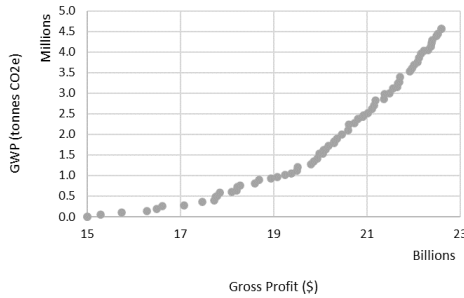


Figure 4: Pareto Front.

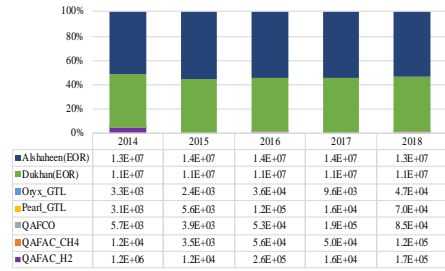


Figure 3: CO<sub>2</sub> Allocation of Scenario 3.

Figure 5 compares the total CO<sub>2</sub> emissions and total profit of scenario 1 and 2 with the mean solution of scenario 3. Scenario 2 witnessed a huge decrease in the profit and CO<sub>2</sub> emission compared to scenario 1; where CO<sub>2</sub> emissions were reduced by 100% and the profit was decreased by around 73%. Scenario 3 on the other hand aims to satisfy both objectives simultaneously. Therefore, CO<sub>2</sub> emissions are reduced significantly from 50.7 million tonnes CO<sub>2</sub> (scenario 1) to 2 million tonnes CO<sub>2</sub> (scenario 3), while the profit fell sharply from 55 billion USD to 20 billion USD for scenarios 1 and 3 respectively. When comparing scenario 2 and 3, the total profit increased slightly from 14.6 billion USD (scenario 2) to 20 billion USD (scenario 3). The total GWP through the lifecycle of LNG in Qatar Gas was 0.46 tonne CO<sub>2</sub>e/tonne of LNG produced. This ratio has been enhanced through this model and reduced to 0.14 (70% cut of emissions) through CO<sub>2</sub> capture and allocation by substitution approach.

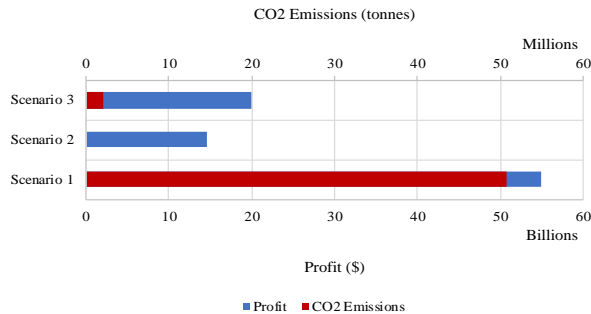


Figure 5: Comparison between scenario 1,2 and 3.

#### 4 Conclusion

The findings of this study aim to provide the most efficient strategies for producing low carbon LNG within a CO<sub>2</sub> utilisation network in the state Qatar by converting carbon dioxide emissions into value added products. The model considers a single CO<sub>2</sub> source (Qatar Gas) and 7 sinks. Furthermore, the optimal CO<sub>2</sub> allocation based on environmental and economic objectives is addressed using a linear programming (LP) model where 3 scenarios were generated. Scenario 1 aimed to maximise the overall profit of the system and had a value of 55 billion USD; whereas scenario 2 aimed to minimise the CO<sub>2</sub> emissions which reached to zero emissions. Scenario 3 generated a profit of 20 billion USD and 2 million CO<sub>2</sub> emissions overall. According to the results of using allocation by substitution, only 0.14 tonne CO<sub>2</sub>e/tonne LNG is emitted whereas 0.32 tonne CO<sub>2</sub>e/tonne LNG is captured and utilised by the sinks. The lack of actual selling market price and plant data for each of the products studied between 2014 and 2018 are some of the

limitations of this study. Also, the production rate is calculated based on assumptions derived from literature that may or may not represent real operating data for the relevant time period. Future work would include: 1) implementation of carbon tax to the model, which can have an effect on both the environmental and economic objective; 2) capture a larger amount of CO<sub>2</sub> from Qatar Gas (85% and 100%).

## 5 Acknowledgement

The authors acknowledge the support of Qatar National Research Fund (a member of Qatar Foundation) provided by GSRA grants no. GSRA9-L-1-0531-22069.

## 6 References

- Al-Yaeshi, A. A., Govindan, R., & Al-Ansari, T. (2020). Techno-economic-based dynamic network design for optimum large-scale carbon dioxide utilisation in process industries. *Journal of Cleaner Production*, 275, 122974. <https://doi.org/10.1016/j.jclepro.2020.122974>
- Aresta, M. (2010). Carbon dioxide: Utilization options to reduce its accumulation in the atmosphere. *Carbon Dioxide as Chemical Feedstock*, 1–13. <https://doi.org/10.1002/9783527629916.ch1>
- Boix, M., Montastruc, L., Azzaro-Pantel, C., & Domenech, S. (2015). Optimization methods applied to the design of Eco-industrial parks: A literature review. *Journal of Cleaner Production*, 87, 303–317. <https://doi.org/10.1016/j.jclepro.2014.09.032>
- Ghiat, I., Alnouss, A., & Al-Ansari, T. (2021). Superstructure Optimisation in Various Carbon Capture and Utilisation Supply Chains, PSE 2021+ Conference.
- Iea. (2018, November 1). Can CO<sub>2</sub>-Eor really provide carbon-negative oil? – analysis. IEA. Retrieved November 29, 2022, from <https://www.iea.org/commentaries/can-co2-eor-really-provide-carbon-negative-oil>
- Intratec Solutions. (2019). Commodity prices - full sample - Historical & Forecast Data in several countries. Intratec.us. Retrieved February 1, 2023, from <https://www.intratec.us/products/primary-commodity-prices>
- IPCC 2005. IPCC special report on carbon dioxide capture and storage. prepared by working group iii of the intergovernmental panel on climate change. In IPCC Special Report on Carbon Dioxide Capture and Storage 2:442. doi: 10.1002/anie.201000431.
- Li, H., & Zhang, Z. (2018). Mining the intrinsic trends of CO<sub>2</sub> solubility in blended solutions. *Journal of CO<sub>2</sub> Utilization*, 26, 496-502.
- Marchese, M., Buffo, G., Santarelli, M., & Lanzini, A. (2021). CO<sub>2</sub> from direct air capture as carbon feedstock for Fischer-Tropsch chemicals and fuels: Energy and Economic Analysis. *Journal of CO<sub>2</sub> Utilization*, 46, 101487. <https://doi.org/10.1016/j.jcou.2021.101487>
- Middleton, R.; Bielicki, J. A Comprehensive Carbon Capture and Storage Infrastructure Model. *Energy Procedia*. 2009, 1, 1611-16.
- Mohammed, S.S., 2016. Qatar's National Emission Inventory Repor. Qatar Environment and Energy Research Institute , Qatar Foundation/Hamad Bin Khalifa University, Figshare open access book license CC BY-4.0. <http://creativecommons.org/licenses/by/4.0>.
- Roman-White, S. A., Littlefield, J. A., Fleury, K. G., Allen, D. T., Balcombe, P., Konschnik, K. E., Ewing, J., Ross, G. B., & George, F. (2022). Response to comment on “LNG Supply Chains: A supplier-specific life-cycle assessment for improved emission accounting.” *ACS Sustainable Chemistry & Engineering*, 10(41), 13552–13554. <https://doi.org/10.1021/acssuschemeng.2c05197>
- Sawaly, R., Ghiat, I., Mohamed, A., Abushaikh, A., & Al-Ansari, T. (2022). Optimal CO<sub>2</sub> allocation for enhanced oil recovery operations within Carbon Utilisation Networks in Qatar. *Computer Aided Chemical Engineering*, 493–498. <https://doi.org/10.1016/b978-0-323-95879-0.50083-7>

# Process Design and Bayesian Optimization of 5-Hydroxymethylfurfural Hydrodeoxygenation

Yuqing Luo<sup>a</sup>, Zhaoxing Wang<sup>a,b</sup>, Prahalad Srinivasan<sup>a</sup>, Dionisios G. Vlachos<sup>a,b</sup>,  
Marianthi Ierapetritou<sup>a,b\*</sup>

<sup>a</sup>*Department of Chemical and Biomolecular Engineering, University of Delaware, 150  
Academy Street, Newark, Delaware, 19716, USA*

<sup>b</sup>*Catalysis Center for Energy Innovation, RAPID Manufacturing Institute, and Delaware  
Energy Institute (DEI), University of Delaware, Newark, Delaware 19716, United States  
[mgi@udel.edu](mailto:mgi@udel.edu)*

## Abstract

The hydrodeoxygenation (HDO) of 5-hydroxymethylfurfural (HMF) is an essential step in lignocellulosic biomass conversion. In this work, we developed process flowsheet simulation models using the hydrodeoxygenation experimental conditions, yields, and separation requirements. Economic and environmental evaluations are then performed and incorporated with Bayesian optimization to search for reaction conditions that minimize production costs and greenhouse gas emissions. The Bayesian optimization demonstrates that the HDO reaction condition with the lowest production cost has a relatively high reaction temperature, no water content, high H<sub>2</sub> pressure, medium HMF loading, and short reaction time. However, the HDO condition with the lowest greenhouse gas emission has a lower temperature but a longer reaction time.

**Keywords:** Bayesian optimization, 5-Hydroxymethylfurfural, Techno-economic analysis, Life cycle assessment.

## 1. Introduction

5-Hydroxymethylfurfural has been established as a promising platform chemical from lignocellulosic biorefinery after hexose dehydration [Bozell and Petersen, 2010]. HMF typically further undergoes hydrodeoxygenation to produce the 2,5-dimethylfuran (DMF) intermediate for the downstream production of other renewable chemicals [Luo and Ierapetritou, 2020]. Over the Ru/C catalyst, HMF is first hydrogenated to 2,5-bishydroxymethylfuran (BHMF), whose alcohol group undergoes hydrogenolysis to produce DMF with the hydrogen produced from the dehydrogenation of a secondary alcohol as the solvent. The furans also undergo etherification, which slowly convert to DMF [Jae *et al.*, 2014]. Water inevitably exists in the feed and process streams as a component of biomass or a byproduct of dehydration and HDO reactions [Esteban *et al.*, 2020]. For instance, in the upstream HMF production (fructose dehydration), many water-organic biphasic systems are developed to prevent rehydration or condensation side reactions [Wang *et al.*, 2020]. As water is partially soluble in the organic phase, the HMF product stream could contain a significant amount of water. The presence of water is observed to impact the ether-alcohol-furan equilibrium and affects the reaction rate. However, removing water from intermediate streams may be energy-intensive and

expensive due to the high heat of vaporization. Therefore, it is important to study the trade-off of the water effects based on process simulation.

Bayesian optimization is an effective active-learning method that combines building the surrogate model and proposing the next experiment/simulation points [Wang and Dowling, 2022]. It thus provides an efficient solution for various practical optimization tasks involving time-consuming computation or experiments, which expedites the discovery of optimal reaction conditions and new materials [Iwama and Kaneko, 2021]. For instance, Diwale *et al.* applied Bayesian optimization under noise to reduce the time spent on computationally expensive non-equilibrium molecular dynamics simulations for the search of additives in polyethylene nucleation. Bayesian optimization also finds its application in the model predictive control parameter tuning [Lu *et al.*, 2021].

In this work, we first built HMF conversion and DMF yield surrogate models as a function of reaction temperature, time, H<sub>2</sub> pressure, HMF loading, and water content using the adaptive design of HDO experiments based on Bayesian optimization. Next, Aspen Plus process flowsheets are simulated using the reaction yields/conversion and appropriate separation process design, which provided information for the techno-economic analysis (TEA) and life cycle assessment (LCA). Next, Bayesian optimization is applied again to search optimal reaction conditions that perform the HDO reaction with the lowest DMF minimum selling price (MSP) and global warming potential (GWP). The results also revealed the effects of water content and reaction conditions on the overall HDO process economics and environmental impacts.

## 2. HMF Hydrodeoxygenation Flowsheet Design and Analysis

First, the NEXTorch (Next EXperiment toolkit in PyTorch) python package was utilized to build accurate HMF conversion and DMF yield surrogate models from the preliminary bench-top HDO reaction experiments *via* Gaussian processes, which ensures the exploration of the design space while exploiting the current best conditions [Wang *et al.*, 2021]. In this case, a multi-variable (reaction temperature, time, H<sub>2</sub> pressure, HMF loading, and water content) ordinary least squares (OLS) regression demonstrated satisfactory fitting performance for DMF yield and HMF conversion. This OLS model was used in section 3 for the process design in each step of the Bayesian optimization.

Process flowsheets are then simulated in Aspen Plus V11 for 29 initial HDO reaction conditions and their corresponding yields/conversions. These conditions affect not only the reaction outcome but also the separation process. The universal quasi-chemical activity coefficient (UNIQUAC) model was chosen for the liquid-vapor and liquid-liquid phase behavior [Abrams and Prausnitz, 1975]. Four different flowsheet configurations are developed for various pentanol/water azeotrope and feedstock compositions of the HDO processes. Figure 1 illustrates the process flowsheet when no water exists in the reaction inlet. Hydrogen gas and HMF solution in pentanol were pressurized and heated to the reaction condition before sending to R-1. After the HDO reaction, flash drum F-1 cooled down the system to 65 °C and recycled hydrogen gas. Then, humins were removed by filtration. It was assumed that most of the liquid phase was retained with a 1% loss in the waste humins stream. The liquid stream was sent to two distillation columns in series (C-1 and C-2) to recycle pentanol and obtain a DMF/water stream. Two decanters cooled down both the bottom (D-1) and distillate streams (D-2) to 40 °C, so that excess water was removed by phase separations. The process capacity was chosen as 10,000 kg HMF/h, which was in line with the scale considered in a recent paper that produced HMF

as an intermediate [Chang *et al.*, 2020]. It is noteworthy that separating the targeted DMF product from water/pentanol azeotrope is not always easy in the other flowsheets, as they have very close boiling points.

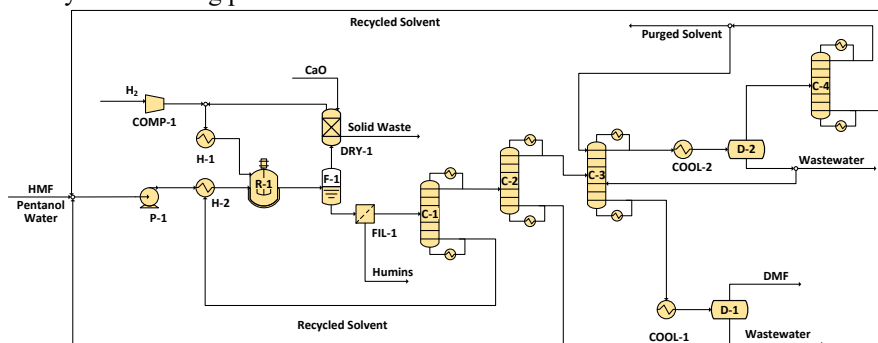


Figure 1: Hydrodeoxygenation process flowsheet without water in the feed

### 3. Bayesian Optimization

Bayesian optimization on the reaction conditions (i.e., water content, HMF loading, temperature, pressure, and reaction time) is conducted through iterative optimal search via NEXTorCh and flowsheet simulation/evaluation. The NEXTorCh package utilizes the Gaussian process surrogate model with the Matern covariance function kernel. In each Bayesian optimization iteration, new reaction condition samples are suggested based on the Monte Carlo  $q$ -expected improvement (qEI) acquisition function of the Gaussian process. The expected improvement acquisition function has the ability to effectively leverage exploration and exploitation aspects of searching [De Ath *et al.*, 2021]. The Monte Carlo sampling approach shown in Eq.(1) further reduces the computational complexity by approximating the integral over posterior distribution [Snoek *et al.*, 2012].

$$qEI(X) = \frac{1}{N} \sum_{i=1}^N \max_{j=1, \dots, q} \{\max(\xi_{ij} - f^*, \mathbf{0})\}, \quad \xi_{ij} \sim \mathbb{P}(f(X) | D) \quad (1)$$

where  $X = (x_1, \dots, x_q)$ , representing the reaction conditions (e.g., temperature, pressure, reaction time, water, and HMF loading),  $q$  is the number of points considered jointly,  $f^*$  is the best value observed so far, and  $\mathbb{P}(f(X) | D)$  is the posterior distribution of the function  $f$  (e.g., MSP or GWP) at  $X$  given the observed data  $D = (X, f)$ . Individual flowsheets are developed based on these trial points, whose economic and environmental performance is evaluated and added to observed data for the next iteration.

## 4. Results and Discussions

### 4.1. Bayesian Optimization of Minimum Selling Prices (MSP)

Techno-economic analysis (TEA) was performed to calculate each case's MSP, the selling price of the product when the net present value is zero at the end of the recovery period, was determined through discounted cash flow analysis for each condition [Athaley *et al.*, 2019]. A 15% internal rate of return (ROR) on investment and a 35% corporate tax was assumed. The straight-line method for depreciation with a 10% salvage value after 20 years was applied [Luo *et al.*, 2022]. The raw material, solvent, and catalyst prices are based on Athaley *et al.* (2019). For instance, the HMF price is taken as \$0.88/kg, and the pentanol cost is assumed to be \$1.5/kg. As illustrated in Figure 2, the best DMF MSP gradually decreases as the iterations of Bayesian optimization continue, reducing from \$5,420/t to \$4,022/t. In all cases, the HMF feedstock is the leading cost

contributor. Hence, the MSP of DMF decreases significantly when the DMF yield increases and less humins byproduct is generated. This is typically the case with little or no water loading, which has high DMF selectivity. However, when the water content increased from 0% to 4% in the feedstock, the MSP only increased by \$ 312/t.

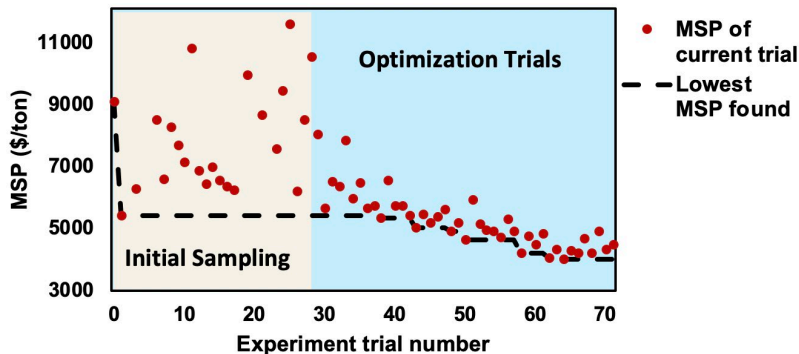


Figure 2: Discovery plot of DMF's MSP

On the one hand, high solvent usage leads to much larger equipment sizes and more solvent loss during filtration. On the other hand, reaction conversion and selectivity generally drop as a result of lower solvent usage. Hence, medium HMF concentration at 3 wt. % was favored by Bayesian optimization as a trade-off between these two effects.

#### 4.2. Bayesian Optimization of GWP

Next, a “cradle-to-gate” LCA was performed to evaluate the environmental impacts by including the upstream activities of raw material extraction (e.g., HMF), utility generation, and the production stages. The functional unit was chosen as 1 kg of DMF produced, and the GWP was selected as the environmental impact of interest as it bears great importance in renewable technology development and policy making. The Ecoinvent v3.7 database was used to provide the background emission data, and the Tool for the Reduction and Assessment of Chemical and other environmental Impacts (TRACI) method was used for the impact assessment [Bare, 2011]. The trajectory of GWP under each explored reaction condition is demonstrated in Figure 3. During these Bayesian optimization runs, the lowest GWP was found to be 22.5 kg CO<sub>2</sub> eq/kg DMF, while the best GWP from the initial 29 experiments was 26.5 kg CO<sub>2</sub> eq/kg DMF.

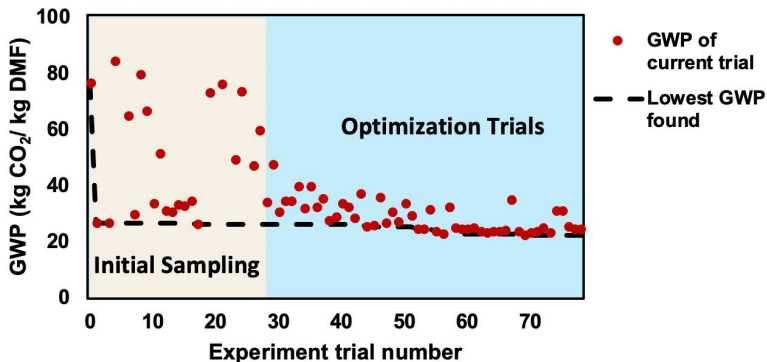


Figure 3: Discovery plot of DMF's GWP

Similar to the improvement trend in the MSP case, the GWP of each process generally matches the increase of HMF conversion and DMF selectivity. Medium HMF loading, high hydrogen gas pressure, and zero water content are still favored. However, as with

other biomass-based chemical production processes, the utility cost only takes up a small fraction of the total MSP while it plays a vital role in the GWP. Based on the conversion and selectivity model, higher temperatures correspond with better yields at fixed water content, HMF loading, and reaction time. Nevertheless, a higher reaction temperature also means more hot utilities to heat up the feedstock and solvent before the reaction. Thus, the optimal temperature for the lowest GWP case (199 °C) is lower than that of the lowest MSP case (210 °C). Other contributors to utility usage include the steam required for distillation and electricity used in the compressor.

Although longer reaction time causes larger reactor sizes and higher capital investment, it has little impact on environmental performance. Thus, the lowest GWP case has increased the reaction time from 3 h to 4.3 h at a lower temperature to improve the HMF conversion without worrying about more expensive equipment. It should be noted that non-zero water content in the feedstock inlet causes a drastic increase in the GWP. The process with 4% inlet water content now has a GWP of 31.1 kg CO<sub>2</sub> eq/kg DMF, which is 38.4% higher than the lowest GWP without any water loading. This is the result of high utility usage during the product separation, as the azeotrope compositions significantly affect the distillation operating conditions. A summary of optimal reaction conditions for the MSP and GWP Bayesian optimization is listed in Table 1.

Table 1: Optimal reaction conditions suggested by MSP and GWP Bayesian optimization

Reaction Conditions	minimum MSP	minimum MSP (with H <sub>2</sub> O)	minimum GWP	maximum DMF yield*
Temperature (°C)	210	208	199	195
H <sub>2</sub> pressure (bar)	10	10	10	10
Water loading (wt. %)	0	4	0	0
HMF loading (wt. %)	3	3	3	1
Time (h)	3	3	4.3	4.6

\* “max DMF yield” is based on the best condition in the initial design of experiment

As illustrated in Table 1, the reaction conditions with the lowest cost and emission are close to the maximum DMF yield condition observed in bench-top experiments. Nevertheless, Bayesian optimization incorporates process design insights to select higher HMF loadings and shorter reaction times that would be otherwise not explored.

## 5. Conclusions

The current work incorporated process simulation, TEA, and LCA into the Bayesian optimization to determine economically and environmentally viable HDO reaction conditions in the early stage. The Bayesian optimization results illustrate that the reaction conditions that minimize the production cost have low water contents, medium HMF loadings, high H<sub>2</sub> pressure, and relatively high reaction temperatures because of low HMF feedstock usage and low solvent loss. The HDO reaction with the minimal GWP chooses a lower reaction temperature but longer reaction time to reduce the emissions associated with utility usage. Moreover, the existence of water has more significant impacts on the process’s GWP than on its production cost. The integrated TEA, LCA, and Bayesian optimization provide candidate HDO conditions and process parameters to explore. However, detailed analysis and validation should be carried out after future scale-up experiments and process improvement. Furthermore, as feedstock and solvent usage are the main contributors to the cost and emissions, more efficient upstream HMF production and effective filtration technologies have great potential to further improve the HMF HDO process.



## Acknowledgement

This work was supported by the Department of Energy's Office of Energy Efficiency and Renewable Energy's Advanced Manufacturing Office under Award Number DE-EE0007888-7.6. The Delaware Energy Institute gratefully acknowledges the support and partnership of the State of Delaware toward the RAPID projects.

## References

- Abrams, D. S. and Prausnitz, J. M., 1975 Statistical thermodynamics of liquid mixtures: A new expression for the excess Gibbs energy of partly or completely miscible systems, *AIChE Journal*, 21, 1, 116-128
- Athaley, A., Saha, B. and Ierapetritou, M., 2019 Biomass-based chemical production using techno-economic and life cycle analysis, *AIChE Journal*, 65, 9, e16660
- Bare, J., 2011 TRACI 2.0: the tool for the reduction and assessment of chemical and other environmental impacts 2.0, *Clean Technologies and Environmental Policy*, 13, 5, 687-696
- Bozell, J. J. and Petersen, G. R., 2010 Technology development for the production of biobased products from biorefinery carbohydrates—the US Department of Energy's "Top 10" revisited, *Green Chemistry*, 12, 4, 539-554
- Chang, H., Bajaj, I., Huber, G. W., Maravelias, C. T. and Dumesic, J. A., 2020 Catalytic strategy for conversion of fructose to organic dyes, polymers, and liquid fuels, *Green Chemistry*, 22, 16, 5285-5295
- De Ath, G., Everson, R. M., Rahat, A. A. M. and Fieldsend, J. E., 2021 Greed is good: Exploration and exploitation trade-offs in Bayesian optimisation, *ACM Transactions on Evolutionary Learning and Optimization*, 1, 1, 1-22
- Diwale, S., Eisner, M. K., Carpenter, C., Sun, W., Rutledge, G. C. and Braatz, R. D., 2022 Bayesian optimization for material discovery processes with noise, *Molecular Systems Design & Engineering*, 7, 6, 622-636
- Esteban, J., Vorholt, A. J. and Leitner, W., 2020 An overview of the biphasic dehydration of sugars to 5-hydroxymethylfurfural and furfural: a rational selection of solvents using COSMO-RS and selection guides, *Green Chemistry*, 22, 7, 2097-2128
- Iwama, R. and Kaneko, H., 2021 Design of ethylene oxide production process based on adaptive design of experiments and Bayesian optimization, *Journal of Advanced Manufacturing and Processing*, 3, 3, e10085
- Jae, J., Zheng, W., Karim, A. M., Guo, W., Lobo, R. F. and Vlachos, D. G., 2014 The Role of Ru and RuO<sub>2</sub> in the Catalytic Transfer Hydrogenation of 5-Hydroxymethylfurfural for the Production of 2,5-Dimethylfuran, *ChemCatChem*, 6, 3, 848-856
- Lu, Q., González, L. D., Kumar, R. and Zavala, V. M., 2021 Bayesian optimization with reference models: A case study in MPC for HVAC central plants, *Computers & Chemical Engineering*, 154, 107491
- Luo, Y. and Ierapetritou, M., 2020 Comparison between different hybrid life cycle assessment methodologies: a review and case study of biomass-based p-xylene production, *Industrial & Engineering Chemistry Research*, 59, 52, 22313-22329

- Luo, Y., Kuo, M. J., Ye, M., Lobo, R. and Ierapetritou, M., 2022 Comparison of 4,4' - Dimethylbiphenyl from Biomass-Derived Furfural and Oil-Based Resource: Technoeconomic Analysis and Life-Cycle Assessment, *Industrial & Engineering Chemistry Research*, 61, 25, 8963-8972
- Snoek, J., Larochelle, H. and Adams, R. P., 2012 Practical bayesian optimization of machine learning algorithms, *Advances in neural information processing systems*, 25,
- Wang, K. and Dowling, A. W., 2022 Bayesian optimization for chemical products and functional materials, *Current Opinion in Chemical Engineering*, 36, 100728
- Wang, Y., Chen, T.-Y. and Vlachos, D. G., 2021 NEX Torch: A Design and Bayesian Optimization Toolkit for Chemical Sciences and Engineering, *Journal of Chemical Information and Modeling*, 61, 11, 5312-5319
- Wang, Z., Bhattacharyya, S. and Vlachos, D. G., 2020 Solvent selection for biphasic extraction of 5-hydroxymethylfurfural via multiscale modeling and experiments, *Green Chemistry*, 22, 24, 8699-8712



# Investigation of biomass blending ratios for optimal biochar's soil application

Farah Obar, Mohammad Alherbawi, Gordon Mckay, Tareq Al-Ansari\*  
*College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation,  
Doha, Qatar*

\*[talansari@hbku.edu.qa](mailto:talansari@hbku.edu.qa)

## Abstract

Due to the rapid growth of population, urbanization and economic development, global municipal waste generation is expected to increase by nearly 70% to 3.4 billion tonnes by 2050. The mismanagement of these wastes results in the deterioration of soil, air, and water quality, causing serious health problems. As such, investigating waste valorization routes becomes more essential to alleviate the concerns associated with the burden of wastes. In this study, 10 different types of biomass used to produce biochar through pyrolysis were investigated; namely camel manure (CM), date pits (DP), sewage sludge (SS), coffee waste (CW), cattle manure (CM), poultry litter (PL), cabbage waste (CB), cucumber waste (CU), tomato waste (TM), and carrot waste (CR). Techno-economic and environmental analyses were performed from “cradle to grave”, covering the key biochar production stages from acquisition of biomass, transportation, plant's construction, pre-processing, pyrolysis, and lastly the end-use of biochar as soil enhancer. The cost and emissions savings using biochar as an alternative to commercial fertilizers were evaluated for the different biomass scenarios. Finally, an optimization model was developed to select the optimal biomass blending ratio for two scenarios to maximize the savings, which was then solved using MATLAB. The findings of this study indicated that the optimal biomass blend for maximizing cost, energy and emissions savings shall consist of the following proportions: 50% TM, 14-15% CW and PL, 5 – 10% CM and CU, 2.5% CT, 1.8-1.9% CB and DP, 1.2% CR and 0 – 0.6% SS. Future studies may involve more types of biomasses, additional optimization objectives, and may target multiple biochar's end-use applications.

**Keywords:** Pyrolysis; Biochar; Soil enhancement; Biomass.

## 1. Introduction

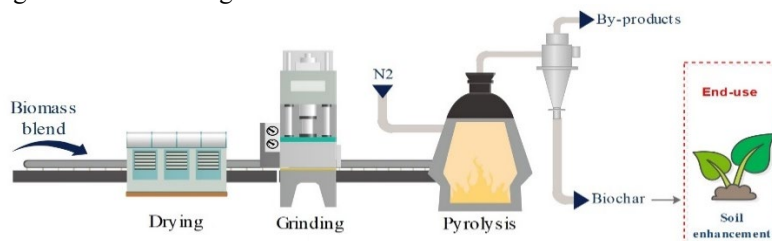
Global municipal solid waste generation is expected to increase from 6.94 million metric tonnes in 2021 to 3.4 billion metric tonnes in 2050 due to the rapid population growth and urbanization (NEA, 2016; Tiseo, 2021). This has accelerated the need to establish new waste-to-energy technologies including pyrolysis. Pyrolysis is a thermochemical process that works on converting biomass, which is a carbon based renewable organic material generated from animals and plants into value added products in the absence of oxygen to produce; biogas, bio-oil and biochar (BC) (Elkhalifa et al., 2019). BC is an incredibly rich source of carbon that can be employed in various disciplines to address the most significant environmental challenges as follows; (1) can be employed as a renewable energy source, (2) can be employed as a soil enhancer reducing the chemical fertilizers requirements, and (3) can be utilized as an innovative adsorbent for water and wastewater decontamination (Haider et al., 2022).

This study focused on biochar's end-use as a soil enhancer as it increases the crop yield, enhances the water retention of soil due to its porous structure, and sequesters carbon in

the soil by absorbing it from the atmosphere to mitigate climate change (Mukherjee & Lal, 2013). Moreover, this study investigates the optimal biomass blends based on 10 different types of biomass feedstocks including; camel manure (CM), date pits (DP), sewage sludge (SS), coffee waste (CW), cattle manure (CT), poultry litter (PL), cabbage waste (CB), cucumber waste (CU), tomato waste (TM), and carrot waste (CR). The aim of this investigation is to target three objectives; maximize cost, energy and carbon emissions savings.

## 2. Methodology

The methodology of this study follows two main steps, which are the data collection from literature, and formulating the equations that measure the savings that can be used for optimizing the biomass blends using the genetic algorithm tool in MATLAB. The process flow diagram is shown in figure 1.



**Figure 1:** Process flow diagram

### 2.1 Data Collection

The production of biochar from the different biomass feedstocks has been simulated using Aspen. This model has been used to evaluate the cost, carbon and energy footprint when producing biochar from each single biomass. The operating parameters of the process have included an operating temperature of 300°C, 5L/min.tonne nitrogen flowrate, 90% drying efficiency and 30 minutes' reaction time. The elemental analysis for the biomass feedstocks was collected from literature as presented in Table 1, whereas this model requires the analysis to be on a dry ash and sulfur-free basis. The price of date pits was quoted locally at 100\$/t. While the prices of the other biomass feedstocks were obtained from literature and was found to be 15\$/t sewage sludge, 88\$/t coffee waste, 25\$/t for all types of animal manure and 10\$/t for vegetable waste (Alherbawi et al., 2021; Kamil et al., 2019; Marufuzzaman et al., 2015; Yun et al., 2018).

**Table 1:** Elemental analysis of biomass feedstock.

Biomass	%C	%H	%O	%N	%S	Ash (%)	Moisture (%)	Ref.
CM	36.7	5.6	34.7	2.6	0.4	19.9	4.3	(Al-Ansari et al., 2020)
DP	46.5	6.5	44.7	0.9	0	1.4	0.3	
SS	46.9	6.8	23.2	7.4	0.5	2.6	82.2	(Poudel et al., 2015)
CW	79.3	5.0	12.5	2.1	1.1	11.1	4.2	(Uwaoma et al., 2021)
CT	41.1	5.9	49.9	2.7	0.4	15.4	10.1	(Yuan et al., 2017)
PL	42.9	5.6	33.4	5.5	0.7	12	18.1	(Tańczuk et al., 2019)
CB	44.5	4.9	35.0	2.3	0.5	11.9	8.6	(Zhang et al., 2022)
CU	56.7	6.5	32.9	3.9	0	1.9	8.1	
TM	49.7	7.4	39.3	3.7	0	4.4	10.0	(Elkhalifa et al., 2022)
CR	41.5	5.4	51.7	1.4	0	5.2	6.6	

A comprehensive comparison has been made between the application of NPK fertilizer and biochar as a soil amendment to evaluate the cost, energy and emissions for each. The model has only estimated the cost, energy and carbon footprint of BC production from each single biomass, whilst the rest of the data relative to BC and NPK application to soil and NPK fertilizer production was obtained from literature. It was found that 0.05 kg/m<sup>2</sup> of NPK fertilizer is required to grow tomato, which was selected as the reference crop for this comparison (Ziane et al., 2021). The production of NPK fertilizer costs 0.8\$/kg, requires energy of 12.45 MJ/kg and emits 0.20 kgCO<sub>2-eq</sub>/kg (Majumder et al., 2019; Samoraj et al., 2022; Wu et al., 2021). Energy and carbon emissions for BC application were found to be equal to 0.434 MJ/m<sup>2</sup> and 0.024 kgCO<sub>2-eq</sub>/m<sup>2</sup>, which are lower than the values of NPK fertilizer application which are 0.615 MJ/m<sup>2</sup> and 0.069 kgCO<sub>2-eq</sub>/m<sup>2</sup> (Brentrup & Palliere, 2008; Jiang et al., 2021).

### 2.2 Problem Formulation

The data collected was inserted in equations 1-3 to find the cost, energy and carbon emissions when utilizing NPK fertilizer and BC, where *X* refers to the percentage of NPK fertilizer being used, *k* refers to the conversion factor between NPK fertilizer and BC, *P* and *A* represent production and application.

$$Cost = (X \cdot NPK_{required} \cdot cost_{NPK-P}) + (k \cdot (1 - X) \cdot NPK_{required} \cdot cost_{BC-P}) \quad (1)$$

$$Energy = X \cdot (NPK_{required} \cdot energy_{NPK-P} + energy_{NPK-A}) + k \cdot (1 - X) \cdot (NPK_{required} \cdot energy_{BC-P} + energy_{BC-A}) \quad (2)$$

$$GWP = X \cdot (NPK_{required} \cdot GWP_{NPK-P} + GWP_{NPK-A}) + k \cdot (1 - X) \cdot (NPK_{required} \cdot GWP_{BC-P} + GWP_{BC-A}) \quad (3)$$

The values were later used to find the savings using equations 4-6:

$$Cost_{savings} = Cost_{NPK} - Cost_{BC} \quad (4)$$

$$Energy_{savings} = Energy_{NPK} - Energy_{BC} \quad (5)$$

$$GWP_{savings} = GWP_{NPK} - GWP_{BC} \quad (6)$$

Based on the savings' results from the previous section, a multi-objective mathematical optimization is performed using genetic algorithm tool within MATLAB. The objectives targeted by this optimization are; maximizing cost, energy and carbon emissions savings, where the mathematical formulation of the multi-objective optimization problem is represented by equations 7-9, and the constraint adopted is shown in equation 10.

$$\text{Maximize cost savings: } \sum_{i=1}^{10} B_i C_i \quad (7)$$

$$\text{Maximize energy savings: } \sum_{i=1}^{10} B_i E_i \quad (8)$$

$$\text{Maximize emission savings: } \sum_{i=1}^{10} B_i G_i \quad (9)$$

$$\text{Constraint: } \sum_{i=1}^{10} B_i = 1 \quad (10)$$

The multi-objective optimization was conducted based on two scenarios as shown in Table 2 to provide broader insight for decision-makers.

**Table 2:** Objectives to be obtained in each scenario.

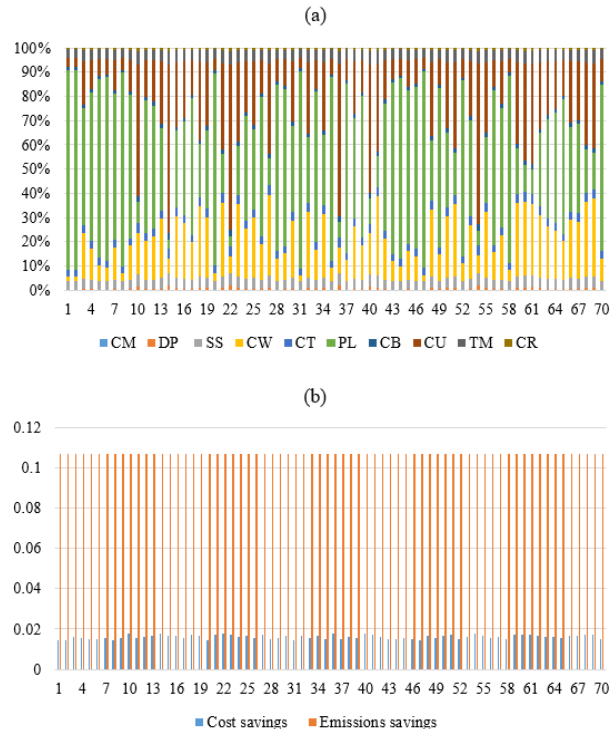
Scenario	Objectives obtained
1	Maximize cost and emissions savings

### 3. Results and Discussion

70 solutions were obtained fulfilling the objectives for each scenario by performing a trade-off between the desired objectives using the genetic algorithm tool in MATLAB. The results of the scenarios are as follows:

#### 3.1 Scenario 1

This scenario aimed to achieve the maximum cost and emissions savings. The findings as illustrated in figure 2 show that all solutions lean more towards achieving higher emissions savings and focused less on the cost savings. Moreover, the solutions show that the most optimal biomass blends consisted of very high proportions of poultry litter, and lower proportions of the other biomass feedstocks except the camel manure that had zero contribution in achieving the objectives. Some optimal blends consisted of around 80% PL, 1.7-4% of CU, TM, SS, CW and CT, 0.2-1% DP, CB and CR, and 0% CM.

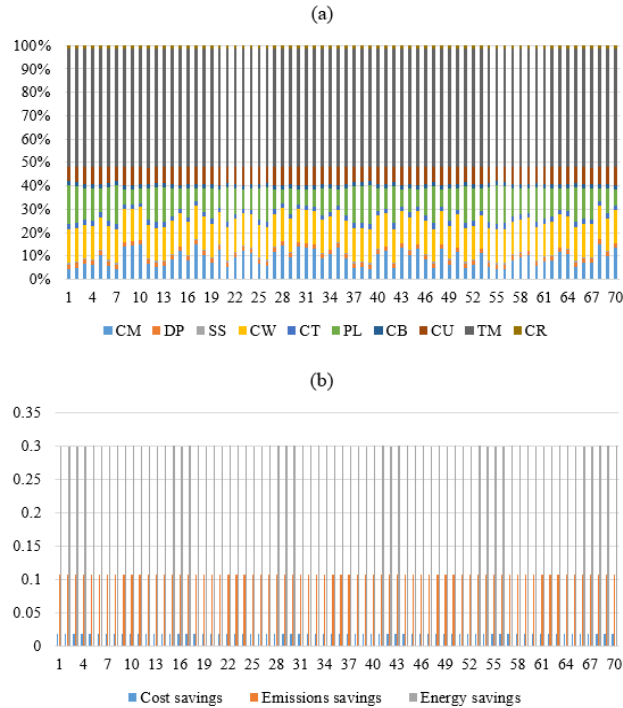


**Figure 2:** Scenario 1 – maximum cost and emission savings: (a) optimal biomass blends, and (b) optimal savings.

#### 3.2 Scenario 2

This scenario was concerned with achieving three objectives; maximize energy, cost and emissions savings and to find the optimal biomass blends that fulfill these objectives. The results as illustrated in figure 3 show that all solutions have contributed mostly to achieving maximum energy savings, then maximizing emissions and cost, respectively.

For optimizing the biomass blends, most of the results show that the optimal blend consisted of around 50% TM, 14-15% CW and PL, 5 – 10% CM and CU, 2.5% CT, 1.8-1.9% CB and DP, 1.2% CR and 0 – 0.6% SS.



**Figure 3:** Scenario 2 – maximum cost, energy and emission savings: (a) optimal biomass blends, and (b) optimal savings.

#### 4. Conclusion

Pyrolysis is a thermochemical process that converts biomass into value-added products, mainly biochar. This carbon-rich product has widely contributed into a diversity of applications including soil enhancement applications.

This study has provided insight into the sustainable solutions for researchers and decision-makers on selecting the most optimal biomass blends for biochar production from an economic and environment perspective. Two scenarios have been studied where each aimed to achieve specific objectives. Scenario 1 focused on maximizing cost and emissions only, where most solutions showed that the most optimal biomass blends consisted of very high proportions of poultry litter, lower proportions of the other biomass feedstocks, and no contribution of camel manure. On the other hand, the results of scenario 2 that aimed to achieve all three objectives, have shown that the most optimal biomass blends have consisted of very high proportions of tomato waste, less proportions of the others, and around 0% of sewage sludge. Future studies may involve more types of biomasses, additional optimization objectives, and may target multiple biochar's end-use applications.

#### Acknowledgement

The authors would like to thank Qatar National Research Fund (QNRF) for their support of this research through NPRP-11S-0107-180216, Hamad Bin Khalifa University (HBKU) and Qatar Foundation (QF).



## References

- Al-Ansari, T., AlNouss, A., Al-Thani, N., Parthasarathy, P., ElKhalifa, S., McKay, G., & Alherbawi, M. (2020). Optimising Multi Biomass Feedstock Utilisation Considering a Multi Technology Approach. In *Computer Aided Chemical Engineering* (Vol. 48).
- Alherbawi, M., Parthasarathy, P., Al-Ansari, T., Mackey, H. R., & McKay, G. (2021). Potential of drop-in biofuel production from camel manure by hydrothermal liquefaction and biocrude upgrading: A Qatar case study. *Energy*, *232*, 121027.
- Brentrup, F., & Palliere, C. (2008). GHG emissions and energy efficiency in European nitrogen fertiliser production and use. *Proceedings - International Fertiliser Society*, *2*(December).
- Elkhalifa, S., Al-Ansari, T., Mackey, H. R., & McKay, G. (2019). Food waste to biochars through pyrolysis: A review. *Resources, Conservation and Recycling*, *144*, 310–320.
- Elkhalifa, S., Mariyam, S., Mackey, H. R., Al-Ansari, T., McKay, G., & Parthasarathy, P. (2022). Pyrolysis Valorization of Vegetable Wastes: Thermal, Kinetic, Thermodynamics, and Pyrogas Analyses. *Energies*, *15*(17).
- Haider, F. U., Wang, X., Farooq, M., Hussain, S., Cheema, S. A., Ain, N. ul, Virk, A. L., Ejaz, M., Janyshova, U., & Liqun, C. (2022). Biochar application for the remediation of trace metals in contaminated soils: Implications for stress tolerance and crop production. *Ecotoxicology and Environmental Safety*, *230*, 113165.
- Jiang, Z., Zheng, H., & Xing, B. (2021). Environmental life cycle assessment of wheat production using chemical fertilizer, manure compost, and biochar-amended manure compost strategies. *Science of The Total Environment*, *760*, 143342.
- Kamil, M., Ramadan, K. M., Olabi, A. G., Shanableh, A., Ghenai, C., Al Naqbi, A. K., Awad, O. I., & Ma, X. (2019). Comprehensive evaluation of the life cycle of liquid and solid fuels derived from recycled coffee waste. *Resources, Conservation and Recycling*, *150*, 104446.
- Majumder, S., Neogi, S., Dutta, T., Powel, M. A., & Banik, P. (2019). The impact of biochar on soil carbon sequestration: Meta-analytical approach to evaluating environmental and economic advantages. In *Journal of Environmental Management* (Vol. 250).
- Marufuzzaman, M., Ekşioğlu, S. D., & Hernandez, R. (2015). Truck versus pipeline transportation cost analysis of wastewater sludge. *Transportation Research Part A: Policy and Practice*, *74*.
- Mukherjee, A., & Lal, R. (2013). Biochar impacts on soil physical properties and greenhouse gas emissions. In *Agronomy* (Vol. 3, Issue 2).
- NEA. (2016). Waste Statistics and Overall Recycling. *National Environmental Agency*.
- Poudel, J., Ohm, T. I., Lee, S. H., & Oh, S. C. (2015). A study on torrefaction of sewage sludge to enhance solid fuel qualities. *Waste Management*, *40*, 112–118.
- Samoraj, M., Mironiuk, M., Witek-Krowiak, A., Izydorczyk, G., Skrzypczak, D., Mikula, K., Baśladyńska, S., Moustakas, K., & Chojnacka, K. (2022). Biochar in environmental friendly fertilizers - Prospects of development products and technologies. *Chemosphere*, *296*.
- Tańczuk, M., Junga, R., Kolasa-Więcek, A., & Niemiec, P. (2019). Assessment of the energy potential of chicken manure in Poland. *Energies*, *12*(7).
- Tiseo, I. (2021). *Global Waste Generation - Statistics & Facts*. Environment, Development and Sustainability.
- Uwaoma, R. C., Strydom, C. A., Bunt, J. R., Matjie, R. H., Okolo, G. N., & Marx, S. (2021). Co-gasification reactivity and kinetics of municipality and coffee waste residue hydrochar and South African density separated coal blends. *Bioresource Technology Reports*, *16*, 100877.
- Wu, H., MacDonald, G. K., Galloway, J. N., Zhang, L., Gao, L., Yang, L., Yang, J., Li, X., Li, H., & Yang, T. (2021). The influence of crop and chemical fertilizer combinations on greenhouse gas emissions: A partial life-cycle assessment of fertilizer production and use in China. *Resources, Conservation and Recycling*, *168*, 105303.
- Yuan, X., He, T., Cao, H., & Yuan, Q. (2017). Cattle manure pyrolysis process: Kinetic and thermodynamic analysis with isoconversional methods. *Renewable Energy*, *107*, 489–496.
- Yun, Y. M., Lee, M. K., Im, S. W., Marone, A., Trably, E., Shin, S. R., Kim, M. G., Cho, S. K., & Kim, D. H. (2018). Biohydrogen production from food waste: Current status, limitations, and future perspectives. *Bioresource Technology*, *248*, 79–87.
- Zhang, Y., Cheng, X., Wang, Z., Tahir, M. H., Wang, Z., Wang, X., & Wang, C. (2022). Full recycling of high-value resources from cabbage waste by multi-stage utilization. *Science of The Total Environment*, *804*, 149951.
- Ziane, H., Hamza, N., & Meddad-Hamza, A. (2021). Arbuscular mycorrhizal fungi and fertilization rates optimize tomato (*Solanum lycopersicum* L.) growth and yield in a Mediterranean agroecosystem. *Journal of the Saudi Society of Agricultural Sciences*, *20*(7).

# Methodology for estimating a country's production of biomethane from agricultural residues: a multiscale and holistic approach

Manuel Taifouris, Mariano Martin

*<sup>a</sup>Salamanca, Department of Chemical Engineering, University of Salamanca, Plza. Caidos 1-5, 37008 Salamanca, Spain*

## Abstract

Dependence on fossil fuels and recent issues on the supply chain from major natural gas providers have highlighted the importance of energy security in the countries of the European Union. This work presents a multi-scale analysis to evaluate the implementation of the circular economy to reduce the region's dependence on fossil natural gas. For this purpose, a techno-economic analysis of technologies (gasification and anaerobic digestion), size, and location of the waste treatment facilities is performed, followed by a facility location problem. This analysis is applied to a case study in Spain, concluding that an investment of 9458M€ and an operating cost of 5000M€ per year are necessary to cover 35% of the natural gas demanded. Up to 19 provinces are totally independent with this investment, favoring the creation of decentralized networks of natural gas, which is a strong incentive to invest in waste management because of the social and economic importance of natural gas as an energy source.

**Keywords:** Agricultural residues, Sustainable process design, Energy, Circular economy

## 1. Introduction

Despite the implementation of decarbonization policies in European countries, the current dependence on fossil fuels together with the centralization of their production in a reduced number of countries put the energy security of most European countries at risk (Mišik, 2022). In addition, the increase in the world's population has led to the intensification of crop, meat, and milk production processes, creating areas of high waste generation. These areas can suffer from nutrient pollution, which can lead to the eutrophication of water bodies and soil deterioration. In addition, municipal solid waste (MSW) and sewage sludge also represent an issue when cities are densely populated and waste is treated inefficiently so in most cases, ending up in landfills or incinerated (European Commission, 2021). The implantation of a circular economy for waste can solve both problems, by using waste management technologies, such as anaerobic digestion and gasification. Studies that analyze the biomethane production potential of a country are especially useful for determining where biomethane production plants should be placed, as well as determining which regions can be energy-independent (Wang et al., 2018). However, these studies directly relate the biomethane production potential to the number of animals, crops, or people, through the empirical yields. This approach decouples the estimation of biomethane (composition and amount) from the processing technology, as well as from the type and composition of waste. On the one hand, the waste composition can be highly variable, even for the same species of animal (Council, 2000). On the other hand, the composition of the waste affects both the design of the treatment equipment and the amount and composition of the biomethane produced, and therefore, the economic and environmental evaluation of its valorization. A multi-scale approach allows for addressing the different scales of the waste management system (Floudas et al., 2016),

from waste characterization and process design to treatment capacity and facility location. By analyzing all scales simultaneously, a better estimate of the biomethane production potential of a location can be determined, as well as the economic and environmental cost of the development and implementation of the process. This allows optimizing the selection of the location, the treatment process, and the type of waste. It also facilitates energy integration between the different stages of the process. This work presents a multiscale study, which uses an optimization framework, to design waste treatment plants (size and type of waste treated), select the best location, and estimate the investment and operating costs. The total amount of biomethane and its cost are also calculated. This framework is applied to the case study of Spain, where the management budget is the main optimization variable. The optimization framework aims at maximizing biomethane production and minimizing production costs.

## **2. Framework development**

This section presents the optimization framework, which addresses waste production characterization (amount and composition), the modeling of processing technologies, the scale-up/down of the facilities, and the selection of the location.

### 2.1-Estimation of the production and composition of waste

Two types of waste are considered in this work, dry waste (lignocellulosic residue) and wet waste (manure, MSW, and sludge). Lignocellulosic residues of a region are estimated from the annual production of crops and the amount of residue grown by the type of crop. Manure is calculated from the amount of waste generated based on the age and the number of animals. The production of MSW and sludge are obtained through the annual production per capita (388 kg of MSW and 105 kg of sludge) and the population. The composition of the residues is obtained from the literature. Nevertheless, it is possible to perform studies for specific residues to increase the accuracy of the estimates.

### 2.2-Process analysis and design

A techno-economic analysis is proposed to analyze and compare gasification and anaerobic digestion as possible technologies to produce biomethane from agro-industrial wastes. In addition, a scale-up study is performed for each waste and for 50 different treatment capacities. Non-linear programming (NLP) is used to optimize each design and reduce the cost of methane production.

An indirect gasification process, which consists of a combination of a gasifier and a furnace, is used for the treatment of dry waste. Both gasification and biomethane purification are modeled following first principles, such as mass and energy balances, empirical yields and correlations, and thermodynamic equilibria. The heat necessary for the gasification process is supplied by the combustion of the char formed in that same gasification when the steady-state regime is reached. Olivine is used as heat transfer media (HTM) to transmit the thermal energy generated in the combustion process to the gasifier. Therefore, the system is energetically self-sustainable. The total combustion of all compounds in the furnace is considered to perform the mass and energy balances. The composition of the syngas is estimated from the gasification temperature (Phillips et al., 2007). Regarding the syngas purification stage, empirical yields are used to estimate the separation efficiency of char, olivine, and ash in the cyclones (99%) and electrostatic precipitators (99%), as well as in the adsorbed amount of H<sub>2</sub>S in the ZnO beds (100%). The hydrocarbon reforming process, the water gas shift reactor, and the methanation

reactor are modeled by using equilibrium constants (Sánchez et al., 2019). Finally, the PSA system is used to reduce the amount of CO<sub>2</sub> down to 2% and remove NH<sub>3</sub> and H<sub>2</sub>O. This system is modeled by empirical yields (León & Martín, 2016). An anaerobic digestion system is designed for the treatment of wet waste. The composition and amount of biogas are estimated from the carbohydrate, lipid, and protein composition of the waste, stoichiometric ratios, and biodegradability (Taifouris & Martín, 2018). The distribution of the gases between the liquid and gas phases is modeled from thermodynamic equilibria. The biogas purification phase consists of an iron bed to remove H<sub>2</sub>S (100%) and a PSA system to reduce the content of CO<sub>2</sub> down to 2% and remove NH<sub>3</sub> and H<sub>2</sub>O. Both systems are modeled from adsorption yields.

The gasification and anaerobic digestion systems are scaled up/down from a minimum to a maximum size of treatment capacity. The minimum corresponds to the minimum amount of waste generated per region, considering all available regions simultaneously, while the maximum can be established by the literature (anaerobic digestion) or by the maximum availability of waste per region (gasification). Up to 50 designs with different capacities for each waste and for each technology are considered. For each design, a techno-economic evaluation is carried out to determine the operating expenses (OPEX) and the capital expenses (CAPEX) following the procedures described in the literature (Sinnott, 2005) and the results of the modeling of the processes.

### 2.3-Facility location problem

A facility location problem is formulated including surrogate models for the facilities (yields, CAPEX, OPEX as a function of their size). A mixed integer linear programming (MILP) problem is solved to select the best location and size of the waste treatment facilities. The objective function aims at maximizing methane production based on the allocated budget. With the information on the amount of natural gas demand, it is possible to determine the fraction of self-sufficiency of each spatial unit.

## **3.Result**

### **3.1-Case of study**

The optimization framework is applied to the case study in Spain. The country is divided into agricultural districts (345 possible locations). Spain is selected by its high agricultural production (Gobierno de España, 2022), high dependency on foreign suppliers of natural gas (Enagas, 2022), and low production of biofuels. The optimization framework consists of two types of models, NLP and MILP. Both are solved in a computer with an Intel Core i7-7700, 32 Gb of Ram, using GAMS.

### **3.2-Properties of the different types of factories**

The designs presented in section 2 are modeled and evaluated economically. The main results are shown in Table 1.

Table 1.- Main results of the techno-economic analysis of the residues (q: capacity of factory(t/ye), DW: dried waste)

Residue	Yield (kgBiomethane/kgDW)	Production Cost (€/kgDW)	Min.Capacity (t/ye)	Max.Capacity(t/ye)
SLUDGE	0.003	$P_{cost}=1.814 \cdot 10^9 \cdot q^{-0.97}$	700624	49754
MANURE	0.012	$P_{cost}=1.603 \cdot 10^8 \cdot q^{-0.885}$	63072	367920
MSW	0.070	$P_{cost}=7.203 \cdot 10^7 \cdot q^{-0.97}$	19657	52560
LIGNO	0.285	$P_{cost}=6.060 \cdot 10^5 \cdot q^{-0.626}$	10000	820000

As it can be seen in Table 1, the relationship between the cost of biomethane production and the treatment capacity of the plant follows a power law, due to the strong economies of scale present in both OPEX and CAPEX. Besides, gasification has the best yield to produce biomethane from lignocellulosic waste while anaerobic digestion of sludge is the worst option.

### **3.3-Total potential of biomethane production in Spain**

The application of gasification and anaerobic digestion to the total amount of waste available per year in Spain shows that the maximum percentage of natural gas that can be covered is 43%. To achieve this, a CAPEX of 21391M€ and an OPEX of 25852M€/ye are required. This allows up to 21 provinces of the 52 available in Spain, to be self-sufficient. It is observed that the areas with the highest biomethane production are those with the highest amount of lignocellulosic residues. This is because gasification has a higher yield than the anaerobic digestion of any of the wet waste. In addition, a significant mismatch is observed between the areas that consume natural gas (industrial and densely populated areas) and the areas that can produce biomethane (rural areas).

### **3.4-Determination of the optimal budget for the reduction of Spain's dependence on fossil natural gas**

The facility location problem is used to establish the optimal selection of location, size, and type of facilities for different budgets, drawing the Pareto (see Figure 1). This curve shows that there are two sections divided by the point of 5000 M€ per year. In the first segment, each 100M€ increase in the budget represents a 2% increment in the self-supply rate. However, in the second section, the self-supply rate remains constant even if the budget is increased. Therefore, the point of 5000M€/ye is selected as the best budget to invest in waste treatment in Spain. This OPEX corresponds to a CAPEX of 9505M€. Besides, the biomethane cost is similar to the current market price of natural gas (10.19 vs 6.24 USD/MMBTU). With this budget, it is possible to cover up to 35% of the country's total natural gas demand. Therefore, through an optimal selection of the budget dedicated to waste treatment, it is possible to reduce the CAPEX and OPEX of the plants by 55.56% and 80.65%, respectively, with respect to the previous section, while the self-sufficiency ratio only decreases by 18.60%. Besides, 19 provinces are totally independent of natural gas from foreign suppliers. As discussed in Section 3.3, these provinces correspond mainly to rural areas where natural gas consumption is low and potential biomethane production is high. However, the areas with the greatest potential for biomethane production are those with a large concentration of lignocellulosic residues (see Figure 2). The results of the techno-economic analysis of the treatment processes show that economies of scale are strongly favored, prioritizing the selection of large treatment plants over small ones.

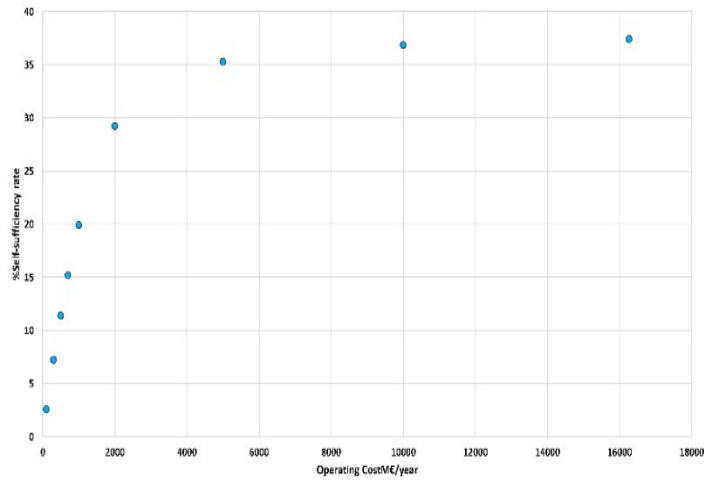


Figure 1.-Relation between OPEX and the self-sufficiency rate

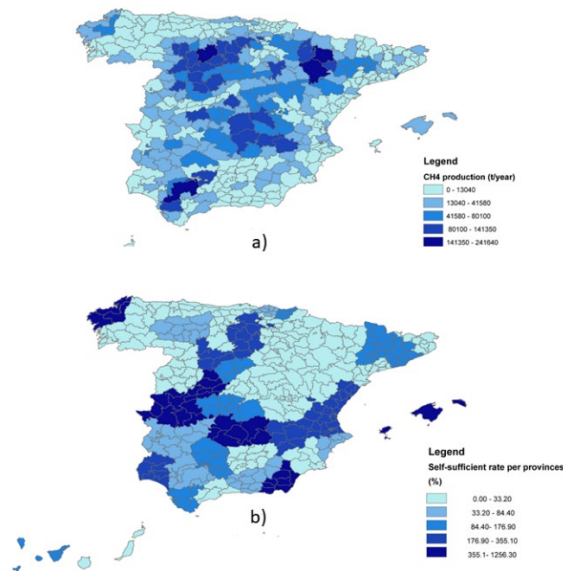


Figure 2.- (a) Production of Biomethane (ton/ye) with an OPEX of 5000 MM/ye and (b) self-sufficient rate of each province

#### 4. Conclusion

The dependence on fossil natural gas and the growing problem of waste generation have pushed European countries to analyze the circular economy of waste as a possible solution to both problems. This work presents a multiscale and holistic analysis to evaluate the biomethane production potential from different types of waste. An optimization framework is formulated to determine the best process, size, and location for the facilities, as well as the system cost. This framework is applied to the case study in Spain, where with an OPEX of 5000M€, it is possible to cover 35% of natural gas consumption and up to 19 provinces can be independent of fossil natural gas. This opens the way to create

decentralized structures to reduce the environmental and economic impacts of transporting natural gas from main pipelines to these rural regions. Therefore, this potential energy independence is a strong incentive to invest in infrastructure for waste treatment in these regions. Most of the biomethane comes from the treatment of dry waste through gasification since it has a performance of 4 times greater than anaerobic digestion. Therefore, the treatment of lignocellulosic waste is prioritized over the others. The optimization framework developed in this work can be applied to other countries, changing the crop, population, and animal databases. Some parameters, such as the composition of the residues, can be modified to improve the estimates, without modifying the models of the optimization framework.

### Acknowledgement

The authors would like to acknowledge Salamanca Research for the optimization licenses, the funding received from the European Union's Horizon 2020 research, innovation program under the Marie Skłodowska-Curie grant agreement no 778168. M.T. ,and appreciates the FPI Ph.D. fellowship from the Junta de Castilla y León.

### References

- Council, N. R. (2000). Nutrient Requirements of Beef Cattle. In Nutrient Requirements of Beef Cattle. National Academies Pres.
- Enagas. (2022). Boletín Estadístico .  
<https://www.enagas.es/content/dam/enagas/es/ficheros/gestion-tecnica-sistema/energy-data/publicaciones/boletin-estadistico-del-gas/Boletin-Estadistico-Gas-agosto-2022.pdf>
- European Comision. (2021). Municipal waste statistics - Statistics Explained. Eurostat Statistics Explained. [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Municipal\\_waste\\_statistics](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Municipal_waste_statistics)
- Floudas, C. A., Niziolek, A. M., Onel, Ö., & Matthews, L. R. (2016). Multi-scale systems engineering for energy and the environment: Challenges and opportunities. *AICHE Journal*, 62(3), 602–623. <https://doi.org/10.1002/AIC.15151>
- Gobierno de España. (2022). PERTE Agroalimentario. <https://www.lamoncloa.gob.es/consejodeministros/resumenes/Documents/2022/080222-ResumenEjecutivoPERTEAgro.pdf>
- León, E., & Martín, M. (2016). Optimal production of power in a combined cycle from manure based biogas. *Energy Conversion and Management*, 114, 89–99. <https://doi.org/10.1016/j.enconman.2016.02.002>
- Mišík, M. (2022). The EU needs to improve its external energy security. *Energy Policy*, 165, 112930. <https://doi.org/10.1016/J.ENPOL.2022.112930>
- Phillips, S., Aden, A., Jechura, J., Dayton, D., & Eggeman, T. (2007). Thermochemical ethanol via indirect gasification and mixed alcohol synthesis of lignocellulosic biomass. <https://doi.org/10.2172/1216397>
- Sánchez, A., Martín, M., & Vega, P. (2019). Biomass Based Sustainable Ammonia Production: Digestion vs Gasification. *ACS Sustainable Chemistry and Engineering*, 7(11), 9995–10007. <https://doi.org/10.1021/acssuschemeng.9b01158>
- Sinnott, R. K. (2005). *Chemical Engineering Design* (6th Edition, Vol. 6, pp. 339–359). Elsevier.
- Taifouris, M. R., & Martín, M. (2018). Multiscale scheme for the optimal use of residues for the production of biogas across Castile and Leon. *Journal of Cleaner Production*, 185, 239–251. <https://doi.org/10.1016/J.JCLEPRO.2018.03.018>
- Wang, S., Jena, U., & Das, K. C. (2018). Biomethane production potential of slaughterhouse waste in the United States. *Energy Conversion and Management*, 173, 143–157. <https://doi.org/10.1016/J.ENCONMAN.2018.07.059>

# Energy management for controlled environment agriculture based on physics informed neural networks and adaptive linearization based data-driven robust model predictive control with AI

Guoqing Hu<sup>a</sup>, Fengqi You<sup>a</sup>

<sup>a</sup>*Cornell University, Ithaca, New York, 14853, USA*

## Abstract

Reducing stress on energy demand is crucial for promoting controlled environment agriculture, the most energy-intensive sector. Refining the modeling approach and control strategies can help reduce energy demand from crop cultivation facilities. This study presents a novel physics-informed neural network that considers both greenhouse energy demand and crop models to minimize weather forecast errors in controlled environment facilities. The adaptive linearization feedback scheme is used to handle nonlinear relationships between environment, crops, and control elements such as heating, ventilation, CO<sub>2</sub> enrichment, and more. The proposed control inputs are determined by the data-driven robust model predictive control, which uses machine learning to address weather forecast uncertainties. The proposed control framework has been demonstrated to effectively reduce energy consumption while maintaining suitable growing conditions for crops, even in challenging conditions.

**Keywords:** controlled environment agriculture, model predictive control, neural network.

## 1. Introduction

Controlled environment agriculture provides a promising solution for the food demand arising from ascending global population (Costantino et al., 2021). Energy-intensive agricultural systems can contribute to the energy crisis and high levels of greenhouse gas emissions during the cultivation process (FAO, 2018). Therefore, effective control is needed to enhance energy usage efficiency and maintain a favorable cultivation environment for crop productivity and environmental sustainability. The conventional control strategies may not be able to adequately be used for energy management since they didn't comprehensively consider the nature of highly coupled nonlinear climate and crops dynamics and intensive environmental disturbances (Chen et al., 2022). The lack of consideration for these could result in excessive/limited energy applied and failure to maintain a favorable cultivation environment. Alternatively, Model predictive control (MPC) has been acknowledged as a suitable control strategy to solve the problem of regulating the climate and growth processes and guarantee the fulfillment of the constraints specification. There are two fundamental notions in tuning an MPC application: (1) high-fidelity model and (2) robustness of control in application (Garriga and Soroush, 2010). An accurate model could prevent the violation of the constraints caused by modeling error during the application and further control efforts to adjust the states.

We propose a novel control framework for regulating energy consumption and boosting crop yield. To construct the comprehensive high-fidelity model for control strategies, we leverage the physics-informed neural network (PINN) in estimating the nonlinear model,



which considers heating/cooling, humidification/dehumidification, ventilation, CO<sub>2</sub> enrichment, lighting, irrigation, and fertilization. In this work, instead of using a single state-space model (SSM) to generalize the dynamics of the indoor production facility, adaptive linearization feedback is adopted to calculate each SSM based on the data predicted from PINN at each control interval. Afterward, each control interval's robust control strategy is determined from the RMPC framework based on the SSM and historical forecast uncertainties. Historical weather forecast data and historical weather measurement data are gathered to form the forecast uncertainties. The data-driven uncertainty set is constructed by machine learning methods, including principal component analysis (PCA) (Wold et al., 1987) coupled with kernel density estimation (KDE) (Zhang et al., 2006). Based on these methods, we further develop a data-driven polyhedron uncertainty set, in which quantile functions are used to describe confidence intervals. Using this modeling and control approach, the nonlinear correlations between environmental disturbances and control effort can be linearized with acceptable accuracy, and control energy can be developed more effectively based on a more accurate model.

## 2. Physics informed neural network and control framework formulation

The physics-informed layer contains all the differential equations to numerically approximate the values of the temperature states described in the section above. However, the nonlinear dynamics of the greenhouse contain stiff ordinary differential equations, which are numerically unstable in the traditional Euler method integration process (Gear et al., 1985). Therefore, to relieve this issue during the integration process, the Radau method (Guglielmi and Hairer, 2001) which is under fully implicit Runge-Kutta class and is essential for the numerical integration over the stiff problems (Hairer and Wanner, 1999), is included in the integration process. Then the values calculated from the physics-informed layer will be fed into the linear weighted-bias layer for fitting to the experimental values. We recognize that the PINN can retain a higher accuracy when the sample data is limited (only 30 samples were used for training). Meanwhile, PINN with stiff ordinary differential equation integration methods can achieve better simulation results than with the Euler methods and save 194% of the simulation time needed. Disjunctive uncertainty sets are built for learning the trend of the uncertainty data. The uncertainty set  $\mathbb{W}$  can be formulated under the introduction of forward and backward deviation variable  $z^+$  and  $z^-$  (Ning and You, 2018; 2019):

$$\mathbb{W} = \left\{ \mathbf{w} \in R^H \left| \begin{array}{l} \mathbf{w} = \hat{\boldsymbol{\mu}} + Q\xi, \quad \xi = \underline{\xi}z^- + \bar{\xi}z^+ \\ \mathbf{0} \leq z^+, z^- \leq \mathbf{1}, \quad z^+ + z^- \leq \mathbf{1}, \quad \mathbf{1}^T(z^+ + z^-) \leq \Gamma \\ \underline{\xi} = [\hat{F}_1^{-1}(\alpha), \dots, \hat{F}_1^{-1}(\alpha)]^T \\ \bar{\xi} = [\hat{F}_1^{-1}(1-\alpha), \dots, \hat{F}_1^{-1}(1-\alpha)]^T \end{array} \right. \right\} \quad (2)$$

With the data-driven uncertainty set, the MPC problem with ADF can be formulated as :

$$\begin{aligned} & \min \sum_{i \in B_u} c_i u_i + \lambda^T L \lambda \\ & \text{s.t. } F_u[Mw + h] \leq f_u, \quad \forall w \in \mathbb{W}^k \\ & \quad F_x[Ax_0 + B_u h + B_v v + (B_w + B_u M)w] \leq f_x + \lambda, \quad \forall w \in \mathbb{W}^k \end{aligned} \quad (3)$$

where  $F_x$ ,  $F_u$ ,  $f_x$ ,  $f_u$  represent the state variable constraints matrix, control input constraints matrix, constraints for state variables, and constraints for the input.  $L$  is the weighted cost matrix that penalizes the violation to the constraints.  $\lambda$  is the slack variable that allows some extent of violation to the hard constraints. The data-driven disjunctive uncertainty

sets are adopted in this work and incorporated into the RMPC framework, because one of the major challenges for control lies in the inherent uncertainty of weather forecasting, due to the stochastic nature of atmospheric processes and the imperfection of the weather model. However, a box uncertainty set can potentially lead to over-conservative solutions to the RMPC problems. Hence, the introduction of data-driven uncertainty sets can alleviate the over-conservatism issue confronted by the most commonly used single uncertainty set-based RMPC frameworks.

### **3. Performance profile of control framework**

Six control strategies are applied to the semi-closed greenhouse year-round control. We can observe that CEMPC could have the violation to the constraints with the disturbance of forecast uncertainties. On the other hand, the RMPC and DRMPC without the PINN framework could still have large violations of the constraints due to the significant model errors in the state prediction, as those model errors could have a non-negligible impact on the system control even though the control scheme left some space for the violation. This result implies the importance of having an accurate prediction model. Alternatively, with the help of PINN in accurately predicting the states, all robust MPC frameworks with the PINN leave some space to prevent constraint violation when weather forecast errors are significant. Thanks to these margins, most constraints are not violated even under extreme scenarios or model errors strike. The results are demonstrated in Table 1. The control frameworks with PINN are observed to have comprehensive advantages over the control schemes which are simply based on the single linear model. Therefore, in comparison to the forecast uncertainties, the model errors are also influential during the control decision-making process. The over-conservatism issue of RMPC could make control frameworks leverage more energy instead of relying on the ambient environment because uncertain forecast errors and purposes of avoiding the violation penalty (Zhao and You, 2022). Therefore, in order to fulfill the state constraints, the control actuators are more often open to maintain the humidity, leading to excessive energy usage. Hence, even though RMPC can achieve a higher yield in the result, the higher energy and resources expenses prevent it from being the most profitable control framework. From this perspective, our control framework can be safely acknowledged as the optimal control strategy compared to other control frameworks by having the highest yield-energy ratio. Despite that the simulation is conducted within Ithaca, in which the climate is relatively hospitable for crop growth, harsh conditions such as Tucson should also be tested. The recent increment in magnitude and frequency of extreme climate in Tucson requires prompt and efficient control strategies to mitigate the effect of climate change (Ghasemi Tousi et al., 2021). To ensure the feasibility and wide applicability of our proposed control framework, the control of the crop cultivation facility is simulated with the summer weather data of Tucson. This location is chosen because of its harsh conditions for crop cultivation. The temperature in Tucson in 2021 was measured as high as 42 °C, and the humidity was less than 50% in summer. As these growing conditions pose tremendous challenges to controlled environment agriculture to minimize the total control cost in maintaining the ideal growing conditions (Alkhalidi et al., 2019), especially for growing tomatoes and leafy green. Therefore, the simulation of growing tomatoes in Tucson was also conducted and served as a case study to demonstrate the performance of our proposed control framework. Our proposed control framework can handle the model error effectively. Similar to the case study above, the control framework with the PINN can noticeably lower the violation rate of 58.07% for temperature control and 2.82% for humidity control. Another observation is that with more severe conditions, the model errors and

forecast uncertainties become more influential in the control decision-making process. Therefore, in order to hedge against the sources of the errors, a more accurate predictive model and more robust control methods are needed. By comparing the results to those simulated with Ithaca weather conditions, we can observe that because of the high ambient temperature and sparse precipitation in Tucson, the climate within the cultivation facility tends to be hotter and drier (Hu and You, 2022). Despite of the harsh cultivation environment in Tucson (Chen et al., 2021), our controller still succeeds in maintaining the hospitable growing environment for the crops. Overall, under both scenarios, our control framework can achieve the highest yield-to-cost ratio, i.e., can achieve the same amount of the yield with the lowest cost, comparing to other control frameworks.

Table 1. The statistical summary of the control results of six controllers in the year 2020 in the indoor cultivation facility in Ithaca, New York. DRMPC is data-driven robust MPC with single linear models based on the training dataset; CEMPC is certainty equivalent MPC with single linear models based on a training dataset; RMPC is robust MPC with single linear models based on the training dataset.; DRMPC\_PINN is data-driven robust MPC with the multiple linear models calculated from PINN; CEMPC\_PINN is certainty equivalent MPC with the multiple linear models calculated from PINN; RMPC\_PINN is robust MPC with the multiple linear models calculated from PINN.

	DRMPC	CEMPC	RMPC	DRMPC PINN	CEMPC PINN	RMPC PINN
Temperature	47.67%	80.21%	45.78%	5.78%	65.21%	1.56%
Relative Humidity	34.56%	70.25%	30.89%	4.21%	60.45%	1.28%
Energy-yield ratio	\$1.14/kg	\$3.23/kg	\$1.29/kg	\$0.83/kg	\$1.82/kg	\$0.99/kg
CPU Time	1.77 s	0.07 s	0.99 s	1.88 s	0.12 s	1.09 s

Table 2. The statistical summary of the control results of six controllers in the year 2020 in the indoor cultivation facility in Tucson, Arizona.

	DRMPC	CEMPC	RMPC	DRMPC PINN	CEMPC PINN	RMPC PINN
Temperature	30.67%	85.41%	49.18%	6.17%	80.21%	5.78%
Relative Humidity	3.56%	80.15%	5.14%	4.56%	70.25%	3.89%
Energy-yield ratio	\$1.28/kg	\$9.09/kg	\$1.78/kg	\$0.93/kg	\$3.57/kg	\$1.09/kg
CPU Time	1.51 s	0.08 s	0.97 s	2.01 s	0.11 s	1.01 s

#### 4. Discussion

In this study, we proposed a novel control framework for minimizing the energy input and sustaining the yield under both forecast uncertainties and model errors. This model takes into not only the greenhouse energy demand model, which includes cover temperature, tray temperature, vegetation temperature, mat temperature, floor temperature, soil temperature, water vapor density, and CO<sub>2</sub> concentration, but also the crop model, which contains both mass of carbohydrate and relative growth rate in leaves, fruit, and stem. This model was incorporated into the physics-informed neural network and integrated with the Radau method to ensure stability throughout the simulation process. The validation results demonstrate that our proposed physics-informed neural network can have high-fidelity prediction over the simulated data which is disturbed under random model error. After the model construction, adaptive linearization feedback is adopted to relieve the computational effort from the nonlinearity within the greenhouse dynamics. Within each linear state-space model, we adopted the data-driven robust MPC to calculate the optimal control solution in each control interval. The case studies are conducted to test the controller's performance under the weather conditions in Ithaca, New York, USA, and Tucson, Arizona, USA. On average, our modeling approach can also help contribute to 85.76% of benefits boost under the same energy cost by balancing

the accuracy from the nonlinear models and computational efficiency from the linear models. On the other hand, both results demonstrate that our proposed data-driven robust MPC framework uses minimal control efforts to achieve the highest yield-to-cost ratio.

## References

- A. Alkhalidi, M. K. Khawaja, D. Abusubaih, 2019. Energy efficient cooling and heating of aquaponics facilities based on regional climate. *International Journal of Low-Carbon Technologies*, 15, 287-298.
- W.-H. Chen, N.S. Mattson, F. You, 2022, Intelligent control and energy optimization in controlled environment agriculture via nonlinear model predictive control of semi-closed greenhouse. *Applied Energy*, 320, 119334.
- W.-H. Chen, F. You, 2021, Smart greenhouse control under harsh climate conditions based on data-driven robust model predictive control with principal component analysis and kernel density estimation. *Journal of Process Control*, 107, 103-113.
- W.-H. Chen, F. You, 2022, Semiclosed Greenhouse Climate Control Under Uncertainty via Machine Learning and Data-Driven Robust Model Predictive Control. *IEEE Transactions on Control Systems Technology*, 30, 1186-1197.
- A. Costantino, L. Comba, G. Sicardi, et al., 2021. Energy performance and climate control in mechanically ventilated greenhouses: A dynamic modelling-based assessment and investigation. *Applied Energy*, 288, 116583.
- FAO, 2018. *The future of food and agriculture: alternative pathways to 2050*.
- J. L. Garriga, M. Soroush, 2010. Model Predictive Control Tuning Methods: A Review. *Industrial & Engineering Chemistry Research*, 49, 3505-3515.
- C. W. Gear, B. Leimkuhler, G. K. Gupta, 1985. Automatic integration of Euler-Lagrange equations with constraints. *Journal of Computational and Applied Mathematics*, 12, 77-90.
- T. E. Ghasemi, W. O'Brien, S. A. Doulabian, T. Shadmehri, 2021. Climate changes impact on stormwater infrastructure design in Tucson Arizona. *Sustainable Cities and Society*, 72, 103014.
- N. Guglielmi, E. Hairer, 2001. Implementing Radau IIA Methods for Stiff Delay Differential Equations. *Computing*, 67, 1-12.
- E. Hairer, G. Wanner, 1999. Stiff differential equations solved by Radau methods. *Journal of Computational and Applied Mathematics*, 111, 93-111.
- G. Hu, F. You, 2022, Renewable energy-powered semi-closed greenhouse for sustainable crop production using model predictive control and machine learning for energy management. *Renewable and Sustainable Energy Reviews*, 168, 112790.
- C. Ning, F. You, 2018. Data-driven decision making under uncertainty integrating robust optimization with principal component analysis and kernel smoothing methods. *Computers & Chemical Engineering*, 112, 190-210.

- C. Ning, F. You, 2019, Optimization under uncertainty in the era of big data and deep learning: When machine learning meets mathematical programming. *Computers & Chemical Engineering*, 125, 434-448.
- L. Sun, F. You, 2021, Machine Learning and Data-Driven Techniques for the Control of Smart Power Generation Systems: An Uncertainty Handling Perspective. *Engineering*, 7, 1239-1247.
- S. Wold, K. Esbensen, P. Geladi, 1987. Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 2, 37-52.
- X. Zhang, M. L. King, R. J. Hyndman, 2006. A Bayesian approach to bandwidth selection for multivariate kernel density estimation. *Computational Statistics & Data Analysis*, 50, 3009-3031.
- N. Zhao, F. You, 2022. Sustainable power systems operations under renewable energy induced disjunctive uncertainties via machine learning-based robust optimization. *Renewable and Sustainable Energy Reviews*, 161, 112428.

# Life Cycle Assessment on Chemical Recycling-aided Cascaded Use of Polypropylene Plastics

Xiang Zhao,<sup>a</sup> Fengqi You<sup>a,b</sup>

<sup>a</sup>*System Engineering, Cornell University, Ithaca, New York 14853, USA*

## Abstract

Plastic pollution posed by material losses and their subsequent chemical emissions is pervasive in nature and varies with age. Cascaded polypropylene plastic use enables the re-manufacturing of virgin polymers from waste and minimizes its generation and environmental exposure. Here we showed the advantages of this cascaded use over other waste end-of-life management pathways by investigating the environmental consequences of plastic losses across the entire life cycle. Plastic losses could form volatile organic chemicals via photo-degradation and pose non-negligible global warming, ecotoxicity, and air pollution effects that worsen by at least 189% in the long run. These environmental burdens could be increased by above 9.96% under high ultraviolet radiation levels and participation rates, which facilitated plastic particulate compartment transport and degradation. Plastic cascaded use aided by fast pyrolysis technologies could effectively cut environmental losses and outperform landfills and incineration in reducing 23.35% ozone formation and 19.91% air pollution by replacing the external monomer and fuel production with onsite manufacturing.

**Keywords:** plastic losses, chemical recycling, cascaded use, life cycle assessment.

## 1. Introduction

Plastics commonly used by humans always come with an environmental cost (Thomas, 2022). Limiting the reuse of plastics can release tons of waste into natural ecosystems (Chin et al., 2022), if not properly addressed through effective waste recovery technologies (Wen et al., 2021), including incineration and recycling, which aim to recover virgin material from waste (Zhao and You, 2021a). These technologies are still not dominantly used in end-of-life (EoL) waste treatment (Fan et al., 2022), and most waste will end up in landfills, posing an environmental burden due to the loss of plastic materials (Min et al., 2020). These chronic plastic losses in the form of micro- and nanoparticles are can easily ingested by organisms and produce volatile organic chemicals through photodegradation (Wu et al., 2022). Explicit assessment remains unmet research needs due to the lack of knowledge about the compartment transport of plastic particles. The cascade use of plastics enables material and waste used in tandem, including re-manufacturing virgin materials from monomers aided by chemical recycling (Zhao et al., 2022), and effectively minimizing material loss release and the derived environmental impacts. However, given the existing knowledge gaps regarding the environmental burden of plastics loss throughout their life cycle, these advantages over other waste management pathways have not been evaluated. This life cycle assessment (LCA) work investigated the advantages and disadvantages of cascade plastic use assisted by chemical recycling processes, which enable material re-production from monomers yielded from effectively cracking plastic waste. Polypropylene (PP) plastic was studied because it embodies similar chemical properties to other plastics (Bora et al., 2020) and is a major source of plastic contamination (Rosenboom et al., 2022). The environmental

burden of plastic losses was assessed based on a novel post-disposal fate model (Wang et al., 2019), and a time-dynamic assessment was considered, reflecting the short- and long-term environmental burden of plastic pollution. Then, a holistic life-cycle assessment approach was used to evaluate the environmental performance of polypropylene cascade use, involving the EoL fate at the life-cycle stage from raw material and resource extraction to plastic loss. Full-spectrum environmental impact results of cascade plastic use over 20, 100, and 500 years were evaluated to reflect pros and cons, point out the future technological innovations, and inform policy implications.

## 2. LCA Methodology

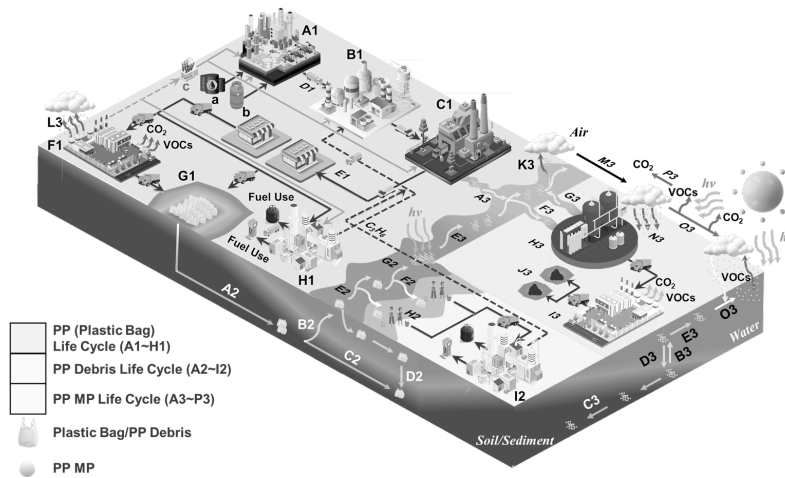


Figure 1. This LCA study's system boundary of PP plastics cascading system.

This LCA study aims to evaluate the life cycle environmental impacts of PP plastics, which are widely used in packaging materials and are also well-known major (micro)plastic (MP) pollution sources (Zhao and You, 2022). The upstream life cycle stages corresponding to the PP manufacturing, use, and EoL waste management, including incineration, landfill, and polypropylene chemical recycling, are included in the system boundary shown in Figure 1, while the downstream processes are related to plastic losses formation, degradation, aquatic waste collection and reclamation by fast pyrolysis-based chemical recycling, MP removals via drinking water treatment (DWTP) (Tian et al., 2020), and incineration. The functional unit was chosen as one-ton waste polypropylene plastic bag treated to align with the mass and energy balance across the entire life cycle (Lal and You, 2022). The life cycle inventories (LCIs) of all polypropylene pre-disposal life cycle stages are extracted from the well-archived Ecoinvent V3.8 Database based on their mass- and energy balance data. The LCIs of onsite energy and chemical production, including the mid-voltage electricity and process heat co-generation from polypropylene waste incineration and virgin chemicals produced from polypropylene chemical recycling, were modelled as their market processes avoided by their internal production, referring to the avoided burden approach (Curran, 2013). The PP waste chemical recycling chemical composition is extracted from relevant PP fast pyrolysis LCA studies (Zhao and You, 2021b). The mass flow rates of the plastic losses and their derived chemical emissions via degradation are evaluated based on the mass balance-based EoL fate model, which is developed based on the random sampling methodology. This methodology has more than 95% accuracy represented by less than

5% relative errors (Wang et al., 2003) in the multimedia transport and exposure dose model. The LCIs of plastic debris removal processes concerning coastal waste collection and drinking water treatment processes are modelled based on the Ecoinvent V3.8 Database (Niaz and You, 2022). The collected LCIs are then interpreted into environmental impact assessment results based on the global warming potential (GWP), ReCiPe 2016, EF3.0, and USEtox indicators in 20, 100, 500 years (Thomassen et al., 2019). The elucidation of the environmental assessment results should help reinforce the effective plastic EoL waste management choices towards pollution reduction by achieving the following goals: 1. Understanding the extent of environmental pollution caused by polymeric material losses and its effect factors; 2. Identification of the environmental pros and cons of current EoL waste management pathways. 3. Comparing the full-spectrum environmental performances of various EoL pathways.

### 3. Result and Discussion

We evaluate the environmental consequences of plastic losses over 20, 100, and 500 years to assess the time-dependent environmental effects across the entire life cycle and inform the disadvantages of plastic pollution. Figure 2 shows that global warming, ecotoxicity, air pollution corresponding to particulate matter (PM) formation, and ozone formation corresponding to photochemical ozone formation and its specific effects on human health and terrestrial ecosystems are four major environmental consequences of the chemical and MP releases. In the long run, the plastic particulates break up and yield smaller MPs and worsen the air pollution effects associated with PM formation by 15 times, while the ozone formation impacts that harm human health and the terrestrial ecosystem shift by 2,105%. The ecotoxicity effects also increase in the long run by at least four times.

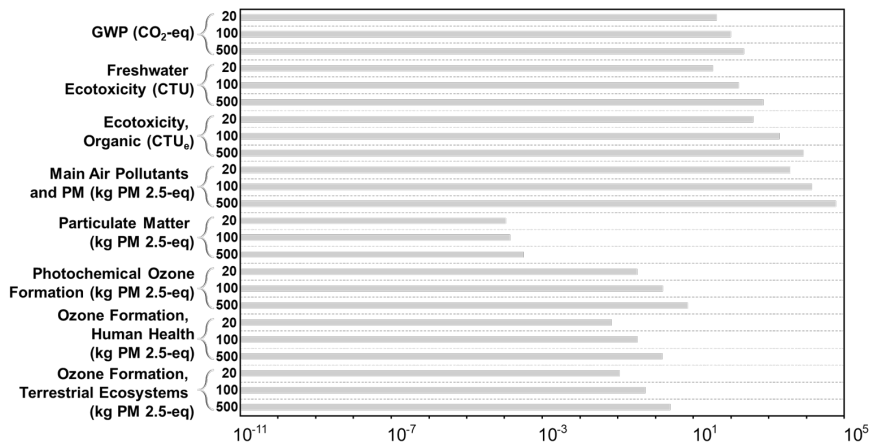


Figure 2. The absolute environmental impacts of plastic losses in 20, 100, and 500 years.

Figure 3 indicates that temperature and UV radiation dose can determine plastic losses and their subsequent chemical releases. High UV radiation levels in summer can enhance the degradation rate of the airborne polymer particulates by two times, resulting in at least 41.48% increments in full-spectrum environmental impacts. During the winter season, these environmental impacts will decrease by over 10.90% at 0 °C. Therefore, as the temperature varies, the plastic degradation rates and their derived chemical emissions can be influenced by the season change and the subsequent long-term environmental burdens. Based on the investigated environmental burdens of the plastic losses, we then evaluate the environmental effects of the PP plastic wastes treated by diverse EoL waste



management pathway, which includes landfill, incineration, and recycling. We consider three typical plastic EoL waste management pathways: 1. U.S. EoL waste management pathway; 2. 50% incineration with 50% recycling; 3. All chemical recycling aided by fast pyrolysis processes. The U.S. pathways are denoted by 2.73% recycling, 19.96% incineration, and 78.31% treated by landfills.

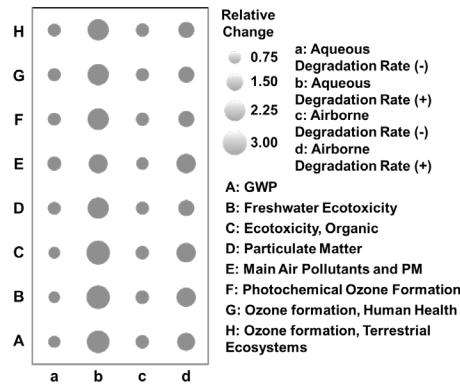


Figure 3. The effect factors of the absolute environmental impacts posed by plastic losses and the relative changes of the environmental impacts with parameter variations.

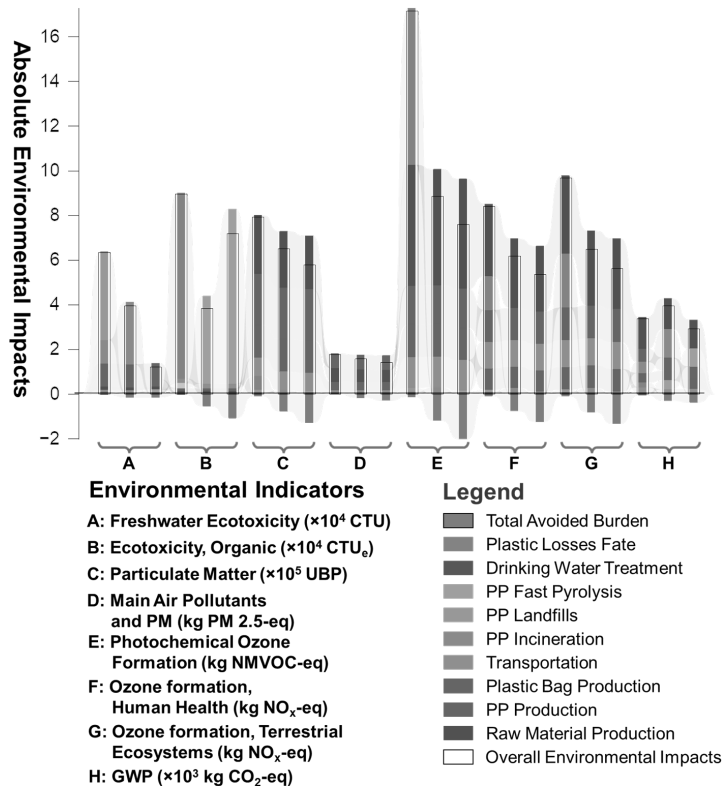


Figure 4. Environmental breakdowns of the life cycle processes of the polypropylene waste treated by U.S. EoL waste management pathway, 50% incineration with 50% recycling, and 100% chemical recycling.

Figure 4 shows that the plastic losses and their environmental impacts can be significantly reduced if other EoL waste management technology options, including incineration or recycling, replace landfilling. Moreover, plastic losses can be cut by above 99.99% if 50% of solid wastes go to incineration and another half by chemical reclamation. However, the onsite incineration processes can contribute 32.15% to the overall GHG emissions due to the intensive CO<sub>2</sub> emissions. On the other hand, fast pyrolysis processes can allocate the carbons to fuels involving gasoline, diesel, and ethylene and propylene to reduce direct GHG emissions from their onsite production. This process can also reduce the environmentally hazardous plastic losses and their derived environmental burdens, as illustrated by 92.58%, 40.87%, and 25% decrements in total ecotoxicity posed by life cycle organic chemical emissions, photochemical ozone formation, and ozone formation that harms terrestrial ecosystems, respectively. Although high organic solvent use in the aromatic extraction process within polypropylene chemically recovery processes can increase ecotoxicity from life cycle organic emissions, this EoL waste management process still outperforms other EoL options, including incineration and landfills, in terms of the full-spectrum environmental impacts. Onsite monomer and fuels and energy production aided by polypropylene chemical recycling can decrease environmental burdens from offsite manufacturing by reducing 19.91% PM formation, 26.49% photochemical ozone formation, and 23.35% ozone formation. The PP can then be used multiple times because the propylene produced from chemical recycling can be re-polymerized into PP plastics. The environmental benefits of this PP cascaded use, as illustrated by Figure 4, are shown as a decrement of 6.74% photochemical ozone formation, 5.70% ozone formation that harms human health, 5.86% ozone formation harming the terrestrial ecosystem, and 4.57% GWP. Moreover, this chemical recycling-aided PP cascaded use can save the raw material and fossil fuels by 11.25% and 25.75%, respectively. Further improving the monomer recovery rate, limiting energy use, and specialized air pollution controls on the exhaust gas can better hinder environmental consequences from plastic losses and reduce fossil resource consumption.

#### **4. Conclusion**

This work investigated the environmental burdens posed by the plastic losses from PP waste mismanagement in current practices to inform the future EoL disposal towards environmental sustainability. The environmental advantages of PP cascaded use, which is aided by chemical recycling, over other waste end-of-life management pathways can be illustrated by an above 99.99% reduction in plastic losses and their derived environmental burdens. Specifically, the volatile organic chemicals formed by plastic losses via photo-degradation can pose non-negligible global warming, ecotoxicity, and air pollution effects that worsen by at least 189% in the long run. These environmental burdens could be increased by above 9.96% under high ultraviolet radiation levels and participation rates. Based on the investigated environmental burden results, the environmental advantages of plastic cascaded use aided by fast pyrolysis technologies are specified as a reduction of 23.35% ozone formation and 19.91% air pollution compared to landfills and incineration. By replacing the external monomer and fuel production with their onsite manufacturing in PP cascade use, 25.75% of fossil fuel is saved.

#### **References**

- R.R. Bora, R. Wang, et al., 2020, Waste Polypropylene Plastic Recycling toward Climate Change Mitigation and Circular Economy: Energy, Environmental, and Technoeconomic Perspectives. *ACS Sustainable Chemistry & Engineering*, 8, 16350-16363.

- H.H. Chin, P.S. Varbanov, et al., 2022, Plastic Circular Economy Framework using Hybrid Machine Learning and Pinch Analysis. *Resources, Conservation and Recycling*, 184, 106387.
- M. A. Curran, 2013. Life cycle assessment: a review of the methodology and its application to sustainability. *Current Opinion in Chemical Engineering* 2(3), 273-277.
- Y.V. Fan, P. Jiang, R.R. Tan, et al., 2022, Forecasting plastic waste generation and interventions for environmental hazard mitigation. *Journal of Hazardous Materials*, 424, 127330.
- A. Lal, F. You, 2022. Targeting Climate - Neutral Hydrogen Production: Integrating Brown and Blue Pathways with Green Hydrogen Infrastructure via A Novel Superstructure and Simulation - based Life Cycle Optimization. *AIChE Journal*, e17956.
- K. Min, J.D. Cuiffi, R.T. Mathers, 2020. Ranking environmental degradation trends of plastic marine debris based on physical properties and molecular structure. *Nature Communications* 11(1), 1-11.
- H. Niaz, M.H. Shams, J.J. Liu, et al., 2022. Mining bitcoins with carbon capture and renewable energy for carbon neutrality across states in the USA. *Energy & Environmental Science* 15(9), 3551-3570.
- J-G. Rosenboom, R. Langer, G. Traverso, 2022. Bioplastics for a circular economy. *Nature Reviews Materials* 7(2), 117-137.
- K. V. Thomas, 2022. Understanding the plastics cycle to minimize exposure. *Nature Sustainability* 5(4), 282-284.
- G. Thomassen, M. Van Dael, S. Van Passel, et al., 2019, How to assess the potential of emerging green technologies? Towards a prospective environmental and techno-economic assessment framework. *Green Chemistry*, 21, 4868-4886.
- X. Tian, R.E. Richardson, J.W. Tester, et al., 2020, Retrofitting Municipal Wastewater Treatment Facilities toward a Greener and Circular Economy by Virtue of Resource Recovery: Techno-Economic Analysis and Life Cycle Assessment. *ACS Sustainable Chemistry & Engineering*, 8, 13823-13837.
- Z. Wen, Y. Xie, M. Chen, C.D. Dinga, 2021. China's plastic import ban increases prospects of environmental impact mitigation of plastic waste trade flow worldwide. *Nature Communications* 12, 1-9.
- X. Zhao, J.J. Klemeš, M. Saxon, F. You, 2022. How sustainable are the biodegradable medical gowns via environmental and social life cycle assessment? *Journal of Cleaner Production*, 380, 135153.
- X. Zhao, F. You, 2021, Consequential Life Cycle Assessment and Optimization of High-Density Polyethylene Plastic Waste Chemical Recycling. *ACS Sustainable Chemistry & Engineering*, 9, 12167-12184.
- X. Zhao, F. You, 2021. Waste high-density polyethylene recycling process systems for mitigating plastic pollution through a sustainable design and synthesis paradigm. *AIChE Journal*, 67, e17127.
- X. Zhao, F. You, 2021. Waste respirator processing system for public health protection and climate change mitigation under COVID-19 pandemic: Novel process design and energy, environmental, and techno-economic perspectives. *Applied Energy* 283, 116129.
- X. Zhao, F. You, 2022, Life Cycle Assessment of Microplastics Reveals Their Greater Environmental Hazards than Mismanaged Polymer Waste Losses. *Environmental Science & Technology*, 56, 11780-11797.
- S-W. Wang, P.G. Georgopoulos, G. Li, H. Rabitz, 2003. Random sampling- high dimensional model representation (RS- HDMR) with nonuniformly distributed variables: application to an integrated multimedia/multipathway exposure and dose model for trichloroethylene. *The Journal of Physical Chemistry A* 107(23), 4707-4716.
- Y. Wang, J.W. Levis, M.A. Barlaz, 2019. An assessment of the dynamic global warming impact associated with long-term emissions from landfills. *Environmental Science & Technology* 54, 1304-1313.
- X. Wu, X. Chen, R. Jiang, J. You, G. Ouyang, 2022. New insights into the photo-degraded polystyrene microplastic: Effect on the release of volatile organic compounds. *Journal of Hazardous Materials* 431, 128523.

# Enviro-economic analysis of tandem and direct processes for ethylene electrosynthesis

*Grazia Leonzio, Benoit Chachuat, Nilay Shah\**

*Sargent Centre for Process Systems Engineering, Department of Chemical Engineering,  
Imperial College London, London SW7 2AZ, UK*

*\*n.shah@imperial.ac.uk*

## Abstract

Ethylene is the most important organic chemical in terms of global demand and production capacity. Of the sustainable alternatives to conventional ethylene production based on steam cracking of natural gas and naphtha, both direct electrochemical reduction of CO<sub>2</sub> as well as a tandem process consisting of CO<sub>2</sub> electro-reduction to CO followed by CO electro-reduction to ethylene have attracted attention. This conference paper presents a comparison between the tandem and direct CO<sub>2</sub> electro-reduction processes both from an economic and environmental point of view, including a global sensitivity analysis of key process parameters on production cost and climate change impact. The results depict a clear trade-off between the economic and environmental performance of both electrochemical routes, although the tandem process remains more favorable at the current carbon price of the EU emission trading system (ETS).

**Keywords:** ethylene, tandem process, CO<sub>2</sub> electro-reduction.

## 1. Introduction

There is an urgency to reduce carbon dioxide (CO<sub>2</sub>) emissions to limit the impact of global warming and climate change. This motivates the development of low-carbon technologies based on CO<sub>2</sub> utilization for the production of major chemicals, with a view to feeding a circular chemical economy.

With global demand and production capacity of 172 and 200 Mt per year in 2021, and a forecast that these values will keep increasing in the near future, ethylene remains the most important organic chemical (Statista, 2022). Most ethylene production is currently via steam cracking of natural gas or naphtha, which is energy intensive (up to 40 GJ heat per ton of ethylene) and emits between 1.8–2 ton of CO<sub>2</sub> per ton of ethylene. About 30% of the total energy consumed by the chemical industry is used to meet the global ethylene demand, leading to the emission of 0.26 GtCO<sub>2</sub>eq (Haribal et al., 2018).

Of the alternatives to produce ethylene in a more sustainable way, direct CO<sub>2</sub> electro-reduction (CO<sub>2</sub>ER) has attracted significant attention (Berkelaar et al., 2022). Nevertheless, the performance and sustainability of this technology at scale is still hindered by low current densities, significant CO<sub>2</sub> losses by crossover in neutral membrane electrode assemblies (MEA), and carbonate formation in alkaline flow cells (AFC). In response to these challenges, Sisler et al. (2021) proposed a tandem process, whereby CO<sub>2</sub> is first reduced to carbon monoxide (CO) in solid-oxide electrolytic cells (SOEC)

followed by CO reduction to ethylene using MEA or AFC. This paper presents a comparison of the tandem and direct CO<sub>2</sub>ER processes. The techno-economic and environmental performance of the two routes is evaluated through scaling up ethylene production based on current performance of electrochemical cells, following Sisler et al. (2021) for the tandem process and Jouny et al. (2018) for the direct process. Multiple CO<sub>2</sub> feed and electricity sources are furthermore considered and uncertainty quantification is conducted for key process parameters using global sensitivity analysis (GSA).

## 2. Methodology

### 2.1. Process modelling

The tandem process is based on two electrolysis stages: CO<sub>2</sub> is first reduced to CO in an SOEC, followed by the reduction of CO to ethylene in an AFC, thereby avoiding CO<sub>2</sub> losses through carbonate formation. Modeling of these electrolytic cells is based on simple material balances and considers the optimistic conditions for the SOEC and AFC, respectively (Sisler et al., 2021): cell voltage of 1.3 V and 1.8 V, Faradaic efficiency of 100% and 90%, current densities of 1000 mA/cm<sup>2</sup>, single-pass conversion of 60% and 53%. Pressure swing adsorption (PSA) units are used after each stage for gas separation. Modeling of the electrolytic cell for direct CO<sub>2</sub>ER to ethylene is also based on simple material balances, considering alkaline conditions and the following assumptions (Jouny et al., 2018): current density of 300 mA/cm<sup>2</sup>, cell voltage of 2 V, Faradaic efficiency of 90%, CO<sub>2</sub> conversion of 50%. As with the tandem process, a PSA unit is used for gas purification.

### 2.2. Techno-economic analysis

The analysis considers a plant producing 100 ton of ethylene per year, with the mainland UK as the geographical location. The CO<sub>2</sub> sources include: natural gas processing/coal to chemicals, ammonia/bioethanol/ethylene oxide, methane steam reforming, iron and steel, cement, power generation, and direct air capture (DAC) with respective costs of \$23.7, \$35.6, \$77.1, \$83, \$106, \$88.9 and \$325 per ton of CO<sub>2</sub> (IEA, 2022); and the electricity sources include: solar, wind, conventional nuclear, and small modular nuclear reactor (SMR), with respective costs of \$0.14, \$0.09 (IRENA, 2019), \$0.059 (NAMRC, 2022) and \$0.049 per kWh (BEIS, 2016).

For the tandem process, capital (CAPEX) and operating (OPEX) expenditures are estimated using the same methodology as Sisler et al. (2021), assuming values of \$1,067 and \$300 per kW for the SOEC and AFC, respectively. For the direct process, the methodology reported in Jouny et al. (2018) is used for the overall CAPEX costs (\$920 per m<sup>2</sup> of electrolyser), while the OPEX costs follow Peter & Timmerhaus (1991). These costs are then used to estimate the production cost of ethylene.

### 2.3. Environmental assessment

The life-cycle assessment (LCA) follows the four phases defined in the ISO 14040 standards: (i) goal and scope, (ii) inventory analysis, (iii) impact assessment, and (iv) interpretation. The aim is to compare the environmental impacts of the two ethylene

production routes, considering 1 kg of ethylene production as the functional unit. The system boundaries are defined for a cradle-to-gate analysis. The foreground inventory data are provided by the process models, for a plant located in the UK. The carbon footprint of the different CO<sub>2</sub> sources is taken from Muller et al. (2020). The impact assessment relies on the Environmental Footprint 2.0 method, as recommended by the European Commission (2018), through the software SimaPro interfaced with ecoinvent.

#### 2.4. Global sensitivity analysis

A GSA is conducted to quantify the overall uncertainty in the KPIs of interest (ethylene production cost and global warming impact) by simultaneously varying all of the uncertain factors through their entire range, then apportioning a KPI's uncertainty to each factor. The uncertain factors for each route assume triangular probability distributions, shown collectively in Table 1. The GSA is conducted using the software SobolGSA, which constructs high dimensional model representation (HDMR) surrogates by drawing 1000 Sobol samples (Sobol, 2001) from the uncertainty distribution and reevaluating the process model inventories, then estimates the first- and total-order Sobol sensitivity indices (Rabitz & Salis, 1999) for the ethylene production cost and global warming impact.

Table 1: Uncertain factors in the GSA of the CO<sub>2</sub>ER direct and tandem processes

Factor	Distribution	Lower	Mode	Upper
Electricity price (\$/kWh), wind	Triangular	0.0675	0.09	0.1125
CO <sub>2</sub> price (\$/ton), NG processing	Triangular	14.22	23.7	33.18
<b>Direct CO<sub>2</sub>ER to ethylene:</b>				
Cell voltage (V)	Triangular	2	2	2.3
Current density (mA/cm <sup>2</sup> )	Triangular	200	300	300
<b>Tandem CO<sub>2</sub>ER to ethylene:</b>				
CO <sub>2</sub> conversion in SOEC (%)	Triangular	50	60	60
Current density in SOEC (mA/cm <sup>2</sup> )	Triangular	475	1000	1000
CO conversion in AFC (%)	Triangular	43	53	53
Current density in AFC (mA/cm <sup>2</sup> )	Triangular	144	1000	1000
Cell voltage in AFC (V)	Triangular	1.8	1.8	2.32
Faradaic efficiency in AFC (%)	Triangular	38	90	90

### 3. Results and discussions

#### 3.1. Economic and environmental analyses

The ethylene production costs are compared in Figure 1 for the various electricity and CO<sub>2</sub> sources. They range between \$1.3–3.7 per kg of ethylene with the tandem process and between \$2.1–8.2 with the direct process. The lowest cost corresponds to nuclear energy from SMR followed by conventional nuclear, wind and solar given the assumed

electricity prices. The contribution of CO<sub>2</sub> price to the production cost is small in comparison to electricity, except in the case of CO<sub>2</sub> from DAC that leads to significantly higher production costs. These production costs are also to be compared with the business-as-usual (BAU) technology, naphtha cracking, at a cost of \$1.3 per kg of ethylene (ICIS, 2022). Under nominal operating conditions, only the tandem process is economically competitive with the BAU, which concurs with the findings by Sisler et al. (2021).

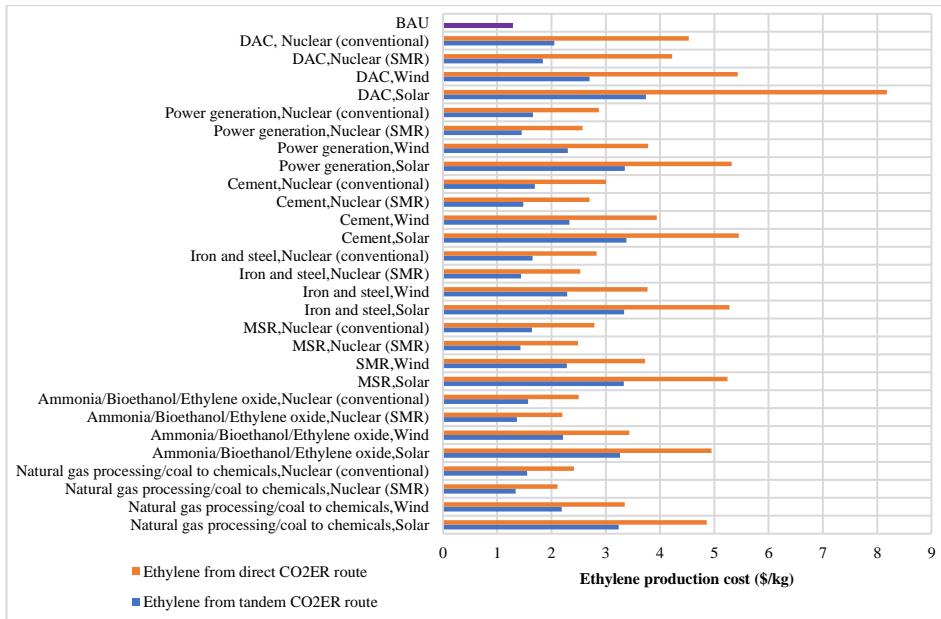


Figure 1: Ethylene production costs for the tandem and direct CO<sub>2</sub>ER processes considering different CO<sub>2</sub> and electricity sources, and comparison with the BAU process

The global warming impacts compared in Figure 2 for the same electricity and CO<sub>2</sub> sources show a different reality, whereby the direct and tandem routes have predicted impacts between -2.5–1.8 and 0.8–2.3 kgCO<sub>2</sub>-eq per kg of ethylene, respectively. Note that the direct route has negative emissions (from a cradle-to-gate accounting) in most electricity and CO<sub>2</sub> source scenarios, which concurs with the findings by Pappijn et al. (2020). Emission from the tandem route are consistently below 1 kgCO<sub>2</sub>-eq per kg of ethylene (for electricity sources other than solar), still making it a positive emitter compared to the direct route, but a significant improvement in comparison with the average global warming impact of naphtha cracking (BAU) at around 1.6 kgCO<sub>2</sub>-eq per kg of ethylene. Interestingly, even with emissions allowances (EUA) traded on ETS at a historically high carbon price of \$100 per ton of CO<sub>2</sub> the direct process fails to be competitive with the tandem process or naphtha cracking. By contrast, the tandem process could be competitive with naphtha cracking in the nuclear scenario, although still 50% more expensive than BAU with wind energy.

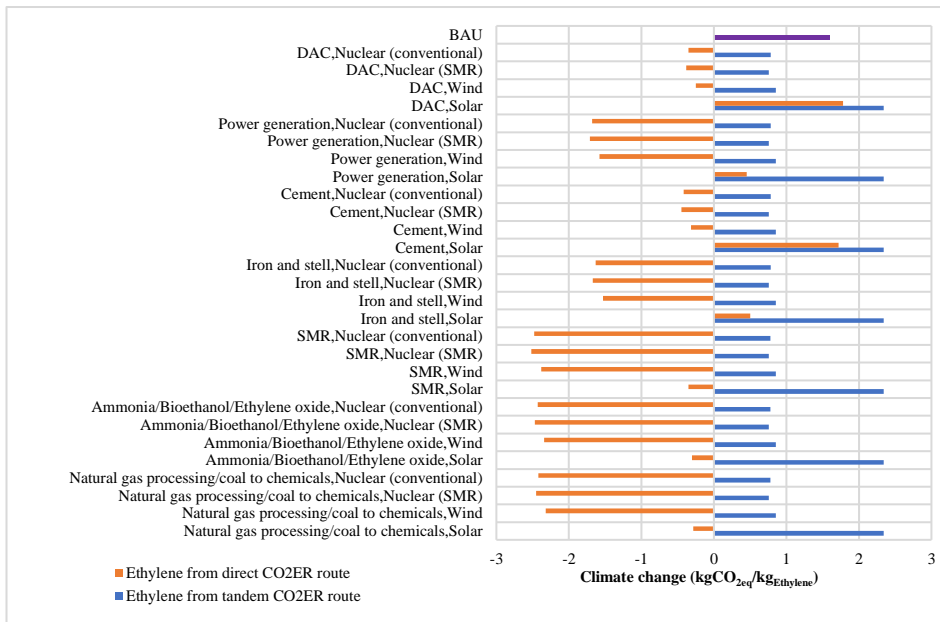


Figure 2: Global warming impact for the tandem and direct CO<sub>2</sub>ER processes considering different CO<sub>2</sub> and electricity sources, and comparison with the BAU process

### 3.2. Global sensitivity analysis

The GSA considers the scenario of wind energy and CO<sub>2</sub> captured from natural gas processing/coal to chemicals (Table 1). For the direct CO<sub>2</sub>ER route, the interquartile (IQ) range for the ethylene production cost is \$3.78–4.19 per kg of ethylene, with electricity price as the dominant factor (total-order Sobol index (TSI): 90%) followed by cell voltage (TSI: 9%); while the global warming impact is seldom affected by the uncertainty with an IQ range of -2.34–2.32 kgCO<sub>2</sub>eq per kg of ethylene. For the tandem CO<sub>2</sub>ER route, the interquartile (IQ) range for the ethylene production cost is \$2.54–3.28 per kg of ethylene, now with Faradaic efficiency in AFC (TSI: 73%) as dominant factor followed by electricity price (TSI: 20%) and cell voltage in AFC (6%); while the uncertainty in global warming impact is again small with an IQ range of 0.862–0.886 kgCO<sub>2</sub>eq per kg of ethylene. Some of these factors provide relevant targets for future technology improvements.

## 4. Conclusions

Economic and environmental analyses have been conducted for two CO<sub>2</sub> electroreduction processes producing ethylene from CO<sub>2</sub>. A comparison under different CO<sub>2</sub> and electricity sources has shown that the tandem process is economically superior to the direct route, and could even be competitive with naphtha cracking with nuclear energy. Nevertheless, the direct process presents a lower global warming impact than the tandem process, itself lower than naphtha cracking. These results depict a clear trade-off between the economic and environmental performance of both CO<sub>2</sub>ER routes, although the tandem



process remains more economically favorable even at a high carbon price of the EU's emission trading system.

**Acknowledgments:** Funding by the Engineering and Physical Sciences Research Council (EPSRC) under grants EP/V011863/1 and EP/V042432/1 is gratefully acknowledged.

## References

- L Berkelaar, J van der Linde, J Peper, A Rajhans, D Tiemessen, L van der Ham, H van den Berg, 2022, [Electrochemical conversion of carbon dioxide to ethylene: Plant design, evaluation and prospects for the future](#), *Chemical Engineering Research & Design* **182**, 194–206
- BEIS, 2016, [Electricity Generation Costs](#)
- European Commission, 2018, [Product Environmental Footprint Category Rules Guidance](#) (version 6.3, May 2018)
- P Haribal, VY Chena, L Neal, F Li, 2018, [Intensification of ethylene production from naphtha via a redox oxy-cracking scheme: Process simulations and analysis](#), *Engineering* **4**, 5, 714-721
- ICIS, 2022, [Ethylene](#)
- IEA, 2022, [Is carbon capture too expensive?](#)
- IRENA, 2019, [Solar Costs](#)
- M Jouny, W Luc, F Jiao, 2018, [General Techno-Economic Analysis of CO<sub>2</sub> Electrolysis Systems](#), *Industrial & Engineering Chemistry Research*, **57**, 2165–2177
- NAMRC, 2022, [Small Modular Reactors](#)
- LJ Muller, A Katelhon, S Bringezu, S McCoy, S Suh, R Edwards, V Sick, S Kaiser, R Cuellar-Franca, AE Khamlichi, JH Lee, N von der Assen, A Bardow, 2020, [The carbon footprint of the carbon feedstock CO<sub>2</sub>](#), *Energy & Environmental Science* **13**, 2979
- CAR Pappijn, M Ruitenbeek, MF Reyniers, KM Van Geem, 2020. Challenges and opportunities of carbon capture and utilization: Electrochemical conversion of CO<sub>2</sub> to ethylene, *Frontiers in Energy Research* **8**, 557466
- MS Peter, KD Timmerhaus, RE West, 1991, [Plant design and economics for chemical engineering](#), McGraw-Hill
- H Rabitz, OF Alis, 1999, [General foundations of high-dimensional model representations](#), *Journal of Mathematical Chemistry* **25**, 197–233
- J Sisler, S Khan, AH Ip, MW Schreiber, SA Jaffer, ER Bobicki, CT Dinh, EH Sargent, 2021, [Ethylene electrosynthesis: A comparative techno-economic analysis of alkaline vs membrane electrode assembly vs CO<sub>2</sub>-CO-C<sub>2</sub>H<sub>4</sub> tandems](#), *ACS Energy Letters* **6**, 997-1002
- IM Sobol, 2001, [Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates](#), *Mathematics & Computers in Simulation* **55**, 271 - 280
- Statista, 2022, [Ethylene demand and production capacity worldwide from 2015 to 2022](#)

# Modeling and Experimental Validation of Poly(3-Hydroxybutyrate-co-3-Hydroxyvalerate) Chain Length Distribution

Stefanie Duvigneau<sup>a,\*</sup>, Annette Wilisch-Neumann<sup>a</sup>, Robert Dürr<sup>b</sup>, Achim Kienle<sup>a,c</sup>

<sup>a</sup> *Otto von Guericke University, Germany*

<sup>b</sup> *Magdeburg-Stendal University of Applied Sciences, Germany*

<sup>c</sup> *Max Planck Institute for Dynamics of Complex Technical Systems, Germany*

\*[stefanie.duvigneau@ovgu.de](mailto:stefanie.duvigneau@ovgu.de)

## Abstract

Biodegradable and bio-based polyhydroxyalkanoates (PHAs) are ecological alternatives to conventional crude oil-based polymers. PHAs can be produced by microorganisms using waste streams from the food industry and domestic households, CO<sub>2</sub>, or other biogenic residues. To increase the current market share of PHAs, producing biopolymers with tailor-made properties by adjusting process parameters is a desirable goal.

In the biotechnology community, much is known about different cultivation conditions, e.g., bioreactor concepts or carbon sources, affecting the thermal and mechanical properties of PHAs (McAdam *et al.*, 2020). Those properties depend mainly on chain length distribution and polymer composition. Mathematical modeling can help investigate the effect of different process conditions on chain length distribution quantitatively.

This contribution outlines a multiscale model that describes the dynamics of chain length distribution and composition of the copolymer Poly(3-Hydroxybutyrate-co-3-Hydroxyvalerate) produced in a fed-batch process with *Cupriavidus necator* using fructose and propionic acid as carbon sources. Theoretical predictions are compared with experimental findings.

**Keywords:** Biopolymer, Multiscale Modelling, Polymerization, *Cupriavidus necator*, Poly(3-Hydroxybutyrate-co-3-Hydroxyvalerate)

## 1. Introduction

Compared to conventional crude oil-based polymers, biopolymers like PHAs are a promising alternative raw material for producing plastic compounds. These polymers can be produced by various bacteria and archaea using different, mostly inexpensive substrates. They are biodegradable under environmental conditions and are open to a wide range of possible applications. In comparison to conventional plastics, PHAs are still expensive. One way to reduce production costs is using organic waste from agriculture and food industries as carbon sources. In addition, polymer properties are essential. For this, monitoring the distribution and composition of PHAs is desirable to assess polymer quality and thus wanted (or unwanted) properties directly within the production process. However, for some important quantities, e.g., the chain length distribution that cannot be measured online, model-based sensors must be applied for the abovementioned goals

(Dürr *et al.*, 2021b). Furthermore, mathematical modeling is a powerful tool for investigating the dynamics of this distribution. Previous works introduced multiscale modeling approaches to simulate chain length distribution of the homopolymer Poly(3-Hydroxybutyrate) (PHB), the best-known representative of the PHA group (Penloglou *et al.*, 2012; Duvigneau *et al.*, 2022).

The current article focuses on the production process of the copolymer Poly(3-Hydroxybutyrate-co-3-Hydroxyvalerate) (PHBV). PHBV outperforms PHB because of better processability and a broader range of possible applications caused, for instance, by decreased crystallinity compared to PHB. The considered production process is a fed-batch process with *Cupriavidus necator* using fructose and propionic acid as carbon sources to produce the copolymer. Here, we present an updated metabolic model based on the model published in Duvigneau *et al.* (2021a) that leads to improved dynamic concentration profiles. By coupling this metabolic model with polymerization kinetics (Dürr *et al.*, 2021a) and calculating an average monomer concentration, we can qualitatively describe changes in PHBV composition and chain length distribution. To demonstrate this, characteristic values of chain length distribution are calculated and compared with measurements to evaluate the multiscale approach.

## 2. Experimental Methods

### 2.1. Microorganism and cultivation

*Cupriavidus necator* (H16, DSM 428) obtained from DSMZ GmbH Braunschweig was used to produce PHBV. The preculture of the bacteria before transfer in a bioreactor is described in detail in Duvigneau *et al.* (2021a). The bioreactor experiment was performed in a DASGIP parallel bioreactor system with an initial optical density of 0.4 at 600 nm. As carbon sources, 21.75 g/L fructose and 0.48 g/L propionic acid are supplemented to an M81 media. During the experiment, the dissolved oxygen (DO) level was kept at 70 % and the temperature was 30 °C. A 20 g/L propionic acid solution and 2 M sodium chloride were used to stabilize the pH at 6.8.

### 2.2. Substrate concentrations

Ammonium chloride and fructose were determined from the supernatant of the samples using an enzymatic test kit (Kit No. 5390 and No. 10139106035, R-Biopharm AG, Germany). Propionic acid concentrations were measured with an Agilent 1260 high-performance liquid chromatography (HPLC). After preparation of the supernatant, 10 µL was loaded on a reverse phase column (Inertsil 100A ODS-3, 5 µm pore size, 250 × 4.6 mm, MZ-Analysentechnik GmbH, Germany) and eluted isocratically with 1 mL/min and 0.1 M NH<sub>4</sub>H<sub>2</sub>PO<sub>4</sub> at pH 2.6 and 40 °C. The propionic acid was detected with a diode array detector (Agilent, Germany) at 210 nm.

### 2.3. Determination of PHBV and biomass concentrations

To validate our modeling approach, it is necessary to determine the distribution of the PHBV polymers. For this, 5 mL of trichloromethane was added to a fixed value of lyophilized biomass (> 4 mg). The mixture was incubated for 48 h at 37 °C and 150 rpm. After extraction, the polymer was filtered and precipitated using 5 mL ice-cold ethanol. The glass tubes were left open under the laboratory hood to ensure solvent evaporation. The extracted polymer was dissolved again in trichloromethane to achieve a concentration of 0.4 g/L. 100 µL of the solution was separated by gel permeation chromatography (GPC, Agilent + PSS, Germany) using four poly(styrol-co-divinylbenzene) columns (PSS, Germany) at 40 °C and trichloromethane as eluent. Detection of the polymer chains was done by conventional calibration with a refractive

index detector (Agilent, Germany). As reference material, 12 polystyrene standards were measured with a weight average molecular weight between 576 Da and  $1.42 \cdot 10^7$  Da. The differential distributions  $W(\log M)$  as well as the weight average molecular weight  $M_w$ , the number average molecular weight  $M_n$ , and polydispersity PDI of the samples were calculated with WinGPC UniChrom Software (PSS, Germany). The biomass concentrations were determined gravimetrically, as described in Duvigneau *et al.* (2021a). A detailed description of the procedure for determining the total 3-Hydroxybutyrate and 3-Hydroxyvalerate monomer concentrations in the PHBV polymer can be found in Duvigneau *et al.* (2021b).

### 3. Mathematical Model

The metabolic model used in the current work is based on Duvigneau *et al.* (2021a) and adapted as described below. Within the metabolic model, state equations for the dynamics of the substrates and products are given.

One key element in the advanced metabolic model approach is using online data for exhaust  $CO_2$  to follow the dynamics of substrate and product concentrations during the process. This is achieved in Duvigneau *et al.* (2021a) by introducing a factor called metabolic activity coefficient:

$$b_{CO_2} = \frac{CO_{2,out}(t)}{CO_{2,in}} \quad (1)$$

In this contribution, we use a slightly adapted version of the metabolic activity coefficient that can include the relative amount of residual biomass:

$$b_{CO_2} = \frac{CO_{2,out}(t)}{CO_{2,in} P_{res}(t)} \quad (2)$$

The relative amount of residual biomass  $P_{res}$  can be defined as follows:

$$P_{res}(t) = \frac{c_{res}(t)}{c_{res}(t) + c_{hb}(t) + c_{hv}(t)} \quad (3)$$

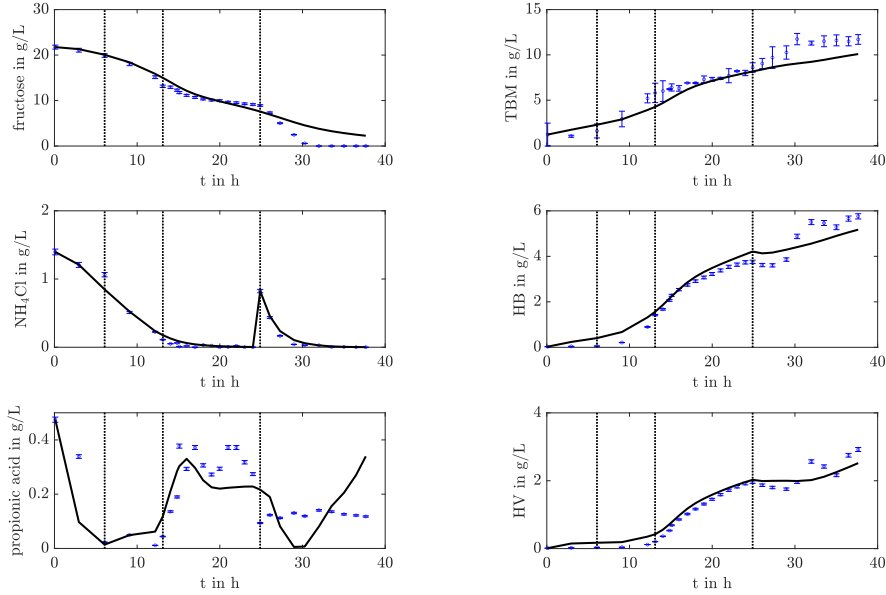
The relative amount of residual biomass  $P_{res}$  is the ratio between the concentration of residual biomass and the total biomass consisting of residual biomass and the total HB and HV monomer concentrations.

Besides the update of the metabolic activity coefficient, we decided to include the dependency not only for biomass growth but also for the uptake of fructose to produce HB (terms with  $k_4$  as rate constant). The updated state equations for fructose and HB are:

$$\frac{dc_{fru}}{dt} = -b_{CO_2} c_{res} c_{fru} [k_1 c_n inh_1 + k_4 inh_2 inh_3 + k_7] - D c_{fru} \quad (4)$$

$$\frac{dc_{hb}}{dt} = c_{res} inh_2 inh_3 [b_{CO_2} k_4 c_{fru} + k_5 c_P] - k_3 c_{res} c_n c_{hv} - D c_{hb} \quad (5)$$

The remaining equations for propionic acid, HV, ammonium chloride, and biomass can be found in our previous publication (Duvigneau *et al.*, 2021a).



**Figure 1: Dynamics of substrate (left) and product concentrations (right) for a PHBV production process.** Propionic acid is added after 10 hours by pH-depended feeding strategy. 3-Hydroxyvalerate (HV) can be included in the chain using this additional odd carbon source. The total biomass (TBM) consists of 3-Hydroxybutyrate (HB), HV, and residual biomass.

An algebraic condition can be applied as described in Duvigneau *et al.* (2022) to couple the metabolic model and polymerization kinetics:

$$\frac{1}{MW_{HA}^P} m_{HA} = \sum_{i=1}^{\infty} i ([LP]_i + [DP]_i). \quad (6)$$

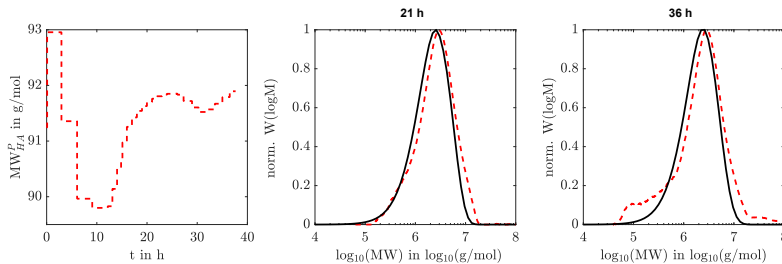
This implies that the number of all HA monomers in all polymers is the same in both model parts. For the polymerization kinetics, all active [LP] and inactive chains [DP] are multiplied by the corresponding chain length  $i$  and added up to calculate the total number of monomers. To calculate the total monomer mass  $m_{HA}$ , the total masses of HB ( $m_{HB}$ ) and HV ( $m_{HV}$ ) monomers from the metabolic model were joined:

$$m_{HA} = m_{HB} + m_{HV} \quad (7)$$

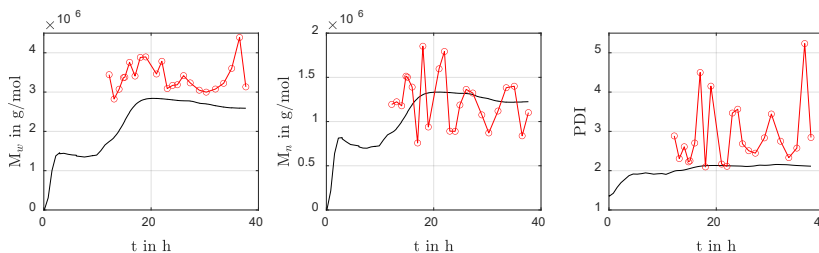
The determination of the average molecular weight of a monomer unit  $MW_{HA}^P$  in equation (6) is much more complex than in the case of the homopolymer PHB. For example, copolymer chains with the same number of monomer units can have different molecular weights caused by varying HB:HV ratios from chain to chain. Conversely, this means that with known molecular weights of the polymer chains, the composition and the number of monomers in the individual chains cannot be determined. For simplification, we assume that the HV fraction in all chains is the same and corresponds to the measured total HV fraction. With these assumptions, an average molecular weight for the monomer can be calculated:

$$MW_{HA}^P = x_{HB} MW_{HB}^P + x_{HV} MW_{HV}^P. \quad (8)$$

Therein,  $x_{HB}$  and  $x_{HV}$  are the fractions of HB and HV in the polymer measured by HPLC. The molecular weight for an HB or HV unit in a polymer ( $MW_{HB}^P$ ,  $MW_{HV}^P$ ) can be found in literature. All other equations to describe the dynamics of all chain and monomer concentrations can be found in Duvigneau *et al.* (2022). Important characteristic values of the chain length distribution to estimate the potential processing properties of the



**Figure 2: Average dynamical molecular monomer weight ( $MW_{HA}^P$ ) and normalized chain length distribution  $W(\log M)$  for chosen time points. Simulated distributions are given as solid black lines and measured distributions are shown as dashed red lines.**



**Figure 3: Characteristic values of the chain length distribution.** The characteristic values are weight average molecular weight ( $M_w$ ), number average molecular weight ( $M_n$ ), and polydispersity (PDI). Simulated values are given as solid black lines, and measured values as red circles.

polymer are the weight average molecular weight ( $M_w$ ), the number average molecular weight ( $M_n$ ), and the polydispersity index (PDI). The calculation of the values is shown in Duvigneau *et al.* (2022).

#### 4. Results

The updated metabolic model based on Duvigneau *et al.* (2021a) is adapted to datasets from experiments with fructose and propionic acid as carbon sources and another experiment with fructose as single carbon source (not shown here). With that model, it is possible to depict the dynamics of substrate and product concentrations. This is shown in Figure 1 for the setup with fructose and propionic acid as carbon sources. Propionic acid is essential to produce the copolymer PHBV but has an inhibitory effect on bacterial growth. Hence, the acid is added after 10 hours with a variable feed rate coupled to pH control to ensure that inhibitory concentrations cannot be reached (Kim *et al.*, 1992). To observe not only accumulation but also the degradation process, ammonium chloride is added after 24 h by the operator. After that, the polymer concentration decreases while consumption rates of the external carbon sources fructose and propionic acid are accelerated. The metabolic model can generally capture the concentration profiles with good agreement. Only in the experiment's last phase (>24 h) mismatch between simulation and experiments can be seen, especially for the carbon sources. One possible explanation for the gap could be the unsecure propionic acid measurement by HPLC that is covered by more apoptotic fragments in the late phase of the experiment. In addition to mapping concentration curves, the chain length distribution is essential for evaluating polymer quality. The dynamics of the chain length distribution can be simulated after coupling the metabolic model to a polymerization kinetics. Using this simple and less detailed approach in terms of the polymerization process, the monomers' average molecular weight must be assumed to simulate the chain length distribution of a PHBV copolymer. Figure 2 shows the average molecular weight of the monomers, as well as

two normalized differential distributions during the PHBV accumulation at 21 h and the end of the process at 36 h. It can be seen that the simulated differential distributions have a similar order of magnitude as the experimental distributions. However, the model simulation cannot reproduce partially occurring chains in the range of  $10^5$  g/mol. The same applies to polymers with long chain lengths. This discrepancy is also reflected in the comparison of the characteristic values shown in Figure 3.

## 5. Conclusion

In this paper, a multiscale model for the PHBV production process using fructose and propionic acid was presented. Using the model, it becomes possible to examine product yields as well as the composition and quality of the copolymer. Future work should deal with adjusting the parameters for the polymerization kinetics to achieve higher accuracy for the prediction of chain length distribution. Furthermore, the preliminary results presented here for the experimental chain length distribution should be repeated and evaluated with universal calibration to examine the influence of measurement noise on the distribution mean values (Figure 3) and simultaneously determine the exact concentrations of the chain lengths. Finally, the influence of certain cultivation conditions, such as pH, temperature, or feed composition, should be experimentally validated and integrated into the multiscale model (Pederson *et al.*, 2006; McAdam *et al.*, 2020). Such knowledge is essential to achieve online control of chain length distribution during microbial PHBV production.

**Acknowledgments** The funding of the EU-program ERDF (European Regional Development Fund) for the projects DIGIPOL (ZS/2018/11/95489) and the Research Center Dynamic Systems (CDS, ZS/2016/04/78155) is greatly acknowledged. Further, we acknowledge the financial support of the state of Saxony-Anhalt within the SmartProSys initiative.

## References

- R. Dürr, S. Duvigneau and A. Kienle. 2021a. "Microbial production of polyhydroxyalkanoates – modeling of chain length distribution." 1975–81. doi: 10.1016/B978-0-323-88506-5.50306-5.
- R. Dürr, S. Duvigneau, C. Seidel, A. Kienle, A. Bück. 2021b. "Multi-rate data fusion for state and parameter estimation in (bio-)chemical process engineering." *Processes* 9(11):1–16. doi: 10.3390/pr9111990.
- S. Duvigneau, R. Dürr, J. Behrens and A. Kienle. 2021a. "Advanced kinetic modeling of bio-copolymer poly(3-hydroxybutyrate-co-3-hydroxyvalerate) production using fructose and propionate as carbon sources." *Processes* 9. doi: <https://doi.org/10.3390/pr9081260>.
- S. Duvigneau, R. Dürr, M. Wulkow and A. Kienle. 2022. "Multiscale modeling of the microbial production of polyhydroxyalkanoates using two carbon sources." *Comput. Chem. Eng.* 160:107740. doi: 10.1016/j.compchemeng.2022.107740.
- S. Duvigneau, A. Kettner, L. Carius, C. Griehl, R. Findeisen and A. Kienle. 2021b. "Fast, inexpensive, and reliable HPLC method to determine monomer fractions in poly(3-hydroxybutyrate-co-3-hydroxyvalerate)." *Appl. Microbiol. Biotechnol.* doi: 10.1007/s00253-021-11265-3.
- J. H. Kim, B. G. Kim and C. Y. Choi. 1992. "Effect of propionic acid on poly( $\beta$ -hydroxybutyric-co- $\beta$ -hydroxyvaleric)acid production by *Alcaligenes eutrophus*." *Biotechnol. Lett.* 14(10):903–6. doi: 10.1007/BF01020626.
- B. McAdam, M. B. Fournet, P. McDonald and M. Mojicevic. 2020. "Production of polyhydroxybutyrate (PHB) and factors impacting its chemical and mechanical characteristics." *Polymers (Basel)*. 12(12):1–20. doi: 10.3390/polym12122908.
- E. N. Pederson, C. W. J. McChalicher and F. Srien. 2006. "Bacterial synthesis of PHA block copolymers." *Biomacromolecules* 7(6):1904–11. doi: 10.1021/bm0510101.
- G. Penloglou, C. Chatzidoukas and C. Kiparissides. 2012. "Microbial production of polyhydroxybutyrate with tailor-made properties: an integrated modelling approach and experimental validation." *Biotechnol. Adv.* 30(1):329–37. doi: 10.1016/j.biotechadv.2011.06.021.

# Predictive Modeling and scale-up of Wet Oxidation for Hydrothermal Liquefaction Process Water treatment

Carolin Eva Schuck,<sup>a,b</sup> Thomas Schäfer,<sup>b</sup> Konstantinos Anastasakis<sup>a,\*</sup>

<sup>a</sup> Department of Biological and Chemical Engineering, Aarhus University, Høngøvej 2, Aarhus 8200, Denmark

<sup>b</sup> Department of Chemical Engineering and Biotechnology, Hochschule Darmstadt – University of applied sciences, Stephanstraße 7, Darmstadt 64295, Germany

\*[kanastasakis@bce.au.dk](mailto:kanastasakis@bce.au.dk)

## Abstract

The present work proposes a general kinetic model for wet oxidation (WO) of hydrothermal liquefaction (HTL) process water (PW) that can aid towards upscaling, integration with HTL and overall optimization of the system. Aspen Plus is used to build a model by incorporating WO kinetics of 13 components selected to represent the composition of HTL-PW derived from sewage sludge. The model showed very good fit with experimental data for WO temperature of 350°C and residence times greater than 2min, with absolute error in COD and TOC removals prediction below 8%. The validated kinetic model was then used for the design of an upscaled WO process to examine the heat performance and energy requirements of the system. The results showed that WO can be operated autothermally with the only external energy input being related to oxygen compression and further cooling of the WO products.

**Keywords:** hydrothermal liquefaction; process water; wet oxidation; kinetics; Aspen plus

## 1. Introduction

Hydrothermal liquefaction (HTL) is not only one of the technologies with the greatest potential for the production of advanced biofuels but is also considered as an advanced technology for efficient waste valorization. However, being a wet thermochemical process, it involves processing of slurries with high water content (typically 80wt.% water), hence resulting in a large fraction of process water (PW) as a by-product stream. The formed PW has high COD (chemical oxygen demand), TOC (total organic carbon) and NH<sub>3</sub> content, hindering direct treatment by conventional biological technologies for wastewater treatment (WWT). Nevertheless, it contains between 20 and 40% of the organics present in the biomass feed (Si et al., 2019) offering the potential for an untapped energy source, that if utilized can potentially improve the overall efficiency of HTL technology.

Several different technologies, including anaerobic digestion, microbial electrolysis, electrochemical oxidation, hydrothermal gasification, etc. have been considered for the efficient valorization/treatment of HTL-PW (SundarRajan et al., 2021). All pose certain advantages and disadvantages related to energy efficiency, cost, catalyst stability, toxicity, inhibition, etc. Recently, we proposed wet oxidation (WO) as a pre-treatment technology to significantly reduce the organic loading of HTL-PW prior directing it in a conventional WWT plant for the final polishing (Silva Thomsen et al. 2022).

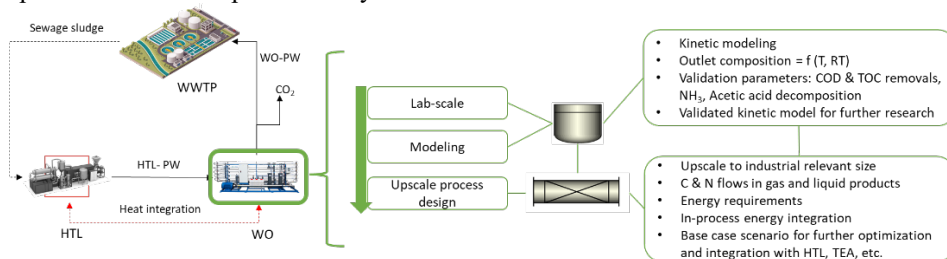
During WO a gaseous form of oxygen (oxygen, air) is used to oxidize the organics present in PW. The process takes place at elevated temperatures and pressures, maintaining PW in its liquid state and enhancing oxygen's solubility in water under these conditions. Upon



total oxidation the carbon and nitrogen present in PW are oxidized to  $\text{CO}_2$  and  $\text{NH}_3$ , respectively. However, partial instead of total oxidation usually takes place giving rise to acetic acid as an intermediate product via a free radical mechanism (Debellefontaine and Foussard, 2000). Both partial and total oxidation however result in significant reduction of the COD, TOC and toxicity contents of the PW enabling the seamless return to the WWT plant for final treatment (**Figure 1**). At the same time, the exothermicity of the oxidation reactions releases heat at an estimated flux of 435 kJ/mol  $\text{O}_2$  reacted (Debellefontaine and Foussard, 2000).

The heat release during WO opens opportunities for heat integration either ‘in-situ’ the WO process by utilizing excess heat for minimizing the overall energy needed to operate the process, or for process-to-process heat integration between HTL and WO. Until now WO has not been subject of flowsheeting and process simulations, both necessary steps for further examining the integration and upscale possibilities.

Hence, the aim of this work is twofold; first to develop a kinetic model in Aspen Plus that can give satisfactory prediction of the WO process of HTL-PW derived from sewage sludge and second, to upscale the model to industrial relevant size and examine the thermal behavior and heat integration options. **Figure 1** depicts the general concept of application of an integrated HTL-WO system in wastewater treatment plants and the specific focus of the present study.



**Figure 1.** General concept of application of wet oxidation (WO) for hydrothermal liquefaction (HTL) process water (PW) treatment and specific focus of present study (in green).

## 2. Methodology

A rigorous kinetic model is built in Aspen Plus (v12) to give realistic prediction for WO of HTL-PW. The PW considered in this study comes from HTL of sewage sludge. Typical organics present in PW from sewage sludge include alcohols, carboxylic acids, aldehydes, ketones, aromatics, N-components and N-containing aromatics (Silva Thomsen et al. 2022). Published kinetic data for WO of individual compounds present in HTL-PW were implemented in individual models in Aspen Plus and validated according to the original experimental kinetic data. Subsequently the individual kinetic reactions were combined to build a multi-reaction kinetic model by incorporating the relevant reaction mechanism. In total 13 compounds (namely: acetic acid, acetaldehyde, acetamide, acetone, butyric acid, cyclopentanone, ethanol, formic acid, formaldehyde, methanol, phenol, propanoic acid and pyridine) were selected to represent the PW stream. Typical products of WO included  $\text{CO}_2$ ,  $\text{H}_2\text{O}$ ,  $\text{NH}_3$ ,  $\text{C}_2\text{H}_4\text{O}_2$ ,  $\text{CO}$  and  $\text{H}_2$ . Own experimental data from batch-scale WO of sewage sludge HTL-PW at different temperatures and residence times (Silva Thomsen et al. 2022) were used for the validation of the multi-reaction kinetic model. The composition of the inlet PW stream was adjusted to match the Chemical Oxygen Demand (COD), ammonia and total N (TN) contents of the PW in the study used for validation. **Table 1** shows the characterization of PW by Silva Thomsen et al. 2022 vs

the one used in the present study according to the adjusted model compounds composition. For the simulation in Aspen Plus, RBATCH was used for the representation of the batch reactor used in the experiments, while the Predictive Soave–Redlich–Kwong (PSRK), which is adequate for water and hydrocarbon mixtures in the presence of light gases, was selected as the property method. An excess amount of O<sub>2</sub>, equivalent to 1.5 times the theoretical O<sub>2</sub> needed for the complete oxidation of the organics present in PW, was provided for the oxidation.

**Table 1.** Characteristics of PW from HTL of sewage sludge as determined experimentally and as represented in the model

	Experimental (Silva Thomsen et al. 2022)	Model
COD (mg/L)	28,300	28,598
NH <sub>3</sub> (mg/L)	208	208
TN (excluding NH <sub>3</sub> ) (mg/L)	477	485
TOC (mg/L)	11,900	9,205

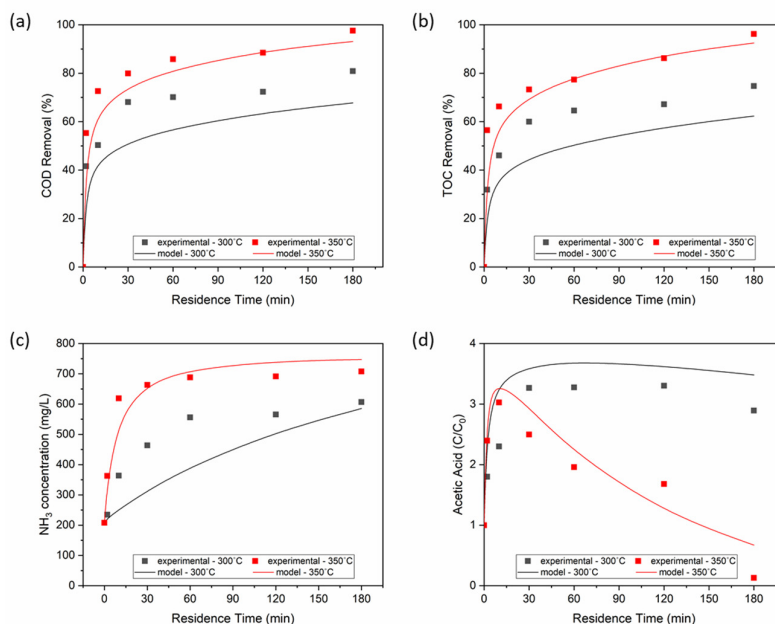
After validation, the model is scaled up to represent mass flows of PW resulting from HTL processing of sewage sludge slurry of 20 wt.% dry matter (DM) content at 1 t/h. The process flowsheet is shown in **Figure 3** where an RPlug reactor and a heat exchanger are implemented to model the scaled-up WO process. The property methods, reaction kinetics, composition of model-feed and reaction conditions remain mostly the same compared to the batch-model. Mass and energy balances have been calculated at 350°C and 15 min residence time (RT). These process conditions were selected since the WO process will likely take place at similar temperature and residence time as the HTL process and represent the base case scenario that can subsequently be further optimized.

### 3. Results & Discussion

#### 3.1. Kinetic multi-reaction model validation

The kinetic multi-reaction model, represented by RBatch reactor, was run under the exact conditions the experimental data were produced to check its validity. **Figure 2** presents predicted (solid lines) versus experimental (data points) COD and TOC removals, as well as final NH<sub>3</sub> concentration and acetic acid decomposition in post-WO PW at two different temperatures (300°C and 350°C) and residence times (0-180min). As it is clear from Figure 2 the model gave satisfactory predictions at 300°C and very good predictions at 350°C for all indicators and RT examined. The absolute error for both COD and TOC removal ranged between 6%-17% at WO temperature of 300°C and between 1%-15% at WO temperature of 350°C and different residence times. For WO temperature of 350°C maximum absolute errors for both COD and TOC removals were observed at 2min RT (9% and 15%, respectively) while the errors were significantly lower (less than 8%) for longer than 2min RT. In general, both COD and TOC removal rates predicted by the Aspen model are lower than the experimental ones for most of the RTs examined, however the general trends are consistent with the experimental data. A similar behavior is observed for NH<sub>3</sub> concentration and acetic acid decomposition in post-WO PW (figure 2-c and 2-d, respectively) with the absolute errors decreasing with increased WO temperature. Overall, the highest agreement with experimental data is reached for WO temperature of 350°C and residence times longer than 2min for all indicators examined. The two prevailing reasons for the deviation between experimental and modeling results (especially at lower temperature) are the assumed composition of the modeled PW and the implemented kinetic parameters of some of the considered compounds. A significant fraction of the components present in PW remains unidentified by conventional analysis

methods employed such as GC-MS. However, even if the complete composition was known, a model feedstock consisting of 13 compounds was assumed for the present study that might not represent the actual PW composition completely. Even though the concentration of the 13 compounds in the model feedstock was adjusted to match the experimental COD, NH<sub>3</sub> and TN contents of PW, the respective TOC content of the model feedstock was just 78% of the one experimentally measured (**Table 1**). Furthermore, the implemented kinetics for the selected 13 compounds were derived during WO of pure components. These can slightly defer when the compounds are present in a multi-component mixture as in the case of the present study. Nonetheless, despite these discrepancies, the developed model showed satisfactory (at 300°C) to very good (at 350°C) predictions for the key performance indicators (KPIs) examined. It was therefore used for the design of an up-scaled WO process that can be used as the base case scenario for further optimization of the overall system.

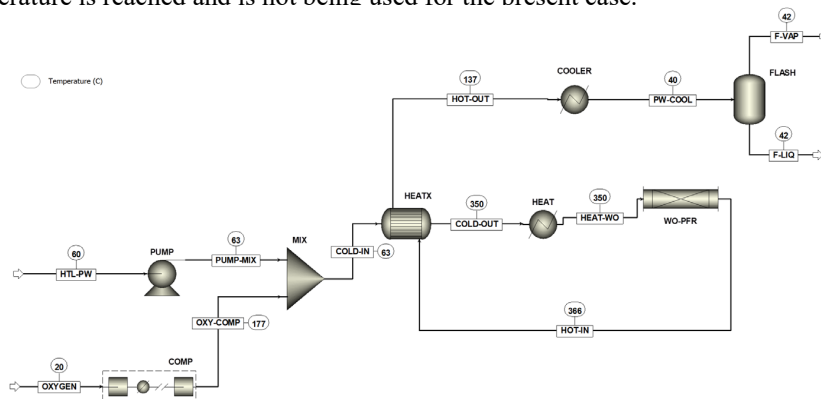


**Figure 2.** Predicted (solid lines) vs experimental (data points) post-WO PW (a) COD removal, (b) TOC removal, (c) NH<sub>3</sub> concentration and (d) acetic acid decomposition at 300°C (black lines and symbols) and 350°C (red lines and symbols) and  $2\text{min} < RT < 180\text{min}$ .

### 3.2. Process design of the scaled-up WO reactor

**Figure 3** presents the flowsheet of the scaled-up WO system. 800 kg/h of HTL-PW (assuming an upstream HTL plant with capacity of 1t/h of sewage slurry at 20wt.% dry matter content) at 60°C is pressurized at 220 bar and mixed with a compressed oxygen stream (at 220 bar) before entering the cold side of the heat exchanger (HEATX). Oxygen is fed at approximately 1.5 times the stoichiometric oxygen needed for the complete oxidation of the organics present in the modeled HTL-PW stream and is compressed in a 5-stage compressor at 220bar. The combined inlet stream (COLD-IN) enters the cold side of the heat exchanger at 63°C and is being heated to reaction temperature (350°C) by exchanging heat with the post-WO stream (HOT-IN). The overall heat transfer coefficient

(U) is assumed as  $500 \text{ W/m}^2 \cdot ^\circ\text{C}$ . Subsequently the heated stream (COLD-OUT) enters an adiabatic PFR reactor (WO-PFR) where the previously validated kinetics have been implemented. The PFR reactor consists of 15 tubes with 10m Length and 0.07m Diameter which ensures residence time of approximately 15min. The hot outlet stream from WO (HOT-IN) is cooled in two stages; first by exchanging heat with the inlet cold stream in HEATX to  $137^\circ\text{C}$  and then by a simple cooler to  $40^\circ\text{C}$ , before separating the stream into vapor (F-VAP) and liquid (F-LIQ) products in a flash drum (FLASH). The HEAT block between the heat exchanger and the WO reactor is used to ensure that the desired reaction temperature is reached and is not being used for the present case.

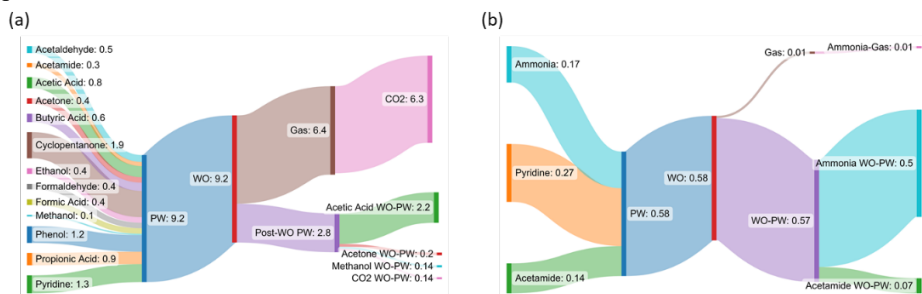


**Figure 3.** Process flowsheet of scaled-up WO process, with stream temperatures indicated on the flowsheet, in Aspen Plus (v12).

The results from the upscaled system indicate that WO can be operated autothermally with the only external energy input being related to oxygen compression and further cooling of the WO products. Autothermal operation is achieved due to the exothermicity of the oxidation reactions, increasing the temperature of the WO outlet stream to  $366^\circ\text{C}$  in the present case. Heat capacity of liquid water increases substantially at elevated temperatures, especially above  $340^\circ\text{C}$ . As an indication  $c_p$  of liquid water increases from  $135.9 \text{ J/mol}\cdot\text{K}$  at  $350^\circ\text{C}$  to  $210.7 \text{ J/mol}\cdot\text{K}$  at  $365^\circ\text{C}$  (Lemmon et al., 2022). Additional external utilities are estimated as  $5.6 \text{ kW/m}^3$  of HTL-PW for oxygen compression (electricity),  $-4.6 \text{ kW/m}^3$  of HTL-PW for compressor cooling and  $-109.9 \text{ kW/m}^3$  of HTL-PW for final cooling of WO products, for the present case with the assumed organic loading of HTL-PW. The results suggest a unique potential for heat integration with HTL process which needs to be further examined. The developed model sets strong fundamentals for the subsequent investigation of the heat integration options.

Carbon and nitrogen flows in HTL-PW assumed in this study and in WO products (gas and WO-PW) are depicted in **Figure 4** (4-a and 4-b, respectively) in the form of a Sankey diagram. The majority of carbon ( $\sim 70\%$ ) present in HTL-PW is oxidized to gaseous products upon WO (figure 4-a). The gaseous stream is composed almost entirely of  $\text{CO}_2$  ( $\sim 98.5\%$ ) offering a relative pure- $\text{CO}_2$  stream that can be potentially further utilized to produce synthetic fuels in a power-to-X approach. The liquid stream from WO still contains approximately 30% of the inlet carbon with most of the carbon ( $\sim 24\%$  of initial TOC) found in the form of acetic acid. Acetic acid is an exceptionally refractory intermediate and is also formed during oxidation reactions of several other organic components (Shende and Levec, 1999). As also shown previously (Figure 2-d) higher temperatures and residence times are required for its decomposition. COD and TOC removals for upscaled WO at  $350^\circ\text{C}$  and 15min were 73.5% and 70.8%, respectively, in

line with both modeling and experimental results obtained from batch scale WO shown in **Figure 2**. Unlike carbon, that is mostly distributed in the gas phase as CO<sub>2</sub>, nitrogen, present in the N-containing components assumed in HTL-PW (pyridine, acetamide), end up as ammonia in WO-PW.



**Figure 4.** (a) Carbon and (b) Nitrogen flows ( $\text{kg}/\text{m}^3$  PW) before and after WO at 350°C and 15min RT as predicted by the developed model.

#### 4. Conclusions

A kinetic multi-reaction model was developed in Aspen Plus to predict the wet oxidation of organics present in process water from hydrothermal liquefaction of sewage sludge. The model showed satisfactory to very good predictions with batch-scale experimental data on COD and TOC removals, NH<sub>3</sub> and acetic acid concentrations in post-WO PW. Especially at higher oxidation temperatures (350°C) and longer than 2 min RT, the absolute error was less than 8% in all key performance indicators examined. Process up-scale revealed that WO can be operated autothermally, leveraging on the excess heat generated by the oxidation reactions. At the same time, the results revealed further potential for heat-integration with HTL technology. The developed kinetic model sets the base case scenario to assist in a subsequent process up-scale optimization design, in terms of tuning WO process conditions to achieve both the desired organic loading reduction and the desired heat release for the different integration options and HTL-PW composition.

#### References

- H. Debellefontaine, J.N. Foussard, Wet air oxidation for the treatment of industrial wastes. Chemical aspects, reactor design and industrial applications in Europe, *Waste Manage.* 2000, 20 (1), 15–25.
- E.W. Lemmon, I.H. Bell, M.L. Huber, M.O. McLinden, 2022, Thermophysical Properties of Fluid Systems, in NIST Chemistry WebBook, NIST Standard Reference Database Number 69, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899
- R.V. Shende, J. Levec, 1999, Wet Oxidation Kinetics of Refractory Low Molecular Mass Carboxylic Acids, *Ind. Eng. Chem. Res.*, 38 (10), 3830–3837
- B. Si, L. Yang, X. Zhou, J. Watson, G. Tommaso, W.T. Chen, Q. Liao, N. Duan, Z. Liu, Y. Zhang, 2019, Anaerobic conversion of the hydrothermal liquefaction aqueous phase: fate of organics and intensification with granule activated carbon/ ozone pretreatment, *Green Chem.* 21 (6), 1305–1318.
- L.B. Silva Thomsen, K. Anastasakis, P. Biller, 2022, Wet oxidation of aqueous phase from hydrothermal liquefaction of sewage sludge, *Water Research*, 209, 117863.
- P. SundarRajan, K. Gopinath, J. Arun, K. GracePavithra, A.A. Joseph, S. Manasa, 2021, Insights into valuing the aqueous phase derived from hydrothermal liquefaction, *Renew. Sustain. Energy Rev.* 144, 111019.

# Process Feasibility for Converting the Waste Plastics to Hydrogen and Methanol

Ali A. Al-Qadri <sup>a</sup>, and Usama Ahmed <sup>a,b\*</sup>

<sup>a</sup>*Department of Chemical Engineering, King Fahd University of Petroleum and Minerals, Dhahran, 31261. Saudi Arabia*

<sup>b</sup>*Interdisciplinary Research Center for Hydrogen and Energy Storage, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia*

[usama.ahmed@kfupm.edu.sa](mailto:usama.ahmed@kfupm.edu.sa)

## Abstract

Hydrogen and methanol are among the most vital fuels and chemicals nowadays. Hydrogen is considered as clean fuel because it produces zero carbon emissions. Comparably, methanol is regarded as a low-carbon alternative fuel. The resources of the two fuels are mainly crude oil, coal, and natural gas. Alternative sources of the two fuels are highly needed to meet the increasing energy demand. On the other hand, plastic waste is also a pollutant and drastically affects the environment. Thermochemical recycling and processing of plastic waste could lead to hydrogen and methanol production. Two comparative models have been developed in Aspen Plus to produce hydrogen and methanol from the waste plastics. In base model (case 1), the waste plastic is first gasified using steam as a gasification agent and the produced syngas containing H<sub>2</sub>, CO, and CO<sub>2</sub> is introduced to the water gas shift unit to maximize hydrogen production. Finally, the CO<sub>2</sub> was captured to recover the pure hydrogen. In the alternative model (case 2), the integration between plastic gasification and steam methane reforming (SMR) is proposed to increase the syngas yield while minimizing the overall process energy requirements. The two cases have been technically and economically evaluated to determine the best case. The results showed that the process efficiency of case 2 is higher than base case by 3%. The case 2 also showed a potential to reduce the fuel production cost by 51.8% as compared to case 1. The CO<sub>2</sub> specific emissions in case 2 are reduced by 34% as compared to case 1. Overall, it has been seen that integrating the SMR process with plastic gasification for the dual hydrogen and methanol production is less energy intensive and represents better project feasibility.

**Keywords:** Gasification; SMR; Waste Plastic; Methanol production; GHG Emissions.

## Introduction

The global production of the plastics over the past 70 years is more than 6.3 billion tons (Okunola A et al., 2019). The huge amount of plastic waste has been contributing to huge environmental problems and distributing the earth ecosystem. The thermochemical conversion of waste plastics is 12%, whereas, 9 % is usually recycled (Okunola A et al., 2019). Therefore, proper recycling of the plastics should be considered to minimize its massive environmental impacts and to utilize them in beneficial applications. The thermochemical recycling of plastic waste is a good option because it eliminates the need for plastic classification and simultaneously produce syngas that could be used to produce several fuels and chemicals with high purity. Plastics in general have a high heating value with high hydrogen and carbon content, thus, syngas with high heating value can be produced by thermochemical conversion. This route offers alternatives to produce clean

fuels from carbon-based raw materials. Hydrogen and methanol are some of the essential fuels that could be produced from waste plastics with a minimum carbon emission. Therefore, the recycling of the plastics should be focused in this approach synthesizing hydrogen and methanol from plastic waste via thermochemical route. AlNouss et al. (2020) performed a techno economic analysis on the conversion of carbonaceous fuels to hydrogen testing several feedstocks and several gasification agents. It was concluded that steam as a gasification agent performed better than others in terms of maximizing the hydrogen production with considerable cost. Another related study conducted by Al-Qadri & Ahmed (2022) on the conversion of waste plastics (polyethylene and polypropylene) to hydrogen via steam gasification. They performed the whole technical and economic analysis investigating an alternative model that integrates the steam methane reforming with plastic gasification to maximize the hydrogen production, enhance the syngas quality, and reduce the cost. They found a net increase in the process efficiency by 4% and a reduction in the levelized hydrogen cost by 29%. Ahmed (2021) also performed a technoeconomic analysis on the dual production of hydrogen and methanol from coal, integrating the steam methane reforming with coal gasification. The study was succeeded in enhancing the fuel production by 22% and reduce the fuel cost by 13% when compared to the conventional methodology. Since there is an interest to produce hydrogen and methanol from alternative resource (plastic waste) to meet the energy demand, the integrated design between plastic gasification and steam methane reforming is an option. Therefore, the objective of this work is to develop two case studies that produce hydrogen and methanol from waste plastics (polyethylene blended with polypropylene) and then performing the technoeconomic analysis for different process alternatives.

### Modelling and Simulation

Aspen Plus (V12) was used to build the simulation models for the dual production of hydrogen and methanol. The thermodynamic package used in the study is Peng Robinson because it is applicable to high temperature and pressure, and several studies had used it in the gasification field. The composition of polyethylene and polypropylene were defined on the basis of their approximate and ultimate analysis. Table 1 represents the assumptions and the modules used for developing the simulation model. The plastic gasification results were validated with the literature (Saebea et al., 2020). Additionally, the methanol synthesis from syngas was also confirmed in terms of methanol yield (Mccaul, 2019). The steam methane results were also validated based on hydrogen to carbon monoxide ratio (Ghoneim et al., 2016).

Table 1: Design Assumptions for case 1 and case 2

Equipment	Aspen Model	Assumption
Flowrate of plastics	RYield	Plastics = 100 kg/h Steam/Plastic = 1.25
Gasifier	RGibbs	Entrained flow gasifier 1500 °C and P = 25 bars
Pre-reformer	RStoic	Methane/Plastics (mass basis) = 1.5
Reformer	RGibbs	1200 °C and 25 bar, steam: methane (mass basis) = 1.6; with nickel-based catalyst
H <sub>2</sub> S Removal	RadFrac	Rectisol process; 0 °C, and 25 bar H <sub>2</sub> S remaining = 1 ppm using 20 stages.
Methanol Synthesis	Plug flow reactor	25 bars and 270 °C Using Graaf. et al kinetics

Water Gas Shift	REquil	Two reactors, Molar steam: CO = 2 using $\gamma$ - $Al_2O_3$ - $\alpha$ - $Ga_2O_3$ catalysts
CO <sub>2</sub> Removal	RadFrac	Rectisol process; 0 °C, and 1 bar CO <sub>2</sub> Removal = 99% using 10 stages.

### 3. Process Description

Two case studies have been developed in this study for dual production of H<sub>2</sub> and methanol. Figure 1 represents the base case (case 1) which consists of several units containing steam plastic gasification, H<sub>2</sub>S removal, methanol synthesis unit, water gas shift (WGS), CO<sub>2</sub> removal unit, and solvent regeneration. The syngas produced from gasification unit is treated in the acid gas removal unit to remove the H<sub>2</sub>S followed by its conversion into methanol synthesis unit, whereas, the remaining gases were fed to the WGS to maximize the hydrogen content followed by the CO<sub>2</sub> removal.

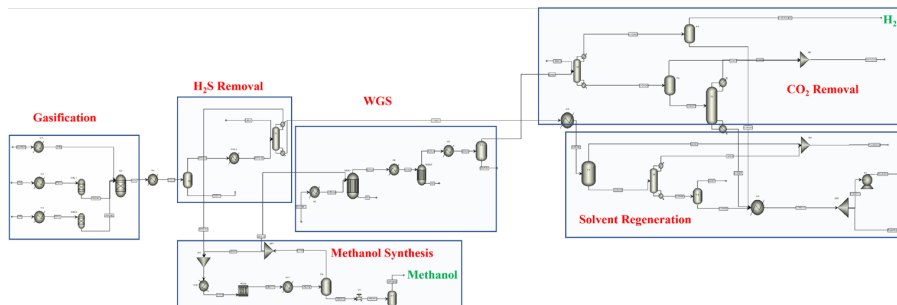


Figure 1: Aspen Plus Model for the Production of H<sub>2</sub> and Methanol from Waste Plastics

The case 2 shown in Figure 2 also represents the dual production of H<sub>2</sub> and methanol. It is similar to case 1, however, this case integrates the steam methane reforming (SMR) technology with plastic gasification, where, the produced syngas from gasification and reforming unit is mixed to enhance the overall syngas production. Then H<sub>2</sub>S is removed from the syngas feed to make it suitable for the methanol synthesis unit and the unconverted syngas is recycled back to WGS unit to get the maximum conversion of unconverted CO to H<sub>2</sub>.

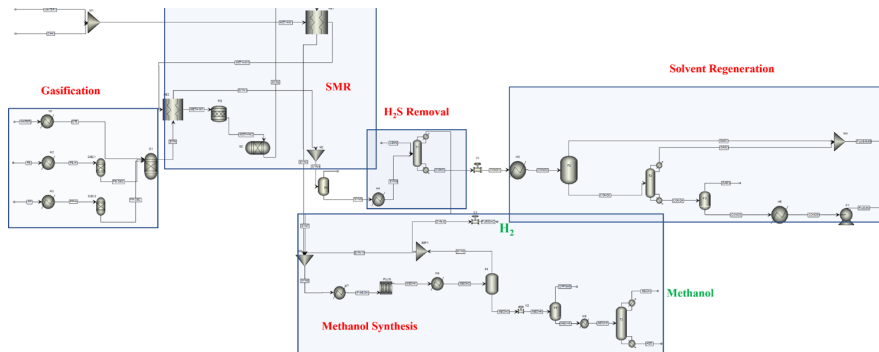


Figure 2: Aspen Plus Model of Waste Plastic Gasification and Methane Reforming models for the Production of H<sub>2</sub> and Methanol



## 4. Results and Discussion

### 4.1. Hydrogen and Methanol Production

The results for the two cases were analyzed in terms of syngas compositions, hydrogen production, methanol production, CO<sub>2</sub> emissions, and overall process efficiency. The case 2 produces syngas with higher hydrogen content of 9% and higher heating value by 11% compared to case 1. The hydrogen production rate on mass per unit feedstock basis (hydrogen yield) has been calculated as 0.06 and 0.10 for case 1 and case 2, respectively. On the other hand, the methanol production rate on mass per unit feedstock basis for case 1 and case 2 has been calculated as 2.12 and 2.18, respectively. From the results, it has been seen that the second case is more efficient than case 1 in term of producing H<sub>2</sub> and CH<sub>3</sub>OH.

### 4.2 Process performance analysis

The following equation (Eq 1) is used to calculate the overall process efficiency.

$$\text{Process Efficiency} = \frac{\text{H}_2 \text{ thermal energy [MWth]} + \text{CH}_3\text{OH thermal energy [MWth]}}{\text{Feed stock thermal energy [MWth]} + \text{Energy consumed [MWth]}} \times 100\% \quad (1)$$

Case 2 utilizes an additional feedstock namely, methane, intaking more feedstock thermal energy of 1995.92 MW. In addition, the second case required 1070.13 MW more utilities for the process needs. While comparing the total energy content, it has been seen that the thermal energy content in terms of produced hydrogen and methanol was higher in the case 2 than case 1 by 2392.96 MW. Developing the heat exchanger network and calculating all the energy requirements using Equation 1, the overall process efficiencies for Case 1 and case 2 is estimated as 73% and 76%, respectively, as represented in the Figure 3. The results also showed that the yield of H<sub>2</sub> and methanol in case 2 is higher than case 1. Additionally, the second case reduces the specific carbon dioxide emissions by 34%. Therefore, case 2 found to be more effective and feasible for dual hydrogen and methanol production from waste plastics. Furthermore, the economic analysis is performed to analyse the overall process feasibility.

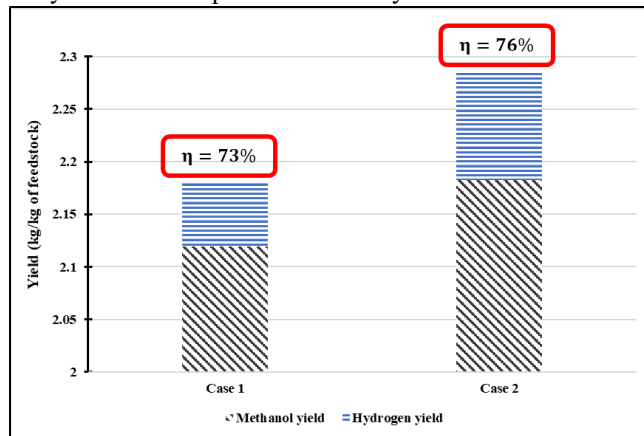


Figure 3: The fuel production and process efficiency for the two cases

## 5. Project Feasibility

The project feasibility is determined via economic analysis which basically included the capital expenditures and operating expenditures for the two cases. The economic analysis was performed based on several assumptions specified by the previous paper (Al-Qadri et al., 2022). The capital expenditure (CAPEX) mainly encompasses the equipment cost,

installation and piping costs, instrumentations, and civil work costs. On the other hand, the operational expenditure (OPEX) consists of utilities, catalyst replacement maintenance, labor costs, and administration costs. The power law (order of magnitude) analysis has been used to estimate the CAPEX, where, the CEPCI of 620 has been used for the year 2022 in the equation (Eq 2).

$$\text{Cost}_{\text{New}} = \text{Cost}_{\text{Old}} \times \left( \frac{\text{Capacity}_{\text{New}}}{\text{Capacity}_{\text{Old}}} \right)^x \times \frac{\text{CEPCI}_{\text{New}}}{\text{CEPCI}_{\text{Old}}} \tag{2}$$

The total investment cost of case 2 is higher than case 1 by 40%. Nevertheless, the total investment cost (TIC) per ton of fuel produced is lower in case 2 than case 1 by 46% which shows the feasibility of the second case in terms of higher fuel yield. The fuel levelized cost over 30 years was reduced in the case 2 by 52% when compared to the base case. The net present value for case1 and case 2 were calculated to be 2.734 M€ and 26.201 M€ respectively. Additionally, the present value ratio for the two cases were 1.614 M€ and 5.301 M€. Table 2 represents the main parameters for the technoeconomic analysis. The economic analysis has been presented in terms of cash flow diagram in Figure 4.

Table 2: Technoeconomic Analysis

Parameter	Units	Case 1	Case 2
Produced hydrogen per feedstock	mass basis	0.06	0.10
Produced methanol per feedstock	mass basis	2.12	2.18
Process Efficiency	%	73	76
Total Investment Cost	€ (10 <sup>3</sup> )	4070	5690
TIC per ton of H <sub>2</sub> and MeOH	M€/ton	18530	9970
Total OPEX/Year	€ (10 <sup>3</sup> ) / Year	839.1	1032.2
Total OPEX/ton H <sub>2</sub> and MeOH	€ (10 <sup>3</sup> ) / ton	16.4	4.9
Revenue	M€/Year	1.7	5.2
Hydrogen and MeOH Cost	€/kg	0.534	0.257

Case 2 is more attractive compared to the case 1 in terms of cash flow and return on investment. This significant enhancement in the return on investment is attributed to the higher revenue generation capacity of case 2 since it represented higher annual income of 3.526 M€ compared to base case. Case 2 can pass the breakeven point in the fifth year, whereas, case 1 requires 12 years to return the initial investment. Therefore, case 2 has been found to be more economically feasible for dual production of hydrogen and methanol.

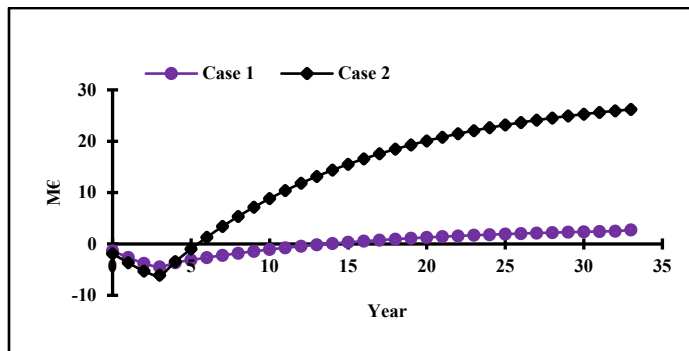


Figure 4: The cash flow diagram for base case and alternative case

## Conclusions

The dual production of hydrogen and methanol from waste plastics has been represented in two different cases. The first case represents on plastic gasification design for the production of hydrogen and methanol. On the other hand, case 2 integrates the steam methane reforming with plastic gasification to enhance the syngas heating value with higher hydrogen content. The results revealed that, case 2 produces 65% and 3% more hydrogen and methanol, respectively, compared to the case 1. The overall process efficiency of the case 2 showed an increment of 3% and reduction of CO<sub>2</sub> specific emissions by 34% compared to case 1. Moreover, Case 2 offers an annual revenue of 5.2 M€ compared to case 1, which is only 1.7 M€. The fuel (hydrogen and methanol) production cost has been found to be lower in the case 2 by 51% as compared to case 1. Keeping in view all the preliminary analysis, case 2 has been found to be more technoeconomically attractive compared to the case 1.

## Acknowledgements

The authors would like to acknowledge support provided by the Deanship of Research Oversight and Coordination (DROC) at the King Fahd University of Petroleum & Minerals (KFUPM) for funding this work through project no. DF201017.

## References

- Ahmed, U. (2021). Techno-economic analysis of dual methanol and hydrogen production using energy mix systems with CO<sub>2</sub> capture. *Energy Conversion and Management*, 228, 113663. <https://doi.org/10.1016/J.ENCONMAN.2020.113663>
- AlNouss, A., McKay, G., & Al-Ansari, T. (2020). Enhancing waste to hydrogen production through biomass feedstock blending: A techno-economic-environmental evaluation. *Applied Energy*, 266. <https://doi.org/10.1016/j.apenergy.2020.114885>
- Al-Qadri, A. A., & Ahmed, U. (2022). Techno-Economic Analysis of the Conversion of Waste Plastics to Hydrogen Fuel. *Computer Aided Chemical Engineering*, 51, 421–426. <https://doi.org/10.1016/B978-0-323-95879-0.50071-0>
- Al-Qadri, A. A., Ahmed, U., Abdul Jameel, A. G., Zahid, U., Usman, M., & Ahmad, N. (2022). Simulation and Modelling of Hydrogen Production from Waste Plastics: Technoeconomic Analysis. *Polymers*, 14(10). <https://doi.org/10.3390/POLYM14102056>
- Ghoneim, S. A., El-Salamony, R. A., & El-Temtamy, S. A. (2016). Review on Innovative Catalytic Reforming of Natural Gas to Syngas. *World Journal of Engineering and Technology*, 04(01), 116–139. <https://doi.org/10.4236/wjet.2016.41011>
- Mccaul, S. C. (2019). *Simulation and Techno-Economic Analysis of the AChT Green Methanol Process*.
- Okunola A, A., Kehinde I, O., Oluwaseun, A., & Olufiropo E, A. (2019). Public and Environmental Health Effects of Plastic Wastes Disposal: A Review. *Journal of Toxicology and Risk Assessment*, 5(2). <https://doi.org/10.23937/2572-4061.1510021>
- Saebea, D., Ruengrit, P., Arpornwichanop, A., & Patcharavorachot, Y. (2020). Gasification of plastic waste for synthesis gas production. *Energy Reports*, 6, 202–207. <https://doi.org/10.1016/j.egyr.2019.08.043>

# Process development for efficient hydrogen transportation and analyses: technical, economic, and environmental perspectives

Byeongchan Ahn,<sup>a</sup> J. Jay Liu,<sup>b</sup> Wangyun Won<sup>a,\*</sup>

<sup>a</sup> *Department of Chemical Engineering (Integrated engineering), Kyung Hee University, 1732 Deogyong-daero, Giheung-gu, yongin-si, Gyeonggi-do 17104, Republic of Korea*

<sup>b</sup> *Department of Chemical Engineering, Pukyong National University, 45, Yongso-ro, Nam-gu, Busan 48513, Republic of Korea*

*\*corresponding author (e-mail: wwon@khu.ac.kr)*

## Abstract

Here, we have designed an integrated hydrogen transport process including hydrogenation and dehydrogenation processes using a toluene-methylcyclohexane (MCH) system. The toluene-MCH system is one of the promising liquid organic hydrogen carrier (LOHC) candidates because it is relatively less toxic and an economical option. For efficient energy utilization, we have carried out heat integration using pinch analysis. To determine the minimum hydrogen transport cost, techno-economic analysis was performed using discounted cash flow analysis. It was estimated that the minimum hydrogen (H<sub>2</sub>) transport cost is \$2,083/kg when 400,000 tons of hydrogen is transported per annum. The sustainability of the proposed process was determined through life-cycle assessment quantifying the direct and indirect carbon emissions. This study can give insight into technological development for the realization of the hydrogen economy.

**Keywords:** Eco-friendly, Hydrogen storage-transportation, Clean energy, Separation process, comprehensive analysis

## 1. Introduction

Recent climate crisis has prompted researchers to take an interest in the transition from a petrochemical-based energy society to a renewables-based energy society. However, regional inconsistency and intermittent energy generation make the commercialization of renewable energy complicated. To resolve this drawback, hydrogen is attracting attention as a carbon-free energy carrier. Although the high gravimetric energy density of hydrogen (120 MJ/kg) is suitable for an energy carrier, its low volumetric energy density (10.9 kJ/L) is a major contributor to increase transport costs (Markiewicz et al., 2019).

There are two major methods for efficient hydrogen storage: physical storage and chemical storage. For physical storage, compression and liquefaction are mainly used. Although they have the advantage of relatively simple processes, their strict storage conditions (400~700 bar and -253°C, respectively) entail safety accident risk and require lots of energy causing indirect carbon emissions. Chemical storage means storing hydrogen in other chemicals by forming chemical bonds. Chemical storage requires sophisticated separation processes to recover high-purity hydrogen, but it consumes relatively low energy (Singh et al., 2021; Teichmann et al., 2011).

Liquid organic hydrogen carriers (LOHCs) are promising candidates for chemical storage systems due to their high hydrogen storage capacity. Furthermore, LOHCs can store hydrogen in the liquid state at ambient conditions with relatively low energy loss and has similar properties with petroleum derivatives, which makes it available to use existing petrochemical infrastructures. LOHCs are well known in industrial set-ups and so they are easy to produce in large quantities.

Among LOHCs, the toluene-methylcyclohexane (MCH) system was selected for the designed process due to its relatively high hydrogen storage capacity (6.2 wt.%) (Niermann et al., 2019). It also has a relatively affordable price and reversible hydrogenation reaction which makes toluene recycle after dehydrogenation to reduce the hydrogen transport cost (Preuster et al., 2018). However, toluene and MCH have high volatilities that make it hard to separate hydrogen from unwanted organics.

In this study, we have designed the integrated hydrogen transport process composed of two sections (hydrogenation and dehydrogenation sections) based on the experiment data from the literature. Especially, the sophisticated separation process including the extraction column and pressure swing adsorption (PSA) column was designed considering the volatile property of toluene and MCH. We have also made heat networks to use waste energy efficiently reducing indirect carbon emissions. Then, techno-economic analysis (TEA) and life-cycle assessment (LCA) were performed to evaluate economic feasibility and sustainability.

## 2. Technology Overview

Toluene in the proposed design is converted to MCH by a hydrogenation reaction and converted back to toluene by a dehydrogenation reaction. For the hydrogenation reaction, we have used a commercially available Ni-SiO<sub>2</sub> catalyst at 120°C with excess hydrogen condition. Although the experimental data from the literature was originally based on ambient pressure, we designed it to react at higher pressure (370 psi) (Atsumi et al., 2020). Generally, hydrogenation reaction at higher pressure results in higher conversion, but we assumed that it has the same yield at ambient pressure due to no experimental data at higher pressure. Instead, the high-pressure reaction can reduce the cost required for pressurization when separating unreacted hydrogen. Unlike the hydrogenation reaction, the dehydrogenation reaction has a higher conversion at lower pressure. MCH dehydrogenates under 1atm at 310°C (Okada et al., 2006). More details including stoichiometric equations and yields are specified in Table 1.

Table 1. Summary of reaction condition.

Reaction	Reaction	Yield
Hydrogenation	$\text{Toluene} + 3 \text{H}_2 \rightarrow \text{MCH}$	0.99
	$\text{Toluene} + \text{H}_2 \rightarrow \text{CH}_4 + \text{Benzene}$	0.01
Dehydrogenation	$\text{MCH} \rightarrow \text{Toluene} + 3 \text{H}_2$	0.9491
	$\text{MCH} \rightarrow \text{CH}_4 + \text{Benzene} + 2 \text{H}_2$	0.0002
	$\text{MCH} + \text{H}_2 \rightarrow \text{Others}$	0.0008

### 3. Method

#### 3.1. Process synthesis

The process is developed to target the white birch handling of 2,000 dry tons per day which is the commercial scale of biorefinery. The process is simulated using the Aspen Plus process simulator V11 and the Soave-Redlich-Kwong (SRK) equation as the thermodynamic method. The heat networks are created based on the pinch analysis method using the Aspen Energy Analyzer V11.

#### 3.2. Techno-economic analysis

The cost of equipment is calculated using the Aspen Process Economic Analyzer V11, and the equipment and material costs are scaled to unify the base year of 2018 using the Chemical Engineering Plant Cost Index. Using the scaled costs, the minimum transport cost is estimated as an indicator of economic feasibility by discounted cash flow method when the internal rate of return is assumed to be 10%. The economic parameters are derived from the National Renewable Energy Lab report, and major economic parameters are summarized in Table 2 (Davis et al., 2018).

Table 2. Major assumed economic parameters.

Economic parameter	Value	Economic parameter	Value
Plant financing by equity	40.0%	Plant life	30.0 years
Income tax rate	21.0%	General plant depreciation	7.0 years
Interest rate for dept financing	8.0%	Steam/electricity plant depreciation	20.0 years
Term for dept financing	10.0 years	On-stream percentage	90.0%

#### 3.3. Life cycle assessment

LCA is performed by following the international guideline presented by the International Organization for Standardization, which is constituted of four steps: 1) goal and scope definition, 2) inventory analysis, 3) impact analysis, and 4) interpretation. The system boundary is defined from the production of input materials and energies (cradle) to the production of hydrogen from the dehydrogenation section (gate), and the functional unit is defined as 1 kg of transported hydrogen. The environmental impacts were estimated using SimaPro V9.1 with the EcoInvent 3.6 database with ReCiPe 1.13V of mid-point from the hierarchical perspective. The environmental impacts were calculated based on the functional unit and allocated based on the mass production rate of the products.

## 4. Process design

### 4.1. Process synthesis

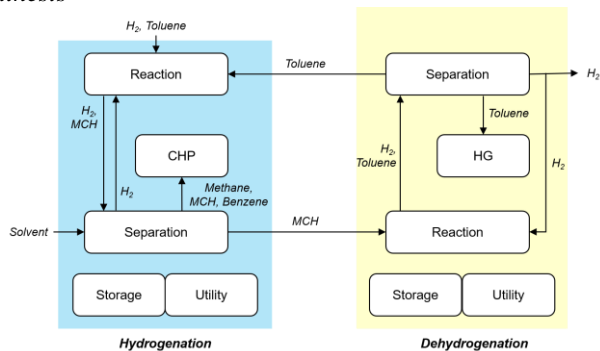


Fig 1. Block flow diagram of the proposed process.

In this study, we designed the integrated process using the Aspen Plus process simulator assuming that 400,000 tons of hydrogen are transported per annum. The integrated process consists of two sections such as hydrogenation and dehydrogenation sections which are 10,000 km apart. including five subsystems respectively. The hydrogenation section is constituted of 1) Reaction, 2) Separation, 3) Combined heat and power (CHP), 4) Storage, and 5) Utility. The dehydrogenation section has the same subsystems except for CHP. Instead of CHP, the dehydrogenation section has a heat generation (HG) subsystem because electricity is not produced in the dehydrogenation section.

#### 4.1.1. Hydrogenation section

Toluene supplemented to compensate for the toluene loss and transported from the dehydrogenation section is hydrogenated to MCH in the reaction subsystem. Since the hydrogenation reaction occurs in the excess hydrogen condition, hydrogen and MCH need to be separated to reduce hydrogen loss. Hydrogen can be separated by depressurization and cooling using the boiling point difference between hydrogen and MCH. Benzene produced as a byproduct from the hydrogenation reaction is easily accumulated in the proposed process. However, eliminating benzene is challenging because its physical property is similar to MCH. Therefore, the extraction column was introduced using N-formylmorpholine (NFM) as a solvent (Chen et al., 2007). The benzene-removed MCH is transported to the dehydrogenation section, and the used NFM is recycled to the extraction column after separating from the benzene using a distillation column.

The separated byproducts are utilized as energy source in the CHP subsystem. After using the heat required for the hydrogenation section, the remaining heat is used to produce the electricity which can be sold as a byproduct or used to produce hydrogen through water electrolysis (Sen et al., 2012). To reduce CO<sub>2</sub> emissions, carbon capture units are also installed in the CHP subsystem (Lee et al., 2021)

#### 4.1.2. Dehydrogenation section

The transported MCH is dehydrogenated in the reaction subsystem with some hydrogen. Since hydrogen (5-20 mol% of MCH) can improve the life period of the catalyst and the yield, the purified hydrogen from the separation subsystem is recycled to the dehydrogenation reactor. Further, the produced hydrogen is purified in the separation

subsystem by pressurization and PSA. Firstly, volatile toluene is liquefied at high pressure and low temperature to recover toluene, then the remaining organics are removed by PSA.

Byproducts are utilized as a heat source like the hydrogenation section to supply the heat required for the dehydrogenation section. However, since the dehydrogenation reaction is endothermic, it is difficult to meet the heat required by byproducts alone. Therefore, some of the produced hydrogen is also burned in the HG subsystem. The HG subsystem also has carbon capture units like the hydrogenation section.

#### 4.2. Heat integration

To lower the operating costs and to efficiently use energy, we have carried out heat integration using pinch analysis. As we assumed that the hydrogenation and dehydrogenation sections are far away, heat integration is conducted separately for each section, and the results are summed up and shown in Figure 2.

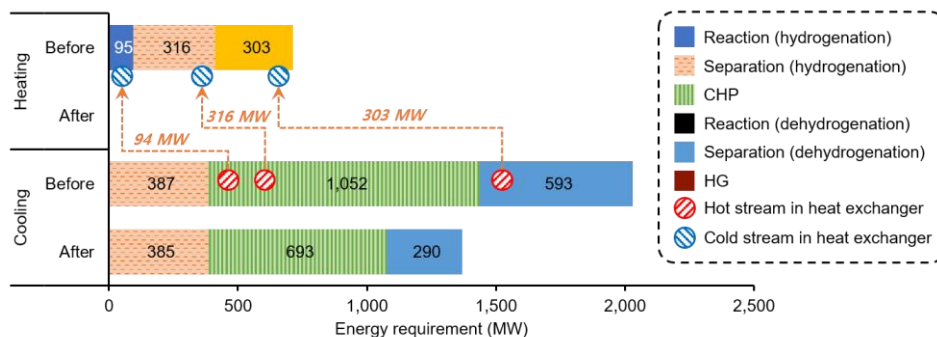


Fig. 2. The result of heat integration and major heat networks

The total heating requirement for the proposed process was 715 MW before heat integration. Most of the heating requirement (44%) was used for separation in the hydrogenation section. However, the heating requirement became zero after heat integration. The hydrogenation section has sufficient heat sources such as the heat from CHP and the hydrogenation reaction which is exothermic. The heat required for the dehydrogenation section is satisfied by burning produced hydrogen.

For the cooling requirement, most of the energy (52%) was required for the power cycle of the CHP subsystem. To liquify the steam used to generate electricity, a huge amount of cooling water was required. Furthermore, the large flow rate of the power cycle made it difficult to form heat networks completely. As a result, only 33% of the cooling requirement was decreased.



## 5. Result and Discussion

### 5.1. Techno-economic analysis

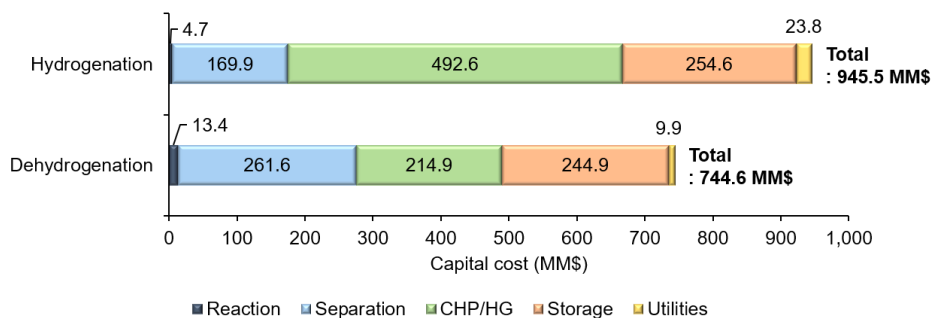


Fig. 3. The result of capital cost breakdown.

As shown in Fig. 3, the hydrogenation section requires a larger capital cost than the dehydrogenation section mainly because of the CHP subsystem. The generator in the power cycle is the most cost-intensive equipment due to the large flow rate. For the dehydrogenation section, the separation subsystem is the most cost-intensive subsystem. Multi-compressor was used to recover toluene, but the high flow rate of a mixture stream required a large-size compressor accounting for a major fraction of the capital cost in the separation subsystem.

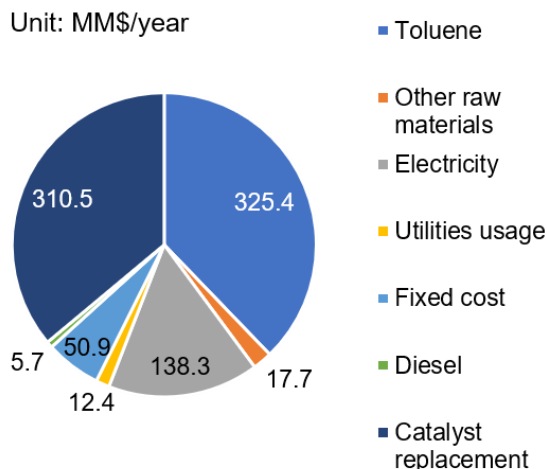


Fig. 4. Operating cost breakdown

As shown in Fig. 4, most of the operating costs come from toluene supplementation and catalyst replacement. Although toluene was recycled in the proposed process, some of the toluene was lost in the separation process of each section. Most of the catalyst replacement cost (89%) comes from the dehydrogenation catalyst because of the high price of platinum. The cost of electricity also accounted for a large portion. While electricity can be self-sufficient in the hydrogenation section because it generates electricity on its own in the CHP subsystem, all electricity required for the dehydrogenation section have to be purchased externally.

### 5.2. Minimum transport cost

We have calculated the minimum hydrogen transport cost which is the value when the total cost of the process and revenues including byproduct (electricity and captured CO<sub>2</sub>) incomes are identical. Fig. 5 shows the total cost and revenue, and the portion of each inventory.

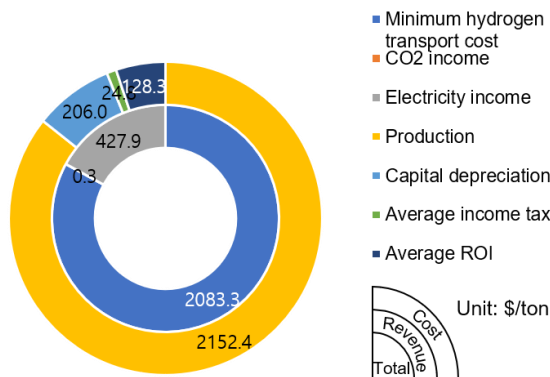


Fig. 5. Minimum transport cost of the process

The total cost of the proposed process is \$2,511/ton-H<sub>2</sub>, and the byproduct incomes are \$428/ton-H<sub>2</sub>. Therefore, the minimum hydrogen transport is found to be \$2083/ton-H<sub>2</sub>.

### 5.3. Life cycle assessment

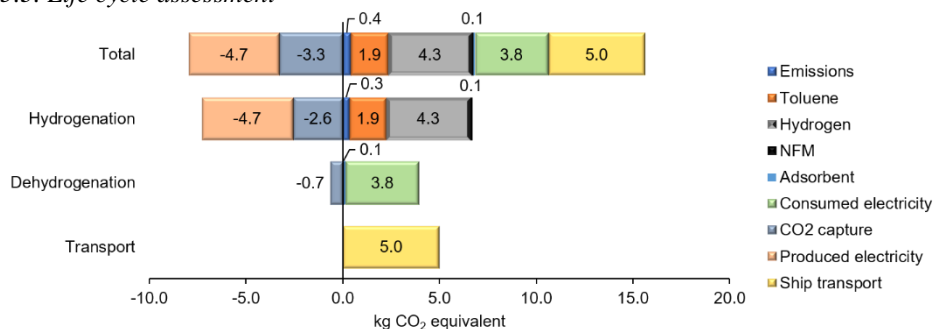


Fig. 6. The result of life-cycle assessment.

The life-cycle assessment was performed to determine the environmental impact of each section. Fig. 6 shows each section's contribution to climate change based on CO<sub>2</sub> emission. The hydrogenation section produced the most carbon emissions mainly because of the solar energy needed for hydrogen production. Though solar energy is sustainable, noticeable amounts of carbon emissions occur when constructing solar power plants. However, since produced electricity and captured from the CHP subsystem are environmentally beneficial, most of the carbon emissions in the hydrogenation section can be offset. The major cause of carbon emission is electricity used in the separation subsystem. The multi-compressor for pressurization consumes 98% of total electricity consumption.

Therefore, carbon emission would be further reduced if pressurized in a less energy-intensive manner.

## 6. Conclusion

In this study, we have proposed an integrated process for long-distance hydrogen transport using a toluene-MCH system and designed heat networks reducing 100% of heating requirement and 33% of cooling requirement. Based on the designed process and heat networks, TEA and LCA were conducted to study economic feasibility and sustainability of the proposed process. It was found that the hydrogenation section incurs higher capital costs than the dehydrogenation section mainly due to the power cycle of the CHP subsystem. It also found that toluene supplementation and catalyst replacement are the major cost contributors. From a sustainable point of view, although more carbon emissions occur in the hydrogenation section compared to the dehydrogenation section, most of the carbon emissions are offset because of the environment friendliness of the produced electricity and captured CO<sub>2</sub> from the CHP subsystem. The carbon emissions of the dehydrogenation section can be decreased by reducing electricity consumption.

## References

- R. Atsumi, K. Kobayashi, C. Xieli, T. Nanba, H. Matsumoto, K. Matsuda, and T. Tsujimura, 2020, Effects of steam on toluene hydrogenation over a Ni catalyst, *Applied Catalysis A: General*, 590, 117374.
- D. Chen, H. Ye, and H. Wu, 2007, Liquid–liquid equilibria of methylcyclohexane–benzene–N-formylmorpholine at several temperatures, *Fluid phase equilibria*, 255, 2, 115–120.
- R. E. Davis, N. J. Grundl, L. Tao, M. J. Bidy, E. C. Tan, G. T. Beckham, D. Humbird, D. N. Thompson, and M. S. Roni, 2018, Process Design and economics for the conversion of lignocellulosic biomass to hydrocarbon fuels and coproducts: 2018 biochemical design case update; biochemical deconstruction and conversion of biomass to fuels and products via integrated biorefinery pathways. NREL/TP-5100-71949.
- S. Lee, H. S. Kim, J. Park, B. M. Kang, C.-H. Cho, H. Lim, and W. Won, 2021, Scenario-Based Techno-Economic Analysis of Steam Methane Reforming Process for Hydrogen Production, *Applied Sciences*, 11, 13, 6021.
- M. Markiewicz, Y.-Q. Zhang, M. T. Empl, M. Lykaki, J. Thöming, P. Steinberg and S. Stolte. 2019, Hazard assessment of quinaldine-, alkylcarbazole-, benzene and toluene-based liquid organic hydrogen carrier (LOHCs) systems, *Energy & Environmental Science*, 12, 1, 366–383.
- M. Niermann, A. Beckendorff, M. Kaltschmitt, and K. Bonhoff, 2019, Liquid Organic Hydrogen Carrier (LOHC)–Assessment based on chemical and economic properties, *International Journal of Hydrogen Energy*, 44, 13, 6631–6654.
- Y. Okada, E. Sasaki, E. Watanabe, S. Hyodo, and H. Nishijima, 2006, Development of dehydrogenation catalyst for hydrogen generation in organic chemical hydride method, *International Journal of Hydrogen Energy*, 31, 10, 1348–1356.
- P. Preuster, Q. Fang, R. Peters, R. Deja, L. Blum, D. Stolten, and P. Wasserscheid, 2018, Solid oxide fuel cell operating on liquid organic hydrogen carrier-based hydrogen–making full use of heat integration potentials, *International journal of hydrogen energy*, 43, 3, 1758–1768.
- S. M. Sen, D. M. Alonso, S. G. Wettstein, E. I. Gürbüz, C. A. Henao, J. A. Dumesic, and C. T. Maravelias, 2012, A sulfuric acid management strategy for the production of liquid hydrocarbon fuels via catalytic conversion of biomass-derived levulinic acid, *Energy & Environmental Science*, 5, 12, 9690–9697.
- R. Singh, M. Singh, and S. Gautam, 2021, Hydrogen economy, energy, and liquid organic carriers for its mobility, *Materials Today: Proceedings*, 46, 5420–5427.
- D. Teichmann, W. Arlt, P. Wasserscheid, and R. Freymann, 2011, A future energy supply based on Liquid Organic Hydrogen Carriers (LOHC), *Energy & Environmental Science*, 4, 8, 2767–2773.

# Integration of metabolic models in biorefinery designs using superstructure optimisation

Lucas Van der Hauwaert<sup>a</sup>; Alberte Regueira<sup>a,b</sup>, Miguel Mauricio-Iglesias<sup>a</sup>

<sup>a</sup> *CRETUS, Department of Chemical Engineering, Universidade de Santiago de Compostela, 15782, Santiago de Compostela, Spain.*

<sup>b</sup> *Center for Microbial Ecology and Technology (CMET), Ghent University, B-9000, Ghent, Belgium.*

## Abstract

In this work the integration of metabolic (structured) models, which describe the entire metabolic network of microorganisms, into a superstructure optimisation framework is presented. These models are of particular interest because they can predict the outcome of a fermentation with different substrates, and without the need of experimental data. In this contribution, a workflow for 2 types of structured models is described: i) genome-scale metabolic models (GEMs) and ii) community models where product yields can be influenced by environmental factors (e.g., pH). To showcase this methodology a simple case study is presented of a superstructure optimisation problem. With this case study it is demonstrated that the described workflow can aid the design of novel biorefineries, by screening potential cultures using metabolic models without the need of experimental data.

**Keywords:** Structured models, biorefinery design, superstructure optimisation, surrogate models, Genome-scale metabolic models

## 1. Introduction

A current hurdle towards the implementation of biorefineries are the various process alternatives which lead to a multitude of possible refinery designs. To aid the design of biorefinery networks, superstructure optimisation tools, like O2V have been created (Gargalo et al., 2022). A superstructure represents all the possible links between process units (and their possible substitutes) formed by flowchart and logic rules, where alternative topologies of processing networks can be visualized. Optimisation techniques can be applied to this superstructure to generate an optimal processing network according to a desired objective (e.g., maximizing profit) (Bertran et al., 2017).

In a superstructure the unit processes are represented by mathematical models, describing the flow and/or transformation of mass. For biorefineries, key unit processes are bioreactors, responsible for the chemical transformation of feedstocks. Metabolic models, simulating intra- and extra-cellular spaces, are commonly used for simulating biotransformations. These models can predict the microbial products from different substrates and can be linked to several inputs, or outputs of other reactors, thereby increasing the amount of new possible network topologies. Additionally, the creation of these models is not dependant on data from time consuming fermentation experiments.

A common type of metabolic model are genome-scale metabolic models (GEMs). A GEM comprehensively reconstructs the organism's entire metabolic network using its genome (Lieven et al., 2020). Another type of metabolic model are community models, such as the bioenergetic model of Regueira et al. (2020). These models are of interest because of their ability to describe open and co-culture fermentations with different

substrates and at different operational conditions (e.g., different pH values). Additionally open culture fermentations are interesting to include because they have economic and operational advantages, such as avoiding sterilisation and being able to handle complex substrates (Kleerebezem et al., 2015).

Unfortunately, structured bioreactor models are often too complex to feasibly incorporate into a superstructure and thus are often not considered. In other words, the integration of GEMs or community models in a superstructure problem is not strait forward. For this reason, the aim of this contribution is to create a workflow to incorporate these models in a superstructure optimization problem. To showcase this workflow a simple superstructure optimization problem is presented to produce propionate and acetate where GEMs and community models are incorporated as surrogate models.

## 2. Methodology

The workflow for the integration of two types of models into a superstructure optimisation problem, by creating surrogate models, will be discussed. The first type of surrogate model is derived from GEMs, where the yield of a product is only dependant on the selected substrate. The second are derived from community models where the yield is dependent on external influences (e.g., pH) other than the type of substrate.

### 2.1. Workflow

A flowchart to incorporate the 2 types of structured models into a superstructure optimization problem is shown in Figure 1. The goal of this flowchart is to find a mathematical expression (i.e., a reaction equation) that can be used in a superstructure optimisation problem, starting from a community model or a GEM. These reaction equations represent the chemical transformation of mass, hence linking the inputs to the outputs (i.e., a yield) of the bioreactors.

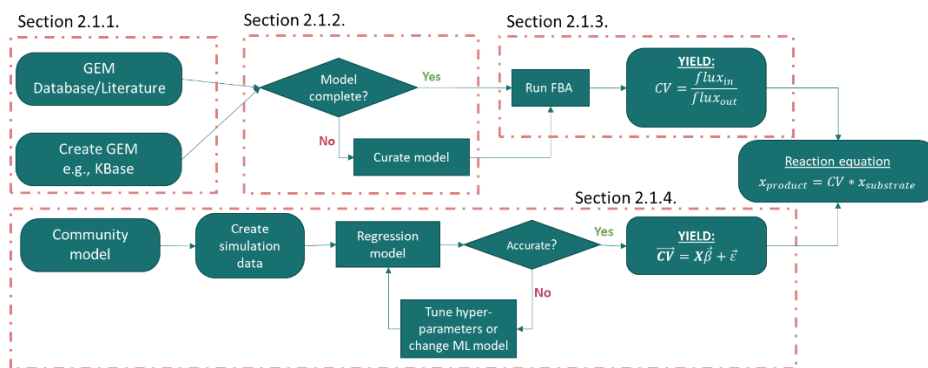


Figure 1. Flowsheet to make reaction equations (i.e., surrogate models) from structured models for superstructure optimisation problems

#### 2.1.1. Making/finding GEMs

To generate a GEM of a particular microorganism(s), software tools and genetic data available on platforms such as KBase and MetaNetX can be used, which enable the creation of metabolic network reconstructions (Arkin et al. 2018). Alternatively, already existing GEMs can be downloaded from databases such as BiGG Models, BioModels, or published work in literature.

### *2.1.2. Validating GEMs*

Many GEMs often lack the required quality to be used reliably and need to undergo a quality assurance and further manual curation. Here the software tool MEMOTE which is an open-source software, representing a unified approach to ensure the formally correct definition of GEMs, was used to check the quality (Lieven et al., 2020). The criteria used to assure the quality of the GEMs are the following: 1) the percentage of metabolic reactions where the chemical elements are correctly balanced; 2) the percentage of metabolic reactions that have a correct charge balance; 3) the percentage of dead-end metabolites (i.e. internal metabolites which do not participate in the internal metabolism); and 4) a leak test, which tests if biomass is created without any substrate. If any of the criteria are not met, the model must be manually curated. For this curation the reader is referred to the work of Thiele & Palsson (2010).

The percentage of correct mass and charge balanced reactions should obviously be as high as possible. In this work a threshold of 95% of correctly mass and charge balanced reactions is chosen as a selection threshold for GEMs. This threshold was taken to minimise the time needed to curate models, as manually correcting the reactions is a very time-consuming procedure. The percentage of dead-end metabolites should be kept as low as possible but are not *per se* indicators of low-quality models. However, a large proportion of dead-end metabolites (e.g., >50%) may indicate problems in reconstruction that need solving.

All non-balanced reactions and the dead-end metabolites need to be checked, and manually corrected if they take part in essential reactions (i.e., essential for the formation of a product). Finally, models producing biomass without substrates indicate an incorrect model formulation and need to be checked thoroughly.

### *2.1.3. Incorporating GEMs into the superstructure problem*

For GEMs, yields are found by solving a flux balance analysis (FBA). An FBA returns the flux of all the metabolites flowing through the metabolic network at steady state (Orth et al., 2010). Using the flux data from the model solution, yields are determined as the ratio of the produced extracellular metabolites and the substrate.

### *2.1.4. Incorporating community models into the superstructure problem*

For community models where the yield is dependent on environmental influences, machine learning (ML) techniques can be used to find simplified algebraic relationships between the yield and the operational variables (e.g., varying pH) (Vollmer et al., 2022).

To create the data needed to calibrate such models, simulations need to be run under varying operational conditions, hence covering the input space. With this data, polynomial equations describing the yield are fitted using lasso regression. This regression technique is used because it effectively avoids overfitting and has a simple algebraic formulation which can easily be handled by superstructure optimization algorithms (such as BARON). To validate the equations, parity plots are made to compare the observed and predicted values of the yields. The normalised mean standard error (NMSE) is also used as a validation metric to make sure the surrogate models reproduce the original model sufficiently. If an insufficient quality of the surrogate model is attained, calibration parameters should be tuned (e.g., the hyperparameter  $\alpha$  determining the degree of regularisation) or the number of polynomials can be changed. If the obtained model still does not sufficiently reproduce the original model, a different ML technique should be used (Vollmer et al., 2022).

## 2.2. Superstructure model

The objective, the flow of mass and logic rules (i.e., the selection of a path in the process network) of the superstructure are represented as mathematical equations (Eq. 1 to Eq. 6):

$$\begin{array}{lcl} \max f(x, y) & (1) & \left| \quad g(x, y) \geq 0 \quad (2) \quad \right| \quad h(x, y) = 0 \quad (3) \\ x \in X & (4) & \left| \quad x^{LO} \leq x \leq x^{UP} \quad (5) \quad \right| \quad y \in \{0; 1\}^n \quad (6) \end{array}$$

where  $f$  is the objective function (e.g., maximising profit),  $x$  is the vector of continuous variables representing operating variables (e.g., the flow of mass) defined by their lower and upper bounds  $x^{LO}$  and  $x^{UP}$  in a continuous feasible region  $X$ ,  $y$  is the vector of binary variables representing the selection of a unit-process,  $g$  and  $h$  are the vectors of inequality and equality constraints, representing the process models and process specifications of a unit process. The equations for the process models are represented by generic process intervals. Every interval in the superstructure is modelled with the same set of generic equations, representing a sequence of processing tasks, namely mixing, reaction, waste removal and product separation, as well as utility consumption (Bertran et al. 2017). The reaction equations are mass conversions between inputs and outputs to a reaction step and are represented by the surrogate models of the structured models.

## 2.3. Case study

The goal of this simplified case study is to find out which type of microbial culture(s), producing propionate and acetate, can generate the largest profit from 2 alternative feedstocks being fructose or glucose. Here 5 different *Propionibacterium* species (*P. acidipropionici*, *P. acnes*, *P. avidum*, *P. freudenreichii*, and *P. propionicum*) and an open fermentation are considered (Figure 2). The surrogate models for the pure cultures are obtained from GEMs from the works of McCubbin et al. (2020), while the open fermentation surrogate model was derived from the bioenergetic model of Regueira et al. (2020). All the cultures use either glucose or fructose as substrate, except for the bioenergetic model that can only consider glucose. In this example, the profit is defined as the revenue generated through sales of the products (i.e., acetate and propionate), minus the operating cost being the cost of raw material (i.e., glucose and fructose). The wholesale prices of the raw materials were found on SELINA WAMUCII (<https://www.selinawamucii.com>) while the wholesale prices for the products were found on ECHEMI (<https://www.echemi.com>). Furthermore, logic constraints were defined to ensure the selection of only one substrate and one type of culture. To further simplify the case study, no utility consumption, waste removal or mixing was considered, and product separation is assumed to be perfect after all reactors.

The formulation of the superstructure was created using PYOMO (version 6.4.2), in the Python environment (version 3.10.1). The formulation of this problem results in a multiple integer non-linear program (MINLP) and was solved using outer-approximation and generalized Benders decomposition (DICOPT) which is accessible through GAMS (GAMS Development Corporation, 2021).

## 3. Results and Discussion

### 3.1. Validating the GEMs

The first step after finding the GEMs from literature, is to assess their quality. The models have been previously curated by McCubbin et al. (2020) and appear to have a good quality where at least 95 % of all reaction are mass and charge balanced. The few metabolic reactions which were not mass or charged balanced were found to not affect essential

reactions producing the products. Furthermore, the percentage of dead-end metabolites are all low (< 5%) and no biomass was created from any of the models without substrate.

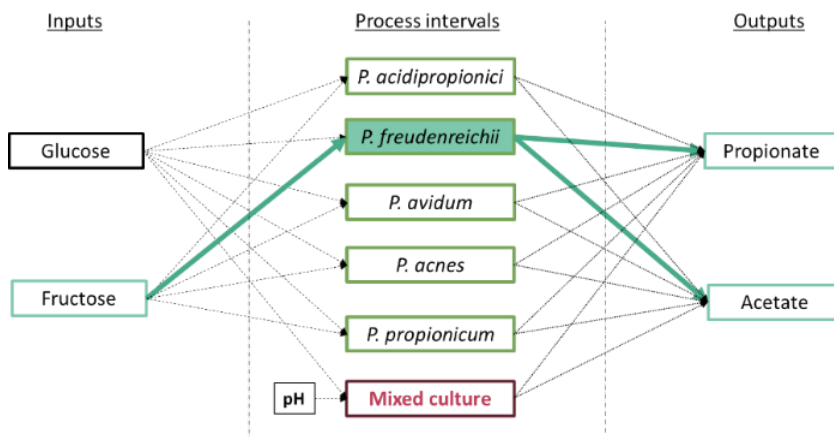


Figure 2. Superstructure representation of the case study. □: process intervals derived from GEMs; □: process intervals derived from community models; → and ■: optimal processing route.

### 3.2. Creating the surrogate reaction models

With the quality of the GEMs assured, the yields can be found (Table 1) by taking ratio of the substrate and product fluxes obtained by solving the FBA for each GEM.

Table 1. Yields (g-carbon<sub>product</sub>/g-carbon<sub>substrate</sub>) of the different substrates (G: Glucose, F: Fructose) and products (P: Propionate, A: Acetate) resulting from the flux balance analysis

Species/yield	P/G	A/G	P/F	A/F
<i>P. acidipropionici</i>	0.32	0.11	0.32	0.11
<i>P. acnes</i>	0.46	0.07	0.48	0.03
<i>P. avidum</i>	0.36	0.05	0.33	0.11
<i>P. freudenreichii</i>	0.34	0.13	0.42	0.16
<i>P. propionicum</i>	0.29	0.15	0.27	0.17

To find the reaction equations for the open fermentation, the bioenergetic model was run to create datapoints of the fermentation outcome at different pH, which were sampled uniformly between 4 and 8.5. Using lasso regression, 5<sup>th</sup> order polynomial equations were obtained to express the yields, with pH being the sole variable of the equations. The NMSE values of the fitted equations were low (0.04 g-carbon<sub>acetate</sub>/g-carbon<sub>glucose</sub> for the yield of acetate and 0.05 g-carbon<sub>propionate</sub>/g-carbon<sub>glucose</sub> for propionate) and the parity plots of the predicted and observed yields closely resemble each other. With these observations it can be concluded that the used regression technique sufficiently reproduces the bioenergetic model and can thus be used in the superstructure model.

### 3.3. Solution of the superstructure

The solution of the superstructure (Figure 2) resulted in fructose being the preferred feedstock with *P. freudenreichii* being the optimal culture to produce acetate and propionate. In this process network 100 kg-C of fructose is transformed into 42.45 kg-C of propionate and 15.56 kg-C of acetate resulting in a profit of 32.39 €.



#### 4. Conclusion

A workflow to incorporate structured models into superstructure optimization problems was developed and successfully showcased using a simplified/conceptual case study. With this workflow more intricate superstructure problems can be defined (starting from different available feedstocks) which can result in the design of novel biorefineries. More specifically this workflow allows for a screening of different microbial cultures in biorefinery designs. This includes open fermentations which are key to handling complex substrates in a biorefinery design.

#### References

- Arkin, A. P., Cottingham, R. W., Henry, C. S., Harris, N. L., Stevens, R. L., Maslov, S., Dehal, P., Ware, D., ... Yu, D. (2018). KBase: The United States department of energy systems biology knowledgebase. In *Nature Biotechnology* (Vol. 36, Issue 7, pp. 566–569). Nature Publishing Group. <https://doi.org/10.1038/nbt.4163>
- Bertran, M. O., Frauzem, R., Sanchez-Arcilla, A. S., Zhang, L., Woodley, J. M., & Gani, R. (2017). A generic methodology for processing route synthesis and design based on superstructure optimization. *Computers and Chemical Engineering*, *106*, 892–910. <https://doi.org/10.1016/j.compchemeng.2017.01.030>
- GAMS Development Corporation. General Algebraic Modeling System (GAMS) Release 36.1.0, Fairfax, VA, USA, 2021
- Gargalo, C. L., Rapazzo, J., Carvalho, A., & Gernaey, K. v. (2022). Optimal Conversion of Organic Wastes to Value-Added Products: Toward a Sustainable Integrated Biorefinery in Denmark. *Frontiers in Chemical Engineering*, *4*, 54. <https://doi.org/10.3389/fceng.2022.837105>
- Kleerebezem, R., Joosse, B., Rozendal, R., & van Loosdrecht, M. C. M. (2015). Anaerobic digestion without biogas? *Reviews in Environmental Science and Bio/Technology*, *14*(4), 787–801. <https://doi.org/10.1007/s11157-015-9374-6>
- Lieven, C., Beber, M. E., Olivier, B. G., Bergmann, F. T., Ataman, M., Babaci, P., Bartell, J. A., Blank, L. M., ... Zhang, C. (2020). MEMOTE for standardized genome-scale metabolic model testing. *Nature Biotechnology*, *38*(3), 272–276. <https://doi.org/10.1038/s41587-020-0446-y>
- McCubbin, T., Gonzalez-Garcia, R. A., Palfreyman, R. W., Stowers, C., Nielsen, L. K., & Marcellin, E. (2020). A pan-genome guided metabolic network reconstruction of five propionibacterium species reveals extensive metabolic diversity. *Genes*, *11*(10), 1–26. <https://doi.org/10.3390/genes11101115>
- Orth, J. D., Thiele, I., & Palsson, B. O. (2010). What is flux balance analysis? In *Nature Biotechnology* (Vol. 28, Issue 3, pp. 245–248). <https://doi.org/10.1038/nbt.1614>
- Regueira, A., Lema, J. M., Carballa, M., & Mauricio-Iglesias, M. (2020). Metabolic modeling for predicting VFA production from protein-rich substrates by mixed-culture fermentation. *Biotechnology and Bioengineering*, *117*(1), 73–84. <https://doi.org/10.1002/bit.27177>
- Thiele, I., & Palsson, B. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols*, *5*(1), 93–121. <https://doi.org/10.1038/nprot.2009.203>
- Vollmer, N. I., Al, R., Gernaey, K. v., & Sin, G. (2022). Synergistic optimization framework for the process synthesis and design of biorefineries. *Frontiers of Chemical Science and Engineering*, *16*(2), 251–273. <https://doi.org/10.1007/s11705-021-2071-9>

# Environmental Impacts of Bio-derived Silicon: Uncertainty in the Benefit of Industrial Transition

Ethan Errington,<sup>a</sup> Miao Guo,<sup>b</sup> Jerry Heng,<sup>b</sup>

<sup>a</sup> *Department of Chemical Engineering, Imperial College London, London, SW7 2AZ, UK*

<sup>b</sup> *Department of Engineering, King's College London, London, WC2R 2LS, UK  
miao.guo@kcl.ac.uk (MG); jerry.heng@imperial.ac.uk(JH)*

## Abstract

Millions of tonnes of silicon are produced each year as part of a global supply chain supporting the manufacture of steel, automobiles, electronics and silicones. However, the supply chain does have its challenges – particularly the high energy intensity of silicon purification methods, and the lack of global supply chain diversity. Consequently, this study investigates the benefit of producing high purity silicon metals (>98 wt% Si) from agricultural biomass wastes from an environmental perspective using Life Cycle Assessment methodology. Taking rice husk as a case study, findings show a reduction in carbon emissions is achievable for co-recovery of silicon with bio-energy. The robustness of these findings is also addressed through uncertainty modelling, which provides further confidence in the major findings. Future agricultural sources of biomass-derived silicon have also been projected for staple crops including corn, sugarcane and wheat.

**Keywords:** silicon, rice husk, life cycle assessment, global warming potential, uncertainty.

## 1. Introduction

8.5 M-tonnes of silicon is produced annually to meet demand for iron alloys in the steel industry, aluminum alloys in the automotive industry, semiconducting materials for electronics products, and a range of wider industrial products such as silicones, silicon carbide and silanes (USGS, 2021). Consequently, a large silicon supply chain (SSC) exists to meet global demand by recovering silicon from mineral deposits of silica (SiO<sub>2</sub>). However, the SSC has its challenges. First, mineral excavation introduces a range of environmental impacts. Second, commercial methods for silicon refinement are energy intensive (Maurits, 2014), making sectoral decarbonization challenging. Third, the SSC lacks diversity as ~70 wt% of global production comes from China alone (USGS, 2021). For these reasons, gaps exist for the development of the economic and environmental sustainability of the SSC. One way this could be achieved is by using biomass-derived silica (bSiO<sub>2</sub>) as a substitute for mineral-derived silica (mSiO<sub>2</sub>). This appeals as it: 1) avoids mineral excavation, 2) provides a pathway for decarbonizing the energy used in silicon purification (via bio-energy co-recovery), and 3) promotes SSC diversity.

Yet, the benefit of using bSiO<sub>2</sub> to produce bio-derived silicon (bSi) is still unknown. Consequently, this work provides the first evaluation of that benefit from the perspective of carbon emissions reductions. Reflecting the SSC, three purity grades of silicon are considered – metallurgical grade (MG-Si), solar grade (SG-Si) and electronic grade (EG-Si). Rice husk-derived bSiO<sub>2</sub> (RH-bSiO<sub>2</sub>) is used as a case study due to its production volume (100s of Mtonnes per year), the availability of relevant impact models

(Errington et al., 2022 and 2023), and because it may be considered an agricultural waste product.

## 2. Overview of the Silicon Supply Chain

Global demand for silicon comes from the functionality that it provides to wider products. Consequently, the SSC involves intermediate products concerning mainly industrial actors. Thus, silicon materials are often categorised based on their purity by weight (wt%). This includes: ferrosilicon (>50 wt%), MG-Si (>98 wt%), SG-Si (>99.9999 wt%), and EG-Si (>99.9999999 wt%) (USGS, 2022; Chigonodo, 2018) - see Figure 1.

Of the 8.5 Mtonnes produced, approximately 70% (5.95 Mtonne) is used in the production of ferrosilicons, with the remainder (2.55 Mtonne) being used for silicon metals (USGS, 2022). Recent statistics suggest as much as 21% of silicon metal use (0.54 Mtonne) may be attributed to the combined production of polycrystalline silicon (p-Si), EG-Si and SG-Si – with the remainder (2.01 Mtonne) finding use as MG-Si in aluminium alloys, silicones and other products (USGS, 2019).

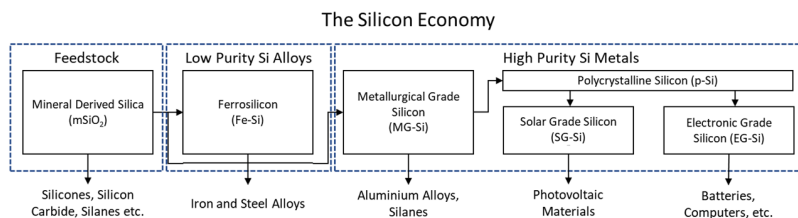


Figure 1: A summary of the major categories of silicon as considered within this work.

## 3. Agricultural Biomass as a Source of Silicon

Many investigations have reported over the past 40 years on the recovery of silicon dioxide from agricultural biomass. This has included a strong focus on rice husk (RH) – a natural waste from Rice Grain Agriculture (RGA) – due to its high silica content. Accordingly, the present authors have recently reported a model for assessing the environmental impact of RH-bSiO<sub>2</sub> in a process including RH combustion for co-recovery and use of RH bio-energy (Errington et al., 2022 and 2023). Having found a potential environmental benefit for such a method, and given the high purity (>90%) of RH-bSiO<sub>2</sub> typically produced, it seems a logical extension that bSi may be produced by using RH-bSiO<sub>2</sub> as a feedstock by extending the method as is summarised in Figure 2.

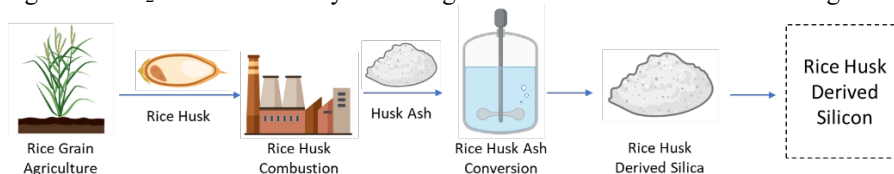


Figure 2: The process for recovery of silicon using rice husk biomass as a resource.

## 4. Method

This work aims to evaluate the benefit of producing MG-Si, EG-Si and SG-Si from bSiO<sub>2</sub>. The “benefit” has been defined based on greenhouse gas emissions reductions associated with producing each grade of Si from bSiO<sub>2</sub> instead of mSiO<sub>2</sub>. This has been done using a Life Cycle Assessment (LCA) methodology to: 1) assess the environmental impact of individual mSi and bSi silicon metal products, 2) calculate the benefit of bSi by

comparison of mSi and bSi material impact predictions, and 3) quantify the uncertainty associated with each benefit prediction. Further information on methods is given below.

#### 4.1. Impact Assessment: Life Cycle Inventories and Impact Characterisation Factors

A summary of the scope and relationship of Life Cycle Inventories (LCIs) for silica and silicon products considered is shown in Figure 3. LCIs for each silicon grade produced by existing industrial methods were taken from the Ecoinvent database (v3.6, Allocation at Point-of-Substitution basis). Inventories for RH-bSiO<sub>2</sub> were adapted from models provided by Errington et al. (2023) - in which RH combustion is followed by the recovery of RH-bSiO<sub>2</sub> via the “wet method”. All LCIs were modelled on a global level.

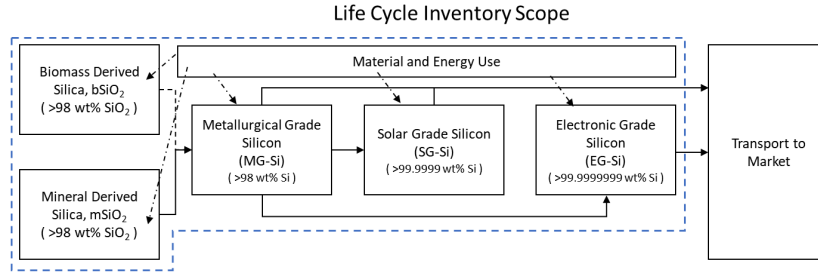


Figure 3: Life cycle scope associated with LCA assessments carried out for the comparison of mineral and biomass derived silicon in this work. Values for energy and material demands are taken from relevant inventories for each metal grade within the Ecoinvent Life Cycle Inventory Database.

The environmental impact,  $EI$  (impact per kg<sub>Si</sub>), of each product was calculated as the sum of impacts from all activities over the entire product life cycle as described by Errington *et al* (2022). Assessment was based on the Global Warming Potential (GWP) of the ReCiPe (H) method (Huijbregts *et al.*, 2017) using calculation cut-off of 0.1%.

##### 4.1.1. Calculation of the Benefit of bio-derived silicon (bSi)

The benefit of bSi production,  $B^{bSi}$  (GWP/kg) was calculated by region based on the assumption that bio-energy co-recovered during bSiO<sub>2</sub> production,  $\hat{Q}^b$  (MJ/kg<sub>Si</sub>), was used to substitute for local electricity in a region with an associated environmental impact,  $EI^e$  (GWP/MJ) – see Equations 1 and 2. Benefits were also converted to a decarbonization extent,  $D^{bSi}$  (%), using Equation 3.

$$B^{bSi} = EI^{bSi} - \hat{Q}^b EI^e \quad (1)$$

$$\hat{Q}^b = (\eta^{BC} / \eta^{SR}) \times [(\hat{Q}^{LHV} (1 - x^{H_2O}) - \lambda x^{H_2O}) / x^{SiO_2} (1 - x^{H_2O})] \quad (2)$$

$$D^{Si} = B^{bSi} / EI^{bSi} \quad (3)$$

Where  $\eta^{RHC}$  is the electrical efficiency of bioenergy recovered during combustion (%),  $\eta^{SR}$  is the efficiency associated with recovering usable bSi from a dry biomass feedstock (%),  $\hat{Q}^{LHV}$  is the lower heating value of the biomass used (MJ/kg<sub>feedstock</sub>),  $x^{H_2O}$  is the moisture content of the biomass feedstock prior to combustion (wt%, wet basis), and  $x^{SiO_2}$  is the silica content of the biomass used (wt%, dry basis). LCIs for the impact of medium-voltage regional electricity were based on the ecoinvent LCI database (Wernet *et al.*, 2016). Regions Europe, USA, Japan and China were considered as case studies for  $EI^e$  due to the extent of contribution that they currently supply to the SSC.

#### 4.1.2. Quantification of Uncertainty in Impact Predictions

The sensitivity of impact evaluations to uncertainty in LCIs (arising from data quality) was quantified by the pedigree matrix approach - in which, the quality of inventory data was rated against five criteria: reliability, completeness, and temporal, geographic and technological representativeness. Ratings were quantified as log-normal probability distributions based on the method of Frischknecht et al. (2004). Monte-Carlo simulations were used to estimate prediction uncertainty (sample size = 10,000).

## 5. Results and Discussion

### 5.1. Impact Predictions and Implications for Industry Decarbonisation by Region

Predicted environmental impacts of different purity grades of bSi and mSi are shown alongside associated decarbonization extents for Europe, USA, Japan and China in Figure 4. Associated estimates for the regional availability of rice husk and regional demands for silicon metal are shown in Table 1.

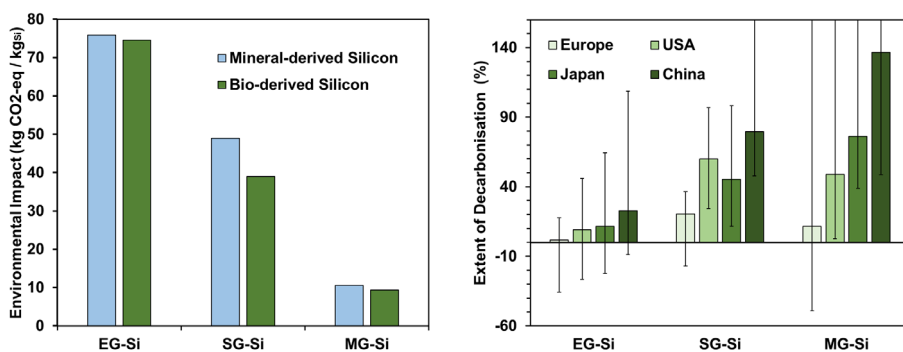


Figure 4: Impact predictions for producing 1 kg of mSi and bSi (left) and the decarbonization expected from a transition to production of bSi (right). Error bars represent a 99% prediction interval.

Table 1: Feasible annual supply of bSi and associated decarbonisation extent

Statistic	Europe	Japan	USA	China	World
RH Production <sup>a</sup> (Mtonne)	0.81	1.94	2.06	42.72	151.35
Available RH-bSiO <sub>2</sub> <sup>b</sup> (Mtonne)	0.10	0.23	0.25	5.09	18.02
Benefit of bMG-Si (kg <sub>CO<sub>2</sub></sub> /kg <sub>Si</sub> )	-1.2	-8.1	-14.4	-5.2	-1.2
Benefit of bEG-Si (kg <sub>CO<sub>2</sub></sub> /kg <sub>Si</sub> )	-10.0	-17.7	-24.9	-14.4	-10.0
Benefit of bSG-Si (kg <sub>CO<sub>2</sub></sub> /kg <sub>Si</sub> )	-1.3	-8.6	-15.4	-5.5	-1.3
Silicon Metal Production <sup>c</sup> (Mtonne)	0.37	N.G.	0.09	1.69	2.40
Possible bMG-Si Supply <sup>d</sup> (Mtonne)	0.04	0.09	0.09	1.88	6.67
Possible bSG-Si Supply <sup>d</sup> (Mtonne)	0.03	0.08	0.09	1.77	6.26
Possible bEG-Si Supply <sup>d</sup> (Mtonne)	0.03	0.08	0.08	1.67	5.91

<sup>a</sup> based on paddy production reported by the United Nations Food and Agriculture Organization (UNFAO, 2022) assuming 20 wt% of paddy is RH, <sup>b</sup> based on a mass ratio of 8.4 kg<sub>RH</sub>/kg<sub>SiO<sub>2</sub></sub> reported by Errington et al. (2023), <sup>c</sup> based on values from U.S. Geological Survey (USGS, 2022), <sup>d</sup> based on silica-to-silicon ratios present for each silicon purity grade in the ecoInvent LCI database (Wernet et al. 2016), N.G.: not given

Results in Figure 4 indicate that significant amounts of decarbonization may be expected to occur in all regions where bSi is produced as an alternative to mSi. However, the extent of decarbonization as well as uncertainty in the extent of decarbonization appear to be sensitive to both region and the silicon grade being produced. These correlations are understandable for two main reasons.

Firstly, correlations between the extent of decarbonization and the region considered are driven by differences in the GWP of electrical energy available in each region. This is due to differences in energy generation fuel blends – as shown by Errington *et al.* (2023).

Secondly, correlations between the magnitude of extent of decarbonization and the silicon grade considered (as well as associated uncertainty in the extent of decarbonization) are attributable to differences in the energy demand associated with each silicon grade considered. Particularly, the demand for energy rises significantly more quickly than the demand for bSiO<sub>2</sub> when increasing the purity of a silicon metal. Consequently, the fraction of bio-energy available to offset regional energy use is significantly diminished for EG-Si and SG-Si when compared to MG-Si – causing the magnitude and certainty in decarbonization to also reduce.

Finally, results in Table 1 imply a reduction in carbon emissions of 25 Mtonne<sub>CO<sub>2</sub>-eq</sub> for silicon metal production in the USA and China (both of which show a reasonable extent of certainty in decarbonization in Figure 4). This is equivalent to approximately 0.05% of global greenhouse gases emitted in 2019 (World Resources Institute, 2022). However, the assumption of complete utilization of global RH waste is an unrealistic scenario. Thus, alternate sources of agricultural bSiO<sub>2</sub> are detailed below.

### 5.2. Alternatives to RH for the recovery of bSiO<sub>2</sub> and bSi from biomass at scale

A summary of estimates for the availability of bSiO<sub>2</sub> associated with five possible alternatives to RH for bSiO<sub>2</sub> and bSi is shown in Table 2. These are: rice straw, wheat straw, corn stalk, corn cob and sugar cane straw. Results in Table 2 indicate 88 Mtonne of bSiO<sub>2</sub> may be practically recoverable from a combination of the agricultural wastes of rice, wheat, corn and sugar globally. This equates to more than 10 times the annual demand for silicon metal in the SSC (assuming 3 kg<sub>SiO<sub>2</sub></sub>/kg<sub>Si</sub> based on the LCI of EG-Si), which may act as a more reasonable scenario for recovery of bSiO<sub>2</sub> from waste biomasses.

Table 2: Amounts of bSiO<sub>2</sub> available annually in major agricultural biomass wastes.

Statistics	Rice Husk	Rice Straw	Wheat Straw	Corn Stalk	Corn Cob	Sugar Cane Straw
Global Production <sup>a</sup> (MT/yr)	757	757	761	1162	1162	1870
Product-to-Biomass Ratio <sup>b</sup> (wt%)	20	20	50	50	38	20
Silica Content <sup>b</sup> (wt% dry basis)	20	10	6	3	2	2.4
High Heat Value <sup>c</sup> (MJ/kg)	14	19	19	18	18	14
Moisture Content <sup>b</sup> (wt%)	12	12	7.5	16	16	12
Theoretical bSiO <sub>2</sub> <sup>d</sup> (Mtonne)	27	13	21	15	7	4
Biomass Use (kg <sub>RH</sub> /kg <sub>SiO<sub>2</sub></sub> , wet basis)	8	23	37	81	121	170
Bioenergy Recovery <sup>d</sup> (MJkg <sub>SiO<sub>2</sub></sub> )	22	86	144	269	403	298

<sup>a</sup> based on 2020 crop production reported by the United National Food and Agriculture Organisation (UNFAO, 2022), <sup>b</sup> assumed values, <sup>c</sup> assumed values, <sup>d</sup> based on a 69% mass recovery efficiency as reported by Errington *et al.* (2023) for RH and assuming a conservative recovery efficiency of 49% (i.e.  $\eta^{SR} = 0.49$ ) for other crops.

### 5.3. Financial Incentives for using bSiO<sub>2</sub> as an alternative to mSiO<sub>2</sub>

Assuming a cost of 100 and 250 \$ tonne<sup>-1</sup> for RH and mSiO<sub>2</sub> respectively, the effective cost of RH-bSiO<sub>2</sub> (800 \$ tonne<sup>-1</sup>, scaled based on “Biomass Use” in Table 2) is larger than that of mSiO<sub>2</sub>. However, the value of co-recovered RH-bSiO<sub>2</sub> bioenergy (22 MJ kg<sub>bSiO<sub>2</sub></sub><sup>-1</sup>, Table 2), equates to 5148 \$ tonne<sub>RH-bSiO<sub>2</sub></sub><sup>-1</sup> - assuming a 0.065 \$ MJ<sup>-1</sup> grid energy cost, EU Commission (2023). Thus, estimates suggest that the cost of RH-bSiO<sub>2</sub> is compensated for in a grid energy substitution scenario. Future work should better account for bioenergy co-recovery costs and different energy substitution scenarios.

## 6. Conclusion

In conclusion, this work aimed to assess the benefit of using biomass-derived silica as an alternative feedstock in production of silicon metals. This has been achieved using rice husk as a biomass case study and led to three main findings. Firstly, results show that a benefit is expected for all silicon grades considered. However, the extent of decarbonization is sensitive to the grade of silicon produced and the region in which it is produced. This is due to differences in the energy required for each grade of silicon, fuel sources used to produce regional energy, and the fraction of energy replaced with co-recovered bioenergy. Secondly, results show a transition to biomass-derived silicon could reduce global emissions by ~25 Mtonne per year (i.e. ~0.05% of global annual carbon emissions), which is driven by the yearly volume of metallurgical grade silicon (2 Mtonne). Finally, preliminary estimates show that as much as 10 times the current demand for silicon metals may be accounted for from agricultural products of rice, wheat, corn and sugarcane. Findings of this study generate new insights into the resource recovery potential from agriculture biomass and sustainable development of the SSC.

## Acknowledgements

The authors acknowledge funding from Scottish Water and the UK EPSRC Science and Solutions for a Changing Planet Doctoral Training Programme. Thanks also to Kimberley Pavier and Alex Bowles for insight on the silicon supply chain and resource recovery.

## References

- Chigondo, F. 2018. From metallurgical-grade to solar-grade silicon: an overview. *Silicon*. 10 (3), 789-798.
- Errington, E., et al. 2022. Environmental Impacts of Rice Husk-Derived Silica under Uncertainty: Is “Bio” better? in Proceedings of the 32nd European Symposium on Computer Aided Process Engineering (eds. Montastruc, L. & Negny, S.) 1615–1620.
- Errington, E. et al. 2023. Synthetic amorphous silica: environmental impacts of current industry and the benefit of biomass-derived silica. *Green Chemistry*. Advance Article.
- EU Commission. [Online]. Electricity industry retail prices 2008-2019. (2023). Available at: [www.energy.ec.europa.eu/](http://www.energy.ec.europa.eu/)
- Food and Agriculture Organization of the United Nations (UNFAO). 2022. Food and Agriculture Data: Crops and Livestock Products– “Rice, Paddy”, “Maize (corn)”, “Wheat”, “Sugar Cane”
- Frischknecht, N. et al. 2004. The ecoinvent Database: Overview and Methodological Framework . *The International Journal of Life Cycle Assessment* 2005 10:1 10 (1), 3–9
- Huijbregts, M. et al. 2017. "ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level." *The International Journal of Life Cycle Assessment* 22.2 138-147.
- Jungbluth, N. et al.. Photovoltaik. Sachbilanzen von Energiesystemen: Grundlagen für den ökologischen Vergleich von Energiesystemen und den Einbezug von Energiesystemen in Ökobilanzen für die Schweiz. ecoinvent report No. 6-XII (2009).
- Maurits, J. E. A. Silicon Production. *Treatise Process Metall.* 3, 919–948 (2014)
- Siddique, R. & Cachim, P. 2018. Waste and Supplementary Cementitious Materials in Concrete: Characterisation, Properties and Applications. Elsevier, Woodhead Publishing.
- U.S. Geological Survey (USGS). Mineral Commodities Summary 2022. (2022).
- U.S. Geological Survey (USGS). Mineral Commodities Summary 2019. (2019).
- Wernet, G., Bauer, C., Steubing, B., Reinhard, J., Moreno-Ruiz, E., and Weidema, B., 2016. The ecoinvent database version 3 (part I): overview and methodology. *The International Journal of Life Cycle Assessment*, 21(9), pp.1218–1230.
- World Resources Institute. 2022. [Online]. Climate Analysis Indicators Tool (CAIT), Climate Watch: "Historical GHG Emissions". Available at: [www.climatewatchdata.org](http://www.climatewatchdata.org)

# Technoeconomic and Life Cycle Assessment of an mRNA Vaccine Integrated Manufacturing Plant

Shang Gao<sup>a</sup>, Brahim Benyahia<sup>a\*</sup>

<sup>a</sup>*Loughborough University, Department of Chemical Engineering, Epinal Way, Loughborough, Leicestershire, LE11 TU, UK*  
*B.Benyahia@lboro.ac.uk*

## Abstract

In this study, a technoeconomic and Life Cycle Assessment (LCA) methodology was developed to identify the best compromises between costs and environmental performance of an integrated mRNA vaccine production plant. The plant-wide simulator was obtained and adopted from the literature which helped generate the technoeconomic data along with the life cycle inventories. The LCA was developed using SimaPro software based on the Ecoinvent database and ReCiPe 2016 method. A prediction tool based on artificial intelligence was used to estimate the environmental footprint of several raw materials. A base case scenario helped identify the environmental hotspots which laid the ground for a new set of scenarios to explore ways to lower the overall environmental footprint. Consequently, several recycles were implemented which helped reduce the endpoint categories by up to 10% while reducing the operating expenditures by 13%. The results gave key insights into the optimization of mRNA vaccine production and waste minimization.

**Keywords:** Plant-Wide Simulator, mRNA vaccine, Life Cycle Assessment (LCA), Technoeconomic Assessment, Recycles.

## 1. Introduction

Over the last few years mRNA technologies have gained a resurgent interest, due to recent technical advances and growing industrial investment putting mRNA vaccines in the position to substitute protein based conventional vaccines. Compared to the conventional vaccines, such as live attenuated and killed virus vaccines, subunit, and DNA-based vaccines, mRNA vaccines are safe, more efficient, and can be produce at large scales (Pardi et al., 2018). The upstream processing of mRNA vaccines is relatively straight forward and exhibits less sources of variability compared to cell-based vaccines. Furthermore, mRNA vaccines production process requires smaller bioreactors (from 30 L to 50 L) compared to cell-based vaccines where typically 2000 L bioreactors are required. With a relatively smaller production scale reduced facility footprints, the capital costs of the mRNA plant may be much lower. In addition, mRNA plants are flexible and versatile and can be rapidly deployed to produce new vaccines to responded for new variants, because different RNA sequences that used into different vaccines can be produced in the same process (Pardi et al., 2018). However, the production of mRNA requires complex recipes and a large set of raw materials, excipients and solvents which may exhibit significant environmental impacts.

The new mRNA vaccines played a prominent role during the covid-19 pandemic. The global mRNA production capacity is skyrocketing and is expected to exceed 8 billion doses per year (Kis et al., 2021). Consequently, it is a crucial time to evaluate the available performance compromises, which include the economic, quality and environmental



criteria, to make well informed decision at this early stage of a new era of mRNA vaccines. Flowsheet and plant wide models are critical in all design stages to help identify the most economic and environmentally friendly options optimize process operation and develop plant wide control strategies (Benyahia et al, 2012; Benyahia 2018; Ramin et al., 2018).

This work is focused on the technoeconomic and Life Cycle Assessment (LCA) of an integrated end-to-end mRNA production plant. The plant-wide simulator was obtained and adopted from the literature (Kis et al., 2021). It is worth emphasizing that process data are still very scarce and so are the environmental data associated with the chemicals, and raw materials. Here, the environmental footprint of the chemicals not available in the data bases were estimated using artificial intelligence-based methods (Wernet et al., 2009). Firstly, a base case LCA was conducted to identify the environmental hot spots which informed on the best opportunities for the implementation of recycles and resulted in 2 new design alternatives. Additional LCA were conducted for the new scenarios along with inherent economics to help evaluate and compare the performance of different design alternatives.

## 2. Material and Methods

### 2.1. Process description

The process was modelled and simulated by using SuperPro Designer V10. The initial SuperPro simulation version was obtained from the literature (Kis et al., 2021) then adopted to the current needs. The mRNA vaccine production process starts with a biochemical reaction called *in vitro* transcription where the mRNA is synthesized, with a series of additions of raw materials including different nucleotides, DNA templates and enzymes to the bioreactor as shown in the upstream processing in Figure 1. This is followed by a series of separation processes to purify the solution and sperate solid and liquids. Firstly, a tangential flow filtration (TFF) along with diafiltration unit is applied to remove RNA fragments, with the ultrafiltration membrane size of 500 kDa and KCl buffer to wash the filter to increase the efficiency. This stage is followed by a chromatography unit which separates the content based on size, total charge, surface hydrophobic groups. After the chromatography unit, another TFF unit is used with different buffer sodium citrate followed by sterile filter to prepare the formulation step. Finally, the mRNA is encapsulated within lipid nanoparticles (LNPs) to deliver the final injectable suspension.

### 2.2. Life cycle assessment

LCA is a holistic approach that can be used to capture the environmental footprint of a process, a product or a service across its life cycle. It provides a comprehensive tool to evaluate the environmental burdens based a large set of impact indicators referred to as midpoint and endpoint categories which are far beyond the capabilities of the green engineering metrics. Here, a gate-to-gate approach is used to evaluate different design and recycle options. The LCA requires the inventory data which can be obtained from the simulators under different scenarios. The implementation of the LCA requires 4 key steps namely:

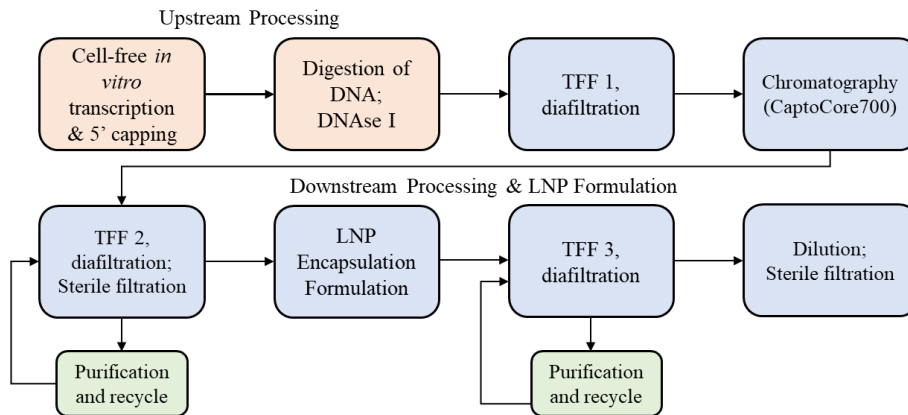


Figure 1. Simplified mRNA plant wide flow diagram (Kis et al., 2021) with recycles.

- *Goal and scope*

Here, the first objective is to perform a LCA for the base case mRNA vaccine production process using a functional unit of 1 kg of mRNA product. The second objective is to evaluate alternative designs by implementing different recycles.

- *Life cycle inventory*

LCI data were obtained from the plant wide simulation of the process flow diagram above using SuperPro Designer.

- *Impact assessment*

In this study, the impact assessment is based on the method ReCiPe 2016 in SimaPro with Ecoinvent 3, environmental footprint database and European and Danish Input/Output data base. ReCiPe 2016 method has three endpoint categories (human, ecosystems, and resources) and seventeen midpoint categories. These different categories will help compare the different design options.

- *Interpretation of results*

Finally, the results obtained from the previous steps are interpreted and the environmental hotspots are identified.

To capture the environmental footprint of the different processing steps, the plant was divided into six main stages: Upstream processing, first filtration TFF1, Chromatography, second filtration TFF2, Encapsulation, and final filtration TFF3. It should be emphasized that several chemicals used in the process are not available in the Ecoinvent database or any other environmental databases. To address this limitation, a prediction tool, based on machine learning Finechem was used to predict the global warming potential (GWP), cumulative energy demand (CED) and Eco-Indicator 99 score (EI99) based on the chemical structure of the molecules (Wernet et al., 2009). These data were then fed to Simapro to obtain the overall impact results.

The proposed LCA was implemented for the following scenarios:

*Scenario 1: Base case* - the initial design presented in Figure 1 without recycles.

*Scenario 2: TFF2 Recycle.* In this design option, the waste from the second filtration is purified using an ultrafiltration and a distillation unit to retrieve pure solvent.

*Scenario 3: TFF2 and TFF3 Recycles.* Here, the wastes from both hotspots are purified and recycled.

To evaluate the economic performance of each of the scenarios, the operating and capital expenditures/costs (OpEX and CaPex) were estimated by using the built-in tool within SuperPro Designer. The economic data were obtained based on the current plant/equipment capacity and material and energy usage for each scenario.

### 3. Results and Discussions

#### 3.1. LCA results

Figure 2 shows the normalized impacts categories for each of the 6 processing stages in the base case scenario. The second tangential flow filtration stage (TFF 2) exhibits the largest contribution in all impact categories. This can be explained by the large amount of solvent used and wastes generated at this stage along with large quantities of sodium citrate which is used as a buffer to adjust the pH of the product solution. The second largest contribution is associated with the diafiltration (TFF3) followed by the encapsulation. These results clearly confirm the environmental hotspots and justify the alternative scenarios based on different recycle options.

#### 3.2. Assessment of scenarios with recycles

As described above, TFF2 is the most important environmental hotspot among all stages. To purify then recycle the waste stream TFF2, a filtration and a distillation unit are proposed and integrated to the flowsheet simulator in SuperPro designer. The reason to choose these units is that in this stream the objective is to recycle pure sodium citrate, so the first ultrafiltration is used to desalt the solution and remove waste RNA transcripts by adding water as solvent to wash the waste. This step is followed by a distillation unit used to separate water and sodium citrate. TFF3 is the second hot spot based on the base case LCA results. In this case, a new recycle scenario is proposed to remove lipids left from encapsulation process and desalt the waste solution to retrieve the buffer from the waste.

With the new recycle scenarios associated with TFF2 and TFF3 process, the LCA were conducted. The corresponding midpoint and endpoint indicators along with the cost

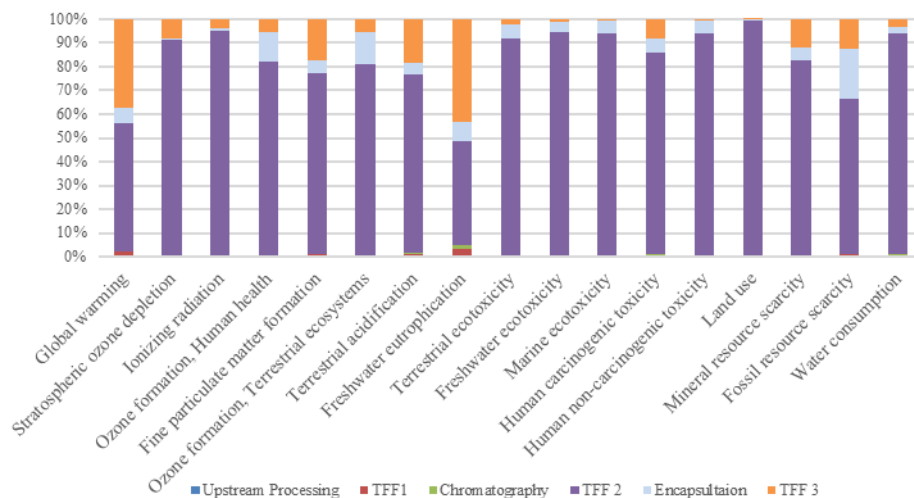


Figure 2. Normalized midpoint categories for the base case scenario.

performance are summarised in Figure 3. Compared to the base case, both recycle scenarios exhibit dramatic reduction in all midpoint categories particularly, the global warming, stratospheric ozone depletion, and land use (Figure 3a). When comparing both recycle seniors, their performances are close with a visible advantage for the senior with combined recycles particularly on the freshwater eutrophication impact category.

The economic performance indicators along with endpoint indicators are shown in figure 3b. Again, the environmental advantages of both recycle scenarios are clear. Although both recycle scenarios require increased capital expenditures associated with the waste purification and recycle equipment, their operating expenditures are reduced by nearly 8% and 13% respectively for the simple and combined recycles respectively. This is a very important outcome which clearly highlights the advantages of the recycle far beyond the environmental consideration and confirms the practicality and feasibility of the recycles in pharma and biopharma which are both notorious for their costly chemicals and additives.

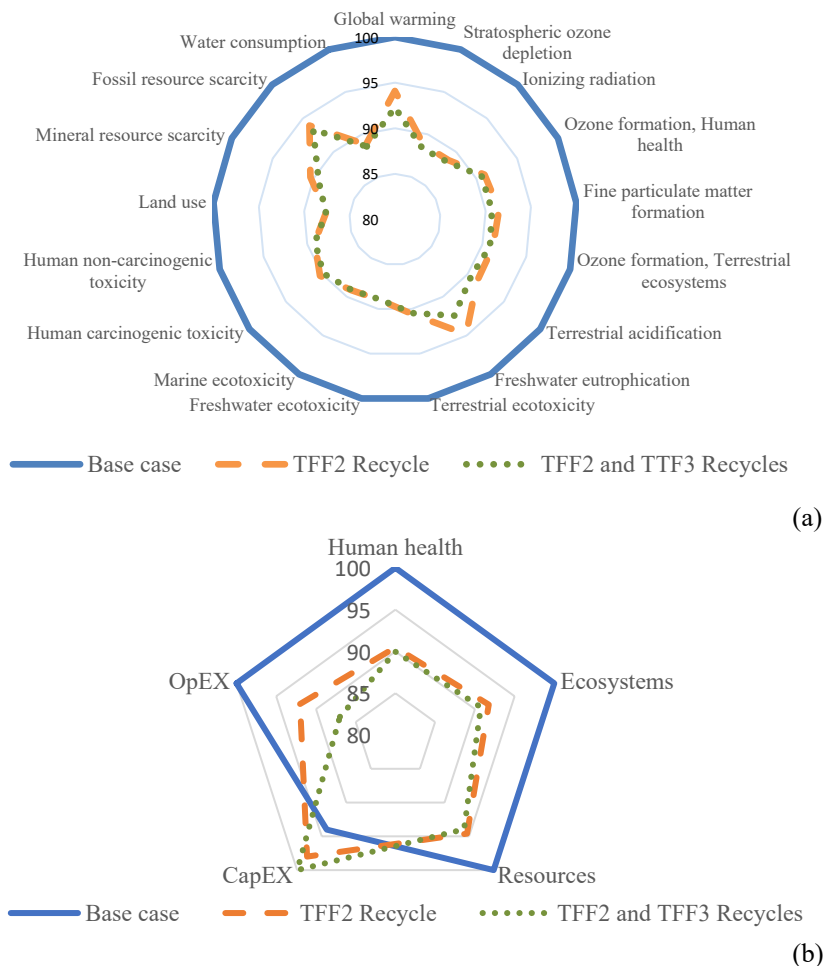


Figure 3. Summary of the environmental and economic key performance indicators for all scenarios. (a) Midpoint indicators. (b) End-point indicators and economic criteria.

#### 4. Conclusions

A technoeconomic analysis and Life Cycle Assessment of a COVID-19 mRNA vaccine production plant was developed. The integrated end-to-end mRNA simulator was developed in SuperPro designer. A first LCA was conducted using the base case scenario to identify the hotspots, based on which new designs with recycles were proposed and simulated. Consequently, new inventory and economic data were obtained for the proposed new designs which helped build new LCA and economic evaluations. As expected, the recycles reduced dramatically the environmental footprint of the process which translates into a decrease in the endpoint indicators, namely human health, ecosystem and resources, from 6% to 10%. Most importantly the results show that the implementation of recycles can also reduce the operating expenditures by up to 13% making the whole process more viable and more environmentally friendly. These results highlight the benefits of the recycles in pharma and biopharma which allows recycling precious or very expensive pharma grade chemicals and reduce inherent wastes. These benefits can be even more important when the production scale or the number of batches is large which is true for the current mRNA vaccine production plants.

#### References

- Benyahia, B., Lakerveld, R., Barton, P.I., 2012 A plant-wide dynamic model of a continuous pharmaceutical process, *Industrial and Engineering Chemistry Research*, 51(47), pp. 15393–15412.
- Benyahia, B., 2018. Applications of a plant-wide dynamic model of an integrated continuous pharmaceutical plant: Design of the recycle in the case of multiple impurities. *Computer Aided Chemical Engineering*, Elsevier B.V. 41, 141-157.
- Kis, Z., Kontoravdi, C., Shattock, R., Shah, N., 2021. "Resources, Production Scales and Time Required for Producing RNA Vaccines for the Global Pandemic Demand" *Vaccines* 9, no. 1: 3. <https://doi.org/10.3390/vaccines9010003>
- Kis, Z., Zain, R., 2021. 'How to Make Enough Vaccine for the World in One Year', *Public Citizen*.
- Pardi, N., Hogan, M., Porter, F., *et al*, 2018. mRNA vaccines — a new era in vaccinology. *Nat Rev Drug Discov* 17, 261–279.
- Ramin, P., Mansouri, S.S., Udugama., I.A., Benyahia, B., Gernaey, K.V., 2018. Modelling continuous pharmaceutical and bio-based processes at plant-wide level: A roadmap towards efficient decision-making. *Chemistry Today* 36 (2), 26-30.
- Wen, E.P., Ellis, R., Pujar, N.S., 2015, *Vaccine Development and Manufacturing*, Wiley Series in Biotechnology and Bioengineering (Hoboken, New Jersey, United States: John Wiley & Sons, Inc.)
- Wernet, G., Stavros P., Stefanie, H., Konrad, H., 2009. "Bridging data gaps in environmental assessments: Modeling impacts of fine and basic chemical production." *Green Chemistry* 11 (2009): 1826-1831.

## Obtaining biofuels from agroindustrial waste of *Fragaria spp* and *Mangifera indica spp*

Ernesto Barrera-Reyes<sup>a</sup>, Luis Alberto Bretado-Aragón<sup>a</sup>, Guillermo Adolfo Anaya-Ruiz<sup>a</sup>, Pascual Eduardo Murillo-Alvarado<sup>a\*</sup>, José Luis Baltazar-Álvarez<sup>a</sup>, José Joel Román-Godínez<sup>a</sup> and Gabriela Guadalupe Esquivel-Barajas<sup>a</sup>.

<sup>a</sup>Universidad de La Ciénega del Estado de Michoacán de Ocampo. Av. Universidad Sur 3000, Lomas de Universidad, Michoacán 59103, México.

### Abstract

Population growth has advanced by leaps and bounds, which has caused the agricultural sector to see the need to meet this demand. Conveniently, the practices of the agro-industrial sectors have directly affected natural resources. Within the discharges from agro-industries, they are mainly conditioned by pollutants with organic and inorganic compounds, being the objects of study of the present investigation organic waste because they are a potential source that compromises human health. In such a way that, the purpose of the present investigation consisted of collecting residues from strawberry and mango agro-industries in Zamora-Jacona, Michoacán, Mexico, such residues were synthesized as biodiesel and were characterized by Fourier transform infrared spectroscopy (FT-IR) and UV-VIS, so it was possible to identify absorption bands, corresponding to the functional groups, which are similar to those of hydrocarbons. Thus, it is concluded that the conversion of agro-industrial remnants can be used in the field of biofuels.

**Keywords:** biodiesel, agroindustry, remnants.

### 1. Introduction

Mexico is one of the main producing countries of both strawberries (*Fragaria spp*) and mango (*Mangifera indica spp*), with the central region of the country being the area with the highest production, this is due to the fact that they were produced in 2016, 468.25 and 1.89 million tons respectively of these fruits, and it is estimated that by 2030 production will increase between 26 and 42% (SAGARPA 2016).

Berries, strawberries or forest fruits, refer to the various fruits of different botanical groups, which are characterized by their short shelf life. They tend to be small, tart-sweet, juicy, and brightly colored, making them very attractive to eat. The production of berries at the national level is concentrated in four states, highlighting the state of Michoacán as the most important producer, followed by Baja California, Jalisco and Guanajuato. The country exported close to 1,000 million dollars of fresh berries to the world in the 2015, which projects this group of fruits as one of those with the greatest growth potential in the agricultural sector (González et al 2019).

In the case of the present investigation, the site of interest focuses on 2 municipalities in the state of Michoacán, which are Zamora and Jacona, since they are the main producers of berry and mango crops nationwide. In such a way that the objective was to elaborate the synthesis of biofuels from the organic waste material generated by agro-industries of Zamora and Jacona of the State of Michoacán de Ocampo.

## 2. Case study

### 2.1 *Geographical characteristics of Zamora-Jacona*

Zamora R., & Salazar in 2019, describe in their manuscript that the municipality of Zamora is among the 6 most important in the state of Michoacán, because it produces 4.6% of the total gross production. It has a population of 170,748 inhabitants (4.3 percent of the total population), of which 26,997 are employed persons, with 6,778 economic units. The city represents one of the three conurbations of the state; Zamora-Jacona-Tangancicuaro. One of the great advantages that Zamora has in strawberry production is its proximity to several of the main trade distribution points nationwide, such as the markets of the Federal District. In addition to its proximity to several of the freezing and packing companies that are involved with the strawberry agro-industrial activities, located mainly in the cities of Irapuato and Celaya, they are also dedicated to freezing and storing strawberries, located in several trading and industrializing companies' leaders in the export of strawberries.

### 2.2 *Contaminating sources of strawberry and mango crops*

At present there is an indiscriminate demand for the consumption of different fruits, among them red fruits stand out, unfortunately high amounts of waste materials produced by these fruits are produced annually.

Fraire-Cordero M., et al (2003), carried out an investigation whose purpose was based on identifying pathogenic fungi in strawberry fruit and their relationship with varieties and the cultivation system, five postharvest samplings were made: three in the region of Zamora, Michoacán, México, in December 2001, and in February and May 2002, in the *Camarosa* and *Aromas* varieties, cultivated with the traditional system, and with plastic cover. In their research they describe the identification of nine pathogenic fungi which were identified from national and imported fruit: *Aspergillus*, *Botrytis*, *Colletotrichum*, *Geotrichum*, *Mucor*, *Penicillium*, *Pestalotiopsis*, *Phytophthora*, and *Rhizopus stolonifer*; and more than eight saprophytic fungi.

Estrada-Loera et al (2018), carried out various samples in mango crops, subsequently carried out microbiological tests, thus in this study they allowed to identify the presence of *Klebsiella spp*, and non-enterohemorrhagic *E. coli*.

Due to the aforementioned, waste from agro-industries is produced, which have different places where they are stored and if they do not have an appropriate treatment, they are a potential source of microorganism reservoirs and these, in turn, can compromise human health. In this way, the purpose of the investigation lies in identifying freezers in the Zamora-Jacona region and sampling the generated waste so that it can be converted into a biofuel.

## 3. Methodology

For the development of the present investigation, it was divided into 4 stages, which consist of: 1) identification of agro-industries, 2) sampling of strawberry and mango agro-industrial residues, 3) extraction of essential oils and obtaining biodiesel, and 4) characterization of oils and biodiesel

*Stage 1.* Identification of agro-industries. Agro-industries that process and distribute strawberry and mango were identified in the municipalities of Zamora-Jacona, Michoacan, Mexico.

*Stage 2.* Strawberry and mango remnants from agro-industries were collected, where they were considered: strawberry pulp and stem, on the other hand mango seeds were collected, the raw material was stored in sterile plastic bags, later they were transferred to a desiccator (Novatech), with the aim of dehydrating the biomass.

*Stage 3.* Extractions of essential oils. For the development of this stage, an adaptation was made to the methodology proposed by Sánchez I. & Huertas K. (2012), which consisted of: 1) dehydrating the biomass of both fruits, 2) grinding the mango seeds and the dehydrated strawberry pulp, and 3) the soxhlet method (KIMAX brand) was used, which consisted of pouring 200 ml of isopropyl alcohol (MEYER), subjecting it to a constant temperature and stirring of 300°C and 1,400 rpm, for 15 cycles. Afterwards, a rotoevaporator was used to separate the solvent from the essential oil and finally the transesterification method was carried out, which was carried out by centrifuging the essential oil for 30 minutes at 6,000rpm and eliminating the degumming or residues to be able to be mixed with the concentration. to 1% of the potassium iodide solution (Reasol brand) with an addition of 3% (p/p) of distilled water and place on the grill to keep stirring for 30 minutes at 60°C to finish it was brought to the centrifuge at 6,000rpm for one hour.

*Stage 4.* Characterization of biodiesel. Biodiesel was characterized by Fourier transform infrared spectroscopy (FT-IR), Perkin Elmer Frontier model in the 4,500-400cm<sup>-1</sup> region and by Perkin Elmer Lambda 25 model UV-VIS spectrophotometry, in the absorbance region of 300-800nm.

#### 4. Results and discussions

Sampling points for agro-industry waste were identified. Once the samples were prepared, the fruits were taken to the drying oven, in which they were dehydrated with the aim that a biomass screening could be carried out (Fig. 1).

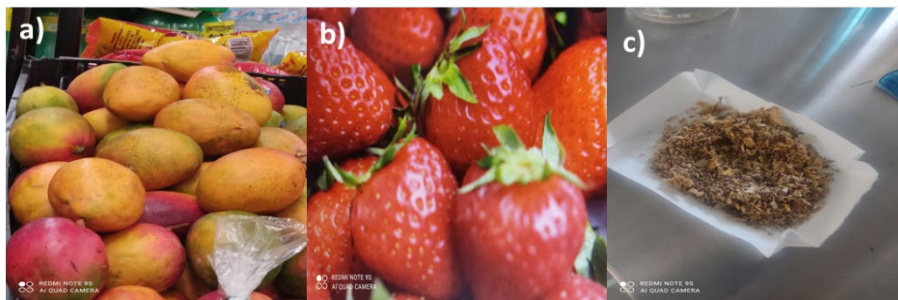


Fig.1. Fruits considered as waste in agro-industries in Zamora-Jacona. a) Decaying fruits of *Mangifera indica spp*, b) Decaying fruits of *Fragaria spp*. and c) Screening for *Mangifera indica spp*.

Subsequently, the extraction of essential oils was carried out (Fig. 2), carrying out an adaptation to the methodology proposed by Sánchez I. & Huertas K. (2012). It started by obtaining essential oils using the Soxhlet method (Fig. 2a and 2b), having the essential oils biodiesel was synthesized (Fig. 2c) which was subjected to refining with the purpose of reducing impurities (Fig. 2d).



Once the biodiesel was obtained, it was characterized by both UV-VIS and FT-IR (Fig. 3 and Fig. 4). Regarding the characterization by UV-VIS, the band corresponding to the reading of mango biodiesel (MD2), presents a greater intensity and absorbance, compared to the band of strawberry biodiesel (FD2), this is attributed to the fact that in the upper part mango seed stores a greater amount of oil.

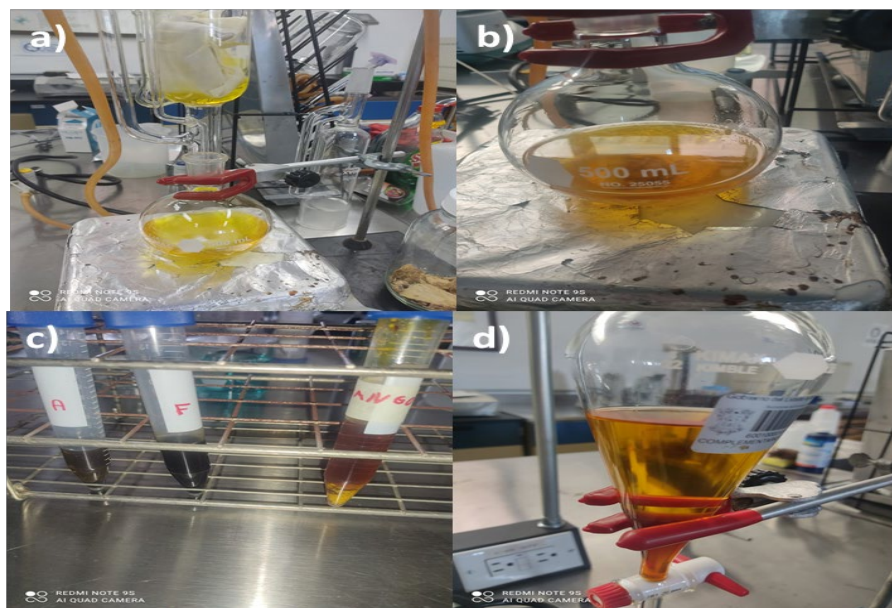


Fig. 2. Conversion of essential oils to biodiesel. a) and b) Essential oil of *Mangifera indica* spp and *Fragaria* spp, c) and d) Refined biodiesel.

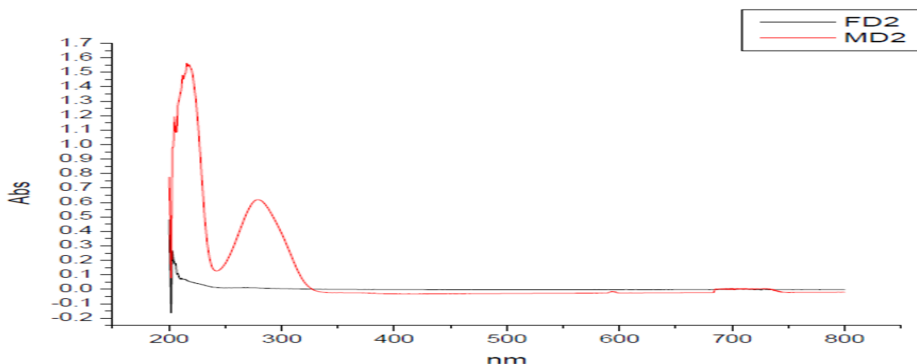


Fig. 3. UV-VIS characterization of strawberry (FD2) and mango (MD2) biodiesel.

Regarding the characterization by FT-IR (Fig.4), it was possible to identify the functional groups, characteristic of the absorption bands of strawberry biodiesel, within the identified groups are: OH, CH<sub>2</sub>, CH and PH.

In the same way, a proposal was developed where this technology can be migrated on an industrial scale, for which the design at the industry level is proposed (Fig. 5).

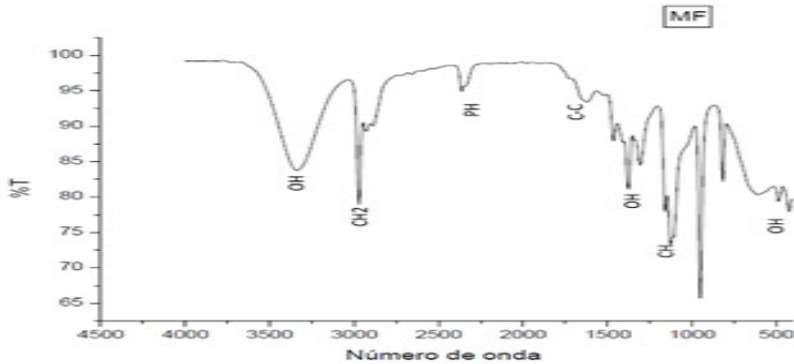


Fig. 4. FT-IR characterization identify the absorption bands of strawberry biodiesel

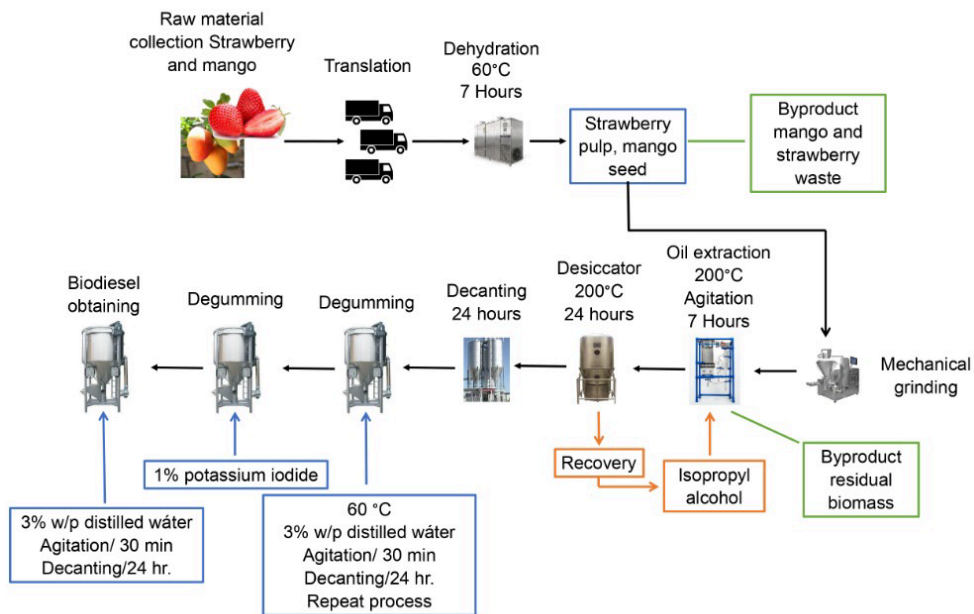


Fig. 5. Biodiesel production scheme from agro-industry remnants.

In the year 2021 the state of Michoacan had a production of 180,871.84 tons of mango, for the proposed methodology only the residual part due to the processing of the fruit is considered, in the case of mango only the interior seed that corresponds to 17.5% is considered of the fruit. In this sense, considering mango production, there are 31,652.57 tons of residual biomass. Based on the data obtained at the experimental level and the process proposed in Fig. 5, it is possible to obtain  $7.0339 \times 10^5$  Lt of biodiesel. In the case of strawberries, in 2021 a production of 326,191.10 tons was obtained, in this case the largest amount of the fruit is pulp, so the residual biomass obtained during

processing corresponds to 10% of the total available. Therefore, a total of 32,619 tons of residual biomass is available. From this flow, it is possible to obtain a total of  $1.63 \times 10^5$  Lt of biodiesel. Through the residual biomass it is possible to obtain a total of  $8.664 \times 10^5$  Lt of biodiesel, considering an approximate price of \$1 dollars, a profit of  $\$8.664 \times 10^5$  USD.

## 5. Conclusions

The remnants generated in the agro-industries can be possible sources of infection for society, in such a way it is necessary to identify ways of using such residues, in this way the alternative of using the obtaining of biofuels and these biofuels can be used by society. The highest yield and production of biodiesel occurs in the mango; this is due to the fact that a greater amount of residual biomass is generated during its processing. The strawberry presents a considerable production, since it is one of the fruits with the highest production and demand in Mexico.

## References

- Estrada-Loera Rosa María, Gallegos-Robles Miguel Ángel, Orona-Castillo Ignacio, García-Hernández José Luis, Osuna-García Jorge Alberto, Sánchez-Lucio Roberto, Ríos-Plaza Juan Luis, Vázquez-Vázquez Cirilo (2018) Prevalence Of Pathogenic Bacteria In Mango (*Mangifera Indica* L) Cv. Tommy Atkins. *Revista de Ciencias Biológicas y de la Salud*. P.p. 6-9.
- Fraire-Cordero María de Lourdes, Yáñez-Morales María de Jesús y Nieto-Ángel Daniel (2003) Hongos Patógenos en Fruto de Fresa (*Fragaria x ananassa* Duch.) en Postcosecha. *Revista Mexicana de FITOPATOLOGIA*. P.p. 285-290.
- González Razo Felipe de Jesús, Rebollar Samuel, Hernández Martínez Juvencio, Morales Hernández José Luis y Ramírez Abarca Orsohe (2019) Situación Actual Y Perspectivas De La Producción De Berries En México. *Revista Mexicana de Agronegocios*, vol. 44. P.p. 261.
- SAGARPA (2016) Planeación agrícola nacional 2017-2030. Mango Mexicano. P.p. 2.
- SAGARPA (2016) Planeación agrícola nacional 2017-2030. Fresa Mexicana. P.p. 2.
- Sánchez Medina y Huertas Greco Karina (2013) Obtención Y Caracterización De Biodiesel A Partir De Aceite De Semillas De *Ricinus Communis*. (Higuerilla) Modificadas Genéticamente Y Cultivadas En El Eje Cafetero. Universidad Tecnológica De Pereira.
- Zamora Jacobo Ricardo y Salazar Mosqueda (2019) Importancia De La Producción De Fresa En El Sector Agrícola En Zamora, Michoacán. *Rev. Realidad Económica*. P.p. 108.

# Life Cycle Assessment of Cellulose Nanocrystals Production in Sugarcane Biorefineries

Gustavo Batista<sup>a</sup>, Ana Carolina Borges Silva<sup>a</sup>, Cristiane Sanchez Farinas<sup>a,b</sup>,  
Antonio José Gonçalves da Cruz<sup>a</sup>

<sup>a</sup> Graduate Program of Chemical Engineering, Federal University of São Carlos,  
13565-905, São Carlos, SP, Brazil.

<sup>b</sup> National Nanotechnology Laboratory for Agribusiness (LNNA), Embrapa  
Instrumentation, Rua XV de Novembro 1452, 13560-970, São Carlos, SP, Brazil.

## Abstract

The purpose of using Life Cycle Assessment (LCA) in this work was to evaluate the cellulose nanocrystals (CNCs) production in sugarcane biorefineries, identifying hotspots that can direct to process design changes to reduce environmental footprint. Modeling and simulation of an incremental unit for CNCs production from sugarcane bagasse (the main residue from the sugarcane industry) was carried out. In order to proceed the LCA, an inventory of raw-materials, products and emissions was made. The analysis employed the cradle-to-gate approach. The Global Warming Potential (GWP) metric was estimated using the SimaPro<sup>®</sup> software and using the EcoInvent v.3.3 database. LCA results showed that daily production of 7.73 ton of CNCs occurred with accumulated GWP of 23.6 tons of CO<sub>2</sub> equivalents. Ethanol from organosolv pretreatment, sulfuric acid from hydrolysis, and sugarcane bagasse burning for steam and energy generation were the main contribution inputs for GWP. An estimated GWP of 0.23 kg of CO<sub>2</sub> equiv. per kg of dry equiv. CNCs (or 13.43 g of CO<sub>2</sub> equiv. / MJ of process outputs) were accounted after energetic allocation. In other impact categories aside from GWP, the environmental impacts of H<sub>2</sub>SO<sub>4</sub> were generally the most relevant, especially in the abiotic depletion.

**Keywords:** Life Cycle Analysis, Cellulose Nanocrystals, CNCs, Sugarcane Biorefineries.

## 1. Introduction

Life Cycle Assessment (LCA) is a standardized tool to quantify the environmental sustainability performance of emerging technology products along all phases of their life cycles, allowing the comparison of different process designs (ÖGMUNDARSON et al, 2020a; ÖGMUNDARSON et al, 2020b; FOROUGHİ et al., 2021). If the methodology is used during a new product development phase, especially during the process methodology planning, it can indicate the process stages or technologies with the highest environmental impacts, and thus provide a guide for improvements in the implementation of the technology (IBICT, 2014). ISO 14040:2006 and ISO 14044:2006 established a methodological framework for conducting LCA studies.

With the implementation of the carbon credits market, the quantification of the Global Warming Potential (GWP) and of other metrics of environmental impacts have now also a strong impact on investment decision-making, once high environmental impact

production methodologies will be neglected due to environmental and economic aspects as well (ÖGMUNDARSON et al., 2020a). Nanocellulose is an example of an emerging, still under development material for which a reduced environmental impact is expected when comparing to other existing materials (ARVIDSSON et al., 2015; PICCINNO et al., 2018). Cellulose nanocrystals (CNCs) are crystalline and highly ordered materials of small diameter, elongated length, and high surface area whose main applications are as reinforcement in polymeric materials and the applications in biomedicine. Consisting in a high added-value renewable biocomposite, but still in limited availability and presenting low yields of obtainment and isolation, the nanocellulose market is in continuous and strong growth (MARKETS AND MARKETS, 2021).

Due to some challenges such as data unavailability and limitations related to the end-of-life applicability and the solvent recycle treatments in lab-scale processes, there are only a few LCA studies related to nanocellulose products. A cradle-to-gate environmental impacts study of three production routes for cellulose nanofibrils (CNFs) was carried out by ARVIDSSON et al. (2015). The authors used wood pulp as feedstock, and three production routes investigated were: enzymatic pretreatment, carboxymethylation pretreatment and no previous pretreatment. The results obtained showed that processes without pre-treatment and using enzymatic treatment had relatively low environmental impact in terms of GWP (0.79 and 1.2 kg of CO<sub>2</sub> equiv. / kg of CNFs produced, respectively). NASCIMENTO et al. (2016) evaluated the LCA of the CNCs production from coconut fibers by extraction with high powered ultrasound, while using the residual lignin as power source for energy cogeneration to the process. The authors quantified the GWP of the process as 0.207 g of CO<sub>2</sub> equiv. / g of CNC produced. Still, to the best of our knowledge, the technical-environmental analysis of the production of CNCs from sugarcane biorefinery residues is a topic not yet well investigated in literature. Thereby, the purpose of using LCA in this work was to simulate CNCs production in sugarcane biorefineries attached to a second-generation ethanol (E2G) biorefinery, and to quantify the GWP of this nanocellulose unit in a 100-year horizon through LCA, identifying hotspots that can direct to process design changes to reduce environmental footprint.

## 2. Methodology

The simulated methodology of production proposed in this work employed sugarcane bagasse as feedstock. The process started with organosolv treatment (190 °C, 2h, 1:1  $V_{\text{ethanol}}/V_{\text{water}}$ , 1:10 dry mass / volume of solution) with ethanol recycle by flashing and distillation, followed by purification with H<sub>2</sub>O<sub>2</sub> (50 °C, 1 h, H<sub>2</sub>O<sub>2</sub> 7% vol, NaOH 5% vol, 10:1 m/m relative to input) and subsequent acid hydrolysis with H<sub>2</sub>SO<sub>4</sub> (45 °C, 45 min, 10:1 m/m relative to dry nanocellulose output, 70% yield). After washing and filtering of the produced CNCs, sulfuric acid was recycled by a system of evaporators operating in three stages, two of which operate at vacuum pressure (1atm, 300 mbar and 100 mbar). The nanocrystals downstream process then proceeded with residual H<sub>2</sub>SO<sub>4</sub> neutralization, dialysis, sonication, and drying of the material to purity of 95% in a spray-dryer. The annual sugarcane bagasse utilization as feedstock was set at 72000 tons per year. The mass and energy balances for discontinuous equipment were discretized, i.e., the input and output streams had mass and energy contents distributed over the time of the equipment usage cycles. Organosolv pretreatment with ethanol was selected over other treatment options because it provides high rates of removal of lignins and residual

hemicelluloses, whilst being environmentally less aggressive than the chemical processes (NASCIMENTO et al., 2016). In its turn, complementary alkaline delignification was selected for its high effectiveness in removing residual lignin content, generating biomasses with high-purity cellulose. Acid hydrolysis by concentrated  $H_2SO_4$  was chosen because it is the state-of art hydrolysis process in the CNCs production (VANDERFLEET & CRANSTON, 2021).

Several sources of literature data, external sources and laboratory data were used as input parameters of the Case Studies analyzed in this work. Mass and energy balances were performed on Microsoft Excel<sup>®</sup> electronic spreadsheets. In the energy balances, the thermodynamic calculations for the equipment were performed using the Aspen Plus<sup>®</sup> software and the results were transposed to the electronic spreadsheets. The NRTL thermodynamic model was used to represent the non-idealities of liquid phase mixtures in each analyzed system. The vapor phases were considered ideal due to the low pressure in the processes. Thermodynamic data related to sugarcane biomasses that was absent in the Aspen Plus<sup>®</sup> were inserted into this simulator using values obtained in DOMALSKI et al. (1986). Figure 1 shows a box diagram that represents the simulated process.

In order to proceed the LCA, an inventory of raw-materials, products and emissions were made based on the process modeling stage. The analysis employed the cradle-to-gate approach, and the functional unit was considered as 1 kg of cellulose nanomaterials produced. The GWP coefficient quantified the Greenhouse Gas Emissions (GHG) as grams of equivalent  $CO_2$  emitted per each produced kg of CNCs for 100 years of operation. The carbon intensity of residues from any process stage was equaled to zero in the cradle-to-gate approach.

In addition to the GWP metric, other environmental assessment categories were also evaluated, as the *CML-IA Baseline v3.04 2000* method was selected in *SimaPro<sup>®</sup> 9.0.0.35 PhD* software and using the *EcoInvent<sup>®</sup> v.3.3* database. For the allocation of environmental impacts to the product and the by-product from the process (lignin), energetic allocation was carried out as the impact's distribution factor.

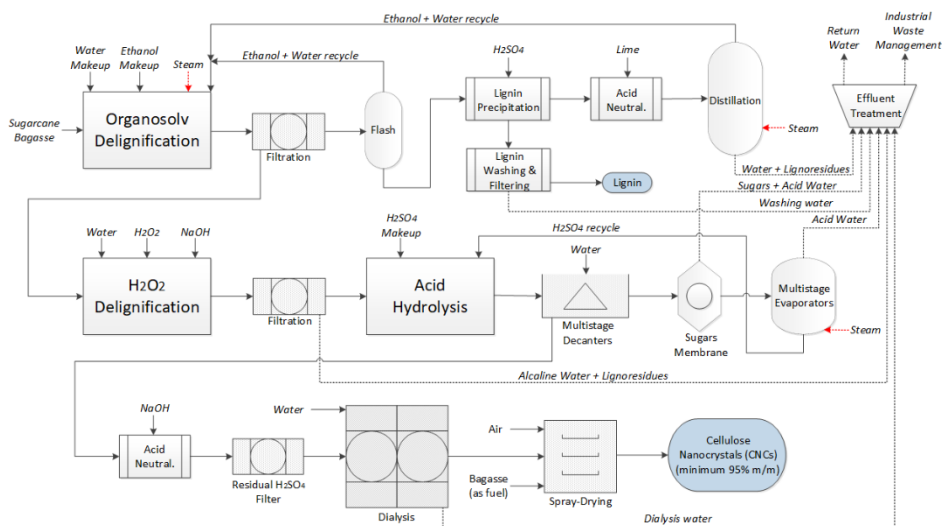


Figure 1 - Block diagram for the CNCs production methodology proposed in this work.

### 3. Results and Discussion

Modeling and simulation results showed that the CNCs production from sugarcane bagasse was water-use intensive, as 209 m<sup>3</sup> water / ton CNCs were consumed. LCA results showed that daily production of 7.73 ton of CNCs occurred with accumulated GWP of 23.6 tons of CO<sub>2</sub> equivalents. Ethanol from organosolv pretreatment, sulfuric acid from hydrolysis, and sugarcane bagasse burning for steam and energy generation were the main contribution inputs for GWP.

An estimated GWP of 0.23 kg of CO<sub>2</sub> equiv. per kg of dry equiv. CNCs were accounted after energetic allocation. This result corresponds to 13.43 g of CO<sub>2</sub> equiv. / MJ of process outputs. In other impact categories aside from GWP, the environmental impacts of H<sub>2</sub>SO<sub>4</sub> were generally the most relevant, especially in the abiotic depletion (96% of the total 1.37 kg of Sb equiv. per day of production), ozone layer depletion, soil acidification (52.8% of total 683 kg of SO<sub>2</sub> equiv. per day of operation) and water ecotoxicity (Figure 2), indicating that research studies may be directed to optimize the use of this acid in the nanocellulose production chain. Other compounds for acid hydrolysis (as organic acids) may represent a hindrance in the economics of the process but are also alternatives to be environmentally analyzed.

Life Cycle Assessment to produce nanocellulose varieties using recycled hydrolysis solvents are scarce studies in literature. Previous works that evaluated LCA of nanocellulose production, such as those by HOHENTHAL et al. (2012) (750 - 3100 g of CO<sub>2</sub> equiv. / kg of CNCs by TEMPO oxidation / high-pressure homogenization of sulfite pulp) presented higher GWP impacts per kg of nanocellulose produced. The relatively low GWP values when compared to processes that use derivatives of fossil fuels (and other biochemicals) indicate that the CNCs production from sugarcane bagasse is an option of interest for industries in the sector that aim to expand their portfolio with a high added-value product while reducing the emission of environmental impacts.

Biochemical production faces economic and environmental challenges that need to be overcome to enable a viable and sustainable bioeconomy. There are combined analysis frameworks indicated in literature (ÖGMUNDARSON et al., 2020b) that consistently combine environmental and economic indicators to support optimized biochemical production at early development stages. These frameworks propose the monetization of environmental impacts that, added to the economic indicators of the process, can generate process indexes that be used in the investment analysis. Given the global appeal for government policies to reduce environmental damage, it is expected that the adoption of this type of combined framework will be important for the expansion of bioeconomy.

### 4. Conclusions

Producing biochemicals from renewable resources is a key driver for moving towards a sustainable society where energetic efficiency and the use of agricultural residues are performance indicators (ÖGMUNDARSON et al., 2020b). A LCA of a nanocellulose production unit annexed to a sugarcane biorefinery was performed. The results showed that the GWP potential of the proposed process is concentrated in specific steps as the organosolv pretreatment and the sulfuric acid from hydrolysis. The identification of these steps can orientate future researchers for the demand of new methods of lignocellulosic biomass purification which can be environmentally friendly withal economically competitive for the nanocellulose production.

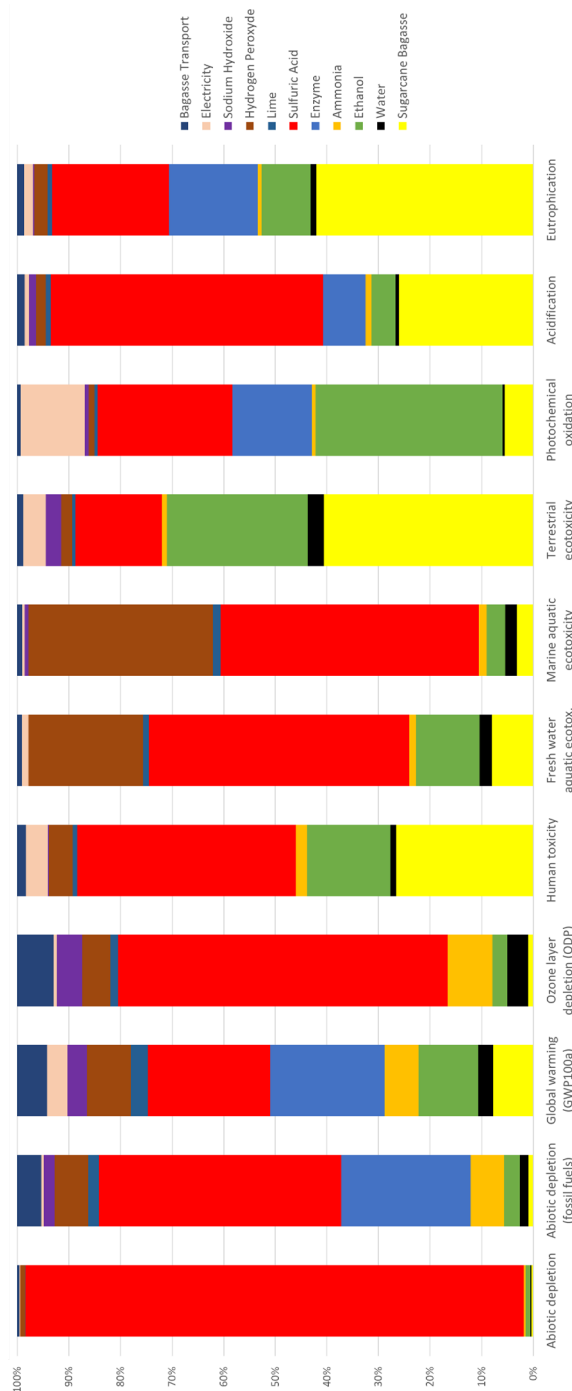


Figure 2 - Relative environmental impacts of each process input for obtained through Life Cycle Assessment (CML-IA Baseline v3.04 2000 method, SimaPro® software).

PERSONAL DATA



## References

- C. Hohenthal; M. Ovaskainen; D. Bussini; P. Sadocco; T. Pajula; H. Lehtinen; J. Kautto; K. Salmenkiv. Final assessment of nano enhanced new products. In SUNPAP (Scale-up Nanoparticles in Modern Papermaking), Oct.31, 2012. CTT Technical Research Center of Finland, InnovHub-SSCCP, Poyry Management Consulting Oy. Funded by European Community's 7th Frame Work Programme under Grant Agreement n° 228802; VTT Technical Research Centre of Finland: Espoo, Finland, 2012.
- D. M. Nascimento, A. F. Dias, C. P. Araújo Junior, M. F. Rosa, J. P. S. Morais, M. C. B. Figueirêdo, 2016. A comprehensive approach for obtaining cellulose nanocrystal from coconut fiber - part II: environmental assessment of technological pathways, *Industrial Crops and Products*, volume 93, pages 58-65.
- E. S. Domalski, T. L. Jobe Jr., T. A. Milne, 1986. Thermodynamic data for biomass conversion and waste incineration, 326 pages, National Bureau of Standards, US Department of Energy.
- F. Foroughi, E. R. Ghomi, F. M. Dehaghi, R. Borayek, S. Ramakrishna, 2021. A review on the life cycle assessment of cellulose: from properties to the potential of making it a low carbon material, *Materials*, volume 14, pages 714-736.
- F. Piccinno, R. Hischier, S. Seeger, C. Som, 2018. Predicting the environmental impact of a future nanocellulose production at industrial scale: application of the life cycle assessment scale-up framework, *Journal of Cleaner Production*, volume 174, pages 283-295.
- IBICT - Brazilian Institute of Information in Science and Technology, 2014. *Desenvolvimento Sustentável e Avaliação do Ciclo de Vida*, 38 pages, CNI, Brasília, Brazil.
- Markets and Markets, 2021. Nanocellulose Market worth \$661.7 million by 2023, available in: <<https://www.marketsandmarkets.com/PressReleases/nanocellulose.asp>>, accessed on August 04, 2021.
- O. M. Vanderfleet, E. D. Cranston, 2021. Production routes to tailor the performance of cellulose nanocrystals, *Nature Reviews Materials*, volume 6, pages 124-144.
- O. Ögmundarson, S. Sukumara, M. J. Herrgård, P. Fantke, 2020a. Combining environmental and economic performance for bioprocess optimization, *Trends in Biotechnology*, volume 38, issue 11, pages 1203-1214.
- O. Ögmundarson, M. J. Herrgård, J. Forster, M. Z. Hauschild, P. Fantke, 2020b. Addressing environmental sustainability of biochemicals, *Nature Sustainability*, volume 3, pages 167-174.
- R. Arvidsson, D. Nguyenand, M. Svanström, 2015. Life cycle assessment of cellulose nanofibrils production by mechanical treatment and two different pretreatment processes, *Environmental Science and Technology*, volume 49, pages 6881-6890.

## Acknowledgements

The authors would like to thank the Brazilian research funding agencies Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP (Grants 2016/10636-8 and 2019/25261-8) and Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (Grant 140761/2017-9). This study was financed in part by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, Brazil (Finance Code 001) with Grant 88887.364443/2019-00.

# Decarbonization of energy-intensive industries: sustainable implementation of CO<sub>2</sub> recycling within the industrial symbiosis

Marta Rumayor\*, Javier Fernández-González, Antonio Domínguez-Ramos,  
Angel Irabien

*Universidad de Cantabria, Departamento de Ingenierías Química y Biomolecular,  
Av. Los Castros s/n, Santander, Spain*

*\*rumayorm@unican.es*

## Abstract

Several decarbonization strategies have been proposed for the cement sector including fuel switching (FS), carbon storage (CS), and carbon recycling (CR). FS and CS are expected in the short-term whereas CR is expected by 2050. Between the CR options, electrochemical reduction (ER) provides the opportunity to close the anthropogenic CO<sub>2</sub> cycle producing chemicals while promoting industrial symbiosis (IS). However, it requires high penetration of renewables. This study evaluates three decarbonization scenarios emphasizing the IS benefits provided by ER. Since ER is at low TRL, we elucidate the prospects to ensure its environmental feasibility within the European cement sector. We analyze the prospects of ER performance to compete with FS, which is viable even under low TRL. However, ER cell energy efficiencies are higher than 50% and methanol concentrations higher than 40%wt. are needed to compete with CS. From the defossilization perspective CR exhibits a wider margin of opportunity to compete even with CS.

**Keywords:** carbon dioxide, electrochemical reduction, decarbonization, defossilization, cement

## 1. Introduction

Since the consequences of climate change become more apparent, industries, scientists, and governments are committed to meet stringent greenhouse gas (GHG) emissions targets. These climate goals together with the recent geopolitical and energy market reality require us to drastically accelerate the energy transition increasing Europe's energy independence from unreliable suppliers and volatile fossil fuels.

Cement, which is an important building material, will pave the way to achieve a climate-resilient infrastructure. However, the success of the energy transition hinges on the effectiveness of its decarbonization. During cement production, carbon dioxide (CO<sub>2</sub>) is emitted directly from the chemical conversion process of clinker production (calcination) but also because of fossil fuel combustion (mainly petcoke, coal, and oil). Given the nexus of cement production with the world's population, cement demand has been projected to increase by 2 or 3 times during the next few decades (Favier et al., 2018). This trend will lead to drastic consequences in terms of CO<sub>2</sub> emissions and fossil resource depletion.

Today, the decarbonization of energy-intensive industries as the cement sector is not straightforward. It employs extremely high temperatures, which can be achieved cost-effectively only by burning fossil fuels. The challenge of finding new, carbon-free approaches to produce low-carbon cement is just as complicated as well as expensive.

The recent cement decarbonization roadmap has predicted an amount of 7.7 Gt CO<sub>2</sub> of cumulative direct CO<sub>2</sub> savings by 2050 under the proper application of reduction strategies (International Energy Agency, 2018). This could be only achieved through improving energy efficiency, fuel switching (FS) to those that are less carbon-intensive, reducing clinker content in cement, and implementing innovative technologies such as Carbon Sequestration (CS) and/or Carbon Recycling (CR). In this sense, innovative CR technologies have arisen in the last decade as promising options to boost the so-called industrial symbiosis (IS) concept. The IS concept was introduced in the fifth Assessment Report from the Intergovernmental Panel on Climate Change (IPCC) as a way to reduce raw material consumption while mitigating GHG emissions on a process level, industry park level, and national level. CR implementation in the cement sector as an IS strategy can contribute not only to reaching the projected cumulative CO<sub>2</sub> emissions reduction but also to reducing the consumption of fossil fuels elsewhere.

The present study evaluates the benefits of the CR strategy in the cement sector in comparison with FS and CS. We have chosen methanol as the CO<sub>2</sub> conversion product given its market capacity and potential uses as fuel and chemical (Olah, 2005). Several CO<sub>2</sub> conversion pathways to methanol have been proposed in the latest decades, including catalytic conversion (CC), thermochemical (TC), and electrochemical reduction (ER) (Roh et al., 2020). We have selected ER route given its potential scalability, mild temperature, and pressure conditions, as well as the final possibility of using sunlight directly in the long-term. In a previous environmental assessment carried out by the authors (Rumayor et al., 2019), it was analyzed the influence of the selected key performance parameters (KPPs) in the methanol carbon footprint (CF), concluding that the feasibility of the ER technology is still hampered by the low methanol production rates, low current densities, and Faraday efficiencies. This fact results in high-energy requirements and therefore, CO<sub>2</sub> savings are restricted to the utilization of a low-carbon electricity source. This study aims to provide a comprehensive perspective on the technical prospects of ER technology to maximize the decarbonization/defossilization synergy in the European cement sector including the renewable requirement. The findings can promote the development of CR within the climate and defossilization transition.

## 2. Methodology

The present study aims to elucidate the environmental trade-offs for a feasible implementation of a CR strategy based on the electrochemical reduction (ER) of CO<sub>2</sub> into methanol within the European cement sector. The assessment is conducted in terms of carbon footprint (CF), fossil abiotic depletion (f-ADP) as well as renewable electricity requirements. The proposed CR scenario is compared with two scenarios, which are expected in a shorter timeframe. These are based on: i) FS; and ii) CS. As the objective is to emphasize the IS benefits involved in the co-production of cement and methanol in the CR scenario, the combined conventional production of both cement and methanol is used as a benchmark for comparison purposes. A Life Cycle Assessment (LCA) tool is applied according to international standards (“ISO 14040:2006,” 2006b; “ISO 14044:2006,” 2006). The LCA is conducted using the software GaBi v9.5 (Sphera) and the Ecoinvent database v3.8. The functional unit (FU) is set as 1 ton of cement plus 429 kg of methanol, which is the stoichiometric amount of methanol that can be converted from the direct emissions of cement (672 kg CO<sub>2</sub>/ton cement) by ER. Figure 1 shows the system boundary of the scenarios analyzed in the present study. The baseline scenario considers the production of 1 ton of cement and the conventional synthesis of 429 kg of methanol according to their Ecoinvent datasets. Briefly, the production of 1 ton of cement

involves the consumption of resources such as limestone, gypsum, metals, and fossil fuels including natural gas (0.158 m<sup>3</sup>), heavy fuel oil (23 kg), petcoke (3.53 kg) and hard coal (32 kg). The CO<sub>2</sub> direct emissions associated with the process account for 672 kg of CO<sub>2</sub>. The production of 429 kg of methanol involves mainly natural gas (280 m<sup>3</sup>), which is the main raw material, together with metals and water, being 2.8 GJ the heating needs that come from natural gas. The CR scenario captures CO<sub>2</sub> assuming membrane separation technology, which has a 90% of efficiency and an energy requirement of 300 kWh/ton CO<sub>2in</sub> (Giordano et al., 2018). The ER of CO<sub>2</sub> yields methanol by direct CO<sub>2</sub> reduction with water and electricity. ER process comprises 4 steps: 1) capture of the CO<sub>2</sub>; 2) electroreduction of CO<sub>2</sub> into methanol and subproducts (O<sub>2</sub> and H<sub>2</sub>); 3) recovery of unreacted CO<sub>2</sub>; and 4) distillation of methanol up to commercial grade. The inventory for the electroreduction is obtained by a mathematical process model described in previous studies (Rumayor et al., 2019, 2020). Given the low technology readiness level (TRL) of the ER route into methanol at its current stage, we have carried out a series of sensitivity analyses. The energy efficiency of the ER cell is ranged between 10% (current) and 90% (optimistic) while the methanol concentration at the outlet stream of the ER cell ranges between 10%wt. and 60%wt. The bivariate analyses show the influence in the categories of CF and f-ADP for the selected FU. It is expected a rapid installation of wind/PV solar in projected national plans as well as the economic incentives for carbon reduction in industries. Therefore, the analyses are conducted using two electricity impacts: i) EU-PV solar mix (2021); and ii) EU-wind-mix (2021). Their impacts may be representative of the environmental performance of renewable electricity in the medium and long-term timeframes, respectively.

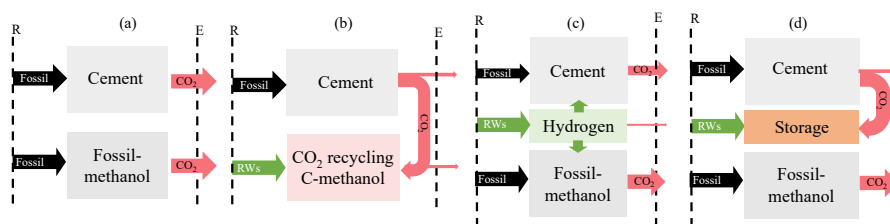


Figure 1. System boundary of a) Baseline; b) CR; c) FS; and d) CS

In the FS scenario, electricity is assumed renewable and the heat to produce 1 ton of cement (3.8 GJ) and 426 kg of methanol is considered to be supplied by 100% green H<sub>2</sub>. Green H<sub>2</sub> production is proposed to be synthesized by the polymer electrolyte membrane (PEM) technology, which is currently at the commercial stage. The electricity consumption of PEM is fixed at 50 kWh/kg H<sub>2</sub>, which is an optimistic value for the medium term. It must be considered that other resources such as limestone, gypsum, metals, and natural gas, which are used as raw materials, are the same that those considered in the Baseline. The CS scenario includes the capture of CO<sub>2</sub>, the compression and its sequestration. Energy consumption for the compression and injection is 117 kWh/t of CO<sub>2</sub> according to the literature (Koornneef et al., 2008). Note that the production of 429 kg of methanol will be kept as in the Baseline. Authors will display FS and CS in the result section only as benchmarking, being the prospective exploration of a combination of CCUS out of the scope of the present study because of the differences in TRL of the proposed technologies. Finally, it is calculated the resources and carbon emissions of an hypothetical decarbonized scenario. This scenario assumes the decarbonization of 10% of EU cement production by the proposed CR pathway. For a fair comparison, the impacts

are also calculated for the Baseline scenario (conventional). Therefore, the resources and impacts are calculated for 10% of the overall EU cement capacity (20 Mton/yr from the overall 200 Mton/yr). The decarbonized scenario is performed according to the figures of merits found in the sensitivity assessment as they will ensure the feasibility of the proposed CR pathway.

### 3. Results and discussion

Figure 2 displays the CF and f-ADP results obtained from the sensitivity analysis for the FU (1 ton of cement+429 kg of methanol) in the CR scenario. Since ER conversion in the CR scenario consumes a significant amount of electricity, the carbon intensity of the electricity to power the conversion process is critical. The CF and f-ADP results include the influence of the environmental impacts of renewable electricity to power the processes using the EU-PV solar mix (Figure 2 (a) and (c)) and the EU-Wind mix (Figure 2 (b) and (d)). There are also shown the CF and f-ADP of the CS (orange line) and FS (yellow line) scenarios as a function of the environmental impacts of renewable electricity. Furthermore, the baseline scenario is presented as a white dashed line and it is calculated keeping fixed the current EU grid mix. It should be mentioned that the contribution of the electricity to the overall CF value per FU (1120 kg CO<sub>2</sub>e) as well as to the overall f-ADP per FU (16500 MJ) in the Baseline are negligible. On one hand, direct CO<sub>2</sub> emissions are the main contributor to the overall CF value representing 85% according to the datasets. On the other hand, natural gas, which is the raw material in methanol production, is the main contributor to the overall f-ADP. From the CF perspective (Figure 2 (a) and (b)), the CR pathway exhibits a wide margin of opportunity to compete with the baseline scenario. CR may be competitive with the FS scenario using renewable electricity with a carbon intensity lower than 0.07 kgCO<sub>2</sub>e/kWh for methanol concentrations higher than 40%wt and energy efficiencies higher than 40%. However, the window of opportunity to compete with CS is narrower. The results indicate that CR may be competitive with CS only using renewable electricity with a CF lower than 0.01 kg CO<sub>2</sub>e/kWh and efficiencies higher than 50%. Note that the current energy efficiency of CO<sub>2</sub> ER is near 10-15%. As can be seen in Figures 2 (c) and (d), the benefits in the f-ADP are clearer. The results indicate a wider window of opportunity even under low-TRL performance of the ER technology compared with Baseline and CS scenarios. Note that 280 m<sup>3</sup> of natural gas is saved in the CR scenario as the amount of methanol (429 kg) is synthesized from the CO<sub>2</sub> direct emissions from the production of 1 ton of cement. According to the results, energy efficiencies higher than 40% and 50% and methanol concentrations higher than 40% could be a positive scenario to compete with CS and FS, respectively.

Figure 3 shows an overview of the resources, in terms of f-ADP, and the carbon emissions, in terms of CF, of the current cement production. It has also analyzed a hypothetical decarbonized scenario based on CR. The values shown in the figure correspond to the production of 10% of the overall EU cement capacity (20 Mton/yr from the overall 200 Mton/yr). The decarbonized scenario is performed according to the figures of merits found in the sensitivity analysis. It is assumed a methanol concentration of 40%wt. and a cell energy efficiency of 50%. EU-Wind-mix electricity impacts have been used to calculate the environmental impacts of the CR plant. According to the results, f-ADP could be reduced by a percentage of 70% with the CR strategy because of the resource savings (natural gas) within the conventional methanol production. On the other hand, CF is reduced by around 80% because CO<sub>2</sub> recycling substitutes natural gas in methanol production. Since ER is an electricity-intensive process, the decarbonization of

10% of EU cement production would require  $15 \cdot 10^{13}$  kWh of low-carbon renewable electricity.

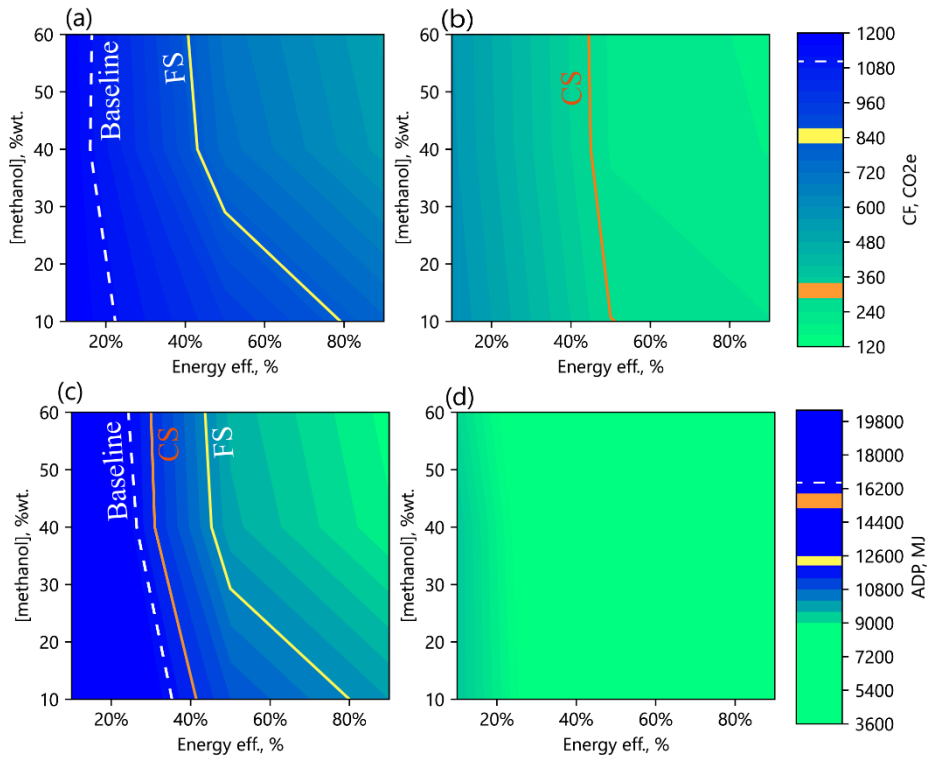


Figure 2. Influence in the CF and f-ADP (per FU) during the sensitivity analysis of CR scenario using electricity based on EU-PV-Mix (a, c) and EU-Wind-Mix (b,d)

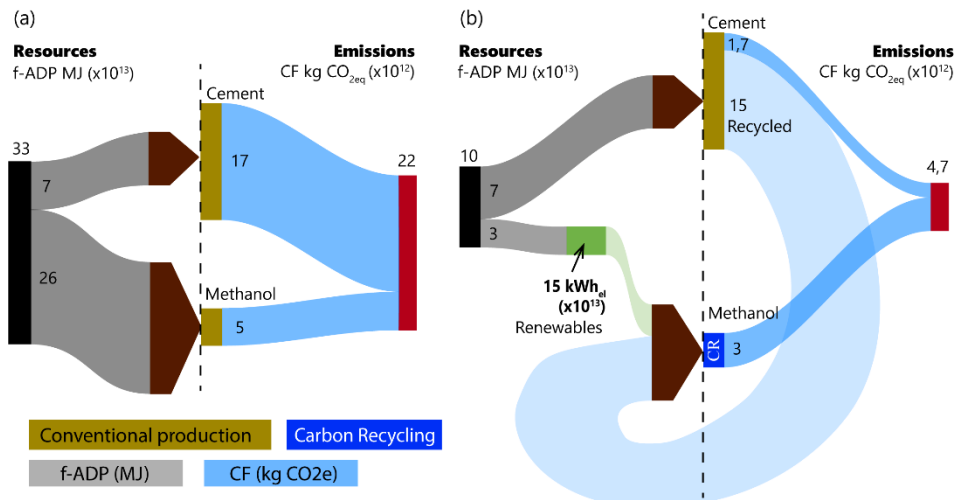


Figure 3. Environmental impacts (f-ADP and CF) of: (a) conventional production; and (b) decarbonized scenario considering 10% of European cement capacity

## 4. Conclusions

This work has provided the environmental prospects that CR strategy through ER could bring to the EU cement sector within the climate and defossilization transition. We have applied a prospective life cycle assessment tool to compare the CF and f-ADP of CR with FS and CS strategies. Since ER conversion is an electricity-intensive process, the environmental impacts of renewable electricity to power the conversion process is critical. In the nearer term, implementing CS seems more beneficial to reach the decarbonization goal of the cement sector. However, the CR strategy, which is an industrial symbiosis approach may boost the circular economy. The pathway involves a 70% of fossil resource savings and a reduction of 80% of carbon emissions.

## Acknowledgments

Authors thank the Spanish Ministry of Science and Innovation for the financial support through the project PID2020-112845RB-I00 funded by MCIN/ AEI /10.13039/501100011033.

## References

- Chauvy, R., & de Weireld, G. 2020. CO<sub>2</sub> Utilization Technologies in Europe: A Short Review. In *Energy Technology*, 8, 12
- Giordano, L., Roizard, D., & Favre, E. 2018. Life cycle assessment of post-combustion CO<sub>2</sub> capture: A comparison between membrane separation and chemical absorption processes. *International Journal of Greenhouse Gas Control*, 68, 146–163.
- International Energy Agency. 2018. *Technology Roadmap - Low-Carbon Transition in the Cement Industry*. [www.wbcsdcement.org](http://www.wbcsdcement.org).
- Koornneef, J., van Keulen, T., Faaij, A., & Turkenburg, W. 2008. Life cycle assessment of a pulverized coal power plant with post-combustion capture, transport and storage of CO<sub>2</sub>. *International Journal of Greenhouse Gas Control*, 2(4), 448–467.
- Olah, G. A. (2005). Beyond oil and gas: The methanol economy. *Angewandte Chemie - International Edition*, 44(18), 2636–2639.
- Roh, K., Bardow, A., Bardow, A., Bardow, A., Bardow, A., Bongartz, D., Burre, J., Chung, W., Deutz, S., Han, D., Heßelmann, M., Kohlhaas, Y., König, A., Lee, J. S., Meys, R., Völker, S., Wessling, M., Lee, J. H., Mitsos, A., Mitsos, A. 2020. Early-stage evaluation of emerging CO<sub>2</sub> utilization technologies at low technology readiness levels. *Green Chemistry*, 22(12), 3842–3859.
- Rumayor, M., Dominguez-Ramos, A., & Irabien, A. 2019. Innovative alternatives to methanol manufacture: Carbon footprint assessment. *Journal of Cleaner Production*, 225, 426–434.
- Rumayor, M., Dominguez-Ramos, A., & Irabien, A. 2020. Toward the Decarbonization of Hard-To-Abate Sectors: A Case Study of the Soda Ash Production. *ACS Sustainable Chemistry and Engineering*, 8(32), 11956–11966.
- Rumayor, M., Fernández-González, J., Domínguez-Ramos, A., & Irabien, A. 2022. Deep Decarbonization of the Cement Sector: A Prospective Environmental Assessment of CO<sub>2</sub> Recycling to Methanol. *ACS Sustainable Chemistry and Engineering*, 10(1), 267–278.

# Gray and hybrid green ammonia price sensitivity to market fluctuations: the Russia-Ukraine war case

Andrea Isella<sup>a</sup>, Alberto Lista<sup>a</sup>, Gabriele Colombo<sup>b</sup>, Raffaele Ostuni<sup>b</sup>,  
Davide Manca<sup>a\*</sup>

<sup>a</sup> *PSE-Lab, Process Systems Engineering Laboratory, Dipartimento di Chimica, Materiali e Ingegneria Chimica “Giulio Natta”, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy*

<sup>b</sup> *Casale SA, Via Giulio Pocobelli 6, 6900 Lugano, Switzerland*

\**davide.manca@polimi.it*

## Abstract

Ammonia synthesis is the biggest emitter of carbon dioxide among chemical processes. Emission regulations will become even stricter over the coming decades and call for innovative sustainable pathways to meet the decarbonization targets. An answer to such a transition may be retrofitting (*i.e.* hybridization) the existing plants by implementing low- or zero-emission technologies. One of the most promising retrofits is introducing green hydrogen from electrolyzers into the synthesis loop. This paper compares the operational expenditures (OPEX) of a conventional gray ammonia plant and a 10%-decarbonized electrolysis-based hybrid-green ammonia plant. The OPEX terms depend primarily on natural gas costs, renewable electric energy prices, and carbon tax values. Specifically, hybrid-green ammonia proves to be competitive at current EU natural gas quotations, which have increased more than five-fold since the first semester of 2021 due to the ongoing Russia-Ukraine war.

**Keywords:** decarbonization; hybridization; electrolysis; renewables; hydrogen.

## 1. Introduction

Ammonia synthesis is the most carbon-intensive process in the chemical industry, with emissions above 440 Mt<sub>CO<sub>2</sub>eq</sub>/y in 2020 (Isella and Manca, 2022). Indeed, current ammonia synthesis heavily relies on fossil fuels to produce the hydrogen needed in the ammonia converters and generate the heat and compression required for the process. This explains the blooming of decarbonization strategies to mitigate the carbon footprint of conventional “gray” ammonia. Among them, we have: (i) blue ammonia, which implements various solutions of carbon capture and utilization or storage (CCUS); (ii) green ammonia, which fully decouples the ammonia production from fossil feedstocks (*e.g.*, hydrogen production through an electrolyzer powered by renewable electricity); (iii) hybrid-green ammonia, which consists in revamping the existing conventional plants by installing an electrolyzer to replace part of the gray, fossil-based hydrogen feedstock with a green, water-based one (MPP, 2022). Concerning the latter technology, it is worth clarifying that within the perspective of a fully decarbonized ammonia scenario, such a production method should be considered a transitional technology due to its reduced CO<sub>2</sub> abatement potential, which is partial by nature. Nonetheless, until near-zero-emissions paths become commercially ready in the following decades, an available-to-date alternative such as hybrid-green ammonia can be the best way to meet emissions



reduction in ammonia synthesis. The main purpose of the paper is to identify the price volatility of hybrid-green ammonia. Specifically, its market trends are reproduced and compared with the far more widely-known and better-characterized conventional ammonia prices. Indeed, the cost of hybrid-green ammonia is shown to be, just as its gray counterpart, a function of several market variables (*e.g.*, the price of fossil fuels and electrical energy). To our knowledge, this assessment is unprecedented in the scientific literature and represents a first attempt to formally describe the market evolution of hybrid-green ammonia. Moreover, the global geopolitical conditions given by current events (*in primis*, the Russia-Ukraine war), add even more novelty to our study.

## 2. Simulation of the hybrid-green ammonia plant

A medium-high capacity 2000 t/d stand-alone natural gas-based ammonia plant was simulated in UniSim® Design R491 to assess its economic hybridization potential. The decarbonization target was set to 10% (a distinctive value of such plants resulting mainly from design considerations, as discussed in the following). The injection of green hydrogen from an alkaline electrolyzer (AEL) powered by renewable electric energy can achieve that target. Indeed, this strategy aims to reduce the “front-end” (*i.e.* the whole of the reforming, shift, and syngas purification sections) load and, consequently, to lower both the natural gas fuel and feed demands of the process (see Figure 1).

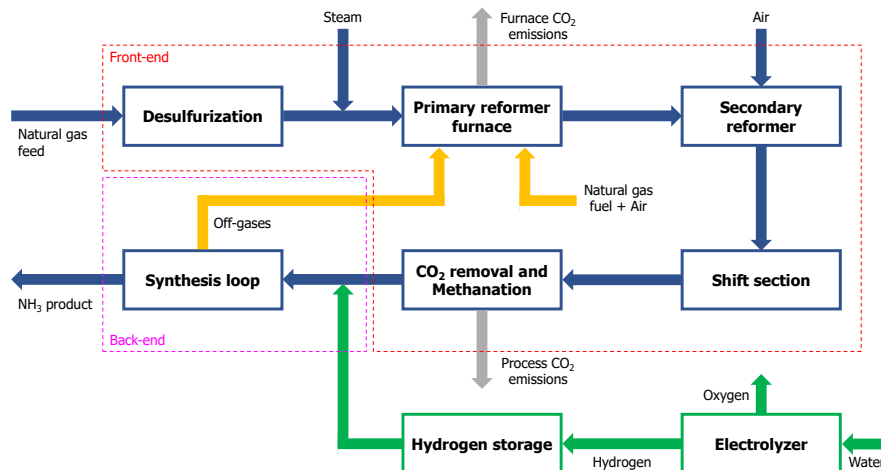


Figure 1: Block flow diagram of the simulated hybrid-green ammonia plant.

An equivalent amount of hydrogen fed to the ammonia synthesis loop must be generated by electrolysis, assuming a 10% average reduction in carbon dioxide emissions ( $\text{DeCO}_2 = 10\%$ ). However, the hydrogen production capacity of the electrolyzer is strongly affected by the fluctuating trends of renewable electric energy profiles. Indeed, renewable energy production is firmly susceptible to the varying availability of energy sources (such as sunlight and wind in the case of solar and wind farms). Since the electrolyzer converts such fluctuating electricity profiles into hydrogen flows that are thus intermittent, it follows that the target green hydrogen productivity should be considered as an average of the oscillations that span from a minimum of zero (when no green electricity is available) to a maximum, which is strictly related to the capacity factor of the renewable power station (that determines the peak power available to the user) and also to some design considerations, precisely the reformers' load. Indeed, considering to keep constant the ammonia productivity of the plant, the injection of an excessive amount

of green hydrogen into the synthesis loop results in substantial drops in the front-end flow rates (*i.e.* the ones linked to the conventional gray route), which may affect a few critical aspects, such as the operability of the fired tubes in the primary reformer and the rise in the temperature of the secondary reformer. To tackle this, hydrogen storage is needed downstream of the electrolyzer to set the maximum amount of green hydrogen fed to the synthesis loop. Specifically, it was assumed that the upper limit of green hydrogen input should be such to require a decrease in the front-end “throughput” (*i.e.* its total processed flow rate) to 80% maximum turndown to ensure reliable operative conditions of the primary reformer. Respecting such a directive leads to an instantaneous decarbonization degree of about 20% ( $\text{DeCO}_2 = 20\%$ ) at the highest. This is a somewhat limiting constraint since, by ignoring all the above design recommendations and basing the calculations exclusively on the peak availability of green electricity (which depends only on the capacity factor of the renewable power source), the maximum achievable decarbonization degree could be, in principle, higher.

Table 1 reports the leading process variables of the simulated hybrid-green ammonia plant in its three main operating conditions: S0, which corresponds to the minimum (*i.e.* null) green hydrogen availability; S1, which corresponds to the average working conditions (which should guarantee a 10% decarbonization target, as requested); and S2, which corresponds to the upper limit of green hydrogen input.

Table 1: Process variables of the simulated hybrid-green plant in its three main operative conditions. The electrolyzer has a specific electricity consumption of 4.5 kWh/Nm<sup>3</sup>H<sub>2</sub>, a specific water consumption of 0.9 kgH<sub>2</sub>O/Nm<sup>3</sup>H<sub>2</sub>, and a hydrogen output temperature and pressure of 35 °C and 30 bar, respectively (Nel Hydrogen, 2021). Note that a natural gas-to-electricity conversion factor of 3 MWh<sub>LHV</sub>/MWh<sub>el</sub> was used to evaluate the net plant energy consumption.

<i>Process variable</i>		<i>Status “S0”</i>	<i>Status “S1”</i>	<i>Status “S2”</i>
		(0% DeCO <sub>2</sub> )	(10% DeCO <sub>2</sub> )	(20% DeCO <sub>2</sub> )
<i>Primary reformer throughput</i>	[t/h]	163	146	129
<i>Natural gas feed to the front-end</i>	[t/h]	38.0	33.9	29.9
<i>Natural gas fuel to the furnace</i>	[t/h]	12.9	11.9	10.7
<i>Green hydrogen production</i>	[t/h]	0	1.74	3.55
<i>Green electricity consumption</i>	[MW]	0	86.8	177
<i>Net plant energy consumption</i>	[MWh <sub>LHV</sub> /tNH <sub>3</sub> ]	8.14	10.4	12.9
<i>Furnace CO<sub>2</sub> emissions</i>	[t/h]	38.9	34.8	30.7
	[tCO <sub>2</sub> /tNH <sub>3</sub> ]	0.467	0.418	0.368
<i>Process CO<sub>2</sub> emissions</i>	[t/h]	97.4	87.9	78.1
	[tCO <sub>2</sub> /tNH <sub>3</sub> ]	1.17	1.06	0.937
<i>Total CO<sub>2</sub> emissions</i>	[t/h]	136.3	122.7	108.8
	[tCO <sub>2</sub> /tNH <sub>3</sub> ]	1.64	1.48	1.31

As expected, an increase in the green hydrogen input to the ammonia synthesis loop leads to a significant decrease in both the furnace and process carbon dioxide emissions at the

expense of the net energy consumption of the plant. Indeed, as the heating demand associated with natural gas streams declines, electricity demand rises significantly. However, also other implications occur, namely: (i) the increase in the operating temperature of the secondary reformer (from 996 °C in S0 to 1043 °C in S1 and 1104 °C in S2) and (ii) the drop in the methane slip (from 0.244 mol% in S0 to 0.098 mol% in S1 and 0.032 mol% in S2) and the inert gases content (from 11.7 vol% in S0 to 8.92 vol% in S1 and 7.28 vol% in S2) in the ammonia synthesis loop. Consequence (i) is mainly due to the choice of keeping the process air flowrate constant (to meet the nitrogen demand of the plant and guarantee the stoichiometric  $H_2/N_2$  ratio at the inlet of the first ammonia converter). This leads to an increase in the oxygen-to-carbon molar ratio at the secondary reformer inlet (from 0.757 in S0 to 0.848 in S1 and 0.963 in S2) and, therefore, in its operative temperature. Concerning consequence (ii), instead, it is strictly related to the first one, (i). Indeed, such a result is due to the higher temperatures in the secondary reformer, which shift the steam reforming reaction towards the products. Finally, it is worth noting that while consequence (i) is very detrimental, as excessive overheating of the secondary reformer and the downstream process boiler must be avoided, consequence (ii) is beneficial to the process as it reduces compression costs and improves the reaction thermodynamics and kinetics in the ammonia converters.

### 3. Gray and hybrid-green ammonia production costs

Through the previous simulations, hybrid-green ammonia proves to be an available-to-date, feasible (although conspicuously electricity-demanding) technique to decarbonize ammonia synthesis. However, the actual testing benchmark to assess the competitiveness of an alternative production route to the conventional one is the cost of the final product. Indeed, since the cheapest process is expected to be the most appealing, the present section focuses on evaluating both gray and hybrid-green ammonia production costs. Precisely, the following assumptions hold:

- both facilities operate 8000 h/y;
- the hybrid-green plant has an average annual decarbonization target of 10%;
- the costs for cooling and demineralized water are 0.04 and 0.75 €/m<sup>3</sup>, respectively (Intratec, 2022);
- the revenues from the medium-pressure steam production within both plants are estimated by considering the natural gas duty that an auxiliary boiler would need for the same task (Noelker and Ruether, 2011). Specifically, using pure water at 35 °C and 40 bar leads to a steam energy export equal to 0.13 MWh<sub>LHV</sub>/t<sub>NH<sub>3</sub></sub> from both facilities;
- the decarbonization degree of the hybrid-green plant is assumed to fluctuate in the 0-20% interval, depending on the instantaneous availability of renewable electricity;
- the installed electrolyzer capacity must be 177 MW to meet the hydrogen demand at the peak of green electricity availability (see Table 1, S2). On top of this, such power capacity value is oversized by 10%;
- the electrolyzer capital expenditures (CAPEX) must be accounted for. Indeed, since the electrolyzer is the most expensive unit that the hybridization retrofit calls for, its cost affects the price of the final product. Specifically, the electrolyzer exhibits a 10-year depreciation time and a fixed fee of 700 €/kW<sub>el</sub> (including engineering, housing, and balance-of-plant components such as the compression and the water and hydrogen purification units). This is a rather low (yet conservative) value compared to the average ones reported in the literature (which typically span from 800 to 1000 €/kW<sub>el</sub>).

Indeed, it is recognized that the economy of scale notably affects AEL modules with an installed power capacity higher than 100 MW (Fraunhofer ISE, 2021). It is worth noting that the main contributions to the final prices of both gray and hybrid-green ammonia are given by natural gas, carbon tax, and (only for the hybrid plant) renewable electricity, whose values are continuously changing at rather high rates. This aspect is very significant since specific fluctuations of such variables affect the costs of both production pathways, eventually overturning their mutual relationship. For example, decreasing renewable power purchase agreement (PPA) costs and increasing values of carbon taxes are expected to promote the shift to environmentally mitigating technologies despite the conventional, carbon-intensive (and otherwise far cheaper) ones. Anyway, the aspect that has predominantly affected the recent evolution of the price trends in ammonia production is the runaway behavior of natural gas quotations. Indeed, with the outbreak of the Russia-Ukraine war, its already rising trend, which started in the summer of 2020, reached extraordinary peaks in a few weeks. Specifically, the remarkable surge in natural gas prices made the hybrid-green ammonia production pathway cheaper than the conventional gray one. Indeed, the reduced fossil fuel consumption in the hybridized facility makes it less susceptible to changes in natural gas quotations.

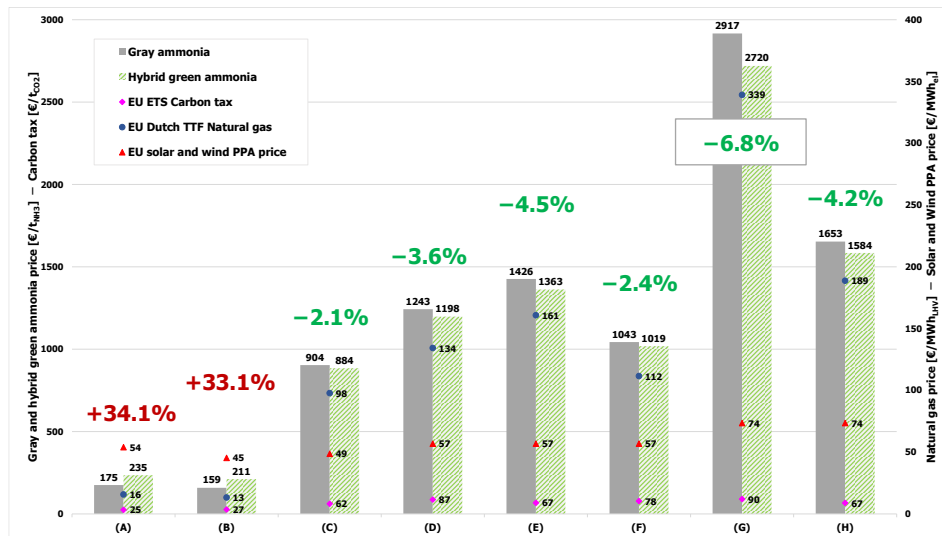


Figure 2: Gray and hybrid-green ammonia production costs and carbon tax values (left y-axis); natural gas and renewable PPA prices (right y-axis) for the A-H scenarios. The percentages at the top of the bars quantify the difference between the hybrid-green vs. gray ammonia total production costs. Input data from Trading Economics (2022), for natural gas and carbon tax, and LevelTen Energy (2022), for PPAs.

To follow the ever-changing panorama of incidents that significantly affected the main prices/costs of both process typologies, we identified eight scenarios that occurred in the last three years (see also Figure 2):

- A. Pre Covid-19 era (30-Sep-19);
- B. After 6 months of Covid-19 (30-Sep-20);
- C. After 18 months of Covid-19 (30-Sep-21);
- D. Start of Russia-Ukraine war (24-Feb-22);
- E. One week into the conflict (03-Mar-22);
- F. One month into the conflict (24-Mar-22);

- G. Highest peak of natural gas price (26-Aug-22);
- H. First EU protective measures (30-Sep-22).

Figure 2 shows the economic assessment, where the prices of gray and hybrid-green ammonia were considered by utilizing as input data the values of natural gas price, renewable PPA price, and carbon tax for the European market at the A-H scenarios. As previously mentioned, the sensitivity of ammonia production prices to market fluctuations is so elevated that the remarkable increase in natural gas quotations has made hybrid-green ammonia cheaper than gray ammonia in the recent C-H scenarios (*i.e.* since 30-Sep-2021). Indeed, hybrid-green ammonia has gradually become cheaper than gray ammonia, despite starting from a 34% higher price than the conventional route in the pre Covid-19 era. This is still true after the highest recorded peak in natural gas quotations (on 26-Aug-2022), with a discounted price of 4.2% to gray ammonia.

#### 4. Conclusions

This paper introduced a novel method to evaluate ammonia market price, aimed at discussing the economic attractiveness of hybrid-green vs. gray ammonia production. The simulation of a 2000 t/d hybridized ammonia plant highlighted the main problems related to the operating limits of that facility. These are the significant demand for renewable power and operability issues in the reformers as far as excessively high decarbonization degrees are concerned. Regarding the economic assessment, the production costs of gray and hybrid-green ammonia proved particularly sensitive to market fluctuations, primarily to natural gas quotations. The steep increase in the natural gas price due to the outbreak of the Russia-Ukraine war made the hybrid-green production route cheaper than the gray one, although being far more expensive just a few months before. This outcome matches the increased adoption of electrified technologies to decarbonize ammonia synthesis.

#### References

- Fraunhofer ISE (2021). Cost Forecast for Low-Temperature Electrolysis. Fraunhofer Institute for Solar Energy Systems ISE.
- Intratec (2022). Cooling Water Current Costs, Historical Series & Forecasts. Available at: <https://www.intratec.us/products/water-utility-cost-database/commodity/cooling-water-cost> (accessed on November 17<sup>th</sup>, 2022).
- Intratec (2022). Demineralized Water Current Costs, Historical Series & Forecasts. Available at: <https://www.intratec.us/products/water-utility-cost-database/commodity/demineralized-water-cost> (accessed on November 17<sup>th</sup>, 2022).
- Isella, A., & Manca, D. (2022). GHG Emissions by (Petro)Chemical Processes and Decarbonization Priorities—A Review. *Energies*, 15(20), 7560.
- LevelTen Energy (2022). Europe's Renewable Energy Shortage Causes PPA Prices to Rise More Than 11% in Q3 and 51% Year Over Year, According to LevelTen Energy. Available at: <https://www.leveltenenergy.com/post/2022q3-ppa-europe> (accessed on November 11<sup>th</sup>, 2022).
- MPP (2022). Making Net-Zero 1.5°C-Aligned Ammonia Possible. Mission Possible Partnership.
- Nel Hydrogen (2021). Nel Hydrogen Electrolyzers The World's Most Efficient and Reliable Electrolysers. Available at: <https://nelhydrogen.com/wp-content/uploads/2020/03/Electrolysers-Brochure-Rev-D.pdf> (accessed on October 30<sup>th</sup>, 2022).
- Noelker, K., & Ruether, J. (2011). Low Energy Consumption Ammonia Production: Baseline Energy Consumption, Options for Energy Optimization. Presented at the Nitrogen & Syngas 2011 International Conference & Exhibition, Duesseldorf, Germany.
- Trading Economics (2022). EU Carbon Permits. Available at: <https://tradingeconomics.com/commodity/carbon> (accessed on November 11<sup>th</sup>, 2022).
- Trading Economics (2022). EU Natural Gas. Available at: <https://tradingeconomics.com/commodity/natural-gas> (accessed on November 11<sup>th</sup>, 2022).

# Paving the way for the integration of synthesis, assessment, and design tools within an ontological framework

Adrián Pacheco-López<sup>a,b</sup>, Kristiano Prifti<sup>a,b</sup>, Flavio Manenti<sup>b</sup>, Ana Somoza-Tornos<sup>c</sup>, Moisès Graells<sup>a</sup>, Antonio Espuña<sup>a\*</sup>

<sup>a</sup> *Department of Chemical Engineering, Universitat Politècnica de Catalunya, Escola d'Enginyeria de Barcelona Est, C/ Eduard Maristany 16, Barcelona 08019, Spain*

<sup>b</sup> *CMIC Department "Giulio Natta", Politecnico di Milano, Piazza Leonardo da Vinci 32, Milan 20133, Italy*

<sup>c</sup> *Department of Chemical Engineering, Delft University of Technology, Van der Maasweg 9, 2629 HZ Delft, Netherlands*

*\*[antonio.espuna@upc.edu](mailto:antonio.espuna@upc.edu)*

## Abstract

The constant development of new alternatives to treat waste aids in closing material loops towards the circular economy and improving sustainability through the use of new renewable materials and energy. This fact leads to the increasing need for decision-making tools for process synthesis and assessment, which can be addressed with an integrated framework that employs ontologies for knowledge management and optimization tools to perform a hierarchical assessment of alternatives. The systematization of these procedures raises the need for tools to automate techno-economic and life cycle analyses. In this work, such a challenge is addressed through the additional integration of add-on modules such as the CapEx-Opex estimation tools and surrogate modeling within this framework. A case study on plastic waste is proposed with the inclusion of several pyrolysis and gasification alternatives. Results show pyrolysis, followed by the subsequent purification of its products, as the best alternative and helped identify main drivers for technologies feasibility such as feedstock purity and energy consumption.

**Keywords:** circular economy, integrated modeling, sustainable development, machine learning, economic optimization.

## 1. Introduction

The Circular Economy appears as a subject of paramount importance toward economically, environmentally, and socially sustainable development. Thus, many entities across the world are working hard to find alternatives that not only bring economic profit but also are environmentally benign and socially responsible. One of the most urgent matters is waste management and, in this specific line, many technical efforts are currently devoted to finding new waste-to-resource processes, recovering valuable resources, and closing material loops. Consequently, decision-makers face the challenge to determine which alternatives are most suitable for each kind of waste. To address this task, Pacheco-López et al., (2022) developed a framework of integrated tools to connect waste sources with valuable products and find the best routes to close material loops. The

application of this framework requires performing economic and environmental analyses and developing models suitable for each one of the alternatives included. To undertake these tasks more systematically, this contribution aims to integrate new tools to make the framework more versatile and minimize the required human intervention for synthesizing and assessing new sustainable approaches. Within the proposed framework, the ontology is used to manage the information and data that are needed as inputs on each module, as well as storing the outputs, therefore centralizing all the knowledge required and produced in the system (orange arrows represent information exchange in Figure 1).

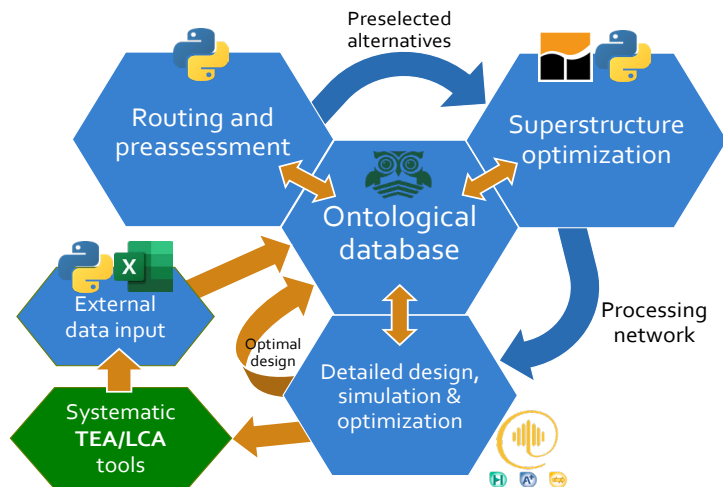


Figure 1. Schematic representation of the proposed framework with an additional module (green).

## 2. Methodology

The methodology used in this contribution is an extended version of a previously proposed one, in which several synthesis, assessment, and optimization tools were integrated to build a decision-making framework (Pacheco-López et al., 2022). There, a sequential approach is used to build alternative paths between waste materials and marketable products and to assess them from several points of view as well as to find the optimal configuration. First stage uses a new proposed algorithm and metric based on different objectives to pre-select alternatives, then a second stage consisting on a Multi-objective network optimization to find best routes based on each objective (for more details of this procedure see (Somoza-Tornos et al., 2021)). One of the identified challenges in this work was the need to perform previous techno-economic (TEA) and life cycle (LCA) assessments manually for every single process in order to introduce their results in the framework to be used by the other modules. To address this challenge, the need for systematization tools was identified to reduce human intervention and standardize the results. In this direction, some tools were already developed and are being implemented in the framework. Prifti et al. (2022) developed a CapEx-OpEx Robust Optimizer (CORO) framework that was intended to assess the cost evaluation of a plant including some sustainability targets, thus, it provides the framework with more dynamism and systematization in the TEA tasks. CORO automatically takes a simulation file (Hysys or ProII) as an input and generates the capital and operational expenditures for each process. Using this tool, the cost estimations for all processes in the ontology have been performed in order to obtain standardized results. Figure 1 shows the original methodology and the proposed additional module for techno-economic/life-cycle systematic assessment tools.

As of now, only systematic TEA is implemented for illustrative purposes, although future developments are intended to include LCA systematization in the same way.

On the other hand, process simulation is also a critical step for the design and optimization of new processes, which, in turn, may require considerable computational effort. Surrogate modeling offers the possibility to speed up simulation convergence times by substituting rigorous models with machine learning methods (Granacher et al., 2021), as the one proposed by Galeazzi et al., (2022) to predict the behavior of several process variables of existing steady-state digital twins of industrial plants. This tool can be accessorially used in the third stage of the methodology to ease the simulation step or to study the processes' sensitivity to variations in feed composition or operating conditions.

Table 1. Deployed process paths for MPW treatment, outputs, and GPI for 24 alternatives. Sor.: Sorting; FBR: Fluidized bed reactor; HGHTR: horizontal gas heated tube reactor; ZSM5: ZSM-5 zeolite; FCC: fluid catalytic cracking; R1: FCC-R1 commercial FCC equilibrium catalysts; HUSY: Ultrastabilized Y zeolite; SAHA: amorphous silica-alumina; HZSM: acid zeolite; SA9Z1: hybrid catalyst (SA: Zeolite, 9:1); LPG: C1-C3 aliphatics.

	Processes	Outputs	GPI
1-3	Pyrolysis 500°C R1/Red Mud/ZSM5/	C1-C5 alkanes, ethylene, H2, hexene, toluene, ethylbenzene, styrene, naphthalene, xylene & C9-C14	723-642
4	Pyrolysis 600°C /Y zeolite/	Methane, ethylene, propene, ethane, butane, butane, hydrogen, benzene, toluene, ethylbenzene & styrene	558
5-8	Acid FCC 390°C /HUSY//SAHA/R1/ZSM5/	Gasoline, LPG, light aromatics mixture, C9-C14 mixture, char & HCl	435-425
9	Gasiforming	Methanol & CO <sub>2</sub> credits	288
10	Sor. + PE pyrolysis 520°C HGHTR /ZSM5/	Gasoline, ethane, propane, butane & char	216
11	Sor. + PE pyrolysis 740°C	LPG, ethylene, benzene, toluene, indane & pyrene	183
12-13	Sor. + PE pyrolysis FBR 375°C /HZSM/SA9Z1/	Gasoline, ethane, propane, butane & char	151-116
14	Sor. + PP pyrolysis 350°C	WPPO (diesel substitute) & char	96
15	Sor. + PE pyrolysis 1000°C	Methane, ethylene, propene, butadiene & benzene	87
16	Sor. + PP pyrolysis 760°C	LPG, ethylene, benzene, toluene & naphthalene	79
17	Idem + Ethylene hydration	LPG, ethanol, benzene, toluene & naphthalene	70
18	Co-gasification 850°C	Methane, ethane, syngas & char	63
19	Gasification 850°C	Methane, ethane, syngas & char	59
20	Electrified Gasiforming	Methanol & CO <sub>2</sub> credits	54
21	Sor. + PS pyrolysis 425°C	Toluene, styrene, cumene & 1,3,5-triphenylbenzene	42
22	Sor. + PE pyrolysis 550°C	WPPO (waste plastic pyrolysis oil – diesel substitute)	35
23	Incineration	Energy Recovery	0
24	Landfill	None	0

### 3. Case Study

The framework is tested on a case study dealing with the recovery of plastic waste materials and the conversion/purification of the resulting products into other valuable carbon-based products. New alternatives are added to the set of alternatives available in the ontology from previous implementations using the same assumptions as presented in Pacheco-López et al., (2022). These new processes consist of gasification followed by methane reforming and conversion to obtain methanol as a final product, either using



traditional energy sources or electrified ones (Prifti et al., 2021). A total of 66 processes (see Figure 2) are implemented in this case, where 200 paths were created and assessed, from which 24 were selected in the first stage (see Table 1).

The preselected alternatives were sent to the multi-objective network optimization module obtaining a set of configurations along with the representation with the corresponding Pareto fronts. The newly added alternatives are compared against those already available in the ontology. The feedstock considered corresponds with a mixture of plastic waste with the following mass composition: 45% polyethylene, 32% polypropylene, 20% polystyrene, and 3% of PVC. A plant to treat an annual amount of 20,000 tonnes of MPW (2.5 tonnes/hour for a plant operating 8000h a year) was considered and all processes were scaled accordingly.

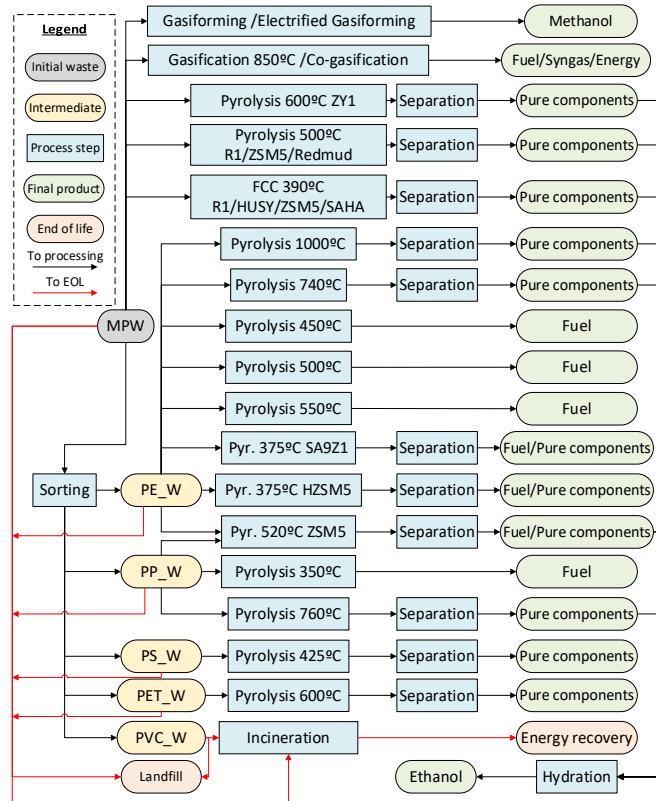


Figure 2. Implicitly generated graph in the pre-assessment stage with all tentative connections. PE: polyethylene, PP: polypropylene, PS: polystyrene, PET: polyethylene terephthalate, PVC: polyvinyl chloride; Pyr.: Pyrolysis; EOL: End of life.

#### 4. Results

After updating the ontology with all the current and new process parameters, the algorithm is run and all the possible connections are made with an input-output matching procedure, resulting in a graph as seen in Figure 2. With the assessment procedure presented in Pacheco-López et al., (2021), all the paths are built and a list of prioritized alternatives is obtained and preselected according to their global performance indicator (see Table 1, for brevity similar routes have been combined). This prioritized list of alternatives is passed to the network optimization stage where a multi-objective

optimization is performed. A set of different process configurations is obtained according to different objectives and intermediate alternatives as well as a set of Pareto points using the  $\epsilon$ -constraint method. As an example, the configuration with the best economic performance is shown in Figure 3 along with the Pareto fronts obtained for the economic profit against each one of the environmental endpoint indicators as shown in Figure 4.

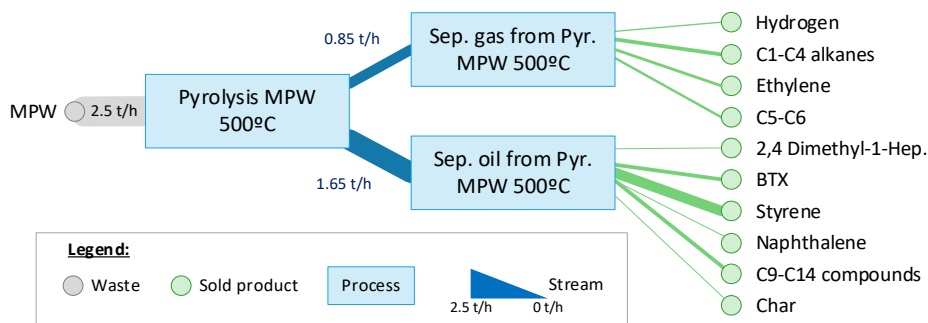


Figure 3. Configuration obtained for maximized profit. Point number 10 in Figure 4.

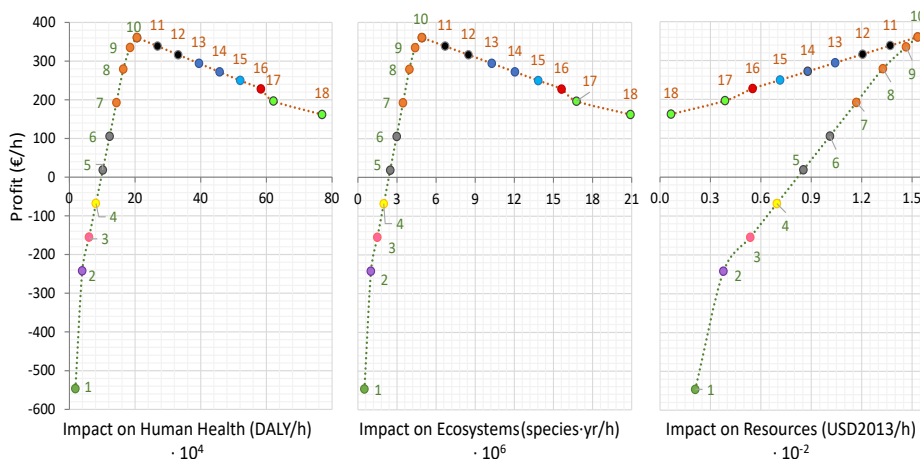


Figure 4. Pareto fronts for the trade-off between profit and the three environmental endpoint indicators. The points are numbered and color-coded to represent different configurations. Each point (as the one seen in Figure 3) consists of a different combination of processes from Figure 2.

Results in this contribution were consistent with those obtained in previous runs of the algorithm (Pacheco-López et al., 2022, 2021). Newly introduced gasification alternatives were selected in the first stage in the positions 9<sup>th</sup>, for the traditionally powered route, or Gasiforming<sup>TM</sup>, and 20<sup>th</sup>, for the Electrified Gasiforming<sup>TM</sup>. In addition, they were not among the best-performing alternatives in the second stage for any of the four obtained anchor points, which may suggest they need further development and more efficient and economically competitive energy sources. It was identified that capital expenditures play an important role in the implementation of these technologies. Electrical energy consumption has been also identified as a critical part since it entails most of the operating expenditures. On the other hand, results show that pyrolysis of mixed plastic waste with or without catalysts at a temperature of 500°C is still among the best-performing alternatives for most objectives since they are the most economically and environmentally competitive or beneficial options for this kind of waste.

## 5. Conclusions

Results show the framework's ability to synthesize and assess routes for the upcycling of waste accounting for economic and environmental objectives. The capability of the framework to be easily expanded and complemented with the integration of other tools to satisfy processes modeling, design, and assessment needs has also been proven. The use of the CapEx-OpEx Robust Optimizer (CORO) has aided in systematically generating the economic performance indicators for all the processes asserted in the ontology, providing a standardized procedure to avoid user-introduced errors and keeping the assessment parameters consistent among all alternatives. The third stage of the methodology, detailed simulation, and design of a chosen alternative is one of the most work-intensive tasks of the framework and stays out of the scope of this study. The use of surrogate modeling has been envisaged as a useful alternative to ease this stage and might be implemented in future developments to study the effect of feedstock composition changes on the economic and environmental performance of the upcycling technologies. Some identified challenges have been approached with the inclusion of new ideas and tools that help advance the automation and systematization of the framework, mainly in terms of TEA/LCA and simulation efforts. The possible inclusion of carbon emission credits has also been included as a potential source of income, therefore improving the overall economic performance of the considered alternatives.

## 6. Acknowledgments

Financial support received from the Spanish "Ministerio de Economía, Industria y Competitividad" and the European Regional Development Fund, both funding the research Project CEPI (PID2020-116051RB-I00) is fully acknowledged. A.P-L thanks the financial support received from the "Ministerio de Ciencia e Innovación" (grant ref. PRE2018-087135) and the Erasmus+ programme of the EU 2021-27 (KA131 call).

## References

- Galeazzi, A., Prifti, K., Gallo, F., Manenti, F., 2022. A Methodology for The Optimal Surrogate Modelling of Digital Twins Using Machine Learning. *Comput. Aided Chem. Eng.* 51, 1543–1548. <https://doi.org/10.1016/B978-0-323-95879-0.50258-7>
- Granacher, J., Kantor, I.D., Lopez, M., Maréchal, F., 2021. Self-learning surrogate models in superstructure optimization, in: *Computer Aided Chemical Engineering*. Elsevier, pp. 439–444. <https://doi.org/10.1016/B978-0-323-88506-5.50069-3>
- Pacheco-López, A., Somoza-Tornos, A., Graells, M., Espuña, A., 2022. Integrated synthesis, modeling, and assessment of waste-to-resource alternatives, in: *Computer Aided Chemical Engineering*. Elsevier, pp. 787–792. <https://doi.org/10.1016/B978-0-323-95879-0.50132-6>
- Pacheco-López, A., Somoza-Tornos, A., Graells, M., Espuña, A., 2021. Synthesis and assessment of waste-to-resource routes for circular economy. *Comput. Chem. Eng.* 153. <https://doi.org/10.1016/j.compchemeng.2021.107439>
- Prifti, K., Galeazzi, A., Barbieri, M., Manenti, F., 2022. A Capex Opex Simultaneous Robust Optimizer: Process Simulation-based Generalized Framework for Reliable Economic Estimations, in: *Computer Aided Chemical Engineering*. Elsevier, pp. 1321–1326. <https://doi.org/10.1016/B978-0-323-95879-0.50221-6>
- Prifti, K., Galeazzi, A., Margarita, I., Papale, A., Miele, S., Bargiacchi, E., Barbieri, M., Petea, M., Manenti, F., 2021. Converting end-of-life plastic waste into methanol: the gasformingtm process as new, efficient and circular pathway. *Environ. Eng. Manag. J.* 9, 1629–1636.
- Somoza-Tornos, A., Pozo, C., Graells, M., Espuña, A., Puigjaner, L., 2021. Process screening framework for the synthesis of process networks from a circular economy perspective. *Resour. Conserv. Recycl.* 164, 105147. <https://doi.org/10.1016/j.resconrec.2020.105147>

# Pyrolysis of livestock manures: Optimal operating conditions and feedstock blending ratios

Prakash Parthasarathy, Mohammad Alherbawi, Snigdhendubala Pradhan,  
Gordon McKay, Tareq Al-Ansari \*

*Division of Sustainable Development, College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar.*

*\*[talansari@hbku.edu.qa](mailto:talansari@hbku.edu.qa)*

## Abstract

Livestock manure contributes to around 18% of global greenhouse gas (GHG) emissions, while the mishandling of manure further pollutes water and soil. As such, the valorisation of these wastes becomes vital for the mitigation of global warming impact. In this context, anaerobic digestion (AD) is commonly used to generate biogas and digestate from manure; however, pyrolysis has recently emerged as a viable alternative to AD for producing cleaner products in much shorter processing times. Therefore, this study investigates the pyrolysis of four types of livestock manures including poultry (PM), dairy (DM), camel (CM) and sheep manure (SM). The air-dried manure samples are characterised with reference to literature reported data, while the process is simulated using existing empirical prediction models. In addition, the process performance in terms of economic feasibility and environmental impact is evaluated from cradle to gate for different manure feedstocks. A mathematical optimisation model is then developed to evaluate optimal blending of manure for the commercial production of bio-oil, considering different operating conditions, manure compositions and availability in the State of Qatar. The model is solved to fulfill seven different objectives independently. The results revealed that PM is linked to the highest amount of bio-oil (41%) at 600°C, while the lowest yield is associated to CM. Finally, the optimal manure blending ratio for highest bio-oil yield and lowest cost is achieved at 52% PM: 2% DM: 3% CM: 43% SM. Whereas the optimal blending ratio for lowest carbon footprint is realised at 42% PM: 21% DM: 30% CM: 7% SM. The designed model provides a thorough insight on optimal technical and environmental valorisation of livestock manures by pyrolysis.

**Keywords:** Pyrolysis; Livestock manures; Optimisation; GHG; Bio-oil; Biochar.

## 1. Introduction

Several billion tonnes of livestock waste/manure are produced worldwide every year. The mismanagement and misuse of manure is severely polluting the environment and contaminating the water supply. As a matter of fact, the livestock sector is estimated to contribute 18% of total Greenhouse gas (GHG) emissions globally (Moran and Wall, 2011). Biomass, including livestock manure, has piqued the interest of many because its carbon neutrality offers a strategic measure to reduce GHG emissions. Biological treatment of livestock manures, such as anaerobic digestion (AD), has traditionally been used. The use of pyrolysis method to convert livestock manure into energy has been emerging as a viable alternative to AD because they can be used to produce value-added

products in a much shorter timeframe than AD. Pyrolysis also significantly reduces waste volume and eliminates pathogens.

This study considers four types of livestock manures: poultry (PM), dairy (DM), camel (CM), and sheep (SM). Due to the large numbers of aforementioned livestock raised in the State of Qatar, these manures were considered. The study also attempts to find an optimal blending ratio of the aforementioned manures for commercial production of bio-oil, considering different operating conditions, manure compositions, and availability in the State of Qatar

In order to predict the pyrolysis product distribution, researchers have created a number of models. In the current study, pyrolytic kinetics model proposed by Song is applied (Song, 2016). The model represents the general trends in product distribution using empirical equations that cover mass and energy balances as well as empirical relationships (based on temperature). Only a few empirical models are based on both feedstock attributes and pyrolysis operation parameters, with the majority of models being based solely on the properties of feedstock materials (Abhijeet et al., 2020; Al-Rumaihi et al., 2022). One of the few studies that considers both the pyrolysis conditions and feedstock qualities is the current one (Parthasarathy et al., 2022). Prediction model-based economic analysis is uncommon, as such, this study is distinctive as it concentrates on both technical and economic assessment. The commercialisation of bio-oil will benefit from the techno-economic analysis of the pyrolysis of the aforementioned manures.

## 2. Methodology

### 2.1. Materials and methods

The proximate and ultimate analyses results of the manures under study are provided in **Table 1**.

**Table 1:** Proximate and ultimate analyses findings of the feedstock manures.

Analysis	PM	DM	CM	SM
Moisture content (%) (as-received basis)				
Moisture	40.00	85.00	58.00	50.00
Proximate analysis (% wt.) (Dry basis)				
Volatile matter	63.58	57.18	60.51	57.30
Fixed carbon*	13.31	10.11	22.74	15.56
Ash	23.09	32.69	16.73	27.13
Ultimate analysis (% wt.) (Dry and ash free basis)				
Carbon	36.67	27.61	37.11	33.14
Hydrogen	5.33	3.48	4.07	4.48
Nitrogen	3.54	1.88	2.27	2.65
Oxygen*	30.61	33.91	39.54	32.21
Sulphur	0.73	0.41	0.24	0.37
Chlorine	0.00	0.00	0.00	0.00

\*Calculated by difference

### 2.2. Model Development

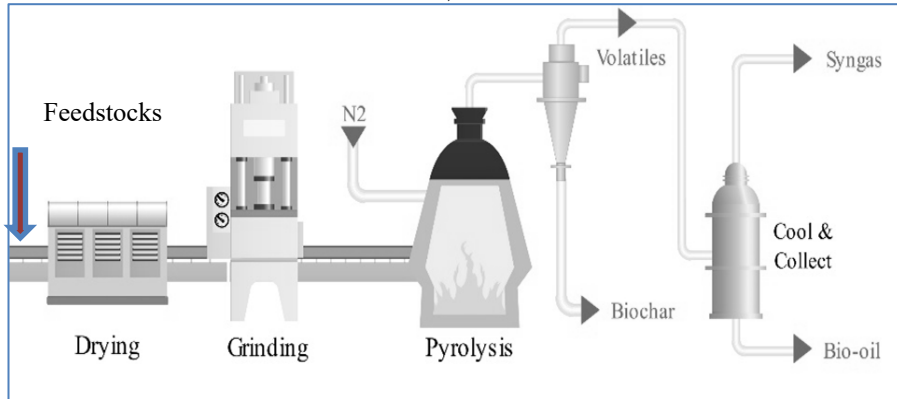
**Figure 1** depicts a simplified process flow diagram of the pyrolysis process. The following assumptions are made for the development of the prediction model. Feedstocks are pyrolysed in an isothermal reactor with a nitrogen environment. As a pre-treatment, the feedstocks are dried, and then they are pyrolysed and transformed into water, bio-oil,

tar, pyrogas (CO<sub>2</sub>, CO, CH<sub>4</sub>, H<sub>2</sub>, hydrocarbons, etc.), biochar, and ash, as denoted by **equations 1-3** (Swagathnath et al., 2019). Further, as the temperature proceeds, the bio-oil is further decomposed into gas components.

$$\text{Bio - char yield} = 0.106 + 2.43 * \exp(-0.66 * T * 10^{-2}) \quad (1)$$

$$\text{Bio - oil yield} = Y_{\text{bio-oil},F} + Y_{\text{H}_2\text{O},F} + \text{samples' moisture composition} \quad (2)$$

$$\text{Pyrogas yield} = Y_{\text{CO}_2,F} + Y_{\text{CO},F} + Y_{\text{H}_2,F} + Y_{\text{CH}_4,F} \quad (3)$$



**Figure 1.** Process flow diagram of the pyrolysis process.

### 2.3. Economic Analysis

The economic analysis of pyrolysis is investigated based on the assumptions listed in **Table 2**. The primary pyrolysis equipment costs and labor requirements have been adapted from previous technical reports (Humbird et al., 2011). Using the Chemical Engineering Plant Cost Index, all prices are scaled up and inflated to the year 2019 (Chemical Engineering Plant Cost Index (CEPCI)).

**Table 2:** Economic analysis assumptions.

Parameters	Values
Pyrolysis plant location	Qatar
Plant capacity	20 t/h
Analysis base year	2019
Plant lifetime (years)	25
Discount rate (%)	20
Feedstock price (USD \$/t)	50

### 2.4. Environmental Analysis

Environmental analysis covers the concept of carbon footprint, which is the total amount of CO<sub>2</sub> emissions that have been generated over the lifecycle of bio-oil, from cradle to gate.

### 2.5. Optimisation

Using the objective functions and the constraints presented below, an optimisation model is created using Excel Solver. For the selection of optimal blends of the four different manures for the pyrolysis process, 7 objectives are adopted independently. The objectives seek to maximise the yields of bio-oil, syngas, and biochar, maximise the heating value of bio-oil and syngas, and minimise cost and carbon footprint. Additionally, the blending must use no less than 10% no more than 50% of each available manure, considering the plant's capacity.

**Objective functions:**

$$\begin{aligned}
 \text{Bio-oil yield } (Z_1) &= \text{Max } \sum_{i=1}^4 BO_i X_i \\
 \text{Syngas yield } (Z_2) &= \text{Max } \sum_{i=1}^4 SG_i X_i \\
 \text{Biochar yield } (Z_3) &= \text{Max } \sum_{i=1}^4 BC_i X_i \\
 \text{Bio-oil HV } (Z_4) &= \text{Max } \sum_{i=1}^4 BHV_i X_i \\
 \text{Syngas HV } (Z_5) &= \text{Max } \sum_{i=1}^4 SHV_i X_i \\
 \text{CO}_2 \text{ emissions } (Z_6) &= \text{Min } \sum_{i=1}^4 E_i X_i \\
 \text{Bio-oil Cost } (Z_7) &= \text{Min } \sum_{i=1}^4 C_i X_i
 \end{aligned}$$

**Constraints:**

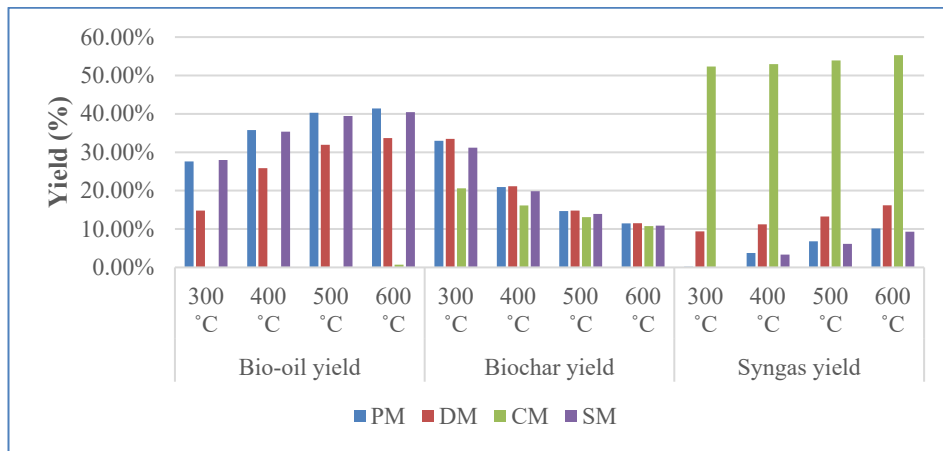
$$\begin{aligned}
 \sum_{i=1}^4 X_i &= 1 \\
 \sum_{i=1}^4 Q_i X_i &\leq \text{feed capacity} \\
 0.5 &\geq X_i \geq 0.1
 \end{aligned}$$

Where:  $X_i$ : blending ratio % of manure ( $i$ );  $Q_i$ : available quantity (t/y) of manure ( $i$ );  $BO_i$ : bio-oil yield (t/y) for manure ( $i$ ),  $SG_i$ : syngas yield (t/y) for manure ( $i$ );  $BC_i$ : biochar yield (t/y) for manure ( $i$ );  $BHV_i$ : Bio-oil heating value (GJ/t) via manure ( $i$ );  $SHV_i$ : syngas heating value (GJ/t) via manure ( $i$ );  $E_i$ : process emissions in kg CO<sub>2</sub>-e /kg manure ( $i$ );  $C_i$ : Bio-oil cost (\$/kg) using manure ( $i$ );  $i=1-4$  represent manure types, whereby: 1: Poultry manure, 2: Dairy manure, 3: Camel manure, 4: Sheep manure.

**3. Results and Discussion**

*3.1. Products Distribution*

**Figure 2.** depicts the effect of temperature on the distribution of pyrolysis products in feedstock samples. In terms of bio-oil and syngas yields, increasing the temperature from 300 to 600°C increased the bio-oil and syngas yields in all feedstocks. However, the increased temperature resulted in a decrease in the biochar yield of all feedstocks. The highest bio-oil and syngas yields were achieved at 600°C, while the highest biochar yield was achieved at 300°C. Among the feedstocks, CM produced the most syngas (55%), DM produced the most biochar (33%), and PM offered the highest bio-oil yield (41%).



**Figure 2.** Effect of temperature on the distribution of pyrolysis products.

*3.2. Bio-oil Levelised Cost*

The bio-oil levelised cost of all manures with respect to temperature is presented in **Figure 3a**. The levelised cost for the bio-oil production from CM is not considered since the bio-oil production from the manure is little. For all manures, the levelised cost for the production of bio-oil is minimum at 500°C, whilst the cost is maximum at 300°C. Of all the manures, PM contributed for the lowest bio-oil levelised cost (0.34 USD/kg of bio-oil). This suggests that pyrolysing PM would result in greater financial gain than pyrolysing other manures.

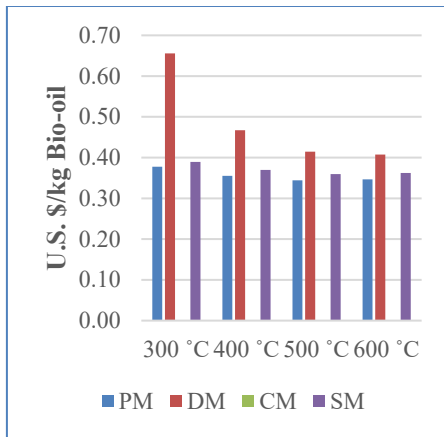


Figure 3a. Bio-oil levelised cost of the manures.

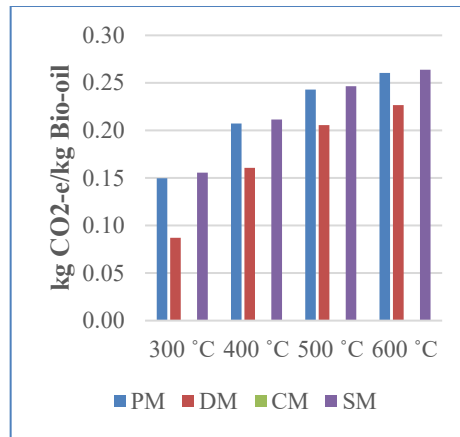


Figure 3b. Carbon footprint of the manures.

### 3.3. Carbon Footprint

Figure 3b. illustrates the carbon footprint (kg CO<sub>2</sub>-e/kg Bio-oil) of all manures in relation to different temperatures. The carbon footprint value for all manures is lowest at 300°C and highest at 600°C. Because the CM produces a meagre amount of bio-oil, the carbon footprint of the manure is negligible. DM had the lowest carbon footprint of the feedstocks, while SM had the highest carbon footprint. This suggests that DM pyrolysis is less polluting than other manures.

### 3.4. Optimal Blend

The effect of manure blending ratio with respect to products' yields, heating value, carbon footprint, and production cost is illustrated in Figure 4. The best manure blending ratio for the maximum bio-oil yield and the lowest cost is achieved at 52% PM: 2% DM: 3% CM: 43% SM. The optimal blending ratio for the lowest carbon footprint is accomplished at 42% PM: 21% DM: 30% CM: 7% SM.

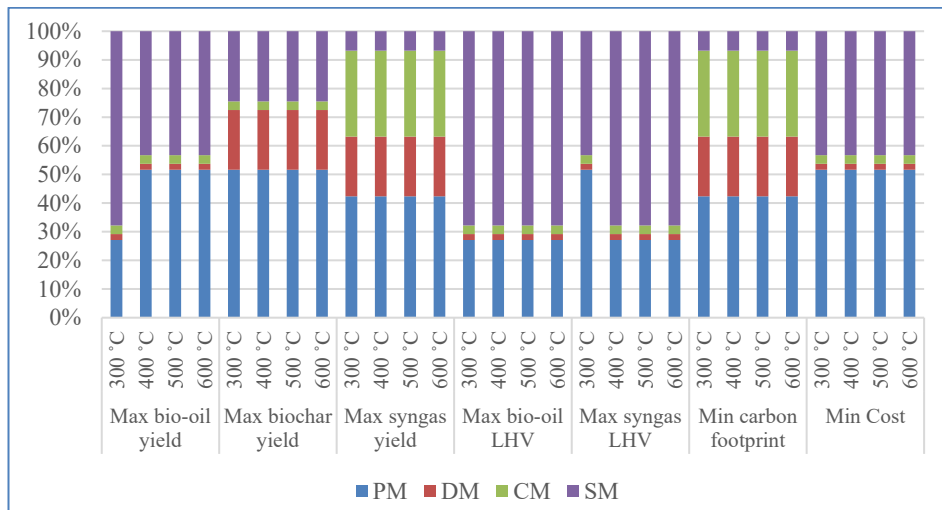


Figure 4. Optimal blending ratios for achieving 7 objectives.



#### 4. Conclusion

The pyrolysis of four different types of livestock manures—poultry (PM), dairy (DM), camel (CM), and sheep manure is examined in this study (SM). The process is simulated using empirical prediction models that are already well established. Additionally, the process performance for the manure feedstocks is assessed from cradle to gate in terms of economic viability and environmental impact. The best manure blend for commercial bio-oil production is then evaluated using a mathematical optimisation model that considers various operating conditions, manure compositions, and availability in the State of Qatar. According to the findings, PM is connected to the highest yield of bio-oil (41%) at 600°C, while CM is responsible for the lowest yield. Finally, 52% PM: 2% DM: 3% CM: 43% SM is found to be the ideal manure blending ratio for the highest bio-oil yield and lowest cost. While 42% PM: 21% DM: 30% CM: 7% SM is the blend that results in the lowest carbon footprint. The created model offers a thorough understanding of the pyrolysis's ideal technical and environmental valorisation of livestock manures.

#### Acknowledgement

The authors would like to thank Qatar National Research Fund (QNRF) for their support of this research through NPRP-11S-0117-180328, the Supreme Committee for Delivery and Legacy (SCDL) and Hamad Bin Khalifa University (HBKU) and Qatar Foundation (QF).

#### References

- Abhijeet, P., Swagathnath, G., Rangabhashiyam, S., Asok Rajkumar, M., Balasubramanian, P., 2020. Prediction of pyrolytic product composition and yield for various grass biomass feedstocks. *Biomass Convers. Biorefinery* 10, 663–674.
- Al-Rumaihi, A., Alherbawi, M., Parthasarathy, P., R.Mackey, H., McKay, G., Al-Ansari, T., 2022. A techno-economic assessment of biochar production from date pits in the MENA region. *Comput. Aided Chem. Eng.* 51, 805–810.
- Alherbawi, M., Parthasarathy, P., Mckay, G., Mackey, H., Al-Ansari, T., 2022. Investigation of optimal blending of livestock manures to produce biocrude via hydrothermal liquefaction. *Comput. Aided Chem. Eng.* 51, 1243–1248.
- Humbird, D., Davis, R., Tao, L., Kinchin, C., Hsu, D., Aden, A., Schoen, P., Lukas, J., Olthof, B., Worley, M., Sexton, D., Dudgeon, D., 2011. Process design and economics for conversion of lignocellulosic biomass to ethanol: Dilute-acid pretreatment and enzymatic hydrolysis of corn stover. United States.
- Moran, D., Wall, E., 2011. Livestock production and greenhouse gas emissions: Defining the problem and specifying solutions. *Anim. Front.* 1, 19–25.
- Parthasarathy, P., Alherbawi, M., Shahbaz, M., Mackey, H., Mckay, G., Al-Ansari, T., 2022. Conversion of oil palm waste into value-added products through pyrolysis: a sensitivity and techno-economic investigation. *Biomass Convers. Biorefinery* 2022 1, 1–21.
- Song, B., 2016. Biomass pyrolysis for biochar production: kinetics, energetics and economics. *Biochar* 227–238.
- Swagathnath, G., Rangabhashiyam, S., Parthasarathi, K., Murugan, S., Balasubramanian, P., 2019. Modeling Biochar Yield and Syngas Production During the Pyrolysis of Agro-Residues, in: H, D., Pillai RG, Tharian MG, M.A. (Eds.), *Green Buildings and Sustainable Engineering: Proceedings of GBSE 2018*. Singapore, pp. 325–336.

# Investigating the feasibility of biomass integrated gasification combined cycle (BIGCC) as a power alternative in Gaza Strip

Rawan Shihada,<sup>a</sup> Rania Salem,<sup>a</sup> Amanda Munia,<sup>a</sup> Nashat Naim,<sup>a</sup> Mohammad Alherbawi<sup>b\*</sup>

<sup>a</sup> Faculty of Engineering and Information Technology, Al-Azhar University, Gaza, Palestine.

<sup>b</sup> College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar

\*[malherbawi@hbku.edu.qa](mailto:malherbawi@hbku.edu.qa)

## Abstract

Gaza strip suffers from severe electricity deficit of 62%, due to the lack of fuel needed to operate the local power plant. As a result, local authorities adopted a rationing plan, through which the power is often cutoff during half of the day. Meanwhile, more than 2,000 tonnes of wastes are generated every day in Gaza, while its municipal capabilities to recycle wastes are limited, leading to the landfilling of most of the generated wastes in a highly populated region. Therefore, this study explores the feasibility of the biomass integrated gasification combined cycle (BIGCC) in Gaza scenario in utilising local wastes and independently generating power at reasonable costs. The BIGCC plant is simulated using the advanced system for process engineering software (Aspen Plus<sup>®</sup>). Furthermore, an economic feasibility analysis is conducted considering the actual costs of similar existing units, as well as local operating costs. The obtained results are encouraging, where the designed BIGCC plant was able to generate ~207 MW, which can fulfill 46% of Gaza's electricity demand, and 86.25% of its current power deficit. The levelised cost of electricity (LCOE) is estimated at 0.045 \$/kWh, which represents 26% only of the current market price. The proposed scenario may contribute to diversifying the power supplies in Gaza and compensate for the power deficit. In addition, a diversion from landfilling can be achieved through wastes valorisation, which reduces the associated land and carbon footprints.

**Keywords:** IGCC; Techno-economics; Clean power; LCOE; Aspen Plus; Gaza.

## 1. Introduction

Gaza Strip has been suffering from a severe electricity shortage since 2007 due to the lack of diesel fuel. The present demand is approximately 450 MW, while the current supplies are no more than 210 MW, ending up with more than 60% shortage of power (PENRA 2022). However, there are 8,900 TJ/year of biomass potential in the Gaza Strip, where considerable amount of biomass is produced every day, including municipal waste (MSW), agricultural residues, and sewage sludge (Al-Najjar et al. 2020). More than 2,000 tonnes of MSW are generated every day, which is expected to reach 2,660 tonnes during the year 2025 (MDLF 2017).

Therefore, this study investigates the feasibility of BIGCC establishment in Gaza to utilise local biomass for the generation of electricity to fill up the current power shortage. In comparison to existing coal power plants, BIGCC plants are known to have a higher thermal efficiency as they convert biomass into syngas and then use its flue gas and heat to generate electricity through a heat-integrated system (Zang et al. 2020).

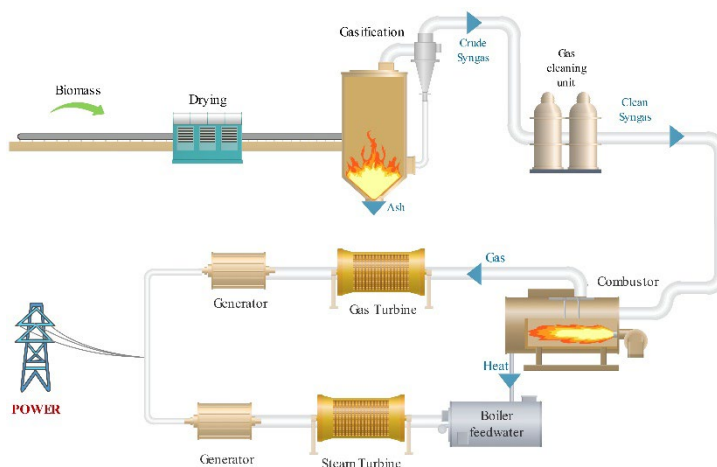
Considering local biomass, a BIGCC plant is modelled and evaluated in Aspen Plus (V.9), while an economic assessment is conducted based on local prices of materials and international equipment costs.

The model in hand may provide insights on power generation alternatives in developing countries through the utilisation of local waste at reasonable costs. Nevertheless, the deployment of BIGCC technology may also contribute towards land filling diversion and emissions mitigation.

## 2. Methodology

### 2.1. Model Development

An integrated gasification combined cycle using municipal solid waste (MSW) is modelled in Aspen Plus V.9. A simplified process flow diagram is shown in Figure 1. Throughout the simulations of BIGCC, mass and energy balance and chemical equilibrium were observed, considering an isothermal and steady state systems.



**Figure 1:** Process flow diagram of BIGCC.

MSW is introduced into the model based on its proximate and ultimate analysis reported in Table 1 (Phyllis 2022). The gasification part of the model was simulated using four Aspen Plus blocks based on Alherbawi et al. (Alherbawi, AlNouss, et al. 2020) model. The process starts by processing the feedstock in a drying unit (RStoic) to reduce its moisture content. The dry MSW which is defined with zero moisture was then passed through an RYield reactor for pyrolysis, where non-conventional components are converted into their equivalent conventional elements and compounds (i.e.,  $H_2$ ,  $N_2$ ,  $O_2$ , C, S, ash). To simplify the model, ash is separated using a solid separator, while useful components are processed along with a gasifying agent (i.e., steam) into two consecutive blocks representing the oxidation and reduction processes following the reactions in Table 2. The reactor calculates the syngas composition by minimising the Gibbs free energy to achieve a chemical and state equilibrium (Alherbawi, Al-Ansari, and Mckay 2020).

**Table 1:** Proximate and ultimate analyses of MSW.

Proximate analysis	(wt. %)	Ultimate analysis	(wt. %)
Moisture	7.56	Carbon	48.23
Ash	17.8	Hydrogen	5.16
Volatile matter	57.99	Nitrogen	1.21
Fixed Carbon	24.21	Oxygen	27.31
		Sulfur	0.29

**Table 2:** Key gasification reactions.

Reaction	Equation	No.
Carbon oxidation	$C + O_2 \leftrightarrow CO_2$	(1)
Partial oxidation	$C + \frac{1}{2} O_2 \leftrightarrow CO$	(2)
Steam generation	$H_2 + \frac{1}{2} O_2 \leftrightarrow H_2O$	(3)
CO oxidation	$CO + \frac{1}{2} O_2 \leftrightarrow CO_2$	(4)
Boudouard reaction	$C + CO_2 \rightarrow 2CO$	(5)
Methanation	$C + 2H_2 \rightarrow CH_4$	(6)
Steam gasification	$C + H_2O \rightarrow CO + H_2$	(7)
Methane reforming	$CH_4 + H_2O \leftrightarrow CO + 3H_2$	(8)
Water gas shift reaction	$CO + H_2O \leftrightarrow CO_2 + H_2$	(9)
Ammonia formation	$N_2 + 3H_2 \rightarrow 2 NH_3$	(10)
Hydrogen sulfide formation	$S + H_2 \rightarrow H_2S$	(11)

Following the gasification reaction, traces of unconverted char are separated from the syngas using solid separator. While a flash unit is used to remove tar and water contents before processing the raw syngas into a methanol-based absorption unit (Alherbawi et al. 2021). To remove CO<sub>2</sub>, NH<sub>3</sub> and H<sub>2</sub>S, cool methanol is introduced into a RadFrac absorber against the flow of syngas, where gas impurities are dissolved in methanol and then stripped in a regenerator column at a higher temperature.

The power generation stage is simulated as a combined cycle system to further utilise the generated heat into electricity production. Where syngas is pumped into a combustion chamber, in which the relevant combustion reactions are defined with adequate flow of air to observe a complete combustion.

The hot and high-pressure gas resulting from the combustion is introduced into a gas turbine and a generator to generate power. In parallel, the combustion's resulting heat is utilised to create steam using a boiler feedwater, where the high-pressure steam is passed through a steam turbine to generate additional power.

### 2.2. Economic Assessment

An economic analysis is conducted based on pre-COVID19 values. An BIGCC plant located in Gaza (Palestine) is considered with a feed capacity of 2000 tonnes/day. The analysis is conducted based on the given parameters and assumptions listed in Table 3. The biomass collection charges, wages, raw materials and utility prices are quoted locally. As for capital expenses (capex), the key equipment prices are adapted from high-level technical reports (National Energy Technology Laboratory 2007; National Renewable Energy Laboratory 2011), scaled up and inflated to the base year of analysis using the Chemical Engineering Plant Cost Index (CEPCI) as expressed in Equation 12 and 13. While the minimum selling price of electricity is calculated using Equation 14.

**Table 3:** Economics parameters and assumptions used in this study.

Parameter	Value/ assumption
Location	Gaza strip (Palestine)
Analysis year	2019
Currency	U.S. \$
Plant feed capacity	2000 t/d
Feedstock gate price	30 \$/t
Water- price	0.26 \$/m <sup>3</sup>
Waste disposal charges	30 \$/t
Electricity price	0.14 \$/kWh
Operating	25% of the labour
Maintenance	2% of equipment
Overhead	60% of the labour
Discount rate	10%
Taxes	25%
Insurance	1.5% of total installed cost (TIC)

$$Cost_{desgin} = Cost_{base} \cdot \left( \frac{Capacity_{desgin}}{Capacity_{base}} \right)^{Scaling\ Factor} \quad (12)$$

$$Cost_{desgin} \$ 2019 = Cost_{desgin} \$ i \cdot \left( \frac{CEPCI_{2019}}{CEPCI\ i} \right) \quad (13)$$

$$LCOE \left( \frac{\$}{kWh} \right) = \frac{Capex + \sum_1^{Lifespan} (Opex(1+Discount\ Rate)^{-Lifespan})}{\sum_1^{Lifespan} generated\ power (1+Discount\ Rate)^{-Lifespan}} \quad (14)$$

### 3. Results and Discussion

#### 3.1. Technical Analysis

At gasification stage, an enhanced syngas is generated with a molar ratio of hydrogen to carbon monoxide at 1.3. A small amount of methane is produced (<0.1%). While the gas impurities (CO<sub>2</sub>, NH<sub>3</sub>, H<sub>2</sub>S) are perfectly removed using methanol absorption. The final molar and mass composition of cleaned syngas is presented in Table 4.

In addition, at combined-cycle power generation stages, 207 MW is generated a year, where 80% of the generated power is obtained from the gas turbine, while the steam turbine contributed with the remaining 20%. The attained power can fulfill 46% of Gaza's electricity demand, and 86.25% of its current power deficit.

**Table 4:** Cleaned syngas composition.

Component	Mass fraction	Mole fraction
H <sub>2</sub>	8.3%	55.7%
CO	88.5%	43%
CO <sub>2</sub>	2.2%	0.6%
Others	1%	0.7%
<b>Sum</b>	<b>100%</b>	<b>100%</b>

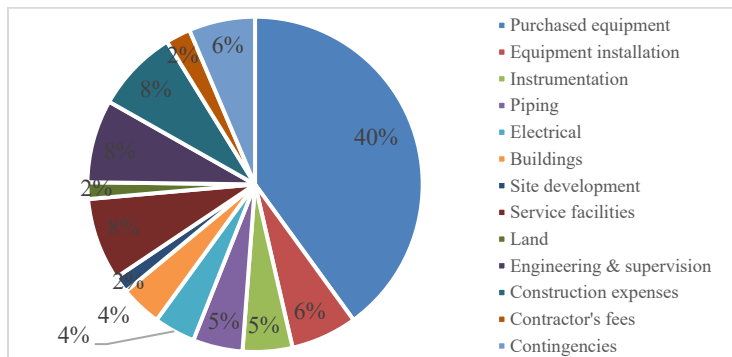
3.2. Economic Analysis

The key findings of the BIGCC economic analysis are presented in Table 5. The capital expenses (capex) are estimated at 608 M U.S.\$, while the plant requires annual operating expenses (opex) of 16.5 M U.S.\$\$. Whereas 190 M U.S.\$ of profit (after tax) can be attained. The return on investment is at 31%, while the investment cost can be paid in 3.2 years. Furthermore, the levelised cost of electricity is achieved at 0.045 \$/kWh, which represents only 26% of the current market price.

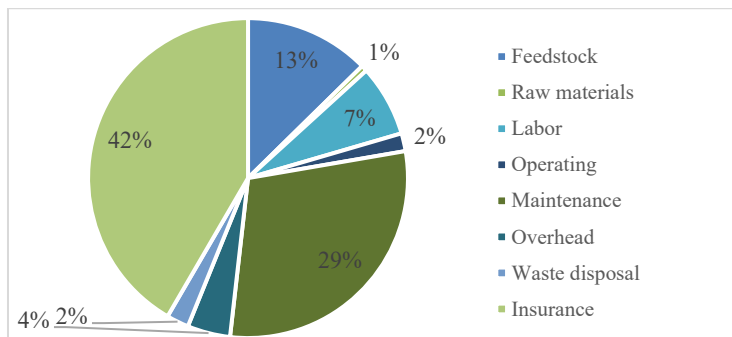
Figures 2 and 3 reveal the breakdown of capex and opex, where purchased equipment occupied 40% of capex, while opex is dominated by raw materials cost at 42%.

**Table 5:** Key economics parameters of Gaza’s BIGCC project.

Item	Value	Unit
Levelised cost of electricity	0.045	\$/kWh
Project capex	608,229,938	\$
Project opex	16,489,484	\$/y
Capacity	207.13	MW
Sales	254,018,100	\$/y
Annual profit	237,528,616	\$/y
Profit after tax	190,022,893	\$/y
Payback time	3.2	years
ROI	31	%



**Figure 2:** Breakdown of Gaza’s BIGCC capex.



**Figure 3:** Breakdown of Gaza’s BIGCC opex.

#### 4. Conclusion

This study has investigated the feasibility of BIGCC establishment in Gaza Strip which is suffering from severe power shortage. The plant was assumed to operate based on local solid waste at 2000 tonnes per day and was simulated and evaluated in Aspen Plus software. The proposed plant generated 207 MW a year, which may fulfill nearly 50% of net Gaza's power demand, and 86% of the present power shortage.

The economic analysis indicated that the plant requires an investment cost of around 600 M U.S.\$, and that it can generate power at a minimum cost of 0.045 \$/kWh, which represents quarter of the current market price.

The model provides insights on local resources-based power generation as a feasible and clean alternative for countries suffering from severe fossil fuel shortage.

#### References

- G. Zang, J. Zhang, J. Jia, E.S. Lora, and A. Ratner, 2020, "Life Cycle Assessment of Power-Generation Systems Based on Biomass Integrated Gasification Combined Cycles." *Renewable Energy* 149:336–46.
- H. Al-Najjar, P. Christoph, R. Al Afif, and H.J. El-Khozondar, 2020, "Estimated View of Renewable Resources as a Sustainable Electrical Energy Source, Case Study." *Designs* 4(3):32.
- M. Alherbawi, T. Al-Ansari, and G. Mckay, 2020, "Potential Integrated Pathways for Jet Biofuel Production from Whole Fruit of *Jatropha*." Pp. 235–40 in 30 European Symposium on Computer Aided Process Engineering. Vol. 48, edited by S. Pierucci, F. Manenti, G. L. Bozzano, and D. B. T.-C. A. C. E. Manca. Elsevier.
- M. Alherbawi, A. AlNouss, G. Mckay, and T. Al-Ansari, 2020, "Optimum Utilization of *Jatropha* Seedcake Considering the Energy, Water and Food Nexus." Pp. 229–34 in 30 European Symposium on Computer Aided Process Engineering. Vol. 48, edited by S. Pierucci, F. Manenti, G. L. Bozzano, and D. B. T.-C. A. C. E. Manca. Elsevier.
- M. Alherbawi, G. McKay, H.R. Mackey, and Tareq Al-Ansari, 2021, "A Novel Integrated Pathway for Jet Biofuel Production from Whole Energy Crops: A *Jatropha Curcas* Case Study." *Energy Conversion and Management* 229:113662.
- MDLF. 2017. Gaza Solid Waste Management Project (GSWMP).
- National Energy Technology Laboratory, 2007, Baseline Technical and Economic Assessment of a Commercial Scale Fischer-Tropsch Liquids Facility.
- National Renewable Energy Laboratory. 2011. Process Design and Economics for Conversion of Lignocellulosic Biomass to Ethanol.
- PENRA, 2022, "Gaza Power Needs." Retrieved August 20, 2022 ([http://www.penra.ps/index.php?option=com\\_content&view=article&id=759:2012-11-14-09-22-07&catid=27:2012-11-14-08-37-30&Itemid=61](http://www.penra.ps/index.php?option=com_content&view=article&id=759:2012-11-14-09-22-07&catid=27:2012-11-14-08-37-30&Itemid=61)).
- Phyllis, 2022, "ECN Phyllis Classification." Retrieved August 15, 2022 (<https://phyllis.nl/Browse/Standard/ECN-Phyllis>).

# The Infinity Reactor: A new conceptual design for a more cost-efficient CO<sub>2</sub> to methanol route

Hilbert Kestra, Edwin Zondervan, Wim Brilman

University of Twente, Drienerlolaan 5, 7522 NB Enschede, The Netherlands

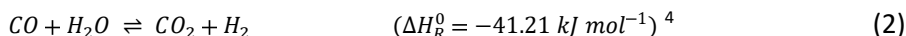
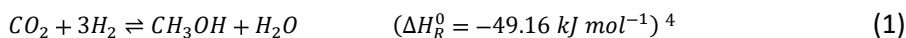
## Abstract

With conventional methanol reactor designs for CO<sub>2</sub>-based methanol synthesis, there is the risk of rapid catalyst deactivation with commercial Cu/ZnO/Al<sub>2</sub>O<sub>3</sub> catalysts at high water partial pressures, related to the use of CO<sub>2</sub> as carbon feed. As catalyst reactivity and exothermicity are influenced by the switch to a CO<sub>2</sub>-rich feed, also heat integration options are affected. A new reactor/process design, the Infinity Reactor, is proposed to both improve catalyst lifetime and reduce catalyst usage. The Infinity Reactor is a shell and tube type of reactor, with catalyst on both sides, utilising the difference in temperature between the inlet and outlet of the reactor to operate as a quasi-gas-cooled reactor. The gas loops over both sides of the reactor, with intermediate cooling and condensation. This modelling study shows around 35% savings in catalyst volume in comparison with adiabatic operation.

**Keywords:** methanol, intensification, infinity, reactor, CO<sub>2</sub>

## 1. Introduction

Methanol serves as an important feedstock for about 30% of industrial chemicals<sup>1</sup>. Methanol is commercially synthesized from fossil-derived syngas (a mixture of CO, CO<sub>2</sub> and H<sub>2</sub>) via the hydrogenation of CO<sub>2</sub> (Equation 1), which is an exothermic reaction. The role of CO at high conversions is to remove excess water via the water-gas shift reaction (WGS) (Equation 2), thereby enhancing CO<sub>2</sub> conversions to methanol<sup>2,3</sup>. However, this mechanism also increases the exothermicity of the process.



Many companies have understood that developing CO<sub>2</sub>-utilizing technologies are an essential step towards a sustainable industrial world<sup>4-6</sup>. However, using



pure CO<sub>2</sub> as a feedstock (e.g. obtained from flue gases to reduce CO<sub>2</sub> emissions) reduces the contribution of the WGS reaction. This reduces the exothermicity of the system but also increases water production during methanol synthesis.

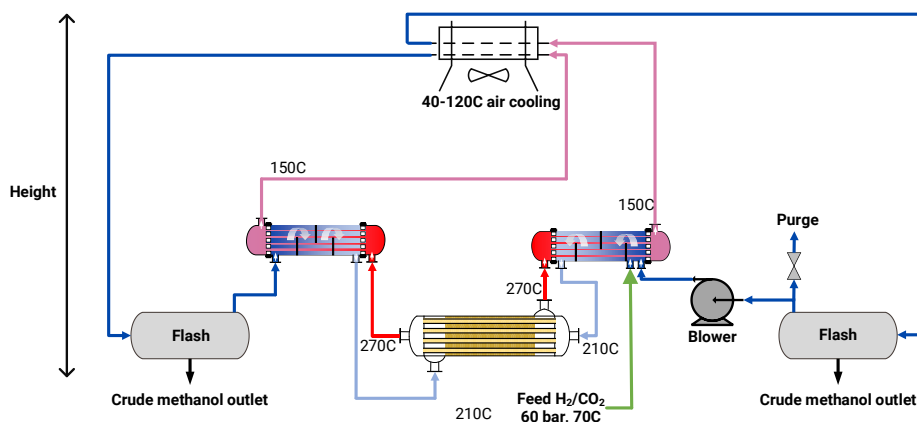
Catalyst deactivation is one of the major hurdles in green methanol production due to a higher water concentration in the process that induces ZnO agglomeration and ZnAl-spinel formation<sup>7,8</sup>. Whereas commercial catalysts have a lifetime of 4-6 years<sup>3,9,10</sup> with CO-rich syngas-based feeds, methanol production with CO<sub>2</sub>-rich feeds suffers from more rapid deactivation in terms of a few months or even weeks<sup>7,11,12</sup>. It is anticipated that this is especially related to higher water partial pressures at high temperatures. Model-based tools like process simulations are necessary for investigating and comparing performances and costs. Therefore, this study has investigated new plant configurations and reactor concepts to improve catalyst lifetime, reduce catalyst usage, and make, through this, the CO<sub>2</sub> to methanol process more cost-efficient.

## 2. Methods

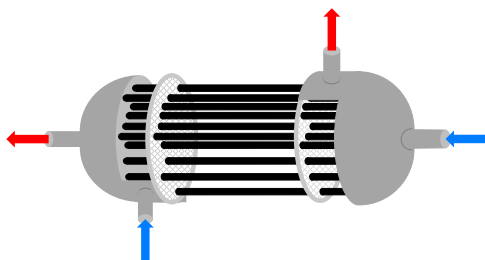
Boiling water reactors (BWR) and quench reactors are the most utilised reactors in commercial methanol synthesis<sup>13,14</sup> and were therefore used as a starting point for this study. Marlin et al.<sup>15</sup> from Carbon Recycling International (CRI) utilise a tube-cooled reactor as a way to save on capital investment because the reaction is less exothermic and progresses slower compared to syngas-based methanol production. The effective wall area for heat exchange between the cooling medium and reaction gas can be decreased by using larger tubes and, thus, heat transfer rates are reduced while still sufficient to remove excess reaction heat.

The models of all the above plant configurations were simulated in Aspen HYSYS and first validated with available plant data from Lurgi<sup>14,16</sup>, which utilises a BWR, and Johnson Matthey<sup>14,17</sup> (formerly ICI), which uses the quench design. All simulations include a recycle stream, heat exchanger, cooler, and flash unit for separating the liquid product mixture.

In this work, an improvement to the above existing reactor concepts and plant configurations is proposed, using in-situ direct heat integration between two different catalyst zones in one (shell and tube) reactor unit to intensify the process, enabled by a lower exothermicity when using CO<sub>2</sub>. Two gas-gas heat exchangers, a combined air cooler, two flash vessels, and a blower complete the (Infinity) reactor loop. Detailed reactor calculations for this reactor were done using Python.



**Figure 1.** Proposed flowsheet of the methanol synthesis loop for the Infinity Reactor.



**Figure 2.** Sketch of an Infinity Reactor. The reactor has catalyst (not drawn) on both sides of the tubes and consists of two stationary heads and two rear heads. The overall flow contacting pattern is counter-current.

In the proposed flowsheet, the fresh feed is mixed with recycle gas and fed to a heat exchanger where it is heated up to 210 °C and then sent to the reactor. The reactor inlet is heated up by the hot reactor outlet (270 °C) of the gas stream on the other side of the tubes, thereby accelerating the reaction rate at the cold feed side and preventing too high temperatures at the hot side. At the other end of the reactor, the reverse happens. Therefore, the reactor is a type of gas-cooled reactor, effectively utilising the temperature gradients inside the reactor. The hot reactor outlet streams are used in a separate heat exchanger to preheat the cold recycle stream from the condenser. The cooled reactor effluent with an outgoing temperature of ~150 °C is subsequently cooled down in a combined air cooler. Any condensed product will be separated in a flash vessel from the unreacted gas. The process can be designed for a low pressure drop to enable the use of a blower instead of a recycle compressor for the recirculation of unconverted syngas.

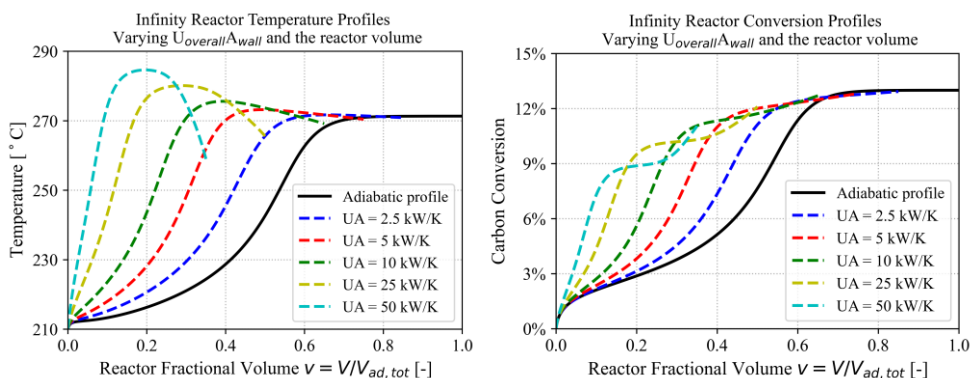
### 2.1. Model development

The kinetic model of Vanden Bussche and Froment<sup>18</sup> (VBF) was chosen because it is one of the most utilised models and serves as a benchmark in the industry<sup>14,19</sup>. Due to input restrictions in Aspen HYSYS, the equilibrium constants as proposed by Graaf et al.<sup>20</sup> were refitted as Arrhenius equations. The kinetic model was first validated with data from VBF<sup>18</sup> and industrial plant data<sup>16,17,21,22</sup>. For heat transfer calculations the lumped value  $UA$  was used, which is the combination of the overall heat transfer coefficient  $U_{ov}$  ( $W/m^2.K$ ), multiplied by the tubular wall area  $A_{wall}$  ( $m^2$ ).

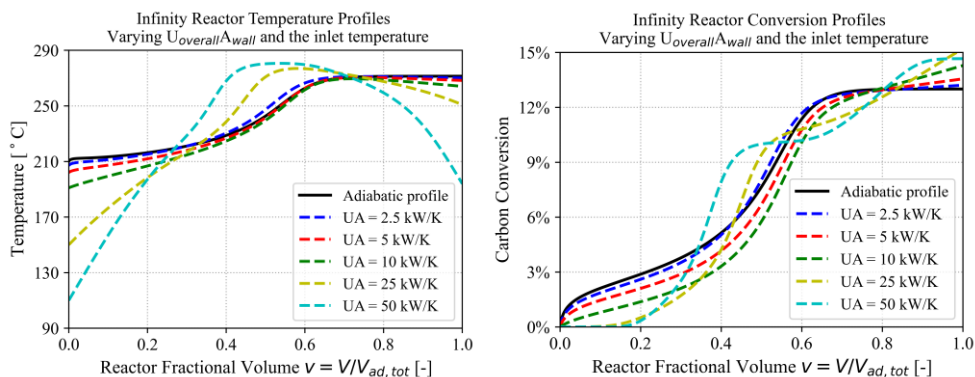
**Table 1.** Main process parameters used in the simulations.

Operating pressure (bar)	60
Cooling temperature ( $^{\circ}C$ )	70
Molar flow rate (kmol/hr)	957
$H_2/CO_2$ ratio	3
Catalyst particle diameter (mm) <sup>14,21</sup>	5.7
Catalyst particle density ( $kg/m^3$ ) <sup>18</sup>	1775
Bed voidage <sup>14,21</sup>	0.39

### 3. Results & discussion

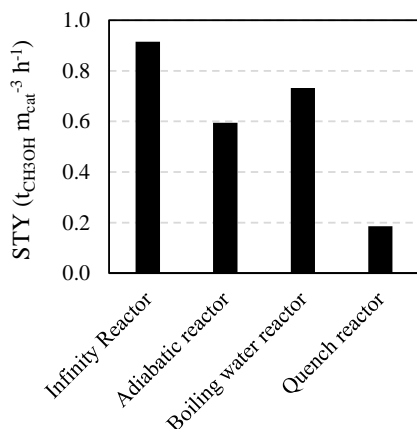


**Figure 3.** Temperature and conversion profiles for the Infinity Reactor with varied  $UA$  and reactor volume but constant inlet temperature. For the adiabatic profile  $V_{tot}=V_{ad,tot}$ ,  $UA=2.5$  kW/K has  $V_{tot}=0.85V_{ad,tot}$ ,  $UA=5$  kW/K has  $V_{tot}=0.75V_{ad,tot}$ ,  $UA=10$  kW/K has  $V_{tot}=0.65V_{ad,tot}$ ,  $UA=25$  kW/K has  $V_{tot}=0.5V_{ad,tot}$ ,  $UA=50$  kW/K has  $V_{tot}=0.35V_{ad,tot}$ .



**Figure 4.** Temperature and conversion profiles of the Infinity Reactor with varied UA and inlet temperature but constant reactor volume. For the adiabatic profile  $T_{in}=210$  °C, UA=2.5 kW/K has  $T_{in}=207$  °C, UA=5 kW/K has  $T_{in}=202$  °C, UA=10 kW/K has  $T_{in}=191$  °C, UA=25 kW/K has  $T_{in}=150$  °C, UA=50 kW/K has  $T_{in}=110$  °C.

Figure 3 shows that the required catalyst amount can be drastically lowered if there is sufficient tube wall area for heat transfer. The average temperature in the reactor is increased and, therefore, the reaction rate is higher on average, resulting in less catalyst required to reach equilibrium. However, too much area for heat transfer will result in a higher peak temperature due to the added heat and results in faster catalyst degradation. Figure 4 shows that the inlet temperature can be lowered while slightly increasing conversions. However, if the temperature is lowered too much, the reaction rate is nearly zero and then it acts only as a cooling medium to lower the reactor outlet temperature of the other reactor. Conversions are ~13% per pass in the Infinity Reactor and adiabatic reactor at the simulated conditions while the BWR and quench reactor are ~20%. Therefore, the water partial pressure is ~35% lower in the proposed concept.



Case-specific conditions and parameters	
Inlet temperature (°C) of the BWR	240
Boiler feed water temperature (°C)	250
UA (kW/K) of the BWR	263
Inlet temperature (°C) of the other designs	210
UA (kW/K) of the Infinity Reactor	10
4-bed volume ratio of the quench reactor	0.2-0.23- 0.23-0.34

**Figure 5.** Space-time yield per simulated reactor concept and the case-specific assumptions.

Figure 5 shows that the Infinity Reactor has a higher catalyst utilisation compared to the other simulated reactor concepts. For the chosen conditions and parameters, catalyst savings are 35% compared to adiabatic operation, 20% compared to BWR, and ~80% compared to a 4-bed quench reactor. In addition, the catalyst lifetime is improved due to the lower water partial pressure compared to a BWR and quench reactor as the single-pass conversion is lower (13% compared to 20%). Furthermore, a multi-stage design with water removal reduces the recycle size and, thus, reduces operational costs.

#### 4. Conclusions

The Infinity Reactor requires significantly less catalyst than adiabatic, boiling water, and quench reactors. It is expected that this will translate to a lower pressure drop, lower investment and lower operational costs. The operational costs are lower due to a smaller gas recycle ratio, caused by a higher conversion per full cycle of the gas loop, related to the two reactor passes per cycle. The somewhat lower single pass conversion leads to a lower water partial pressure inside the reactor compared to the boiling water and quench reactor concepts, by which the catalyst lifetime is improved.

#### References

1. Tountas, A. A. *et al.* Towards Solar Methanol: Past, Present, and Future. *Advanced Science* **6**, (2019).
2. Liaw, B. J. & Chen, Y. Z. Liquid-phase synthesis of methanol from CO<sub>2</sub>/H<sub>2</sub> over ultrafine CuB catalysts. *Applied Catalysis A: General* **206**, 245–256 (2001).
3. Dieterich, V., Buttler, A., Hanel, A., Spliethoff, H. & Fendt, S. Power-to-liquid via synthesis of methanol, DME or Fischer–Tropsch-fuels: a review. *Energy and Environmental Science* **13**, 3207–3252 (2020).
4. Ott, J. *et al.* Methanol. *Ullmann's Encyclopedia of Industrial Chemistry* (2012) doi:10.1002/14356007.a16\_465.pub3.
5. Mitsui Chemicals. CSR Report 2010. [https://jp.mitsuichemicals.com/en/sustainability/report/pdf/csr2010\\_e.pdf](https://jp.mitsuichemicals.com/en/sustainability/report/pdf/csr2010_e.pdf).
6. Bayer. 2010 Sustainable Development report. 2 (2010).

7. Lunkenbein, T. *et al.* Bridging the Time Gap: A Copper/Zinc Oxide/Aluminum Oxide Catalyst for Methanol Synthesis Studied under Industrially Relevant Conditions and Time Scales. **55**, 12708–12712 (2016).
8. Prašnikar, A. & Likozar, B. Sulphur poisoning, water vapour and nitrogen dilution effects on copper-based catalyst dynamics, stability and deactivation during CO<sub>2</sub> reduction reactions to methanol. *Reaction Chemistry & Engineering* **7**, 1073–1082 (2022).
9. Cul, Y. *et al.* Latest catalyst provides more methanol for longer. (2022).
10. Hirotani, K., Nakamura, H. & Shoji, K. Optimum catalytic reactor design for methanol synthesis with TEC MRF-Z<sup>®</sup> reactor. *Catalysis Surveys from Asia 1998 2:1* **2**, 99–106 (1998).
11. Schlögl, R. Chemical Batteries with CO<sub>2</sub>. *Angewandte Chemie International Edition* **61**, e202007397 (2022).
12. Liang, B. *et al.* Investigation on Deactivation of Cu/ZnO/Al<sub>2</sub>O<sub>3</sub> Catalyst for CO<sub>2</sub> Hydrogenation to Methanol. *Industrial & Engineering Chemistry Research* **58**, 9030–9037 (2019).
13. Dieterich, V., Buttler, A., Hanel, A., Spliethoff, H. & Fendt, S. Power-to-liquid via synthesis of methanol, DME or Fischer–Tropsch-fuels: a review. *Energy & Environmental Science* **13**, 3207–3252 (2020).
14. Bisotti, F. *et al.* Impact of Kinetic Models on Methanol Synthesis Reactor Predictions: In Silico Assessment and Comparison with Industrial Data. *Industrial and Engineering Chemistry Research* **61**, 2206–2226 (2022).
15. Marlin, D. S., Sarron, E. & Sigurbjörnsson, Ó. Process Advantages of Direct CO<sub>2</sub> to Methanol Synthesis. *Frontiers in Chemistry* **6**, 446 (2018).
16. Chen, L., Jiang, Q., Song, Z. & Posarac, D. Optimization of Methanol Yield from a Lurgi Reactor. *Chemical Engineering & Technology* **34**, 817–822 (2011).
17. Al-Fadli, A. M., Soliman, M. A. & Froment, G. F. Steady State Simulation of a Multi-Bed Adiabatic Reactor for Methanol Production. *Journal of King Saud University - Engineering Sciences* **7**, 101–132 (1995).
18. Vanden Bussche, K. M. & Froment, G. F. A steady-state kinetic model for methanol synthesis and the water gas shift reaction on a commercial Cu/ZnO/Al<sub>2</sub>O<sub>3</sub> catalyst. *Journal of Catalysis* **161**, 1–10 (1996).

19. Mbatha, S. *et al.* Power-to-methanol process: a review of electrolysis, methanol catalysts, kinetics, reactor designs and modelling, process integration, optimisation, and techno-economics. *Sustainable Energy & Fuels* **5**, 3490–3569 (2021).
20. Graaf, G. H., Sijtsema, P. J. J. M., Stamhuis, E. J. & Joosten, G. E. H. Chemical Equilibria in Methanol Synthesis. *Chemical Engineering Science* **41**, 2883–2890 (1986).
21. Rahmatmand, B., Rahimpour, M. R. & Keshavarz, P. Introducing a novel process to enhance the syngas conversion to methanol over Cu/ZnO/Al<sub>2</sub>O<sub>3</sub> catalyst. *Fuel Processing Technology* **193**, 159–179 (2019).
22. Keramat, F., Mirvakili, A., Shariati, A. & Rahimpour, M. R. Investigation of anti-condensation strategies in the methanol synthesis reactor using computational fluid dynamics. *Korean J. Chem. Eng.* **38**, 2020–2033 (2021).

# Optimal design of a microbial electrolysis cell for biohydrogen production

Uriel Roberto Pedroza-Medina<sup>a</sup>, Luis Felipe Cházaro-Ruiz<sup>b</sup>, Ricardo Femat<sup>b</sup>,  
Alicia Román-Martínez<sup>a\*</sup>

<sup>a</sup>*Facultad de Ciencias Químicas, UASLP, M. Nava 6, San Luis Potosí, 78210, MÉXICO*

<sup>b</sup>*Division de Ciencias Ambientales, Instituto Potosino de Investigación Científica y Tecnológica A. C., Camino a la Presa San José 2055, San Luis Potosí, 78216, MÉXICO*

\**alicia.romanm@uaslp.mx*

## Abstract

Biohydrogen is considered as an alternative to replace fossil fuels and is being researched as a possible alternative to reduce conventional non-renewable hydrogen technologies. Towards a circular economy, microbial electrolysis cells have been suggested as a tool to both reduce organic pollutants from wastewater as well as to produce desirable hydrogen gas. Research efforts have focused on experimental design and computational simulation. The work presented here aims to develop an optimization model capable of estimating an optimal design and operational criteria to maximize cell's hydrogen output.

**Keywords:** microbial electrolysis cell, hydrogen, modeling, optimization

## 1. Introduction

The Microbial Electrolysis Cell (MEC) is a fairly new technology that has emerged and got scientific attention. The cell exploits the conductor-like properties of electrogenic microorganisms that catalyze the electrochemical oxidation of organic substrate, converting electrical energy into chemical energy stored in the H<sub>2</sub> molecule that is produced through the electrochemical reduction of the protons of the electrolytic medium. The main advantage of this technology is the ability to both treat wastewater effluents (source of organic matter) and synthesize biohydrogen. Numerous wastewater effluents have been demonstrated as potential substrate sources for MEC operation with varying degrees of efficiency in both organic-waste removal and hydrogen generation.

MEC-reactor design is known to be influenced by effluent's substrate concentration, reactor geometry and sizing, electrode material and configuration, anode-cathode distance, presence of a separator between them, composition of the electrolytic medium, as well as the microbial communities in the anodic chamber (Rousseau, 2021). Experimental studies have been performed to comprehend and optimize the processes occurring in these systems, however, so far there is no consensus to determine a feasible reactor configuration that would lead to competitive hydrogen production rates.

One approach that could be used to oversee these hindrances is mathematical modelling, to which fewer studies have been carried out. The first model proposed for a MEC was reported by Pinto *et al.* in 2011. This work studied the effect of several microbial communities: electrogenic, methanogenic, acetoclastic and hydrogenotrophic distributed in bioanode layers to understand their competition for the substrate in a continuous MEC. Later, Karimi, *et al.* in 2015 reported a generalized model that assumes coexisting microbial populations and the effects of a dynamic biofilm growth on substrate-potential distribution within it considering a conduction-based biofilm model in a batch reactor. A



similar model was proposed by Flores-Estrella, *et al.*, in 2019 considering a continuous operation MEC. Hernández-García, *et al.*, in 2020 reported a model that considers the influence of geometry with primary and secondary current-substrate distributions in 2D, and was re-studied, in 3D considering the effects of hydrodynamic behavior in a fed-batch reactor. Other works have developed models, which include stability and sensibility analysis and process control. As for optimization, some studies have been performed by trial-error and online methods, with little focus in a mathematical problem statement. Design optimization is an engineering tool that utilizes a mathematical formulation of a design problem to support the selection of an optimal design among several alternatives. This work proposes a process model for a MEC based in a conduction-based bioanode considering operational and topological constraints, stated as discrete decision variables in order to find a design that maximizes hydrogen production rate.

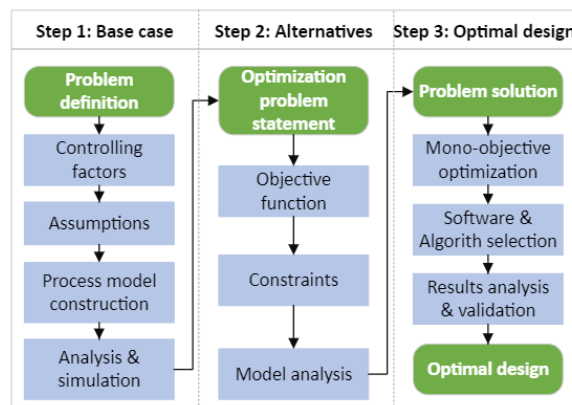
## 2. Methodology

### 1.1. Modeling framework

A systematic modeling approach has been adapted from Aguiñaga-Morales *et al.* (2017) which allows the construction of a process model that will be set as equality constraint in an optimization model, described in Figure 1.

### 1.2. Problem statement

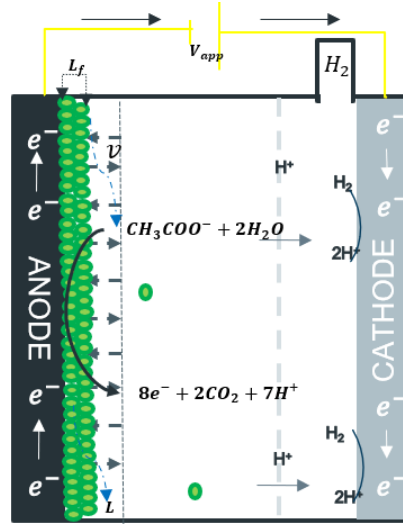
A Microbial electrolysis cell (MEC) consisting of two chambers, separated by a cation exchange membrane, contains acetate as substrate in the anodic chamber, a carbonaceous material anode colonized by electroactive bacteria and an abiotic platinum cathode. The electrodes are linked by an external circuit. The Ag/AgCl/KCl<sub>(sat)</sub> system is used as the reference electrode. The substrate oxidation releases CO<sub>2</sub>, electrons, and protons, which will be electrochemically reduced at the cathode. The system is illustrated in Figure 2.



**Figure 1.** Modeling approach adopted to solve the optimal MEC design problem.

#### 1.2.1. Identification of limiting factors and assumptions

(1) Substrate mass transport is mediated by diffusion, here described in 1D from and to the bioanode. (2) The biofilm acts as a porous conductor with conductivity  $\kappa_{bio}$ . (3) Electroactive bacteria exhibit extracellular electron transfer mechanisms through conductive pili and cytochrome C in their cell wall. (4) Cell operates in batch mode with uniform substrate distribution in bulk liquid. (5) Biofilm is mediated by advection. (6) Cathodic chamber is free of microorganisms. (7) An external power supply provides



**Figure 2.** Schematic representation of phenomena in the modeled MEC. On the bioanode the model substrate (acetate) is diffused and oxidized by electroactive bacteria for cell growth. The generated protons go through the membrane to be reduced at the cathode.

the cell with a voltage that drives the electrochemical production of  $H_2$ . (8) The substrate oxidation generates the electrons needed to reduce the protons at the cathode, in terms of produced current or consumed electric charge. (9) This model is limited to hydrogen generation only, and does not consider methane production by other bacteria.

### 1.3. Construction of process model

#### 1.3.1. Substrate mass balance in biofilm domain

Substrate accumulation in biofilm is mediated by diffusion and the consumption term, expressed as the following equation:

$$\frac{\partial S_d}{\partial t} = D_{S,f} \frac{\partial^2 S_d}{\partial z^2} - X_{f,a} q \quad (1)$$

With boundary conditions in anode surface ( $z=0$ ) and biofilm-liquid interface ( $z=L_f$ ):

$$D_{ED,f} \frac{\partial S_d}{\partial z} \Big|_{z=0} = 0$$

$$D_{ED,f} \frac{\partial S_d}{\partial z} \Big|_{z=L_f} = \left(\frac{D_{S,l}}{L}\right)(S_{d,l} - S_{d,n})$$

Where  $S_d$  is the substrate concentration,  $D_{S,f}$  is the substrate diffusion constant at the biofilm,  $X_{f,a}$  is the electroactive biomass density and  $q$  is the substrate consumption rate.  $q$  is expressed as Monod-like kinetics derived from the Nernst equation, named Nernst-Monod kinetic equation:

$$q = q_{max} \phi_a \frac{S_d}{S_d + K_S} \frac{1}{\exp[-F\eta/RT] + 1} \quad (2)$$

$q_{max}$  is the maximum rate of substrate consumption,  $\phi_a$  is the volumetric fraction of electroactive biomass,  $F$  is the Faraday constant,  $T$  is the anodic chamber temperature and  $\eta$  is the local potential, and  $K_S$  is the Monod's half-saturation constant.

#### 1.3.2. Charge balance in biofilm

Electron transfer through the biofilm matrix is allowed by the local potential difference across the biofilm; the biofilm potential's change over time is expressed by this conduction (derived from the Ohm's law) and the addition of electrons generated by substrate oxidation and endogenous respiration:

$$\frac{\partial \eta}{\partial t} = \kappa_{bio} \frac{\partial^2 \eta}{\partial z^2} + \frac{F\gamma_1}{\tau} L_f f e^0 X_{f,a} q + \frac{F\gamma_2}{\tau} L_f X_{f,a} r_{res} \quad (3)$$

With boundary conditions in anode surface ( $z=0$ ) and biofilm-liquid interface ( $z=L_f$ , no conduction):

$$\eta|_{z=0} = V_{anode} \quad \frac{\partial \eta}{\partial z}|_{z=L_f} = \frac{j|_{z=L_f}}{\kappa_{bio}} = 0$$

Here  $\eta$  is the local potential,  $\kappa_{bio}$  is the biofilm's conductivity,  $\gamma_1$  is the yield of electrons from substrate oxidation,  $L_f$  is the biofilm thickness,  $X_{f,a}$  is the biomass density,  $f e^0$  is the fraction of electrons used for electrosynthesis,  $\gamma_2$  is the electron yield from endogenous respiration,  $F$  is the Faraday's constant,  $\tau$  is the time conversion and  $r_{res}$  is the endogenous respiration rate:

$$r_{res} = b_{res} \frac{1}{\exp[-F\eta/RT]+1} \quad (4)$$

### 1.3.3. Biomass balance in biofilm

The microbial population exists as a consortium of electroactive bacteria, capable of metabolic extracellular electron transfer, and non-electroactive bacteria, which are capable of substrate consumption. In volume terms, the sum of their fractions is given by:

$$\phi_a + \phi_i = 1 \quad (5)$$

$\phi_a$  and  $\phi_i$  are the electroactive fraction and the non-electroactive fraction, respectively. Their mass balances are stated as:

$$\frac{\partial \phi_a}{\partial t} + \frac{\partial(v\phi_a)}{\partial z} = Yq - r_{res} - r_{rina} = \mu_a \quad (6)$$

Which represents the transient electroactive fraction with respect to its position (a function of the advective velocity),  $v$ , by metabolic growth, loss by endogenous respiration and loss by inactivation.

$$\frac{\partial \phi_i}{\partial t} + \frac{\partial(v\phi_i)}{\partial z} = \frac{X_{f,a}}{X_{f,i}} r_{rina} = \mu_i \quad (7)$$

This equation gives the transient volumetric fraction of non-electroactive biomass, and its growth is given by the rate of inactivation,  $r_{rina}$ , which is stated as:

$$r_{rina} = b_{ina} \phi_a \quad (8)$$

Where  $b_{ina}$  is the inactivation constant. The advective velocity may also vary on biomass growth, so we can state:

$$\frac{\partial v}{\partial z} = \mu_a + \mu_i \quad (9)$$

This can be integrated along the biomass specific growth rate to get the advective velocity at any point  $z'$ , considering that the net growth enhances the advective velocity.

The biofilm growth is thus a function of the advective velocity and the detachment phenomenon (which reduces biofilm's thickness) and can be formulated as:

$$\frac{dL_f}{dt} = v(t, L_f) - b_{det} L_f \quad (10)$$

Where  $b_{det}$  is the detachment constant.

### 1.3.4. Mass balances in liquid bulk

The substrate fed to the cell is accumulated on its volume, and fluxes to the biofilm domain as it's also consumed. The macroscopic mass balance can be stated as:

$$\frac{d(S_l V_b)}{dt} = A_s \frac{D_{S,l}}{L} (S_{d,n} - S_l) - V_b q_b \quad (11)$$

Here  $S_l$  represents the substrate concentration in the bulk liquid,  $V_b$  is the liquid volume in the anodic chamber,  $A_s$  is the anode surface area,  $D_{S,l}$  is the diffusion coefficient in the liquid,  $L$  is the thickness of the diffusion layer,  $S_{d,n}$  is the substrate concentration in the biofilm's surface and  $q_b$  is the substrate consumption rate by suspended biomass.

For the suspended biomass, we express:

$$\frac{dx_b}{dt} = Y_b q_b - b_b x_b \quad (12)$$

$x_b$  is the biomass concentration in the anodic chamber,  $b_b$  is the death rate and  $Y_b$  is the substrate/biomass yield. The consumption rate law for suspended biomass, which doesn't contribute to electron transport, is:

$$q_b = q_{max,b} x_b \frac{S_l}{K_{s,l} + S_l} \quad (13)$$

### 1.3.5. Current generation and hydrogen productivity

The anode potential is function of the applied cell potential by the external power supply.

$$V_{an} = V_1 V_{app} + V_2 \quad (14)$$

Where  $V_{an}$  is the anode potential,  $V_{app}$  is the applied potential, and  $V_1$  and  $V_2$  are constant relationships of linear applied potential.

By Ohm's Law, the current density ( $j$ ) might be stated as the product of the biofilm's conductivity times the quotient of the potential/biofilm thickness gradients, and the cell current ( $I$ ) is known by the product of the anode's surface area times  $j$ :

$$j = -\kappa_{bio} \frac{\Delta \eta}{\Delta z} \quad \text{and} \quad I_{MEC} = A_s j \quad (15)$$

So, we can state the hydrogen productivity as a function of the cell current that conducts to the cathode:

$$Q_{H_2} = \frac{RT}{V_b P} \left( \frac{\tau I_{MEC} Y_c}{nF} \right) \quad (16)$$

$Q_{H_2}$  is the hydrogen volumetric productivity ( $\text{m}^3 \text{H}_2 \text{m}^{-3} \text{anode day}^{-1}$ ) and  $Y_c$  is the cathode's dimensionless efficiency.

## 1.4. Analysis and simulation

Kinetic and electrochemical parameters are obtained from previous MFC and MEC modeling works (Pinto, 2011; Karimi, 2015). To implement, discretize (1) and (3) and their B.C., in space through finite differences, solve non-linear system through Newton's method. Calculate advective velocity (9) by numerical integration (trapezoid rule) then solve for biofilm thickness (10). Biomass balances (5, 6, 7) at the biofilms are solved by Backward-Euler. Subtract detachment rate from biofilm thickness. Interpolate the remaining biomass onto a new grid and iterate. Use a similar approach for (11) and (12).

## 2.5. Optimization problem statement

### 2.5.1. Objective function

The present work aims to optimize the energy efficiency of an MEC considering spatial and operational factors. The cell efficiency is stated as the energy load achieved by hydrogen gathered, a function of its enthalpy of combustion *versus* the applied energy (applied electric potential) and any other applied energy sources as mixing and pumping. We can express this mathematically as:

$$\eta_{H_2} = \frac{W_{H_2}}{W_E} = \frac{Q_{H_2} \rho_{H_2} \Delta H_{H_2}}{I_{MEC} V_{app}}$$

Thus, we can state our objective function as:

$$\max \eta_{H_2} = \frac{Q_{H_2} \rho_{H_2} \Delta H_{H_2}}{I_{MEC} V_{app}}$$

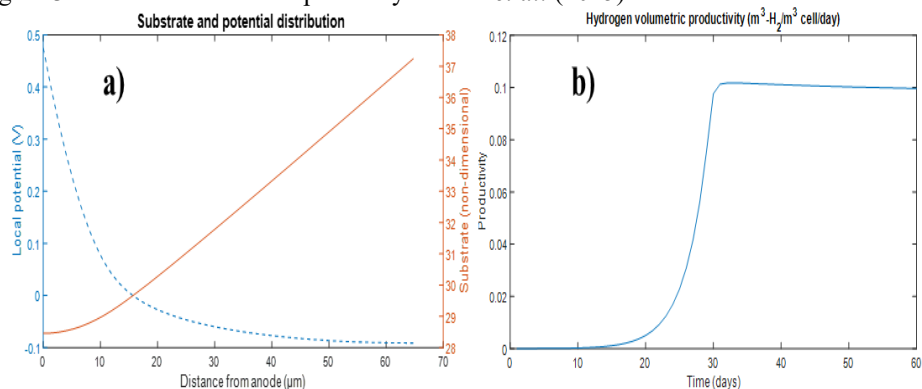
s. t.  $\mathbf{h}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$  (equality constraints);  $\mathbf{g}(\mathbf{x}, \mathbf{y}) \leq \mathbf{0}$  (inequality constraints)

$\mathbf{x} \in \mathbf{X} \subseteq \mathbb{R}^n$ ;  $\mathbf{y} \in \mathbf{Y}$  integer

## 3. Process model simulation results

Thus far, the process model returns spatial distributions of substrate, voltage, and current density in the  $z$  direction across the biofilm. Figure 3 shows that the substrate concentration is highest in biofilm's vicinity with liquid, opposite to the potential which decreases with biofilm length, while the current density behaves similarly, achieving a current density of  $2.9 \text{ A/m}^2$  at tested substrate load of  $1.5 \times 10^{-3} \text{ mol/L}$ . The max. current reached in the  $10 \text{ cm}^2$ -bioanode is  $3.1 \times 10^{-4} \text{ A}$ , so the max. hydrogen productivity is  $0.102$

$\text{m}^3\text{H}_2/\text{m}^3\text{cell day}$  at 32 days, when microbes reach stationary phase and substrate starts depleting. For validation, potential/substrate distribution and  $\text{H}_2$  productivity profiles in Figure 3 are similar to those reported by Karimi *et al.* (2015).



**Figure 3.** (a) Distribution of potential (dashed) and substrate (solid) across biofilm at max length 70  $\mu\text{m}$ . Simulated time, 60 days, initial substrate concentration:  $1.5 \times 10^{-3}$  mol Acetate/L. (b) Hydrogen productivity, reached max. at 32 days.

#### 4. Conclusions

The developed process model is suitable to be formulated as an equality constraint for the objective function and may be tested in different parameter scenarios. It demonstrates microbial processes are limiting to MEC performance. Further alternative analysis will be carried out to complete the optimization problem formulation.

#### 5. Acknowledgments

U.R. Pedroza-Medina gratefully acknowledges CONACyT in Mexico for granting the scholarship 1143427 to pursue his M. Sc. Degree.

#### References

1. R. Rousseau, 2021, Microbial electrolysis cell (MEC): Strengths, weaknesses and research needs from electrochemical engineering standpoint, *Applied Energy*, V. 257, 113938.
2. F. Aguiñaga-Morales, 2017, Model identification, calibration and validation of the aerobic stage in water remediation of a paper mill effluent, *Computer Aided Chem. Eng.*, V40, 547-552.
3. R. Pinto, 2011, Multi-Population Model of a Microbial Electrolysis Cell, *Environ. Sci. Technol.*, 45, 5039–5046.
4. M. Karimi-Alavijeh, 2015, A Generalized Model for Complex Wastewater Treatment with Simultaneous Bioenergy Production Using the Microbial Electrochemical Cell, *Electrochim. Acta*, V. 167, 84-96.
5. R.A. Flores-Estrella, 2019, A Dynamic Biofilm Model for a Microbial Electrolysis Cell, *Processes*, 183, 7, 1-20.
6. K.M. Hernández-García, 2020, Modeling 3D current and potential distribution in a microbial electrolysis cell with augmented anode surface and non-ideal flow pattern, *Biochem. Eng. Journal*, V. 162, 107714.

# Simulation and Exergoeconomic Analysis of Brewer's Spent Grains convective drying process

Mero-Benavides M.,<sup>a</sup> Enriquez-Posligua J.,<sup>a</sup> Barcia-Quimi A.F.,<sup>a</sup> Tinoco-Caicedo Diana L.<sup>a,b\*</sup>

<sup>a</sup> *Facultad de Ciencias Naturales y Matemáticas (FCNM), Escuela Superior Politécnica del Litoral Ecuador, 090903 Guayaquil, Ecuador*

<sup>b</sup> *Centro de Energías Renovables y Alternativas (CERA), Escuela Superior Politécnica del Litoral Ecuador, 090903 Guayaquil, Ecuador*

\**dtinoco@espol.edu.ec*

## Abstract

A low-cost and highly exergetic efficient drying process assures a sustainable transformation of Brewer's Spent Grains (BSGs) into value-added product. This work aims to perform an exergoeconomic analysis at a level component of the convective simulated drying process of BSGs. To validate the simulated process, characterization, and a drying kinetic study of BSGs obtained from a factory in Guayaquil, Ecuador, were carried out. The simulated drying process was validated, reaching a maximum relative error of 4.64%. The exergy destruction rate varied from 12.6 MJ/kg to 3.42 MJ/kg when the air inlet temperature varied from 60 °C to 120 °C. Results suggest that the increment of the air inlet temperature could reduce the specific cost of the dry BSGs up to a 68.3% and increase the exergetic efficiency by 24.6%.

**Keywords:** kinetic study, exergoeconomic analysis, process simulation, modeling, Brewer's Spent Grains (BSGs)

## 1. Introduction

Currently, reducing waste is considered a critical need in the brewing industry. The main waste in this production process is the brewer's spent grains (BSGs) (Rachwał et al., 2020). Usually, it is necessary to dry the BSGs due to the high moisture levels before converting them into value-added products (Mallen & Najdanovic-Visak, 2018). Hence, the drying process needs to be sustainable and have low operating costs. The exergoeconomic analysis is considered a vital decision-making tool to reduce production costs and increase the sustainability of a process by reducing the exergy destruction rate (Bejan & Tsatsaronis, 2021). Although there are some drying experimental studies of this waste at temperatures from 40 °C to 95 °C (Capossio et al., 2022) (Castro et al., 2022), these studies do not analyze the best-operating conditions to diminish the exergy destruction rate and the operating cost of the drying process. Thus, there has been no exergoeconomic analysis of the drying process of BSGs. In this context, this work presents an exergoeconomic analysis of a convective drying process of BSGs. The characterization and drying kinetic study were performed using BSGs from a factory located in Guayaquil, Ecuador. The drying process was simulated using Aspen Plus V12.1, and the obtained experimental data was used to validate the process simulation.

The exergetic, economic, and exergoeconomic balances were performed for the dryer using Engineering Equation Solver (ESS) software following the methodology from (Bejan et al., 1996). The exergy destruction rate, investment and operational costs were determined for drying temperatures from 60 °C to 120 °C.

## 2. Materials and Methods

### 2.1. Experimental Studies

#### 2.1.1. Characterization of BSGs

The BSGs were obtained from a brewing company in Guayaquil, Ecuador. The samples were dried at 105°C based on ISO 18123:2015. The proximate analysis (ash and volatile material (VM)) was done according to ISO 18122:2015, respectively. The difference between ash and volatile material determined the fixed carbon (FC) in dry basis. The ultimate analysis was performed by a CHNS Elemental Analyzer (Elementar). The particle size distribution was analyzed according to the ASTM D1511-10 procedure.

#### 2.1.2. Drying Kinetics

The drying kinetic study was carried out at temperatures 60, 70, 80, 90, 105, and 120°C. Higher temperatures were not selected, because a previous study shows that the degradation rate increases (Mishra et al., 2017). Dry BSGs were submerged in distilled water at room temperature during 2 h to reach their initial moisture (Cuevas et al., 2019). Samples were placed on a plate up with thickness of 1.30 cm inside a convective dryer (UNB 500, Memmert, Germany). Periodically the samples were removed from the dryer, weighing them on digital balance in order to measure their average water content.

The critical moisture was determined as transition moisture, where the first constant speed dry deviation was observed. The equilibrium moisture was estimated as the constant behavior for an extended period (Srikiatden & Roberts, 2007). The effective diffusivity was determined from the slope of the natural logarithm of the moisture ratio ( $M_R$ ) versus drying time according to the second diffusion Fick Law for the infinite dry plate (Cuevas et al., 2019). The heat and mass coefficients were determined from the dimensionless numbers based on Incropera et al. (Incropera et al., 2007).

### 2.2. Process Simulation

Figure 1 shows the BSGs simulated drying system. It consists of an air preheater (HX-101) and a batch-type convection drying chamber (D-101). Flue gases at 300 °C (stream 6) with a mass composition of 18.5% CO<sub>2</sub>, 6.8% H<sub>2</sub>O, 7.09% N<sub>2</sub>, and 3.8% O<sub>2</sub> enter the HX-101. Then, ambient air enters the HX-101 with a relative humidity of 73% at 25 °C (stream 1) and exits at the drying temperature (stream 2), which varies between 60 °C to 120 °C. Then, the flue gases exit the HX-101 (stream 7) and they are discharged to the environment. The hot air enters the D-101 to remove the water from the BSGs (stream 3), and the moist air exits the dryer (stream 4). After the process, the dry BSGs (stream 5) is discharged.

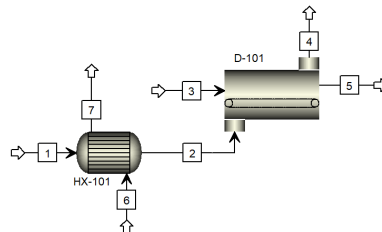


Figure 1. Process flow diagram of the drying system

The process simulation was performed in Aspen Plus V12.1. BSGs were simulated as an unconventional component. The ultimate and proximate analysis results obtained in section 2.1 were entered into the HCOALGEN and DCOALIGT models. First, the ideal thermodynamic model was used since it worked at atmospheric pressure, and the MIXCNCPD for handling unconventional solids configured the flow class. Next, the dryer was simulated using a convective dryer. Finally, the simulation was configured in batch mode, where the total drying time was set as the maximum iteration and operation time. Drying data such as heat and mass transfer coefficients, drying curve, equilibrium, and critical moisture were defined in the component.

2.3. Model Validation

The absolute moisture of the BSGs obtained in the simulation was compared with the results obtained experimentally for each drying air temperature. The same operating conditions were maintained to ensure the validity of the modeled process.

2.4. Exergoeconomic analysis

The exergoeconomic analysis was developed according to the methodology offered by (Bejan et al., 1996) using equations Eq.(1)-(4). Engineering Equations Solver (EES) software was used to determine the thermophysical properties for the calculation of enthalpy and entropy of each state.

$$\begin{aligned} \dot{E}_{D,k} &= \dot{E}_{F,k} - \dot{E}_{P,k} \quad (1) & \dot{C}_{F,k} + \dot{Z}_k &= \dot{C}_{P,k} \quad (3) & PEC_{HX-101} &= 1900 + 2500 \cdot S \quad (5) \\ \dot{Z}_k &= \dot{Z}_k^{O\&M} + \dot{Z}_k^{CI} \quad (2) & \dot{C}_{D,k} &= c_{F,k} \dot{E}_{D,k} \quad (4) & PEC_{D-101} &= 10000 + 7900 \cdot S^{0.5} \quad (6) \end{aligned}$$

The economic analysis was carried out following the revenue requirement method. The economic parameters were taken from a previous study by authors (Tinoco-Caicedo et al., 2021). The purchase equipment costs (PEC) of the HX-101 and the D-101 were determined using Eq. (5)-(6) (Towler & Sinnott, 2012), respectively, for 2010 (CEPCI of 532.9). Where S is the area of each tray from the dryer (20 m<sup>2</sup>) and the heat exchange area for the HX-101 (obtained in the simulator). The chemical engineering plant cost index (CEPCI) of 607.5 was used to adjust the PEC for 2019.

Table 1 shows the fuel exergy ( $\dot{E}_F$ ), product exergy ( $\dot{E}_P$ ), fuel cost ( $\dot{C}_F$ ), product cost ( $\dot{C}_P$ ) and auxiliary equations at the component level.

Table 1. Equations used for the exergoeconomic analysis

Component	$\dot{E}_{F,k}$	$\dot{E}_{P,k}$	$\dot{C}_{F,k}$	$\dot{C}_{P,k}$	Auxiliary equation
HX-101	$\dot{E}_6 - \dot{E}_7$	$\dot{E}_2 - \dot{E}_1$	$\dot{C}_6 - \dot{C}_7$	$\dot{C}_2 - \dot{C}_1$	$c_6 = 0$ $c_1 = 0$
D-101	$\dot{E}_2 - \dot{E}_4 + \dot{E}_3$	$\dot{E}_5$	$\dot{C}_2 - \dot{C}_4 + \dot{C}_3$	$\dot{C}_5$	$c_3 = 0$
Overall System	$\dot{E}_F = (\dot{E}_6 - \dot{E}_7) + \dot{E}_3 + (\dot{E}_2 - \dot{E}_4)$ $\dot{E}_P = \dot{E}_5$				

3. Results and Discussion

3.1. Experimental studies

The characterization and drying kinetic study was done for the BSGs. Table 2 shows the results obtained for the ultimate and proximate analysis and the particle size distribution.

Table 2. BSGs characterization required for convective dryer simulation

	Ultimate analysis (% in dry basis)	Proximate analysis (% in dry basis)	Particle size distribution (mm, %)	
C	47.55	Ash	4.03	PS>2 68.48



H	8.09	VM	88.31	1.4<PS<2	16.24
N	5.56	FC	7.66	1<PS<1.4	6.03
O	38.13	Moisture (wb)	81.77	0.85<PS<1	1.28
S	0.67			PS<0.85	7.97

Figure 2 a) shows the drying kinetic of BSGs in a batch type convective dryer at different drying temperatures between 60 °C and 120 °C. The drying time reduction from 4 to 1.25 hours was observed when the temperature increased from 60 °C to 120 °C. This change occurs because the increase in temperature causes a faster water evaporation rate from 0.43 to 2.45 g/kg-s, which can be graphically seen in the slope of the curve from the first period of constant rate. The critical and equilibrium moistures vary from 53.9 % to 72.4 % and 0.63 % to 3.86 % in wet basis, respectively.

Figure 2 b) shows the mass and heat transfer coefficients estimated from the results of the drying kinetic study. It is observed that the heat and mass transfer coefficients increased at high temperatures because the increase in the inlet drying temperature intensifies the water evaporation rate (Beigi et al., 2017a). Similar results were obtained using a tray dryer for pumpkin (Guiné et al., 2012) in which the mass transfer coefficient increased from  $0.1 \times 10^{-5}$  m/s to  $4.1 \times 10^{-5}$  m/s when the inlet air drying temperature increased from 30 °C to 70 °C.

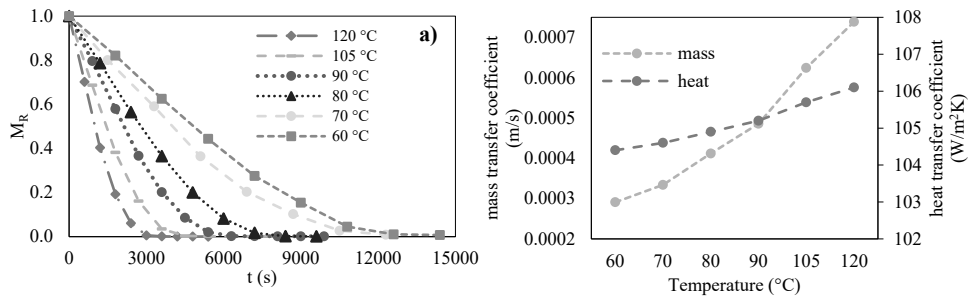


Figure 2. a) Drying kinetic curves for BSGs, b) mass and heat transfer coefficients at different temperatures.

### 3.2. Model validation

The drying process was simulated in Aspen Plus using the data obtained from the characterization of the BSGs and the kinetic drying study. The final moisture obtained in the simulation and in the drying experiments were compared to validate the model developed for the dryer. Table 3 shows that the simulation results have relative errors between 3.0 and 4.6%. Therefore, the developed drying model can describe the behavior of the drying process between 60 °C and 120 °C.

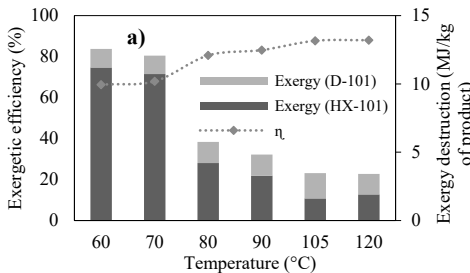
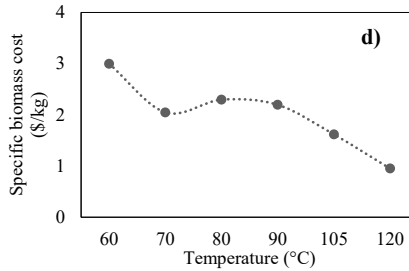
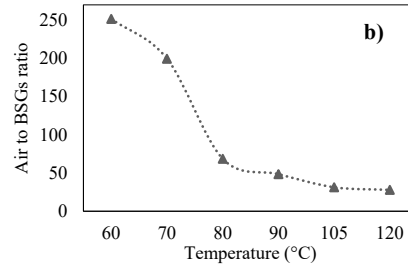
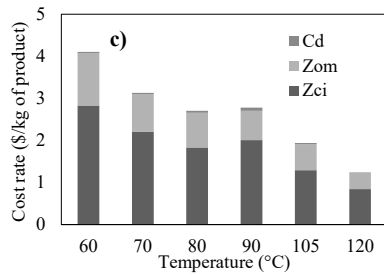
Table 3. Model validation of the batch type convective dryer.

Temperature (°C)	Moisture (wet basis)		Relative Error (%)
	Experiment	Simulation	
60	2.610	2.729	4.56
70	3.861	3.979	3.05
80	0.635	0.656	3.28

90	0.819	0.856	4.55
105	0.715	0.749	4.64
120	1.231	1.279	3.90

### 3.3 Exergoeconomic Analysis

The exergoeconomic indicators of the simulated BSGs convection drying process at different drying air temperatures were determined. Figure 3 a) shows that the exergetic efficiency increases significantly between 70 °C and 80 °C, while exergetic efficiency varies insignificantly at higher temperatures. This behavior is similar to the overall exergy destruction rate. The greatest reduction of destroyed exergy (52.3%) occurs between 70 °C and 80 °C because the required drying air flow decreases by 66%, according to Figure 3



b). Moreover, at higher temperatures, less drying air is required. Therefore, the drying airflow is directly proportional to the destroyed exergy rate. (Beigi et al., 2017b) obtained similar results, where rice is dried in a convective dryer and the exergetic efficiency increased from 33% to 68% when the drying temperature increased from 40 °C to 60 °C.

Figure 3. The effect of drying temperature in a) global exergy destruction and exergetic efficiency per component, b) air to BSGs ratio, c) exergy destruction, investment, operational and maintenance cost rates of the drying system, and d) specific biomass cost.

Figure 3 c) shows that at 120 °C the operating costs are the lowest and the cost per destroyed exergy. This change occurs because, at 120 °C, the lowest air/flue gas ratio was used. Thus, there is a need to burn less fuel for flue gas production. Also, previous studies have demonstrated that implementating other technologies, such as microwave-assisted dryers, would reduce operating costs (Ranjbaran & Zare, 2013). Furthermore, (Agbede et al., 2020) recommended implementating renewable energy sources such as solar energy in the air heater of drying processes due to their lower energy requirement. (Atalay, 2019)

compared conventional heat pump drying with solar-assisted heat pump and found that this reduces investment and operating costs by 20.3% and 34.5%, respectively. Figure 3 d) shows that the specific cost of biomass decreases significantly with the temperature increment. At 120 °C, it is possible to reduce the cost of dry biomass by 69.3%. This cost reduction responds to the low operating costs achieved at 120 °C. The minimum cost of BSGs was \$0.95/kg, which is half that obtained in a previous study for dry spent coffee grounds (Tinoco-Caicedo et al., 2021).

#### 4. Conclusions

This study aimed to simulate and perform an exergoeconomic analysis of a batch type convective dryer of BSGs based on experimental studies, carried out for air drying temperatures between 60 °C to 120 °C. This research shows that the exergetic efficiency varies between 66.4% and 88.0% when increasing the temperature from 60 °C to 120 °C. Production costs can be reduced by 69.1% by increasing the temperature from 60 °C to 120 °C. The specific cost of biomass varies between 3.00 \$/kg and 0.95 \$/kg and is at 120 °C when the lowest cost is reached. Finally, the results suggest that convective drying should operate at high temperature because it benefits exergetic efficiency and operating costs. Future studies should focus on experimental studies using other drying technologies that increase the drying rate and reduce both costs and the rate of destroyed exergy.

#### References

- Agbede, O. O., Oke, E. O., Akinfenwa, S. I., Wahab, K. T., Ogundipe, S., Aworanti, O. A., Olatunde Arinkoola, A., Agarry, S. E., Ogunleye, O., Nihinlola Osuolale, F., & Babatunde, A. (2020). Thin layer drying of green microalgae (*Chlorella* sp.) paste biomass: Drying characteristics, energy requirement and mathematical modeling. *Bioresource Technology Reports Journal*, 11.
- Atalay, H. (2019). Comparative assessment of solar and heat pump dryers with regards to exergy and exergoeconomic performance. *Energy*, 189.
- Beigi, M., Tohidi, M., & Toriki-Harchegani, M. (2017a). *Exergetic analysis of deep-bed drying of rough rice in a convective dryer*.
- Beigi, M., Tohidi, M., & Toriki-Harchegani, M. (2017b). Exergetic analysis of deep-bed drying of rough rice in a convective dryer. *Energy*, 140, 374–382.
- Bejan, A., & Tsatsaronis, G. (2021). Purpose in thermodynamics. *Energies*, 14(2).
- Bejan, A., Tsatsaronis, G., & Moran, M. (1996). *Thermal Design & Optimization*. John Wiley & Sons, Inc.
- Capossio, J. P., Fabani, M. P., Reyes-Urrutia, A., Torres-Sciancalepore, R., Deng, Y., Baeyens, J., Rodriguez, R., & Mazza, G. (2022). *Sustainable Solar Drying of Brewer's Spent Grains: A Comparison with Conventional Electric Convective Drying*.
- Castro, L. E. N., Matheus, L. R., & Colpini, L. M. S. (2022). Optimization of brewers' spent grain drying process. *Brazilian Journal of Development*, 8(2), 14481–14488.
- Cuevas, M., Martínez-Cartas, M. L., Pérez-Villarejo, L., Hernández, L., García-Martín, J. F., & Sánchez, S. (2019). Drying kinetics and effective water diffusivities in olive stone and olive-tree pruning. *Renewable Energy*, 132, 911–920.

- Guiné, R. P. F., Henriques, F., & Barroca, M. J. (2012). Mass Transfer Coefficients for the Drying of Pumpkin (*Cucurbita moschata*) and Dried Product Quality. *Food and Bioprocess Technology*, 5(1), 176–183.
- Incropera, F. P., Dewitt, D. P., Bergman, T. L., & Lavine, A. S. (2007). *Fundamentals of Heat and Mass Transfer* (J. Hayton (ed.); 6th ed.). John Wiley & Sons, Inc.
- Mallen, E., & Najdanovic-Visak, V. (2018). Brewers' spent grains: Drying kinetics and biodiesel production. *Bioresource Technology Reports*, 1, 16–23.
- Mishra, P. K., Gregor, T., & Wimmer, R. (2017). Utilising brewer's spent grain as a source of cellulose nanofibres following separation of protein-based biomass. *BioResources*, 12(1), 107–116.
- Rachwał, K., Waško, A., Gustaw, K., & Polak-Berecka, M. (2020). Utilization of brewery wastes in food industry. *PeerJ*, 8, e9427.
- Ranjbaran, M., & Zare, D. (2013). Simulation of energetic- and exergetic performance of microwave-assisted fluidized bed drying of soybeans. *Energy*, 59, 484–493.
- Srikiatden, J., & Roberts, J. S. (2007). Moisture transfer in solid food materials: A review of mechanisms, models, and measurements. *International Journal of Food Properties*, 10(4), 739–777.
- Tinoco-Caicedo, D. L., Mero-Benavides, M., Santos-Torres, M., Lozano-Medina, A., & Blanco-Marigorta, A. M. (2021). Simulation and exergoeconomic analysis of the syngas and biodiesel production process from spent coffee grounds. *Case Studies in Thermal Engineering*, 28(October), 101556.
- Towler, G., & Sinnott, R. (2012). *Chemical Engineering Design - Principles, Practice and Economics of Plant and Process Design* (2nd ed.). Elsevier Inc.



# Incorporating negative emissions technologies with policy instruments for net-zero emissions

Elizabeth J. Abraham, Patrick Linke, Dhabia M. Al-Mohannadi\*

*Department of Chemical Engineering, Texas A&M University at Qatar, Education City,  
PO Box, 23874, Doha, Qatar*

## Abstract

The ongoing climate crisis requires the reduction and removal of carbon to avoid global catastrophic consequences to the environment. As such, efforts presently focus on understanding the impact of negative emissions technologies in actively removing carbon dioxide from the atmosphere. While these emerging technologies bring a distinctive set of feasibility challenges, their integration into existing systems is essential due to the inability of decarbonization strategies to solely mitigate greenhouse gas emissions. The objective of this work is to develop strategies for integrating negative emissions technologies with carbon capture utilization and storage options, spurred by the increasing relevance of circular economies. The multi-period resource integration approach will determine the optimal technology portfolios best suited to a specific industrial landscape to achieve net-zero emissions in the presence of environmental policy instruments across time. The analysis will provide decision makers with a preliminary understanding of the viability of these technologies and the progress needed in regulatory frameworks to enable their effective implementation. The applicability of the proposed approach is demonstrated through the deployment of negative emissions technologies, such as bioenergy carbon capture and storage, direct air carbon capture and storage, enhanced weathering of minerals, and utilization of biochar, in an oil and gas based industrial setting. The optimal strategies determined indicate that carbon caps and taxes are fundamental in reducing the creation of emissions, while carbon credits encourage the active removal of carbon dioxide from the atmosphere.

**Keywords:** Negative emissions technologies, Process integration, Optimization, Carbon management

## 1. Introduction

With the climate crisis at the forefront of the sustainable development agenda, oil and gas economies that account for 15-20% of global greenhouse gas emissions (Gatto et al., 2021) will play a critical role in achieving net-zero emissions by 2050. To meet these targets, negative emissions technologies or NETs will need to be integrated into these economies along with more rampant decarbonization efforts that focus on the improvement of process efficiencies, transition to renewable energy, and deployment of carbon capture and storage (CCS) options. NETs fundamentally remove greenhouse gases from the atmosphere through various mechanisms, with their primary target being carbon dioxide (CO<sub>2</sub>) (McLaren, 2012). Some examples of NETs include technological solutions such as bioenergy with carbon capture and storage (BECCS) and direct air carbon capture and storage (DACCS), natural solutions such as afforestation and

reforestation, and enhanced natural solutions such as the use of biochar and enhanced weathering (IEA, 2020). Recently, these technologies have gained significant traction in industrial and academic research and development domains due to the major roles they play in the scenarios analyzed by integrated assessment models towards attaining carbon neutrality, where both carbon removal and reduction are of paramount importance (Tavoni & Socolow, 2013).

A significant portion of the extensive literature in this direction belongs to BECCS and DACCS, detailing the critical feasibility challenges that prevent their large-scale deployment and the measures to overcome them. Of the two, DACCS shows promising potential for deployment in oil and gas economies due to their large energy reserves, which can power the capture process and provide a means through the depleted reserves to store the captured carbon (Babonneau et al., 2022). While storage concerns for BECCS can also be managed similarly, there are concerns surrounding the sustainable procurement of the biomass required that additionally needs to be addressed. However, if obtained in a viable manner, the bioenergy it produces is a valuable commodity that can provide power for various operations such as those of other NETs, as suggested by Abraham et al. (2022c), who investigated the feasibility of such synergistic exchanges between NETs. Alternatively, instead of storing the carbon removed in geological storage sites, principles of circular economy suggest that it can be redirected for utilization in CCUS options (Abraham et al., 2022a). Therefore, to create an integrated system that embodies circular economy, all material and energy resources in a system must be analyzed. Towards enabling such circularity in systems, Abraham et al. (2022b) developed a multi-period optimization approach based on the resource integration method proposed by Ahmed et al. (2020) for designing processing clusters. This work adopts the multi-period resource integration approach to analyze the viability of integrating NETs with CCUS options and policy instruments in energy-rich economies to further devise deployment strategies over time where the carbon removed is specifically stored in environmental compartments in their customary sense to achieve net-zero emissions.

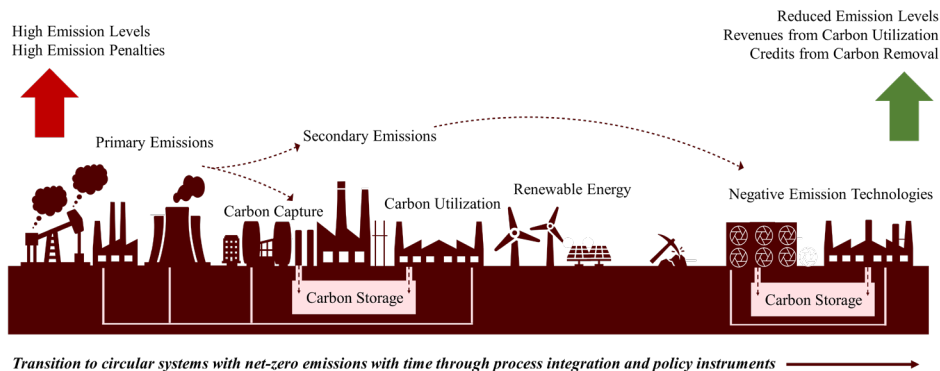


Figure 1: Overview of the proposed strategy for net-zero emissions

## 2. Approach

The multi-period resource integration model adopted in this work modifies the approach proposed by Abraham et al. (2022b) to develop planning strategies set to achieve carbon neutrality using networks of NETs and CCUS options. The model formulation takes the form of a mixed-integer linear program (MILP) to develop an integrated deployment strategy where further synergistic opportunities that are unapparent when solely focusing

on carbon can be explored. For a given carbon removal and reduction target, it identifies the optimal network that achieves the maximum net present value (NPV) over a planning horizon as given by Eq(1):

$$NPV_{Deployment\ strategy} = \sum_{Period} \frac{\sum_{Process} Revenue - Operating\ Cost - Capital\ Cost}{(1 + Discount\ Rate)^{Number\ of\ periods\ considered}} \quad (1)$$

The optimization determines which processes become active in a period and the capacities they operate at to meet the specified target. The carbon removal and reduction options are each represented as processes that consume certain resources to produce certain others using explicit resource parameters. Each of the resources is then represented through a resource line defined by a unique set of specifications to facilitate integration between processes that share common resources. The total costs incurred depend on the capital and operating costs of the processes, along with any stipulated inflation rates, resource penalties, and credits considered. Furthermore, equality and inequality constraints can be placed on the operational capacities and resource flows throughout the network according to the specific details of the scenario analyzed, along with resource balances that monitor the flow of resources to, from, and between processes.

### 3. Illustrative Example

To illustrate the application of this approach, a hypothetical scenario is investigated over a planning horizon that consists of three periods spanning 10 years each. The processes considered for capturing, storing, utilizing, and removing CO<sub>2</sub> are listed in Table 1, along with their reference products and CAPEX parameters. The process parameters for the carbon removal options, namely, the BECCS, biomass pelletization unit, biochar production unit, DACCS, and EW mineral processing unit, were retrieved from Abraham et al. (2022c), while those of the sequestration unit and the production and capture units of methanol and urea, were determined from Ahmed et al. (2020). Each process operates at a maximum capacity of 25 million tons of their respective reference products in the first period, after which the capacities of the NETs and biomass pelletization unit increase by 20% from the previous period.

Table 1: Reference products and CAPEX parameters

Processes (Process ID)	Reference Product	CAPEX Parameter (Million \$)
Bioenergy Carbon Capture and Storage (BECCS)	CO <sub>2</sub> removed	2,478
Biomass Pelletization Unit (BMP)	Biomass pellets	130
Biochar Production Unit (BCP)	CO <sub>2</sub> removed	1,601
Direct Air Carbon Capture and Storage (DACCS)	CO <sub>2</sub> removed	2,175
EW Mineral Processing Unit (EW)	CO <sub>2</sub> removed	71
Fischer-tropsch Unit (FTU)	Fischer-tropsch fuels	248
Methanol Production Unit (MPU)	Methanol	262
Methanol Carbon Capture Unit (MCC)	CO <sub>2</sub> captured	7
Synthetic Natural Gas Production Unit (SNG)	SNG	39
Sequestration (SQU)	CO <sub>2</sub> captured	12
Urea Production Unit (UPU)	Urea	135
Urea Carbon Capture Unit (UCC)	CO <sub>2</sub> captured	7

The data for the SNG production unit was obtained from Chauvy et al. (2021), while those for the FT unit were obtained from Panahi et al. (2018) and Al-Yaeshi et al. (2020). The resource parameters for these processes are summarized in Table 2, along with the prices of their associated resources. The prices per ton of the resources not listed here were set at \$294 for ammonia, \$1,383 for biochar, \$112 for biomass pellets, \$509 for calcium carbonate, \$14 for low pressure steam, \$305 for methanol, \$484 for propane, \$24 for



residues, \$26 for rock in, and \$288 for urea. All prices were set to vary at an inflation rate of 5%.

Table 2: Resource parameters of the FT and SNG units along with their prices (2020 US\$)

Processes (Resource unit)	FT Unit	SNG Production Unit	Prices (\$/ Resource unit)
CO <sub>2</sub> emitted (tons)	3.85	-	-
CO <sub>2</sub> utilized (tons)	-2.07	-2.41	-
Cooling water (tons)	-	-23.69	0.03
Electricity (kWh)	-	-17.14	0.03
FT Fuels (tons)	1.00	-	1,029
High pressure steam (tons)	3.88	-	16
Hydrogen (tons)	-	-0.45	1,178
Medium pressure steam (tons)	4.70	3.71	15
Natural gas (tons)	-2.01	-	100
Saturated steam (tons)	-3.56	-	15
Synthetic natural gas (tons)	-	1.00	100
Wastewater (tons)	-	23.69	-
Water (tons)	-	-3.71	5
Water out (tons)	4.22	1.86	-

Certain resources in the processes considered contain CO<sub>2</sub> that must be accounted for to ensure climate targets are met in time. As such, the compositions of these resources, specified in Table 3, will also be incorporated into the overall carbon flow analysis.

Table 3: CO<sub>2</sub> compositions of the carbon carrying resources

Resource	CO <sub>2</sub> Composition
CO <sub>2</sub> emitted, captured, utilized, removed	1.00
Methanol process emissions	0.30
Methanol capture unit emissions	0.06
Urea process emissions	0.26
Urea capture unit emissions	0.03

Two cases were investigated with this setup to maximize the NPV of the deployment strategy analyzed at a discount rate of 10%. Every period receives a feed of CO<sub>2</sub> captured from various industrial processes associated with the oil and gas industry. With each period, this feed decreases at a rate of 25% as changes to conventional processes are expected to lower the emissions they release. To achieve net-zero emissions at the end of the planning horizon, constraints are placed on the overall flow of CO<sub>2</sub> from the system. While the first period has no specific constraints, the CO<sub>2</sub> that leaves the network in the second and third periods, defined as the difference between the carbon created and removed by it, are constrained with carbon caps, as specified in Table 4.

Table 4: Carbon feed and cap limits over the three periods

Periods	Period 1	Period 2	Period 3
CO <sub>2</sub> fed (tons)	175,000,000	131,250,000	98,437,500
CO <sub>2</sub> cap (tons)	No constraint	65,625,000	0

The processes activated for the two cases and their operational capacities in each period are given in Table 5. In the first case, net-zero emissions are achieved by transitioning from focusing solely on carbon utilization to additionally incorporating carbon removal. The optimal strategy determined has an NPV of \$32 billion and removes nearly 128 million tons of CO<sub>2</sub> using the BECCS, DACCS, and EW processes. These NETs only become active when constraints are placed on the emissions leaving the system. In contrast, periods such as the first, without any constraints, focus on generating revenues through carbon capture and utilization through the production of FT fuels, methanol, and

urea. In later periods, there is a decrease in carbon utilization processes with high operating costs, such as SNG production, and those which lead to further emissions, like methanol production. This decrease, when simultaneously implemented with an increase in carbon removal through NETs, permits the system, over time, to lower sequestration capacities and leads to the cumulative release of about 295 million tons of CO<sub>2</sub>.

Table 5: Operational capacities in million tons of the reference product determined for the three cases in each period and their set carbon tax and credit rates

Case	Case 1			Case 2		
	tp1	tp2	tp 3	tp1	tp2	tp 3
Time periods	tp1	tp2	tp 3	tp1	tp2	tp 3
Carbon tax	0	0	0	150	300	450
Carbon credit	0	0	0	200	400	600
<b>Processes</b>						
BMP	-	22.21	26.65	18.51	22.21	26.65
BECCS	-	30.00	36.00	25.00	30.00	36.00
DACCS	-	-	27.10	25.00	27.12	27.10
EW	-	1.26	33.66	25.00	30.00	36.00
FTU	25.00	25.00	25.00	25.00	0.10	0.10
MPU	25.00	25.00	17.95	25.00	25.00	25.00
MCC	-	2.56	1.84	2.56	2.56	2.56
SNG	14.98	0.10	0.10	16.69	19.92	6.31
SQU	25.00	21.23	0.10	25.00	25.00	25.00
UPU	25.00	25.00	25.00	25.00	25.00	25.00
UCC	0.10	1.66	1.66	1.66	1.66	1.66

In the second case, carbon taxes and credits that doubled and tripled from their rates in the first period are implemented as described in Table 5, along with the emission caps detailed previously. The increasing carbon tax rates further limit the release of emissions, while the increasing carbon removal credits incentivize the deployment of NETs with the overall network emitting and removing about 99 and 261 million tons of CO<sub>2</sub>, respectively. Over time, emissions were drastically reduced from 97 million tons in the first period to nearly only a million tons of CO<sub>2</sub> in the second and third periods. The reduction was realized by decreasing carbon utilization in the FT unit, increasing carbon capture from the two capture units, and increasing sequestration to operate at its maximum capacity throughout the periods considered. The operations of the FT unit were specifically reduced while methanol and urea production remained constant as these processes have capture units to manage the emissions they create. On the other hand, the NETs operating at or near their maximum capacities enable the network to achieve the increased carbon removal required. The only exception here is the biochar process, which fails to be deployed in any period as it releases some emissions. An overview of the different CO<sub>2</sub> flows in the two cases explored through the proposed approach is illustrated in Figure 2.

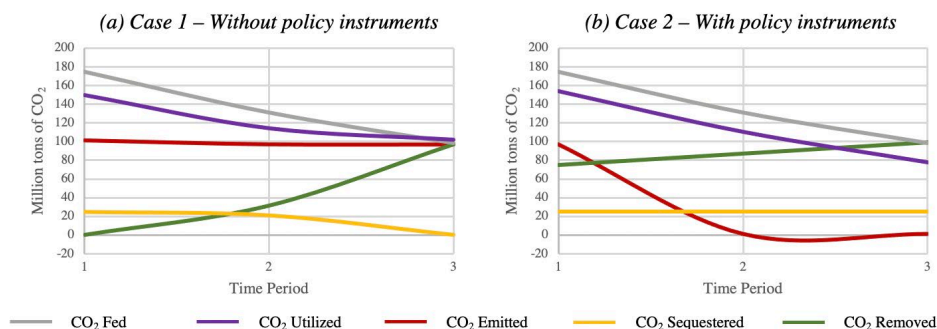


Figure 2: Overview of CO<sub>2</sub> flows (a) without and (b) with policy instruments

With time, the carbon caps ensure that the amount of carbon removed from the atmosphere increases while carbon utilization decreases as it still leads to the creation of emissions. Additionally, the carbon tax and credit implemented in the second case further reduce emissions and promotes the deployment of NETs. Of the policy instruments in place, the carbon credits were particularly pivotal in achieving net-zero emissions and generating profits in the second case with an NPV of \$45 billion. On further analysis, the network can break even with carbon credits set as low as \$53.37 per ton CO<sub>2</sub> removed in the first period. Furthermore, with the great uncertainty encompassing the capital required for the NETs and their associated processes, adjusting these carbon credit rates appropriately can encourage their deployment. In fact, with their CAPEX parameters doubled, the second case can break even at rates as low as \$68.73 per ton CO<sub>2</sub> removed in the first period.

#### 4. Conclusion

The approach described in this work enables the assessment of policy instruments such as carbon caps, taxes, and credits with CCUS and NETs options while identifying potential for material and energy integration between them. Future work should integrate renewable energy with these technologies and additionally consider their land and supply chain feasibility. Furthermore, the non-linear expressions used in pricing the carbon taxes and credits can be simplified to allow for the analysis of more network configurations to achieve climate targets. Thus, the multi-period multi-resource approach described here allows policymakers to propose guidelines to achieve net-zero emissions.

#### References

- E.J. Abraham, D.M. Al-Mohannadi, P. Linke, 2022a, Integrating Carbon Negative Technologies in Industrial Clusters, *Computer Aided Chemical Engineering*. 49, 763–768.
- E.J. Abraham, D.M. Al-Mohannadi, P. Linke, 2022b, Resource integration of industrial parks over time, *Computers & Chemical Engineering*. 107886.
- E.J. Abraham, P. Linke, D.M. Al-Mohannadi, 2022c, Optimization of low-cost negative emissions strategies through multi-resource integration, *Journal of Cleaner Production*, 372, 133806.
- R. Ahmed, S. Shehab, D.M. Al-Mohannadi, P. Linke, 2020, Synthesis of integrated processing clusters, *Chemical Engineering Science*, 227, 115922.
- A.A. Al-Yaeeshi, A. AlNouss, G. McKay, T. Al-Ansari, 2020, A simulation study on the effect of CO<sub>2</sub> injection on the performance of the GTL process, *Computers & Chemical Engineering*. 136, 106768.

- F. Babonneau, M. Benlahrech, A. Haurie, 2022, Transition to zero-net emissions for Qatar: A policy based on Hydrogen and CO<sub>2</sub> capture & storage development, *Energy Policy*, 170, 113256.
- R. Chauvy, D. Verdonck, L. Dubois, D. Thomas, G. De Weireld, 2021, Techno-economic feasibility and sustainability of an integrated carbon capture and conversion process to synthetic natural gas, *Journal of CO<sub>2</sub> Utilization*, 47, 101488.
- A. Gatto, W. Loewenstein, E.R. Sadik-Zada, 2021, An extensive data set on energy, economy, environmental pollution and institutional quality in the petroleum-reliant developing and transition economies, *Data in Brief*, 35, 106766.
- IEA, 2020. Going carbon negative: What are the technology options?
- D. McLaren, D, 2012, A comparative global assessment of potential negative emissions technologies, *Process Safety and Environmental Protection*, 90, 489–500.
- M. Panahi, E. Yasari, A. Rafiee, 2018, Multi-objective optimization of a gas-to-liquids (GTL) process with staged Fischer-Tropsch reactor, *Energy Conversion and Management*, 163, 239–249.
- M. Tavoni and R. Socolow, 2013, Modeling meets science and technology: an introduction to a special issue on negative emissions, *Climatic Change*, 118, 1–14.



# Waste valorization of non-commercialized edible mushrooms

Brenda I. Lazaro-Molina,<sup>a</sup> Teresa Lopez-Arenas<sup>a</sup>

<sup>a</sup> *Departamento de Procesos y Tecnología, Universidad Autónoma Metropolitana – Cuajimalpa, Vasco de Quiroga 4871, Mexico City 05348, Mexico*  
*mtlopez@cua.uam.mx*

## Abstract

The production of edible mushrooms generates two types of waste: the spent substrate from growing mushrooms and a loss of non-marketed mushrooms. Currently, great interest has been placed in the recovery of the spent substrate through the development of biorefineries for bioethanol production, mainly. However, the use of non-marketable mushrooms as raw material has not been extensively investigated. Therefore, the main objective of this work is the proposal for the synthesis, design and evaluation of a biorefinery of non-marketable edible mushrooms to obtain chitin, citric acid and animal feed, as high value-added products. The methodology proposes a technical-economic-environmental evaluation for decision-making. As a result, a viable conceptual design is obtained, with positive profitability and reduction of environmental impact factors.

**Keywords:** biorefinery, edible mushrooms, waste valorization, process simulation.

## 1. Introduction

Edible mushrooms constitute a functional food with nutritional and medicinal properties that promote health, which can strengthen food security in times of crisis. Their production has increased more than 30 times since 1978, reaching a world production of 34 million tons (Royse et al., 2017). The most widely cultivated edible mushroom worldwide is *Agaricus bisporus* (common white mushroom), followed by *Lentinus edodes*, *Pleurotus spp.* and *Flammulina velutipes* (Kumar et al., 2021), which are regulated according to different standards to guarantee organoleptic qualities and consumer satisfaction. Depending on its quality, there is an amount of non-marketable waste that ranges between 5 - 20% of the production.

It is important to highlight that there are proposals for the use of waste lignocellulosic substrates (such as cereal straws) employed in mushroom cultivation (Awasthi et al., 2022), but there are no biorefineries that take advantage of mushroom waste, since they are currently used directly as animal feed or fertilizer. Thus, the objective of this work is to design a biorefinery to produce high value-added products from these (non-marketable) edible mushroom residues, under a circular economy perspective.

Agro-industrial residues are unavoidable and should be considered as a new opportunity for their reuse within a biorefinery, using environmentally friendly and profitable bioprocess technologies. In such a way that these wastes from biorenewable resources can be used for the production of food, animal feed, fertilizers, fuels, energy, amino acids, biopolymers, among other industrial chemical products. From a review of the literature, it was found that the product with the highest value that can be obtained from edible

mushrooms is chitin, which, together with its derivatives, show low immunogenicity and therefore have a great use in biomedical applications. As a second product of the biorefinery, citric acid is considered due to its wide use as an additive in the food, pharmaceutical, cosmetic, and cleaning industries. And since citric acid can be processed by microbial culture, then the biomass resulting from cell growth with high nutritional value can be processed for animal feed.

In this work, first, a biorefinery conceptual design is proposed for the waste valorization of non-commercialized edible mushrooms to produce chitin, citric acid and animal feed. Then the biorefinery flowsheet diagram is implemented and evaluated in a modular process simulator. Afterwards, a parametric sensitivity analysis is performed, considering equipment dimensions, amount of waste fed, design configurations, etc., as a result: the technical evaluation demonstrates the feasibility of using the waste, while the economic evaluation presents a positive profitability of the proposal using as indicators the unit production cost, the return on investment and its payback time; and the environmental impact is evaluated considering the consumption of water, utilities and energy.

## 2. Methodology

The methodology consists first of determining the composition of the edible mushroom residues, then identifying the best reaction routes to obtain value-added products, then determining a process flow diagram, and finally evaluating the biorefinery in a comprehensive manner considering technical, economic and environmental impact aspects. Each of these stages is briefly described below.

### 2.1. Mushroom composition

The bodies of the fruits of *A. bisporus* are rich in proteins (29.14%), carbohydrates (51.05%) and fats (1.56%); and also include proteins, lipids, vitamins and minerals (Ahlawat et al., 2016). For simulation purposes, the composition is established according to the compounds of interest (chitin and glucans present), while the rest of the compounds were grouped under the name of other solids. Therefore, the composition used was: 90% water, 0.47% chitin, 1.01% glucans, and the rest of other solids (Nitschiske et al., 2011; Khan et al., 2017).

### 2.2. Conceptual design of the biorefinery

A biorefinery is proposed to obtain three high-value products, through the following processing stages:

- (a) Conditioning of the mushroom waste, which consists of washing, drying, grinding and sieving the mushrooms to obtain dry flour.
- (b) Chitin extraction, where the deproteinization process is first carried out by enzymatic hydrolysis using proteases. Immediately a thermal hydrolysis allows extracting compounds such as lipids, minerals, vitamins and remaining proteins. As a result, supernatant rich in soluble glucans and a precipitate rich in chitin are obtained (Urbina-Salazar et al., 2020). The latter is sent to dry to obtain the chitin.
- (c) Saccharification and fermentation, where first the soluble glucans are converted to glucose by enzymatic hydrolysis. Then, glucose is converted to citric acid by fermentation with *Aspergillus niger*. Subsequently, the solid phase (residual biomass) is separated, and the liquid phase is subjected to a citric acid recovery process.
- (d) Animal feed production, by using the residual biomass from fermentation, which is dried and packaged for its sale.
- (e) Production of citric acid, where this product is first precipitated as calcium citrate, then treated with sulfuric acid and the gypsum precipitate formed (calcium sulfate) is

removed. Finally, the recovery of the product is carried out through a crystallization process to obtain dry crystals of citric acid.

Figure 1 shows the process flow diagram for the biorefinery of non-marketable edible mushrooms, according to this proposed conceptual design. Here it is important to mention that all liquid waste is sent to a wastewater treatment plant (WWT), which is not included but is considered in the evaluation.

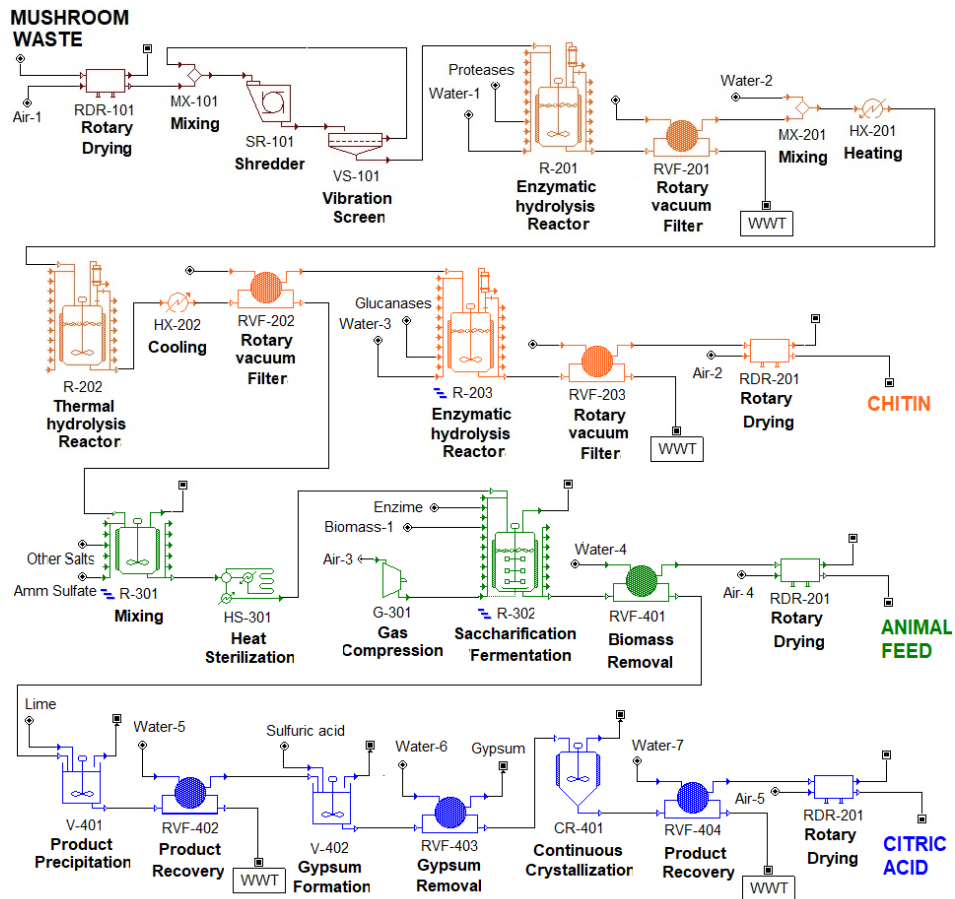


Figure 1. Process flow diagram of the proposed biorefinery.

### 2.3. Biorefinery assessment

The comprehensive evaluation of the biorefinery includes technical, economic and environmental aspects, as detailed below.

#### 2.3.1. Technical evaluation

The implementation of the conceptual diagram of the biorefinery considers technical and design aspects for an industrial scale, with which the mass and energy balances are evaluated taking into account the reactions in the processes and their conversions; the sizing of equipment and its efficiency; the working times of the equipment and processing cycles; as well as the installed capacity of the plant in terms of the amount of raw material fed and products generated.



### 2.3.2. Economic evaluation

Undiscounted cash flow profitability criteria provide an immediate assessment of process economics and are often used for the early stages of process synthesis and design (Seider et al., 2009). So, in this work, the economic criteria used to determine if a project is profitable are the unit production cost (UPC), the return on investment (ROI) and the payback time (PT) of the investment.

### 2.3.3. Environmental evaluation

A successful economic development also depends on the rational use of resources and the reduction of the adverse environmental impact of the development project (Dixon and Pagiola, 1998). Therefore, this work evaluates (a) the consumption of water in the process, due to problems of water scarcity and damage to ecosystems; (b) energy consumption, as it is a key element in climate change; and (c) the demand for heat transfer agents, such as steam and cooling water.

## 3. Results

The biorefinery flow diagram was implemented in *SuperPro Designer*, a modular process simulator to evaluate the mass and energy balances that allow to calculate the technical-economic-environmental criteria. It is important to mention that the evaluation of the economic criteria depends on the unit production revenues, which were estimated at 200, 2.0, 0.2 USD/kg for chitin, citric acid and animal feed, respectively (Pighinelli et al, 2019; Heinzle et al. al., 2017; Anaya-Reza and Lopez-Arenas, 2018). With the purpose of carrying out a profitability analysis of the biorefinery, first, the feeding of edible mushroom waste was set at 50 T/batch, and later, a sensitivity analysis was carried out based on the number of reactors installed.

### 3.1. Base Case

Considering a 50 T/batch mushroom feed and an annual operating time of 336 days, the recipe batch time was calculated at 14.7 days, processing 50 batches per year and producing 301.6 kg of chitin, 21.3 kg of animal feed and 178.6 kg of citric acid per batch. The direct fixed capital, working capital, and startup cost were estimated as 13 176 000, 212 000, and 659 000, USD, respectively, giving a total investment of 14 047 000 USD for the biorefinery installation. The calculated value for UPC was 390.13 USD/kg, which was too high compared to the unit production revenue, so the ROI was -2.88% and it was not possible to assess the PT. The latter are summarized in the first line of Table 1.

Table 1. Profitability analysis based on the number of reactors installed in R-203, 301 and 302

Number of reactors	UPC (USD/kg)	ROI (%)	PT (years)	Amount of mushroom waste (T/year)
1	390	-2.9	NA	2 500
2	305	2.6	37.8	4 950
3	278	6.7	15.0	7 450
4	261	9.6	10.4	9 900
5	252	11.9	8.4	12 400
6	256	10.7	9.4	12 450

### 3.2. Profitability analysis

Since the initial base case was not profitable, the equipment occupancy was analyzed and it was determined that the reactors R-203, R-301 and R-302 generate a bottleneck as shown in Fig. 3(a). So, a sensitivity analysis was carried out increasing the number of these reactors installed, managing to increase the profitability of the biorefinery as shown

in Table 1. As can be seen by increasing the number of reactors it is possible to increase the profitability of the process, achieving a maximum with 5 reactors. The increase in profitability is explained by the fact that dead times decrease, reaching an optimal equipment occupancy with 5 reactors. The comparison of processing times for 1 and 5 reactors can be seen in Fig. 3. So, increasing to 6 reactors, a bottleneck starts again and the annual amount of mushroom waste increases little (as observed in the last line of Table 1), so that cost effectiveness tends to decrease.

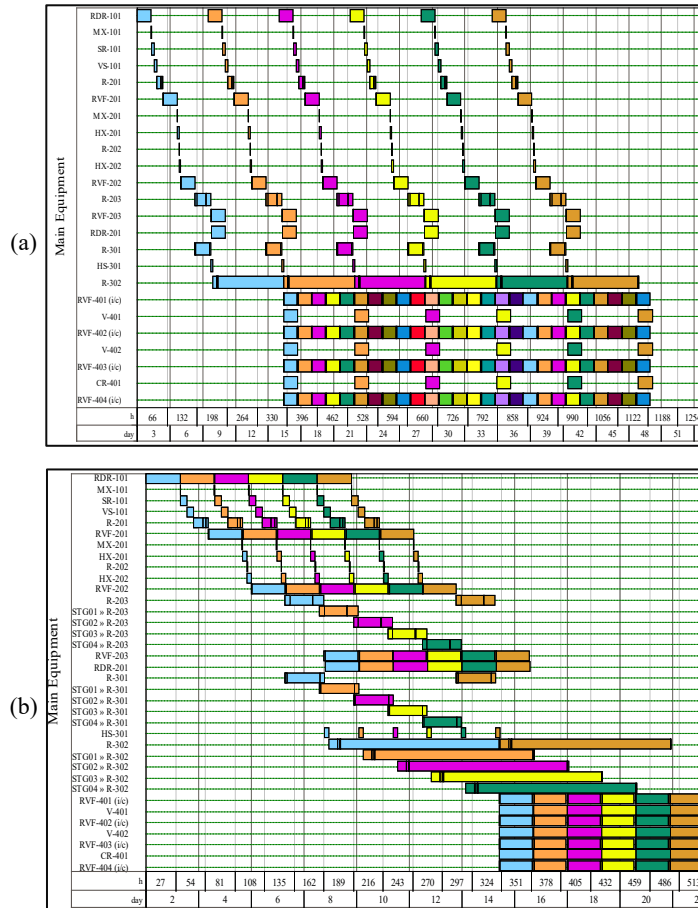


Figure 3. Equipment occupancy (R-203, R-301, R-302) considering (a) one and (b) 5 reactors..

### 3.3. Environmental results

Regarding the environmental impact indicators, Table 2 shows the values obtained for the base case (with 1 reactor) and the improved case (with 5 reactors). According to these results, it can be seen in general that a large amount of utilities is required, however, by increasing the number of reactors it was possible to reduce the use of heat transfer agent resources and energy consumption. Regarding the process water, it is important to highlight that although this water is required for all processing, the residual water is sent to a WWT plant and later recirculated. Therefore, the consumption of water after start-up is minimal, only to recover losses due to venting and evaporation.

Table 2. Environmental criteria for the base and improved cases.

Case study	No. reactors	Process water (T/kg chitin)	Utilities		
			Steam (T/kg chitin)	Cooling water (T/kg chitin)	Std. power (kW-h / kg chitin)
Base	1	349	0.45	22.8	51.5
Improved	5	271	0.35	17.3	46.4

#### 4. Conclusions

The need to create alternatives for the synthesis, design, and operation of biorefineries for the processing of high value-added products from non-commercialized edible mushrooms is a promising, profitable, and sustainable alternative. In this work, a conceptual design of a biorefinery has been proposed, and a comprehensive evaluation methodology has been presented that allows decision-making at an early design stage. Although the results show a positive profitability of the project, there are still various aspects to be studied, such as the modification of the installed capacity, the synthesis of other bioproducts (for example, lactic acid instead of citric acid), mixing various types of edible mushroom waste, alternatives to reduce the consumption of utilities (such as heat integration), etc.

#### References

- O. P. Ahlawat, K. Manikandan, M. Singh, 2016, Proximate composition of different mushroom varieties and effect of UV light exposure on vitamin D content in *Agaricus bisporus* and *Volvariella volvacea*. *Mushroom Res*, 25(1), 1-8.
- O. Anaya-Reza, T. Lopez-Arenas, 2018, Design of a sustainable biorefinery for the production of lactic acid from sugarcane molasses, *Revista Mexicana en Ingeniería Química*, 17(1), 243-259.
- M.K. Awasthi, S. Harirchi, T. Sar, et. al, 2022, Myco-biorefinery approaches for food waste valorization: Present status and future prospects, *Bioresource Technology*, 360, 127592.
- J. Dixon, S. Pagiola, 1998, Economic analysis and environmental assessment. *Environmental assessment sourcebook update*, 23, 1-21.
- E. Heinzle, A. Biber, C. Cooney, 2007, *Development of Sustainable Bioprocesses: Modeling and Assessment*. Chichester: Wiley
- A.A. Khan, A. Gani, F.A. Masoodi, U. Mushtaq, A.S. Naik, 2017, Structural, rheological, antioxidant, and functional properties of  $\beta$ -glucan extracted from edible mushrooms *Agaricus bisporus*, *Pleurotus ostreatus* and *Coprinus atramentarius*. *Bioactive Carbohydrates and Dietary Fibre*, 11, 67-74.
- H. Kumar, K. Bhardwaj, R. Sharma, E. Nepovimova, N. Cruz-Martins, D.S. Dhanjal, R. Singh, et al., 2021, Potential Usage of Edible Mushrooms and Their Residues to Retrieve Valuable Supplies for Industrial Applications. *Journal of Fungi*, 7(6), 427. MDPI AG.
- J. Nitschke, H.J., Altenbach, T. Malolepszy, H. Mölleken, H., 2011, A new method for the quantification of chitin and chitosan in edible mushrooms. *Carbohydrate Research*, 346(11), 1307-1310.
- L. Pighinelli, J. Broquá, J., B.G. Zanin, A.M. Flach, C. Mallmann, F.G.D. Taborda, L.E.L. Machado, S.M.L. Alves, M.M. Silvia, R.J.S.P. Dias, 2019, Methods of chitin production a short review. *Am. J. Biomed. Sci. Res*, 3(4), 307-314.
- D.J. Roysse, J. Baars, Q. Tan, 2017, Current Overview of Mushroom Production in the World, in *Edible and Medicinal Mushrooms*, 5-13.
- D.W. Seider, J. D. Seader, D. R. Lewin, S. Widagd, 2009. *Product and Process Design Principles: Synthesis, Analysis and Design*. John Wiley & Sons, New York.
- A.R. Urbina-Salazar, A.R. Inca-Torres, M.P. Carbonero-Aguilar, J. Bautista-Palomas, J., 2020, Preparación de quitina fúngica a partir de subproductos de hongos comestibles (*Agaricus bisporus*), *Polo del Conocimiento: Revista científico - profesional*, 5, 115-140.

# Enviro-economic assessment of sustainable aviation fuel production from direct CO<sub>2</sub> hydrogenation

Andrea Bernardi<sup>a,b,\*</sup>, Daniel Bagan Casan<sup>b</sup>, Andrew Symes<sup>c</sup>, Benoit Chachuat<sup>a,b</sup>

<sup>a</sup>*The Sargent Centre for Process Systems Engineering, Imperial College London, UK*

<sup>b</sup>*Department of Chemical Engineering, Imperial College London, UK*

<sup>c</sup>*OXCCU Tech Ltd, Oxford, UK*

\**a.bernardi13@imperial.ac.uk*

## Abstract

The aviation industry is responsible for 2% of the total GHG emissions and 10% of the fuel consumption worldwide and sustainable aviation fuel (SAF) is considered a key step towards achieving net-zero aviation. In this work, we carry out an enviro-economic comparison of a one-step Fischer-Tropsch process (1sFT), based on a novel Mn-Fe-K catalyst, whereby CO<sub>2</sub> and H<sub>2</sub> are directly converted to liquid hydrocarbons, with a two-steps FT process (2sFT), in which a reverse water gas shift reactor is used to produce syngas, followed by a conventional FT process. Our analysis considers 1 MJ of liquid fuel as functional unit and the following key performance indicators: levelized cost of production, global warming potential, and monetized end-point environmental impacts. Our results suggest that the fuel blend from 1sFT has a minimum selling price 20% lower than the fuel blend from 2sFT, due to a lower capital cost and a higher selectivity towards liquid hydrocarbons. 1sFT is also found to be superior to 2sFT from an environmental point of view, with 30% lower GWP and 70% lower externalities cost.

**Keywords:** sustainable aviation fuel, technoeconomic analysis, LCA, Fischer-Tropsch

## 1. Introduction

Decarbonisation of the transport sector is a key challenge in the path towards net-zero and a sustainable future. The aviation industry is responsible for 2% and 10% of the global GHG emissions and fuel consumption, respectively. Sustainable aviation fuel (SAF) is called upon to play a pivotal role in a net-zero aviation, with up to 80% GHG emissions reduction (IEA 2022a). According to IEA (2020) low-carbon hydrogen-derived liquid fuels will gain significant market share in the sector, as batteries and hydrogen are not feasible alternatives, especially for long-range flights, due to the low energy content per unit of mass and volume respectively (Karadotcheva *et al.* 2021). Moreover, hydrogen-fuelled aircrafts would require a complete redesign and advanced materials to avoid hydrogen embrittlement and leakage (Airbus 2020).

Fischer-Tropsch (FT) synthesis is a well-established chemical route to produce hydrocarbons of different chain length distributions from syngas, a mixture of CO and H<sub>2</sub>, and can produce a drop-in alternative fuel for the aviation industry. However, syngas is nowadays produced from steam methane reforming or coal gasification, carries a high environmental burden. Significant research efforts have been devoted to identifying routes to use CO<sub>2</sub> as the carbon source in the FT process and deliver sustainable fuels. One option is to convert CO<sub>2</sub> and H<sub>2</sub> in a reverse water gas shift (RWGS) reactor and use

the resulting syngas in a traditional FT process (Zang *et al.* 2021). The main limitations of this two-step approach are: (i) RWGS reaction requires high temperatures to achieve a good conversion (600–1000 °C); and (ii) the production of wax, a mixture of heavy hydrocarbons, requires additional upgrading units, such as a hydrocracking (HC) reactor to improve the yield of liquid fuels. The direct hydrogenation of CO<sub>2</sub> to liquid fuels in a single step is more appealing, but the selectivity towards liquid fuels is usually low, and the formation of short chain hydrocarbons is favored (Chen *et al.* 2018; Gao *et al.* 2017). Recently, Yao *et al.* (2020) synthesized a novel Mn-Fe-K based catalyst capable of converting CO<sub>2</sub> and H<sub>2</sub> with excellent yield and selectivity towards liquid hydrocarbons in the jet fuel range (C<sub>8</sub>–C<sub>16</sub>) and minimal wax production.

The main objective of this paper, therefore, is to conduct a comparative assessment of one-step (1sFT) and two-step (2sFT) FT synthesis from CO<sub>2</sub> and H<sub>2</sub> to produce liquid fuels, also with a fossil-based alternative, in terms of economic and environmental performance. The methodology is described in the next section, followed by results and discussions, before concluding the paper.

## 2. Methodology

Figure 1 presents block flow diagrams of the two processes considered in this paper, which are designed to utilize 100 t/h of CO<sub>2</sub> coming from a direct air capture (DAC) plant and H<sub>2</sub> from water electrolysis. We further assumed that the fuel blend is sent to an existing refinery.

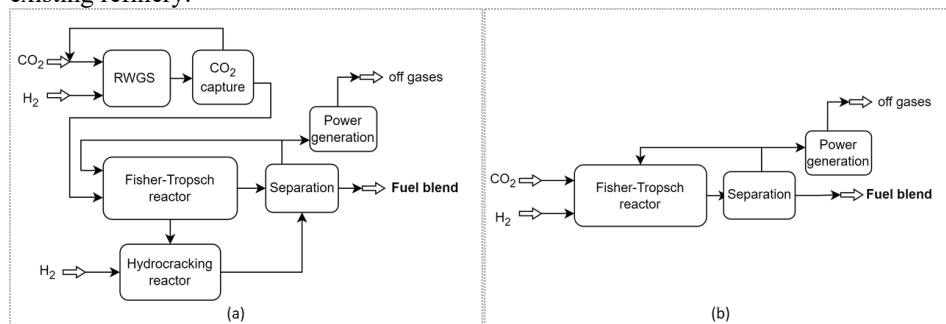


Figure 1: block flow diagrams of the two-step FT synthesis (a) and one-step FT synthesis (b). A wastewater stream is produced by both processes but is not represented in the diagram.

### 2.1. Process simulation and economic analysis

Flowsheet of both 1sFT and 2sFT routes were developed using Aspen HYSYS (version 11) and simulated in steady state, considering the Peng Robinson fluid package. Heat integration was performed using the Aspen Energy Analyzer. In the following the processes are described in more details.

**Feed streams.** We consider CO<sub>2</sub> from DAC as carbon source, available at 25 °C and 1 bar. According to IEA (2020), the cost of DAC CO<sub>2</sub> capture ranges between 135–345 \$/tCO<sub>2</sub> (note that all costs throughout the paper are expressed in US\$<sub>2021</sub>), in our base case the cost is 160 \$/tCO<sub>2</sub> as reported in Keith *et al.* (2016). H<sub>2</sub> is obtained from water electrolysis using wind energy, and it is available at 20 bar and 25 °C (Al Qahtani *et al.* 2021). The H<sub>2</sub> price is 4.2 \$/kg as reported in IRENA (2019) for wind-based water electrolysis. Both feed streams are assumed to have no impurities. An optimistic best-case scenario is also briefly discussed, considering 2050 predicted prices to compute the raw material costs (1.5 \$/kg<sub>H<sub>2</sub></sub>, IRENA 2019; 50 \$/tCO<sub>2</sub>, IEA 2022).

**One-step FT synthesis (1sFT).** The FT reactor uses the new Mn-Fe-K based catalyst and operates at 260 °C and 10 bar (Yao *et al.* 2020). A two-stage compressor with intercooling is used to bring the CO<sub>2</sub> at 10 bar, while no compression is needed for H<sub>2</sub>. The H<sub>2</sub> is fed to the process in a H<sub>2</sub>:CO<sub>2</sub> molar ratio of 3:1. The outlet stream enters a sequence of two 3-phase separators operating at 70 °C and 40 °C under 10 bar. A small (2%) fraction of the gaseous stream coming out of the second separator is burned to produce heat and electricity in a gas and steam cycle, while the rest is recycled to the FT reactor. The other two streams from the 3-phase separators are wastewater and liquid hydrocarbons. The liquid hydrocarbons are mixed, expanded to 1 bar and cooled down to 35 °C. Lastly, the vapor phase from a flash separator (CO<sub>2</sub> and C<sub>1</sub>–C<sub>4</sub> hydrocarbons) is burned to produce additional electricity, the liquid phase contains the fuel blend product to be sent to the refinery. This blend contains 29%<sub>wt</sub> gasoline, 48%<sub>wt</sub> SAF, 19%<sub>wt</sub> diesel, 2.5%<sub>wt</sub> light hydrocarbons, and traces (<1%<sub>wt</sub>) amount of wax.

**Two-step FT synthesis (2sFT).** The RWGS reactor operates at 600 °C under 30 bar (Zang *et al.* 2021). A three-stage compressor with intercooling is used to bring the CO<sub>2</sub> to 30 bar, and a single stage compressor is used for the H<sub>2</sub> feed. An amine-based capture system is used to capture and recycle the unreacted CO<sub>2</sub> (Danaci *et al.* 2020). The hydrogen inlet flowrate is adjusted to have a CO:H<sub>2</sub> ratio of 2 in the syngas. The syngas is fed to the FT reactor operating at 220 °C and 25 bar. The wax produced by the FT reactor is sent to the HC reactor where it reacts with H<sub>2</sub> fed in stoichiometric amount at 350 °C and 25 bar. The downstream separation units are identical to the 1sFT, except for the pressure of the 3-phase separators set to 25 bar. The purge rate is set to 5% to avoid excessive accumulation of CO<sub>2</sub> in the FT section. The hydrocarbon mixture from the 3-phase separators is mixed with the HC reactor outlet stream before expanding to 1 atm and cooling to 35 °C, and feeding to a flash separator. The liquid fuel blend produced contains 25%<sub>wt</sub> gasoline, 40%<sub>wt</sub> SAF, 33%<sub>wt</sub> diesel, 1.4%<sub>wt</sub> light hydrocarbons, and traces (<1%<sub>wt</sub>) amount of wax.

**Economic analysis.** The RWGS reactor is modeled as an equilibrium reactor, while the FT and HC reactors are modeled as simple conversion reactors in Aspen HYSYS. In 2sFT, yields and selectivities of the FT and HC reactors as well as the cost correlations for RWGS, FT and HC reactors are obtained from Zang *et al.* (2021). In 1sFT, the reactor yields and selectivities are obtained from Zao *et al.* (2020), with correlations from Zang *et al.* (2021) used to estimate the cost. The levelized cost of production (LCP) is calculated assuming a 15% interest rate and 20-year lifetime. LCP is compared with the fossil fuel cost, which is assumed to be 14 \$/GJ<sub>fuel</sub> in the base case, and 23 \$/GJ<sub>fuel</sub> in the optimistic best-case scenario considering a carbon tax of 150\$/tCO<sub>2</sub> (IEA 2020).

## *2.2. Life cycle analysis*

The functional unit (FU) for the LCA analysis is chosen as “1 MJ of liquid fuel blend” and the assessment is done considering a Well-to-Tank (WtT) approach. GWP is also analyzed from a Well-to-Wake (WtW) perspective, assuming that all the carbon in the fuel blend is released as CO<sub>2</sub> during the use phase. Following the recommendation by von der Assen *et al.* (2013), system expansion via substitution is adopted to deal with multi-functionality herein.

**Data collection and life-cycle inventories.** The life-cycle inventories (LCIs) of 1sFT and 2sFT are computed from the simulated flowsheets described in Section 2.1 and are summarized in Table 1. The LCIs for DAC are taken from Keith *et al.* (2018) and those for H<sub>2</sub> are taken from Al-Qahtani, *et al.* (2021). The electricity source for H<sub>2</sub> production

is assumed to be onshore wind from the UK, and the credits given for the electricity sold as a byproduct are attributed based on the electricity production mix of the UK.

Table 1: life-cycle inventories of 1sFT and 2sFT processes. FH = Fired Heater, CW = Cooling water, WW = Wastewater. The FU is 1 MJ of liquid fuel blend.

	1sFT	2sFT			1sFT	2sFT	
	Inputs	Units			Outputs	Units	
CO <sub>2</sub>	$8.25 \cdot 10^{-2}$	$9.99 \cdot 10^{-2}$	kg	FU	$1.00 \cdot 10^0$	$1.00 \cdot 10^0$	MJ
H <sub>2</sub>	$1.14 \cdot 10^{-2}$	$1.23 \cdot 10^{-2}$	kg	Electricity	$6.70 \cdot 10^{-3}$	$6.00 \cdot 10^{-3}$	kWh
FH	-	$3.04 \cdot 10^{-1}$	MJ	CO <sub>2</sub>	$1.03 \cdot 10^{-2}$	$3.27 \cdot 10^{-2}$	kg
CW	$2.29 \cdot 10^{-4}$	$3.12 \cdot 10^{-4}$	m <sup>3</sup>	CO	$6.01 \cdot 10^{-4}$	$2.32 \cdot 10^{-4}$	kg
				WW	$6.41 \cdot 10^{-5}$	$6.95 \cdot 10^{-5}$	m <sup>3</sup>

**Environmental impact assessment.** The environmental assessment is conducted in SIMAPRO 9.3, using Ecoinvent 3.8 Cut-Off database for the background process inventories. The business-as-usual (BAU) fossil-based liquid fuel alternative is assumed to be “Petroleum {GLO} | market for”, a liquid fuel with a LHV of 43.2 MJ/kg. The environmental KPIs are GWP100 alongside the monetized endpoint impacts to human health, ecosystem quality and resource scarcity, all computed with the ReCiPe2016 impact assessment methodology (Huijbregts *et al.* 2017) and using the monetization factors from Dong *et al.* (2019) converted to \$<sub>2021</sub>.

### 3. Results and discussion

#### 3.1. Midpoint environmental impacts – global warming potential

Figures 2a and 2b compare the GWP of 1sFT and 2sFT against BAU, from a WtT (2a) and WtW (2b) perspective. Both 1sFT and 2sFT have negative GWP because of the credits from DAC, with 1sFT outperforming 2sFT by c. 30%.

The molar-based carbon efficiency<sup>1</sup> is c. 68% in 2sFT and 76% in 1sFT. Therefore, 2sFT has both a higher negative contribution from DAC ( $-8.5 \cdot 10^{-2}$  kg<sub>CO<sub>2</sub>-eq</sub>) and a higher positive contribution from direct emissions ( $3.3 \cdot 10^{-2}$  kg<sub>CO<sub>2</sub>-eq</sub>) compared to 1sFT. The positive contributions to the overall GWP are dominated by the direct emissions (65%), followed by H<sub>2</sub> production (19%) and heat generation (16%). In 1sFT credits from DAC are  $-7.0 \cdot 10^{-2}$  kg<sub>CO<sub>2</sub>-eq</sub>, while the total positive emissions ( $1.9 \cdot 10^{-2}$  kg<sub>CO<sub>2</sub>-eq</sub>) are determined by direct emissions (54 %) and H<sub>2</sub> production (46%)

Since the FU is a fuel blend, it is useful to assess the WtW GWP as well. We can observe from Figure 2b that c. 95% of the total GWP for BAU ( $7.6 \cdot 10^{-2}$  kg<sub>CO<sub>2</sub>-eq</sub>) is attributed to the use phase. Both 1sFT and 2sFT are significantly better than BAU, with 1sFT outperforming 2sFT, with a 75% and 57% GWP reduction, respectively.

#### 3.2. Endpoint analysis and total cost

Figure 2c compares the monetized values of all three endpoint environmental impacts. The total externalities cost of 1sFT and 2sFT is lower than BAU, but the reduction achieved by 1sFT is significantly larger (80% vs. 30%). For all the alternatives impact on resource scarcity (RS) represents the largest share of externalities cost. Compared to

<sup>1</sup> ratio between moles of carbon in the liquid fuel blend and total moles of CO<sub>2</sub> fed to the process

BAU, 1sFT shows 40% reduction and 2sFT a 15% reduction. In both 1sFT and 2sFT the DAC energy requirements are an environmental hotspot, being responsible for 93% and 80% of the total impact on RS, respectively. In terms of damage to ecosystem quality (EQ) both processes have a negative impact, while damage to human health (HH) is almost zero for 1sFT and slightly larger than BAU for 2sFT.

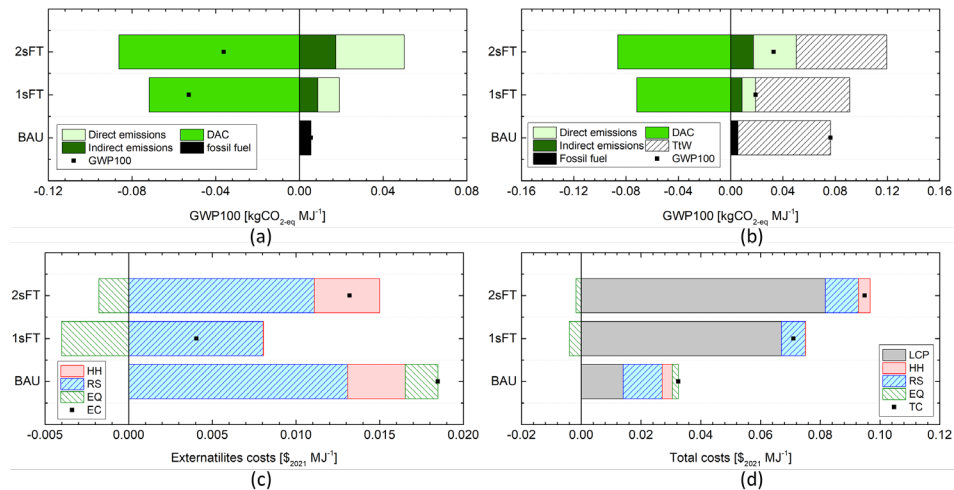


Figure 2: Results of the enviro-economic analysis for the 1sFT, 2sFT and BAU routes. (a) Well-to-Tank (WtT) global warming potential; (b) Well-to-Wake (WtW) global warming potential; (c) monetized externalities cost (EC); and (d) total cost (TC).

Turning to economic considerations, both options have a higher LCP compared to the fossil alternative, as expected. The LCP of 1sFT is 20% lower than that of 2sFT, but 5 times higher than the LCP of fossil fuel, nonetheless. In terms of total cost including externalities, 1sFT outperforms 2sFT by 25 % but has a total production cost 2 times higher than the BAU. It is worth noting that 1sFT has a 50% lower CAPEX than 2sFT, but the LCP is largely determined by the raw materials cost, where the difference between 1sFT and 2sFT per MJ<sub>fuel</sub> is only 10% (Table 1). In the base case, H<sub>2</sub> constitutes 75-78% of the raw material costs in 2sFT and 1sFT, respectively, with CO<sub>2</sub> accounting for the rest. However, in an optimistic scenario where prices drop to 1.5 \$/kg<sub>H<sub>2</sub></sub> and 50 \$/t<sub>CO<sub>2</sub></sub>, the LCP of 1sFT would decrease to 25 \$/GJ, only slightly higher than the cost of fossil fuels with a CO<sub>2</sub> tax of 150 \$/t (23 \$/GJ).

#### 4. Conclusions

This paper presented an enviro-economic assessment of two FT-based processes to produce liquid hydrocarbon fuels from CO<sub>2</sub> and H<sub>2</sub>, considering DAC and water electrolysis using wind energy as source of CO<sub>2</sub> and H<sub>2</sub> respectively. A one-step process using a novel catalyst was compared to a two-step process, assuming that the product blend is sent to an existing refinery to produce gasoline, SAF, and diesel. Our results show that the one-step process is superior both in economic and environmental terms to the two steps process, due to a lower capital cost, higher selectivity towards liquid hydrocarbons and lower energy requirements. Both processes have a negative GWP potential from a well-to-tank perspective, which compares favorably against fossil fuel production. The one-step process achieves a 75% reduction of well-to-wake GHG emissions, which is aligned with the reduction targets for SAF reported in the literature.



Our environmental assessment also considered an endpoint impacts analysis, confirming a better environmental performances of the synthetic fuels compared to fossil fuels, with an 80% reduction for the one-step process. The largest share of externalities cost in both the one-step and two-step processes is attributed to the energy requirements for DAC, and therefore alternative DAC processes should be considered to move further away from the fossil-dependency of the aviation sector. Nevertheless, the proposed low-carbon synthetic fuels are significantly more expensive than fossil fuels, and even if the prices of H<sub>2</sub> and CO<sub>2</sub> were to decrease, policy interventions such as carbon taxation, or reduced taxes on SAF would likely remain necessary for the low-carbon fuel to gain market shares. Future work will focus on adding more options for H<sub>2</sub> and CO<sub>2</sub>, including blue hydrogen, alternative DAC processes, biomass gasification and other biogenic carbon sources. On the methodological side, an uncertainty analysis will be carried out to confirm the environmental and economic results.

**Acknowledgements:** AB and BC are grateful to the Engineering and Physical Sciences Research Council (EPSRC) for funding the research under the UKRI Interdisciplinary Centre for Circular Chemical Economy programme (EP/V011863/1).

## References

- Airbus, 2020, Airbus reveals new zero-emission concept aircraft. <https://tinyurl.com/AirbusNet0> [Accessed 29 11 2022].
- A. Al-Qahtani *et al.* 2021. Uncovering the true cost of hydrogen production routes using life cycle monetisation. *Applied Energy*, 281, 115958.
- L. Chen *et al.* 2018. Entropy generation minimization for CO<sub>2</sub> hydrogenation to light olefins. *Energy* 147, 187–196.
- D. Danaci *et al.* 2020. Exploring the limits of adsorption-based CO<sub>2</sub> capture using MOFs with PVSA—from molecular design to process economics. *Molecular Systems Design & Engineering*, 5(1), 212–231.
- Y. Dong *et al.* 2019. Evaluating the monetary values of greenhouse gases emissions in life cycle impact assessment. *Journal of Cleaner Production*, 209, 538–549.
- P. Gao *et al.* 2017. Direct conversion of CO<sub>2</sub> into liquid fuels with high selectivity over a bifunctional catalyst. *Nature Chemistry*, 9, 1019.
- M.A. Huijbregts *et al.* 2017. ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level. *The International Journal of Life Cycle Assessment*, 22(2), 138–147.
- IEA 2020, CCUS in Clean Energy Transitions, IEA, Paris <https://www.iea.org/reports/ccus-in-clean-energy-transitions>, License: CC BY 4.0
- IEA 2022a, Aviation, IEA, Paris <https://www.iea.org/reports/aviation>, License: CC BY 4.0
- IEA 2022b, Direct Air Capture 2022, IEA, Paris <https://www.iea.org/reports/direct-air-capture-2022>, License: CC BY 4.0
- IRENA 2019, Hydrogen: A renewable energy perspective, International Renewable Energy Agency, Abu Dhabi
- E. Karadotcheva *et al.* 2021. Structural power performance targets for future electric aircraft. *Energies*, 14(19), 6006.
- D. W. Keith *et al.* 2018. A process for capturing CO<sub>2</sub> from the atmosphere. *Joule*, 2(8), 1573–1594.
- B. Yao *et al.* 2020. Transforming carbon dioxide into jet fuel using an organic combustion-synthesized Fe-Mn-K catalyst. *Nature Communications*, 11(1).
- G. Zang *et al.* 2021. Performance and cost analysis of liquid fuel production from H<sub>2</sub> and CO<sub>2</sub> based on the Fischer-Tropsch process. *Journal of CO<sub>2</sub> Utilization*, 46.

# Ammonia powered solid oxide fuel cells for general aviation propulsion systems: challenges and opportunities

Giampiero Di Legge,<sup>a</sup> Seyed Ali Nabavi,<sup>a</sup> Lorenzo Mazzei,<sup>b</sup> Riccardo Da Soghe,<sup>b</sup> Cosimo Bianchini,<sup>b</sup> Amirpiran Amiri,<sup>c</sup> Theoklis Nikolaidis,<sup>d</sup> Soheil Jafari,<sup>d</sup>

<sup>a</sup>*School of Water, Energy and Environment (SWEE), Cranfield University, Cranfield, MK43 0AL, United Kingdom*

<sup>b</sup>*Ergon Research srl, 50127, Firenze, Italy*

<sup>c</sup>*Aston University, Birmingham, B4 7ET, United Kingdom*

<sup>d</sup>*School of Aerospace, Transport and Manufacturing (SATM), Cranfield University, Cranfield, MK43 0AL, United Kingdom*

## Abstract

The aviation sector contributes 3% of global CO<sub>2</sub> emissions as it still relies heavily on fossil fuels. The use of low carbon hydrogen as a fuel, in both compressed or liquified forms, would drastically reduce its emissions. However, hydrogen has low volumetric energy density, and imposes a mass penalty due to storage tank requirements if used in compressed form. On the other hand, the liquefaction of hydrogen is energy-intensive and can consume ~40% of chemical energy stored in hydrogen. Ammonia, as a carbon-free hydrogen carrier, benefits from ease of storage and transport, and can be potentially used as an energy vector to enable advanced propulsion systems. Given the possibility of achieving zero-NO<sub>x</sub> in fuel cells, recent advances in fuel cell electric-powered aircraft, as well as fuel flexibility of Solid Oxide Fuel Cells (SOFCs), this work will explore the feasibility and challenges of direct ammonia-fed SOFCs as the power source for air aviation propulsion system. A model based on semi-empirical equations is developed for ammonia-powered fuel cells. A 130 kW general aviation case study is modelled and simulated in the paper to confirm the effectiveness and potential of the ammonia-powered fuel cell technology to be used as the next generation of propulsion systems in this size of aircraft. The comparative analyses demonstrated that the use of ammonia can decrease the mass of fuel consumed by 5.9% (in comparison with kerosene fueled aircraft) in a one-hour mission and allow achieving the emissions level targets set by Advisory Council for Aeronautics Research in Europe (ACARE) Flight Path 2050. Finally, the remaining practical challenges that need to be addressed to enable this important technology to be adopted by air vehicle manufacturers have been identified and discussed.

**Keywords:** Ammonia, Fuel cells, General aviation, Net zero targets, Flight Path 2050, Emission

## 1. Introduction

Anthropogenic greenhouse gas emissions from fossil fuel use are among the main contributors to climate change. It is estimated that the aviation sector, which is still highly dependent on petroleum-derived products, contributes to 3% of total carbon dioxide emissions [1], with an annual production of nearly 1000 million tons. Additional non-carbon pollutants are generated during the combustion process in gas turbines, such as nitrogen and sulfur oxides (NO<sub>x</sub> and SO<sub>x</sub>), which can contribute into contrail formation and global warming [2]. Not to mention the particulate matter (soot), known for its detrimental effects on human health, causing breathing issues, including asthma, bronchitis, coronary heart disease, and even cancer. Therefore, numerous efforts are being made to make the aviation sector more sustainable through its electrification. There are currently three main ways to obtain electrical energy on board. The first, represented by batteries, is currently limited by energy density. Bills *et al.* [3] estimated the distributions of specific energy required by different classes of aircraft. The mean requirements were determined as 600 (regional), 820 (narrow-body) and 1280 Wh/kg-pack (wide-body). However, despite the steady improvement over the last two decades, the specific energy of current generation Li-ion batteries is about 250 Wh/kg-cell [3]. Future developments are considering novel active materials and lithium metal anode, pushing the specific energy up to 400-500 Wh/kg-cell [4]. Post-lithium-ion technologies such as lithium/sulfur (Li/S) or lithium/oxygen (Li/O<sub>2</sub>) batteries might provide a further improvement up to 550 and 750 Wh/kg-cell [4]. Despite that, it is clear that batteries will not represent a viable solution for regional aircraft even in the short term and will require several decades to be deployed in long-range aircraft. Combustion engines (and in particular gas turbines), which are the second option, have low efficiencies in power production (around 45% [5]) and are currently relying on liquid fossil fuels. Despite the plans to develop hydrogen combustors, storage and safety are still concerning aspects. Sustainable aviation fuel (SAF) is the most promising solution in the short term. Despite recent tests by Airbus CITE [6] and Boeing CITE [7] proved its feasibility, limited availability and cost (two to eight times as much as kerosene) are currently hindering its massive deployment. Onboard fuel cells represent the last choice, which benefits from higher energy efficiency and enables fuel flexibility. The use of fuel cells such as solid oxide fuel cells (SOFC) in aviation is receiving much attention but is still in the early stages of development. Given the issues associated to hydrogen storage, aviation industries are considering the exploitation of ammonia, which offers no carbon content, slow reactivity and the possibility to liquify it below -33 °C at atmospheric pressure. This paper explores the feasibility of using onboard direct ammonia-fed SOFCs as one of the potential solutions for the next generation of disruptive propulsion technologies.

Currently, most studies of SOFCs to power air vehicles are based on the integration with a gas turbine. These simulations differ in architecture (new components can be implemented into the process), type of fuel used, and the type of system powered (all-electric aircraft or electric APU). Many of these studies concern large aircraft and consequently long-range missions that require fuel cells with the power higher than 400 kW. One of the main obstacles to the use of SOFC onboard is represented by the weight of the stack that causes an excessive fuel consumption. Many studies, in fact, in addition to analysing the performance, have focused on the power density of the system [8-12]. Tornabene *et al.* [8] delivered a preliminary estimation of the impact on the system's mass and volume of all the components required for an APU powered by a SOFC-GT. Whyatt and Chick [9] provided a similar analysis for a Boeing 787 and concluded that the system weighed 1500 kilograms beyond the break-even point. Steffen, Freeh and Larosiliere [10]

developed a model in which a SOFC-GT powered the APU of a 300-passenger aircraft. Their results, which showed an improved efficiency, but an insufficient energy density compared to combustion engines, are in line with other studies. Dollmayer et al.[11] analysed the performance of a two-engines all-electric aircraft powered by hydrogen for a 4000 nautical miles mission. They examined the mass penalties and fuel consumption, discussing strategies to lower fuel consumption such as cabin air for processing, a partial thrust recovery, or water re-utilisation. The influence of the exhaust temperature was also examined because it can enable considerable energy recovery. Their results also established that 4kg/kWh would be the starting mass-to-power ratio for achieving fuel consumption that would justify the presence of SOFC. Lim et al. [12] designed, built, and compared the performance of a SOFC as a stand-alone unit and a SOFC-GT. The study, which also included a comparison between hydrogen and methane, showed that high pressures have a positive effect on efficiency and power density. In a National Aeronautics and Space Administration (NASA) technical publication, Srinivasan et al. [13] replaced the APU of a Boeing 777-200ER with a SOFC-GT (450 kW) and explored different levels of integration with the electrical power supply (EPS), thermal management (TMS), and environmental control (ECS) systems. SOFC-GT integration reduced nitrogen oxides (NO<sub>x</sub>), carbon monoxide, and unburned hydrocarbons (UHC) emission by 15.45% (NO<sub>x</sub>), 3.25% (CO), and 4.72% (UHC) respectively for ground operations. In addition, the authors showed that higher integration increased fuel economy for long-range missions by up to 0.77%. However, the highest reduction in fuel consumption (3%) was obtained in another analysis that targeted short-range aircraft, supporting the hypothesis that SOFCs are currently best suited for small aircraft or Unmanned Air Vehicles (UAVs). In this regard, Ji et al. [14] modelled a UAV powered by a propane-fuelled SOFC hybrid engine, achieving a thermal efficiency of 57.6%. As discussed above, most of previous studies used the SOFC as an APU. This paper is now exploring the feasibility of applying this technology to primary propulsion. In other words, to bridge the above-mentioned research results and the stringent targets and limitations set in the ACARE's flight path 2050 for emissions level reduction, this study will explore the feasibility of using ammonia-powered SOFC as the main power source for General Aviation (GA) applications. For this purpose, an experimentally validated ammonia-fed SOFC model will be developed in section two. The developed model will be used as the main source of the power for propulsion system in section three where a small aircraft (Cessna 172 size) will be simulated as the case study. The simulation results are presented and discussed in this section followed by the identified remaining research challenges in the way of using onboard ammonia powered SOFCs as the propulsion systems for the future air vehicles. Conclusion remarks are summarized in section four. It should also be mentioned that although the GA contributions to the overall climate impact of aviation is not huge, it does not question the methodological approach presented in this paper.

## **2. Ammonia Powered SOFC Modelling**

This section describes the developed tool for performance simulation of the ammonia powered SOFC which could be used to generate the required power for future air vehicles. First, the mathematical modelling of the SOFC will be detailed through a semi-empirical physics-based approach. The developed process model in ASPEN Plus<sup>®</sup> software will then be explained, and the hierarchies of the calculation method are presented in a

flowchart format. Finally, the developed model will be validated against experimental data to confirm the validity and effectiveness of the proposed modelling approach.

### 2.1. Mathematical Model

The common technique for modelling SOFCs' performance in steady conditions starts from the Open Circuit Voltage (OCV) obtained from the Nernst equation as follow:



$$E_N = \frac{RT}{2F} * \ln\left(\frac{p_{H_2} * p_{O_2}^{0.5}}{p_{H_2O}}\right) \quad (2)$$

Where R is the gas constant, T is the temperature, F is Faraday's constant. The polarisation losses (so-called overpotentials) occurring within the FC (i.e., activation, ohmic and concentration losses) are then subtracted from the OCV to determine the actual voltage. The activation overpotential ( $\eta_{act}$ ) is due to the irreversibility of the charge transfer that take place at the electrodes. The ohmic overpotential ( $\eta_{ohm}$ ) is caused mainly by the electrolytic resistance to the ions flow. Finally, the concentration overpotential ( $\eta_{conc}$ ) depends on the mass transfer limitation to the kinetics of the electrochemical process.

$$V_{final} = E_N - (\eta_{act} + \eta_{ohm} + \eta_{conc}) \quad (3)$$

In addition to being influenced by the working conditions of the fuel cell, these potential variations are closely dependent on the geometric and chemical-physical characteristics of the fuel cell. Therefore, this modelling approach requires parameters that regard the activation energy, interaction factors, anodic and cathodic composition, and cell structure. The model adopted in this work is based on semi-empirical equations derived from interpolations of experimental data. Besides providing greater flexibility, this method excludes the need to determine the microscopic characteristics of the FC and the mentioned parameters. This study is based on a generator produced by Siemens-Westinghouse and composed of multiple tubular SOFCs with internal reforming connected in bundles. The single cell has an active area of 834 cm<sup>2</sup> and can produce about 130 kW at a temperature of 900-950 °C. YSZ was used as material for the electrolyte, while the anode is made of Ni-YSZ cermet [15].

For validation case, when natural gas is used, fuel is reformed with the steam coming from the anode exhaust gas recirculation. The products, mainly CO and H<sub>2</sub>, are sent to the anode and oxidised by oxygen coming from the cathode. The fraction of exhaust gas that is not recycled is sent to a combustion plenum (called afterburner), where the total oxidation of the remaining syngas or unburnt hydrocarbons takes place. Before being released into the atmosphere, the afterburner exhaust gases are used to pre-heat the incoming air. For ammonia, this compound, being carbon-less, is not subjected to SR but undergoes thermal cracking generating hydrogen and nitrogen ( $2NH_3 \rightarrow N_2 + 3H_2$ ). Since the process is carried out with Ni-based catalysts, it can occur directly at the FC inlet. For this reason, the temperature used to model the reforming is 950 °C, even though

previous studies have established that the decomposition of ammonia is already complete at 800 °C. The thermal cracking does not require steam; hence there is no need for recirculation. It should be noted that since the ammonia is thermochemically decomposed to hydrogen and nitrogen, and hydrogen undergoes electrochemical reaction, Eq. (2) is written based on hydrogen.

## 2.2. ASPEN Model

The described process has been reproduced in Aspen Plus using the model illustrated in figure 1. This section gives a detailed description of the technical specifications used to design the model. Aspen terminology is represented in italic>

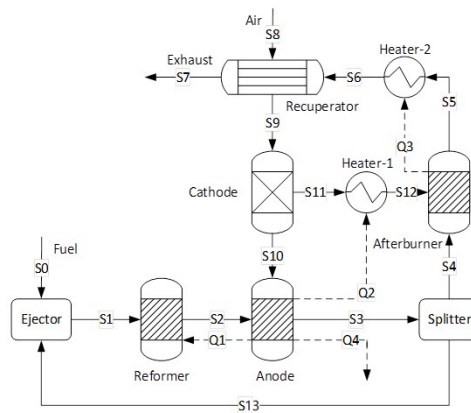


FIGURE 1 – ASPEN PLUS MODEL

The reforming and reactions occurring inside the anode were simulated using the *Rgibbs* block. This equilibrium reactor determines the final concentrations based on the chemical potential of the species involved. In other words, the Gibbs reactor predicts the product compositions by minimising the total Gibbs free energy ( $G_{total}$ ) (Equation 4)

$$G_{total} = \sum_{i=1}^N n_i \Delta G_{fi}^0 = \sum_{i=1}^N n_i RT \ln \frac{f_i}{f_i^0} \quad (4)$$

Where  $f_i^0$  and  $f_i$  represent the fugacity of the species  $i$  in the reactor and standard conditions, respectively. Two different algorithms have been developed for the model: the first (Figure 2a) allows the power produced by the fuel cell to be calculated from the amount of fuel consumed. The second (Figure 2b) determines the potential, current and fuel required starting from the expected power. These two algorithms are based on the same equations, which are rearranged to follow the structures described in Figure 1 (In a)  $V$ ,  $I$  and  $P$  are determined starting from the amount of fuel consumed. In b)  $V$ ,  $I$  and required fuel are obtained from a pre-set power [16]).

It should be mentioned that in case b), after an initial estimation of the voltage from the process parameters, the current is easily calculated using the relationship:

*Current = Power / Voltage*

(5) The next step is to calculate the amount of fuel required (mol/h) to generate that current. In the case of using hydrogen as the fuel, the required fuel could be calculated as follow:

$$H_{2,required} = \frac{Current}{2 * U_f * F} * (3600) = \frac{0.018655 * I}{U_f} \quad (6)$$

Where  $U_f$  is the utilisation factor determining the percentage of hydrogen consumed at the anode, and  $F$  is the Faraday constant. By calculating the amount of hydrogen required, the consumption of any suitable fuel can be determined just by knowing its composition. Finally, the efficiency can be obtained through the low calorific value ( $LHV_{fuel}$ ) of the used fuel and its molar flow ( $n_{fuel}$ ) (Equation 7)

$$Efficiency = \frac{Power}{(n_{fuel} * LHV_{fuel})} \quad (7)$$

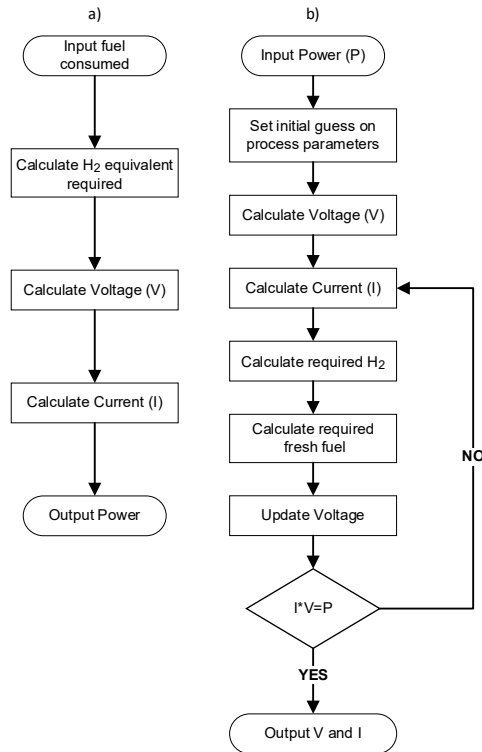


FIGURE 2 – HIERARCHIES OF CALCULATION USED IN THE MODEL

### 2.3. Validation of the Model

The developed model is validated with the experimental results of Zhu and Kee [17] SOFC fed with natural gas at a temperature of 910°C and pressure of 1 bar). The utilisation factor and steam to carbon ratio were set to 0.85 and 2.5, respectively. The natural gas used as fuel consists of the following components: methane 81.3%, ethane

2.9%, propane 0.4%, butane 0.2%, carbon monoxide 0.9% and nitrogen 14.3%. The proximity between experimental data and simulation results confirms the validity of the model (Table 1).

TABLE 1- VALIDATION OF THE DEVELOPED SOFC MODEL

	Literature data [17]	Simulation results
Power (kW)	120	120
Current density (mA/cm <sup>2</sup> )	180	175
Reforming temperature (°C)	550	537
<b>Anode outlet composition (mol%)</b>		
Water	48%	51.1%
Carbon dioxide	28%	25.1%
Hydrogen	14%	11.4%
Carbon monoxide	5%	7.3%
Nitrogen	5%	5.1%
<b>Exhaust stream composition (mol%)</b>		
Nitrogen	77%	77.3%
Oxygen	16%	16.1%
Water	5%	4.3%
Carbon dioxide	2%	2.2%

### 3. Results and Discussion

After having the validated SOFC model up-and-running, the next step is to assess the performance of the ammonia-powered SOFC to be used as the power source for general aviation. As a case study, a light aircraft with the take off power of 130 kW (Cessna 172 size aircraft) is used for simulation purpose. A short-range mission lasting a total of one hour is utilized in the simulations as shown in figure 3. The mission was simulated in all its phases, from take-off to landing, considering the different power requirements and duration of the phases. To calculate the required power for the propulsion system at each flight phase, the mission is simulated in Turbomatch (TM) - a Cranfield in-house developed software – for the Cessna 172 with Kerosene as the fuel. As a state-of-the-art gas turbine performance simulation and diagnostics code, TM is a key proprietary tool of the Propulsion Engineering Centre (PEC) at Cranfield University. This has been widely used in aircraft engine research projects and within the Cranfield postgraduate community [18]. The required power calculated by TM at Take-Off, Climb, Cruise, Descend, approach, and taxi phases are 130, 104, 83.2, 65, 52, and 19.5 kW respectively.



The mission with associated required power is then implemented in the developed SOFC model with ammonia as the fuel. Table 2 shows the results of the fuel consumption for different flight phases. The total fuel burn during the mission is 25.87 kg for ammonia powered SOFC. It is 5.9% less than the fuel consumption for the Kerosene powered aircraft with similar size (27.51 kg). The other advantages of using ammonia as the fuel is it does not generate any carbon dioxide. So, it is a high potential candidate for next generation of sustainable propulsion systems. However, the remained research challenges should be explored and addressed in order to enable this technology to be adopted by air vehicle manufacturers.

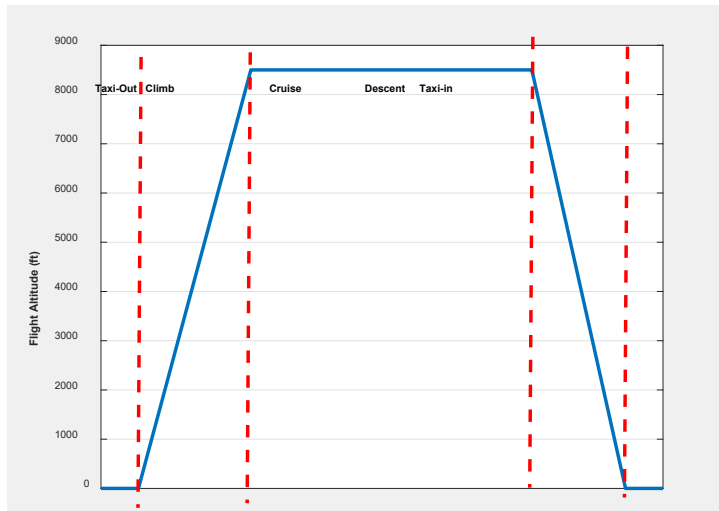


FIGURE 3 – DEFINED FLIGHT MISSION

TABLE 2 – SIMULATION RESULTS FOR 1-HOUR MISSION. FUEL: AMMONIA, T=950°C.

	Power (kW)	required	Duration (s)	Ammonia		
				kg/h	L/min	Fuel consumed (kg)
Take off	130		30	46.37	1.06	0.39
Climb	104		600	36.49	0.83	6.08
Cruise	83.2		1800	28.96	0.66	14.48
Descend	65		480	22.46	0.51	2.99
Approach	52		210	17.88	0.41	1.04
Taxi	19.5		480	6.62	0.15	0.88
<b>Total</b>						<b>25.87</b>

In summary, the feasibility and potential of using ammonia-fed SOFC as the propulsion system of future light air vehicles is studied and the simulation results reported in table 2. This technology is pretty aligned with Net Zero 2050 and Flight Path 2050 targets as it reduces the fuel consumption and does not produce carbon dioxide. However, it should be mentioned that this is just the first shot in this topic, and the study just took the power required of a kerosene-powered Cessna 172 and limited the comparison to the fuel used by a SOFC providing said power profile. In other words, different aspects of mass and aerodynamics considerations should be studied in future work to fully explore the pros and cons of using SOFC as the main source of power in propulsion systems. It may result in either increasing power required for a given performance or limiting performance when maintaining power levels. Other details like drag penalties and tank-system design and integration should also be discussed accordingly. Moreover, the following challenges are still remained:

- The first challenge is the thermal management system (TMS) architecture design and development. As the temperature of the ammonia powered SOFC is 950°C, an advanced and effective TMS should be designed and implemented to manage the thermal loads generated in the propulsion system during the mission.
- NO<sub>x</sub> can be generated, besides from ammonia oxidation, also by thermal processes that take place at temperatures above 1300°C and convert diatomic nitrogen. It is worth noting that, in some simulations, the exhaust gases leaving the afterburner contain nitrogen and oxygen at a temperature close to 1300°C. However, introducing a heat exchanger to slightly cool the gases entering the combustion chamber can solve the problem.
- Most of the global production of ammonia is based on steam methane reforming (SMR) to produce (grey) hydrogen, followed by Haber–Bosch synthesis. Although the state-of-the-art SMR technology offers the most economical approach, the process is associated with significant CO<sub>2</sub> emissions. So, potential solutions for green ammonia production are welcomed.
- The conventional SOFCs are developed mainly for stationary applications, in which the system weight may not be a key indicator. However, for air transportation, their energy-to-weight ratio is far below that of turboprops. Although the higher efficiency of SOFC may potentially compensate for this to some extent, for further deployment of SOFC-powered air vehicles, the energy-to-weight ratio should match or exceed that of turbines.
- Synchronization of the SOFC with other elements of the propulsion system should also be considered to address the implementation challenges like slow response time of SOFCs.

#### **4. Conclusion**

This research represents a first step in the feasibility study and conceptual design for introducing SOFCs with reforming onboard air vehicles. This technology makes it possible to overcome the obstacles posed by difficult hydrogen storage conditions. A physics-based model based on semi-empirical equations is developed for ammonia-powered SOFC and validated against experimental data. The performance of the developed model is assessed as the propulsion system for a light aircraft in Cessna 172

size with the Take-Off power of 130kW. Simulation results for a 1-hour mission confirm that using SOFC with ammonia fuel will reduce the mass of fuel consumed by the aircraft by 5.9% as well as omitting the carbon dioxide. So, the high potential of using ammonia powered SOFC as a disruptive propulsion technology for future aircraft is obvious. However, there are still remaining research challenges including thermal management system design, dealing with NO<sub>x</sub> production, and green ammonia production that need to be addressed in order to enable this important technology to be adopted by air vehicle manufactures in the future.

## References

- [1]. Grewe V., Gangoli Rao A., Grönstedt T., Xisto C., Linke F., Melkert J., et al. Evaluating the climate impact of aviation emission scenarios towards the Paris agreement including COVID-19 effects. *Nature Communications* 2021 12:1. Nature Publishing Group; 22 June 2021; 12(1): 1–10. Available at: DOI:10.1038/s41467-021-24091-y
- [2]. Lee D.S., Fahey D.W., Skowron A., Allen M.R., Burkhardt U., Chen Q., Doherty S.J., Freeman S., Forster P.M., Fuglestedt J., Gettelman A., De León R.R., Lim L.L., Lund M.T., Millar R.J., Owen B., Penner J.E., Pitari G., Prather M.J., Sausen R., Wilcox L.J., The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018, *Atmospheric Environment*, Volume 244, 2021, 117834, ISSN 1352-2310, <https://doi.org/10.1016/j.atmosenv.2020.117834>.
- [3]. Bills A., Sripad S., Fredericks W. L., Singh M., and Viswanathan V., “Performance Metrics Required of Next-Generation Batteries to Electrify Commercial Aircraft,” *ACS Energy Lett.*, vol. 5, no. 2, pp. 663–668, 2020, doi: 10.1021/acseenergylett.9b02574
- [4]. Placke, T., Kloepsch, R., Dühnen, S. et al. Lithium ion, lithium metal, and alternative rechargeable battery technologies: the odyssey for high energy density. *J Solid State Electrochem* 21, 1939–1964 (2017). <https://doi.org/10.1007/s10008-017-3610-7>
- [5]. Langston, L. S. "Aspects of Gas Turbine Thermal Efficiency." *ASME. Mechanical Engineering*. September 2020; 142(09): 54–55. <https://doi.org/10.1115/1.2020-SEP4>
- [6]. This A319neo is the latest to test 100% SAF. An in-flight study is focusing on emissions performance of single-aisle Aircraft. 30 October 2021. <https://www.airbus.com/en/newsroom/stories/2021-10-this-a319neo-is-the-latest-to-test-100-saf>
- [7]. 2 Aerospace Pioneers Partner For Sustainability. 2021. <https://www.boeing.com/features/2021/11/aerospace-pioneers-partner-for-sustainability.page>
- [8]. Tornabene R., Wang X-Y., Steffen CJ., Freeh JE. Development of Parametric Mass and Volume Models for an Aerospace SOFC/Gas Turbine Hybrid System. *Proceedings of the ASME Turbo Expo. American Society of Mechanical Engineers Digital Collection*; 11 November 2008; 5: 135–144. Available at: DOI:10.1115/GT2005-68334
- [9]. Whyatt GA., Chick LA. Electrical Generation for More-Electric Aircraft Using Solid Oxide Fuel Cells. Richland, WA (United States); 1 April 2012; Available at: DOI:10.2172/1056768
- [10]. Steffen CJ., Freeh JE., Larosiliere LM. Solid Oxide Fuel Cell/Gas Turbine Hybrid Cycle Technology for Auxiliary Aerospace Power. *Proceedings of the ASME Turbo Expo. American Society of Mechanical Engineers Digital Collection*; 11 November 2008; 5: 253–260. Available at: DOI:10.1115/GT2005-68619
- [11]. Dollmayer J., Bundschuh N., Carl UB. Fuel mass penalty due to generators and fuel cells as energy source of the all-electric aircraft. *Aerospace Science and Technology. Elsevier Masson*; 1 December 2006; 10(8): 686–694. Available at: DOI:10.1016/J.AST.2006.08.001
- [12]. Lim TH., Song RH., Shin DR., Yang J il., Jung H., Vinke IC., et al. Operating characteristics of a 5 kW class anode-supported planar SOFC stack for a fuel cell/gas turbine hybrid system. *International Journal of Hydrogen Energy. Pergamon*; 1 February 2008; 33(3): 1076–1083. Available at: DOI:10.1016/J.IJHYDENE.2007.11.017
- [13]. Srinivasan H., Yamanis J., Welch R., NASA ST-., CR-2006 undefined., 2006 undefined. Solid oxide fuel cell APU feasibility study for a long range commercial aircraft using UTC

- ITAPS approach. ntrs.nasa.gov. 2006; Available at: <https://ntrs.nasa.gov/api/citations/20070004892/downloads/20070004892.pdf?attachment=true>
- [14]. Ji Z., Qin J., Cheng K., Liu H., Zhang S., Dong P. Thermodynamic analysis of a solid oxide fuel cell jet hybrid engine for long-endurance unmanned air vehicles. *Energy Conversion and Management*. Pergamon; 1 March 2019; 183: 50–64. Available at: DOI:10.1016/J.ENCONMAN.2018.12.076
- [15]. Fuerte A., Valenzuela RX., Escudero MJ., Daza L. Ammonia as efficient fuel for SOFC. *Journal of Power Sources*. Elsevier; 1 July 2009; 192(1): 170–174. Available at: DOI:10.1016/J.JPOWSOUR.2008.11.037
- [16]. W. Zhang, E. Croiset, P.L. Douglas, M.W. Fowler, E. Entchev, Simulation of a tubular solid oxide fuel cell stack using AspenPlus™ unit operation models, *Energy Conversion and Management*, Volume 46, Issue 2, 2005, Pages 181-196, ISSN 0196-8904, <https://doi.org/10.1016/j.enconman.2004.03.002>.
- [17]. Zhu H., Kee RJ. Thermodynamics of SOFC efficiency and fuel utilization as functions of fuel mixtures and operating conditions. *Journal of Power Sources*. Elsevier; 27 October 2006; 161(2): 957–964. Available at: DOI:10.1016/J.JPOWSOUR.2006.05.006
- [18]. Apostolidis, A, Sampath, S, Laskaridis, P, & Singh, R. "WebEngine: A Web-Based Gas Turbine Performance Simulation Tool." *Proceedings of the ASME Turbo Expo 2013: Turbine Technical Conference and Exposition*. Volume 4: Ceramics; Concentrating Solar Power Plants; Controls, Diagnostics and Instrumentation; Education; Electric Power; Fans and Blowers. San Antonio, Texas, USA. June 3–7, 2013. V004T08A007. ASME. <https://doi.org/10.1115/GT2013-95296>



33<sup>rd</sup> EUROPEAN SYMPOSIUM ON  
COMPUTER AIDED PROCESS  
ENGINEERING

VOLUME 3



**COMPUTER-AIDED CHEMICAL ENGINEERING, 52**

**33<sup>rd</sup> EUROPEAN SYMPOSIUM ON  
COMPUTER AIDED PROCESS  
ENGINEERING**

**VOLUME 3**

*Edited by*

**Prof. Antonios C. Kokossis**  
*National Technical University of Athens  
School of Chemical Engineering  
Athens, Greece*

**Prof. Michael C. Georgiadis**  
*Aristotle University of Thessaloniki  
School of Engineering  
Department of Chemical Engineering  
Greece*

**Prof. Efstratios Pistikopoulos**  
*Chemical Engineering  
Texas A&M Energy Institute  
TX, USA*



**ELSEVIER**

**Amsterdam – Boston – Heidelberg – London – New York – Oxford  
Paris – San Diego – San Francisco – Singapore – Sydney – Tokyo**



Elsevier  
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands  
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK  
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, USA

Copyright © 2023 Elsevier B.V. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: [www.elsevier.com/permissions](http://www.elsevier.com/permissions).

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

### Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

### British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

### Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

ISBN (Volume 3): 978-0-443-23555-9

ISBN (Set) : 978-0-443-15274-0

ISSN: 1570-7946

For information on all Elsevier publications visit our website at <https://www.elsevier.com/>

		<b>Working together to grow libraries in developing countries</b>
<a href="http://www.elsevier.com">www.elsevier.com</a> • <a href="http://www.bookaid.org">www.bookaid.org</a>		

*Publisher:* Candice Janco  
*Acquisition Editor:* Anita Koch  
*Editorial Project Manager:* Lena Sparks  
*Production Project Manager:* Paul Prasad Chandramohan  
*Designer:* Greg Harris

Typeset by STRAIVE

# Contents

- 375. Simulation-based design of regional biomass thermochemical conversion system for improved environmental and socio-economic performance**  
*Leonardo L. Corradini, Aya Heiho, Yuichiro Kanematsu, Ryoko Shimono, Satoshi Ohara, Yasunori Kikuchi* 2361
- 376. Conceptual Design of a Biorefinery to Use Brown Seaweed Sargassum**  
*Omar Flores-Mendoza, Teresa Lopez-Arenas* 2367
- 377. Design and Production of Biodiesel and Porous Filaments from Endemic Species of the Cienega de Chapala Region**  
*María Fernanda Torres-Magaña, José Luis Baltazar-Álvarez, Pascual Eduardo Murillo-Alvarado and Gabriela Guadalupe Esquivel-Barajas* 2373
- 378. Real-time control and scheduling decision making in fermentation-based food process operations**  
*Isuru A. Udugama, Keegan K. Hall, Christoph Bayer, Brent R. Young, Timothy G. Walmsley* 2379
- 379. Production of  $\text{Ca}(\text{HCO}_2)_2$  from formate solution through  $\text{CO}_2$  hydrogenation**  
*Taeksang Yoon, Hayoung Yoon, Sungho Yoon, Chul-Jin Lee* 2385
- 380. Techno-Economic Process Analysis of the Chemical Recycling of Nylon 6 Using Phosphoric Acid**  
*Ann-Joelle Minor, Ruben Goldhahn, Liisa Rihko-Struckmann<sup>a</sup>, Kai Sundmacher* 2391
- 381. Multiscale High-throughput Screening for Membrane-based Green Hydrogen Separation Process**  
*Nahyeon An, Seongbin Ga, Hyungtae Cho, Boram Gu, Junghwan Kim* 2397
- 382. Discrete Element Simulation and Economics of Mechanochemical Grinding of Plastic Waste at an Industrial Scale**  
*Elisavet Anglou, Yuchen Chang, Arvind Ganesan, Sankar Nair, Carsten Sievers, Fani Boukouvala* 2403
- 383. Multi-scale Modelling and Experimental Investigation of Hydrogen Sulphide Thermal Decomposition**  
*Anna Nova, Francesco Negri, Flavio Manenti* 2409

- 384. Reuse and recycling of waste material: design of a pyrolysis-based process**  
*Gabriel Rodríguez Garrido, Pablo Marinangeli, Jorge A. Ressia, Maria A. Volpe, Patricia M. Hoch* **2415**
- 385. Life Cycle Assessment of bio-oil produced through lignocellulosic biomass liquefaction**  
*Ana S. Augusto, Catarina G. Braz, Duarte M. Cecílio, Maria M. Mateus, Henrique A. Matos* **2421**
- 386. Impacts of defossilising downstream derivatives in petrochemical clusters – MTBE case study**  
*Inna Stepchuk, Mar Pérez-Fortes, Andrea Ramírez Ramírez* **2427**
- 387. Methane-to-X: an economic assessment of methane valorisation options to improve carbon circularity**  
*Ben Lyons, Andrea Bernardi, Nilay Shah, Benoît Chachuat* **2433**
- 388. Purification of Raw Bio-Methanol from Renewable Feedstocks: A Flexibility Assessment**  
*Matteo Fedeli, Alessandro di Pretoro, Ludovic Montastruc, Flavio Manenti* **2439**
- 389. Life cycle assessment of Pine needle-based electricity generation in India**  
*Pratham Khaitan, Ankush Halba, Pratham Arora* **2445**
- 390. Evaluation of alternative carbon based ethylene production in a petrochemical cluster: Technology screening & value chain impact assessment**  
*James Tonny Manalal, Mar Pérez-Fortes, Paola Ibarra Gonzalez and Andrea Ramirez Ramirez* **2451**
- 391. Syngas fermentation to ethanol: the effects of gas recycling on economics**  
*Haneef Shijaz, Fausto Gallucci, Adrie Straathof,<sup>a</sup> John Posada* **2457**
- 392. Modeling, simulation and techno-economic analysis of an integrated biorefinery based on halophytes**  
*Tutku Taşçı Çilak, Sanketkumar Raval, Sylvia Fasse, Mette H. Thomsen, Axel Gottschalk* **2463**
- 393. Optimal Control for Deriving Policies for Global Sustainability**  
*Urmila Diwekar, Apoorva Nisal, Yogendra Shastri, and Heriberto Cabezas* **2469**

- 394. Application of CAPE Tools into Prospective Life Cycle Assessment: A Case Study in Feedstock Recycling of Waste Plastics**  
*Yasunori Kikuchi, Yuki Nomura, Takuma Nakamura, Shoma Fujii, Aya Heiho, Yuichiro Kanematsu* **2475**
- 395. Designing roadmaps for transitioning to value chains with net-zero emissions: Case of the chemical industry**  
*Amrita Sen, Vyom Thakker, George Stephanopoulos, Bhavik R Bakshi* **2481**
- 396. Data Envelopment Analysis of Ammonia and Methanol as Sustainable Energy Storage Vectors**  
*Sebastiano C. D'Angelo, Michael Bregy, Philipp Steiner, Raul Calvo-Serrano, Gonzalo Guillén-Gosálbez* **2487**
- 397. Simulation of thermochemical recycling of waste plastic to produce dimethyl ether**  
*Mohammaed Awwad and Umer Zahid* **2493**
- 398. Sustainable recycling of refrigerants: Analysis of alternatives**  
*J. Moreno, I. Karpov, A. Ahmed, J. Foglia, H. Wu, A. Gadekar, P.L.J. Swinkels, R. Gani* **2499**
- 399. Thermodynamic study and simulation of the process of separation of the IPA+water mixture by heterogeneous azeotropic distillation**  
*Elvira Spatolisano, Camilla Barbieri, Laura A. Pellegrini, Stefania Moioli* **2507**
- 400. Sustainability assessment of poly(butylene succinate) production and End-of-Life options from wheat straw**  
*Sofia-Maria Ioannidou, Dimitrios Ladakis, Ricardo Rebolledo-Leiva, Maria Teresa Moreira, Ioannis K. Kookos, Apostolis Koutinas* **2513**
- 401. Computer simulation of a plastic waste gasification-solid oxide fuel cell power generation system using Aspen Plus**  
*Mahmudul Hassan Riyad, Sadman Fakid, Nahid Sanzida* **2519**
- 402. Enhancement of hydrochar from coal discards and sewage sludge (co)-HTC for transition to a low-carbon economy**  
*Gentil Kahilu Mwendu, Orevaoghene Eterigho-Ikelegbe, Jean Mulopo, Samson Bada* **2525**

<b>403. Fully Electrified Conversion of Low-Quality Plastic Waste to Polymer Precursors</b> <i>Kristiano Prifti, Andrea Galeazzi, Adrián Pacheco-López, Antonio Espuña, Flavio Manenti</i>	<b>2533</b>
<b>404. Deep neural network-based multi-objective optimization of NO<sub>x</sub> emission and profit by recovering lignocellulosic biomass</b> <i>Kim Y., Park J., Lim J., Joo C., Cho H., Kim J</i>	<b>2539</b>
<b>405. Decarbonization of gray ammonia production through the Allam power cycle: a techno-economic assessment and optimization</b> <i>Andrea Isella, Mattia Forina, Mario Spagna, Davide Manca</i>	<b>2547</b>
<b>406. Into the Valley of Death Rode the Green Transition</b> <i>Robert Pujan and Heinz A. Preisig</i>	<b>2555</b>
<b>407. A novel industrial biotechnology approach to valorize fatty acids to bioplastics: scope for scale-up and process efficiency using an integrated approach</b> <i>Pantelis Vasilakis, Kostas Pyrgakis, Melina Psycha, Antonino Biundo, Antonis Kokossis</i>	<b>2561</b>
<b>408. On the re-deployment of lignocellulosic biorefineries with solvent-free designs and process-to-process integration</b> <i>Iraklis Sarris, Kostas Pyrgakis, Antonis Kokossis</i>	<b>2567</b>
<b>T5: Systems methods in industrial biotechnology and biomedical applications</b>	
<b>409. Digital-twin development for a novel vibrating membrane aiming at fractionating fermentation broths</b> <i>Oscar A. Prado-Rubio, Wai Fung Hui, Mads Stevnsborg, Manuel Pinelo, Jakob Kjøbsted Huusom</i>	<b>2573</b>
<b>410. Post-Combustion Carbon Capture and Utilization by Sodium Hydroxide Aqueous Solution for Bicarbonate Microalgae Cultivation</b> <i>Jialin Liu, Yoong Kit Leong, Jo-Shu Chang</i>	<b>2579</b>
<b>411. Comparing different modelling approaches for metabolic network dynamic simulation under uncertainty</b> <i>Oliver Pennington, Dongda Zhang</i>	<b>2587</b>

- 412. Integrating hybrid modelling and transfer learning for new bioprocess predictive modelling**  
Sam Kay, Harry Kay, Alexander W. Rogers, Dongda Zhang **2593**
- 413. Machine learning-supported cybergenetic modeling, optimization and control for synthetic microbial communities**  
*Sebastián Espinel-Ríos, Katja Bettenbrock, Steffen Klamt, José L. Avalos, Rolf Findeisen* **2599**
- 414. Model-based optimization of a recombinant *E. coli* fed-batch process for best time-space yield and conversion efficiency**  
*Julian Kager, Nora Horst, Christoph Herwig and Ioannis V.Skiadas* **2605**
- 415. Closed-loop optimization of high-throughput robotic platforms for reproducible bioprocess development**  
*Federico M. Mione, Judit Aizpuru, Martin F. Luna, Pablo Rodriguez Bahamon, Jong Woo Kim, Ernesto C. Martinez, M. Nicolas Cruz B* **2611**
- 416. Dynamic microbial-community metabolic modeling for yogurt fermentation based on the metagenome of starter culture**  
*Sizhe Qiu, Zhijie Yang, Hong Zeng, Bei Wang, Aidong Yang* **2617**
- 417. Multilevel Discrete Modelling of Microbial Cell Cultures using Population Balances**  
*Menwer Attarakih, Hans-Jörg Bart, Khaled Rawajfeh* **2623**
- 418. Binary Classification of the Endocrine Disrupting Chemicals by Artificial Neural Networks**  
*Zahir Aghayev, George F. Walker, Funda Iseri, Moustafa Ali, Adam T. Szafran, Fabio Stossi, Michael A. Mancini, Efstratios N. Pistikopoulos, Burcu Beykal* **2629**
- 419. Digital Twins in Pilot Scale Fermentation: Non-Linear State Estimation for Improving Induction Timing**  
*Mads Stevnsborg, Kurt Selle, Ryan Barton, Oscar A. Prado-Rubio, Carina Gargalo, Krist V. Gernaey, Gary Gilleskie, Jakob K. Huusom* **2635**
- 420. Optimal dosing of thyroid hormones in hypothyroid patients with an individualized compartmental model**  
*Davide Manca, Federico Appiani, Giovanni Colombo* **2641**
- 421. Technoeconomic process modelling of waste remediation within a supply chain network**  
*Robert Milton, Jagroop Pandhal and Solomon F. Brown* **2647**

- 422. Spatio-temporal analysis of milk safety under climate change**  
*Lydia Katsini, Satyajeet S. Bhonsale, Styliani Roufou, Sholeem Griffin, Vasilis Valdramidis, Simen Akkermans, Monika Polanska, Jan F.M. Van Impe* **2653**
- 423. Ensemble Kalman Filter for estimation of intracellular nucleotide sugars from extracellular metabolites in monoclonal antibodies**  
*Luxi Yu, Ehecatl Antonio del Rio Chanona, Cleo Kontoravdi* **2659**
- 424. Towards a unified multi-scale strategy for bio-manufacturing process development**  
*Thomas Bisgaard, Nima Nazemzadah, Eduardo Krebs Kleingesinds, Negin Yousefi, Christian Beenfeldt, Seyed Soheil Mansouri* **2665**
- 425. Development of a framework for simulation of biotechnological processes**  
*Priscila Marques da Paz, Caroline Satye Martins Nakama and Galo Antonio Carrillo Le Roux* **2671**
- 426. Knowledge modelling framework for per-and poly-fluoroalkyl substances (PFAS) treatment solutions**  
*Madeleine Bussemaker, Nikos Trokanas, Ioannis Kavakiotis, Franjo Cecelja* **2677**
- 427. Combined metabolic modeling and experimental data for enhanced biotechnological production**  
*Anita L. Ziegler, Melanie Filbig, Johannes Parschau, Till Tiso, Lars M. Blank, Alexander Mitsos* **2683**
- 428. Machine learning models for predicting membranolytic anticancer peptides**  
*Fatemeh Alimirzaei, Chris A. Kieslich* **2689**
- 429. Data-driven prediction of peptide-MHC binding using oscillations of physicochemical properties**  
*Hyeju Song, Chris A. Kieslich* **2695**
- 430. A Computational Pipeline to Optimize 3D Scaffolds for Cancer Immunotherapy**  
*Lucy Todd, Matthew Chin, Marc-Olivier Coppens* **2703**
- 431. Developing Cole-Cole model for *Bacillus subtilis* fermentation**  
*Kumar Rajan Gopa, Tom Wenzel, Sankalp Jena, Supasuda Assawajaruwan, Valentin Khaydarov, Leon Urbas* **2709**

- 432. Improving accuracy scores of neural network driven QSAR models of mutagenicity**  
*Alexander D. Kalian, Emilio Benfenati, Olivia J. Osborne, Jean-Lou C.M. Dorne, David Gott, Claire Potter, Miao Guo, Christer Hogstrand* 2715
- 433. Investigating tunable experiment variable effects on hiPSC-CMs maturation via unsupervised learning**  
*Shenbageshwaran Rajendiran, Mohammadjafar Hashemi, Ferdous Frinklea, Nathan Young, Elizabeth Lipke, Selen Cremaschi* 2721
- 434. On the systematic development of large-scale kinetics using stability criteria and high-throughput analysis of curated dynamics from genome-scale models**  
*Konstantinos Mexis, Stefanos Xenios, and Antonis Kokosis* 2727
- T6: Multi scale energy systems engineering (organized by the EFCE energy section)**
- 435. Development and assessment of intensification alternatives on the lignocellulosic bioethanol production process**  
*Yulissa M. Espinoza-Vázquez, Fernando Israel Gómez-Castro, Eduardo Sánchez-Ramírez, Araceli G. Romero-Izquierdo* 2733
- 436. Optimal ship-based CO<sub>2</sub> transport chains from Mediterranean emission points to the North Sea**  
*Federico d'Amore, Matteo C. Romano* 2739
- 437. DynHeat: Heat Exchanger Network Design for Batch Processes via Dynamic Optimization**  
*Dörthe Franzisca Hagedorn, Sören Demandt, Florian Joseph Baader, Christiane Reinert, Niklas von der Aßen* 2745
- 438. Process Simulation Approaches to Reduce Commercial Risk of CCS Projects**  
*Ryan Muir, Ralph Cos, Chloe Smith, Stelios Papastratos* 2751
- 439. Incorporation of non-linear efficiency constraints in green ammonia production design for cost minimisation**  
*Nicholas Salmon,<sup>a</sup> René Bañares-Alcántara* 2757
- 440. Strategic participation of a gas-fired power plant in interdependent electricity and natural gas markets under carbon emission trading schemes**  
*Christos N. Dimitriadis, Evangelos G. Tsimopoulos, Michael C. Georgiadis* 2763



- 441. Modeling and optimization of hybrid P/VSA-membrane separation processes for CO<sub>2</sub> capture from post combustion flue gas**  
*Christos C. Chatziasteriou, Michael C. Georgiadis, Eustathios S. Kikkinides* 2769
- 442. Integration of renewable energy and CO<sub>2</sub> capture and utilization technologies for decarbonization of energy intensive process industries**  
*Calin-Cristian Cormos, Letitia Petrescu, Ana-Maria Cormos, Simion Dragan, Cristian Dinca, Marius Sandru* 2775
- 443. Simultaneous optimisation of energy recovery and water recirculation in a ceramic plant**  
*Miguel Castro Oliveira, Muriel Iten,<sup>a</sup> Henrique A. Matos* 2783
- 444. Embedding Flexibility to the Design of Pressure-Vacuum Swing Adsorption Processes for CO<sub>2</sub> Capture**  
*Steven Sachio, Adam Ward, Ronny Pini, Maria M. Papathanasiou* 2789
- 445. Multiple-Model Fault Detection of PEMFC**  
*A.Allam, M.Mangold, P.Zhang* 2795
- 446. Implications of Heat and Work Integration in Biomass to Fischer-Tropsch Fuels: A simulation Approach**  
*Samukeliso Dube, Baraka Celestin Sempuga, Xinying Liu* 2801
- 447. Modeling and Assessment of Electrochemical Reduction of CO<sub>2</sub> in Amine-based Capture Solvents**  
*Alexios-Spyridon Kyriakides, Panagiotis Kazepidis, Athanasios I. Papadopoulos, Panos Seferlis* 2807
- 448. Design of CO<sub>2</sub> Capture and Mineralization Systems: Integrated Process Optimization and Controllability Assessment in Parallel Infrastructures**  
*Thomas Prousalis, George Gkizas, Athanasios I. Papadopoulos, Panos Seferlis* 2813
- 449. Accurately Modeling Hydropower in the USA**  
*Amanda Farnsworth, Emre Gençer* 2819
- 450. Biogas valorisation to liquid fuels: modelling and setting up process targets**  
*Selusiwe Ncube, Baraka C Sempuga, Xinying Liu* 2825
- 451. Modeling Commercial Fleet Charging and Regulation Strategies via Data-driven Optimization**  
*James Owens, Emre Gençer* 2831

- 452. Economic Benefits from Planned Renewable Installations in the US using Hydrogen and Modular Ammonia Production Units**  
*Apoorv Lal, Fengqi You* 2837
- 453. Data-Driven Robust Model Predictive Control on Building Climate Control with Renewable Energy**  
*Wei-Han Chen, Fengqi You* 2845
- 454. Thermodynamic Modelling and Performance Analysis of Power Plant for Switching from Coal to Biomass**  
*Romuald Coupan, Ahmed Baccouche, Pierre Szymanski* 2851
- 455. Bottleneck-identification methodology and debottlenecking strategy for heat exchanger network with disturbance**  
*Liwon Zhao, Guilian Liu* 2859
- 456. A Power-to-Gas energy system: modeling and operational optimization for seasonal energy supply and storage**  
*Yifan Wang, Luka Bornemann, Christiane Reinert, Niklas von der Aßen* 2865
- 457. Liquid organic hydrogen carriers (LOHC) as a solution for the energy transition: a reactor design assessment**  
*Carlos Prieto, Antonio Sánchez, Mariano Martín* 2871
- 458. Powering Clean and Affordable Energy through Geothermal Integration: A Case Study on Dominica**  
*Keeara Bhagaloo, Rehannah Ali, Anastasia Baboolal and Keeran Ward* 2877
- 459. Optimal design of a water energy food nexus system through a multi-objective approach**  
*Daniel Peña-Torres, Marianne Boix, Ludovic Montastruc* 2883
- 460. Techno-economic comparison of peaking power plants: OCGT+CCS vs Lithium-ion batteries**  
*Mathew Dennis Wilkes, Diarmid Roberts, Solomon Brown* 2889
- 461. Bioenergy with carbon capture and storage (BECCS) – power generation coupled with hydrogen production**  
*Mathew Dennis Wilkes, Oludayo Asuni, Solomon Brown* 2895
- 462. Optimizing the Sustainable Energy Transition: A Case Study on Trinidad and Tobago**  
*Sherard Sadeek, Dhurjati Chakrabarti, Maria M. Papathanasiou, Keeran Ward* 2901

- 463. Optimal operation of a sustainable water desalination unit integrated a green energy system**  
*Ali Ahmadian, Ali Almansoori, Ali Elkamel* **2907**
- 464. Optimal production of green hydrogen with grid assistance for enhanced flexibility**  
*Christopher Varela, Mahmoud Mostafa, Edwin Zondervan* **2915**
- 465. Water sources evaluation for green hydrogen production: a case study in Brazil**  
*Leonardo O. S. Santana, Artur S. Bispo, Gustavo S. Santos, Chrislaine B. Marinho, Ewerton E. S. Calixto, José L. G. Almeida, Fernando L. P. Pessoa* **2921**
- 466. Sustainable Development Goals assessment of carbon capture on-board**  
*Valentina Negri, Margarita A. Charalambous, Juan D. Medrano-García, Gonzalo Guillén-Gosálbez* **2927**
- 467. Parametric study of the reactive absorption of CO<sub>2</sub> for soda ash production**  
*Maria F. Gutierrez, Peter Schulze, Andreas Seidel-Morgenstern, Heike Lorenz* **2933**
- 468. Machine learning enabled modelling and sensitivity analysis for the power generation from a 660 MW supercritical coal power plant**  
*Waqar Muhammad Ashraf, Vivek Dua* **2939**
- 469. Hydrogen production using renewable energy: solar PV and offshore wind power – An economic evaluation in Bahia**  
*Gustavo de S. dos Santos, Chrislaine do B. Marinho, Leonardo O. S. Santana, Artur S. Bispo, Fernando L. P. Pessoa, José L. G. Almeida, Ewerton E. S. Calixto* **2945**
- 470. Resilience-aware multi-scale integration of distributed energy systems**  
*Natasha J. Chrisandina, Shivam Vedant, Eleftherios Iakovou, Efstratios N. Pistikopoulos, Mahmoud M. El-Halwagi* **2951**
- 471. Techno-enviro-economic analysis of H<sub>2</sub> economy in China from H<sub>2</sub> production to utilization**  
*Xiaodong Hong, Zuwei Liao, Yao Yang, Jingdai Wang, Yongrong Yang* **2957**
- 472. Towards a novel concept for solid energy storage**  
*Adrian Straub, Bogdan Dorneanu, Harvey Arellano-Garcia* **2963**

- 473. Multiobjective optimization of distributed energy systems design through 3E (economic, environmental and exergy) analysis**  
*Bogdan Dorneanu, Sayeef Miah, Evgenia Mechleri, Harvey Arellano-Garcia* 2969
- 474. Calcium looping for CO<sub>2</sub> capture, H<sub>2</sub> and electricity coproduction in coal fired power plants**  
*B. Basant Kumar Pillai, Ramsagar Vooradi, Sarath Babu Anne* 2975
- 475. 3-E (Energy, Exergy & Ecological) analyses of Chemical Looping Air Separation integrated Oxy-Coal Combustion based Power plant**  
*Agarwal Shruti Sumankumar, Venkata Suresh Patnaikuni, Praveen Kumar Bommineni* 1981
- 476. Reliability analysis for robust power-to-X design and storage sizing under renewable uncertainty**  
*Jeongdong Kim, Meng Qi, Il Moon* 2987
- 477. Systematic Development of Strategies for the Decarbonization of Process Utility Systems**  
*Julia Jiménez-Romero, Adisa Azapagic, Robin Smith* 2993
- 478. Energy and Exergy analysis of Biomass Direct Chemical Looping Combustion for CO<sub>2</sub> capture and Utilization**  
*Shailesh Singh Sikarwar, Ramsagar Vooradi, Venkata Suresh Patnaikuni* 2999
- 479. End-effect mitigation in multi-period stochastic programming of energy storage operations**  
*Teemu J. Ikonen, Dongho Han, Jay H. Lee, Iiro Harjunkoski* 3005
- 480. How to accurately fast-track sorbent selection for post-combustion CO<sub>2</sub> capture? A comparative assessment of data-driven and simplified physical models for screening sorbents**  
*Sai Gokul Subraveti, Luca Riboldi, Hao Yang Xu, Yannick Jooss, Simon Roussanaly, Leif Erik Andersson, Rahul Anantharaman* 3011
- 481. A framework for decision-making to encourage utilization of residential distributed energy systems in Brazil**  
*Ana Paula Alves Amorim, Bogdan Dorneanu, Karen Valverde Pontes, Harvey Arellano-Garcia* 3017
- 482. Hydrogen infrastructure planning for heat decarbonisation in Great Britain**  
*Margarita E. Efthymiadou, Vassilis M. Charitopoulos and Lazaros G. Papageorgiou* 3023

- 483. Optimal Concentrated Solar Plant (CSP) location accounting for social and environmental impact: A three-location study in Spain**  
*José A. Luceño Sánchez, Mariano Martín, Sandro Macchietto* **3029**
- 484. Inventory data generation for prospective lifecycle design through process modeling of energy recovery from waste plastics**  
*Shoma Fujii, Yuichiro Kanematsu, Yasunori Kikuchi* **3035**
- 485. Process intensification in a fixed bed reactor for a small-scale process in the stranded assets**  
*Adrian R. Irhamna and George M. Bollas* **3041**
- 486. Optimal capacity planning integrating ammonia storage for power and heat decarbonisation**  
*Georgios L. Bounitsis, Vassilis M. Charitopoulos* **3047**
- 487. European Union's biomass availability for Sustainable Aviation Fuel production and potential GHG emissions reduction in the aviation sector: An analysis using GIS tools for 2030**  
*Sivaramakrishnan Chandrasekaran, Nora B. Salah, John A. Posada* **3053**
- 488. Statistical modeling of electrodes manufacture for CO<sub>2</sub> electroreduction to value-added products**  
*Jose Antonio Abarca, Guillermo Díaz-Sainz, Iván Merino-García, Jonathan Albo, Angel Irabien* **3059**
- 489. Comparative exergoeconomic analysis of two Organic Rankine Cycle (ORC) configurations**  
*Ladislao E. Méndez-Cruz, Miguel A. Gutiérrez-Limón, Raúl Lugo-Leyte, Mauricio Sales-Cruz* **3065**
- 490. Hybrid polymeric membrane – chemical absorption system for pre-combustion CO<sub>2</sub> capture**  
*Maytham Alabid, Nela Slavu, Marius Sandru, Cristian Dinca* **3071**
- 491. Electrification of ethylene production: exploring the potential of flexible operation**  
*Julia L. Tiggeloven, André P.C. Faaij, Gert Jan Kramer and Matteo Gazzani* **3077**
- 492. An MINLP model for the optimal design of CO<sub>2</sub> transportation infrastructure in industrial clusters**  
*Jude O. Ejeh, Sergey B. Martynov, and Solomon F. Brown* **3083**

- 493. Data Driven Leak Detection in a Real Heat Exchanger in an Oil Refinery**  
*Aslı Yasmal, Gizem Kuşoğlu Kaya, Emirhan Oktay,  
Ceylan Çölmekci, Erdal Uzunlar* **3089**
- 494. Optimization of flash-separation based CO<sub>2</sub> purification units**  
*Lorenzo Sala, Syed Ali Zaryab, Paolo Chiesa, Matteo Romano,  
Emanuele Martelli* **3095**
- 495. Big data analysis of solar energy fluctuation characteristics and integration of wind-photovoltaic to hydrogen system**  
*Yaoxi Chen, Siyu Yang, Yu Qian* **3101**
- 496. Dynamic Modeling and Control of CO<sub>2</sub> Capture Systems with a Biphasic Solvent**  
*Panagiotis Kazepidis, Athanasios I. Papadopoulos, Panos Seferlis* **3109**
- 497. Development of integrated liquid air energy storage systems based on air separation units using waste energy from power plant: A case study of South Korea**  
*Jaerak Ko, Seolji Nam, Soonho Hwangbo* **3115**
- 498. A systematic method for performing pinch analysis of the Liquid Air Energy Storage (LAES) process**  
*Vuppanapalli Chaitanya, S. Narasimhan, G. Venkatarathnam* **3121**
- 499. Climate-zone-based Techno-Economic Assessment of Different Carbon Capture Technologies for Coal-Fired Power Plants**  
*Mohammadamin Zarei, Ha-Jun Yoon, Chul-Jin Lee* **3127**
- 500. Numerical tool for dynamic simulation of anaerobic digesters including an air-inflated double membrane gasholder**  
*Giberto M. Yuki Junior, Sabine Sochard, Elio Dinuccio,  
Frédéric Marias* **3133**
- 501. An iterated double auction model for peer-to-peer electricity trading**  
*Timothy Huty, Prof Solomon Brown* **3139**
- 502. Comparing operational strategies for alkaline electrolysis systems considering a probabilistic wind power distribution**  
*Lucas Cammann, Johannes Jäschke* **3145**
- 503. Conceptual design of pre-treatment units for co-electrolysis of CO<sub>2</sub> and water**  
*Josephine Vos, Andrea Ramirez, Mar Pérez-Fortes* **3151**

<b>504. A Process Design, Intensification, and Modularization Approach for Membrane-Assisted Reaction Systems</b> <i>Yuhe Tian, Ayooluwa Akintola, Benjamin Akoh</i>	<b>3157</b>
<b>505. Multiscale design and analysis of templated zeolite for Li-O<sub>2</sub> battery with improved discharge capacity</b> <i>Khizar Hayat, Daniel Bahamon, Lourdes F. Vega, Ahmed AlHajaj</i>	<b>3163</b>
<b>506. Numerical Evaluation of Fuel Utilisation Enhancement Strategies in a Solid Oxide Fuel Cell System Fuelled by Hydrocarbons Mixtures</b> <i>Clyde-Theodore Nguimbous Batista, Amirpiran Amiri, Robert Steinberger-Wilckens</i>	<b>3169</b>
<b>507. Heat pumping and renewable energy integration for decarbonizing brewery industry and district heating</b> <i>Daniel Flórez-Orrego, Meire Ribeiro Domingos, François Maréchal</i>	<b>3175</b>
<b>508. Dynamic Simulation to Verify Operability of LNG Plants with Post Combustion Carbon Capture</b> <i>Jaleel Valappil</i>	<b>3181</b>
<b>509. Optimization of wind farm layout and cable network topology under generalized wake effects</b> <i>Demian J. Presser, M. Josefina Fiorini, Diego C. Cafaro</i>	<b>3189</b>
<b>510. Mixed-integer recourse in industrial demand response scheduling with interruptible load</b> <i>Jnana Sai Jagana, Satyajith Amaran, Qi Zhang</i>	<b>3195</b>
<b>511. Technoeconomic Assessment of Digestate Valorization for Biogas Plant</b> <i>Shivom Sharma, Rafael Graf, Jaroslav Hemrle, Adrian Schneider, Francois Maréchal</i>	<b>3203</b>
<b>512. Optimization of Hydrogen systems for prospective Life Cycle Assessment: Well-to-Tank approach</b> <i>Rocky Mashi, Yohan Vincotte, Sofia De-León Almaraz, Catherine Azzaro-Pantel</i>	<b>3209</b>
<b>513. A techno-economic assessment of biomass combustion with CO<sub>2</sub> capture technology</b> <i>Nela Slavu, Maytham Alabid, Marius Sandru, Cristian Dinca</i>	<b>3217</b>

- 514. Process Modelling and Pinch Analysis for an Integrated System of Anaerobic Digestion with Digestate Recycling via Hydrothermal Gasification**  
*Fadilla Noor Rahma, Khanh-Quang Tran, Roger Khalil* 3225
- 515. Optimization of a continuous multi-stage fluidized bed system for CO<sub>2</sub> capture utilizing temperature swing adsorption**  
*Yuri Souza Beleli, José Luis de Paiva, Marcelo Martins Seckler, Galo Antonio Carrillo Le Roux* 3231
- 516. Regional sustainability of food-energy-water nexus considering water stress using multi-objective modeling and optimization**  
*Anupam Satyakam, Rashi Dhanraj, Yogendra Shastri* 3237
- 517. Toluene as effective LOHC: detailed techno-economic assessment to identify challenges and opportunities**  
*Elvira Spatolisano, Angelo Matichecchia<sup>a</sup>, Laura A. Pellegrini, Alberto R. de Angelis, Simone Cattaneo, Ernesto Roccaro* 3243
- 518. Evaluation of Methane Mitigation Technologies in the Upstream Oil and Gas Sector using the TEAM Digital Platform**  
*Marcelo Mathias, Ahad Sarraf Shirazi, Jairo Duran, Alberto Alva-Argaez* 3249
- 519. Thermodynamic description of the CO<sub>2</sub>-AMP-H<sub>2</sub>O system by ENRTL-RK model**  
*Stefania Moiola, Laura A. Pellegrini, Ricardo R. Wanderley, Hanna K. Knuutila* 3257
- 520. An Automated Approach for Emission Reduction Cost Calculation**  
*Maria Victoria Migo-Sumagang, Raymond R. Tan, Kathleen B. Aviso, Dominic C. Y. Foo* 3263
- 521. Lumped-Parameter Heat Exchanger Models for the Robust Dynamic Modelling of Power Generation Cycles**  
*Oliver Ward, Federico Galvanin, Nelia Jurado, Chris Clements, Mohamad Abdallah, Daniel Blackburn, Eric Fraga* 3269
- 522. Energy and economic analysis of integration of water electrolysis with an oxygen-enriched combustion power plant**  
*Yuxing Peng, Pei Liu, Zheng Li* 3275
- 523. CO<sub>2</sub> water-lean capture: Mathematical modelling and analysis of the industrial process**  
*Ulderico Di Caprio, Emine Kayahan, Min Wu, Peter Hellinckx, Tom Van Gerven, Steffen Waldherr and M. Enis Leblebici* 3281



<b>524. Stochastic pinch analysis to target for resources in batch processes</b>	
<i>Md Alquma Haider, Nitin Dutt Chaturvedi</i>	<b>3287</b>
<b>T7: Sustainable supply chains and ecosystems</b>	
<b>525. MINLP modelling and optimization of the supply chain for the renewable production of methanol in Mexico</b>	
<i>Nereyda Vanessa Hernández-Camacho, Fernando Israel Gómez-Castro, José María Ponce-Ortega, Mariano Martín</i>	<b>3293</b>
<b>526. Supply chain optimal planning for revalorization of empty plastic pesticide and fertilizer containers</b>	
<i>Sergio Iván Martínez-Guido, Rubén Bernabé-Martínez, Claudia Gutiérrez-Antonio, Salvador Hernández</i>	<b>3299</b>
<b>527. Developing a Comprehensive Decision Support Optimization Model for Biofuel Supply Chain</b>	
<i>Brook Tesfamichael, Ludovic Montastruc, Stéphane Negny</i>	<b>3305</b>
<b>528. Investigation of economic and environmental tradeoffs of spatially explicit biomass supply chains towards the production of net-negative biofuels</b>	
<i>Eric G. O'Neill, Caleb H. Geissler, Christos T. Maravelias</i>	<b>3311</b>
<b>529. A MILP-based approach to manage logistics in large industrial gas supply chains</b>	
<i>Sergio G. Bonino, Luis J. Zeballos, Akash Moolya, Jose Lainez, Jose M. Pinto, Ignacio E. Grossmann, Carlos A. Méndez</i>	<b>3317</b>
<b>530. Modelling of reverse supply chains in the context of circular economy</b>	
<i>Christiana M. Papapostolou, Emilia M. Kondili</i>	<b>3323</b>
<b>531. Modular vs centralized manufacturing supply chain: identifying the best solution under uncertainty</b>	
<i>Alessandro Di Pretoro, Stéphane Negny, Ludovic Montastruc</i>	<b>3329</b>
<b>532. Stochastic Optimization of Agrochemical Supply Chains with Risk Management</b>	
<i>Saba Ghasemi Naraghi, Zheyu Jiang</i>	<b>3335</b>
<b>533. Supply Chain Optimization Considering Disruption Demand Uncertainty</b>	
<i>Oluwadare Badejo, Marianthi Ierapetritou</i>	<b>3343</b>

- 534. Integrating economic, environmental, and social concerns into the design and planning of supply chains using monetization strategies**  
*Cátia da Silva, Ana Paula Barbosa-Póvoa, Ana Carvalho* 3349
- 535. Strategic low-carbon hydrogen supply chain planning under market price uncertainty**  
*Tushar Rathi, Jose M. Pinto, Qi Zhang* 3355
- 536. Ontology Modelling for Valorisation of Sugarcane Bagasse**  
*Maureen Chiebonam Okibe, Michael Short, Franjo Cecelja, Madeleine Bussemaker* 3361
- 537. Modelling availability and affordability concerns in the design and planning of pharmaceutical supply chains**  
*Inês Duarte; Bruna Mota; Tânia Pinto-Varela; Ana Amaro; Ana Paula Barbosa-Povo* 3367
- 538. Optimization of Decentralized Pine Needle-to-Electricity Supply Chain for Almora District of Uttarakhand, India**  
*Gayatri Gawande, Ankush Halba, Pratham Arora* 3373
- 539. Life Cycle Assessment of Two Liquid Organic Hydrogen Carriers**  
*Camille Bontron, Diogo Rodrigues, Catarina G. Braz, Henrique A. Matos* 3379
- 540. Distribution planning of medical oxygen supply chains under uncertainty**  
*Georgios L. Bounitsis, Yena Lee, Karthik Thyagarajan, Jose M. Pinto, Lazaros G. Papageorgiou, Vassilis M. Charitopoulos* 3385
- 541. Life-cycle assessment of marine biofuels from thermochemical liquefaction of different olive residues in Spain**  
*Sivaramkrishnan Chandrasekaran, Puck Wammes, John A. Posada* 3391
- 542. An optimisation approach for the design of distributed solar farms with shared energy storage in Pakistan**  
*Obaid Khan, Ishanki De Mel, Robert A. Steven, Michael Short* 3397
- 543. Integrating environmental sustainability in next-generation biopharmaceutical supply chains**  
*Miriam Sarkis, Jesslyn Fung, Ming Hei Lee, Andrea Bernardi, Nilay Shah, Maria M. Papathanasiou* 3403

<b>544. Techno-economic assessment of sustainable energy planning on renewable electricity demand-supply networks: A deep learning approach</b>	
<i>Byeongmin Ha, Taehyun Kim, Jou-Hyeon Ahn, Soonho Hwangbo</i>	<b>3409</b>
<b>545. Evaluation of centralized/decentralized configuration schemes of CO2 electrochemical reduction-based supply chains</b>	
<i>Thijmen Wiltink, Stijn Yska, Andrea Ramirez, Mar Pérez-Fortes</i>	<b>3415</b>
<b>546. Optimization of large-scale energy systems to achieve carbon emissions neutrality</b>	
<i>Sanja Potrč, Andreja Nemet, Lidija Čuček, Petar Varbanov, Zdravko Kravanja</i>	<b>3421</b>
<b>547. Multi-objective sustainable supply chain design under uncertainty in energy price</b>	
<i>F.L. Garcia-Castro, R. Ruiz-Femenia, R. Salcedo-Diaz, J. A. Caballero</i>	<b>3427</b>
<b>548. An Ecologically Safe and Socially Just Supply Chain Design for Li-battery</b>	
<i>Ying Xue, Yazeed M. Aleissa, Bhavik R. Bakshi</i>	<b>3433</b>
<b>549. Leveraging Semantics and Machine Learning to Automate Circular Economy Operations for the Scrap Metals Industry</b>	
<i>Manolis Vasileiadis, Konstantinos Mexis, Nikolaos Trokanas, Theodoros Dalamagas, Thomas Papageorgiou, and Antonis Kokossis</i>	<b>3439</b>
<b>550. System identification of scrap metal shredders under minimal process and feedstock information</b>	
<i>Manolis Vasileiadis, Thomas Papageorgiou, Antonis Kokossis</i>	<b>3445</b>
 <b>T8: Education and knowledge transfer</b>	
<b>551. Development of a user-friendly platform for binary interaction parameter estimation</b>	
<i>Oscar D. Lara-Montaño, Fernando I. Gómez-Castro, Emilio Alba-Robles</i>	<b>3451</b>
<b>552. Liquid-Liquid Equilibrium Data Correlation: Predicting a robust and consistent set of initial NRTL parameters</b>	
<i>Juan A. Labarta, José A. Caballero, Antonio F. Marcilla</i>	<b>3457</b>

- 553. Practical learning activities to increase the interest of university applicants in STEM careers in the era of Industry 4.0**  
*Monica Tirapelle, Dian Ning Chia, Fanyi Duanmu, Konstantinos Katsoulas, Alberto Marchetto and Eva Sorensen* **3463**
- 554. Integrating Python in the (bio)chemical engineering curriculum: challenges and opportunities**  
*Fiammetta Caccavale, Carina L. Gargalo, Krist V. Gernaey, Ulrich Krühne* **3469**
- 555. An Online Course for Teaching Process Simulation**  
*Daniel R. Lewin, Assaf Simon, Sapir Lifshiz Simon, Asia Matatyaho Ya'akobi, and Abigail Barzilai* **3475**
- 556. Designing an interactive environment to share educational materials and resources. Application to the Geomatics Hub at UniLaSalle Beauvais**  
*Réjanne LE BIVIC, Sébastien OTTAVI, Pierre SAULET, Pauline LOUIS, Arnaud COUTU* **3481**
- 557. Immersive learning through simulation: implementing twin screw extrusion in Unity**  
*Pedro Santos Bartolomé, Daniel Just, Ariana Bampouli, Simon Kemmerling, Aleksandra Buczko, Tom Van Gerven* **3487**
- 558. An Educational Workshop for Effective PSE Course Development**  
*Daniel R. Lewin, Edwin Zondervan, Meik Franke, Anton A. Kiss, Stefan Krämer, Mar Pérez-Fortes, Artur M. Schweidtmann, Petronella M. (Ellen) Slegers, Ana Somoza-Tornos, Pieter L.J. Swinkels, and Bart Wentink* **3493**
- 559. Teaching strategies for the effective use of computational tools within the Chemical Engineering curriculum**  
*A. Tsatse, E. Sorensen* **3499**
- 560. Teaching courses to STEM students – Lessons learned from the Pandemics**  
*Nelson Chibeles-Martins, Lourdes B. Afonso* **3505**
- 561. Enhancing Human Machine Interface design using cognitive metrics of process operators**  
*Mohammed Aatif Shahab, Babji Srinivasan, Rajagopalan Srinivasan* **3511**

<b>562. A Cloud-based Collaborative Interactive Platform for Education and Research in Dynamic Process Modelling</b>	
<i>Vinay Gautam, Alberto Rodríguez-Fernández, and Heinz A. Preisig</i>	<b>3517</b>
<b>INDEX</b>	<b>3523</b>

# Simulation-based design of regional biomass thermochemical conversion system for improved environmental and socio-economic performance

Leonardo L. Corradini,<sup>a</sup> Aya Heiho,<sup>b</sup> Yuichiro Kanematsu,<sup>b</sup> Ryoko Shimono,<sup>b</sup> Satoshi Ohara,<sup>c</sup> Yasunori Kikuchi<sup>a,b,c</sup>

<sup>a</sup>*Department of Chemical Sytem Engineering, the University of Tokyo, 7-3-1 Hongo, Bunkyo-Ku, Tokyo 113-8656, Japan*

<sup>b</sup>*Presidential Endowed Chair “Platinum Society”, the University of Tokyo, 7-3-1 Hongo, Bunkyo-Ku, Tokyo 113-8656, Japan*

<sup>c</sup>*Institute for Future Initiatives, the University of Tokyo, 7-3-1 Hongo, Bunkyo-Ku, Tokyo 113-8654, Japan*

*ykikuchi@ifi.u-tokyo.ac.jp*

## Abstract

Utilizing regional biomass instead of imported fossil fuels can support sustainability; but the presence of multiple options for products and conversion methods creates a complex scenario. In this study, we are tackling the simulation-based design of a regional system utilizing local biomass by thermochemical conversion technologies. By choosing a remote island, Tanegashima, in Japan, as a case study, the availability of local biomass and the demand for its products was determined. A model on Aspen Plus<sup>TM</sup> was developed to examine the possible yield of biomass products in different scenarios. Life cycle assessment (LCA) and input-output analysis (IOA) were conducted to calculate environmental and socio-economic indicators. Almost all scenarios considered saw greenhouse gas emission reductions and improved economic circularity, with byproduct reutilization proving beneficial for improving resource self-sufficiency.

**Keywords:** Biofuels, process simulation, sugarcane bagasse, woody biomass

## 1. Introduction

Japan’s self-sufficiency ratio is only around 12.1% (IEA, 2020). In 2020, almost 90% of the country’s primary energy supply was fulfilled by fossil resources and only about 3.9% of Japan’s primary energy mix originated from biomass. Although local biomass use in Japan could make the country more self-sufficient and environmentally friendly (Goh et al., 2020), biomass must be converted into multiple different products of interest, like liquid fuels and chemicals. Hydrogen, ethanol, methanol, charcoal, and hydrocarbons usable as diesel are just a few examples of biomass-derived products (Kang et al., 2021). There are also various existing processes that can be utilized to perform this valorization; in some cases, the same end product can be created using different processes. In particular, thermochemical methods are the ones in which the biomass is decomposed by heat in addition to auxiliary chemicals such as gasification and pyrolysis (Wang et al., 2022). In addition to that, when designing a biomass valorization system, there is very likely to be a mismatch in the supply and demand quantities. For these reasons, if local governments and residents in a region make the decision to employ biomass as a fuel source due to the benefits outlined previously, a complex situation arises: a certain

biomass-derived product may have to be prioritized over the others and it is necessary to determine what final product and technology choice would maximize the potential benefits to the population (Sun and Fan, 2020). Previous works in the literature have explored the matching of supply and demand of biomass within regions (Ayoub and Yuji, 2012; Sun and Fan, 2020). To quantify such “potential benefits”, life cycle assessment (LCA) can be employed as a quantitative technology assessment tool (Kikuchi et al., 2010). Input-output analysis (IOA) (Leontief, 1936) could be applicable to illustrate how different sectors of a biomass-based economy interact with each other and to quantify the economic benefits of regional resource utilization (Kikuchi et al., 2020).

In this study, we are tackling the simulation-based design of a regional system utilizing local biomass by thermochemical conversion technologies. By choosing a remote island, Tanegashima, in Japan, as a case study, the availability of local biomass and the demand for its derived products was determined. A model on Aspen Plus™ was developed to examine the possible yield of biomass products in different scenarios. LCA and IOA were conducted to calculate environmental and socio-economic indicators.

## 2. Materials and methods

### 2.1. Tanegashima case study: biomass supply and fuel demand

In order to discuss the design of a biomass valorization system, the technologies and products considered in this study are applied to a case study in Tanegashima. The system boundary in this study is represented on Figure 1.

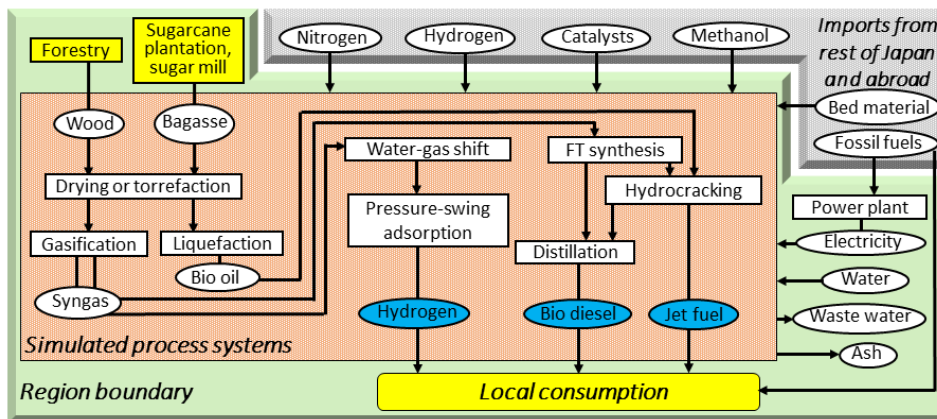


Figure 1: Biomass conversion system and its boundaries

Actual values were investigated and applied for estimating the supply of idle biomass from local forestry and sugarcane industries (Kikuchi et al. 2020; Kikuchi et al. 2016). As for the end product demand, hypotheses were made regarding the adoption of three biomass-derived fuels: hydrogen, jet fuel and biodiesel. Ten representative scenarios were generated to evaluate the effects of different technology and product choices, as listed in Table 1. In the first five, a single product is produced, and different technologies are used. In the other five scenarios, multiple products are manufactured. Scenarios Mix 1, 2 and 3 assume one third of the biomass is allotted to each product, altering the conversion technology between liquefaction, Fischer-Tropsch (FT) synthesis and liquefaction plus torrefaction pre-treatment, respectively. In Mix 4, 70% of the biomass is used for jet fuel production via liquefaction and 30% for hydrogen. In Mix 5, half of the raw materials are allocated for jet fuel production via liquefaction and the rest for diesel production via FT synthesis.

When possible, byproducts like char or flue gases are recycled into different unit operations, or combusted to generate steam and subsequently electricity. Thus, while electricity is always consumed to run some equipment, in some cases there is net generation of electricity.

Table 1: Scenarios of products and technologies applied for local biomass resources (Gas.: Gasification, FT: Fischer-Tropsch, Tor.: torrefaction, Liq.: liquefaction)

Allocation (wt%)	H <sub>2</sub>	Jet (FT)	Jet (Liq)	DL (FT)	DL (Liq)	Mix 1	Mix 2	Mix 3	Mix 4	Mix 5
H <sub>2</sub>	100					33	33	33	30	
Jet Fuel		100	100			33	33	33	70	50
Bio Diesel				100	100	33	33	33		50
Technology	Gas.	FT	Liq.	FT	Liq.	Liq.	FT	Liq. with Tor.	Liq.	Liq. And FT

## 2.2. Modeling of biomass conversion technologies in Aspen Plus<sup>TM</sup>

The process inventory data was extracted from the process simulation in Aspen Plus (Kikuchi et al., 2010). Biomass is initially described as a “non-conventional” component.

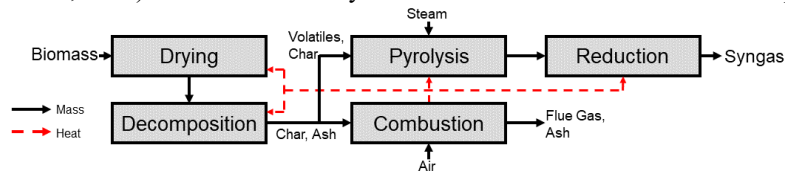


Figure 2: Gasification model scheme

As an example of biomass conversion modeling, the gasification model shown on Figure 2 is explained here. Dual fluidized bed gasification (Müller, 2013) was adopted in this study. The first stage, drying, was modeled with the RStoic reactor. (Patrocínio, 2017) The next block is the decomposition of biomass into conventional components modeled by an RYield reactor according to elemental balance (Abdelouahed et al., 2012). The pyrolysis occurs after decomposition process. The steam utilized as the gasifying agent is also an input in an RGibbs reactor at 400°C and atmospheric pressure. In this case, the final gasifier reactor is modeled as a CSTR in which reduction reactions occur in the gas phase. The combustion bed is also simulated on Aspen Plus by an RStoic reactor. Additionally, liquefaction is modeled with the RYield reactor taking data from previous studies (Wright et al., 2010). The torrefaction model developed by Peduzzi et al. was adopted (Peduzzi et al., 2014). Downstream processes were modeled by more conventional models (Saeidi et al., 2017).

## 2.3. Settings for LCA and IOA

The LCA boundary encompasses the conversion process itself, with biomass as the main input and the three potential products as the main outputs. Other auxiliary inputs necessary for operation such as electricity are also included. The inventory data was from IDEA v3.1 (AIST, 2022), a Japanese LCA database. The functional unit was defined as the utilization of local biomass for energy. The reference flow was set as unit mass of dry biomass input. IOA was performed based on the IO table for Tanegashima (Kikuchi et



al., 2020). IOA in this study was executed by adding a new sector(s) to expand the IO table representing the new products being generated in the island. This requires calculating all the monetary inputs to the new sector (money spent by the manufacturing units) and the outputs; product sales. This is done by estimating the fixed and operational costs (Turton, 2009). The costs of plant construction are divided evenly throughout the 20-year lifetime of the project. Selling price was calculated assuming internal return rate of 11%. The circularity of the economy before and after the changes is then determined.

### 3. Results and Discussion

#### 3.1. LCA results

The life cycle greenhouse gas (GHG) emissions are shown on Figure 3. The contribution of electricity represents a large share of the impact for the single product scenarios.

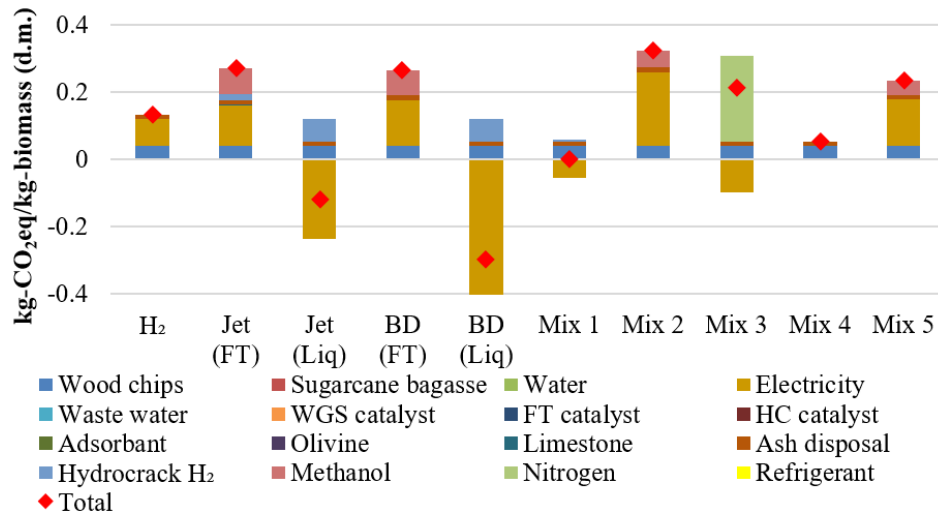


Figure 3: LCA results, life cycle

For FT processes, a relatively large amount of electricity is required in order to compress the syngas before it undergoes CO<sub>2</sub> removal and FT synthesis. Additionally, the operating temperature of gasification is high (>900°C). Therefore, all of the off gases released which could potentially produce electricity were re-directed into the gasifier to supply heat. On the other hand, for liquefaction, a smaller amount of the byproducts (char and flue gases) is required to heat up the pyrolyzer, since it operates at 480°C. Thus, net electricity generation is possible. Besides electricity, it can be seen that the liquefaction routes also stand out due to the large effects of hydrogen used in hydrocracking.

#### 3.2. Input-Output Analysis

The first step of the IOA was the calculation of the monetary flows due to raw materials and utilities induced by the technology implementation. These values are ascribed to pre-existing sectors in the Tanegashima IO table, and utilized to describe the monetary flows in the renewable fuel production sector. After calculating the final effects on the economy, the change in internal and external flows before and after the production of fuels from regional biomass resources are presented on Figure 4. The overall value of the imports into Tanegashima actually only decreases on scenarios H<sub>2</sub> and Mix 2. In all other scenarios, after establishing these new sectors, the monetary value of imports increases.

It can be seen that, if limited to fossil fuels, the values of the imports decrease in all scenarios, however there are increases in other sectors such as petrochemicals.

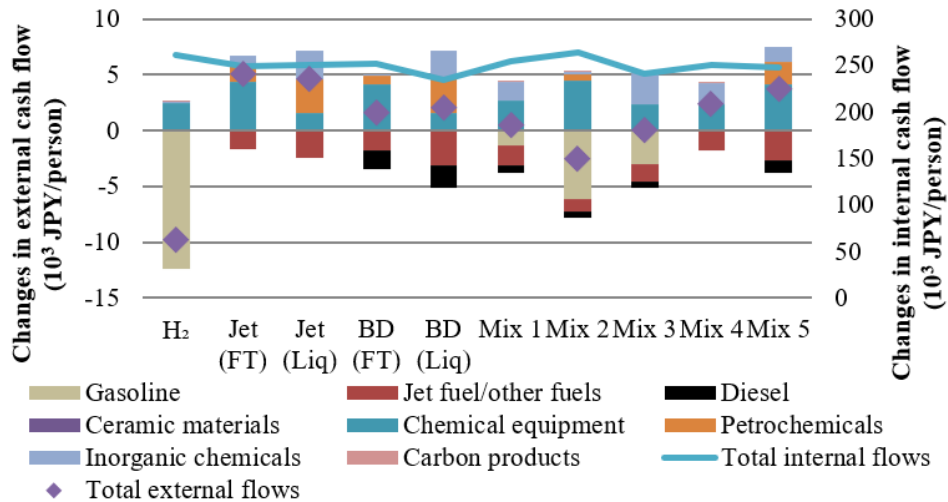


Figure 4: IOA results, where the bar and dots show the changes in external cash flows (left vertical axis) and the line shows those in internal cash flows (right axis).

One of the constants in all scenarios is the chemical equipment category. The capital costs for the plant, as would be expected, proved to be a significant factor in cash outflow. However, overall, the changes in internal flows were much larger, as the high added value of the fuels trickles down throughout many other sectors in the economy.

### 3.3. Interpretation of the multiple indexes for technology implementation

Using the results, three indexes were examined for interpretation: resource self-sufficiency (ratio of product demand fulfilled by the local biomass), GHG reduction potential and economic circularity improvement (increase in the ratio of internal monetary flows to total monetary flows).

Scenario H<sub>2</sub> was the optimal one in this study regarding resource self-sufficiency. This speaks to the efficiency of this conversion route itself and also to the lack of extra inputs required when compared to other scenarios. In fact, the Mix 4 scenario, in which mostly jet fuel is produced with the rest of the biomass generating hydrogen which fulfills the in-house demand for hydrogen in the hydrocracking process, performed better than the Jet (Liq) scenario, also due to the fact that demand for external auxiliary inputs was lower. As for the GHG reduction potential, the Jet (Liq) scenario stands out as the best performing one. This illustrates that, in addition to low GHG emission results from LCA, the process yield is also important as the amount of imported fuel substituted also plays a large role in the GHG emissions reduction. For circularity, scenarios with a reduction in export flows performed better but the difference between scenarios is not as large.

## 4. Conclusion

Production of multiple fuels in mixed scenarios was shown to be more efficient due to the synergy between the thermochemical conversion processes. In general, all cases, especially the ones more reliant on liquefaction, showed potential to cause significant reduction in the GHG emissions in the energy sector. Particularly, electricity was shown to be a significant factor in GHG emissions, thus integrating this sort of biomass valorization with other projects for a greener energy grid could also be beneficial.

## Acknowledgement

We thank Dr. Yuko Oshita for the discussion on the input-output analysis on Tanegashima and Dr. Shoma Fujii for discussions throughout the project. This work was supported by MEXT/JSPS KAKENHI Grant Number JP20K20016, 21K17919, and JST COI-NEXT JPMJPF2003. Activities of the Presidential Endowed Chair for “Platinum Society” at the University of Tokyo are supported by Mitsui Fudosan Corporation, Sekisui House, Ltd., the East Japan Railway Company, and Toyota Tsusho Corporation.

## References

- L. Abdelouahed, O. Authier, G. Mauviel, J. P. Corriou, G. Verdier, A. Dufour, Detailed Modeling of Biomass Gasification in Dual Fluidized Bed Reactors under Aspen Plus, *Energy and Fuels*, 26, 3840 (2012)
- N. Ayoub, N. Yuji, Demand-driven optimization approach for biomass utilization networks, *Computer and Chemical Engineerin*, 36, 129 (2012)
- C. S. Goh, T. Aikawa, A. Ahl, K. Ito, C. Kayo, Y. Kikuchi, Y. Takahashi, T. Furubayashi, T. Nakata, Y. Kanematsu, O. Saito, Y. Yamagata, Rethinking sustainable bioenergy development in Japan: Decentralised system supported by local forestry biomass, *Sustainability Science*, 15, 1461-1471 (2020)
- IEA, World Energy Balances: Overview (2020)
- K. Kang, N. B. Klinghoffer, I. Elghamrawy, F. Berruti, Thermochemical conversion of agroforestry biomass and solid waste using decentralized and mobile systems for renewable energy and products, *Renewable and Sustainable Energy Reviews*, 149 (2021)
- Y. Kikuchi, K. Mayumi, M. Hirao, Integration of CAPE and LCA Tools in Environmentally-Conscious Process Design: A Case Study on Biomass-Derived Resin, *Computer-Aided Chemical Engineering*, 28, 1051-1056 (2010)
- Y. Kikuchi, M. Nakai, Y. Kanematsu, K. Oosawa, T. Okubo, Y. Oshita, Y. Fukushima, Application of technology assessments into co-learning for regional transformation: A case study of biomass energy systems in Tanegashima, *Sustainability Science*, 15, 1473-1494 (2020)
- Y. Kikuchi, Y. Kanematsu, M. Ugo, Y. Hamada, T. Okubo, Industrial Symbiosis Centered on a Regional Cogeneration Power Plant Utilizing Available Local Resources: A Case Study of Tanegashima, *Journal of Industrial Ecology*, 20(2) 276-288 (2016)
- W. W. Leontief, Quantitative Input and Output Relations in the Economic Systems of the United States, *Review of Economics and Statistics*, 18, 105 (1936)
- S. Müller, Hydrogen from Biomass for Industry, Vienna University of Technology (2013)
- LCI database IDEA version 3.1; National Institute of Advanced Industrial Science and Technology (2022)
- F. Patrocínio, Aspen simulation of biomass conversion processes, *IST* (2017)
- E. Peduzzi, G. Boissonnet, G. Haarlemmer, C. Dupont, F. Maréchal, Torrefaction modelling for lignocellulosic biomass conversion processes, *Energy*, 70, 58 (2014)
- S. Saeidi, F. Fazlollahi, S. Najari, D. Iranshahi, J. Klemeš, L. Baxter, Hydrogen production: Perspectives, separation with special emphasis on kinetics of WGS reaction: A state-of-the-art review, *Journal of Industrial and Engineering Chemistry*, 49, 1-25 (2017)
- O. Sun, N. Fan, A Review on Optimization Methods for Biomass Supply Chain: Models and Algorithms, Sustainable Issues, and Challenges and Opportunities, *Process Integration and Optimization for Sustainability*, 4, 203 (2020)
- R. Turton, R. C. Bailie, W. B. Whiting, J. A. Shaeiwitz, *Analysis, Design and Synthesis of Chemical Processes* (2009)
- W. C. Wang, Y. C. Liu, R. A. A. Nugroho, Techno-economic analysis of renewable jet fuel production: The comparison between Fischer-Tropsch synthesis and pyrolysis, *Energy*, 239, 121970 (2022)
- M. M. Wright, J. A. Satrio, R. C. Brown, D. E. Daugaard, D. D. Hsu, Techno-Economic Analysis of Biomass Fast Pyrolysis to Transportation Fuels (2010)

# Conceptual Design of a Biorefinery to Use Brown Seaweed *Sargassum*

Omar Flores-Mendoza,<sup>a</sup> Teresa Lopez-Arenas<sup>a</sup>

<sup>a</sup>*Departamento de Procesos y Tecnología, Universidad Autónoma Metropolitana – Cuajimalpa, Vasco de Quiroga 4871, Mexico City 05348, Mexico*  
*mtlopez@cua.uam.mx*

## Abstract

The massive arrival of sargassum to the Caribbean beaches has generated a problem for its collection, treatment, disposal, and use. Therefore, the recovery of sargassum has become a topic of interest from the social, economic, and environmental point of view. The objective of this work is to propose a conceptual design of a sargassum biorefinery to produce three value-added products: alginate, a biopolymer and a biofertilizer. The methodology consisted mainly of the synthesis, design, and techno-economic evaluation of the biorefinery, using modeling and simulation tools. Comprehensive evaluation of the biorefinery design shows its feasibility (with moderate product yields) and profitability (with satisfactory return on investment and payback time) on an industrial scale.

**Keywords:** sargassum, biorefinery, circular economy, process simulation.

## 1. Introduction

Since 2011, the arrival of certain brown seaweed in the Caribbean, known as *Sargassum*, has been a problem not only for the tourism and fishing sectors, but also for the environment, given that the leachate produced by the seaweed saturates the habitat of nutrients and its presence in the sea prevents the photosynthesis of seagrasses. Moreover, the sargassum degradation by microorganisms in the open air generates greenhouse gases (GHG). Currently, the collected sargassum is stored in landfills and is used mainly as fertilizer. However, in some areas of the Caribbean, sargassum is also used to make bricks, cosmetics, and paper products; but they must be viewed with caution due to their ability to accumulate toxic metals. Thus, the implementation of the circular economy through sargassum biorefineries (Azcorra-May et al., 2022) can help cushion the secondary effects of its arrival, as well as obtain a benefit for society by transforming seaweed into useful bioproducts such as precursors for food, biofuels, enzymes, or biopolymers, generating a minimum of waste and emissions (Amador-Castro et al., 2021).

The species of *Sargassum natans* and *Sargassum fluitans* reproduce in the subtropical West Atlantic, which includes the Mexican Caribbean coasts, and in the North Equatorial Recirculation Region, which is located in the tropical area between Brazil and Africa (Wang, 2019). The main components of the cell wall of this brown seaweed are anionic polysaccharides, of which are alginates and sulfated polysaccharides that contain fucose; as well as cellulose, hemicellulose, lignin, arabinogalactan proteins and polyphenols (Kloareg, 2021). On the one hand, alginate is a compound widely used in various industries such as food (to maintain viscosity or retard crystallization), dental (as an impression material), paper (to provide smooth, continuous films with great ink and print

retention quality), and medical (for the creation of biomedical and clinical materials such as tissue engineering). While on the other hand, hemicellulose and cellulose are a source for the generation of sugars (pentoses and hexoses) that can be fermentable to produce biofuels, such as bioethanol, biogas, or biodiesel, as well as other bioproducts with high added value such as biopolymers, organic acids, amino acids, among others.

This work presents the design and evaluation of a sargassum biorefinery to produce alginate, PHB (polyhydroxybutyrate) and a biofertilizer, using process modeling and simulation tools. The case study considers the brown seaweed sargassum that arrives at one of the main beaches of the Riviera Maya, Mexico. First, a conceptual design of the biorefinery is proposed, then the process implementation and the technical evaluation are carried out in a modular process simulator, and then simulations are performed to determine the best process configuration, the equipment dimensions, the operating conditions, as well as the productivity and performance of the products. At this early stage of the process design, a non-cash flow economic evaluation is performed to determine the profitability of the process.

## 2. Simulation methodology

The methodology consisted mainly of the synthesis, design, and techno-economic evaluation of the biorefinery. To carry out the synthesis of the biorefinery, first a bibliographic review was carried out to determine the composition of the sargassum and the best processing alternatives, identifying the physical, chemical, and biological phenomena. So that the unitary operations, the reaction routes, physicochemical properties of the components, etc. were established. Subsequently, the design of the biorefinery was implemented in a process simulator, considering an industrial scale, and taking into account equipment sizing and design, operating conditions, production programming, cost information (raw materials, equipment, operation, investment), etc. And finally, the simulation of the process was carried out through material and energy balances to assess technical-economic aspects, such as: product yields, installed production capacity, utilities consumption and profitability.

### 2.1. Availability and composition of sargassum

Currently the sargassum is collected on the beaches, followed by a drying process for storage and use. In particular, the amount of sargassum on the beach of Tulum, in the Mexican Caribbean, in recent years has ranged between 4 500 and 9 000 tons per year. The changes in its composition of cellulose, calcium alginate, hemicellulose and lignin vary depending on the season in which the sargassum is harvested. For simulation purposes, an average of each of the compounds was obtained (Bertagnolli et al., 2014; Lopez et al., 2020), corresponding to a composition on a dry basis of: 23% calcium alginate, 11.15% cellulose, 9.6% lignin, 8.2% hemicellulose, and 48.05% other solids.

### 2.2. Selection and selling cost of products

Alginate has been selected as the main product due to its diversity of applications and its high selling cost, which was set in the simulation at 20 USD/kg (Galindo et al., 2007). The PHB was selected as the second product, since from the point of view of the circular economy, the reduction of plastic consumption and its replacement with biopolymers should be encouraged. The PHB income was established at 4 USD/kg (Lopez-Arenas et al., 2017). While organic solids, a by-product of the PHB fermentation process, can be recovered as a biofertilizer due to its nutrient content. The unit selling cost was set at 0.35 USD/kg (Praveen and Singh, 2019).

### 2.3. Synthesis and design of the biorefinery

As part of the synthesis of the conceptual design, five processing sections of the biorefinery were identified: raw material conditioning, alginate extraction, alginate purification, PHB production, and PHB and biofertilizer purification. The corresponding flowsheet diagram of the sargassum biorefinery is shown in Figure 1, and each section is briefly described below.

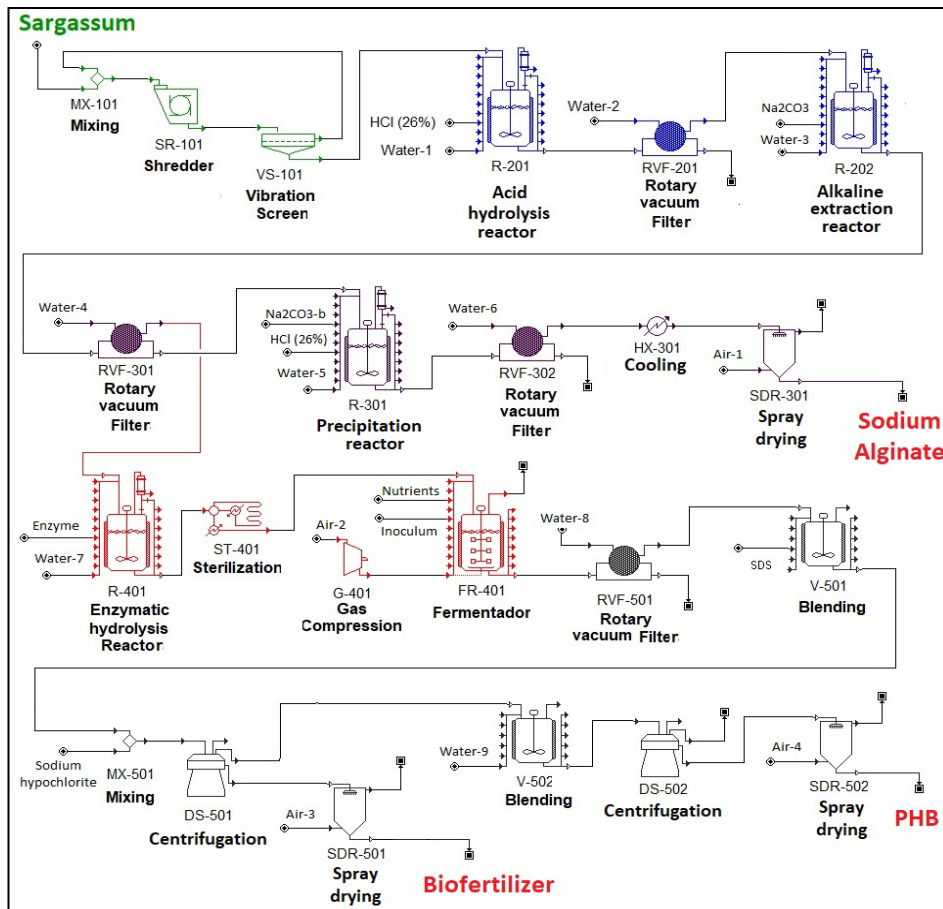


Figure 1. Process flowsheet diagram of the sargassum biorefinery.

#### 2.3.1. *Sargassum* conditioning

First, the size of the dried seaweed is reduced to increase the exposed area and therefore the mass transfer. This is done by shredding (SR-101) and subsequent screening (VS-101) to ensure maximum particle size.

#### 2.3.2. Alginate extraction

The crushed seaweed enters an acid hydrolysis reactor (R-201) to separate the cellulose and hemicellulose from the lignin, and to transform the calcium alginate into alginic acid. After removing the acidic liquid phase through a rotary filter (RVF-201), the alkaline extraction is carried out in the reactor R-202 with the addition of calcium carbonate, converting the alginic acid into sodium alginate. Subsequently, the liquid phase containing the soluble alginates is separated from the rest of the lignocellulosic biomass.

### 2.3.3. Alginate purification

The liquid phase rich in alginates is introduced into the reactor R-301, where hydrochloric acid and calcium carbonate are added to precipitate sodium alginate. Finally, the product is filtered (RVF-302) and dried (SDR-301).

### 2.3.4. PHB production

Lignocellulosic biomass is converted into glucose through an enzymatic hydrolysis reaction (R-401) using glucanases. Then, the hydrolyzed product is sterilized (ST-401) to avoid any contamination and production of unwanted compounds. The hydrolyzate passes to a fermenter (FR-401) for the intracellular production of PHB, using *Azohydromonas australica* as bacterial strain.

### 2.3.5. PHB and biofertilizer purification

Cell granules are recovered by filtration (RVF-501), and then PHB is extracted (V-501) by cell lysis by adding dodecyl sulfate surfactant. A sodium hypochlorite wash is then performed to achieve maximum cell disruption. Residual biomass and PHB are separated (DS-501). On the one hand, the residual biomass is dried for sale as a biofertilizer. And on the other hand, the stream with PHB is washed to remove the surfactant and sent for centrifugation (DS-502). Finally, the product is dried (SDR-502) to obtain the PHB in solid form.

## 2.4. Technoeconomic evaluation

Technoeconomic evaluation plays a crucial role in helping to select the best biorefinery conceptual design alternatives. From the technical point of view, product yields, processing times, installed capacity in terms of raw material and production rate are verified. In particular, the amount of required utilities is calculated, such as the consumption of heat transfer agents (cooling water and steam) and standard energy. While, to carry out the profitability analysis, an economic model without discounted cash flow was used to calculate the unit production cost, the rate of return on investment (ROI) and its payback time (PT). These cost effectiveness criteria provide a snapshot of the economics of the process and are often used to make preliminary estimates during the synthesis stage of process design. However, more rigorous profitability criteria could be used, in which cash flow is considered (such as the net present value method, NPV, or the internal rate of return, IRR) (Seider et al., 2009).

## 3. Simulation results

In this work, the biorefinery flowsheet diagram was implemented in *SuperPro Designer*, a commercial modular process simulator that allows the evaluation of material and energy balances, as well as the economic balance. This modular process simulator (like other commercial simulators such as *Aspen Plus*) includes databases of physicochemical properties of components, thermodynamic models, unit operations models, costs of raw materials and equipment, economic models, etc.

The biorefinery was simulated in batch mode and considering an annual operating time of 48 weeks. Each process equipment takes into account the times for material loading, process operation, material unloading and equipment cleaning, according to the case. The conceptual design proposal for the biorefinery considers a feed of 15 tons of sargassum per batch. Each process batch has a duration of 161.6 hours; however, each new production cycle can start after 37.3 hours of the previous cycle. Such that 212 batches are produced per year. A summary of the technical-economic evaluation is presented in Table 1.

Table 1. Technoeconomic report for the sargassum biorefinery.

<b>Technical assessment</b>			<b>Economical assessment</b>		
Concept	Value	Units	Concept	Value	Units
Sargassum consumption	3 180	t/yr	Direct fixed capital	32 822 000	USD
Sodium Alginate production	764.3	t/yr	Working capital	314 000	USD
PHB production	153.1	t/yr	Startup cost	1 641 000	USD
Biofertilizer production	0.472	t/yr	Total investment	34 777 000	USD
Sodium Alginate yield	0.2403	kg/kg S*	Operating cost	8 941 000	USD/yr
PHB yield	0.0485	kg/kg S*	ROI	19.98	%
Biofertilizer yield	0.0001	kg/kg S*	PT	5.0	yr
<b>Water and utilities consumption</b>					
Concept	Value	Units			
Process water	142.9	kg/kg SA*			
Steam	55.3	kg/kg SA*			
Cooling Water	5.5	t/kg SA*			
Standard Power	6.8	kW-h/kg SA			

\*S and SA to Sargassum and Sodium Alginate, respectively.

According to the technical results, it can be seen that the installed capacity of the plant in terms of annual sargassum consumption is within the availability of the Mexican beach of Tulum, as reported in the introduction section. The best product yield is obtained for sodium alginate, since it is the first product extracted; followed by the PHB yield, and finally, the yield of biofertilizer is quite low, which means that little residual biomass is recovered and its contribution to profitability is minimal.

Regarding the economic results, it can be observed that despite the estimated total investment being high, the biorefinery design shows positive and satisfactory profitability values for ROI and PT. According to Seider et al. (2009), an acceptable profit margin should be between 15-20% for ROI and 3 years for PT. It is worth mentioning that the operating cost considers various cost items, whose distribution is shown in Fig. 2(a). As can be seen, the utility cost is 15% of the annual operating cost, corresponding to the utility consumption reported in Table 1. The utility cost distribution is shown in Fig. 2(b), where the standard power (i.e., electricity) has the largest contribution. Therefore, alternatives to reduce energy consumption should be studied to reduce the cost and its environmental impact.



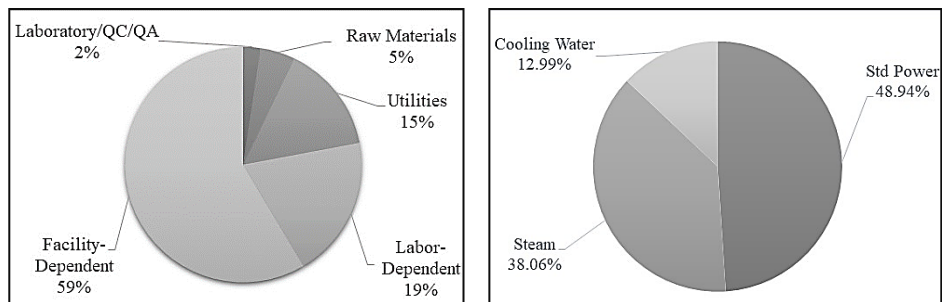


Figure 2. (a) Annual operating cost distribution and (b) utility cost distribution.

#### 4. Conclusions

The results showed that the implementation of a biorefinery from the perspective of the circular economy allows addressing the problem of sargassum, not only to mitigate the secondary effects of its arrival, but also that a benefit could be obtained by transforming this seaweed into useful bioproducts such as precursors for biomaterials and food, biopolymers, fuels, among others. In particular, the profitability of the proposed biorefinery is supported by the production of sodium alginate, which is a product with high added value and great potential for use in biomedical areas. Even though the conceptual design presented shows satisfactory profitability, several aspects can be improved and studied further, such as analyzing other process configurations, optimizing operating conditions, evaluating different installed capacities, and more.

#### References

- F. Amador-Castro, T. García-Cayuela, H.S. Alper, 2021, Valorization of sargassum biomass into sustainable applications: Current trends and challenges, *J. Environ. Manage.*, 283, 112013.
- K. J. Azcorra-May, E. Olguin-Maciel, J. Dominguez-Maldonado, 2022, Sargassum biorefineries: potential opportunities towards shifting from wastes to products, *Biomass Convers. Biorefin.*
- C. Bertagnolli, A.P. Espindola, S.J. Kleinübing, L. Tasic, M.G. da Silva, 2014, Sargassum *filipendula* alginate from Brazil: Seasonal influence and characteristics. Elsevier.
- E. Galindo, C. Peña, C. Núñez, D. Segura, G. Espín, 2007, Molecular and bioengineering strategies to improve alginate and polyhydroxyalkanoate production by *Azotobacter vinelandii*. *Microb Cell Fact* 6, 7 (2007).
- B. Kloareg, Y. Badis, J. M. Cock, G. Michel, 2021, Role and Evolution of the Extracellular Matrix in the Acquisition of Complex Multicellularity in Eukaryotes: A Macroalgal Perspective, *Genes*, 12 (7), 1059.
- L.B. Lopez, J.J. Alvarado, J.C. Corral, A. Aguilera, R.E. Rodríguez, S.J. Guevara, J.V. Alcaraz, J.G. Rutiaga, J. Zárata, M.L. Ávalos, M. Morales, 2020, A prospective study of the exploitation of *pelagic Sargassum spp.* As a solid biofuel energy source. *Applied Sciences*, 10(23), 8706.
- T. Lopez-Arenas, M. González-Contreras, O. Anaya-Reza, M. Sales-Cruz, 2017, Analysis of the fermentation strategy and its impact on the economics of the production process of PHB (polyhydroxybutyrate), *Computers & Chemical Engineering*, 107, 140-150.
- K.V. Praveen, A. Singh, 2019, Realizing the potential of a low-cost technology to enhance crop yields: evidence from a meta-analysis of biofertilizers, *Agricultural Economics Research Review*, 32, 77-91.
- D.W. Seider, J. D. Seader, D. R. Lewin, S. Widagd, 2009. *Product and Process Design Principles: Synthesis, Analysis and Design*. John Wiley & Sons, New York.
- M. Wang, C. Hu, B. Barnes, G. Mitchum, B. Lapointe, J.P. Montoya, 2019, The great Atlantic Sargassum belt, *Science*, 365(6448), 83-87.

# Design and Production of Biodiesel and Porous Filaments from Endemic Species of the Ciénega de Chapala Region

María Fernanda Torres-Magaña<sup>a</sup>, José Luis Baltazar-Álvarez<sup>a</sup>, Pascual Eduardo Murillo-Alvarado<sup>a</sup> and Gabriela Guadalupe Esquivel-Barajas<sup>a</sup>

<sup>a</sup>Universidad de La Ciénega del Estado de Michoacán de Ocampo. Av. Universidad Sur 3000, Lomas de Universidad, 59103 Sahuayo de Morelos, Michoacán México.

## Abstract

The production of biofuels has contributed to reducing the use of fossil fuels, however, in the means of obtaining biofuels, waste materials are generated, which must be given extra treatment for their storage. In such a way that the objective of the present investigation was to obtain biodiesel through the transesterification method, using the seeds of 3 plants (*Ricinus communis*, *Pithecellobium dulce* and *Prosopis laevigata*), which are endemic to the Ciénega Region of the State of Michoacán, Mexico. The generated residual biomass served as raw material for the synthesis of bioplastics. Both biodiesel and bioplastics were characterized by Fourier Transform Infrared Spectroscopy (FT-IR) and UV-VIS, thus it was possible to identify absorption bands, corresponding to the functional groups, which are similar to those of hydrocarbons. Thus, it is concluded that the remnants obtained during biodiesel production processes can be used for a second use as bioplastics.

**Keywords:** biodiesel, bioplastic, seeds.

## 1. Introduction

Due to anthropogenic actions, the planet has manifested various changes, due to the effect of greenhouse gases produced by the consumption of fossil fuels, which emit CO<sub>2</sub> and sulfur, for which reason it has ventured into the use of alternative fuels, such is the case of obtaining and applying biofuels.

Biofuels are fuels produced from raw materials of vegetable or animal origin. Another way of obtaining them is through the processing of agro-industrial products or organic waste. Arellano C. (2015), describes biofuels as an alternative of interest to reduce the accumulation of greenhouse gases and thus reduce dependence on the use of fossil fuels, in the same way in his manuscript he describes that when using biomass from In the food sector, an alternative is the production of lignocellulosic materials, since these do not compete directly with food intended for both human and animal consumption, which makes it possible to take advantage of biomass generated from agro-industrial processes.

Mentioned the above, the objective of the research has been based on the use of 3 endemic plants of the Ciénega region, which are: castor plant (*Ricinus communis*), mezquite (*Pithecellobium dulce*) and guamúchil (*Prosopis laevigata*), of which two processes are derive, the first is to obtain biodiesel from plants and the second is to take advantage of the lignocellulosic remnants so that they have become biofilms, which have properties similar to bioplastics.

## 2. Biopolymers

Biopolymers are macromolecular materials, which can be synthesized naturally or synthetically. The classification of biopolymers is based on three families, which are: proteins, polysaccharides and nucleic acids. Among the most abundant biopolymers are cellulose and starch, which are of interest for the present investigation.

### 2.1 Bioplastics

Bioplastics are materials that have an organic composition, the main polymers that characterize them being both starch and cellulose (Riera A., & Palma R., 2018). It is necessary to have chemical structures that allow the degradation of materials from native biota where they have been discarded (mainly they must be degraded by soil microorganisms).

### 2.2 Endemic plants and trees of the Ciénega de Michoacán Region

The use of the material that is used for the generation of a biofuel must be careful not to break the fine line of balance between the use of organic waste and the use of organic resources for human consumption (Merino P., & Nonay M., 2009) that is why for this project in the generation of two bioproducts. In biodiesel, the decision was made to use seeds of fruits that have been considered as plants of low nutritional value for humans such as Guamúchil and Mezquite which are characterized by be abundant trees in the Ciénega de Chapala region.

## 3. Methodology

For the development of the present investigation, it was divided into 5 stages which consist of: 1) sampling of endemic plant remnants, 2) extraction of essential oils and obtaining biodiesel, 4) characterization of oils and biodiesel 5) synthesis and characterization of thin films from residual biomass.

*Stage 1.* Identification of endemic plants. Castor weed plants were identified, in addition to mezquite and guamúchil trees, in the municipalities of Sahuayo-Jiquilpan, Michoacan, Mexico.

*Stage 2.* Castor berries were collected, in addition to sampling the remnants of guamúchil and mezquite berries, approximately 2 kg of each berry, which were stored in sterile plastic bags, later stored in a desiccator (Novatech), with the aim of facilitating the extraction of seeds.

*Stage 3.* Extractions of essential oils. For the development of this stage, an adaptation was made to the methodology proposed by Demirbas 2008, which consisted of: grinding the seeds and using the soxhlet method (KIMAX brand) in which 200ml of isopropyl alcohol (MEYER) were poured, subjecting at constant temperature and stirring at 200°C and 200 rpm. In order to extract the essential oil from the seeds, it was necessary to recirculate the alcohol in the extractor 15 times with a duration of approximately 45 minutes each. Upon completion of the 15 cycles, the sample was removed from the grill and placed in a desiccator for a period of 24 hours to remove as much water and alcohol as possible to subsequently obtain the pure essential oil from the seed. To obtain biodiesel, the essential oil was taken to the centrifuge (Hettich brand, EBA 20

model) for 30 minutes at 3,500 rpm and removing the degumming or residues to be able to be mixed with the 1% concentration of the potassium iodide solution (Reasol brand) with an addition of 3% (p/p) of distilled water and place on the grill to maintain agitation for 30 minutes at 60°C to finish it was placed in the centrifuge at 60 rpm for one hour.

*Stage 4.* Characterization of biodiesel. Biodiesel was characterized by Fourier transform infrared spectroscopy (FT-IR), Perkin Elmer Frontier model in the 4,500-400  $\text{cm}^{-1}$  region and by Perkin Elmer Lambda 25 model UV-VIS spectrophotometry, in the absorbance region of 300-800 nm.

*Stage 5.* Synthesis and characterization of biofilms. For the development of this stage, an adaptation was made to the methodology proposed by Salazar, 2013. To make use of the waste generated from obtaining biodiesel, a bioproduct has been generated that is formed from the remains of seeds that passed by the transesterification method with an adaptation to the methodology carried out by Salazar M., & Valderrama A. (2013) in which the previous conditioning was carried out in an extraction hood (Tecnolab brand) with a grill that preheats a Porcelain evaporation dish (Duve brand) at 75°C in which 30 ml of distilled water (Golden Bell) and 3 ml of natural glycerin are added, shake for 5 minutes; add 0.9 ml of a 1 molar solution of glacial acetic acid (Golden Bell brand) keep stirring for 10 minutes and then pour it into a glass petri dish to take it to the preheated desiccator at 75 °C for 8 hrs. And finally they were characterized by UV-VIS and FT-IR.

#### 4. Results and discussions

Seeds and berries were collected in the municipalities of Sahuayo and Jiquilpan Michoacan Mexico (Fig.1), later the essential oils were obtained from each of the samples, which were characterized by UV-VIS, to identify the band of greater absorption, which oscillated for the 3 samples in the interval of 700-800 nm (Fig. 2).



Fig.1. Seeds and berries of endemic plants of Sahuayo-Jiquilpan, Mich. a) *Prosopis laevigata*, b) *Ricinus communis* and c) *Pithecellobium dulce*.

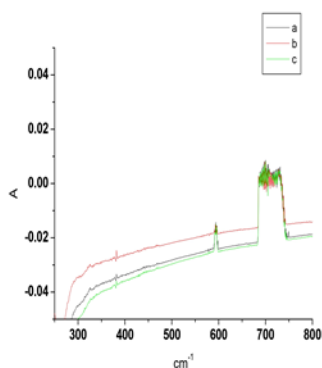


Fig.2. UV-VIS characterization of endemic plants from Sahuayo-Jiquilpan, Mich. a) *Prosopis laevigata*, b) *Ricinus communis* and c) *Pithecellobium dulce*.

Biodiesel was synthesized through the transesterification method, which was characterized by FT-IR, with the aim of identifying its functional groups (Fig. 3). Where it can be seen that both in *Ricinus communis* and in *Pithecellobium dulce* there is a greater predominance by the OH group, since their absorption bands are greater compared to that of *Prosopis laevigata*.

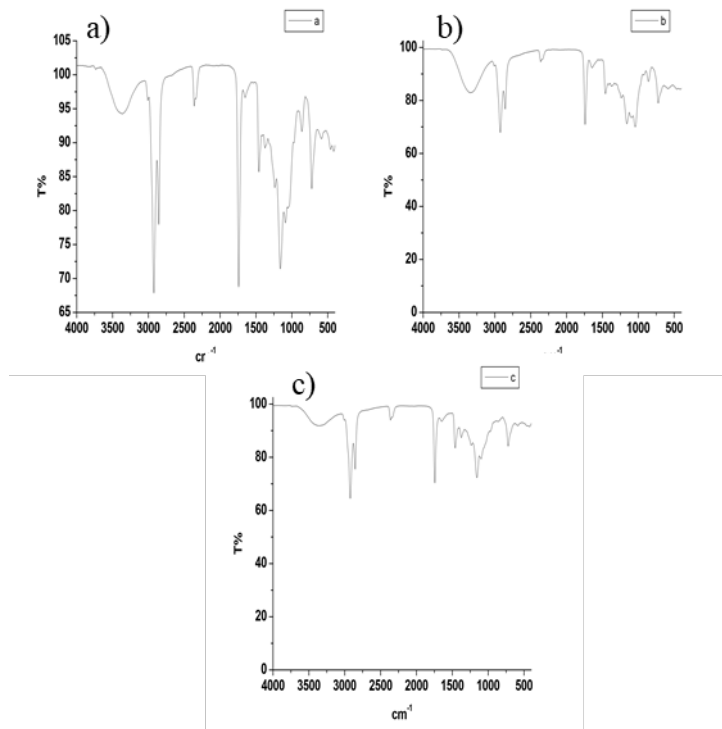


Fig.3. FT-IR characterization of endemic plants from Sahuayo-Jiquilpan, Mich. a) *Prosopis laevigata*, b) *Ricinus communis* and c) *Pithecellobium dulce*.

However, during the synthesis of biodiesel, remnants are generated during its elaboration, in such a way that it was decided to give this residual material a conversion, for which bioplastics were synthesized from it (Fig. 4).

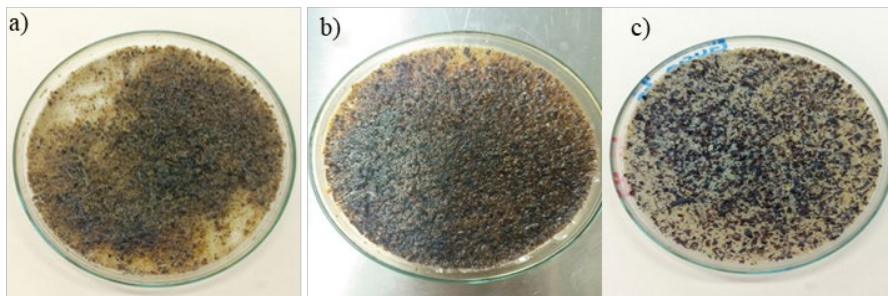


Fig. 4. Bioplastics obtained from biodiesel remnants. a) *Prosopis laevigata*, b) *Ricinus communis* and c) *Pithecellobium dulce*.

The bioplastics were characterized by FT-IR (Fig. 5), where it can be observed that there is a decrease, compared to the characterization of biodiesel, in the absorption bands belonging to the esters and amides groups.

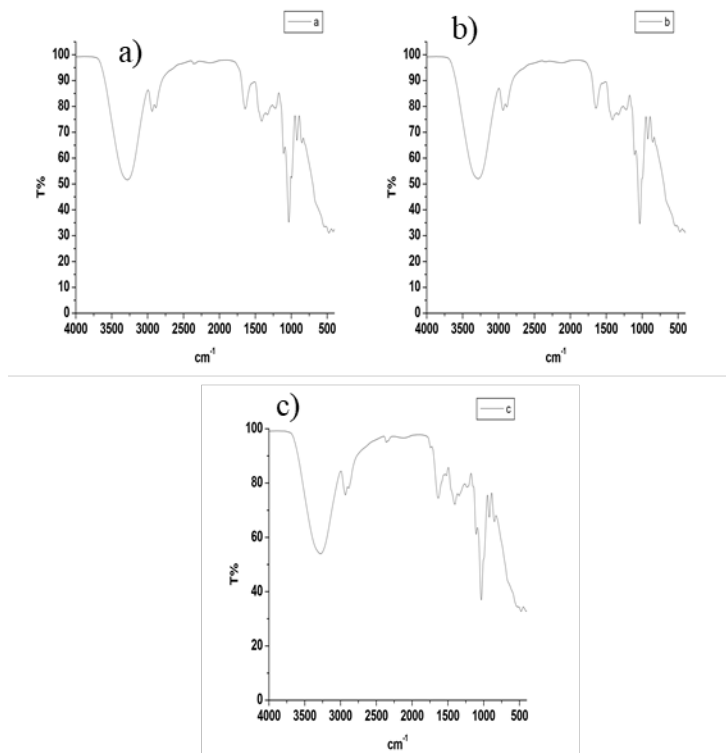


Fig. 5. FT-IR characterization of bioplastics, Mich. a) *Prosopis laevigata*, b) *Ricinus communis* and c) *Pithecellobium dulce*.

In the same way, a proposal was developed where this technology can be migrated on an industrial scale, for which the design at the industry level is proposed (Fig. 6).

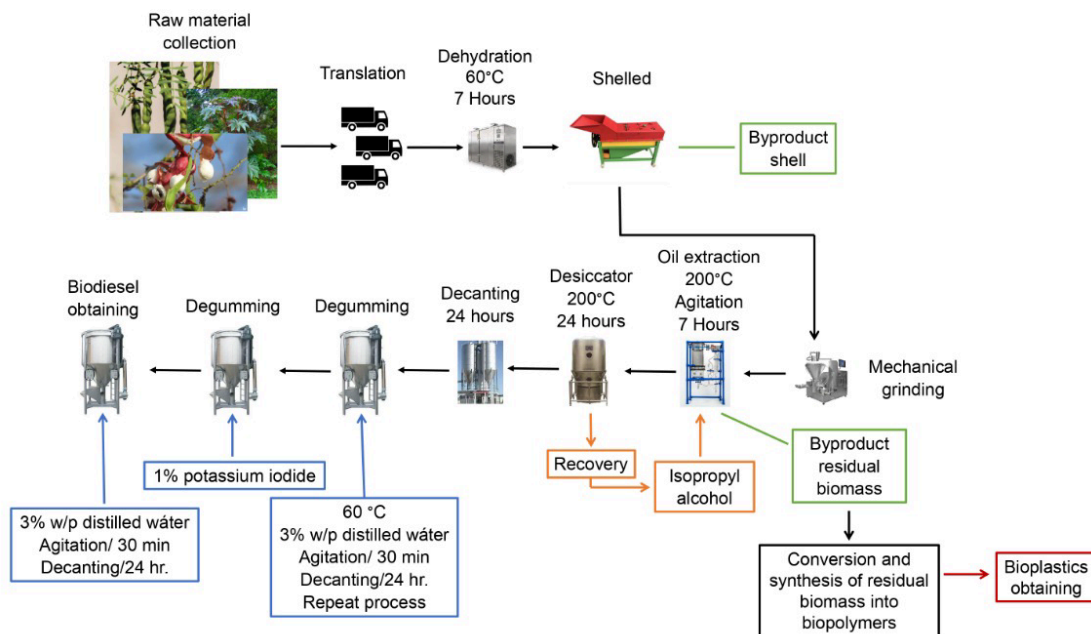


Fig. 6. Biodiesel and bioplastics industrial production scheme.

## 5. Conclusions

The transesterification method is suitable for obtaining biofuels, in particular it is the case of biodiesel. However, the synthesis process involves various reagents, which in turn produce materials of waste, for which it is considered viable to use them to obtain bioplastics, which can serve as environmentally friendly materials.

## 6. References

- Arellano Perales Carla (2015) Obtención de bioetanol a partir de materiales lignocelulósicos sometidos a hidrólisis enzimática. Facultad de ciencias químicas. P. p. 3-5.
- Demirbas Ayhan 2008, Biodiesel from Triglycerides via Transesterification, Biodiesel A Realistic Fuel Alternative for Diesel Engines, P.p. 121-139.
- Merino Garcia Pedro Antonio y Nonay Domingo Maria Teresa (2009) Descripción, evolución y retos del sector de los biocombustibles. Boletín ICE Economico. No. 2971. P.p. 11-20
- Riera Maria Antonieta and Palma Ricardo R., 2018, Obtención de bioplásticos a partir de desechos agrícolas. Una revisión de las potencialidades en Ecuador, Avances en Química, Vol. 13, No. 3, P.p. 69-78.
- Salazar Max Carlos and Valderrama Negrón Ana, 2013, Preparation And Characterization Of Depolymerised Chitosan Films And Crosslinked With Sodium Tripolyphosphate. Rev. Soc. Quím. Perú Vol.79 No.3. P.p. 197.

# Real-time control and scheduling decision making in fermentation-based food process operations

*Isuru A. Udugama<sup>a,\*</sup>, Keegan K. Hall<sup>a</sup>, Christoph Bayer<sup>b</sup>, Brent R. Young<sup>c</sup>, Timothy G. Walmsley<sup>d</sup>*

*<sup>a</sup>Ahuora – Centre for Smart Energy Systems, School of Engineering, University of Waikato, Hamilton, 3240, New Zealand*

*<sup>b</sup>Department of Process Engineering, TH Nürnberg, Nürnberg, 90489, Germany*

*<sup>c</sup>Industrial Information and Control Centre, Department of Chemical & Materials Engineering, The University of Auckland, Auckland, 1010, New Zealand*  
[isuru.udugama@waikato.ac.nz](mailto:isuru.udugama@waikato.ac.nz)

## Abstract

With government policy encouraging decarbonisation, the New Zealand food processing industry is rethinking its process operations and their improvement. This work focuses on identifying opportunities to decarbonise industrial sites using simultaneous multi-process control and scheduling optimisation methods. A common issue for multi-process sites is the disconnection between the unit operation level process controls that operate under closed loop regulatory, or advanced regulatory, principles and the scheduling layer that ideally operates under real-time optimisation principles. As the first step towards solving this issue, an illustrative case of a cream cheese production process consisting of a pasteurisation unit, two fermentation units and a separation unit are studied by constructing a plant wide scheduling model in Python, integrated with a fermentation model, to explicitly model the interactions between two layers. Compared to typical plant operation, controlling the dosing of substrate in the fermenter allows for more uniform fermenter end times leading to improved scheduling of successive batches which results in less discarding of batches due to poor quality. Future work will embed the trade-off between overall energy usage, product quality and product throughput constraints (supply chain) within a holistic objective function centred on carbon footprint.

**Keywords:** process control, fermentation modelling, real-time decision making

## 1. Introduction

Food and dairy production processes account for over 40% of New Zealand export revenue but represent a significant proportion of New Zealand's natural gas and coal consumption [1]. Like many other countries [2], the New Zealand government has implemented the Emission Trading Scheme to encourage the process industries to decarbonise while supporting them to maintain their competitive edge through co-investment in low-emission technology. From an engineering perspective, tackling the decarbonisation challenge requires a mix of capital investments into new technologies, e.g., integrating heat pump technologies in supplying process heat for milk spray dryers [3], and real-time control optimisation of industrial site operations.

Food production processes are often operated batch-wise where a given batch is fed through multiple unit operations over time. As such, the control of these food production



processes requires both process control (carried out at the unit operation level) and scheduling of unit operation utilisation throughout multiple batches. However, the simultaneous multi-process control and scheduling optimisation methods are not an established practice in the process systems engineering domain. At present, a hierarchical process control approach is taken where plant-wide scheduling and unit operation level control are executed at different layers, i.e., sequentially. In this approach, instructions are passed down the process control hierarchy from the plant-wide scheduling layer to advance process control layer and onwards to regulator controls. The idea being that the process control layer will focus on ensuring aspects such as product quality while the scheduling layer focuses on process optimisation [4].

However, food processing has large inherent production variations due to availability of the natural raw materials (e.g., milking seasons for dairy processes) and changes in consumer demands. As a result the schedules need to be constantly updated, for example, using short term and continuously updating schedules [5]. Another option of tackling these variations is to change set-points in the underlying supervisory and regulatory level control structure in “real-time”. This entails speeding up or slowing down the unit operations such that the overall production schedule is not influenced. In practice, a combination of both “real-time” control actions and timely re-scheduling will be needed to ensure a production process maintain quality, production targets while actively ensuring decarbonisation goals are met.

The aim of this study is to investigate the decarbonisation potential of simultaneous optimisation of process control and multi-plant scheduling actions in the food industry. Figure 1 expresses the current industry practice in contrast to the proposed approach.

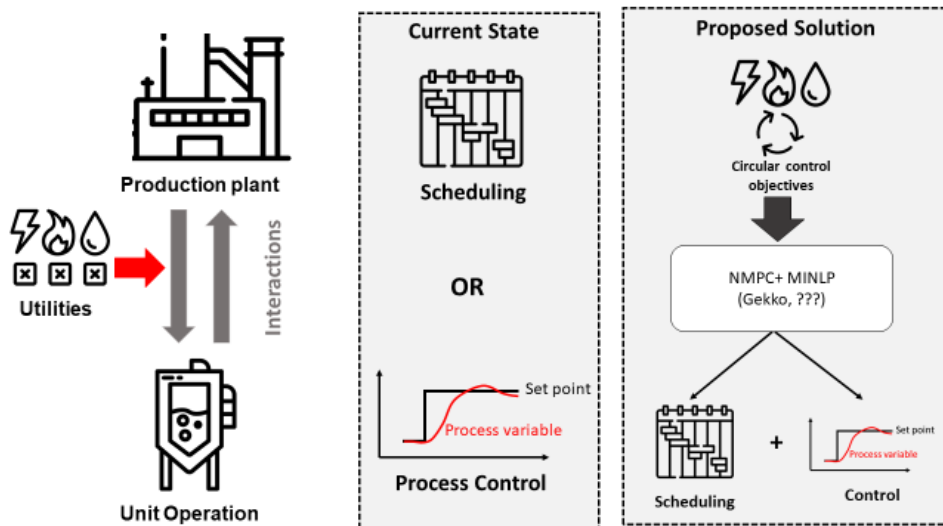


Figure 1: Visualisation of the challenge, current state and the proposed solution

Solvers such as active-set based APOPT have the potential to be a useful tool in bridging the gap between sequential and simultaneous methods for scheduling and control optimisation. For example, by translating advanced process control concepts such as model predictive control from differential algebraic equations to algebraic form allows

the simultaneous solution of the scheduling and the control problem as a mixed-integer non-linear optimisation problem [6]. As a realistic test case, this paper presents a fermentation-based process to illustrate the potential of simultaneous decision-making.

## **2. Model Development**

This section details the current state of the test case model of a simple cream cheese production process with a pasteurisation, fermentation and downstream processing unit operations. The test model is implemented in Python and solved using the GEKKO library with its APOPT solver.

### *2.1. Plant level model*

A sequential scheduling model is used to represent the dynamic dependency of unit operations at the plant level. The scheduling model consists of three-unit operations; pasteurisation, fermentation and separation. Two fermenters are used in parallel while only a single pasteuriser and separation unit are present. This resembles industrial plant design since the fermentation time is longer than the processing times of the other two units. The pasteurisation and separation processing times are assumed to be static while the fermentation processing time varies between batches. A fermentation model is used to determine fermentation time where model parameters are randomly varied to simulate variation within a real process. More detail on the model is provided in Section 3.2. The model also considers cleaning in place (CIP) procedures as well as transfers between unit operations. These procedures are also assumed to have static times.

The model is initially setup to replicate reality where the fermentation end time is not known at the beginning, which means successive operations can only be scheduled when the fermentation ends. This can then be compared with a scenario where the fermentation is perfectly controlled to finish within a target time, and the next batch can begin preparation in advance allowing for maximum production. The time series data was stored in data frames from the Pandas package which is then visualised in a Gantt chart using the Plotly package in Python.

### *2.2. Fermentation model*

The mechanistic fermentation model reported in [7] is used as the basis of modelling the fermentation process. The model tracks substrate (lactose), biomass (*Lactococcus lactis*) and product (lactic acid) evolution over time. The model utilises a Monod type biomass growth kinetics with substrate and product inhibition thus production rates are dependent on substrate concentration and availability. To represent inherent variabilities that are present in fermentation processes, the model parameters are randomly sampled from a uniform distribution within the reported upper and lower parameter bounds in [7]. The initial concentration of substrate, product and biomass were assumed constant throughout batches to represent feedstock standardisation. To this end, all variation in fermentation concentration profiles between batches is due to random variations in model parameters.

The fermentation model is implemented into Python as a series of ordinary differential equations (ODE) representing time dependent evolution of substrate (S), biomass (B) and product (P). In the case of no fermentation control the model is solved simultaneously with the Scipy numerical integration package across a series of discretised time points. Integration is performed until the product concentration reaches 5 g/L which represents experimental observations in [7] that ensures the target product quality is achieved.

To allow for fermentation end time control the model must be formulated as a dynamic non-linear programming (NLP) that is solved in Python GEKKO using the IPOPT solver [6]. The time dependent concentration growth rates of the species are governed by Equations 1-3 where the substrate dosing flowrate,  $m_s$ , has been included in Equation 2 as a manipulated variable. However, the flowrate must be divided by the fermenter volume to convert mass flow to change in concentration per unit time. A terminal constraint is applied in Equation 4 which enforces that the product concentration at the target final time,  $t_f$ , is equal to the target concentration,  $P_{target}$ . The model is solved across a consecutive series of time points ranging from 0 to  $t_f$ . Lastly the objective function is applied to maximize the final product concentration pursuant to the terminal constraint.

$$\frac{dX}{dt} = \mu_{max} \left( \frac{S}{K_{sx} + S} \right) \left( 1 - \frac{P - P_{ix}}{P_{mx} - P_{ix}} \right) \left( \frac{K_{ix}}{K_{ix} + S} \right) X \quad 1$$

$$\frac{dS}{dt} = -q_{s,max} \left( \frac{S}{K_{ss} + S} \right) \left( 1 - \frac{P - P_{is}}{P_{ms} - P_{is}} \right) \left( \frac{K_{is}}{K_{is} + S} \right) + \frac{m_s}{V} \quad 2$$

$$\frac{dP}{dt} = \alpha \frac{dX}{dt} + q_{p,max} \left( \frac{S}{K_{sp} + S} \right) \left( 1 - \frac{P - P_{ip}}{P_{mp} - P_{ip}} \right) \left( \frac{K_{ip}}{K_{ip} + S} \right) \quad 3$$

$$P \times t_{final} - P_{target} = 0 \quad 4$$

$$t = [0, t_{final}] \quad 5$$

$$\max\{P \times t_{final}\} \quad 6$$

### 2.3. Model integration and decarbonisation

The scheduling and fermentation model are integrated such that simultaneous decisions can be performed between changing manipulated variables in the fermentation process, and schedule changes at the plant wide level. More specifically, the fermentation end time applied to the scheduling model is dynamically updated from the fermentation model which is influenced by control actions. Regarding decarbonisation, the combined modelling of the schedule and fermentation enables the calculation of energy intensive actions such as CIP and material flows in unit operations. In future work, further investigation of explicit minimization of CO<sub>2</sub>-e will be investigated. Explicitly considering decarbonisation allows identification of effective actions both in the scheduling layer and unit operation layer (e.g., mitigating scheduling actions that may require batch disposal or situations where CIP must be redone).

## 3. Model output

Figure 2 illustrates the process schedule for the two use cases and Figure 3 shows the concentrations of species within each fermenter batch. In both use cases 10 fermentation batches are processed. Figure 2a illustrates the progression of production batches through the cream cheese production process with no fermentation control and Figure 3a shows the corresponding fermentation concentrations.

Each batch is first heat treated in the pasteuriser prior to transfer to one of the two fermenters, then the batch is sent for separation. The inherent variations in the fermentation time clearly influence the schedule as the next batch cannot be pasteurised until a fermenter is available, nor separated until the separator is available. For example, in batch number 3 (second batch on fermenter 1), the fermentation process has finished

early but cannot be transferred out for another two hours until the separator is available. This can result in product quality degradation leading to batch disposal. In contrast, Figure 2b shows the production schedule where the fermentation end time is perfectly controlled by substrate dosing to finish (i.e. reach the target lactic acid concentration) at 6 hours. The concentration profile in each batch is shown in Figure 3b where the substrate concentration changes considerably between batches due to control induced dosing.

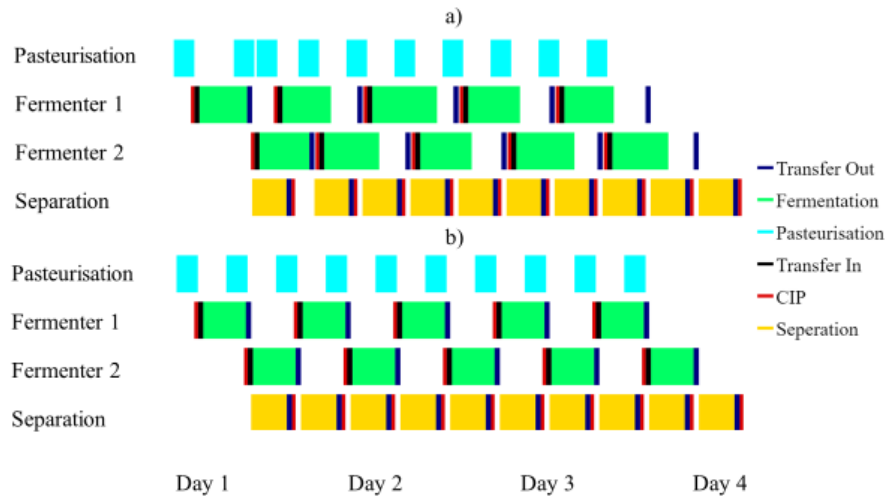


Figure 2: a) Current schedule of cream cheese process where fermentation end times are not controlled b) Improved schedule where the fermentation is controlled to finish at a specified end time.

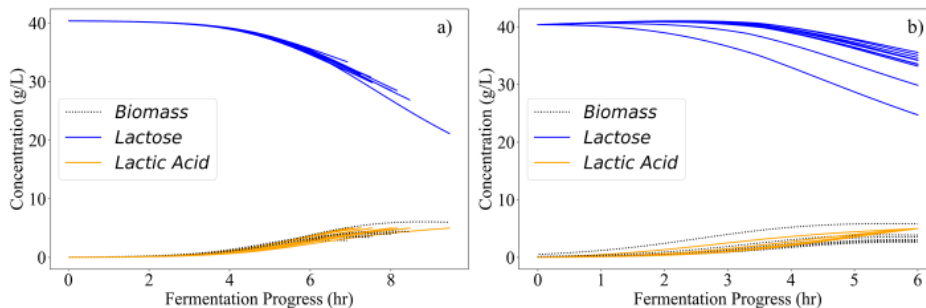


Figure 3: a) Concentration profiles in a cream cheese fermenter with random variation and no control b) Concentration profiles in a fermenter with random variation but controlled to achieve the target lactic acid concentration at 6 hours.

Fermentation control leads to improved scheduling since fermentation end times are known so successive batches can be scheduled so that fermentation ends as soon as a separator is available. As such, batches can immediately be transferred for separation resulting in no potential loss of product. It should be noted that in the improved case, the fermentation operation is changed from a batch to fed-batch.

Despite likely improvements in the form of reductions in batch disposal, the optimized schedule finishes the 10 batches the same time as the un-optimized schedule due to the separator unit availability. However, tightly controlling the fermentation end time means extra substrate dosing is needed. From a decarbonisation perspective, this optimisation

does not consider the carbon trade-off from substrate dosing (which increases the CO<sub>2</sub>-e of a given batch) that may be needed to save the batch, and the cost of losing the batch which has implications on successive batch scheduling. As an immediate future step, the simultaneous optimization of batch scheduling and fermentation substrate dosing must be considered. For example, re-looking at batch number 3 in Figure 2a (un-optimized case), a fed-batch fermentation could reduce the initial dosage of substrate thereby reducing the biomass growth rate and lactic acid production rate. As such, the end time of the fermentation will be prolonged which reduces the waiting time for separation without product degradation. In this example, the reduced substrate dosing will not only improve the overall CO<sub>2</sub>-e of the specific batch, but also reduce the overall CO<sub>2</sub>-e of a unit of finished product.

#### 4. Conclusions and Future work

This work presented a cream cheese fermentation process where a plant wide scheduling model was integrated with a fermentation model to illustrate the propagation of variations in fermentation times at unit operation level and scheduling level. This work is the first step towards solving the key issue involving the disconnection between the operation level process controls to maintain product quality and the scheduling layer that attempts to maximise product throughput. It was seen that unit operation level decisions such as dynamic dosing of substrate can reduce the variation experienced at the scheduling layer thereby allowing for improved scheduling of successive operations. In future work, the test case will be used to develop a simultaneous multi-process control and scheduling optimisation methods that explicitly consider the CO<sub>2</sub>-e trade-off for all decisions.

#### References

- [1] I. & E. MBIE NZ (Ministry of Business, “Dairy Manufacturing – Process Heat and Greenhouse Gas Emissions.”
- [2] B. K. Sovacool, M. Bazilian, S. Griffiths, J. Kim, A. Foley, and D. Rooney, “Decarbonizing the food and beverages industry: A critical and systematic review of developments, sociotechnical systems and policy options,” *Renew. Sustain. Energy Rev.*, vol. 143, p. 110856, Jun. 2021, doi: 10.1016/j.rser.2021.110856.
- [3] Q. Chen, D. J. Cleland, J. K. Carson, and T. G. Walmsley, “Integration of desiccant wheels and high-temperature heat pumps with milk spray dryers,” *Appl. Therm. Eng.*, vol. 216, p. 119083, Nov. 2022, doi: 10.1016/j.applthermaleng.2022.119083.
- [4] I. A. Udugama et al., “The Role of Big Data in Industrial (Bio)chemical Process Operations,” *Ind. Eng. Chem. Res.*, vol. 59, no. 34, pp. 15283–15297, Aug. 2020, doi: 10.1021/acs.iecr.0c01872.
- [5] M. G. Ierapetritou and C. A. Floudas, “Effective Continuous-Time Formulation for Short-Term Scheduling. 1. Multipurpose Batch Processes,” *Ind. Eng. Chem. Res.*, vol. 37, no. 11, pp. 4341–4359, Nov. 1998, doi: 10.1021/ie970927g.
- [6] L. Beal, D. Hill, R. Martin, and J. Hedengren, “GEKKO Optimization Suite,” *Processes*, vol. 6, no. 8, p. 106, Jul. 2018, doi: 10.3390/pr6080106.
- [7] B. Li, Y. Lin, W. Yu, D. I. Wilson, and B. R. Young, “Application of mechanistic modelling and machine learning for cream cheese fermentation pH prediction,” *J. Chem. Technol. Biotechnol.*, vol. 96, no. 1, pp. 125–133, 2021, doi: 10.1002/jctb.6517.

# Production of $\text{Ca}(\text{HCO}_2)_2$ from formate solution through $\text{CO}_2$ hydrogenation

Taeksang Yoon,<sup>a</sup> Hayoung Yoon,<sup>b</sup> Sungho Yoon,<sup>b</sup> Chul-Jin Lee<sup>a,c</sup>

<sup>a</sup>*School of Chemical Engineering and Material Science, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul, Republic of Korea*

<sup>b</sup>*Department of Chemistry, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul, Republic of Korea*

<sup>c</sup>*Department of Intelligent Energy and Industry, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul, Republic of Korea*

## Abstract

With growing levels of  $\text{CO}_2$  emission and severity of global warming, interest in carbon capture and utilization (CCU) has increased. However, the amount of  $\text{CO}_2$  reduction that can be realized is lower than that of  $\text{CO}_2$  emissions, and most of the existing CCU technologies cannot be commercialized owing to their limited economic feasibility. Therefore, it is necessary to identify new chemicals as CCU products. When  $\text{CaO}$  is added to a formate aqueous solution, it reacts with the formate to produce calcium formate ( $\text{Ca}(\text{HCO}_2)_2$ ). In this study,  $\text{Ca}(\text{HCO}_2)_2$  was synthesized using dolomite and  $\text{CaO}$  ash, an industrial waste. Process design and evaluation were performed based on the experimental data. When  $\text{CaO}$  ash and dolomite were used as the raw material, the production costs were 16% lower and 8% higher than those of the conventional process, respectively. Moreover, for both materials, the  $\text{CO}_2$  emissions were 20% lower than those of the conventional processes.

**Keywords:**  $\text{CO}_2$  hydrogenation, Calcium formate, Carbon capture and utilization

## 1. Introduction

Carbon capture and utilization (CCU) technologies, which convert  $\text{CO}_2$  to value-added chemicals, have emerged as promising solutions to address the growing severity of global warming related issues (Aresta and Dibenedetto 2007; Peters et al. 2011). However, most of the existing CCU processes have not been commercialized because of their low economic feasibility compared with the conventional process (Chauvy et al. 2019; Lindsey and Jeskey 1957). Furthermore, the amount of  $\text{CO}_2$  that can be eliminated through CCU is lower than that of the  $\text{CO}_2$  emissions. Therefore, novel CCU chemical products must be identified (Rafiee et al. 2018).

Recently, Park et al. (2020) developed a method for producing formic acid by hydrogenating  $\text{CO}_2$  through a heterogeneous Ru catalyst and an aqueous amine solution. This method can overcome the disadvantage of heterogeneous catalysts, i.e., a low conversion rate. However, the process involves multiple distillation processes in separation (Figure 1(a)).

Calcium formate,  $\text{Ca}(\text{HCO}_2)_2$ , is used in various applications such as animal feed, cement production, and leather tanning as an additive (EFSA 2015).  $\text{Ca}(\text{HCO}_2)_2$  is typically produced by the carbonylation of calcium hydroxide with carbon monoxide (Ma et al.

2016). Recently, Yoon et al. (2022) reported the production of  $\text{Ca}(\text{HCO}_2)_2$  using  $\text{CaO}$  ash (waste from a desulfurization process) and  $\text{CO}_2$  through hydrogenation. This method can instantaneously convert  $\text{CaO}$  ash, an industrial waste, to  $\text{Ca}(\text{HCO}_2)_2$  and separate products through simple evaporation in mild conditions. Moreover, the main raw material,  $\text{CaO}$ , is readily produced through the calcination of dolomite, an abundant natural resource.

In this study,  $\text{Ca}(\text{HCO}_2)_2$  was synthesized through  $\text{CO}_2$  hydrogenation, with process design performed based on experimental data (Figure 1(b)). The feasibility of the process was confirmed through a techno-economic analysis (TEA) and lifecycle assessment (LCA), in comparison with the conventional process (Ma et al. 2016).

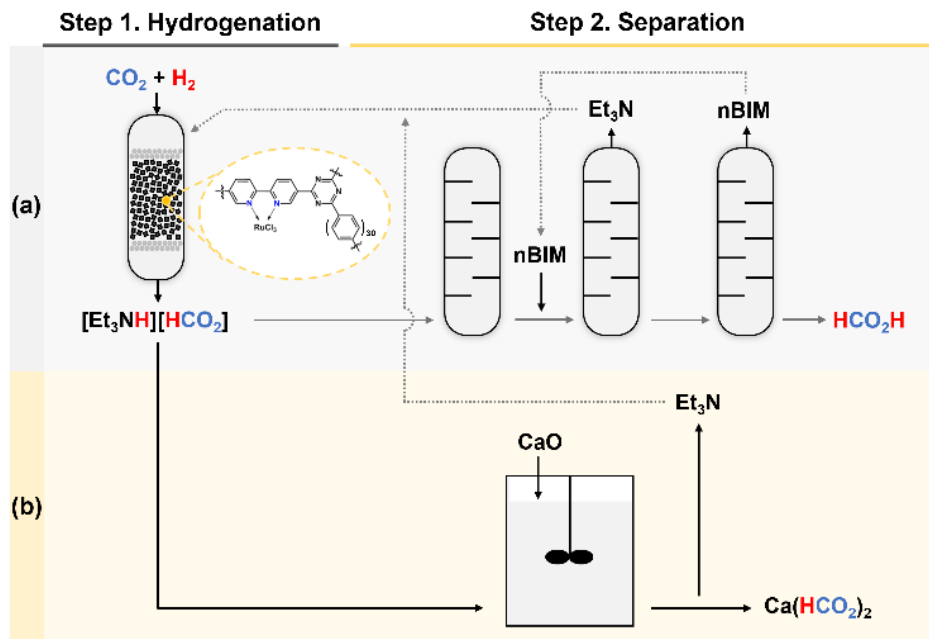
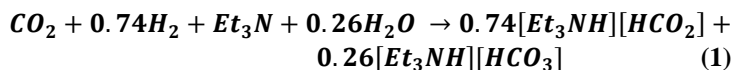


Figure 1 Schematic of (a) existing and (b) proposed (Yoon et al. 2022) processes for  $\text{HCO}_2\text{H}$  and  $\text{Ca}(\text{HCO}_2)_2$  production

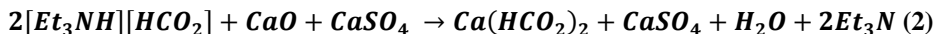
## 2. Experimental

### 2.1. Hydrogenation of $\text{CO}_2$ to $[\text{Et}_3\text{NH}][\text{HCO}_2]$



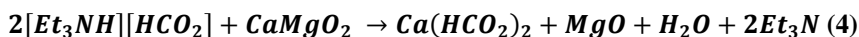
$\text{CO}_2$  hydrogenation was performed in a stainless-steel tubular reactor packed with 1.5 g of Ru/bpyTN-30-CTF catalyst. The reactor was pressurized to 120 bar and heated to 120 °C before the reaction.  $\text{H}_2\text{O}$  and  $\text{Et}_3\text{N}$  were supplied to the reactor, and an  $\text{Et}_3\text{N}$  concentration of 3 M was maintained using a high-pressure liquid pump. The liquid product was analyzed to determine the formate concentration. The results indicated that the solution contained 2.03 M  $[\text{Et}_3\text{NH}][\text{HCO}_2]$  and 0.73 M  $[\text{Et}_3\text{NH}][\text{HCO}_3]$  (Equation (1)). The solution was heated at 90 °C and ambient pressure for 2 h to remove  $\text{CO}_2$  that remained as the  $\text{HCO}_3$  adduct and could react with  $\text{CaO}$  to produce  $\text{CaCO}_3$ .

## 2.2. Synthesis of $\text{Ca}(\text{HCO}_2)_2$ using CaO ash



Flue gas desulfurization ash (CaO/CaSO<sub>4</sub> ash) can be generated as a product of the desulfurization process using CaO. Because the solubility of CaSO<sub>4</sub> (0.2 g per 100 mL H<sub>2</sub>O at 20 °C) is considerably lower than that of Ca(HCO<sub>2</sub>)<sub>2</sub> (16.6 g per 100 mL H<sub>2</sub>O at 20 °C), CaO and CaSO<sub>4</sub> are expected to be separated in a facile manner. When CaO/CaSO<sub>4</sub> ash was added to the [Et<sub>3</sub>NH][HCO<sub>2</sub>] solution at 90 °C for 5 min, a clear solution with a white precipitate, CaSO<sub>4</sub> (Equation (2)), was obtained. CaSO<sub>4</sub> and Ca(HCO<sub>2</sub>)<sub>2</sub> were separated from the solution through filtration and evaporation, respectively.

## 2.3. Synthesis of $\text{Ca}(\text{HCO}_2)_2$ using dolomite



First, dolomite was calcined in a furnace at 950 °C and ambient pressure for decomposition to CaMgO<sub>2</sub> (Equation (3)). Results of an X-ray diffraction analysis indicated that all of the dolomite was completely decomposed to CaO and MgO. Allen et al (2009) highlighted that reaction kinetics between Ca and Mg ions are considerably different. Therefore, when calcined dolomite was added to the [Et<sub>3</sub>NH][HCO<sub>2</sub>] solution at 90 °C, and a reaction was allowed to occur for 5 min, only CaO was expected to react with formate and produce Ca(HCO<sub>2</sub>)<sub>2</sub> while MgO remained as the oxide (Equation (4)). Because MgO is sparingly soluble in water, it remained as a precipitate in the solution, which could be filtered to yield a high purity product. This filtrate could be separated into Ca(HCO<sub>2</sub>)<sub>2</sub> and solvents by simple evaporation, as described in Section 2.2.

# 3. Process design and analysis

## 3.1. Process design

Based on the experimental results, two Ca(HCO<sub>2</sub>)<sub>2</sub> processes were designed using CaO ash and dolomite. Figure 2 illustrates the process flow. Gaseous CO<sub>2</sub> and H<sub>2</sub> were compressed and heated to 120 °C at 120 bar before being supplied to the hydrogenation reactor. The unreacted gases were separated through degassing and recycled. The degassed solution was introduced to the reactor with a CaO mixture, thereby synthesizing Ca(HCO<sub>2</sub>)<sub>2</sub>. Because the reaction kinetics and solubility of CaO and MgO were considerably different, only CaO transformed to Ca(HCO<sub>2</sub>)<sub>2</sub>, and MgO remained as solid in the solution. The solid component, as a byproduct with high purity, was separated by filtering. The remaining solution was separated into the solvent and Ca(HCO<sub>2</sub>)<sub>2</sub> through evaporation.

## 3.2. TEA

The economic feasibility of the process was evaluated in terms of the production cost. Table 1 lists the prices of the raw materials and utilities, and Table 2 summarizes the results of the economic analysis. When dolomite and CaO ash were used as the raw materials, the production cost was approximately 8% higher and 16% lower than that of the conventional process, respectively. Figure 3(a) shows the breakdown of the production cost for each process component: raw material, utility, pretreating and reaction, and separation. The cost components are defined in the following text.



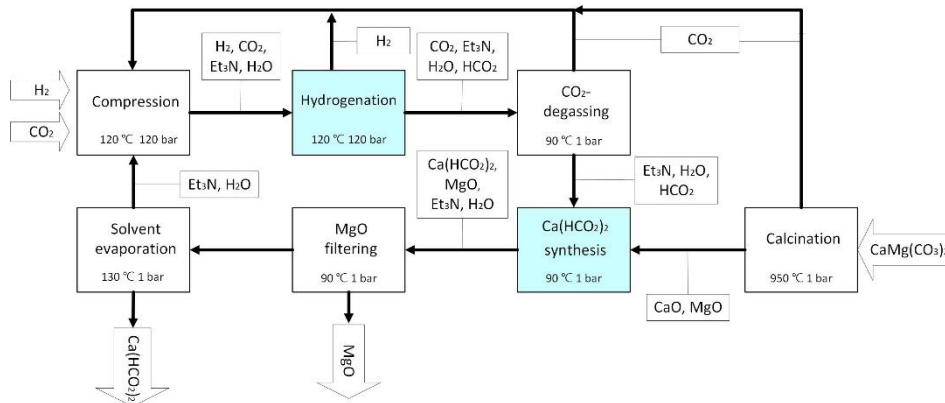


Figure 2 Block flow diagram of  $\text{Ca}(\text{HCO}_2)_2$  production using dolomite.

- Raw material: Summation of all costs associated with the raw materials.
- Utility costs: Summation of all costs for utilities.
- Pretreating and reaction: Summation of capital expenditures (CAPEX) and fixed operational expenditure (fixed OPEX) for the pretreatment of material (e.g., compression and heating) and reaction.
- Separation: Summation of CAPEX and fixed OPEX for separation equipment/processes such as a column.

The ratio of utility costs was higher than that of the raw material costs, which indicated that  $\text{Ca}(\text{HCO}_2)_2$  production is energy-intensive. The high energy consumption is attributable to the evaporation of water (Yoon et al. 2022).

Table 1 Material and utility costs for the economic analysis

Component	Price (unit)	Reference
$\text{CO}_2$	35 (USD/tonne)	(Kim et al. 2011)
$\text{H}_2$	1630 (USD/tonne)	(Lee et al. 2018)
LP steam	13.28 (USD/GJ)	(Turton and Bailie 2012)
Electricity	0.066 (USD/kWh)	(Turton and Bailie 2012)
Cooling water	0.354 (USD/GJ)	(Turton and Bailie 2012)
Natural gas	11.34 (USD/GJ)	(Turton and Bailie 2012)

### 3.3. LCA

LCA was performed to evaluate the  $\text{CO}_2$  emissions resulting from the processes. The Intergovernmental Panel on Climate Change 2013 GWP 100a method was applied for the analysis. Data from Ecoinvent 3.8 database and the literature were used to calculate the influence of the product and  $\text{CO}_2$  emission factors (Lee et al. 2018; Wernet et al. 2016). The functional unit of analysis was 1 kg of  $\text{Ca}(\text{HCO}_2)_2$ . Because the processes associated with CCU are multi-functional and yield more than one product, the  $\text{CO}_2$  emission of the byproduct was excluded from the total  $\text{CO}_2$  emission. Table 2 summarizes the LCA results. The  $\text{CO}_2$  emissions associated with both the proposed processes were 20% lower than those of the conventional process. Figure 3(b) shows the breakdown of the cost and  $\text{CO}_2$  emissions. Similar to the cost, the  $\text{CO}_2$  emissions corresponded to the following four components:

- Raw material: Summation of  $\text{CO}_2$  emissions from raw materials.
- Utility: Summation of  $\text{CO}_2$  emissions for utilities consumed in the process.

- Purge: Summation of  $\text{CO}_2$  emission from the purge gas.
- Byproduct: Summation of  $\text{CO}_2$  emissions from the byproducts, excluded from the total  $\text{CO}_2$  emission.

For all processes, the contribution of the utility to the total  $\text{CO}_2$  emission was the most notable (~50% of the total value), and it accounted for more than 80% of the total  $\text{CO}_2$  emissions in the CCU processes.

Table 2 Results of economic and lifecycle analysis of each process.

	Conventional	Proposed (dolomite)	Proposed (CaO ash)
Production cost (USD/tonne)	678	729	567
$\text{CO}_2$ emission ( $\text{kgCO}_{2\text{eq}}/\text{kgCa}(\text{HCO}_2)_2$ )	3.97	3.18	3.19

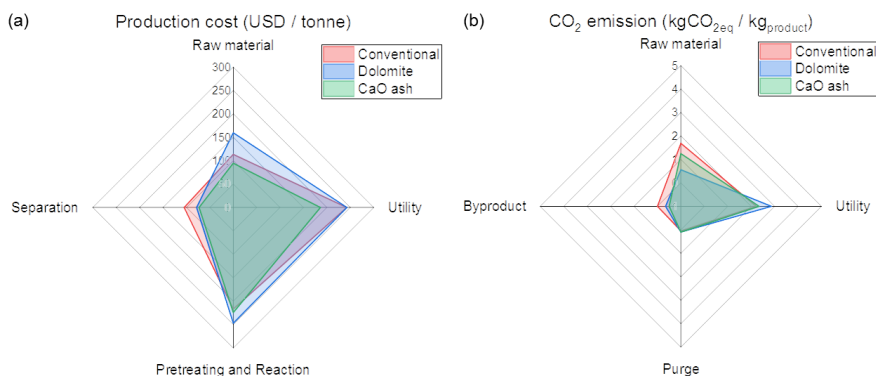


Figure 3 Breakdown of (a) production cost and (b)  $\text{CO}_2$  emission for  $\text{Ca}(\text{HCO}_2)_2$  production.

#### 4. Conclusions

This paper proposes two processes for  $\text{Ca}(\text{HCO}_2)_2$  production based on CCU, using CaO and formate adduct. The novel processes were compared with the conventional process. The proposed processes could selectively produce  $\text{Ca}(\text{HCO}_2)_2$  and byproducts with high purity while utilizing  $\text{CO}_2$  as a value-added chemical. Although the production cost of the CCU process using dolomite was 8% higher than that of the conventional process, it emitted 20% lower  $\text{CO}_2$ , demonstrating environmental feasibility. When CaO ash was used as the CaO source, the economic and environmental feasibility were comparable to those of the conventional approach. The results highlight the potential of  $\text{Ca}(\text{HCO}_2)_2$  production based on CCU method as a replacement for conventional processes. The energy consumption can be reduced to enhance the economic feasibility of the CCU-based processes.

#### References

EFSA Panel on Additives and Products or Substances used in Animal Feed (2015). "Scientific opinion on the safety and efficacy of calcium formate when used as a technological additive for all animal species." *EFSA Journal* **13**(5): 4113.

- Allen, J. P., S. C. Parker and D. W. Price (2009). "Atomistic simulation of the surface carbonation of calcium and magnesium oxide surfaces." The Journal of Physical Chemistry C **113**(19): 8320-8328.
- Aresta, M. and A. Dibenedetto (2007). "Utilisation of CO<sub>2</sub> as a chemical feedstock: opportunities and challenges." Dalton Transactions (28): 2975-2992.
- Chauvy, R., N. Meunier, D. Thomas and G. De Weireld (2019). "Selecting emerging CO<sub>2</sub> utilization products for short- to mid-term deployment." Applied Energy **236**: 662-680.
- Kim, J., C. A. Henao, T. A. Johnson, D. E. Dedrick, J. E. Miller, E. B. Stechel and C. T. Maravelias (2011). "Methanol production from CO<sub>2</sub> using solar-thermal energy: process development and techno-economic analysis." Energy & Environmental Science **4**(9): 3122-3132.
- Lee, D.-Y., A. Elgowainy and Q. Dai (2018). "Life cycle greenhouse gas emissions of hydrogen fuel production from chlor-alkali processes in the United States." Applied Energy **217**: 467-479.
- Lindsey, A. S. and H. Jeskey (1957). "The Kolbe-Schmitt reaction." Chemical Reviews **57**(4): 583-620.
- Ma, H., X. Feng, Y. Yang, Z. Zhang and C. Deng (2016). "Preparation of feed grade calcium formate from calcium carbide residue." Clean Technologies and Environmental Policy **18**(6): 1905-1915.
- Park, K., G. H. Gunasekar, S.-H. Kim, H. Park, S. Kim, K. Park, K.-D. Jung and S. Yoon (2020). "CO<sub>2</sub> hydrogenation to formic acid over heterogenized ruthenium catalysts using a fixed bed reactor with separation units." Green Chemistry **22**(5): 1639-1649.
- Peters, M., B. Köhler, W. Kuckshinrichs, W. Leitner, P. Markewitz and T. E. Müller (2011). "Chemical technologies for exploiting and recycling carbon dioxide into the value chain." **4**(9): 1216-1240.
- Rafiee, A., K. Rajab Khalilpour, D. Milani and M. Panahi (2018). "Trends in CO<sub>2</sub> conversion and utilization: A review from process systems perspective." Journal of Environmental Chemical Engineering **6**(5): 5771-5794.
- Turton, R. and R. C. Bailie (2012). Analysis, synthesis, and design of chemical processes. Upper Saddle River, NJ, Prentice Hall.
- Wernet, G., C. Bauer, B. Steubing, J. Reinhard, E. Moreno-Ruiz and B. Weidema (2016). "The ecoinvent database version 3 (part I): overview and methodology." The International Journal of Life Cycle Assessment **21**(9): 1218-1230.
- Yoon, H., T. Yoon, H.-J. Yoon, C.-J. Lee and S. Yoon (2022). "Eco-friendly and techno-economic conversion of CO<sub>2</sub> into calcium formate, a valuable resource." Green Chemistry **24**(4): 1738-1745.

# Techno-Economic Process Analysis of the Chemical Recycling of Nylon 6 Using Phosphoric Acid

Ann-Joelle Minor<sup>a,\*</sup>, Ruben Goldhahn<sup>a</sup>, Liisa Rihko-Struckmann<sup>a</sup>,  
Kai Sundmacher<sup>a,b</sup>

<sup>a</sup>*Max Planck Institute for Dynamics of Complex Technical Systems, Process Systems Engineering, Sandstorstr. 1, 39106 Magdeburg, Germany*

<sup>b</sup>*Chair for Process Systems Engineering, Otto-von-Guericke University, Universitätsplatz 2, 39106 Magdeburg, Germany*

\*E-mail address: [aminor@mpi-magdeburg.mpg.de](mailto:aminor@mpi-magdeburg.mpg.de)

## Abstract

Chemical recycling has increasingly gained attention in order to counteract environmental pollution and support the transition to a circular economy. Today, only around two percent of polyamides are chemically recycled each year. This is mainly due to the economic performance of the process. The conventional depolymerisation process of polyamide 6 using phosphoric acid was simulated to analyse the techno-economic potential and identify its existing constraints. A conceptual process design was carried out using Aspen Plus and Python. The results show that the process becomes profitable at 4600 t/year of caprolactam. Assuming a reasonable plant capacity of 12.6 kt/year, the payback time would be between three and seven years in the best- and worst-case scenarios. The disposal of the waste stream containing unreacted polyamide 6, oligomers and phosphoric acid, as well as phosphoric acid raw material costs together contribute 47% of the total production cost. Accordingly, a sensitivity analysis showed that the total annualised costs depend significantly on the amount of phosphoric acid raw material used. Overall, the results indicate great industrial potential for the depolymerisation of polyamide 6. Still, the focus of further research on this conventional process should rely on phosphoric acid and its efficient recovery along with the treatment of the waste stream.

**Keywords:** chemical recycling, circular economy, Nylon 6, process simulation, techno-economic analysis.

## 1. Introduction

The massive production of plastics caused undeniable issues such as environmental pollution and depletion of petroleum-based resources. To meet the recycling target of 50% set by the European Commission for 2025, chemical recycling has increasingly gained attention (Pohjakallio et al., 2020). As opposed to the most popular mechanical recycling strategy, chemical recycling avoids quality loss of the polymer (also called downcycling), thereby acting in favour of a circular economy with multiple polymerisation–depolymerisation cycles. Chemical recycling involves breaking down the polymer chain into monomers through a high-temperature and solvolytic environment (Coates and Getzler, 2020).

According to a recent review, polyamide 6 (PA6), also known as Nylon 6, is an ideal candidate for chemical recycling as it is not thermodynamically and kinetically limited (Coates and Getzler, 2020). Despite having been investigated since the early

1960s, still today, only a minor fraction of the annual PA6 production is recycled each year (Coates and Getzler, 2020). Reportedly, the reasons for this are the low profitability and economic competitiveness of depolymerisation processes (Kamimura et al., 2014). Several companies have been employing the chemical recycling of PA6. The acid-catalysed hydrolysis of PA6 to its respective monomer caprolactam (CL) using phosphoric acid ( $\text{H}_3\text{PO}_4$ ) and steam can be considered as a conventional process which is covered by several patents, e.g. of BASF (Corbin et al., 1999). Furthermore, acid-catalysed hydrolysis of polyamides has been investigated experimentally (Chen et al., 2010). However, the cost contributions, bottlenecks, most sensitive parameters, and the reason for lacking economic performance of the process have not been identified yet. In order to improve the understanding and find possible opportunities for optimisation of the process, the chemical recycling process of PA6 using  $\text{H}_3\text{PO}_4$  was simulated, and a techno-economic analysis was performed, both of which have never been published in the literature. Subsequent to the next chapter giving the background information and methods, in the results section, the scale for profitability, the payback period, as well as the leading fixed and variable cost contributions are estimated. Additionally, the effect of the most influential parameters is studied within a sensitivity analysis and case study.

## 2. Background and Methodology

In the following, the continuous conventional process described by BASF will be referred to as the “representative process” (Corbin et al., 1999). This representative process was simulated and analysed with a model connecting Aspen Plus and Python. To study it in the most realistic manner, the changes to this process were kept at a minimum.

### 2.1. Polyamides and Depolymerisation Background Knowledge

PA6 is most commonly synthesised through hydrolytic ring-opening polymerization, involving reversible reactions, e.g. the hydrolysis of CL to aminocaproic acid (ACA) and polycondensation and addition reactions (Penczek et al., 1985). However, due to dipole-dipole and hydrogen bonding interactions of the amide groups, required depolymerisation temperatures are really high. The ceiling temperature of 277 °C can be lowered by adding a solvent such as  $\text{H}_3\text{PO}_4$  in the conventional process (Coates and Getzler, 2020).

### 2.2. Thermodynamics and Kinetics

Vapour-liquid equilibrium (VLE) data of water and CL are retrieved from the NIST database by AspenPlus, and the NRTL model was used to describe the VLE behaviour.  $\text{H}_3\text{PO}_4$  and PA6 were assumed to be non-volatile.

Owing to the excess of water and catalyst, the literature approximates the hydrolytic depolymerisation of PA6 to CL by a pseudo-first-order reaction in PA6 (Chen et al., 2010). This is displayed in Eq. 1, where  $[\text{PA6}]$ ,  $k'$  and  $X$  are the PA6 concentration, the rate constant and conversion, respectively.

$$-\frac{d[\text{PA6}]}{dt} = k' \cdot [\text{PA6}] \Leftrightarrow dt = \frac{dx}{k' \cdot (1-X)} \quad \text{with } [\text{PA6}] = [\text{PA6}]_0 \cdot (1-X) \quad (1)$$

The reaction kinetics for chemical recycling using  $\text{H}_3\text{PO}_4$  are not available. Therefore, the required kinetic parameter is derived from Eq. 1 using the batch data point with the final time and conversion given by the representative process (Eq. 2).

$$k' = \frac{1}{t_f} \cdot \ln\left(\frac{1}{1-X_f}\right) \quad (2)$$

### 2.3. Data Base for Component Input Into Aspen Plus

A database needed to be established for the component input of PA6. Scalar properties, such as heat of formation and degree of polymerisation, as well as temperature-dependent

properties, such as solid and liquid heat capacity values, were taken from the Polymer Handbook, and parameters were regressed when necessary (Brandrup et al., 2003). As reaction enthalpy, the negative of the standard polymerisation enthalpy (-12.5 kJ/mol) was assumed (Brandrup et al., 2003).

#### 2.4. Process and Equipment Design

The process is designed to be as similar as possible to the representative process. The rate-based continuous stirred tank reactor (CSTR) model was chosen because of the liquid-phase reaction and slow reaction kinetics. The type of pumps, heat exchangers, vacuum equipment, columns or vessels was determined by heuristics. For the economic evaluation, equipment dimensions such as the vessel height, diameter, wall thickness or weight need to be known. The equipment design was done by a combination of empirical equations, e.g. the determination of heat exchanger areas or column diameters, and heuristics such as tray sizing or the choice of the heat exchanger tube length (Seider et al., 2009).

#### 2.5. Utilities

Cooling water (CW), low-pressure steam (LP), medium-pressure steam (MP), high-pressure steam (HP) and No. 2 fuel oil (OIL) were chosen as utilities. The temperature (T), enthalpy of evaporation ( $h_e$ ), or higher heating value (HHV) of those are displayed in Table 1.

Table 1: Details of the chosen utilities.

CW	LP	MP	HP	OIL
$T_{in} : 30^{\circ}\text{C}$	$T : 139.9^{\circ}\text{C}$	$T : 186^{\circ}\text{C}$	$T : 250^{\circ}\text{C}$	HHV :
$T_{out} : 45^{\circ}\text{C}$	$h_e : 2144 \text{ kJ/kg}$	$h_e : 1991 \text{ kJ/kg}$	$h_e : 1715.2 \text{ kJ/kg}$	43046.4 kJ/kg

Near-optimal approach temperatures were defined according to temperature levels, e.g. 5 °C for ambient temperature levels. The pipeline and valve pressure drop was neglected, and the heat exchanger pressure drop was set dependent on the phase of the stream.

#### 2.6. Economics

The capital investment expenditures (CAPEX), total operational expenditures (OPEX) and cash flow were calculated in U.S. dollars. The CAPEX was computed using the Lang factor method with an accuracy of 35%. The Lang factor  $F_L$  is multiplied by the sum of the free onboard (f.o.b.) equipment purchase costs  $C_{pi}$  and accounts for all other direct and indirect investment costs (Eq. 3).

$$\text{CAPEX} = 1.05 \cdot F_L \cdot \sum_i \frac{I_i}{I_{bi}} \cdot C_{pi} \quad (3)$$

The purchase costs were calculated using regressed correlations as a function of equipment size factors and updated to the current year by applying the chemical engineering plant cost index I (Seider et al., 2009).

The OPEX comprise variable costs, manufacturing costs, and general and administrative (G&A) costs. The variable costs, such as raw materials, utility and waste treatment costs, were determined using current industrial market prices and a plant operating factor of 0.9 was assumed. Direct wages and benefits were calculated by accounting for the typical labour requirements dependent on the type of equipment and assuming five shifts (J. Harmsen et al., 2018). The other manufacturing costs, such as maintenance, overall plant overhead and fixed costs, as well as the G&A costs, are typical fractions of the labour costs, capital investment expenses or sales (Seider et al., 2009).

For the cash flow analysis, net earnings have been calculated with the sales (S) and the income tax rate according to the U.S. federal income tax rate schedule for corporations. Subsequently, the annual cash flow was calculated by incorporating the depreciation (D) and assuming one year for the plant construction (Eq. 4 and 5) (Seider et al., 2009).

$$\text{Net earnings} = (S - \text{OPEX}) \cdot (1 - \text{income tax rate}) \quad (4)$$

$$\text{Annual Cash Flow} = (\text{Net earnings} + D) - \text{CAPEX} \quad (5)$$

The total annualised costs (TAC) are shown in Eq. 6 and include a reasonable return in investment  $i_{\min}$  of 0.2 (Seider et al., 2009).

$$\text{TAC} = \text{OPEX} + i_{\min} \cdot \text{CAPEX} \quad (6)$$

### 3. Results and Discussion

#### 3.1. Choice of Separation System and Process Flowsheet

In contrast to the representative process, pure PA6 enters the depolymerisation reactor. Hence, the required separation steps after depolymerisation to purify CL from water need to be adapted to this simplification. It was investigated whether a flash evaporator would be sufficient to achieve the patents' CL recovery constraints (see Figure 1).

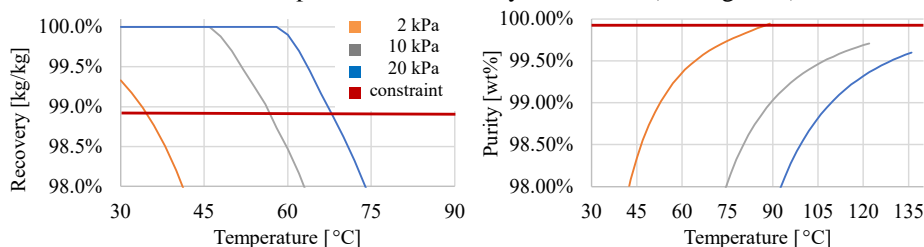


Figure 1: CL recovery (left) and purity (right) versus temperature at different flash unit pressures.

The graphs show that within the pressure range, the recovery and purity constraints cannot be met at the same time. Hence, a distillation column was chosen with an operating pressure of 16 kPa, which allows the use of a single liquid ring pump as well as CW as a cold utility. The process flowsheet for a reasonable CL production capacity of 14.0 kt/year, including the equipment and utility choices, is shown in Figure 2.

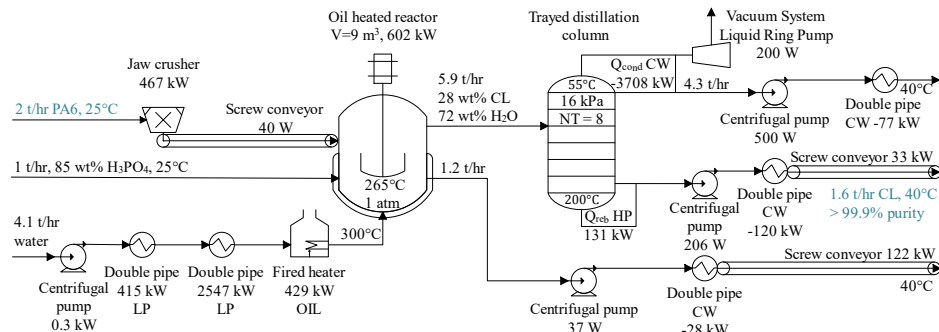


Figure 2: Schematic representation of process flowsheet and choice of equipment and utility.

As can be seen, crushed PA6 reacts to CL using  $\text{H}_3\text{PO}_4$  and superheated steam in a jacketed and agitated reactor with a conversion of 80%. The formed CL is obtained at the top as a vapour stream along with steam. Superheated steam functions in several ways, as it shifts the reaction equilibrium by carrying CL, supports mixing and supplies heat. As described in chapter 2.1,  $\text{H}_3\text{PO}_4$  acts as a catalyst by activating the carbonyl group and facilitating the attack of water. However, it is also consumed in the simultaneous protonation of ACA and salt formation (Češarek et al., 2020). Those products are discharged at the bottom of the reactor and have to be treated as hazardous waste.

The product stream is obtained at the bottom of the vacuum distillation column with a recovery of 99.9% and a CL purity of over 99.9 wt%. Subsequently, the product stream is cooled and solidified to 40 °C by a heat exchanger and screw conveyor.

### 3.2. Economic Analysis

It was found that the described process is profitable, as the break-even point is already reached at a CL production of 4.6 kt/year.

#### 3.2.1. Cost Contributions

To identify the most cost-intensive process units, the OPEX and CAPEX structure was investigated (Figure 3).

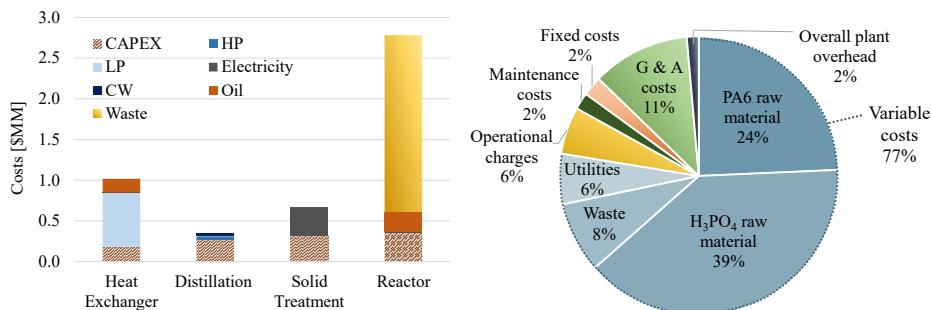


Figure 3: Costs of the process units in \$MM (left bar chart) and OPEX (right pie chart).

Despite the fact that the CAPEX of \$5.5M is rather small compared to the OPEX of \$26.3M per year, none of these process units stands out regarding their capital expenditures. On the other hand, the graphs reveal that variable costs are dominant. Even more interesting is that H<sub>3</sub>PO<sub>4</sub> raw material costs and its waste treatment as reactor discharge make up 47% of the overall OPEX. The cost contribution of H<sub>3</sub>PO<sub>4</sub> is, therefore, further investigated through a sensitivity analysis.

#### 3.2.2. Sensitivity Analysis and Case Studies

Next to the amount of H<sub>3</sub>PO<sub>4</sub>, possible effects of electricity price volatilities were analysed. Additionally, the kinetic pre-exponential factor was varied to study the consequences of potential inaccuracies on the results, which are displayed in Figure 4.

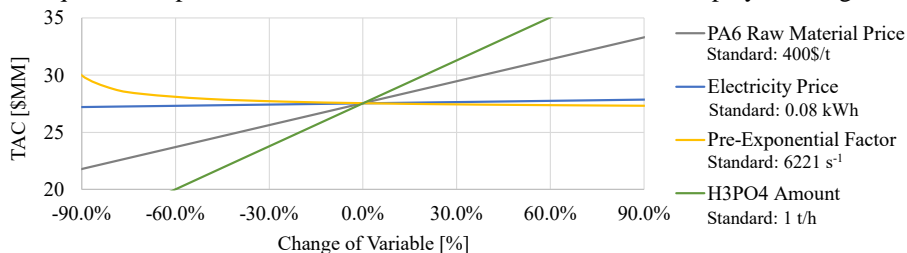


Figure 4: Sensitivity analysis where TAC is displayed over the percentual change of certain variables.

The change of the H<sub>3</sub>PO<sub>4</sub> amount has notably the most significant impact on the TAC, which aligns with the results of the previous chapter. This is because increasing the amount of H<sub>3</sub>PO<sub>4</sub> does not only result in high raw material costs but also in huge waste treatment costs. This further proves that H<sub>3</sub>PO<sub>4</sub> is the main bottleneck of the process. On the other hand, the electricity and OIL price variations barely have any impact on the TAC. Similarly, in a range of 30%, any inaccuracy of the pre-exponential factor would have minor effects. However, the profitability of the process heavily depends on the PA6 waste raw material purchase cost, with an upper profitability boundary of \$650 per tonne.



### 3.2.3. Cash Flow Analysis in Best and Worst Case

To account for inaccuracies and model errors, a best and worst-case scenario was established for the profitability and cash flow analysis (see Figure 5). Utility prices and the pre-exponential factor were all altered by 20%. Additionally, the Lang-Factor method for calculating CAPEX has an accuracy of  $\pm 35\%$ , which was incorporated as well.

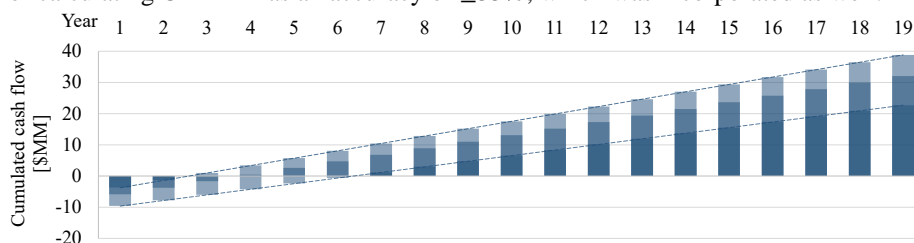


Figure 5: Cumulative cash flow in \$MM for 19 years in the best, standard, and worst-case scenarios.

It can be predicted that the process becomes profitable within a minimum of three to a maximum of seven years, and cumulative cash flow stays within certain boundaries. This proves that the established model and process analysis are solid.

## 4. Conclusions

In this work, a techno-economic analysis of the conventional chemical recycling route of PA6 was performed. It showed that the process stays profitable even after the incorporation of realistic parameter inaccuracies and worst-case assumptions. The results indicate that neither the reaction nor the purification part of the process has a significant impact on the TAC. Instead, the amount of  $H_3PO_4$  and the waste stream containing unreacted PA6, oligomers and  $H_3PO_4$  are the critical variables. Eventually, it can be concluded that depolymerisation processes of PA6 can be very promising and have a great potential for optimisation, especially regarding the catalyst and solvent amount.

## References

- M. Pohjakallio, T. Vuorinen, A. Oasmaa, 2020, Chemical routes for recycling -dissolving, catalytic, and thermochemical technologies, *Plastic Waste and Recycling*, pp. 359–384.
- G. W. Coates, Y. D. Y. L. Getzler, 2020, Chemical recycling to monomer for an ideal, circular polymer economy, *Nature Reviews Materials*, Vol. 5, No. 7, pp. 501–516.
- A. Kamimura, K. Ikeda, S. Suzuki, K. Kato, Y. Akinari, T. Sugimoto, K. Kashiwagi, K. Kaiso, H. Matsumoto, M. Yoshimoto, 2014, Efficient conversion of polyamides to  $\omega$ -hydroxyalkanoic acids: a new method for chemical recycling of waste plastics, *ChemSusChem*, Vol. 7, No. 9, pp. 2473–2477.
- T. F. Corbin, A. Handermann, R. Kotek, W. D. Porter, 1999, Reclaiming epsilon-caprolactam from nylon-6 carpet, BASF, US005977193A.
- J. Chen, Z. Li, L. Jin, P. Ni, G. Liu, H. He, J. Zhang, J. Dong, R. Ruan, 2010, Catalytic hydrothermal depolymerisation of nylon 6, *J Mater Cycles Waste Manag*, Vol. 12, No. 4, pp. 321–325.
- W. D. Seider, J. D. Seader, D. R. Lewin, 2009, *Product and Process Design Principles: Synthesis, Analysis and Evaluation*, New York:Wiley.
- S. Penczek, P. Kubisa, K. Matyjaszewski, 1985, Chapter 11: Polyamides, *Cationic Ring-Opening Polymerization, Advances in Polymer Science*, Springer, Vol. 68/69, pp 201–208.
- U. Češarek, D. Pahovnik, E. Žagar, 2020, Chemical Recycling of Aliphatic Polyamides by Microwave-Assisted Hydrolysis for Efficient Monomer Recovery, *ACS Sustainable Chemistry & Engineering*, Vol. 8, No. 43, pp. 16274–16282.
- J. Harmsen, A. B. de Haan, P. L. J. Swinkels, 2018, *Product and Process Design: Driving Innovation*, De Gruyter.
- J. Brandrup, E. H. Immergut, E. A. Grulke, 2003, *Polymer Handbook*, Wiley.

# Multiscale High-throughput Screening for Membrane-based Green Hydrogen Separation Process

Nahyeon An<sup>a,b</sup>, Seongbin Ga<sup>a,c</sup>, Hyungtae Cho<sup>a</sup>, Boram Gu<sup>d,†</sup>, Junghwan Kim<sup>a,b,†</sup>

<sup>a</sup>*Green Materials and Processes RD Group, Korea Institute of Industrial Technology, 55, Jongga-ro, Jung-gu, Ulsan, 44413, Republic of Korea*

<sup>b</sup>*Department of Chemical and Biomolecular Engineering, Yonsei University, 50, Yonsei-ro, Seoul, 03722, Republic of Korea*

<sup>c</sup>*Department of Chemical Engineering, College of Engineering, University of Ulsan, 93, Daehak-ro, Nam-gu, Ulsan, Republic of Korea*

<sup>d</sup>*School of Chemical Engineering, Chonnam National University, 77, Yongbong-ro, Gwangju, 61186, Republic of Korea*

<sup>†</sup> *Corresponding authors.*

## Abstract

Hydrogen purification is essential to use ammonia as a carrier of green hydrogen since the green hydrogen produced from the transported ammonia inevitably results in nitrogen as a byproduct. The membrane-based separation process has been considered an option for green hydrogen purification because of its various advantages. The selection of proper membrane materials is an important task to efficiently separate gas mixtures. Because it is difficult to compare numerous membranes through experiments, the membrane has been selected using the molecular simulation-based high-throughput screening (HTS) method. However, in the previous studies, the physical properties used to evaluate the membrane materials were not directly related to the separation process performances. To address this limitation, this study entailed the use of multiscale HTS to select the covalent organic frameworks (COFs) best performing in the membrane-based green hydrogen separation process. The proposed method combines the process and molecular simulations to evaluate the performances of COF membranes. The proposed multiscale HTS method was applied to a COF database, and the 648 COFs in the database were explored to select the COFs with the highest performance in hydrogen separation.

**Keywords:** green hydrogen, membrane separation process, covalent organic framework, high-throughput screening

## 1. Introduction

The recent enthusiasm for green technology has put green NH<sub>3</sub> as a hydrogen transportation medium in the spotlight. Separation of N<sub>2</sub> and H<sub>2</sub> is a critical step in converting NH<sub>3</sub> transported from another location into hydrogen for use (Morlanés et al., 2021). Among the various hydrogen purification methods, the membrane-based hydrogen separation process is an active topic of research owing to its ease of use and low energy cost (Ahmad et al., 2012). Various polymer membranes have been designed for efficient gas separation, but the trade-off between selectivity and permeability remains challenging. To address this issue, new nanoporous materials, such as metal organic frameworks and covalent organic frameworks (COFs), have been actively developed. Due to their chemical diversity and structural tunability, COFs have the highest potential for use as

membranes (Aydin et al., 2022).

Selecting an appropriate material among the many porous materials developed to date is one of the most important tasks in the design of a membrane-based separation process because membranes significantly impact the membrane separation process performance. Previous studies on molecular simulation-based high-throughput screening (HTS) have evaluated various materials (Gulbalkan et al., 2022). However, the inability to evaluate the performance in the actual membrane process is a limitation in the previous HTS studies because the computational cost was too high and because only molecular simulation was used.

Therefore, to provide a practical evaluation of the membranes, this study proposes a multiscale evaluation framework. Using the multiscale evaluation framework, this study aimed to select high-performing COF membranes for the membrane-based green hydrogen separation process. In this multiscale evaluation framework, molecular and process simulations were performed to evaluate each COF with process performance. The 20 COFs with the highest  $H_2$  recovery were selected based on the evaluation results. Finally, by examining the  $H_2$  recovery and purity of the membrane process for various process conditions, the performances of the selected COFs were presented for different situations.

## 2. System description

The concept of ammonia as hydrogen transportation media is a future technology for transporting green hydrogen from renewable resource-rich countries to other countries. Green hydrogen is produced in areas with abundant renewable energy resources, and green hydrogen is transported to other areas in the form of ammonia. In the destination, the ammonia is decomposed into hydrogen. The hydrogen should be purified to high purity since its applications as an energy source or as a chemical stock require high purity of hydrogen.

To produce high purity of hydrogen from the transported green ammonia,  $NH_3$  decomposition,  $NH_3$  separator, and  $H_2$  purification processes are required (Figure 1). First, the  $NH_3$  decomposition reactor decomposes  $NH_3$  into  $N_2$  and  $H_2$ . The trace amount of unreacted  $NH_3$  is assumed to be completely removed in the  $NH_3$  separator. The mixture of  $N_2$  and  $H_2$  gases is used as the feed into the membrane separation process. The process conditions are set at a temperature of 313.15 K and a pressure of 8 bar. The ratio of hydrogen to nitrogen ( $H_2:N_2$ ) in the composition was assumed to be 3:1, as in  $NH_3$ . In this study, a co-current membrane separator is used, assuming the same flow directions in the feed and permeate sides. Other required parameters are listed in Table 1.

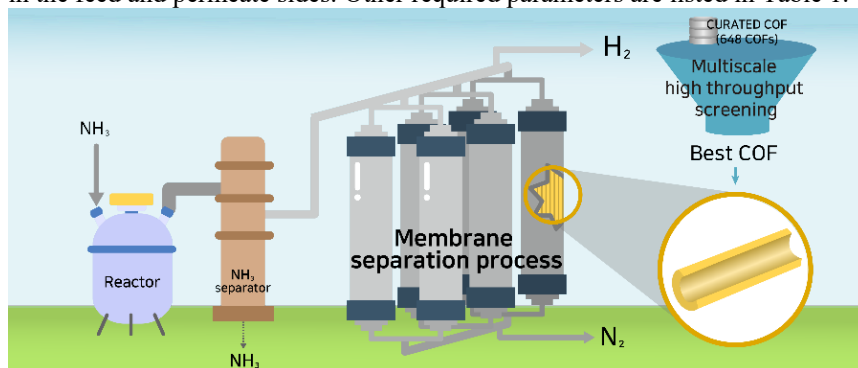


Figure 1. The schematic of integrated  $NH_3$  decomposition and  $H_2$  purification

Table 1. Model parameters and simulation conditions

Definition	Symbol	Value (unit)
Feed flowrate	$F_{f,0}$	3.02 mol/s
Feed composition, H <sub>2</sub> :N <sub>2</sub>	$x_{H_2,0}, x_{N_2,0}$	0.75: 0.25
Membrane inner diameter	$D_i$	$2 \times 10^{-3}$ m
Membrane outer diameter	$D_o$	$2.5 \times 10^{-3}$ m
Membrane module diameter	$D_m$	0.1 m
Number of fibers	$N$	600
Fiber length	$L$	0.6 m
Mass transfer coefficient	$k$	$1 \times 10^{-4}$ m/s

### 3. Multiscale membrane screening method

Figure 2 illustrates the steps involved in the screening method used in this study to discover the membranes with the highest process performances. The proposed multiscale HTS method consists of structural analysis, molecular simulation, process simulation, and results analysis under a wide range of operating conditions (i.e., assess the H<sub>2</sub> recovery and purity of each component and compare the different membranes). To select the best-performing COF membranes, the proposed methodology is applied to the CURATED COF database. Each step of the methodology is explained in the following section.

#### 3.1. Structural properties analysis

This step involves screening work based on the structural property. For gas molecules to permeate, the pore limiting diameter (PLD), which is the smallest pore in the nanoporous material, must be greater than the gas molecule diameter. The COFs with a PLD larger than 3.31 Å, the diameter of the N atom, were chosen because this study requires that H<sub>2</sub> and N<sub>2</sub> be permeable. The PLDs were calculated using the Zeo++ software (ver. 0.3).

#### 3.2. Molecular simulation

In this step, the Grand Canonical Monte Carlo (GCMC) and molecular dynamic (MD) simulations were performed to acquire the parameters for membrane permeance calculations. Permeance, defined by the ratio of membrane permeability ( $P$ ) to thickness ( $d$ ), is an essential parameter to evaluate the process performance. Permeance is obtained using the following equation:

$$a_i = \frac{P}{d} = \frac{q_i \cdot \mathcal{D}_i}{p_i \cdot \rho \cdot d} \quad (1)$$

where  $q$ ,  $\mathcal{D}$ ,  $p$ , and  $\rho$  denote gas uptake, self-diffusion coefficient, partial pressure, and density of COF, respectively. The subscript  $i$  denotes each gas component. Parameters  $q$  and  $\mathcal{D}$  are calculated from the GCMC and MD simulations, respectively. The simulations were conducted using the RASPA 2.0 package.

The N<sub>2</sub> and H<sub>2</sub> partial pressures were set to 2 and 6 bar, respectively, and the temperature at 313.15 K for the GCMC simulations. The GCMC simulation used 10,000 and 20,000 cycles for equilibration and ensemble averages, respectively. The van der Waals (vdW) interaction, which was modeled using the Lennard–Jones (LJ) 12-6 type equation, was used to approximate the interaction between atoms in the systems (Altintas and Keskin, 2019). The LJ parameters for N<sub>2</sub> and H<sub>2</sub> molecules were obtained from the transferable potential for the phase equilibria (TraPPE) force field. A vdW cutoff distance of 12.8 Å was used to truncate the vdW interactions between atoms.

MD simulations were performed to calculate  $\mathcal{D}$  for both H<sub>2</sub> and N<sub>2</sub> (Azar et al., 2019). Thirty molecules for each gas component were simulated in infinite dilution. The number of initialization, production, and total cycles was set to 1000, 10,000, and 1,000,000,

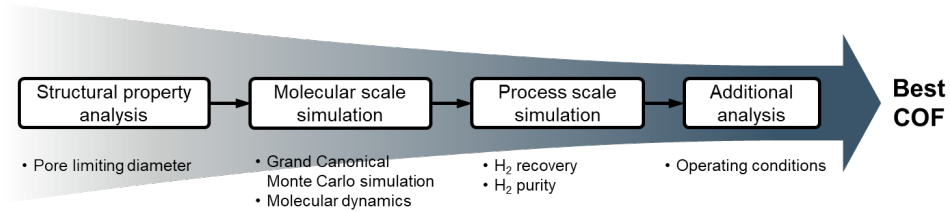


Figure 2. The multiscale high-throughput screening method

respectively. MD simulations were performed in the NVT ensemble with a time step of 0.5 fs. The mean square displacement of gas molecules was used to calculate  $D_i$ .

### 3.3. Process simulation

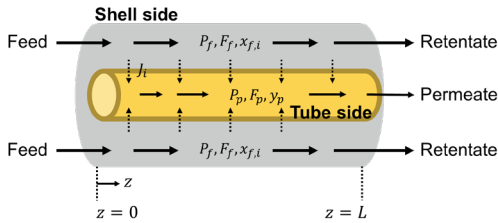


Figure 3. An illustration of a hollow fiber membrane separation module

Using the permeance obtained in the previous section, the  $H_2$  recovery was calculated through process simulation in this step to evaluate each COF membrane. To simulate the membrane separation process, the system is configured as shown in Figure 3.  $P$ ,  $F$ , and  $y$  represent pressure, flowrate, and mole fraction, respectively. The subscripts  $f$ ,  $p$ , and  $i$  represent the feed, permeate side, and gas component, respectively. The partial pressure difference between the feed and permeate sides, which is calculated using Eq. (2), determines the flux of gas passing through the membrane.

$$J_i = a_i(P_{f,i} - P_{p,i}) \quad (2) \quad P_{j,i} = x_{j,i}P_j \quad \text{for } j = f \text{ or } p \quad (3)$$

where  $J$  denotes the permeate flux in  $\text{mol}/(\text{m}^2 \cdot \text{s})$ ,  $a$  denotes the membrane permeance in  $\text{mol}/(\text{Pa} \cdot \text{m}^2 \cdot \text{s})$  derived from the molecular simulations, and  $P$  denotes the partial pressure in Pa. Partial pressure is calculated using the total pressure and molar fraction of each component, as shown in Eq. (3), where  $x_{j,i}$  denotes molar fraction of component  $i$  in  $j$  stream. The component balances for each feed and permeate side are expressed as follows:

$$\frac{dF_{f,i}}{dz} = -J_i W_{int} \quad (4) \quad \frac{dF_{p,i}}{dz} = J_i W_{int} \quad (5)$$

where  $W_{int}$  denotes the interfacial membrane width in m, calculated using  $W_{int} = \pi D_o N$ . The detailed calculation method is described in the literature (Gu, 2022). The flowrate and pressure for each side and component are derived from solving the above ordinary differential equation along with  $z$ -axis. Then, the  $H_2$  purity and recovery are defined as follows:

$$\text{Purity} = \frac{F_{p,H_2}}{\sum F_{p,i}} \times 100 (\%) \quad (6) \quad \text{Recovery} = \frac{F_{p,H_2}}{F_{f,H_2}} \times 100 (\%) \quad (7)$$

The  $H_2$  product must be at least 97% pure to be used as an energy source; the higher the recovery, the better. A parametric study with varied operating conditions was carried out to identify the optimal conditions for maximized  $H_2$  recovery while satisfying the purity standard.

## 4. Results and discussion

### 4.1. Multiscale high-throughput COF screening results

In this section, the results of each screening step are reported. Figure 4(a) presents the PLD distribution of the 648 COFs in the CURATED COF database, where the PLD threshold of 3.31 Å is indicated by a red dashed line. Among the target COFs, 16 COFs were identified to have PLD values below the threshold. GCMC, MD, and process simulations were performed for the remaining 632 COFs.

The process simulation results are displayed in Figure 4(b). While most COF membranes can produce H<sub>2</sub> with a high purity greater than or equal to 80%, the H<sub>2</sub> recovery varies widely. The red dots in the graph represent the top 20 COFs that achieved the highest recovery with H<sub>2</sub> purity above the 97% threshold.

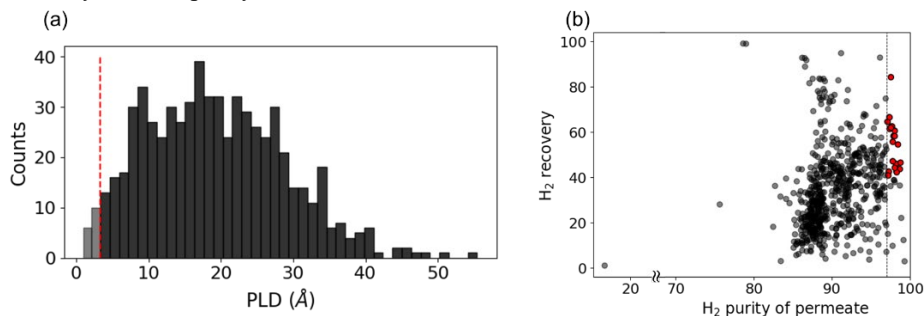


Figure 4. Pore limiting diameter distribution of CURATED COFs database (a), and H<sub>2</sub> purity and recovery results from process simulation (b).

### 4.2. Analysis of membrane module performance

This section analyzes the influence of feed flow rate on the membrane process performances, which reveals the correlation between H<sub>2</sub> recovery and H<sub>2</sub> purity of the 20 selected COFs. Figure 5(a) displays the variations in H<sub>2</sub> purity and H<sub>2</sub> recovery for varied feed flowrates into each fiber. It is shown that the recovery decreases with increasing feed flowrate, while the purity is higher at a higher flowrate. In other words, by increasing the feed flowrate, the purity constraint can be met at the expense of recovery. When each COF was set to have an H<sub>2</sub> purity of 97% by adjusting the feed flowrate individually, Tp-DBD turned out to achieve the highest H<sub>2</sub> recovery (synthesized by 2,5-diamino-1,4-

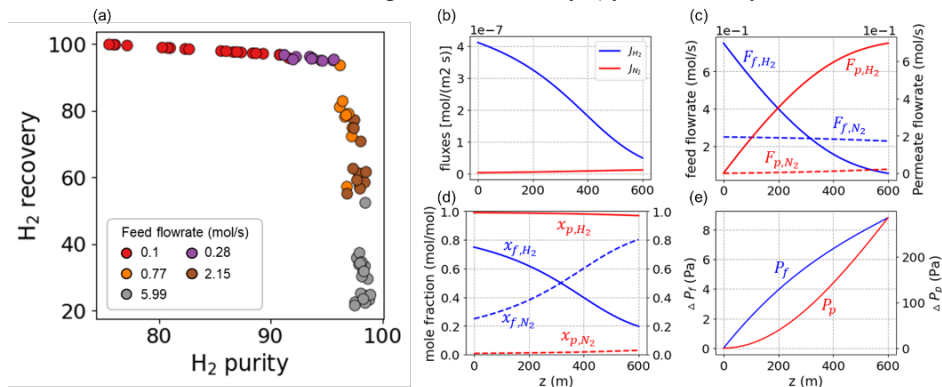


Figure 5. Process simulation results with different feed flowrate and (a) the spatial profiles of fluxes, (b) flowrates, (c) mole fractions (d) pressure drop, and (e) in the feed and permeate sides in the module

benzenediol dihydrochloride and 4,6-triformylphloroglucinol) (Cui et al., 2021). The spatial profiles of the state variables are provided in Figure 5(b)–5(e). Therefore, Tp-DBD COF, which demonstrated an increase in H<sub>2</sub> recovery (46.54% to 92.11%) with a change in feed flowrate (3.02 mol/s to 1.00 mol/s), is the highest-performing COF.

## 5. Conclusion

This study demonstrated the use of the multiscale HTS method to select the best membrane material for the membrane separation process for green hydrogen purification. The proposed multiscale HTS method, consisting of structural analysis, GCMC simulation, MD simulation, and process scale simulation, is used to evaluate the process performance of each membrane. The method was applied to the CURATED COF database to identify the top 20 COFs. By analyzing the H<sub>2</sub> purity and H<sub>2</sub> recovery for different feed flowrates, the COF membrane with the highest recovery was determined. The parametric study with varied feed flowrate demonstrated the trade-off relation between H<sub>2</sub> recovery and purity. The feed flowrate was reduced from 3.02 mol/s to 1.00 mol/s, yielding the highest performance (H<sub>2</sub> recovery of 92.11%).

## Acknowledgement

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2021R1C1C2011236) and by the Korean Institute of Industrial Technology within the framework of the following projects: “Development and application of carbon-neutral engineering platform based on carbon emission database and prediction model [grant number KM-22-0348]”.

## References

- Ahmad, F., Lau, K.K., Shariff, A.M., Murshid, G., 2012. Process simulation and optimal design of membrane separation system for CO<sub>2</sub> capture from natural gas. *Computers and Chemical Engineering* 36, 119–128. <https://doi.org/10.1016/j.compchemeng.2011.08.002>
- Altintas, C., Keskin, S., 2019. Molecular Simulations of MOF Membranes and Performance Predictions of MOF/Polymer Mixed Matrix Membranes for CO<sub>2</sub>/CH<sub>4</sub> Separations. *ACS Sustainable Chemistry and Engineering* 7, 2739–2750. <https://doi.org/10.1021/acssuschemeng.8b05832>
- Aydin, S., Altintas, C., Keskin, S., 2022. High-Throughput Screening of COF Membranes and COF / Polymer MMMs for Helium Separation and Hydrogen Purification. <https://doi.org/10.1021/acssami.2c04016>
- Azar, A.N.V., Velioglu, S., Keskin, S., 2019. Large-Scale Computational Screening of Metal Organic Framework (MOF) Membranes and MOF-Based Polymer Membranes for H<sub>2</sub>/N<sub>2</sub> Separations. *ACS Sustainable Chemistry and Engineering* 7, 9525–9536. <https://doi.org/10.1021/acssuschemeng.9b01020>
- Cui, W., Zhang, C.-R., Xu, R.-H., Chen, X.-R., Yan, R.-H., Jiang, W., Liang, R.-P., Qiu, J.-D., 2021. Low Band Gap Benzoxazole-Linked Covalent Organic Frameworks for Photo-Enhanced Targeted Uranium Recovery. *Small* 17, 2006882.
- Gu, B., 2022. Mathematical Modelling and Simulation of CO<sub>2</sub> Removal from Natural Gas Using Hollow Fibre Membrane Modules. *Korean Chemical Engineering Research* 60, 51–61. <https://doi.org/10.9713/kcer.2022.60.1.51>
- Gulbalkan, H.C., Haslak, Z.P., Altintas, C., Uzun, A., Keskin, S., 2022. Assessing CH<sub>4</sub>/N<sub>2</sub> separation potential of MOFs, COFs, IL/MOF, MOF/Polymer, and COF/Polymer composites. *Chemical Engineering Journal* 428, 131239. <https://doi.org/10.1016/j.cej.2021.131239>
- Morlanés, N., Katikaneni, S.P., Paglieri, S.N., Harale, A., Solami, B., Sarathy, S.M., Gascon, J., 2021. A technological roadmap to the ammonia energy economy: Current state and missing technologies. *Chemical Engineering Journal* 408. <https://doi.org/10.1016/j.cej.2020.127310>

# Discrete Element Simulation and Economics of Mechanochemical Grinding of Plastic Waste at an Industrial Scale

Elisavet Anglou<sup>a</sup>, Yuchen Chang<sup>a</sup>, Arvind Ganesan<sup>a</sup>, Sankar Nair<sup>a,b</sup>, Carsten Sievers<sup>a,b</sup>, Fani Boukouvala<sup>a</sup>

<sup>a</sup>Department of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA

<sup>b</sup>Renewable Bioproducts Institute, Georgia Institute of Technology, Atlanta, Georgia 30332, United States

## Abstract

Efficient and sustainable chemical recycling pathways for plastics are vital for addressing the negative environmental impacts associated with their end-of-life management. Mechanochemical depolymerization in ball mill reactors is a new promising route to achieve solid-state conversion of polymers to monomers, without the need for additional solvents. Physics-based models that accurately describe the reactor system are necessary for process design, scaling up, and reducing energy consumption. Motivated by this, a Discrete Element Method (DEM) model is developed to investigate the ball milling process at laboratory and industrial scales. The lab-scale model is calibrated and validated with data extracted from videos using computer vision tools. Finally, scaled-up ball mill designs capable of depolymerizing varying feeds of PET waste were simulated, and their capital and operating costs are estimated to assess the economic potential of this route.

**Keywords:** ball milling, plastic recycling, discrete-element-method, technoeconomics

## 1. Motivation and background

Traditional economies rely on linear production processes in which raw materials are manufactured into useful products and later discarded as waste. However, this economic model places a significant strain on natural resources and the environment, particularly in the case of single-use plastics. In 2018, 360 million tons of plastics were produced globally, with 80% ending in landfills or in the sea, whereas only 10% were recycled (Tricker et al., 2022). To alleviate the environmental stress on natural resource degradation, there is increasing interest in transitioning to a circular economic model in which waste materials, such as plastics, will be recycled back into the economy.

Plastic recycling methods can be categorized into mechanical and chemical pathways. Most recycling infrastructures currently focus on mechanical processes in which waste is physically reshaped into new products from a melt. However, mechanical and thermal stresses acting on the polymer melt during processing degrade the integrity of the plastic product; therefore, each plastic product can only be recycled a limited number of times, and only if mixed with large quantities of virgin polymers (Tricker et al., 2022). A promising alternative is the chemical recycling of polymers, which bypasses the limitation of material degradation entirely, by converting polymers directly to monomeric molecules through depolymerization. Common chemical recycling pathways such as pyrolysis and solvolysis involve high energy consumption or the use of solvents, therefore hindering their economic viability. The development of alternative depolymerization



processes for plastic waste is crucial for implementing sustainable practices and reducing negative environmental effects. In a previous study (Tricker et al., 2022), it was shown that mechanochemical depolymerization in ball mill reactors is a promising route for converting solid polymers into monomers without the need for additional solvents. Mechanochemical depolymerization pathways utilize the mechanical energy supplied by collisions between grinding bodies (balls and walls) of a ball mill to drive the reaction between solid reactant particles. Thus, accurately simulating the movement and interactions between these entities can provide crucial information for explaining mechanisms of mechanochemical depolymerization and for optimizing process design. Two modeling frameworks have predominantly been used to simulate ball mill systems: semi-empirical population balance and mechanistic Discrete Element Method (DEM) models. DEM models for ball milling have received increasing attention owing to their ability to describe the kinematics of moving entities and the corresponding energy involved in their collisions during ball milling. Additionally, ball milling operations at the industrial scale can be energy-intensive; hence, even a marginal improvement in efficiency can lead to a significant reduction in expenses. For instance, approximately 110 kWh of electric energy is consumed during the production of 1 ton of cement, of which 70% is required for comminution, whereas only 1-5% is explicitly used for particle breakage; the remaining energy is wasted (Boemer & Ponthot, 2017; Tavares, 2017).

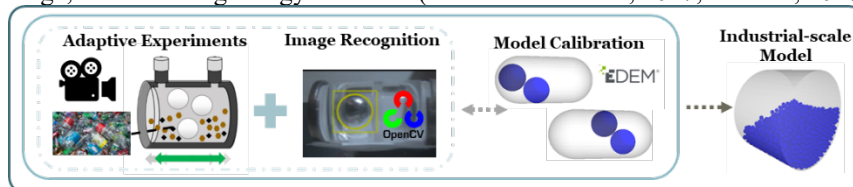


Figure 1: The main components of this study include: (a) experimental design and machine learning computer vision tools for the calibration of a lab-scale DEM model, and (b) DEM simulations of scaled-up ball mill configurations for various operating settings to estimate capital and operating costs

In this work, a DEM model was developed to investigate the depolymerization of poly(ethylene terephthalate) (PET) powder. First, high-speed video recordings of a lab-scale ball mill were obtained at operating conditions that allowed for quantitative PET depolymerization in corresponding experimental studies in the same mill (Tricker et al., 2022), and a computer vision algorithm was used to track the moving balls in the reactor. The extracted data were used to parameterize the DEM model and predict the kinematic interactions of moving entities in the mill. Once the mechanistic model parameters were optimized, the geometry of an industrial-scale ball mill reactor was created, and various operating settings were simulated. The resulting DEM simulation was used to investigate the tradeoffs between achieved PET yields and energy requirements at the industrial scale. Finally, capital, and operating costs of the process were calculated for various scenarios of waste feeds, using data obtained from US state levels of plastic waste collection.

## 2. Methods

### 2.1. High-fidelity model for the laboratory ball mill reactor

PET depolymerization experiments were performed in a Retch MM400 vibratory ball mill with stainless steel grinding bodies. PET in powder form was milled with a stoichiometric amount of sodium hydroxide (NaOH) to form ethylene glycol (EG) and disodium terephthalate salt ( $\text{Na}_2\text{TPA}$ ). Details regarding experimental conditions and results have been reported in (Tricker et al., 2022). A transparent PMMA milling vessel

identical to that used in the depolymerization experiments was manufactured and used in motion-tracking experiments, but without reactants (PET + NaOH) present. The mill's operation at various conditions (ball numbers and sizes, milling frequencies) was filmed using a high-speed camera (2134 fps) to study the collision kinematics between balls and the vessel wall. A Python script based on the OpenCV computer vision library was used to analyze the video frames and track the position of the ball(s). The coordinates of the ball(s) in space and time were identified and used to calculate their velocity, which is defined as the change in position between two consecutive frames.

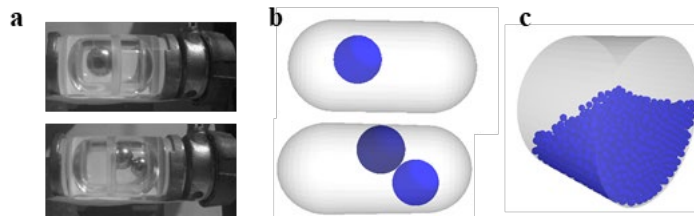


Figure 2: (a) Experimental setup including a vibratory 25mL reactor and stainless-steel balls (b) Replication of the experimental setup in DEM simulation with one and two balls (c) The geometry of the industrial-scale reactor as simulated in the DEM software

Next, the geometry of the laboratory setup was developed in Solidworks while the DEM model was developed in EDEM. The Hertz-Mindlin contact model was applied, as it is most appropriate for non-cohesive spherical shapes (Metta, Ierapetritou, & Ramachandran, 2018). Typical steel properties were used as inputs to the DEM model (density 7900 kg/m<sup>3</sup>, Poisson's ratio 0.3, coefficient of restitution 0.5, shear modulus 77 GPa). The coefficient of static and rolling friction are taken as 0.5 and 0.01 respectively based on (Metta et al., 2018). For PMMA-steel interactions, a grid search of the coefficient of restitution and the friction parameters was performed, and optimal values that match the velocity and the number of collisions between the simulation and the motion-tracking experiments were identified. The laboratory ball milling simulation was validated based on the recorded experiments under different operating conditions. Once the lab-scale simulation is validated by experiments, the identified parameters that govern ball-to-ball and ball-to-wall interactions can be used to simulate a scaled-up geometry.

### 2.2. Scaling up and technoeconomic analysis of ball mill reactor

A typical industrial ball milling setup comprises a rotating cylindrical vessel filled with balls and PET particles. Electric energy is used to satisfy the energy requirements necessary to drive the electric motor that rotates the mill and grinding media. To establish the operating costs, a new geometry and scale of DEM simulations was developed. A rotating cylindrical vessel was designed, and the simulation was developed (Figure 2c) using the validated material and contact parameters. The resulting torque requirements loads were extracted from the DEM simulation, and the operating costs were evaluated based on Equation (1) for 7920 hours of operation annually.

$$\text{Electricity Cost} = \text{Price} \times \text{Power} = \text{Price} \times (T_M \times \omega_M) \quad \text{Equation (1)}$$

$$C_{\text{Ball Mill}} = 64640 \times W^{0.64} \quad \text{Equation (2)}$$

$$Y_{\text{PET}}(\%) = 6.38 \text{ BPR} - 46.35 \text{ for } 10 \leq \text{BPR} \leq 23 \quad \text{Equation (3)}$$

Here,  $T_M$  is the torque associated with the rotation of the reactor vessel while  $\omega_M$  is the mill's angular velocity, which in this work is equals to 15 rpm. The angular velocity was chosen such that it lies between 60-80% of the critical speed depending on the reactor

volume (Wang, Yang, & Yu, 2012). The fill level of the ball mill was kept constant at 30 %, and the radius of the balls was set to 80 mm for all cases. An electricity cost of \$0.07/kW-h was used while the purchase cost of the ball mill equipment was estimated based on the feed flowrate ( $W$  [ton/hr]) according to Equation (2) (Seider, Seader, Lewin, & Widagdo, 2004), and then adjusted to the 2022 dollar value via the CEPCI Index. According to our previous kinetics study on the PET depolymerization reaction (Tricker et al., 2022) for the lab-scale ball milling system, monomer yield follows a linear relationship (Equation (3)) with the ball-to-powder mass ratio (BPR) parameter, that is, the ratio of the ball mass to the mass of the reactants. Assuming that the same BPR values can be maintained at the industrial scale, the corresponding conversion, and costs can be calculated. Higher BPR ratios results in higher PET monomer yields, and full conversion is achieved when  $BPR = 23$  with a 20 min reaction time (Tricker et al., 2022). The BPR parameter was used to scale-up the reactor and to evaluate the corresponding conversion, number of balls, and reactor volume for a certain PET waste feed.

### 3. Results and Discussion

#### 3.1. Laboratory-scale model validation

The average ball velocity is compared in Figure 3a, which reveals that the values are in near-perfect agreement for the different operating conditions (vibration frequencies, diameters) tested and can thus validate the lab-scale model. Stainless-steel balls with diameters equal to 17.6 mm and 20.6 mm were used for the model validation while the milling frequency ranged from 22.5 to 30 Hz. Once the DEM parameters were estimated, the simulation can be executed for different conditions and reactor sizes, and the results can be used to further analyze the ball milling system.

#### 3.2. Simulations at the industrial scale

DEM simulations were used to define the scaling-up procedure for PET waste depolymerization in ball mills and to evaluate the associated energy requirements. To illustrate the resulting capital and operating costs of plastic waste depolymerization at the industrial scale, two separate cases were considered.

##### 3.2.1. Case-study 1: Influence of BPR to the operating costs

The BPR ratio significantly influenced the extent of PET depolymerization. To this end, the first case study depicts how the operating costs change with respect to the BPR value for the same feed of 1000 kg/hr PET waste (+400 kg/hr NaOH). The reactor volume and number of balls required to depolymerize the PET waste were calculated based on the BPR parameter.

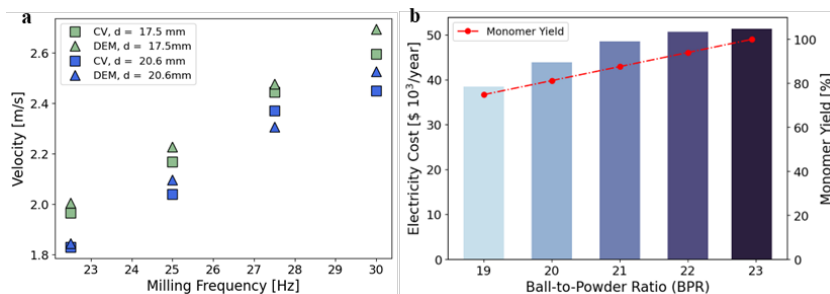


Figure 3: (a) Comparison of the average velocities of the stainless-steel ball at different vibration frequencies (22.5 – 30 Hz) extracted from the DEM simulations (triangles) and the recorded ball mill experiments (squares), (b) Comparison of annual electricity cost and achieved monomer yield for reactors processing 1000 kg/hr of PET (+400 kg/hr NaOH) at different BPR ratios.

The results of this analysis are shown in Figure 3b, which illustrates how the electricity cost changes with different BPR values. For higher BPRs, more balls are added inside the mill to grind the same amount of waste, resulting in increased power requirements for the mill rotation. More specifically, to achieve a 75% monomer ( $\text{Na}_2\text{TPA}$ ) yield (BPR=19), \$1.62/tonne<sub>PET</sub> are spent on electricity, whereas for full depolymerization (BPR=22.9) the cost is increased to \$2.16/tonne<sub>PET</sub> for 20 min milling time (Tricker et al., 2022). The purchase cost of the ball mill vessel was estimated to be \$119,719 and was constant for all BPR combinations (assuming a negligible cost for the additional grinding balls). However, as PET conversion is reduced, subsequent unit operations will be necessary to separate and recycle unreacted PET waste which will increase capital and operating expenditures.

### 3.2.2. Case study 2: Influence of PET waste feedstock flowrate

The second case study illustrates the dependence of the total cost of ball milling to the mass flow rate of PET waste. Different PET waste recycling feed flow rates were evaluated, and the corresponding costs were estimated for a BPR value of 20 resulting in 81.2 % of PET converted to monomers. The PET waste for different US states were used as input to the DEM simulation based on recycling quantities for 2018 as described in the report of the National Association for PET Container Resources (*NAPCOR 2018*). The study region encompasses the South Atlantic region of the US, which in turn comprises nine states (Florida, Georgia, North Carolina, Virginia, Maryland, South Carolina, West Virginia, and District of Columbia). It is assumed that the PET recycling quantities are equally distributed to each state in the region based on the corresponding population for the same year. The resulting PET waste flowrates for each state are used as inputs to the DEM simulation using stoichiometric NaOH feed.

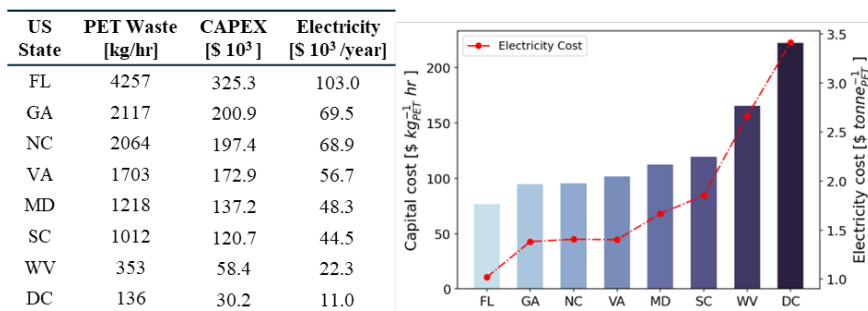


Figure 4: PET waste collected at each state in 2018 scaled based on population, and the resulting capital and operating costs. The graph illustrates the capital and operating costs per kg of PET waste fed in the ball mill reactor. The bars denote the capital cost per feed flow rate in \$ kg<sub>PET</sub><sup>-1</sup>hr, whereas the electricity cost per tonne is shown on the secondary axis for 20 min operation (red dotted line).

The results of this analysis are shown in Figure 4, which illustrates how the required expenditure per unit feed changes for different PET waste feeds. As expected, both capital and operating costs increased with feed flow rate, because maintaining the same BPR value requires larger numbers of balls and a larger reactor volume, which would require greater motor power to achieve depolymerization conditions for that entire mass of grinding bodies. In contrast, per unit costs decreases as the feed rate increases highlighting the benefits of economy of scale.

The economic potential of the mechanochemical route based solely on raw material (PET bale: \$150/tonne, NaOH: \$608/tonne, H<sub>2</sub>SO<sub>4</sub>: \$96/tonne) and product (vTPA: \$1143/tonne, EG: \$961/tonne) prices (Singh et al., 2021) was evaluated at \$978/tonne of

TPA, which is an indication of the upper bound of the potential profits. This value, in combination with the low electricity costs that were estimated in both case studies at \$2-3/tonne<sub>PET</sub>, reveals that there is still a significant profit margin to account for mechanical pretreatment, required to grind PET waste to micro sizes, as well as downstream processes to recover the products of mechanochemical hydrolysis of PET waste that can be capital- or energy-intensive, i.e., rTPA crystallization, EG distillation, salt recovery (Singh et al., 2021).

#### 4. Conclusions

In this work, we built a DEM simulation required to design and cost PET depolymerization via ball milling for a variety of feedstock flow rates and operating settings. Two case studies demonstrate the dependence of the operating and capital costs on the achieved PET conversion and feed flow rates. The results indicate that this can be viable process economically, especially for states that recycle more PET where there would be a benefit from economy of scale. Also, there are economic trade-offs to be considered in the future with respect to more costly depolymerization at high conversions, versus partial depolymerization and design of subsequent separation/purification steps and recycling streams. In future work, a more detailed technoeconomic analysis of plastic waste depolymerization via mechanochemical reactions will be performed to account for pre-treatment costs to shred plastic waste into powder form, as well as downstream treatment to purify the mechanocatalysis products. This analysis will be critical for comparing this technology with alternative recycling routes and aid policymakers and industry stakeholders in making informed decisions regarding its overall potential.

#### 5. Acknowledgments

This work is supported by Kolon Industries, Inc., through the Kolon Center for Lifestyle Innovation at Georgia Tech and the U.S. National Science Foundation Emerging Frontiers in Research and Innovation program under grant 2028998.

#### References

- Boemer, D., & Ponthot, J.-P. (2017). DEM modeling of ball mills with experimental validation: influence of contact parameters on charge motion and power draw. *Computational Particle Mechanics*, 4(1), 53-67.
- Metta, N., Ierapetritou, M., & Ramachandran, R. (2018). A multiscale DEM-PBM approach for a continuous comilling process using a mechanistically developed breakage kernel. *Chemical Engineering Science*, 178, 211-221.
- Postconsumer PET Recycling Activity in 2018. Retrieved from <https://napcor.com/wp-content/uploads/2021/07/Postconsumer-PET-Recycling-Activity-in-2018.pdf>
- Seider, W. D., Seader, J., Lewin, D. R., & Widagdo, S. (2004). Product and process design principles: synthesis. *Analysis and Evaluation*, 4.
- Singh, A., Rorrer, N. A., Nicholson, S. R., Erickson, E., DesVeaux, J. S., Avelino, A. F., . . . Avery, G. (2021). Techno-economic, life-cycle, and socioeconomic impact analysis of enzymatic recycling of poly (ethylene terephthalate). *Joule*, 5(9), 2479-2503.
- Tavares, L. M. (2017). A Review of Advanced Ball Mill Modelling. *KONA Powder and Particle Journal*, 34(0), 106-124. doi:10.14356/kona.2017015
- Tricker, A. W., Osibo, A. A., Chang, Y., Kang, J. X., Ganesan, A., Anglou, E., . . . Sievers, C. (2022). Stages and Kinetics of Mechanochemical Depolymerization of Poly (ethylene terephthalate) with Sodium Hydroxide. *ACS Sustainable Chemistry & Engineering*.
- Wang, M., Yang, R., & Yu, A. (2012). DEM investigation of energy distribution and particle breakage in tumbling ball mills. *Powder technology*, 223, 83-91.

# Multi-scale Modelling and Experimental Investigation of Hydrogen Sulphide Thermal Decomposition

Anna Nova,<sup>a,b</sup> Francesco Negri,<sup>a,c</sup> Flavio Manenti<sup>a,b</sup>

<sup>a</sup>*Politecnico di Milano, CMIC Dept. “Giulio Natta”, Piazza Leonardo da Vinci 32, Milan 20133, Italy*

<sup>b</sup>*Consorzio Interuniversitario Nazionale per la Scienza e Tecnologia dei Materiali, Via Giusti 9, 50121 Firenze, Italy*

<sup>c</sup>*Itelyum Regeneration spa, Via Tavernelle 19, Pieve Fissiraga 26854, Lodi, Italy*

\*[flavio.manenti@polimi.it](mailto:flavio.manenti@polimi.it)

## Abstract

The splitting of hydrogen sulphide into hydrogen and elemental sulphur results to be a promising route for the treatment of this dangerous waste gas and, at the same time, for the sustainable production of hydrogen. To reach relevant production rates a comprehensive approach should be applied to the modelling of the process. This work investigates how the thermal decomposition of hydrogen sulphide can be described at different scales: the kinetic scale, the reactor scale, and the chemical process scale. In particular, a detailed kinetics is implemented into DSmoke, a reactor simulation suite for complex kinetics and non-ideal reactive systems. The obtained reactor model is validated through an experimental lab campaign. At the process scale, the splitting step is integrated into a complete plant design (considering a 10 t/h H<sub>2</sub>S feed, the mass flow of H<sub>2</sub> produced from the simulated plant is equal to 543.8 kg/h, with a production cost of 2.39 \$/kg).

**Keywords:** hydrogen sulphide, H<sub>2</sub>S decomposition, multi-scale model, orange hydrogen.

## 1. Introduction

Hydrogen sulphide (H<sub>2</sub>S) is a compound that is naturally present in many resources (both fossil and renewable) and is also artificially produced in some industrial processes. The main source are the processes for the treatment and desulphurization of oil and gas. Since hydrogen sulphide is pollutant, corrosive and flammable, it cannot be released in the environment, and it must be conveniently processed.

At present, H<sub>2</sub>S is usually sent to sulphur recovery units (e.g., Claus process), which perform a partial oxidative conversion able to transform H<sub>2</sub>S into sulphur and water via sulphur dioxide (SO<sub>2</sub>) as intermediate and, at the same time, to recover thermal energy. The process leads to the production of elemental sulphur, a more environmentally friendly and less hazardous compound than H<sub>2</sub>S. In this case hydrogen is not recovered and is completely wasted in the form of water. An interesting alternative could be its conversion into sulphur and hydrogen (Burra et al., 2018), according to the overall splitting reaction ( $\Delta H_r^0 = 84.9$  kJ/mol):



The production of hydrogen results in great interest since this compound is a highly valuable product with a wide range of applications, especially in the energy and chemicals industry. Hydrogen is usually identified through different colours, according to the

production process. Black refers to coal gasification, grey to natural gas reforming, turquoise to natural gas pyrolysis, blue to reforming/gasification + CCUS, green to electrolysis of water with renewable electricity, purple to electrolysis of water with electricity from nuclear plants and yellow to electrolysis of water with mixed-origin grid energy). H<sub>2</sub>S thermal splitting represents a promising path to produce relevant amounts of hydrogen. This H<sub>2</sub> has not been yet classified in the chromatic scale, therefore, it will be named “orange hydrogen”. Obtaining this product from the H<sub>2</sub>S generated in hydrotreating processes may enable the recovery and recycling of H<sub>2</sub>, significantly reducing the need to acquire new H<sub>2</sub> sources and, in general, reducing costs and environmental impacts. In addition, H<sub>2</sub> has the potential to be integrated into the fuel cycle for energy storage carriers, leading to the complete suppression of air pollutants emissions.

A variety of splitting techniques aimed at producing hydrogen and sulphur from H<sub>2</sub>S, have been studied. To date, the developed H<sub>2</sub>S splitting methodologies include thermal (catalytic/membrane-assisted), photocatalytic, non-thermal plasma, thermochemical, electrolysis and biological processes (De Crisci et al., 2019). At 1200 °C, 1 atm and with a steady gas flow rate of 50 mL/min, H<sub>2</sub>S conversion was 35.6 % (97.5 % of its thermodynamic limit). To decrease the high-temperature requirements, catalytic processes have been investigated. Low-temperature processes have been performed with the use of MoS<sub>2</sub>-based catalysts, resulting in 95 % conversion to H<sub>2</sub> at 500 °C. Other materials have also been studied, including, for example, perovskites (Burra et al., 2018), alumina oxides (Cha et al., 2022) and vanadium-decorated titanium carbides (Zhou et al., 2021). Photocatalytic reactions have also been considered. TiO<sub>2</sub> is the most studied semiconductor because of its nontoxicity, low cost, and high stability (Uesugi et al., 2022).

Nevertheless, H<sub>2</sub> generation from H<sub>2</sub>S has not yet been scaled up to bulk production due to economic and technical reasons (Gupta et al., 2016). Some of the open issues are related to the complete capture of pure H<sub>2</sub>S, its energy-intensive regeneration step, the management of sulphur vapours, and the H<sub>2</sub>S effective conversion into hydrogen and elemental sulphur. The problem requires a multi-scale approach, able to model the process from the kinetic scale, to the reactor scale, to the complete process scale. In this work, the hydrogen sulphide thermal splitting is analysed at different scale levels, proposing in this way a comprehensive approach to the study of a process.

## 2. Materials and methods

### 2.1. Experimental set up

The study of H<sub>2</sub>S splitting is first approached by conducting an experimental campaign. An ex-novo laboratory apparatus has been set up. It consists of an open circuit system fed by a gaseous H<sub>2</sub>S/N<sub>2</sub> mixture with a H<sub>2</sub>S concentration up to 8 vol%. The reactor is made of quartz and is designed according to the ideal Plug Flow Reactor (PFR) model. To ensure the adequate thermal conditions required by the process, the reactor is placed inside a tubular oven. Upon exiting the reactor, the sulphur is condensed thanks to the bubbling of the reaction gases into a flask containing distilled water. The gaseous stream is then sent to a stilling chamber to remove all the entrained droplets. A filter ensures then the complete removal of any solid particle from the stream. The gases are finally analysed with a Micro GC (Micro Gas Chromatograph).

### 2.2. DSmoke simulation

DSmoke is a reactor network simulation suite for complex kinetics and non-ideal reactive systems, developed by the research group at Politecnico di Milano. The software allows

the modelling of sequences of reactors (in series or in parallel), possibly coupled with mixers and/or splitters. It is composed by a chemical interpreter, that processes the kinetic mechanism and the thermodynamic properties of the species, and by a reactor model developer.

The lab reactor employed in the experimental campaign has been modelled in DS<sub>Smoke</sub> with a network of three reactors in series. This sequence is composed by a first isothermal plug flow reactor, corresponding to the homogeneously heated length of the quartz tube inside the oven, a second plug flow reactor, displaying a linear temperature profile (going from the temperature value inside the oven to room temperature) and corresponding to the part of the tube at the exit of the oven, where the temperature drops rapidly, and finally a third isothermal plug flow reactor, simply corresponding to the remaining length of the tube. Table 1 shows the details of the three reactors composing the series.

Table 1 - Details of the three reactors composing the DS<sub>Smoke</sub> model sequence.

	Length [mm]	Diameter [mm]	Volume [mm <sup>3</sup> ]
Reactor 1	180	22	68389.2
Reactor 2	5	22	1899.7
Reactor 3	355	22	134878.7

The kinetic scheme introduced in the program for the thermal splitting of hydrogen sulphide has been developed by the group in the past years (Manenti et al., 2013). It is a revised model, introducing some modifications on H<sub>2</sub>S pyrolysis literature schemes (Binoist et al., 2003). The complete mechanism, validated on literature and industrial data, is presented in Table 2.

Table 2 - Kinetic parameters for H<sub>2</sub>S pyrolysis. Rate eq.:  $k=A \cdot T^{\beta} \cdot \exp(-E_a/RT)$  [m-kmol-s-K] (Manenti et al., 2013).

n	Reactions	A	$\beta$	E <sub>a</sub>	n	Reactions	A	$\beta$	E <sub>a</sub>
R1	S+H+M=SH+M	.62E+17	-6	.0	R10b	HSS+H=SH+SH	.11E+14	.353	210.0
R2	S+H <sub>2</sub> =SH+H	.14E+15	.0	19300	R11	HSS+H=S <sub>2</sub> +H <sub>2</sub>	.12E+09	1.653	-1105.0
R3	S <sub>2</sub> +M=S+S+M	.48E+14	.0	77000	R12	HSS+H=H <sub>2</sub> S+S	.44E+14	.000	6326.0
R4	S <sub>2</sub> +H+M=HSS+M	.40E+15	2.84	1665	R13	HSS+S=S <sub>2</sub> +SH	.42E+07	2.200	-600.0
R5	SH+SH=S <sub>2</sub> +H <sub>2</sub>	.50E+12	0.0	.0	R14	HSS+SH=H <sub>2</sub> S+S <sub>2</sub>	.63E+04	3.050	-1105.0
R6	SH+S=S <sub>2</sub> +H	.30E+14	0.0	.0	R15	HSS+HSS=HSSH+S <sub>2</sub>	.96E+01	3.370	-1672.0
R7	H <sub>2</sub> S+M=S+H <sub>2</sub> +M	.16E+25	2.613	89100	R16	HSSH+M=SH+SH+M	.14E+16	1.000	57030.0
	N <sub>2</sub> /1.5 SO <sub>2</sub> /10 H <sub>2</sub> O/10				R17	HSSH+H=HSS+H <sub>2</sub>	.50E+08	1.933	-1408.0
R8	H <sub>2</sub> S+H=SH+H <sub>2</sub>	.35E+08	1.94	904	R18	HSSH+H=H <sub>2</sub> S+SH	.20E+15	.000	.0
R9	H <sub>2</sub> S+S=SH+SH	.83E+14	0.0	7400	R19	HSSH+S=HSS+SH	.29E+07	2.310	1204.0
R10a	HSS+H=SH+SH	.97E+08	1.62	-1030	R20	HSSH+SH=HSS+H <sub>2</sub> S	.64E+04	2.980	-1480.0

### 2.3. Process simulation

The detailed process simulation has been developed using Aspen HYSYS V11 suite. As a thermodynamic model, the SRK (Soave-Redlich-Kwong) equation of state has been chosen for the reactive section of the process, while the Aspen HYSYS amine package has been chosen for the separation section. The reactors present in the process are modelled as Gibbs Reactors.

The reaction is highly endothermic and takes place with an increase in the number of moles. So, it is favoured at high temperatures and low pressure. An H<sub>2</sub>S mass flow of 10 t/h enters the splitting reactor at 1000 °C and 180 kPa. Sulphur is removed by condensation. Residual sulphur vapours are then hydrogenated to produce H<sub>2</sub>S. The obtained gas stream (containing hydrogen in an amount equal to 83 mol% and unreacted H<sub>2</sub>S and H<sub>2</sub>S produced by the hydrogenation of sulphur vapours in an amount equal to 17



mol%) is treated in an amine scrubbing process employing  $2 \times 10^6$  kg/h of a DEA (25 wt%) water solution (Jamekhorshid et al., 2021). After a water removal step, the sweet stream contains pure  $H_2$  and can be sent to the compression train to reach the market pressure of 200 bar.  $H_2S$  is instead recirculated to the splitting reactor. The detailed process flow diagram is reported in Figure 1.

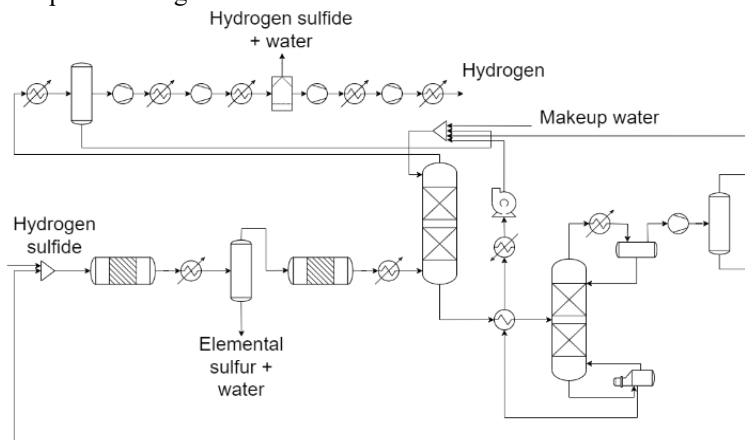


Figure 1 - Process flow diagram for  $H_2S$  conversion.

### 3. Results and discussion

#### 3.1. Experimental results

The reaction has been performed at 1 bar and with a stream of 2 L/h  $H_2S$  (1.5 wt%) in  $N_2$  as feed (to ensure a long residence time and therefore a thermodynamic regime of reaction). Tests at different temperatures has been carried out, obtaining the results summarized in Table 3. The conversions achieved are quite in line with the expected thermodynamic conversions (Kaloidas & Papayannakos, 1987). The error is due to the recombination that takes place at the exit of the reactor: above  $800^\circ C$   $H_2$  reacts with sulphur to give back  $H_2S$ .

Table 3 - Experimental  $H_2$  concentrations and conversions.

Temperature [ $^\circ C$ ]	$H_2$ concentration [mol%]	Experimental conversion [%]	Conversion at thermodynamic equilibrium [%] (Kaloidas & Papayannakos, 1987)
640	0.06	5	6
740	0.18	14	18
831	0.35	28	31
929	0.54	43	47
1016	0.71	56	61
1107	0.81	64	73

#### 3.2. DSmoke lab reactor simulation

The models realized are simulations of the  $H_2S$  splitting reaction performed in the lab reactor. Two temperatures have been considered: 1016 and  $1107^\circ C$ . The results reported in the previous section, 3.1, are an experimental validation of the models obtained with the software.

In Figure 2 and 3 it is possible to notice that along the length of the reactor the concentration of  $H_2$  increases and reaches a plateau, corresponding to the thermodynamical equilibrium. This condition is reached at a short length of the reactor (equal to 12 mm at  $1016^\circ C$  and 4 mm at  $1107^\circ C$ ), suggesting that, in the considered operative regime, a more compact reactor could be a convenient design choice. At the

point of temperature drop (after a residence time inside the reactor of 30 s), the concentration of  $H_2$  starts to decrease, due to the recombination reaction. The temperature drop takes place in a 5 mm length; therefore, the quench is rapid, and the recombination reaction interrupts, fixing a final value for the  $H_2$  concentration.

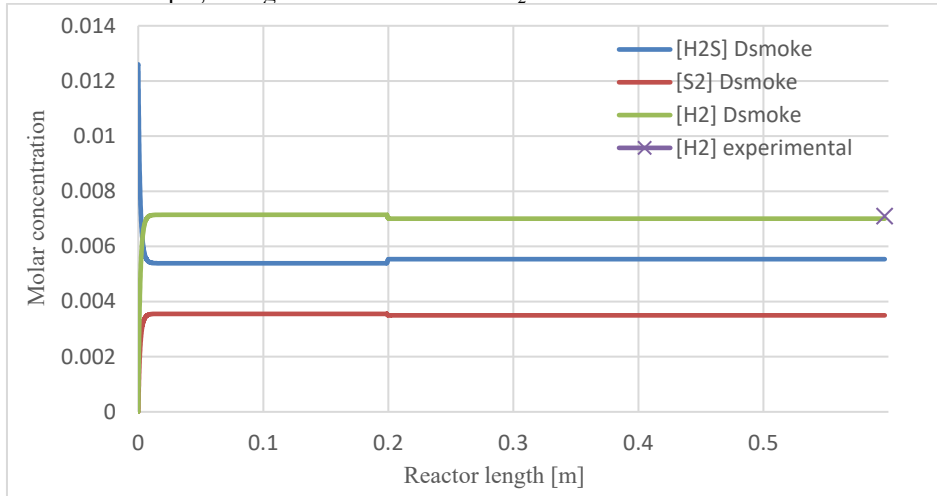


Figure 2 - Molar concentrations of  $H_2S$ ,  $S_2$ ,  $H_2$  at 1016 °C.

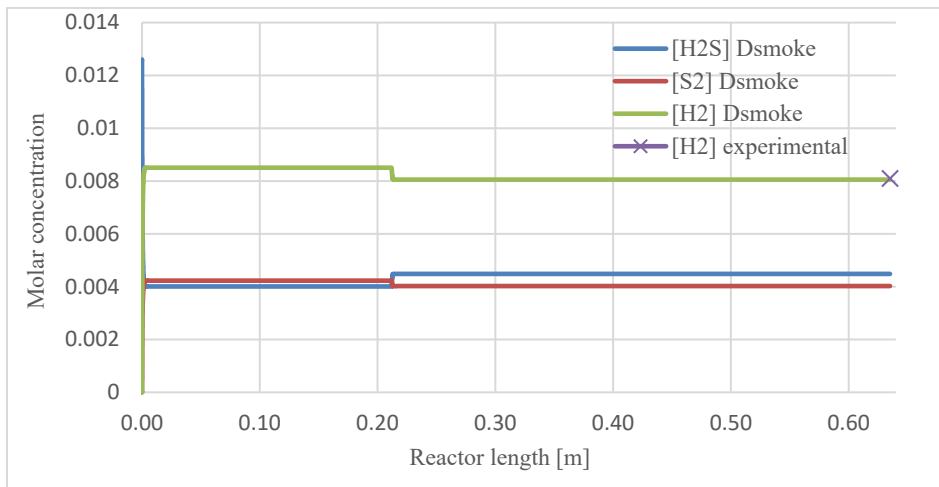


Figure 3 - Molar concentrations of  $H_2S$ ,  $S_2$ ,  $H_2$  at 1107 °C.

### 3.3. Process design

Considering the process scale, a conceptual design for a plant integrating a  $H_2S$  splitting step results to be significantly different from that of classical  $H_2S$  sequestration and conversion processes. This is due to the high concentration of  $H_2S$  that characterises the process streams. In particular,  $H_2S$  is present as a pure component in the feed. This results in a difficult separation of the unreacted  $H_2S$  and in an extremely energy-intensive regeneration step.

In a simulation with the characteristics described in section 2.3, the production is equal to 8666 kg/h of elemental sulphur and 543.8 kg/h of hydrogen. The process is characterized by a considerable  $H_2S$  recycle, consisting of 48340 kg/h. This aspect has a

huge impact on the operative costs of the plant, but at the same time makes the complete conversion of  $\text{H}_2\text{S}$  into  $\text{H}_2$  and  $\text{S}_2$  possible. Water as well can be recirculated inside the plant, exploiting the streams coming from the condensation steps. This allows to reduce the needed amount of makeup water to 110.5 kg/h. The steam consumption, particularly the one related to the regeneration column reboiler, is high, equal to  $1.843 \times 10^5$  kg/h. A possible point of improvement could be represented by an optimisation able to reduce the energy consumption at the regeneration step.

An economic analysis has also been performed, by calculating the cost of the utilities associated with the treatment of  $\text{H}_2\text{S}$  in the simulated plant. The total OPEX are 1301.18 \$/h. Considering the obtained  $\text{H}_2$  productivity of 543.8 kg/h,  $\text{H}_2$  production cost is 2.39 \$/kg. In Europe, this price is lower than the average price of  $\text{H}_2$  produced through electrolysis (3.6 \$/kg) and is almost equal to the price of  $\text{H}_2$  produced from natural gas (EIA, 2019).

#### 4. Conclusions

In this work a multi-scale approach has been applied to the development of a process layout for  $\text{H}_2\text{S}$  thermal splitting, including the kinetics, the reactor modeling and the plant design. As future development, DSmoke will be integrated into Aspen Hysys as CAPE-OPEN tool, in order to better combine the different scales and to optimize the process.  $\text{H}_2\text{S}$  splitting could really contribute to improving the performances of refineries, biorefineries and other industrial plants, since it changes the role of  $\text{H}_2\text{S}$  from waste to commodity.

#### References

- M. Binoist, B. Labégorre, F. Monnet, P. D. Clark, N. I. Dowling, M. Huang, D. Archambault, E. Plasari, P. M. Marquaire, 2003, Kinetic study of the pyrolysis of  $\text{H}_2\text{S}$ , *Industrial and Engineering Chemistry Research*, 42(17), 3943–3951.
- K. R. G. Burra, G. Bassioni, A. K. Gupta, 2018, Catalytic transformation of  $\text{H}_2\text{S}$  for  $\text{H}_2$  production, *International Journal of Hydrogen Energy*, 43(51), 22852–22860.
- B. J. Cha, J. Y. Choi, Y. Ji, S. Zhao, S. Y. Kim, S. H. Kim, Y. D. Kim, 2022, Fe-oxide/ $\text{Al}_2\text{O}_3$  for the enhanced activity of  $\text{H}_2\text{S}$  decomposition under realistic conditions: Mechanistic studies by in-situ DRIFTS and XPS, *Chemical Engineering Journal*, 443(April), 136459.
- A. G. De Crisci, A. Moniri, Y. Xu, 2019, Hydrogen from hydrogen sulfide: towards a more sustainable hydrogen economy, *International Journal of Hydrogen Energy*, 44(3), 1299–1327.
- EIA, 2019, The Future of Hydrogen, Report prepared by the IEA for the G20 Japan, Issue June.
- A. K. Gupta, S. Ibrahim, A. Al Shoaibi, 2016, Advances in sulfur chemistry for treatment of acid gases, *Progress in Energy and Combustion Science*, 54, 65–92.
- A. Jamekhorshid, Z. Karimi Davani, A. Salehi, A. Khosravi, 2021, Gas sweetening simulation and its optimization by two typical amine solutions: An industrial case study in Persian Gulf region, *Natural Gas Industry B*, 8(3), 309–316.
- V. E. Kaloidas, N. G. Papayannakos, 1987, Hydrogen production from the decomposition of hydrogen sulphide. Equilibrium studies on the system  $\text{H}_2\text{S}/\text{H}_2/\text{Si}$ , ( $i = 1, \dots, 8$ ) in the gas phase, *International Journal of Hydrogen Energy*, 12(6), 403–409.
- F. Manenti, D. Papisidero, E. Ranzi, 2013, Revised kinetic scheme for thermal furnace of sulfur recovery units, *Chemical Engineering Transactions*, 32(January), 1285–1290.
- Y. Uesugi, H. Nagakawa, M. Nagata, 2022, Highly Efficient Photocatalytic Degradation of Hydrogen Sulfide in the Gas Phase Using Anatase/ $\text{TiO}_2$ (B) Nanotubes, *ACS Omega*, 7(14), 11946–11955.
- J. Zhou, Z. Ao, T. An, 2021, DFT Study of the Decomposition Mechanism of  $\text{H}_2\text{S}$  on V-Decorated  $\text{Ti}_2\text{CO}_2$  Single-Atom Catalyst, *Wuli Huaxue Xuebao/ Acta Physico - Chimica Sinica*, 37(8), 1–9.

# Reuse and recycling of waste material: design of a pyrolysis-based process

Gabriel Rodríguez Garrido<sup>a</sup>, Pablo Marinangeli<sup>b</sup>, Jorge A. Ressia<sup>c,d</sup>,  
Maria A. Volpe<sup>c</sup>, Patricia M. Hoch<sup>c</sup>

<sup>a</sup>*Instituto Petroquímico Argentino, Avda. Córdoba 629, CABA, Argentina*

<sup>b</sup>*CERZOS – Departamento de Agronomía – UNS, Bahía Blanca, Argentina.*

<sup>c</sup>*PLAPIQUI, Chem Eng, Dept., Universidad Nacional del Sur. Bahía Blanca, Argentina*

<sup>d</sup>*Comisión de Investigaciones Científicas de Buenos Aires, La Plata, Argentina.*

[p.hoch@plapiqui.edu.ar](mailto:p.hoch@plapiqui.edu.ar)

## Abstract

This paper addresses the use of pyrolysis to reduce carbon footprint of several waste material as agricultural waste plastic (AWP). The destination of such residues was usually landfills, but there is a high potential for the obtention of high added value products after the pyrolysis process is completed, with the added benefit of reducing the carbon footprint and then accomplishing objectives of circular economy. Argentina's energetic matrix relies on the use of natural gas, which is abundant, so there is no compromise with the energy demand, however, the production of fuels from waste brings financial and environmental benefits, thus being highly attractive. Pyrolysis is a multifunctional process, and the indirect "avoided" impacts associated with the production of energy that substitute primary one can be considered (Giugliano et al., 2011).

Different paths for the treatment of the pyrolyzed material can be taken, as well as different pretreatment paths. Also, the possibility of mixing the feedstocks is considered, being the composition one of the key variables for the economics of the process. The objective function for the optimization considers environmental aspects as the reduction of CO<sub>2</sub> emissions. The potential carbon footprint emissions of pyrolysis process are calculated and compared to incineration and recycling processes. Results show that pyrolysis rises as the best alternative for reducing the carbon footprint, not only because CO<sub>2</sub> emissions are highly reduced, but also because the products are a fossil fuel replacement when the appropriate paths for pre and post treatment are used. Limitations are related to the composition of the residue to be treated.

**Keywords:** Pyrolysis, GHG emissions, circular economy

## 1. Introduction

The management and disposal of agricultural waste plastic (AWP) in the southwest region of Buenos Aires, Argentina, has become an environmental challenge. This waste is mainly composed of empty agrochemical containers and silo bags. As it is relatively easy to separate both sources of residues, silo bags can be physically recycled as they are made of polyethylene, and they do not contain any toxic component. Agrochemical containers, however, cannot be recycled with ease because they are contaminated with the pesticides or fertilizers used and no real solution has been found for the end of life of this waste. Recently, industry was interested in chemical recycling of waste plastic, mainly to decrease the carbon footprint and to be in line with circular economy (Kruger, 2020).

According to ISO 15270 (2008), pyrolysis is one of the four methods of chemical recycling, which transforms plastics back into basic chemicals by thermal decomposition at elevated temperatures, where usually all the plastic additives and contaminants are also converted back into basic chemicals. The pyrolysis products, liquid and gas, could be used as feedstock in chemical industry substituting conventional fossil-based energy or product, but technologies are still in early stages for industrial scale use. The advantages are the potential to improve the recycling rate of plastics, as well as reducing fossil fuels demand for energy and GHG emissions. All of these is aligned with circular economy principles.

The goal of this study is to analyze pyrolysis of AWP as an end-of-life treatment, comparing it with other popular options, i.e. incineration. The carbon footprint of each option will also be calculated, depending mainly on the temperature of the thermal process if applicable (Gasification, incineration, pyrolysis) but the carbon footprint for disposing the AWP in landfills will also be calculated. For the pyrolysis process the optimal process temperature, the amount of utilities used for heating (e.g. natural gas for the pyrolysis reactor, if the pyrolysis gas results insufficient as fuel for this heating requirement) and for the condensing units will be found solving a nonlinear programming problem. Being the feedstock a residual material from other economic activities the price is very low or can be neglected in some cases. A comparison between two other different treatments for the residues (incineration and mechanical recycling) is presented. Between the main objectives, the calculation of CO<sub>2</sub> emissions and carbon footprint together with energy production and economic aspects was considered. Calculations of emissions were made using the 2006 Intergovernmental Panel on Climate Change (IPCC) guide, published data and data obtained in the context of the present study. Manfredi et al (2011) show that GHG emissions are lowered using pyrolysis opposite to other thermal treatments.

## 2. Materials and methods

Experimental data are required to adjust parameters of the models used for the simulation and optimization of the process. Pyrolysis of AWP is carried out in a bench scale reactor, at 500°C, under 5cc/min of N<sub>2</sub> flow (Figure 1).



Figure 1: Pyrolysis bench scale reactor.

The products are a liquid, a gas, and a carbonaceous co-product. The products are characterized in the *context* of their use for generating thermal and electrical energy. The solid co-product was studied in order to determine the possible presence of toxics (metals, Polyunsaturated Aromatic Hydrocarbons, Polycyclic Aromatic Hydrocarbons PHAs) The characterization tools are thermal analysis (DSC/TGA), elemental analysis, lower heating value (LHV) and viscosity measurements and gas chromatographic associated to mass spectroscopy detection

(GC/Mass) analysis. CO<sub>2</sub> equivalent emissions are calculated using these results. Pyrolysis technology is compared to both mechanical recycling and incineration of plastic, considering CO<sub>2</sub> emissions. In addition, the energetic efficiency of the pyrolysis

process of AWP is calculated, considering the liquid and gas products as alternatives to fuel- oil and natural gas respectively.

The liquid and the char were analyzed to find their composition. The carbonaceous solid was practically free from transition metals, with only 120 ppm of iron. The presence of PAHs was not detected.

Liquid and gas are rich in saturated and non-saturated hydrocarbons, showing only trace amounts of oxygenated compounds. While gas is postulated to be a thermal source for the pyrolysis endothermic process, liquid is envisaged as a substitute of heavy fuel-oil, or at least for the formulation of blends with liquid fossil-based fuels.

From experimental values a simulation model is built for the pyrolysis reactor together with the mill used to chip the containers and the separation sector to obtain char, liquor and gas. Mass and energy balances are posed for all the equipment. This model is used for the optimization of the process. To model the process, yields of liquid, gas and char were correlated vs. temperature of reaction, so this temperature could be used as optimization variable.

The yields % as a function of the temperature are then used to find the flowrates of char, liquid and gas products. Those are found from the experimental values as functions of temperature  $T[^\circ\text{C}]$  as:

$$\text{Liquid yield \%: } y_{LIQ} = -1.50\text{E-}03 T^2 + 1.47 T - 2.99\text{E+}02 \quad (1)$$

$$\text{Char yield \%: } y_{CHAR} = -1.00\text{E-}05 T^3 + 1.63\text{E-}02 T^2 - 8.85 T + 1.60\text{E+}03 \quad (2)$$

$$\text{Gas yield \%: } y_{GAS} = 1.65\text{E-}03 T^2 - 1.61 T + 4.29\text{E+}02 \quad (3)$$

Char yield has been approximated with a cubic function as it presents a plateau between 500-550 °C, to represent with fidelity the experimental behavior. A process schematic is shown in Figure 2. For illustrative purposes only one condenser is shown, but due to the high temperatures of the gas three units are used, with different cooling media (Dowtherm© for the first heat exchanger, where the gas is cooled with no phase change, Dowtherm© to condense up to 100°C, and cooling water for the subsequent unit. Final temperature of the process streams leaving the condenser is approximately 80°C).

Global mass balances for the pyrolysis products:

$$M_j = M_{feed} y_j / 100; j = LIQ, GAS, CHAR \quad (4)$$

being  $M_j$  the amount of liquid, gas or char obtained from the pyrolysis reactor after the condensation step takes place.

Energy balances for the process are also posed, together with the equations that allow calculating the energy requirements of the rest of the equipment.

### 3. Results

The analysis of the scenario corresponding to the accumulation and location of AWP in the region under study indicates that a plant for processing 4t per day is feasible, so this will be used as a calculation basis ( $M_{feed} = 4$  t/day).

Equipment dimensions were calculated for the requirements of the process. The condensing unit is composed of several heat exchangers as the gas leaves the pyrolysis unit at the pyrolysis temperature (between 450-650 °C) as different refrigeration utilities will be used in each step. This total cost of the process can be calculated using this information.

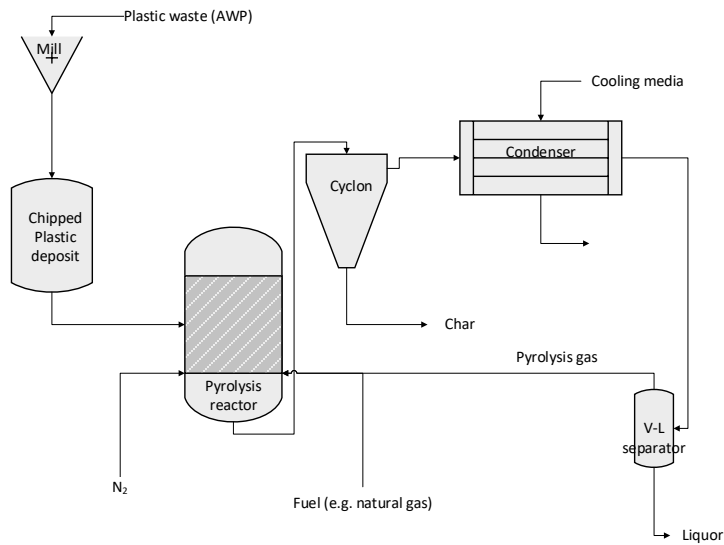


Figure 2: Process schematic for a standard pyrolysis process.

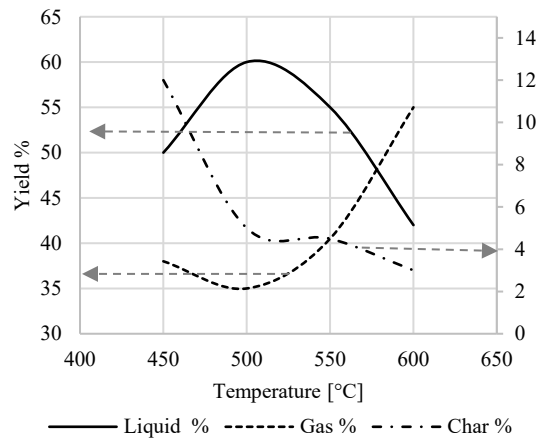


Figure 3: Yield of products from pyrolysis

Figure 3 shows the global liquid, gas and char yields for different pyrolysis temperatures as presented in Eqs. 1 to 3. The liquid product is maximized for a temperature around 500°C. It will be assumed for the purpose of obtaining the process model that the yields for a pilot scale reactor will be like the experimental ones. For approximately 500 °C, which resulted to be the optimum temperature for maximizing the liquid production, it was obtained experimentally that LHV was 49,6 MJ/kg for the liquid and 46,9 MJ/Kg for gas. It was found that those values do not change significantly between process temperatures ranging from 450 to 600°C as the amount of products change with temperature, but not the composition. Gas LHV is used to calculate the heat generated for the pyrolysis process. This could replace partially the equivalent of fossil-based fuel-oil traditionally used for energy production in thermoelectric plants in the region, with a consequent decrease in CO<sub>2</sub> eq. emissions (7.74 t of CO<sub>2</sub> avoided per day).

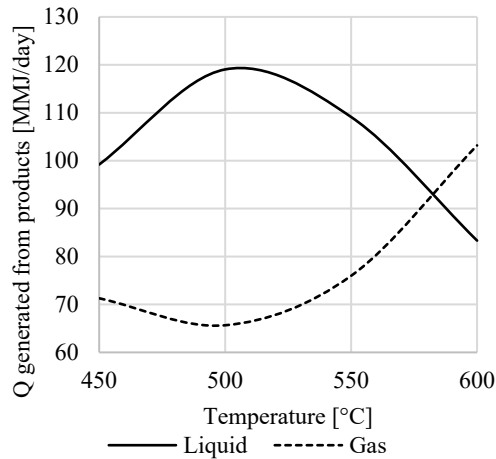


Figure 4: Energy obtained from the products

Table 1: Products distribution for the whole process

T (°C)	Liquid [kg/day]	Gas [kg/day]	Char [kg/day]
450	2000	1520	480
500	2400	1400	200
550	2200	1620	180
600	1680	2200	120

The incineration of AWP, without energy recovery produces 4,61 [t CO<sub>2</sub>/t of plastic]. (O. Eriksson and G. Finnveden, 2009). Thus, the pyrolysis gas, used as the energy source of the recycling process originates 3.76 t CO<sub>2</sub> / t AWP per day, being the net CO<sub>2</sub> emission -4.0 t CO<sub>2</sub> per day.

For our process it is found that 4 t CO<sub>2</sub>/day are avoided compared to incineration.

Regarding the use of the products as energy source, from Figure 4 it is shown the compromise between the production of liquid (which can be further used to produce higher added value products or can be injected in oil ducts as its characteristics are similar to light crude) and the production of gas that can be used as a fuel substitute to process the AWP. In any case, it must be noticed that producing plastics requires less energy than disposing them. Regarding the Sustainable Development Goals, it is tempting to conclude that the solution for the problem of pollution due to plastics should be banning them. But this will collapse with another objective of Zero hunger, as the petrochemical industry is tightly linked to the food related industry, apart from being counterproductive (Creadore and Castaldi, 2022)

#### 4. Alternatives for processing

The study has been conducted using chipped agrochemical containers, made from polyethylene. Alternative sources for the feedstock can be considered:



1. Gas can be used to generate energy for the reactor instead of using fossil-based fuel, as shown in this paper
2. Liquid can be used as a fuel for the reactor, in which case there will be a remnant of gas.
3. As the amount of energy required for the pyrolysis process can be between 4 and 10 MJ/t, an analysis can be made of the external requirements as the energy required increases (for example when more gas is produced to decrease or even eliminate the need of an external source for fuel).
4. If there is biomass present in the residue (sometimes the containers are labeled and the paper is not separated beforehand, so it is chipped together with the plastic giving a mixed residue) LHV decreases. For example, for 10% biomass, LHV decreases 5%.

Several scenarios can be included within the model to compare the emissions and the cost of the process. If the objective is to decrease emissions, the objective function is to minimize t CO<sub>2</sub>/t AWP as shown in this work.

## 5. Conclusions

Chemical recycling by pyrolysis represents a solution to the environmental problem associated with the end of life of AWP.

The impacts on environment are positive since liquid and gas products generate energy with a negative CO<sub>2</sub> emission balance.

Emissions due to incineration are avoided and at the same time, replacement of fossil-based fuel by pyrolysis liquid originates a significant decrease in CO<sub>2</sub> emissions.

Apart from the numerical results, from the experiments the carbonaceous co-product has no toxic components, thus it could be disposed in soils without producing negative impacts. These results can be compared to the other possible processes to dispose AWP,

- Landfills: 0,25 kg CO<sub>2</sub>/kg plastic,
- Waste-to-energy processes 2,20 kg CO<sub>2</sub>/kg plastic.
- Mechanical recycling emissions are alike to those of landfills, 0,2 kg CO<sub>2</sub>/kg plastic.

If the pyrolysis gas is used to generate electricity, emissions are neglected as it is used to replace a fossil fuel.

Studies are being conducted to analyze the impact of this solid product from pyrolysis on the crops when used to improve soil which are beyond the scope of this work. Future work is also related to the development of processes for adding value to the liquor. Also, studies for scaling up the process are being conducted, because of the availability of data.

## References

- Creadore, L, Castaldi M, 2022 Quantitative Comparison of LCAs on the Current State of Advanced Recycling Technologies. Report from Earth Engineering Center Chemical Engineering Department, The City College of New York.
- Eriksson O., G. Finnveden, 2009, "Plastic waste as a fuel - CO<sub>2</sub>- neutral?", *Energy Environ. Sci*, 2, 907-914
- Giugliano, M., Cernuschi, S., Grosso, M., Rigamonti, L., 2011. *Waste Management*. 31, 2092–2101.
- ISO 15270, 2008.
- Kruger C., BASF (Sphera Solutions GmbH), 2020, Evaluation of pyrolysis with LCA: 3 case studies.
- Manfredi, S., Tonini, D., Christensen, T.H., 2011., *Conserv. Recycl.* 55, 995–1004.

# Life Cycle Assessment of bio-oil produced through lignocellulosic biomass liquefaction

Ana S. Augusto<sup>a</sup>, Catarina G. Braz<sup>a</sup>, Duarte M. Cecílio<sup>a,b</sup>, Maria M. Mateus<sup>a,b</sup>,  
Henrique A. Matos<sup>a</sup>

<sup>a</sup>*Centro de Recursos Naturais e Ambiente, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal*

<sup>b</sup>*Fábrica Secil - Outão, Apartado 71, 2901-864 Setúbal, Portugal*  
*catarina.braz@tecnico.ulisboa.pt*

## Abstract

The objective of the present work was to assess the impact on the environment of a biofuel produced through lignocellulosic biomass liquefaction through a Life Cycle Assessment (LCA). Using *SimaPro* software, the Product Environmental Footprint (PEF) methodology was used to calculate the global environmental impact. The system was expanded to consider the credits of petroleum coke replacement. The effect of the type of solvent used in the liquefaction process was studied, as well as the possibility of operating the pilot plant in continuous mode.

**Keywords:** Biomass liquefaction, biofuels, environmental impact, Life Cycle Assessment, *SimaPro*.

## 1. Introduction

In response to the challenges of climate change, the European Green Deal (2019) set the goal of transforming the European Union (EU) into a resource-efficient and carbon-neutral economy, mobilising the industry for a clean and circular economy. In this context, the EnerGreen project, developed by Cimentos Maceira e Pataias, S.A., built a pilot plant to produce advanced biofuels from biomass-derived waste to replace the usage of petroleum coke in a clinker's kiln.

The main objective of this work was to carry out an LCA study that analysed the environmental impact of bio-oil production through the liquefaction of lignocellulosic biomass using the *SimaPro* software. The effect of the life cycle stages of the bio-oil was studied and evaluated to define which steps of the process are the main contributors to the environmental impact and which impact categories are the most affected. Two different biomass sources were compared, eucalyptus and sludges from the paper industry, at BATCH and later at continuous operating conditions.

The LCA method used for this purpose was the Product Environmental Footprint (PEF), available in *SimaPro* software. This methodology is the most suitable to use in this study since its results consider the European citizens' average impact as a reference.

## 2. Methods

### 2.1. Goal and Scope

The LCA study presented in this work has been conducted according to the ISO 14040/44 guidelines (2006a, 2006b), following an attributional approach. The study was conducted in a cradle-to-gate scope, considering the operation of the units (material and electric and thermal energy usage). On the other hand, energy and materials related to the construction of the processing units are not included in data inventories.

The inventory data was obtained from the biomass liquefaction pilot plant installed at Secil – Maceira e Pataias to produce 1GJ of energy. The Ecoinvent 3.6 database is used for the background data assuming the Portuguese context.

### 2.2. Life Cycle Inventory

In the liquefaction process (Figure 1), the biomass is fed to an endless screw, undertaking a pre-treatment process. Here, the biomass is pre-heated and sprayed with the solvent, 2-ethyl hexanol (2EH), undergoing a process of swelling, where the volume of the cells increases, breaking the cellulosic and lignocellulosic membranes, facilitating the access of the catalyst to all components of the residue. The liquefaction process takes place inside the stirring reactor at a temperature between 130 and 250 °C in the presence of an acid catalyst, p-toluene sulfonic acid.

The reactor is equipped with an external jacket, where the gases from liquefaction (mainly steam) circulate. The gases also go through a hollow conveyor outside the endless screw to pre-heat the feed and are condensed afterwards.

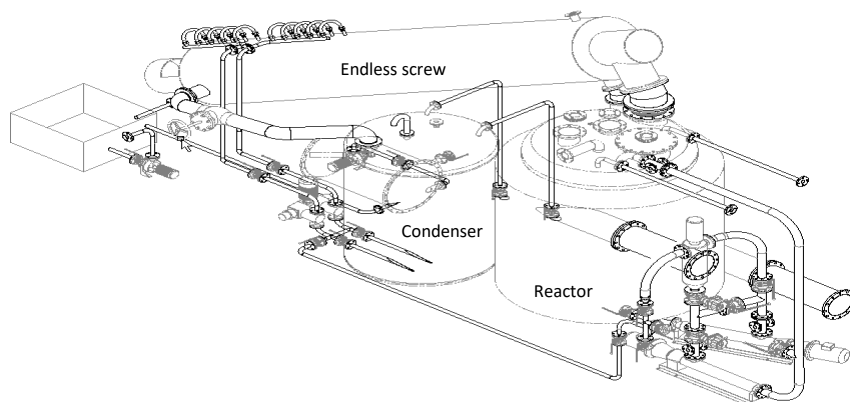


Figure 1 – Biomass liquefaction facility installed in Secil – Maceira e Pataias.

The lower calorific value (LCV) of the liquefied biomass was calculated as the pondered average of the components' LCV present in the reactive mixture and used to assess the amount of bio-oil equivalent to 1 GJ of energy. The results are presented in Table 1, together with the values for petroleum coke.

Table 1 - Mass of fuel to obtain 1 GJ of energy.

	LCV (GJ/ton)	Mass (kg) to 1 GJ
Bio-oil from Eucalyptus	21.1	47.5
Bio-oil from Paper Sludges	15.7	63.7
Petcoke	31.3	32.0

Table 2 presents the data inventory introduced in *SimaPro* to produce bio-oil from eucalyptus and paper sludges in BATCH and continuous operating conditions. Since biomass sources are residues, they were considered to have zero impact and weren't introduced in *SimaPro*.

Table 2 – Data introduced in *SimaPro* for the liquefied biomass from eucalyptus and paper sludges for BATCH and continuous mode.

Inputs and outputs	Eucalyptus		Paper Sludges	
	BATCH	Continuous	BATCH	Continuous
<i>Inputs</i>				
Solvent (kg)	29.8	3.7	208.7	3.9
Catalyst (kg)	1.1	0.5	4.0	0.4
Fuel (kg)	0.8	6.8	1.2	0.8
Transport (tkm)	5.3	43.3	15.9	10.1
Electricity (kWh)	2.0	16.3	3.0	1.9
<i>Outputs</i>				
Bio-oil (kg)	47.5	47.5	63.7	63.7
Condensed water (l)	19.9	210.3	139.1	132.8
<i>Avoided products</i>				
Petroleum coke				32.0

### 2.2.1. Introduction of p-toluene sulfonic acid and 2-ethyl hexanol in *SimaPro*

As p-toluene sulfonic acid and 2EH are not available in the Ecoinvent database, these two components' environmental impact must be estimated. To create this molecule, the RREM method was used (Huber et al., 2022), which consists of four steps: general research on the chemical process, setting up the reaction equation, researching the required thermal energy, and modelling the dataset. Toluene sulfonic acid, the catalyst, is synthesised from a reaction between toluene and sulfuric acid with a molar ratio of 1:1.



The values introduced in *SimaPro* (Table 3) correspond to their respective molar masses, considering a yield of the reaction of 100 %. As the energy for this reaction was not found in the literature, it was assumed that 1 kg of product needed 1.8 MJ to be created (Huber et al., 2022).

Based on the synthesis of 2EH via butanol (BuOH), the same procedure applied to the catalyst was used. The synthesis of 2EH via butanol through the Guerbet reaction starts with a dehydrogenation of butanol followed by aldol condensation and hydrogenation of the unsaturated condensation product. A bifunctional catalyst, characterised by a dehydrogenating/hydrogenating metal species and a basic component, is required for this reaction to occur. Homogeneous catalysts at 120°C based on phosphine complexes of different transition metals (Rh, Ru, Ir, Pt, Pd, and Au) are used in the Guerbet self-condensation of BuOH.



Based on stoichiometric production reactions of p-toluene sulfonic acid and 2EH and molar masses, the following inputs in Table 3 were inserted in *SimaPro*.

Table 3 – Inputs and outputs of p-toluene sulfonic acid and 2-ethyl hexanol in *SimaPro*.

p-toluene sulfonic acid system		2-ethyl hexanol system	
<i>Inputs</i>		<i>Inputs</i>	
Toluene	92.14 g	Palladium	0.58 mg
Sulfuric acid	98.08 g	1-butanol	74.12 g
Heat from steam	0.05 MJ	Heat from steam	0.7 MJ
<i>Outputs</i>		<i>Outputs</i>	
p-toluene sulfonic acid	172.02 g	2-ethyl hexanol	130.23 g
Wastewater (water emission)	1.25 ml	Hydrogen (air emission)	2 g
		Wastewater (water emission)	18 g

### 3. Results

The environmental impact of the liquefied biomass was evaluated using the PEF methodology, and the results are presented in Figure 2. The total impact score of both cases is positive. Since the petcoke replacement credits were considered, this means that the impact of both bio-oils is higher than that of petcoke, and thus, they are not suitable as a substitute.

Bio-oil derived from eucalyptus presents a much lower impact score than the one derived from paper sludges. This fact is justified mainly because the first presents a much higher LCV than the latter, meaning that more significant amounts of paper sludge-derived bio-oil are necessary to produce 1 GJ.

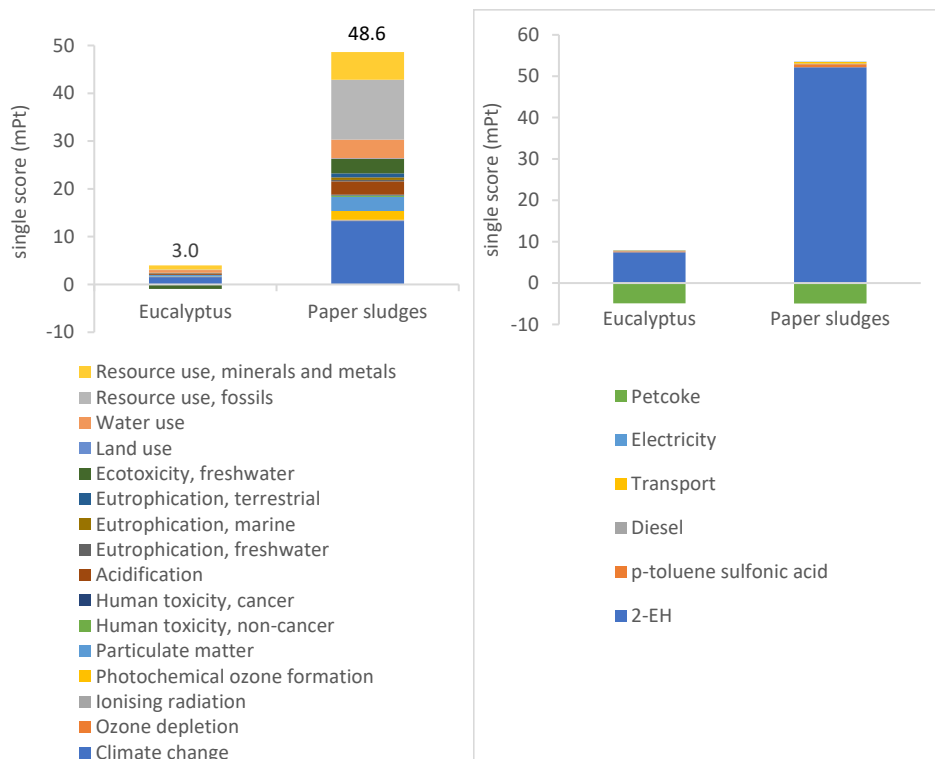


Figure 2 – Comparison of the weighting values of impact categories of liquefied biomass with PEF method detailed by impact categories and contribution of processes with the petcoke credits.

In both cases, the impact categories that stand out the most are the use of fossil resources, climate change and the use of minerals and metals resources. These results are due to the manufacturing process of 2EH, which uses minerals, metals and fossil resources, and consequently will release prejudicial compounds to the environment that contribute to climate change.

### 3.1. Influence of glycerine in the solvent

Since the solvent used in the liquefaction process is the main contributor to the environmental impact of the bio-oil produced, new alternatives were searched to substitute 2EH. A component that may be part of the solvent is glycerine, which influences the solvent's behaviour. Using glycerine with 2EH improves the homogenisation of the medium and makes it possible to increase the process conversion and obtain added-value products with nitrogen. Glycerine derived from waste cooking oil biodiesel plants was selected since it had the lowest impact of all the available options in *SimaPro*.

Figure 3 presents the results of the environmental impact score of bio-oil with different proportions of 2EH and glycerine mixtures.

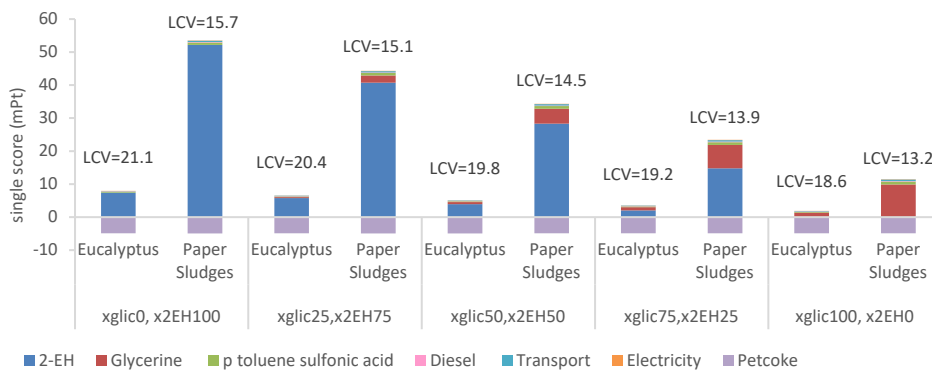


Figure 3 – Impact of glycerine from waste cooking oil on the bio-oil environmental impact. LCV is in GJ/ton.

As expected, the substitution of 2EH is beneficial to decrease the environmental impact of the liquified biomass. On the other hand, adding glycerine reduces the LCV of the resulting bio-oil, which might limit its utilisation in a clinker's kiln.

### 3.2. Continuous operation

A new environmental impact evaluation was performed, this time considering the liquefaction process operates in continuous mode (currently, the liquefaction pilot plant operates in BATCH), where less solvent is necessary per kg of bio-oil.

The results obtained in continuous operation (Figure 4) have a lower impact, associated with other benefits like less labour-intensive or less processing and holding time.

Considering the environmental impacts of bio-oil, the liquified biomass from eucalyptus has a higher impact. This is a consequence of the smaller amount of solvent used per kg of bio-oil.

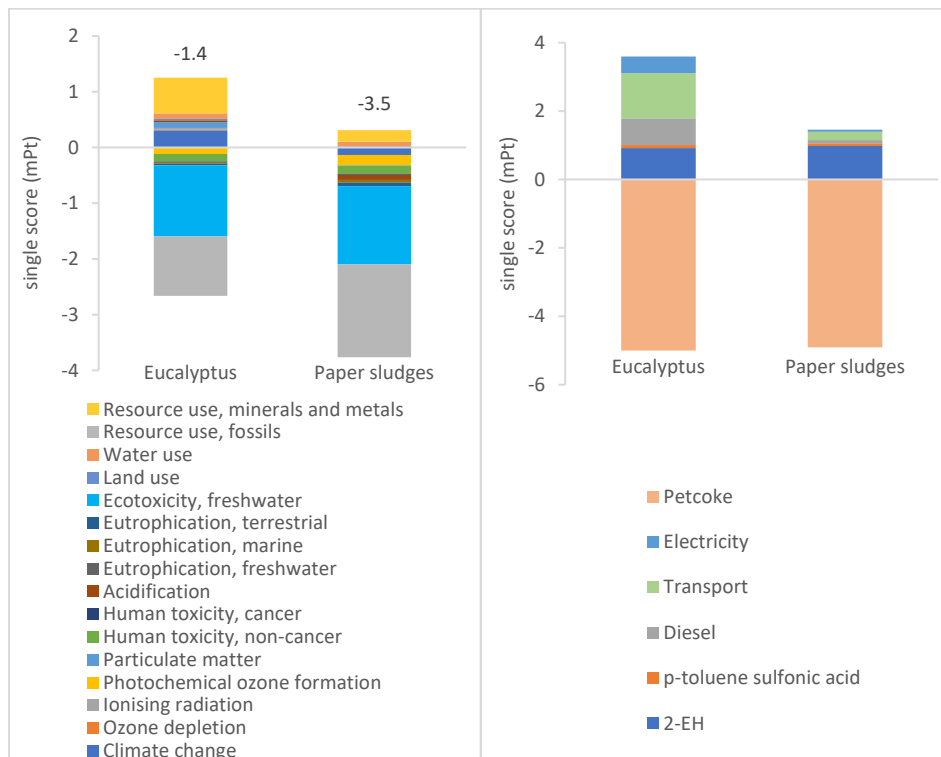


Figure 4 – Comparison of the weighting values of impact categories of liquefied biomass with PEF method detailed by impact categories and contribution of processes with the petcoke credits.

#### 4. Conclusions and Future Work

The main objective of this work was to evaluate the impact of the liquefied biomass production process. Different scenarios, such as working conditions, biomass source and glycerine presence, were studied. Based on the analysis, it can be concluded that it is possible to obtain a bio-oil suitable to replace petroleum coke if the unit is operating continuously. Eucalyptus residues are preferred as biomass sources because they form a bio-oil with a higher LCV. Adding glycerine as solvent up to 75 % is beneficial for eucalyptus-derived liquefied biomass, reducing the environmental impact while maintaining a reasonable LCV of the bio-oil. In the future, more studies on the impact of p-toluene sulfonic acid and 2EH should be done to decrease the uncertainty of the results.

#### References

- 2006a, ISO 14040:2006 Environmental Management – Life Cycle Assessment – Principles and Framework, International Standardization Organisation, Geneva.
- 2006b, ISO 14044:2006 Environmental Management - Life Cycle Assessment – Requirements and Guidelines, International Standardization Organisation, Geneva.
- 2019, The European Green Deal
- C. Carlini, A. Macinai, A. M. Raspolli Galletti, and G. Sbrana, 2004, Selective synthesis of 2-ethyl-1-hexanol from n-butanol through the Guerbet reaction by using bifunctional catalysts based on copper or palladium precursors and sodium butoxide, *J Mol Catal A Chem*, vol. 212, no. 1–2, pp. 65–70
- E. Huber, V. Bach, P. Holzapfel, D. Blizniukova, M. Finkbeiner, 2022, An Approach to Determine Missing Life Cycle Inventory Data for Chemicals (RREM), *Sustainability*, 14, 3161

# Impacts of defossilising downstream derivatives in petrochemical clusters – MTBE case study

Inna Stepchuk<sup>a</sup>, Mar Pérez-Fortes<sup>a</sup>, Andrea Ramírez Ramírez<sup>a</sup>

<sup>a</sup>*Department of Engineering Systems and Services, Faculty of Technology, Policy and Management, Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands*  
[i.stepchuk@tudelft.nl](mailto:i.stepchuk@tudelft.nl)

## Abstract

Using alternative carbon sources (ACS) to produce downstream derivatives (DDs) is a promising option for defossilising the chemical industry. However, the potential consequences of using ACS in interconnected petrochemical clusters are generally overlooked. This paper aims to develop a methodological approach for systematically analysing defossilisation impacts at the value chain level. For this, a single value chain for producing methyl-tert-butyl-ether (MTBE) was used as a case study. The individual components of the value chain were modelled in Aspen Plus v12. Both ACS- and fossil-based value chains were compared in terms of (i) changes in the structure of the value chain and (ii) the magnitude of the impacts. The results show that the defossilisation of a single value chain causes additional impacts at the cluster level.

**Keywords:** Industrial defossilisation, downstream derivatives, interconnected cluster.

## 1. Introduction

The chemical sector is highly dependent on the usage of fossil carbon as feedstock for the production of chemicals (Nesbitt 2020). The launch of the EU chemicals strategy for sustainability has triggered the chemical industry to look for innovative solutions toward a green transition. One of the possible solutions is to defossilise the chemical industry by replacing fossil-based feedstock with alternative materials (EC 2020). Alternative carbon sources (CO<sub>2</sub>, biomass, waste) are increasingly considered as alternative feedstocks for bulk chemicals production (i.e. ethanol, methanol, ammonia). However, there is less research on the potential of defossilising downstream derivatives (DDs) (Moncada, Posada, and Ramírez 2015). DDs are the chemicals present at the end of value chains before they are distributed to potential customers, for instance, styrene, acetone, or polyvinyl chloride (Kramer et al. 2017).

DD's value chains are usually integrated into large petrochemical clusters (Porter 1998), through symbiotic relationships inside and between value chains within the clusters (Chertow 2000). The defossilisation of one DD process will likely affect other processes inside the value chain and within the cluster. Thereby, making uncertain what the total impact on the performance of the value chain and cluster could be. The goal of the current work is to develop and test a methodology to analyse how using alternative carbon materials for defossilising a single DD value chain can affect its existing interactions and performance in a petrochemical cluster. For this paper, MTBE (methyl tert-butyl ether) was selected as a case study.



## 2. Methodology

This paper departs from an in-house model of a representative petrochemical cluster. The model mimics a real petrochemical cluster and it is based on the 6 industrial sites of the Port of Rotterdam (PoR) petrochemical cluster (Ramirez-Ramirez 2019). The model includes a total of 22 plants with the production of 52 chemicals classified in 9 subclusters. Each plant was modelled in Aspen Plus v12, and the interconnections are modelled in an interface. The subclusters are chlorine, methanol, ethylene, propylene, ammonia, aromatics, olefins, bio-based and other chemical-based subclusters.

The representative cluster contains a diverse range of chemicals. In this paper, chemicals produced in the cluster are presented as follows (Figure 1): feedstock (F), auxiliary chemicals (AC), chemical building blocks (CBBs), commodity chemicals (CCs), and downstream derivatives (DDs). All these chemicals are interconnected by mass, energy and or waste flows. The connectivity of the value chain is represented by the number of mass flows interconnections within value chain boundaries (i.e. horizontal connections) and the number of subclusters involved in the production of a single DD (i.e. vertical connections). The methodology contains four main steps which are discussed below.

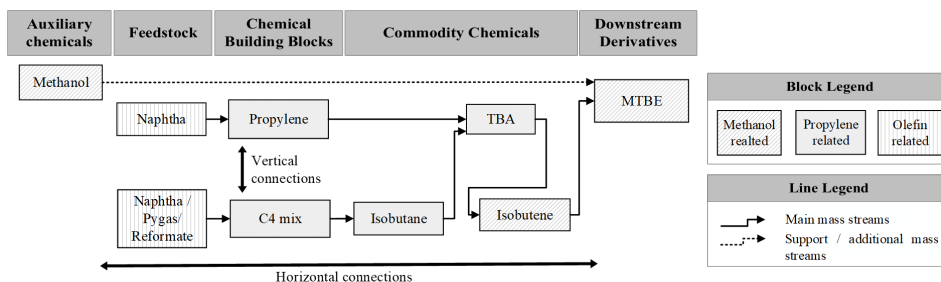


Figure 1: Schematic representation of MTBE value chain with existing mass flow interconnections.

### 2.1. Selection of key downstream derivatives

The representative cluster contains 13 DDs but not all of them have a large potential for defossilisation. Therefore, the selection of the most promising DDs was made based on three criteria: (i) embodied carbon, (ii) connectivity of the components in the value chains and (iii) availability of data of the ACS-based routes in the literature.

First, the amount of embodied carbon was assessed for each DD to identify the most promising derivatives for defossilisation. Embodied carbon (EC) is defined as the elemental carbon embodied in material flows of chemicals entering specific stages of production to meet process needs (Eq.(1)). For the assessment of embodied carbon, data was collected from each DDs' Aspen Plus model of the representative petrochemical cluster. The embodied carbon of each DD was calculated as follows:

$$EC_{DD} = \sum_{i=1}^n M_i * C_{wt\%_i} \quad (1)$$

where  $i = (1 \dots n)$  – number of material flows entering the production process of the DD;  $M_i$  – inlet capacity of particular material flow, ktonne/year;  $C_{wt\%_i}$  – weight percent of fossil-based carbon embodied in the chemical, %.

The next step involved identifying the existing connectivity of mass flows within value chains and representative clusters. This was done by mapping the network, i.e.,

identifying the interconnections (i.e. mass flows) between chemicals, and the connectivity of each value chain was assessed in terms of horizontal and vertical interconnections.

Last, a literature screening was performed to identify ACS-based production routes (i.e. from CO<sub>2</sub>, waste, biomass). Information was collected for each ACS-based route with a technology readiness level (TRL) of 4 or higher.

Based on three selection criteria, potential DDs were selected for further research. To evaluate the available defossilisation pathways (i.e. ACS-based routes) and develop the methodology to assess subsequent impacts at the value chain level, first, a single DD value chain (i.e. case study) was selected.

### *2.2. Selection of ACS-based route for the case study*

This step involved the selection of the most promising ACS-based production routes for the case study. Routes were further screened in terms of data availability (i.e. the more data, the better), type of options (i.e. thermochemical, electrochemical, biochemical and the process conditions, e.g., mild conditions were preferred over stringent conditions), complexity (i.e. the number of process steps, the fewer steps the better), and scalability. The routes were then ranked and the most promising option was selected for further study.

### *2.3. Process modelling of ACS-based route*

The selected ACS-based route was then modelled at the scale (i.e. capacity) and purity of the current fossil process. Here it is assumed that the alternative route will replace one-to-one the amount of product produced by the fossil-based route. For this, the development of the process flow diagram and Aspen Plus v12 model was done based on data collected for the selected ACS-based route. The model was assumed to be in continuous mode with a whole year of operation (i.e. 8000 hours per year). Properties for biomass components were introduced into Aspen Plus simulation by the usage of biomass property databases developed by the National Renewable Energy Laboratory (NREL).

### *2.4. Defossilisation impacts analysis*

Value chains can be represented graphically as a network, with nodes and links. Nodes refer to chemical processes involved in the value chain, while links represent material, waste and or energy flows between the nodes. Networks were developed for fossil- and ACS-based value chains of the case study, which allow to study changes that would occur when replacing a fossil-based process, a node, in the network with ACS-based process. Structural changes in nodes and links were categorized, as (i) removed - no longer required, (ii) affected - required, but capacities are changed, (iii) unchanged - required, capacities remain the same, and (iv) added – required, to be added (i.e. new chemicals/flows).

It is expected that structural changes in the value chain will affect its total performance. To quantify the magnitude of these impacts the following techno-economic and environmental KPIs were selected: (i) production rates, (ii) fossil-based carbon offset (i.e. the amount of fossil embodied carbon replaced by alternative carbon), and (iii) additional investment needed. Based on these indicators, the ACS-based value chain of the case study was assessed and compared to the fossil-based counterpart.

## **3. Results and discussion - MTBE case study**

### *3.1. Selection of case study – MTBE*

Based on the methodology for the selection of DDs presented in Section 2.1, 6 DDs were selected (see Table 1). First, priority was given to DDs with embodied carbon of

> 40 ktonne/year. Then, chemicals were examined in terms of connectivity and availability of ACS-based routes. DDs with a higher number of vertical connections were preferred because changes in the value chains of those DDs will more likely result in a higher probability of impacts at the cluster level. MTBE was selected as a case study. As can be seen from Table 1, it has high EC, is highly interconnected within the value chain and cluster, and has a wide range of ACS-based routes present in the literature.

Table 1: Key downstream derivatives.

Downstream Derivatives <i>name</i>	EC* <i>kt/y</i>	Connectivity		ACS-based routes			Research <i>selected</i>
		<i>vertical</i>	<i>horizontal</i>	<i>CO<sub>2</sub></i>	<i>biomass</i>	<i>waste</i>	
Styrene	914	3	11	-	+	+	+
Methyl tert-butyl ether (MTBE)	281	2	8	+	+	+	+
Polyvinyl chloride (PVC)	181	2	6	+	+	+	+
Polyethylene terephthalate (PET)	155	2	8	+	+	+	+
Carbon Black	132	0	1	+	+	+	-
Phthalic anhydride	71	0	3	-	+	-	-
Propylene glycol ether (PGE)	58	4	14	+	+	+	+
Polyols	44	4	13	+	+	-	+
Dimethyl ether (DME)	35	0	2	+	+	+	-
Acetone	33	0	4	+	+	+	-
Biodiesel	4	0	4	n.a.	n.a.	n.a.	-
Hydrochloric acid (HCl)	0	0	8	n.a.	n.a.	n.a.	-
Bioethanol	0	0	2	n.a.	n.a.	n.a.	-

\*secondary data for calculations is taken from Aspen models

### 3.1.1. Fossil-and ACS-based MTBE

MTBE is used in the fuel industry as an additive to increase the octane number. Conventionally, it is produced from a chemical reaction of methanol and isobutene. In the representative cluster, MTBE is produced in the amount of 400 ktonne/year. For its production, methanol and C4 mix are purchased from outside the cluster. Propylene is obtained from the cluster's olefin plant. Isobutane and tert-butyl alcohol (TBA) are produced within the propylene subcluster of the representative cluster (Section 2).

From the literature, the biomass-to-isobutanol-to-MTBE synthesis process was selected based on economic and technical constraints. The process involves different stages: (i) conversion of biomass to sugars through ethanol-based organosolv, (ii) enzymatic hydrolysis and fermentation of isobutanol, (iii) dehydration of isobutanol to isobutene, and (iv) conversion of isobutene to MTBE (Moncada, Posada, and Ramírez 2017). Ethanol is required as a solvent in a quantity of 5 L per kg of dry biomass and is completely recycled in the process. Simplified block diagrams of both fossil- and ACS-based value chains of MTBE are presented in Figure 2.

### 3.2. Defossilisation impacts analysis

Figure 2 shows a qualitative comparison of major changes in the structure of mass flows of MTBE value chain after defossilisation. Nearly 60% of the structure of MTBE value chain is changed. TBA, isobutane, C4 mix, and propylene are no longer required inside value chain and therefore their use is reduced by 100%. However, an additional chemical, ethanol, is needed as a solvent for isobutanol production. Although the quantity of ethanol is high, it is fully recovered within the value chain. Here, it was assumed that it does not affect the value chain carbon offset rate. The fossil-based carbon offset was calculated as the amount of alternative carbon replacing fossil embodied carbon of MTBE. The latter is a sum of fossil-based carbon embodied in material flows of isobutene and imported methanol entering the MTBE process. The fossil-based carbon offset is 80 %, as fossil-

based isobutene is completely replaced by an alternative one. Economic performance is also affected. Based on ongoing modelling work, additional investment is roughly 15% of the bare equipment costs of MTBE fossil-based value chain.

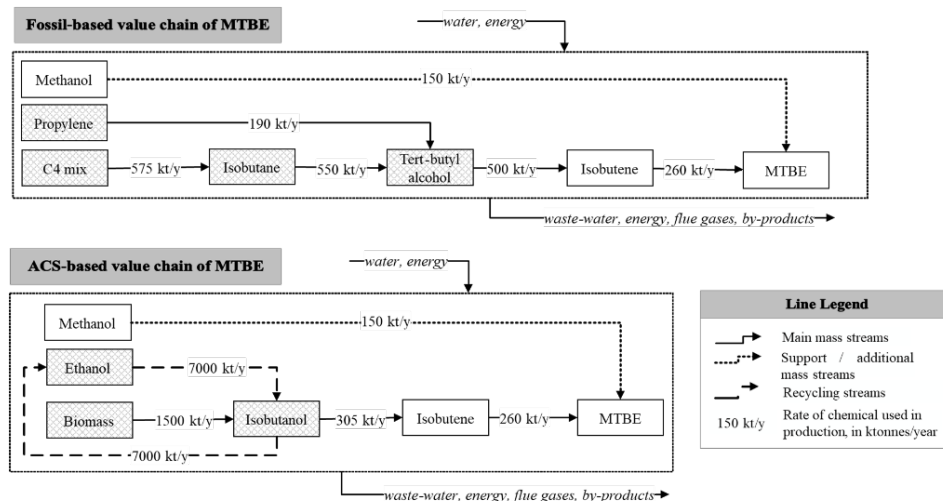


Figure 2: Simplified MTBE value chains with main mass flow interconnections for the production of 400 ktonne/year.

The majority of chemicals involved in fossil-based MTBE production are currently produced within the cluster and are involved in the value chains of other DDs (Table 2). Therefore, structural changes in the MTBE value chain will affect other value chains. For instance, tert-butyl alcohol is derived as a coproduct of propylene oxide production, requiring both C4 mix and propylene as feedstock materials. The defossilisation of the MTBE value chain will therefore affect the amount of C4 mix imported into the cluster as well as the isobutane and propylene capacities. Thereby impacting other DDs, such as styrene, and propylene glycol ether which also require propylene. The latter are highly interconnected within the cluster, resulting in further propagation of defossilisation impacts.

Table 2: Structural changes at the value chain and cluster level.

Chemical <i>name</i>	Present in the cluster				Required for MTBE value chain		Defossilisation impacts	
	<i>yes/no</i>	<i>*</i>	<i>**</i>	<i>kt/y</i>	<i>fossil</i>	<i>ACS</i>	<i>value chain</i>	<i>cluster</i>
Methanol	yes	B	no	150	150	150	unchanged	unchanged
Propylene	yes	P	yes	497	190	-	removed	affected
C4 mix	yes	B	yes	575	575	-	removed	affected
Isobutane	yes	P	yes	550	550	-	removed	affected
Tert-butyl alcohol	yes	P	no	603	500	-	removed	affected
Isobutene	yes	P	no	260	260	260	unchanged	unchanged
Biomass	yes	B	yes	2000	-	1500	added	affected
Ethanol	no	B	no	-	-	7000	added	added
Isobutanol	no	-	-	-	-	305	added	added

*\*B – is bought from the market, P- is produced within the representative cluster;  
\*\* involved in value chains of other DDs, yes/no.*

The impact assessment in this paper was done based on the changes in mass flows. However, to develop the full picture of potential impacts, it is crucial to also include impacts on energy and water flows due to defossilisation. This is part of the current research, which will be available at the time of the conference.

#### 4. Conclusions and future work

The defossilisation of a single DD can significantly have impacts beyond the process itself, especially in highly interconnected petrochemical clusters. This paper proposed a novel methodology for systematically analysing defossilisation impacts at the value chain level. The findings highlight the need for methodologies to be able to assess changes at the value chain level and their consequences at the cluster level (e.g. impacts on fossil-based carbon offset of other DDs due to defossilisation of MTBE). In future work, the ACS-based MTBE value chain will be assessed in terms of changes in energy and water flows and the methodology will be further developed to identify where the largest impacts in terms of defossilisation can be obtained in different DD value chains.

#### Acknowledgements

This publication is part of the project “Unravelling the impacts of using alternative raw materials in industrial clusters”, project number VI.C.183.010 of the research programme VICI which is financed by the Dutch Research Council (NWO).

#### References

- Chertow, Marian R. 2000. “INDUSTRIAL SYMBIOSIS : Literature and Taxonomy.” *Industrial symbiosis* 25(November): pp 313-337.  
<https://www.annualreviews.org/doi/pdf/10.1146/annurev.energy.25.1.313>.
- EC. 2020. Brussels, 14.10.2020 COM(2020) 667 final *Chemicals Strategy for Sustainability Towards a Toxic-Free Environment*. Brussels.
- Kramer, Jan-Philipp et al. 2017. no *Trilateral Strategy for the Chemical Industry*. Düsseldorf. [www.mwide.nrw.de](http://www.mwide.nrw.de).
- Moncada, Jonathan, John A. Posada, and Andrea Ramirez. 2015. “Early Sustainability Assessment for Potential Configurations of Integrated Biorefineries. Screening of Bio-Based Derivatives from Platform Chemicals.” *Biofuels, Bioproducts and Biorefining* 9(3): 722–748.  
<https://onlinelibrary.wiley.com/doi/epdf/10.1002/bbb.1580>.
- Moncada, Jonathan, John A. Posada, and Andrea Ramirez. 2017. “Comparative Early Stage Assessment of Multiproduct Biorefinery Systems: An Application to the Isobutanol Platform.” *Bioresource technology* 241: 44–53.  
<http://dx.doi.org/10.1016/j.biortech.2017.05.074>.
- Nesbitt, Elizabeth R. 2020. 1 Working Paper ID-065 *Using Waste Carbon Feedstocks to Produce Chemicals*.
- Porter, Michael. 1998. “Clusters and the New Economics of Competition.” *Harvard business review* Reprint 98(November-December): 3–8.  
[http://backonline.apswiss.ch/6001/porter\\_clusters\\_and\\_the\\_new\\_economics\\_of\\_competition.pdf](http://backonline.apswiss.ch/6001/porter_clusters_and_the_new_economics_of_competition.pdf).
- Ramirez-Ramirez, C. 2019. “Unravelling the Impacts of Using Alternative Raw Materials in Industrial Clusters.” *NWO*. <https://www.nwo.nl/en/projects/ttw-vic183010> (June 2, 2022).

# Methane-to-X: an economic assessment of methane valorisation options to improve carbon circularity

Ben Lyons,<sup>a</sup> Andrea Bernardi,<sup>a</sup> Nilay Shah,<sup>a</sup> Benoît Chachuat<sup>a,\*</sup>

<sup>a</sup> *The Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, SW7 2AZ, UK*

*\*b.chachuat@imperial.ac.uk*

## Abstract

Methane side streams are produced in many different chemical processes and are normally combusted to provide process heat or to generate electricity. However, this practice is becoming less and less attractive as the industry strives towards net-zero targets and increasing the circularity of chemicals. Methane could instead be recovered and used as a valuable feedstock to produce other platform chemicals, such as H<sub>2</sub> or ethylene, which could be beneficial both for the economic performance and the carbon circularity of the system. In this work, seven different methane valorisation routes to produce additional chemicals are investigated. The considered routes include: i) five syngas-based routes combined with methanol synthesis and a methanol-to-olefins process; ii) plasma methane pyrolysis; and iii) oxidative coupling of methane. The results suggest that oxidative coupling of methane is the most profitable, with methane pyrolysis, tri-reforming and autothermal reforming also being more profitable in the base case. All routes have lower scope 1 and 2 emissions than the base case, however, dry-reforming and bi-reforming have the lowest emissions thanks to credited CO<sub>2</sub> feed streams.

**Keywords:** methane, olefins, hydrogen, techno-economic analysis, circular chemicals

## 1. Introduction

Methane is a by-product of many different chemical processes, including steam cracking for the production of olefins, pyrolysis and gasification processes, aromatic dealkylation and the anaerobic digestion of wastewater sludge. The most common use of these methane side streams is as a fuel, where it is combusted to either directly provide process heat or to produce additional electricity in a turbine. This use of methane helps chemical plants to close their energy balance and aids in their profitability. However, the chemical industry is moving towards a more sustainable future with both net-zero ambitions and carbon circularity at the forefront. The combustion of methane therefore becomes the antithesis of these goals due to the release of CO<sub>2</sub> and thus loss of carbon from the system (Lange, 2021). Additionally, methane itself is a potent greenhouse gas and a recent study discovered that flaring is less efficient than expected for the destruction of methane in oil and gas operations (Plant et al., 2022). The combination of these two factors reveals a critical need for alternative outlets for methane in the chemical sector.

A potential solution for this problem is using the methane as a feedstock to produce additional chemicals, such as H<sub>2</sub> or olefins. This would avoid the direct emissions from burning the side stream, but it would also maximise the value obtained from the feedstock in the original chemical process and improve the overall carbon circularity. In this paper, seven alternative processes are compared that convert methane into additional chemicals.

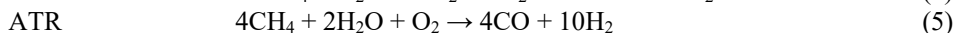
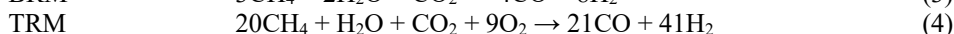
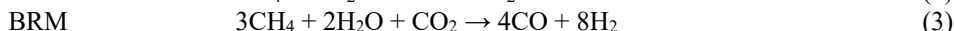
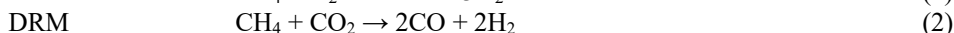
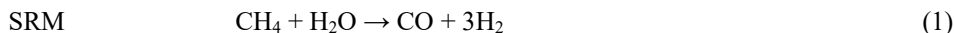
The routes considered are: i) five syngas-based routes which comprise of a methane reforming technology, methanol synthesis and finally a methanol-to-olefins conversion section; ii) a methane pyrolysis plant which decomposes methane into solid carbon and hydrogen; and iii) oxidative coupling of methane to produce ethylene. The main aim of this work is to conduct a comparative assessment of these seven methane valorisation routes by considering both their economic performance and their scope 1 and 2 emissions.

## 2. Methodology

Process flowsheets for the syngas-based routes, methane pyrolysis and oxidative coupling of methane were simulated using Aspen HYSYS® and are summarized in Figure 1. The results from these models are then used to perform a techno-economic analysis to compare the profitability of each route. A methane feed flowrate of 8 t/h was used in all routes, which is approximately the methane flowrate in an ethane steam cracker producing 700kt/y of ethylene (Belohlav et al., 2003). Note that a critical assumption made in this work is that separation units are already present and available in the base plant where the methane feed originates. Separation requirements were therefore assumed to be provided by the base plant and were not considered in this analysis.

### 2.1. Process Modelling

**Syngas-Based Routes:** These routes are comprised of three main sections: methane reforming; methanol (MeOH) synthesis; and MeOH-to-Olefins (MTO). In the reforming section, methane reacts with different oxidants to produce syngas, a mixture of carbon monoxide (CO) and hydrogen (H<sub>2</sub>). In this work, five different methane reforming options were considered: steam reforming of methane (SRM); dry-reforming (DRM); bi-reforming (BRM); tri-reforming (TRM); and auto-thermal reforming (ATR) (Arora et al., 2016). These reforming routes produce syngas at different CO:H<sub>2</sub> ratios as shown below:



All five of these reforming reactions were modelled in Gibbs reactors with the oxidants fed in their respective stoichiometric ratios. After the syngas has been produced, it is sent to the MeOH synthesis section, which is based on the work by Van-Dal and Bouallou (2013). Here, the syngas is cooled and then compressed to 78 bar via multistage compression with interstage cooling. The high-pressure syngas is sent to a packed bed isothermal reactor filled with a copper-based catalyst, where the syngas is converted into methanol via the following primary reaction:



Methanol and unconverted syngas are separated in a series of two flash vessels, with the syngas recycled to the reactor's inlet and the methanol sent to the MTO section. A purge is also present on the syngas recycle, which is sent to the base plant to recover any excess H<sub>2</sub>. The MTO section is based on the UOP/Norsk Hydro process, which consists of a recirculating fluidized bed reactor with a SAPO-34 catalyst along with a regenerator (Kuechler et al., 2000). Ethylene and propylene are the main hydrocarbon products, along with higher olefins, coke, H<sub>2</sub>, CO, CO<sub>2</sub>, alkanes and water. The higher olefins are sent to an additional olefin cracking process (OCP), which consists of a catalytic packed bed

reactor that converts the olefins into additional ethylene and propylene. The yields of the MTO and OCP reactors were taken from Hannula et al. (2015).

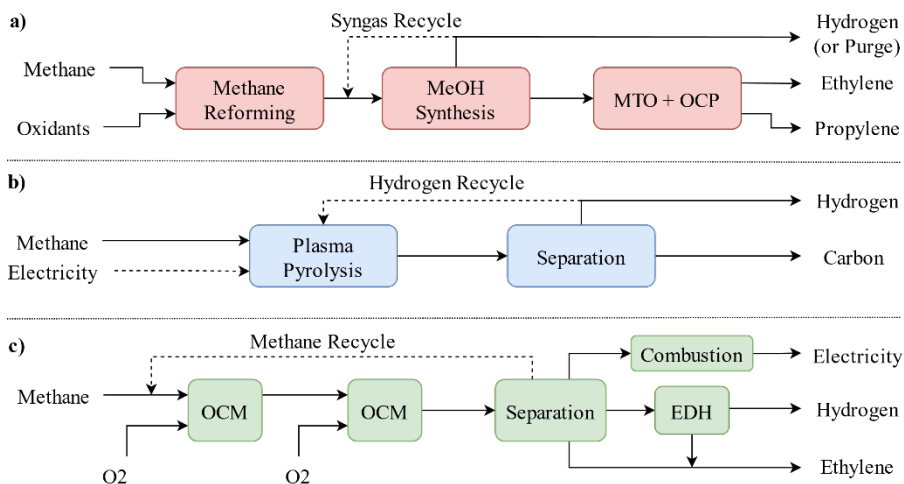


Figure 1: Schematic diagrams for the: a) syngas; b) plasma pyrolysis; and c) OCM routes

**Methane Pyrolysis (MP) Route:** In this route, methane is decomposed into its elemental components: solid carbon and H<sub>2</sub>. Multiple methods exist for the pyrolysis of methane and include using plasma, molten metals, catalytic conversion and thermal decomposition. In this work, the plasma pyrolysis route was modelled as it is the technology closest to commercialisation (Schneider et al., 2020). Methane is injected into a reactor vessel between carbon electrodes, which are fed with recycled H<sub>2</sub> and electricity to generate a plasma gas. The plasma reacts with the injected methane and decomposes into solid carbon and H<sub>2</sub> and was modelled using a conversion reactor achieving a 94% per-pass conversion (Hardman et al., 2020). Temperatures within the plasma reactor reach approximately 2100°C so the products are first cooled and then separated using cyclones and filters. If it is of sufficient quality, the solid carbon produced can be sold as ‘carbon black’, which is commonly used as a pigment and a reinforcing filler for tyre production.

**Oxidative Coupling of Methane (OCM):** This route has long been the subject of high industrial and academic interest for being able to produce ethylene directly from methane and for overcoming the highly endothermic nature of steam cracking. OCM is an exothermic process that is conducted over metal oxide catalysts at temperatures between 500-1000°C. The two primary reactions in OCM are first the formation of ethane via the oxidative coupling of methane (Eq. 8) and then the formation of ethylene via the oxidative dehydrogenation of ethane (Eq. 9) (Stansch et al., 1997):



OCM also forms CO, CO<sub>2</sub> and H<sub>2</sub> as unwanted side products and has typically low yields and selectivity towards both ethane and ethylene, which has restricted its commercialisation potential. In this work, the reaction flowsheet developed by Penteado et al. (2018) is used as a basis. Methane is fed into a series of two packed bed adiabatic reactors with oxygen in a CH<sub>4</sub>:O<sub>2</sub> molar ratio of 10:1, which improves selectivity towards C<sub>2</sub> components at the expense of per-pass-conversion rate. The reactors were sized so that



O<sub>2</sub> was completely consumed at the end of each reactor. Additional O<sub>2</sub> was added before the second reactor to restore the 10:1 ratio. The products are cooled and separated, and unreacted methane is recycled back to the start of the process. Unreacted ethane is sent to an ethane dehydrogenation reactor (EDH), where it is catalytically dehydrogenated into additional ethylene and hydrogen. For this reactor, 60% was assumed for both the per-pass-conversion and selectivity towards ethylene (Salerno et al., 2016). The H<sub>2</sub> and CO side products are combusted and used to generate electricity using a steam cycle.

### 2.2. Assessment

Purchased equipment cost curves and the factorial method were used to obtain the total capital investment (TCI) for each route and a plant life of 20 years and interest rate of 10% were used to annualize the TCI (Towler et al., 2022). Due to the lack of reliable data on the cost of a plasma pyrolysis reactor, the MP route was costed using the six-tenths rule on the TCI of a similar plant (Keipi et al., 2016). The levelised profit, which is defined as the profit obtained per kg of methane feed, was used to determine the economic performance of each route. A business-as-usual (BAU) case was used as a benchmark, which is the price of methane if it was sold to the National Grid as a natural gas substitute (0.17 USD/kg, Feb 2021). Utility and raw material costs were taken as market prices in 2021, with heating provided by blue H<sub>2</sub> at 2.01 USD/kg (Al-Qahtani et al., 2021). The methane feed is assumed to be free as it originates from the base plant. Sensitivity analysis on the natural gas price will be considered in future work to account for its volatility. The scope 1 and 2 emissions were calculated considering direct CO<sub>2</sub> emissions from each route as well as the carbon intensity of the electricity supply (UK 2021 average at 0.265 kg<sub>CO2</sub>/kWh) and the blue H<sub>2</sub> for heating (4.3 kg<sub>CO2</sub>/kg<sub>H2</sub>, Al-Qahtani et al., 2021). Credits were given for any CO<sub>2</sub> that was consumed as a feedstock.

## 3. Results and Discussion

The levelised profit and scope 1 and 2 emissions of each route are presented below in Figure 2 (a) and (b) respectively. For the economic assessment, initial observation reveals that four out of the seven routes are more profitable than the BAU case. OCM shows the highest profit at 0.38 USD/kg, followed by MP and TRM at 0.30 USD/kg and ATR at 0.23 USD/kg. The promising results for OCM are due to both its relatively low capital investment and the electricity generation from the side product combustion, which negates the utility cost. However, its total revenue is limited by its low ethylene selectivity as only 47% of the carbon from the methane feed ends up in ethylene. This carbon efficiency becomes 75% when ethane is included, however it presents a negligible contribution to the route's total revenue. Like OCM, MP benefits from a simple flowsheet with few units leading to a low capital investment, but it is hindered by the high electricity requirement in the plasma reactor. It is also dependent on the carbon black sale to remain more profitable than the BAU case. TRM produces syngas with a H<sub>2</sub>:CO ratio slightly greater than 2:1, which meets the minimum stoichiometric requirement for methanol synthesis. This maximises both the production of methanol from syngas and thus olefin yield in the MTO and OCP section. TRM also has the lowest utility requirement out of the syngas-based routes due to its very low heat of reaction in the reforming section, which in turn enables it to be less vulnerable to the high H<sub>2</sub> heating costs. ATR has a higher total revenue than TRM thanks to the excess H<sub>2</sub> produced; however, this is not sufficient to offset the increase in capital and utility costs due to a higher heating requirement in the reforming section and a higher cost for the heat exchange network.

The other three syngas-based routes are less profitable than the BAU case: SRM at 0.11 USD/kg; BRM at 0.08 USD/kg; and DRM at -0.24 USD/kg. All three of these routes have

highly endothermic reforming reactions and thus have significant heating requirements. Despite SRM and BRM having the highest revenues, this is not sufficient to offset the large utility cost. BRM performs slightly worse than SRM due to the additional cost of the CO<sub>2</sub> feed. Finally, DRM is the only route that is unprofitable. Since it generates syngas in a 1:1 ratio, most of the carbon in the methane feed is lost in the purge as excess CO. Further analysis revealed that only 42% of carbon in the DRM feed was converted into olefins. Another contributing factor to DRM's poor performance is the 1:1 CH<sub>4</sub>:CO<sub>2</sub> molar feed requirement, which generates a substantial total raw material cost.

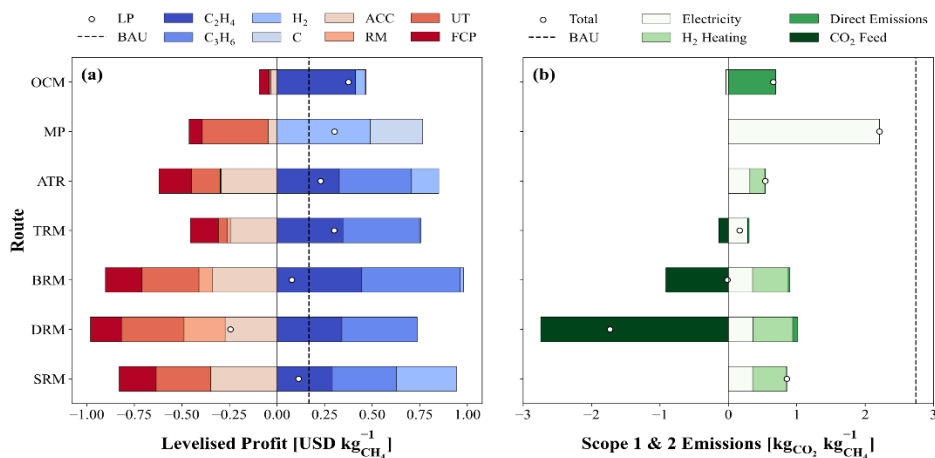


Figure 2: (a) Levelised profit and (b) Scope 1 & 2 emissions of each methane valorisation route. (LP = levelised profit; BAU = business-as-usual case; ACC = annual capital charge; RM = raw materials; UT = utilities; FCP = fixed costs of production)

Figure 2 (b) reveals that all routes have lower scope 1 and 2 emissions compared to the BAU case. However, the relative performance of the routes has somewhat reversed when compared to the economic results. MP has the highest emissions at 2.21 kg CO<sub>2</sub>, followed by SRM and OCM at 0.86 and 0.66 kg CO<sub>2</sub> respectively. All the emissions for MP come from the electricity source, whilst OCM is the only route that has a substantial contribution from direct emissions due to the side-product combustion. ATR and TRM are the other routes with positive scope 1 and 2 emissions at 0.54 and 0.17 kg CO<sub>2</sub> respectively. As seen in the economic assessment, the minimal heating requirement for TRM results in better performance compared to ATR. BRM achieves net-zero operation at -0.01 kg CO<sub>2</sub> and DRM shows net-negative emissions at -1.74 kg CO<sub>2</sub> thanks to the substantial CO<sub>2</sub> feed requirement. However, the promising results for DRM will likely be negated when scope 3 emissions are considered, as the unconverted CO stream would likely be combusted. Note that accounting for the product separations may impact both the economic and emissions results and will be considered in future work.

#### 4. Conclusions

This paper presented a model-based comparison between seven different methane valorisation routes to produce additional chemicals. The routes considered were i) five syngas-based routes which used different methane reforming technologies (steam, dry, bi, tri and autothermal reforming) followed by methanol synthesis and methanol-to-olefins; ii) methane plasma pyrolysis; and iii) oxidative coupling of methane. The results

show that OCM has the highest profit due to its low capital and operating costs, despite its low ethylene selectivity. MP, TRM and ATR are also more profitable than the base case, whilst SRM, BRM and DRM are less profitable in part due to their highly endothermic reforming reactions. However, DRM and BRM have the lowest scope 1 and 2 emissions owing to their large CO<sub>2</sub> feed streams, whilst MP has the highest thanks to its high electricity requirement. Nevertheless, all seven routes have lower scope 1 and 2 emissions compared to the base case. Further investigation is required to assess the overall environmental impact of each route (including analysing alternative heating and electricity sources), the robustness of each route to fluctuations in market conditions, and the scale effect of the methane side stream. The low-capital OCM and MP routes are expected to remain feasible at lower methane flowrates whilst the syngas routes will only be applicable to large operations to leverage economies of scale.

**Acknowledgements:** The authors are grateful to the Engineering and Physical Sciences Research Council (EPSRC) for funding the research under the UKRI Interdisciplinary Centre for Circular Chemical Economy programme (EP/V011863/1). BL thanks the Dpt. of Chemical Engineering at Imperial College London for a PhD scholarship.

## References

- A. Al-Qahtani et al., 2021, Uncovering the true cost of hydrogen production routes using life cycle monetisation, *Applied Energy*, 281, 115958
- S. Arora et al., 2016, An overview on dry reforming of methane: strategies to reduce carbonaceous deactivation of catalysts, *RSC Advances*, 6, 110, 108668-108688
- Z. Belohlav et al., 2003, The kinetic model of thermal cracking for olefins production, *Chemical Engineering and Processing: Process Intensification*, 42, 6, 461-473
- I. Hannula et al., 2015, Light olefins and transport fuels from biomass residues via synthetic methanol: performance and cost analysis, *Biomass Conversion and Biorefinery*, 5, 1, 63-74
- N. Hardman et al., 2020, Carbon black from natural gas, US10808097B2 (US Patent)
- T. Keipi et al., 2016, Techno-economic analysis of four concepts for thermal decomposition of methane: Reduction of CO<sub>2</sub> emissions in natural gas combustion, *Energy Conversion and Management*, 110, 1-12
- K. Kuechler et al., 2000, Process for converting oxygenates to olefins with direct product quenching for heat recovery, US6121504A (US Patent)
- J. P. Lange, 2021, Towards circular carbo-chemicals – the metamorphosis of petrochemicals, *Energy & Environmental Science*, 14, 8, 4358-4376
- A. Penteado et al., 2018, Techno-economic evaluation of a biogas-based oxidative coupling of methane process for ethylene production, *Frontiers of Chemical Science and Engineering*, 12, 4, 598-618
- G. Plant et al., 2022, Inefficient and unlit natural gas flares both emit large quantities of methane, *Science*, 377, 6614, 1566-1571
- D. Salerno et al., 2016, Techno-Economic Evaluation of an Oxidative Coupling of Methane Process at Industrial Scale Production, *Computer Aided Chemical Engineering*, 38, 26, 1785-1790
- S. Schneider et al., 2020, State of the Art of Hydrogen Production via Pyrolysis of Natural Gas, *ChemBioEng Reviews*, 7, 5, 150-158
- Z. Stansch et al., 1997, Comprehensive Kinetics of Oxidative Coupling of Methane over the La<sub>2</sub>O<sub>3</sub>/CaO Catalyst, *Industrial & Engineering Chemistry Research*, 36, 7, 2568-2579
- G. Towler, R. Sinnott, 2022, *Chemical engineering design: principles, practice and economics of plant and process design*, Elsevier
- E. S. Van-Dal, C. Bouallou, 2013, Design and simulation of a methanol production plant from CO<sub>2</sub> hydrogenation, *Journal of Cleaner Production*, 57, 38-45

# Purification of Raw Bio-Methanol from Renewable Feedstocks: A Flexibility Assessment

Matteo Fedeli<sup>a,b</sup>, Alessandro di Pretoro<sup>b</sup>, Ludovic Montastruc<sup>b</sup>, Flavio Manenti<sup>a</sup>

<sup>a</sup>*Politecnico di Milano, Dipartimento di Chimica, Materiali ed Ingegneria Chimica  
"Giulio Natta", Piazza Leonardo da Vinci 32, 20133 Milano, Italy*

<sup>b</sup>*Laboratoire de Génie Chimique, Université de Toulouse, CNRS/INP/UPS, Toulouse,  
France*

*flavio.manenti@polimi.it*

## Abstract

This paper investigates the production of bio-methanol from different origins feedstocks such as renewable and biological ones. The focus is on the analysis of the raw methanol composition's perturbations, dependent on the chosen feedstock, which impacts the design, the economics, and the environmental emissions of the downstream section. A thermodynamic analysis is computed to determine the final raw methanol purification depending on the utilized feedstocks. The conventional distillation column configuration is chosen as the process design for methanol purification. The purified methanol is AA grade, according to the ASTM D1152-97 standard specification. The entire section is simulated with ProSim software with a chosen Non-Random-Two-Liquid (NRTL) thermodynamic package. A two-directional flexibility analysis is performed to assess the impact of different perturbations under uncertain conditions in raw methanol compositions. The results of the flexibility analysis highlight optimal plant flexibility with around 7 % of the total cost deviation.

**Keywords:** Flexibility analysis; Biomethanol; Crude Methanol Purification.

## 1. Introduction

Methanol is one of the most produced bulk chemicals worldwide, with a global production of 160 Mton/y (Fedeli et al., 2022). Its production is a well-established technology which involves basically (i) syngas production; (ii) CO<sub>x</sub> hydrogenation (depicted in Eqs. 1-2); and (iii) raw methanol purification. The most utilized feedstocks for production are coal and natural gas. Recently biomethanol is gaining interest since is a promising advanced fuel for gasoline blends and especially for the disposal of marine fuels. Thus, could be helpful for the actual European policies and dictates. European RED II is currently set, to tackle the problem of increasing GreenHouse Gas (*GHG*) emissions, a target share of biofuel's final consumption of energy in the transport sector shall be at least 3,5 % in 2030 (EU SCIENCE HUB, 2022). Whatever feedstock wastes contain a good content of carbon and hydrogen is exploitable for the biomethanol production. Biogas, products waste gasification, and flue gases are considered the first candidates for this purpose. These new process technologies will increase worldwide biomethanol production despite several problems that are to figure out. The most highlightable obstacle of the feedstock wastes is the seasonability fluctuations and the produced raw biomethanol composition. Above mentioned feedstocks have a high CO<sub>2</sub> content compared to the natural gas one. Thus, will led to a higher production of methanol with the CO<sub>2</sub> hydrogenation pathway (Eq.2). Consequently, the produced raw methanol will

have higher water content depending by the feedstock quality. Even a slight variation could impact the design of the plant purification section.



This work aims to analyze the effects of the perturbation due to different feedstock utilization in the conventional raw methanol purification. Flexibility analysis is introduced to supply an optimum design which includes the probability that a perturbation could change the process performance.

## 2. Materials and Methods

The flexibility assessment consists of three different steps: (i) a thermodynamic analysis to classify feedstock quality (based on  $\text{CO}_2$  content) and the expected raw methanol composition; (ii) the design and the model of the biomethanol purification section in pre-defined nominal conditions; (iii) the evaluation of the process flexibility introducing the bi-directional system perturbations. In the first task, an Aspen HYSYS v11 process simulation is computed to understand how the  $\text{CO}_2$  content in the feedstock affects the outlet composition. The feedstock composition is split into three main compounds:  $\text{H}_2$ ,  $\text{CO}$ , and  $\text{CO}_2$ . The hydrogen molar fraction is kept constant while  $\text{CO}$  and  $\text{CO}_2$  are changed respectively respecting the constraint of constant overall flowrate. The thermodynamic model selected is SRK-Twu. An isothermal Gibbs reactor with a temperature of  $260^{\circ}\text{C}$  is used as a model to predict the final methanol composition. The final target is to correlate the  $\text{CO}_2$  to  $\text{CO}$  ratio with the final water content in the crude methanol. In this step, the by-products of the reaction such as higher alcohol, DME, and acetone are not considered. However, this final assumption is neglected in the second step where is important to model, with a good detail grade, the purification section.

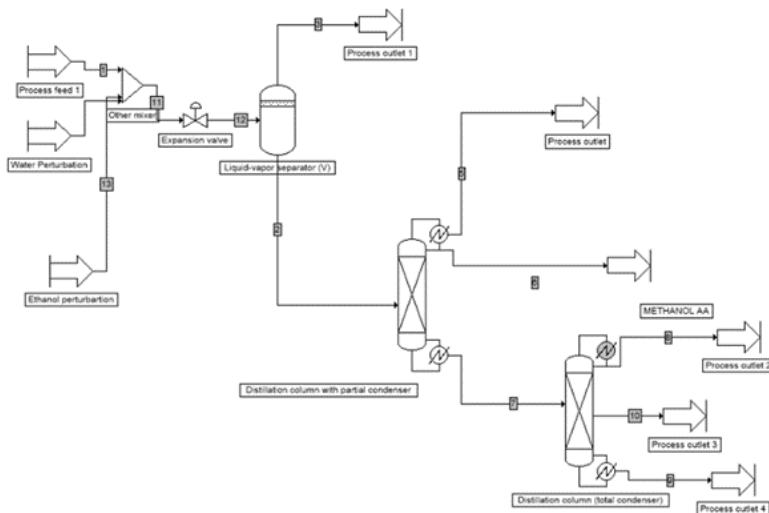


Figure 1: ProSim layout of raw methanol purification.

This is developed with ProSim software with a chosen Non-Random-Two-Liquid (NRTL) thermodynamic package. The typical impurities data is taken from the work of Xue et al., (2021). The crude biomethanol is characterized by 19 different compounds (Xue et al., 2021).

The purification process is modelled with the most used industrial applications: distillation column trains. Figure 1 depicts the process flow simulation. The purified methanol target is AA grade, according to the ASTM D1152-97 standard specification (ASTM, 1998). As a nominal case study to design the unit is chosen the raw methanol from biogas processing. Finally, in the third step, two-directional perturbations are applied to system to assess the plant flexibility. Water and ethanol content in the crude are selected as perturbation variables. Resilience index ( $RI$ ) and Swaney and Grossmann ( $F_{sg}$ ) are selected to evaluate the impact concerning economical assessment.  $RI$  expresses the largest total disturbance load and is independent of the direction of the perturbation.  $F_{sg}$  represents the maximum fraction of the expected deviations  $\delta$  that can be accommodated by the systems (Di Pretoro et al., 2022).

### 3. Results and Discussion

The thermodynamic analysis is useful to characterize and forecast the crude methanol composition among a wide range of syngas quality. This property is identified with the ratio between the inlet flowrates of  $\text{CO}_2$  and  $\text{CO}$ . Figure 2 shows the expected water production (in molar fraction) varying the  $\text{CO}_x$  ratio. It is quite clear that the chemical pathway  $\text{CO}_2$  hydrogenation is favoured increasing the carbon dioxide content. Considering syngas produced from biogas-based reforming processes, the  $\text{CO}_x$  ratio is close to the unity (Chein et al., 2021). In this case study, the estimated water molar fraction will be around 15% mol/mol<sub>tot</sub>. When the molar ratio tends to infinite values, i.e., methanol produced from flue gas and renewable hydrogen, the water produced overcomes the value products. In this case, is useful to add in the process a pre-treatment with a CO-shift unit to increase the feedstock CO quantity.

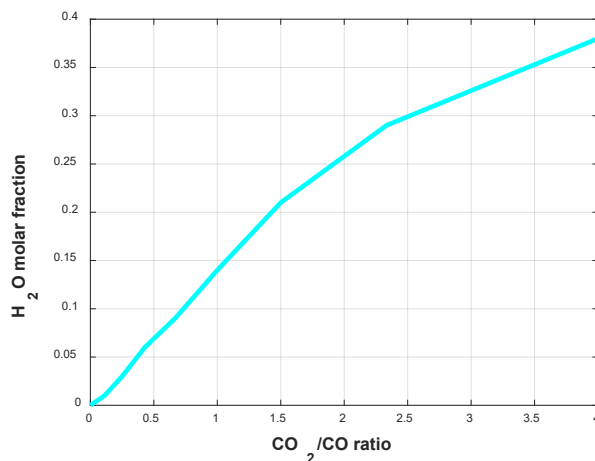


Figure 2: Thermodynamic analysis results, the x-axis represents the  $\text{CO}_2/\text{CO}$  ratio while the y-axis is the expected water molar fraction in the produced crude biomethanol.

The second step of this work is to develop a process simulation of the methanol purification section which includes all the crude impurities and the useful unit operation to reach the AA-grade methanol specification. The nominal case study is chosen as production of biomethanol from biogas-based reforming process. The syngas is laminated through a valve and passes through a separator which removes the uncondensable gases ( $\text{CO}$ ,  $\text{H}_2$ ). These are recycled back to enhance methanol production. After this operation,

the crude mixture enters a stabilizer column to remove the CO<sub>2</sub>. The last unit operation is a rectifier, with a total condenser, useful to refine until the desired specification. Both distillation columns work in atmospheric conditions. Table 1 describes the results of the process simulation.

Table 1: Main output variables of the ProSim process simulation.

<b>Variables</b>	<b>Stabilizer</b>	<b>Rectifier</b>
<b>Temperature</b>	64-67°C	64-67°C
<b>Pressure</b>	1 atm	1 atm
<b>Reflux ratio</b>	16.74	2.17
<b>Reboiler duty</b>	80.01 kW	575.25 kW
<b>Condenser duty</b>	30 kW	575.09 kW
<b>Recovery ratio</b>	100 %	70 %
<b>Number of trays</b>	5	51
<b>Specifications</b>	---	Methanol purity 0.99878 w/w  Ethanol recovery 98%

The actual operating conditions are low-energy demanding, the utilities are Cooling Water (CW) for the condensers and Low Pressure Steam (LPS) for the reboilers. The first column (stabilizer) works with only five trays to remove the CO<sub>2</sub> from the crude biomethanol. The rectifier requires fifty-one equilibrium stages to achieve very high purity separations. This one presents a stage with an ethanol spill useful to reach the solution convergence. CAPital EXpenditure (CAPEX) and OPERational EXpenditure (OPEX) related to the nominal case study have been estimated using Guthrie's method (Baillie and Whiting, 1998). 10 years of plant lifetime is chosen as parameter to normalize the fixed costs in a specified time range. In this way is possible to compute the Total Annual Cost (TAC) following Eq. 3:

$$TAC = \frac{CAPEX}{10 \text{ years}} + OPEX \quad (3)$$

The main OPEX contributions are the CW and LPS utility costs. The final step of the work is the economical flexibility assessment of the purification section in different operating conditions. As mentioned in section 2, the perturbations variables are the water and the ethanol content in the crude mixtures. The aim is to study how the perturbations affect the nominal design and its flexibility. The output variables to analyze are the column duties which are directly correlated to the CAPEX/OPEX voices. If a distillation system requires more LPS to work, this is reflected in higher OPEX or a different sizing of the kettle reboiler (CAPEX). The perturbation variables are written as vectors as reported in Table 2,  $\delta$  is the step deviation for each vector iteration.

Table 2 Property of the perturbation vectors

Perturbated variable	Value	$\delta$	Size
Water feed	250 kg/h	$\pm 50$ kg/h	11
Ethanol feed	4 kg/h	$\pm 0.8$ kg/h	11

The combination of each value associated with the vectors represents a case study. The size of the perturbation matrix is 121 case studies. Each of them is evaluated in the process simulation with the tool “case studies”, reporting convergence in the entire domain. The TAC is evaluated for each cell of the matrix, and the corresponding deviations are calculated concerning the nominal case study TAC. Figure 3 is the surface plot which correlates the TAC deviation (%) with the two flowrate perturbations.

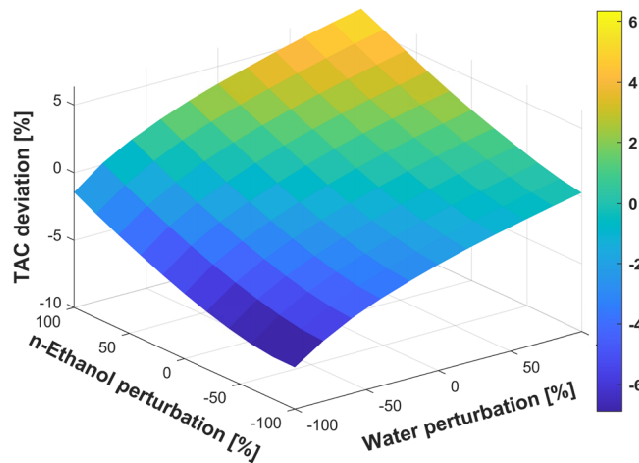


Figure 3: Matrix of TAC deviation with water and ethanol flowrates perturbation.

The maximum deviation value corresponds to 6.34% of the nominal TAC, the associated case study is related to the maximal perturbation of both vectors. The perturbation matrix and its relative TAC deviation are useful to assess the purification section flexibility via deterministic indexes computation. The plot represented in Figure 4 correlates the deterministic index,  $F_{sg}$ , and  $RI$ , in the function of the relative cost deviation. It is easier to observe the economic investment to effort different probability of perturbations are quite feasible. The proposed and simulated system has an optimistic value of flexibility following the deterministic indexes. Considering the  $F_{sg}$ , with a total investment of 6.3% the plant could afford the entire range of selected perturbations. The  $RI$  value is lower than  $F_{sg}$  because considers only the effort load with a specific perturbation direction. The negative TAC deviations mean less investments in case of feedstock with minor content of water and ethanol. That area of TAC saving includes the conventional process plants, where natural gas and coal are utilized for the production.



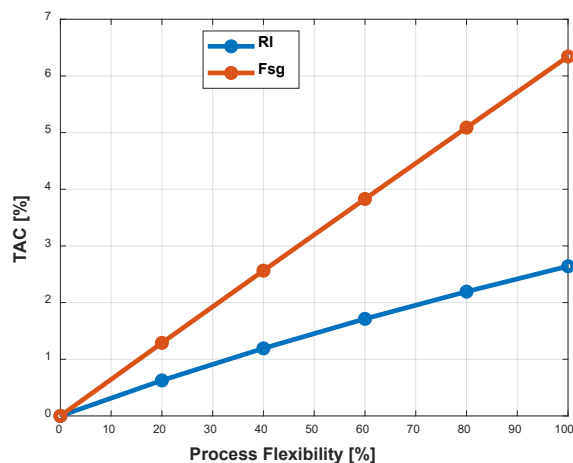


Figure 4: Deterministic index comparison to additional costs.

#### 4. Conclusions

New processes to produce biomethanol consider the utilization of renewable and biological feedstock replacing conventional ones such as natural gas and coal. Different material input significantly affects the production introducing oscillating perturbations. Water and ethanol composition mostly impacts the crude methanol purification section. This work highlights the wide range of flexibility under uncertain conditions of conventional distillation columns. The results of flexibility analysis are quite good indicators of the potentiality to retrofit, revamp, and build process plants able to afford feedstock perturbations with relatively small investments.

#### References

- ASTM, 1998. Standard Specification for Methanol (Methyl Alcohol).
- R.C. Bailie, W.B. Whiting, 1998. Analysis, synthesis, and design of chemical processes, Choice Reviews Online, <https://doi.org/10.5860/choice.36-0974>
- R.Y. Chein, W.H. Chen, H. Chyuan Ong, P. Loke Show, Y. Singh, 2021, Analysis of methanol synthesis using CO<sub>2</sub> hydrogenation and syngas produced from biogas-based reforming processes, Chem. Eng. J. 426, 130835, <https://doi.org/10.1016/j.cej.2021.130835>
- A. Di Pretoro, M. Fedeli, F. Ciranna, X. Joulia, L. Montastruc, F. Manenti, 2022, Flexibility and environmental assessment of process-intensified design solutions: A DWC case study, Comput. Chem. Eng. 159, <https://doi.org/10.1016/j.compchemeng.2022.107663>
- EU SCIENCE HUB, 2022, URL <https://ec.europa.eu/jrc/en/jec/renewable-energy-recast-2030-red-ii>
- M. Fedeli, F. Negri, F. Manenti, 2022, Biogas to advanced biofuels: Techno-economic analysis of one-step dimethyl ether synthesis, J. Clean. Prod. 376, 134076, <https://doi.org/10.1016/j.jclepro.2022.134076>
- X. Xue, Q. Gu, H. Pascal, O.M. Darwesh, B. Zhang, Z. Li, 2021, Simulation and optimization of three-column triple-effect methanol distillation scheme, Chem. Eng. Process. - Process

# Life cycle assessment of Pine needle-based electricity generation in India

Pratham Khaitan,<sup>a</sup> Ankush Halba,<sup>a</sup> Pratham Arora,<sup>a,\*</sup>

<sup>a</sup>*Hydro and Renewable Energy Department, Indian Institute of Technology Roorkee, Roorkee-247667, India*

*\*pratham.arora@hre.iitr.ac.in@Indian Institute of Technology Roorkee, Roorkee, India.  
+91-1332284921*

## Abstract

Electrification in the rural and remote areas and constant forest fires due to Pine needles are the major problems of the Uttarakhand state of India. Pine forests comprise 16.15% of the state's total forest area, producing  $1.91 \times 10^6$  tonnes of Pine needles annually. Gasification of Pine needles could be employed to generate localized electricity, thereby diminishing forest fire frequency. Nevertheless, a life cycle assessment (LCA) must be conducted to comprehend the environmental advantages of this process. Therefore, using the GaBi software, the LCA of Pine-needles to the electricity gasification supply chain is performed for two biomass pretreatment cases (1-shredding and 2-briquetting) for the Almora district of Uttarakhand state. The predicted global warming potential (GWP)-100-year emissions for case 1 and case 2 were 0.679 and 0.578 kg CO<sub>2</sub>-eq for 1 kWh of electricity production, respectively, compared to the national grid (0.86 kg CO<sub>2</sub>-eq kWh).

**Keywords:** LCA, Pine needles, Gasification, Electricity, Supply chain.

## 1. Introduction

Almost 3.5 million people live close to significant tracts of Pine forests, where firewood and electricity are in short supply in Uttarakhand (Dhaundiyal and Tewari, 2016). According to the Forest Department of Uttarakhand, Pine forests are spread over around 3,430 km<sup>2</sup> in 17 forestry divisions across 13 districts, producing two million tonnes of Pine needles (PNs) annually (Bisht et al., 2014). Since PNs are extremely flammable by nature, they can contribute to or even start forest fires, especially in the summer when temperatures are high and relative humidity is low. In addition, PN creates a covering on the earth that prevents rainwater from entering the soil and restricts nearby plants' growth, reducing the food available for cattle and resources such as wood, medicinal herbs, and flowers. Therefore, collecting PNs from forest floors and using them for energy production can promote energy sustainability in rural and remote parts of Uttarakhand, which will also aid in minimizing forest fires (Dhaundiyal and Gupta, 2014). Pine needle gasification is a promising route for harnessing its energy potential (Kala and Subbarao, 2017). Additionally, decentralized PN gasification could be a sustainable source of electricity in areas far from the main power system. However, an ineffective biomass supply chain is a major hindrance to utilizing biomass gasification-based plants for energy generation (Halba et al., 2022). Generally, the biomass supply chain consists of three major components: collection, transportation, and storage, where all three stages, directly or indirectly, impact the environment (González and Sandoval, 2020). Therefore, the environmental assessment of the biomass supply chain is mandatory to make any

biomass-based energy project successful in producing green energy and reducing greenhouse gas (GHG) emissions.

Life cycle assessment (LCA) is a commonly employed approach used to evaluate the environmental performance of any system or process. Whittaker et al. (2016) carried out the LCA using biomass (short rotation coppice) to predict the GHG and methane emissions from outside biomass storage. The study concluded that there were up to 20% dry matter losses during biomass storage and the biomass used provided up to 90% savings in GHG emissions. Another study by Beagle and Belmont (2019) used the LCA to evaluate the effects of transporting biomass (wood chips and wood pellets) on the overall amount of GHG emissions produced by bioenergy production. It was found that an integrated biomass facility could cut emissions by as much as 76% compared to coal. The research concluded using either of the two biomasses for electricity generation would result in lower GHG emissions in comparison to coal-powered stations. Costa et al. (2022) conducted the economic, environmental, and technical performance of the complete woodchips-to-energy supply chain using a commercially available biomass gasification-based system in Southern Italy. The study summarized that the impact of combustion and gasification accounted for almost 50% of the overall environmental impact. Gupta et al. (2022) employed LCA for bioenergy conversion via pyrolysis to produce bio-oil from the PNs, including the pyrolysis co-products (non-condensable gases and biochar). They integrated LCA with parameter sensitivity and economic analysis to evaluate processes more thoroughly. The study's main drawback was that the data came from a lab setting experimental investigation and not a commercial plant. The catalytic pyrolysis used led to a 42% decrease in the overall GWP of the process.

To the best of the author's knowledge, no researcher has examined the LCA of the PNs supply chain for electricity generation, including the biomass pretreatment scenario such as shredding and briquetting. Therefore, the current study has been undertaken to conduct the LCA of the PNs-to-electricity gasification supply chain, including the two biomass pretreatment cases (1-shredding and 2-briquetting) to check their environmental viability.

## 2. Methodology

A "cradle-to-gate" LCA analysis approach is adopted for the present study, which considers the environmental impacts from a product's life cycle, which start from the Pine tree plantation (cradle) to the factory gate (electricity). The PNs are assumed to be collected on foot in the forests of Uttarakhand's Almora district. Since the fallen PNs contain very little moisture throughout the summer, it is presumed that the PN was gathered during the 90-day interval (April to June). The biomass storage duration is assumed to be nine months, so the plant can function continuously throughout the year. The collected PNs are sent to primary storage (S1) via truck transportation (T1) in the next step. The biomass pretreatment routes (shredding and briquetting) for PNs are performed in a preprocessing plant (P1) located in the primary storage. The primary distinction between shredding and briquetting is that biomass briquettes often have larger shapes and sizes, calorific values, and densities than shredded biomass. The preprocessed PNs are then transported to a secondary storage facility (S2), where a gasification plant (P2) is installed. The gasification power generation system consists of a cooling and cleaning system, a gasifier, and a gas-to-electricity generator (Halba et al., 2022).

### 2.1. Life Cycle Analysis

The LCA process consists of four main steps: setting objectives and boundaries, conducting an inventory analysis, evaluating the effects of the process, and interpreting the results.

### *2.1.1. Goal and scope definition*

The LCA process begins with establishing its goals and scope in the first phase. According to ISO 14040, this phase must specify environmental impacts, functional units, assumptions, and system boundaries (Kalinci et al., 2012). The study's goal was to compare and evaluate the potential environmental footprints of the two different preprocessed PNs (Shredded and briquettes) for electricity production based on gasification. The LCA model was developed using GaBi (version: 10.6.0.110). For the present study, 1 kWh of electricity production from pine needle biomass has been considered the functional unit (FU) for estimating the environmental impacts. The following steps are considered for conducting the LCA: biomass transportation, pretreatment of biomass, storage of biomass, and power generation from gasification.

### *2.1.2. Inventory analysis*

This LCA combines and quantifies major inputs and outputs for the FUs. Energy inputs and raw materials for the LCA model were compiled from various sources like the GaBi-integrated Eco-invent database, research literature, and information acquired from local vendors and people. The India-specific database (available in GaBi) is utilized for creating the LCA model.

The PNs are assumed to be collected by foot in the forests surrounding Almora, Uttarakhand. It is assumed that raw PNs are delivered by diesel-powered Bharat Stage IV trucks (with 7.5, 11.4, and 25-ton payload capacities). The transportation distance from the collecting location to the preprocessing and storage site is assumed to be 11 km, whereas the transportation distance from the preprocessing site to the plant site is assumed to be 56 km for a decentralized plant.

The PNs shredding requires around 7.15 kWh of electricity to process per tonne of PNs (Gupta et al., 2022). In contrast, briquetting uses around 380.69 kWh of electricity to process per tonne of PNs (Singh et al., 2021). The factor used to calculate methane emissions from biomass storage is 0.0175 kg of CH<sub>4</sub> per kg of biomass (C. Whittaker et al., 2016). Raw, shredded, and briquette PNs are assumed to have a bulk density of 80 kg/m<sup>3</sup> (Houssami et al., 2018), 175 kg/m<sup>3</sup> (Dhaundiyal and Tewari, 2016), and 1200 kg/m<sup>3</sup> (Mandal et al., 2019), respectively. GaBi's Northern-India grid mix provided all the electricity input needed for chopping and briquetting (Gupta et al., 2022).

### *2.1.3. Impact assessment*

In this stage, the data of the inventory analysis (the measured inputs and outputs) are evaluated for their ecological impact. The life cycle impact assessment (LCIA) is used to identify potential environmental consequences of life cycle inventory (LCI) and to determine their magnitude. The current investigation applies an LCIA technique based on the CML 2016-Aug 2016 version.

### *2.1.4. Interpretation*

During the final phase of the LCA process, interpretation, it is crucial to confirm that the assessment's data align with its stated objective. In this regard, the two processes that yield useful information are the analysis of inventories and the evaluation of their environmental consequences (Kalinci et al., 2012). The qualitative interpretation of impact analysis data is made possible by addressing each effect category separately. The environmental effects of producing electricity from pine needles are evaluated and compared using impact categories such as Global Warming Potential (GWP), Acidification Potential (AP), and Eutrophication Potential (EP).

### 3. LCA Results and discussion

#### 3.1. GWP

It is observed that the shredded and briquette PNs gasification results in an overall GHG emission of 0.679 and 0.578 kg CO<sub>2</sub> equivalent per kWh, as illustrated in Figure 1. It is revealed that the overall GWP for both the gasification of briquettes and the shredding of PNs is less than that of India's Northern grid electricity (0.86 kg CO<sub>2</sub> equivalent per kWh). The Pine tree absorbs CO<sub>2</sub> for its growth to produce products (such as PNs) by photosynthesis, resulting in -3.24 kg CO<sub>2</sub> equivalent per kWh and -2.61 kg CO<sub>2</sub> equivalent per kWh emissions for shredded and briquette biomass, respectively. The disparity between the values could be explained by the variation in energy density between the shredded and briquette PNs. In both cases, the emissions of exhaust gases have the largest impact on the surrounding environment, with 2.23 kg CO<sub>2</sub> equivalent per kWh for shredded PNs and 1.36 kg CO<sub>2</sub> equivalent per kWh for briquette PNs. The reason for the difference in emissions between shredded PNs and PN briquettes is due to the difference in their bulk density and biomass quantity. 1.5 kg of shredded PNs were required to produce 1 kWh of electricity (Avani, 2016) while only 1.2 kg of PN briquettes were required for the same (Kala and Subbarao, 2017). Based on the methane emissions induced by the decomposition of the biomass over time, the storage phase is accounted for as the second greatest contributor to overall GWP. It is found that the emissions from transportation are minimal due to shorter travel distances (11 and 56 km) and higher bulk density of shredded and briquette PNs.

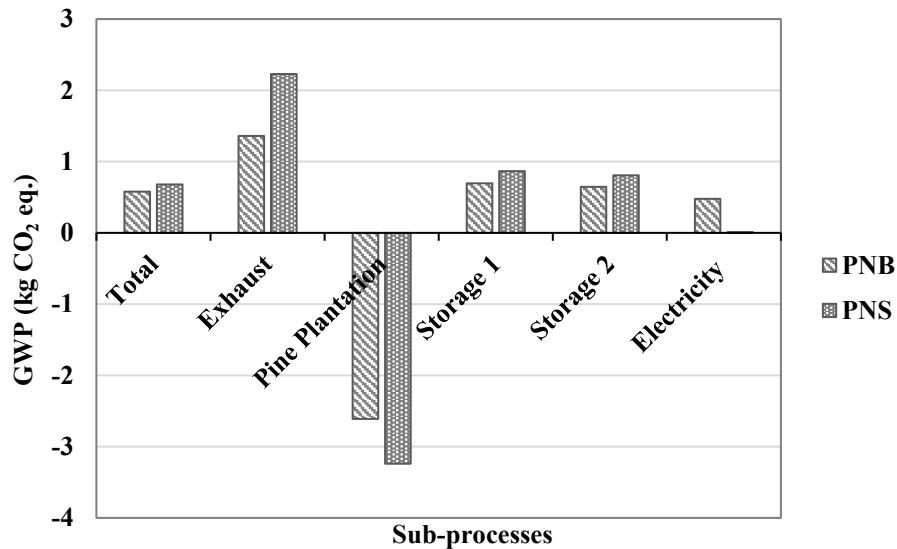


Figure 1: GWP results for shredded pine needles (PNS) and pine needle briquettes (PNB)

#### 3.2. AP and EP

The potential for acidification across both LCA scenarios is shown in Figure 2. Shredded PNs have an overall AP of  $0.35 \times 10^{-3}$  kg SO<sub>2</sub> equivalent, with exhaust gases ( $0.25 \times 10^{-3}$  kg SO<sub>2</sub> equivalent) as the primary contributor. Compared to shredded PNs, PNs briquettes have a higher AP of  $4.59 \times 10^{-3}$  kg SO<sub>2</sub> equivalent, with the most significant contribution resulting from the grid electricity required in the briquetting process. North Indian grid electricity has an AP of  $7.76 \times 10^{-3}$  kg SO<sub>2</sub> equivalent, which is higher in comparison to

the PNs used. Briquetting requires a substantial amount of electricity, due to which the emissions of PN briquettes were higher than shredded PNs.

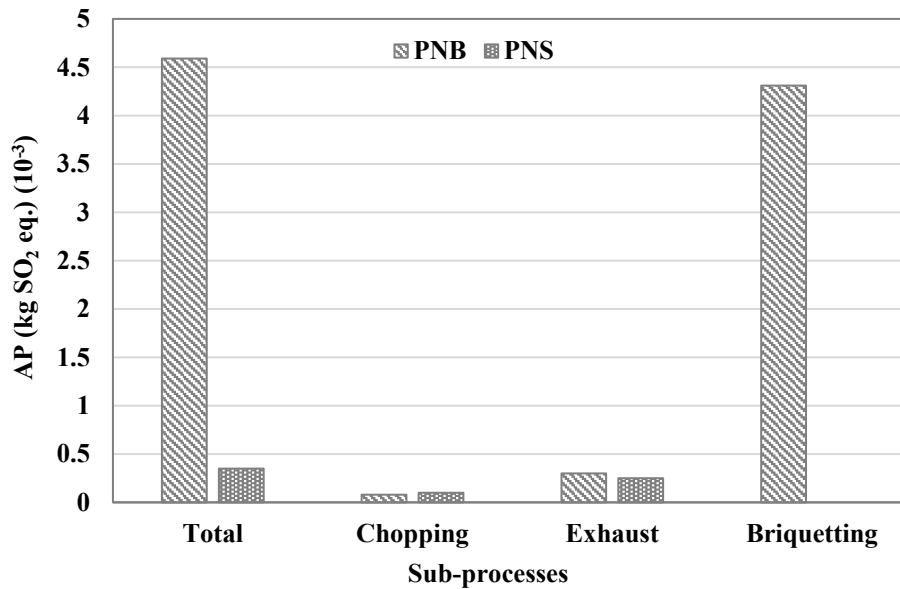


Figure 2: AP results for shredded pine needles (PNS) and pine needle briquettes (PNB)

In comparison to the Northern grid ( $4.8 \times 10^{-4}$  kg PO<sub>4</sub> equivalent), the EP values of both PNs briquettes ( $3.23 \times 10^{-4}$  kg PO<sub>4</sub> equivalent) and shredded pine needles ( $0.71 \times 10^{-4}$  kg PO<sub>4</sub> equivalent) are lower. Figure 3 depicts the overall EP of both LCA scenarios. The difference in environmental impacts was due to the higher consumption of electricity in the case of briquetting.

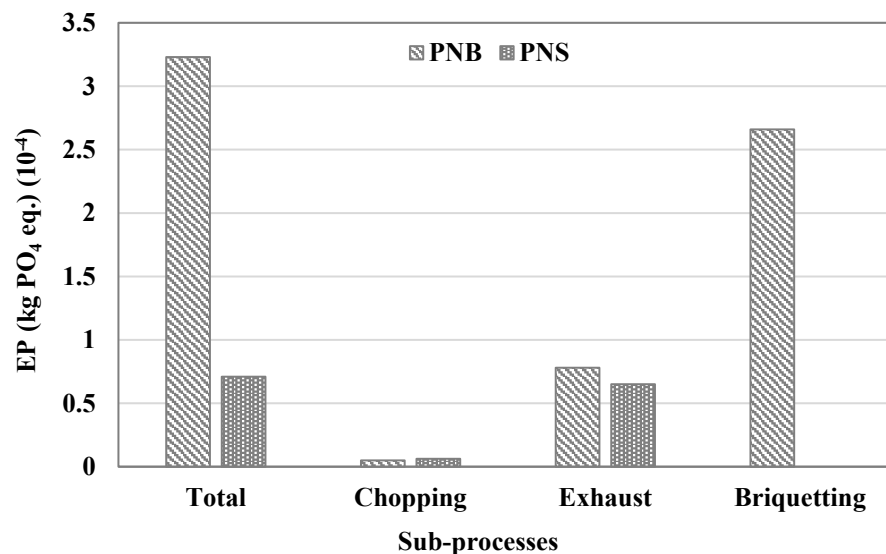


Figure 3: EP results for shredded pine needles (PNS) and pine needle briquettes (PNB)

#### 4. Conclusions

The LCA approach was applied to carry out an analysis of the effects that either of the pre-treated biomasses has on the surrounding environment when gasified to produce electricity. In addition to this, the GWP, along with AP and EP, were investigated. For GWP, the exhaust gases from the gasification facility were responsible for the greatest amount of CO<sub>2</sub> emissions, while transportation had the least influence on the surrounding environment. Electricity used in the process was the leading contributor to acidification and eutrophication potentials. Both PN-based processes outperform the North-Indian grid in terms of GWP, AP, and EP for producing 1 kWh of electricity. Reduced emissions were a result of using plant biomass as an energy source instead of coal.

#### References

- Y. Kalinci, A. Hepbasli, & I. Dincer, 2012. Life cycle assessment of hydrogen production from biomass gasification systems. *International journal of hydrogen energy*, 37(19), 14026-14039.
- S. Gupta, P. Patel, & P. Mondal, 2022. Life cycle analysis (LCA) and economic evaluation of catalytic fast pyrolysis: implication of co-product's end-usage, catalyst type, and process parameters. *Sustainable Energy & Fuels*.
- S. Mandal, G. V. Prasanna Kumar, T. K. Bhattacharya, H. R. Tanna, & P. C. Jena, 2019. Briquetting of pine needles (*Pinus roxburgii*) and their physical, handling, and combustion properties. *Waste and Biomass Valorization*, 10(8), 2415-2424.
- P. Singh, T. P. Singh, R. K. Sharma, Y. K. Negi, & R. Pal, 2021. Potential of Pine Needle Biomass as an Alternative Fuel to Mitigate Forest Fire in Uttarakhand Himalayas-A Review. *Journal of Agricultural Engineering*, 58(2), 192-203.
- M. El Houssami, A. Lamorlette, D. Morvan, R. M. Hadden, & A. Simeoni, 2018. Framework for submodel improvement in wildfire modeling. *Combustion and flame*, 190, 12-24.
- M. Costa, D. Piazzullo, D. Di Battista, & A. De Vita, 2022. Sustainability assessment of the whole biomass-to-energy chain of a combined heat and power plant based on biomass gasification: biomass supply chain management and life cycle assessment. *Journal of Environmental Management*, 317, 115434.
- L. D. Kala & P. M. V. Subbarao, 2017. Pine needles as potential energy feedstock: availability in the central Himalayan state of Uttarakhand, India. In *E3S Web of Conferences* (Vol. 23, p. 04001). EDP Sciences.
- E. Beagle & E. Belmont, 2019. Comparative life cycle assessment of biomass utilization for electricity generation in the European Union and the United States. *Energy Policy*, 128, 267-275.
- C. A. D. González & L. P. Sandoval, 2020. Sustainability aspects of biomass gasification systems for small power generation. *Renewable and Sustainable Energy Reviews*, 134, 110180.
- C. Whittaker, W. Macalpine, N. E. Yates, & I. Shield, 2016. Dry matter losses and methane emissions during wood chip storage: the impact on full life cycle greenhouse gas savings of short rotation coppice willow for heat. *BioEnergy Research*, 9(3), 820-835.
- A. Dhaundiya & P. C. Tewari, 2016. Performance evaluation of throatless gasifier using pine needles as a feedstock for power generation. *Acta Technologica Agriculturae*, 19(1), 10-18.
- A. Dhaundiya & V. K. Gupta, 2014. The analysis of pine needles as a substrate for gasification. *Hydro Nepal: Journal of Water, Energy, and Environment*, 15, 73-81.
- A. Halba, P. K. Vidyarthi, & P. Arora, 2022. Gasification as a potential solution for forest fires in the Indian Himalayan Region: A review. *Bioresource Technology Reports*, 101162.
- A. S. Bisht, S. Singh, & M. Kumar, 2014. Pine needles a source of energy for Himalayan region. *International Journal of Scientific & Technology Research*, 3(12), 161-164.
- Avani Bio Energy, 2016. Transforming Uttarakhand's Pine Needle Problem into an Innovative Energy Solution. <[http://www.millenniumalliance.in/imgs/Avani\\_Case\\_study.pdf](http://www.millenniumalliance.in/imgs/Avani_Case_study.pdf)> (Accessed on 12/09/2022).

# Evaluation of alternative carbon based ethylene production in a petrochemical cluster: Technology screening & value chain impact assessment

James Tonny Manalal,<sup>a</sup> Mar Pérez-Fortes,<sup>a</sup> Paola Ibarra Gonzalez<sup>a</sup> and Andrea Ramirez Ramirez<sup>a</sup>

<sup>a</sup>*Department of Engineering, Systems and Services, Faculty of Technology Policy and Management, Delft University of Technology, Jaffalaan 5, 2628BX, Delft, Netherlands.  
E-mail: [j.t.manalal@tudelft.nl](mailto:j.t.manalal@tudelft.nl)*

## Abstract

Due to the heavy dependence on fossil-fuels as raw materials, the defossilization of feedstocks in the petrochemical industry represents a challenge. A large number of possible process routes that use alternative carbon sources (ACS) like CO<sub>2</sub>, biomass, and waste are being developed for the feedstock replacement. For instance, to produce ethylene, more than 40 ACS process routes were identified. These multiple options make the selection of the promising process route a complex task. By replacing feedstocks, a process can change significantly and the impacts related to these changes in a highly interconnected industrial cluster can create cascading effects due to system interdependencies. This work aims to understand the cascading impacts in carbon flows and prices of implementing an ACS production process in an ethylene cluster. The results show that PVC will be the highest impacted and defossilizing one value-chain can have cascading effect on other value-chains as observed for PET.

**Keywords:** technology screening, feedstock defossilization, alternative raw materials, value chain impacts.

## 1. Introduction

Europe has set the ambitious goal to be carbon-neutral by 2050 and feedstock defossilization of carbon-based industrial processes is challenging (J.Rissman et al, 2020). For instance, the petrochemical industry depends on fossil fuels like naphtha, natural gas and ethane as carbon sources. Using alternative carbon sources (ACS) such as CO<sub>2</sub>, biomass and waste is considered an option for the replacement of oil and gas feedstocks (D. Saygain et al, 2021). A large number of possible processes using ACS are being developed to produce chemicals through electrochemical, biochemical, and thermochemical routes (M. Stork, 2018). The existence of multiple processing routes makes the decision-making process of selecting the promising alternative process route a complex task that relies on multiple techno-economic and environmental criteria. It can be expected that replacing feedstocks might significantly change process requirements like equipment, energy, water, utilities and safety, and production outputs like products, by-products, and waste. These changes can affect the overall system, for instance in terms of product prices and import-export dependencies, due to system interdependencies. This



work presents a systematic screening methodology for selecting promising ACS process technologies and aims to study cascading impacts (carbon and price), for the selected ACS ethylene production in an ethylene cluster as case study.

## 2. Methodology

For the production of ethylene using CO<sub>2</sub> and biomass as feedstocks, more than 40 different process routes were identified from the literature. In order to select the most promising alternative carbon-based ethylene production technologies a screening methodology was developed as explained in section 2.1. Then the value chain impacts of ACS technologies was studied for a base-case ethylene cluster, as explained in section 2.2 and section 2.3.

### 2.1. Technology screening

The screening methodology was developed based on a stage-gate concept and the selection was divided into five stages. The amount of information needed is reduced with this approach, as the processes are eliminated as stages progress. In stage-1, technology readiness level (TRL) is used for selection and technologies with TRL > 3 are selected to the next stage. In stage-2, the ideal stoichiometric reactions of each process route were used. Each stoichiometric reaction required to produce a chemical was assumed as individual unit operation or process step. Here, only ideal reactions are considered and side reactions are not taken into account. In stage-2, technologies with less than 4 process steps were selected. It was because as the number of steps increased, the number of routes as well as the level of complexity increased drastically. Then in stage-3, thermodynamic state functions; standard enthalpy change ( $\Delta H^0$ ), standard Gibbs energy change ( $\Delta G^0$ ) and standard entropy change ( $\Delta S^0$ ) for the overall reaction, were calculated at standard conditions using inputs from Aspen properties. Using these thermodynamic state functions, the theoretical overall heat need or generation and electricity need were calculated as:

$$\text{Heat production/need} = \sum_{\text{Biochemical+Thermochemical+Catalytic steps}} \Delta H^0 \quad (1)$$

$$\text{Electricity need} = \sum_{\text{Electrochemical step}} \Delta G^0 \quad (2)$$

At this stage, as the theoretical input and output components are known, the carbon utilization efficiency of the process route was calculated using Equation 3:

$$\text{Carbon utilisation efficiency (CUE)} = \frac{\text{Moles of carbon atom in product}}{\text{Moles of carbon atom in feedstock}} \quad (3)$$

Based on the energy need and CUE, technologies were then ranked. Using a comparative assessment, technologies with electricity need <1500 kJ/mol ethylene and CUE >50% were selected to the next stage. In stage-4, a basic economic constraint was calculated (Equation 4), using mass flow, component price and energy requirements. An in-house compiled price database with prices adjusted to 2018 as base-year using the PPI (producer prices indices) and price data from ICIS chemicals outlook was used.

$$\text{Economic constraint (EC)} = \frac{\sum_{\text{reactants}} \text{mass flow} * \text{Component price} + (\Delta H \text{ or } \Delta G)_{\text{endergonic}} * \text{Utility price}}{\sum_{\text{products}} \text{mass flow} * \text{Component price}} \quad (4)$$

Technologies with an economic constraint < 1 were selected. An EC ratio > 1 indicates that the input costs are higher than potential revenue and hence process is considered non-

profitable for assumed product prices. In stage-5, the process technologies which passed the previous gates were ranked based on the number of process steps and economic ratio, and one process route from each feedstock category (CO<sub>2</sub> and biomass) were selected.

## 2.2. Base-case ethylene cluster model

This paper used an in-house developed ethylene cluster model part of the project “Unravelling the impacts of using alternative raw materials in industrial clusters”, created in Aspen plus and based on existing processes in the Port of Rotterdam. The cluster model includes olefin, ethylene dichloride (EDC), vinyl chloride monomer (VCM), polyvinyl chloride (PVC), chlorine waste incineration (CKI), ethylene oxide (EO), ethylene glycol (EG), polyethylene terephthalate (PET), ethylbenzene (EB) and propylene oxide (PO)/styrene monomer (SM) synthesis units. The cluster has PVC, PET and SM value chains and these value-chains will be studied in this paper. The corresponding mass flows used as reference in this study are depicted in Figure 1.

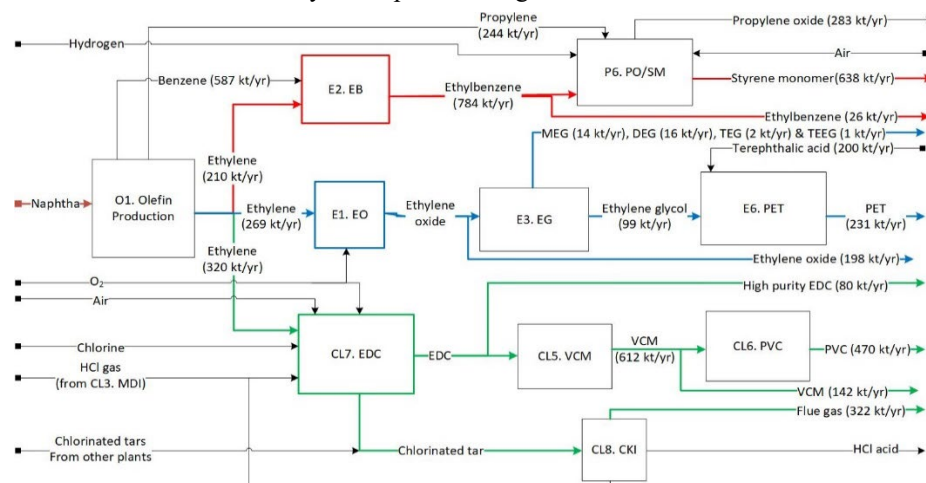


Figure 1: Simplified ethylene cluster with main mass flows

## 2.3. Value chain impact propagation

### 2.3.1. Feedstock defossilization impact on value chain

In this study, value chain defossilization is defined as the replacement of fossil-based raw materials with ACS carbon-based raw materials. To understand the extend of the defossilization impact on a value chain, Equation 5 and Equation 6 are used. The carbon contribution of a chemical building block (CBB) is defined as the amount of carbon in the value chain that originates from the CBB molecule (see Equation 5).

$$CBB \text{ carbon contribution} = \frac{CBB \text{ carbon atoms}}{\text{Value chain product carbon atoms}} \quad (5)$$

As in a chemical process, all the carbon used in a process does not end-up in the required product as by-products or waste are also formed. Hence, to understand the carbon flow along a value chain, the carbon utilization efficiency was also evaluated. The value chain carbon utilization efficiency (CUE) is defined as the ratio of the carbon mass flow in the value chain to the carbon mass flow of raw materials, as given in Equation 6. This equation helps to understand how efficiently the CBB carbon is used in the value chain to make the target product.

$$\text{Value chain CUE} = \frac{\text{Value chain chemical carbon mass flow}}{\sum_{\text{CBB}} \text{Carbon mass flow}} \quad (6)$$

### 2.3.2. Price change impact on value chain

It was assumed that the new ACS ethylene plant has the same capacity as the base-case ethylene production and the downstream units acquire the ethylene from the new ACS plant as it happens in the base-case scenario. For the downstream plants, it is considered their CAPEX remains the same (as the same product is produced and therefore there are no changes in equipment) but their OPEX changes due to changes in raw material price. Hence, in order to maintain the same gross margin, it is proposed that the downstream products will increase their corresponding prices (see Equation 7, and Equation 8).

$$\begin{aligned} \text{Gross margin} &= \text{Revenue}_{\text{base case}} - \text{OPEX}_{\text{base case}} = \\ &= \text{Revenue}_{\text{ACS case}} - \text{OPEX}_{\text{ACS case}} \end{aligned} \quad (7)$$

$$\sum_{\text{Products}} \Delta \text{Price} * \text{Mass flow} = \sum_{\text{Raw materials}} \Delta \text{Cost} * \text{Mass flow} \quad (8)$$

For multi-product processes, the increase in a raw material cost needs to be allocated to different products. In this study, a constant revenue ratio between products was assumed for all of them and, based on Equation 9, product prices were allocated.

$$\begin{aligned} &\left[ \frac{\text{Product Price}_p * \text{Product mass flow}_p}{\sum_{\text{All products}} \text{Product Price} * \text{Product Mass flow}} \right]_{\text{base case}} \\ &= \left[ \frac{\text{Product Price}_p * \text{Product mass flow}_p}{\sum_{\text{All products}} \text{Product Price} * \text{Product Mass flow}} \right]_{\text{ACS case}} \end{aligned} \quad (9)$$

## 3. Results and discussion

For the production of ethylene from CO<sub>2</sub> and biomass feedstocks, multiple technologies at different TRL were identified (see for example in Figure 2 the overview of CO<sub>2</sub>-based routes).

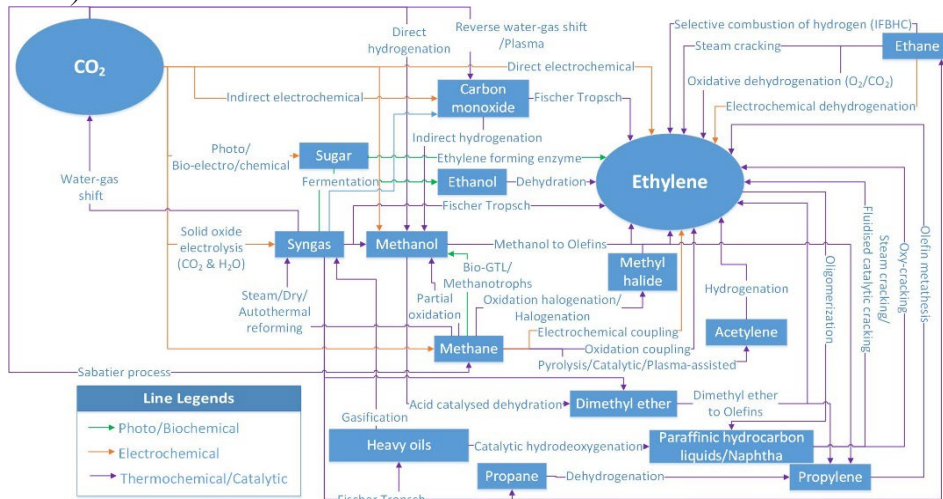


Figure 2: CO<sub>2</sub>-based process routes for ethylene production

The identified technologies were evaluated using the screening methodology and a comparative assessment for a total of 48 process routes was performed (see for example in Table 1 the comparison of four process routes). It was observed that for CO<sub>2</sub>-based routes, the needed theoretical electricity for the direct electrochemical route is lower than for the indirect water electrolysis route. This increase in electricity is reflected in the EC

ratio, as the input cost is higher in the indirect route. For biomass routes, there is no electrochemical process step and hence the electricity need is zero. But these routes have lower CUE than electrochemical routes as biomass does not have sufficient inherent H<sub>2</sub> for complete conversion of carbon in biomass. Hence, carbon is not fully utilized although these routes theoretically seem economically better than electrochemical routes because of their lower EC ratio. Hence, based on the screening methodology as explained in section 2.1 using criteria: number of process steps, energy usage, CUE and EC; the selected process routes for the production of ethylene from biomass and CO<sub>2</sub> feedstocks respectively were: Biomass steam gasification (BSG) with Fischer-Tropsch (FT) process and direct electrochemical reduction (DER) of CO<sub>2</sub> to ethylene.

Table 1: Process route comparison using the described screening methodology for ethylene production (Abbreviations: DER-Direct electrochemical reduction, MTO-Methanol to olefins, BSG- Biomass steam gasification, FT- Fischer Tropsch)

Technology	No. of process steps	kJ/mol ethylene		CUE	EC
		Electricity need	Heat production		
DER of CO <sub>2</sub> to ethylene	1	1331	0	1.00	0.90
Water electrolysis + CO <sub>2</sub> to methanol + MTO	3	1423	-304	1.00	0.95
BSG + FT process	2	0	-82	0.67	0.14
BSG + syngas to methanol + MTO	3	0	-82	0.67	0.14

In this paper, only downstream plant impacts are studied using the preliminary results. The defossilization impact of ACS ethylene on an ethylene cluster, were studied as explained in section 2.3.1 and the results are tabulated in Table 2. The PVC value chain had the highest carbon impact due to the ACS ethylene production, as 100% of PVC carbon comes from ethylene. The least carbon impact was found for PET value chain, as significant part of the carbon in PET is provided by xylene, not ethylene. Hence, the impact of ethylene defossilization is non-identical for different value chains, despite sharing the same CBB. It means that different value chains based on the same carbon contribution molecule may require different defossilization strategies. The value chain carbon utilization efficiency shows that in the PVC value chain, only 66% of the ethylene entering ends-up in the PVC product. The CUE varies for different value chains based on the main reaction selectivity, conversion, product recovery rate and by-product demand. The significance of by-product demand can be observed in the PET value chain as some of the ethylene entering is used to make ethylene oxide (EO) which is an intermediate product used in other value chains. Therefore, for multi-product value chains, the defossilization impact is not just limited to the target product but can have wider impact on other dependent value chains.

Table 2: Ethylene cluster carbon impact

Value chain	Value chain theoretical carbon flow	CBB carbon contribution	Value chain CUE
PVC	$C_2H_4 + Cl_2 \rightarrow C_2H_4Cl_2 \xrightarrow{-HCl} C_2H_3Cl$	Ethylene= 100%	PVC= 66%, EDC=7%, VCM=20%, Waste=7%
PET	$C_2H_4 + 0.5 O_2 \rightarrow C_2H_4O + H_2O$ $\rightarrow C_2H_6O_2 + C_8H_6O_4$ $\rightarrow C_{10}H_8O_4 + 2H_2O$	Ethylene= 20% P-xylene= 80%	PET=33%, PTA=19%, EG=3%, EO= 25%, Waste=20%

SM	$C_2H_4 + C_6H_6 \rightarrow C_8H_{10}$ $\rightarrow C_8H_8 + H_2$	<b>Ethylene= 25%</b> <b>Benzene= 75%</b>	<b>SM= 82%</b> , EB= 3%, Waste=15%
----	---	---	---------------------------------------

The price impact of an ACS based ethylene plant on the value chains were studied as explained in section 2.3.2 and an ACS ethylene price increase of 67% was assumed as per literature (L.Berkelaar et al, 2022). As shown in Table 3, the PVC value chain will have the highest price impact due to the higher ACS ethylene price. This is because ethylene has the highest raw material purchase cost contribution in PVC due to its price and mass flow. However, in the PET and SM value chains, the respective product prices only increased by 17%. This is because ethylene is not the main raw material in these value chains, which can also be observed from the CBB carbon contribution values given in Table 2. This results also highlights how differently the price change impacts propagate in different value chains of the same CBB due to feedstock defossilization.

Table 3: Ethylene cluster price change impact

Value chain	Price change impact propagation (delta)
Ethylene→ EDC→ VCM→ <b>PVC</b>	67%→ 62%→ 62%→ <b>49%</b>
Ethylene→ EO→ EG→ <b>PET</b>	67%→ 47%→ 47%→ <b>17%</b>
Ethylene→ EB→ <b>SM</b>	67%→ 17%→ <b>17%</b>

#### 4. Conclusion

A screening methodology based on stage-gate concept was developed to select promising ACS technologies for ethylene production. From 48 process routes, the selected technologies were: DER (for CO<sub>2</sub> feedstock) and BSG with Fischer Tropsch (for biomass feedstock). The methodology showed how the concept of stage-gate can be used to screen large number of process routes. Then the value chain impacts of feedstock defossilization for an ethylene cluster in terms of product carbon flow and price change was studied. It was observed that for ethylene value chain, PVC will have the highest carbon and price impacts due to the mass and price significance of ethylene in PVC production. It was also observed that as in the case of PET value chain, how defossilization of one value chain can effect multiple value chains due to system interdependencies.

#### Acknowledgements

This publication is part of the project Unravelling the impacts of using alternative raw materials in industrial clusters (with project number VI.C.183.010) of the research programme Vici DO which is (partly) financed by the Dutch Research Council (NWO).

#### References

- J. Rissman et al., "Technologies and policies to decarbonize global industry: Review and assessment of mitigation drivers through 2070," *Appl. Energy*, vol. 266, no. March, p. 114848, May 2020, doi: 10.1016/j.apenergy.2020.114848.
- D. Saygin and D. Gielen, "Zero-Emission Pathway for the Global Chemical and Petrochemical Sector," *Energies*, vol. 14, no. 13, p. 3772, Jun. 2021, doi: 10.3390/en14133772.
- M. Stork, J. de Beer, N. Lintmeijer, and B. den Ouden, "Chemistry for Climate: Acting on the need for speed. Roadmap for the Dutch Chemical Industry towards 2050," *Vnci*, pp. 1–52, 2018.
- L. Berkelaar et al., "Electrochemical conversion of carbon dioxide to ethylene: Plant design, evaluation and prospects for the future," *Chemical Engineering Research and Design*, vol. 182, pp. 194-206, 2022, doi: 10.1016/j.cherd.2022.03.034

# Syngas fermentation to ethanol: the effects of gas recycling on economics

Haneef Shijaz,<sup>a,b</sup> Fausto Gallucci,<sup>b</sup> Adrie Straathof,<sup>a</sup> John Posada<sup>a</sup>

<sup>a</sup> *Department of Biotechnology, Faculty of Applied Sciences, Delft University of Technology, van der Maasweg 9, 2629 HZ, Delft, the Netherlands.*

<sup>b</sup> *Department of Chemical Engineering and Chemistry, Eindhoven University of Technology, 5600 MB, Eindhoven, the Netherlands*

## Abstract

Syngas fermentation is a biochemical pathway to produce ethanol and has been commercialized successfully. The economic viability of this process could be further improved to become more competitive in the existing ethanol market. Improving gas utilization is the key, and can be done by recycling the unreacted syngas. This work is an early-stage techno-economic assessment of recycling in producing ethanol from Basic Oxygen Furnace (BOF) gas. Economic viability is measured in terms of Relative Competitive Percentage (RCP) and is a measure of closeness to the current market. Two scenarios, firstly a once-through process, and secondly a process with recycling (0.9 split ratio: recycle/purge) of gas is considered. None of them showed a positive RCP as compared to the current ethanol market. Comparing these scenarios, beyond the single pass conversion of 60%, the additional production costs due to recycling become dominating and lead to a lower RCP compared to once-through systems.

**Keywords:** Syngas fermentation, Ethanol, Basic oxygen furnace Gas, Recycling, Profitability

## 1. Introduction

Syngas fermentation is a ground-breaking industrial biotechnology platform to produce Ethanol from various feedstocks. Lanzatech, a carbon recycling company has Commercialized this technology at various locations around the world. At the same time, the produced ethanol is not competitive enough to the existing ethanol market based on thermo-catalytic processes (Benalcázar 2017). One of the promising ways to improve the profitability of ethanol is by increasing gas utilization. This is possible via improving the mass transfer, genetic engineering to enhance the microbial rate or recycling the unreacted reactants. The latter is a possible scenario at the commercial scale, but the same has not been studied in detail in the large-scale syngas fermentation processes. Therefore, this work is an attempt to give an early-stage techno-economic evaluation of the recycling of unreacted reactants in the gas fermentation process.

## 2. Methodology

The process concept for ethanol from BOF is shown in figure 1. A bubble column fermenter and a distillation-based downstream process were considered as they are the most used types for ethanol production in large-scale syngas fermentation industries. Aspen Plus V8.8 was used to model this process concept in a steady state. The BOF gas

with a composition (mole %) of 0.65% CO, 0.03% H<sub>2</sub>, 0.16% CO<sub>2</sub>, and 0.16% N<sub>2</sub> was the feedstock and a process reaction to ethanol was developed by the black box thermodynamic approach (Heijnen 1992). Gas-liquid mass transfer rates of the reactants were also integrated into the model using the empirical relations of the bubble column reactor for the mass transfer coefficient, and Henry's law for the compositions. The sizing of the units for economic evaluations was carried out either manually or by using Aspen Process Economic Analyzer. The economic viability of ethanol via syngas fermentation was checked in terms of relative competitive percentage (RCP), which is the measure of its marketability compared to the current ethanol market. The effect of recycling unreacted gas was studied via sensitivity analysis, and RCPs were measured in each case.

### 2.1. Process Description

As shown in figure 1, a mixture of fresh BOF gas and recycle gas is adjusted to the inlet pressure and temperature of the fermenter. In the bubble column fermenter, the gas (syngas + NH<sub>3</sub>) and the medium (water+NH<sub>4</sub>OH) are fed counter currently. The ethanol is fully stripped off by the gas leaving the fermenter, and it mainly contains product, water, and other unreacted components of syngas. Therefore, the gas is condensed first and sent to a phase separator. The liquefied stream mainly contains ethanol and water with other dissolved gases. Ethanol is further purified in a distillation-based downstream separation section. A part of the unreacted gas is purged (10%) to avoid the accumulation of H<sub>2</sub> in the reactor. The rest of the gas (90%) is sent back to the fermenter after separating N<sub>2</sub>, and CO<sub>2</sub>. The fermenter broth mainly contains water, biomass, & ammonium acetate, and the solids are separated before reusing the water in the fermenter. The fermenter operating conditions (Temperature at 37°C, pH at 6) are selected based on the optimal growth conditions for acetogens. The top pressure in the fermenter is 1 atm and the bottom pressure was calculated based on the hydrostatic pressure as follows;

$$p_b = p_t + \rho_{broth} * g * h_g \quad (1)$$

Where,

$p_b$	Bottom Pressure (atm)
$p_t$	Top Pressure (atm)
$\rho_{broth}$	The density of broth (kg/m <sup>3</sup> )
$g$	acceleration due to gravity (m/s <sup>2</sup> )
$h_g$	Height of the broth (m)

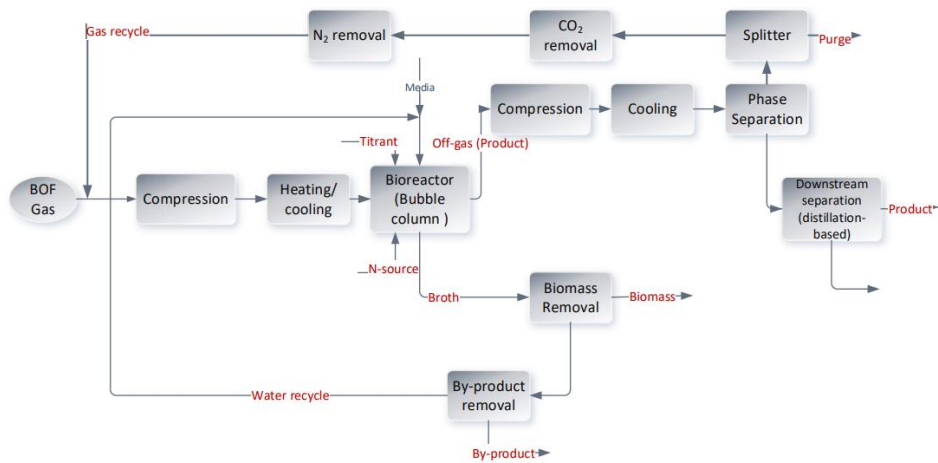


Figure 1 schematic block diagram of ethanol from BOFG via syngas fermentation

### 2.2. Economic evaluation

To calculate the economic viability of the process, the capital expenditure (CAPEX), and the annual operating cost (OPEX) are estimated using the factorial method proposed by Peters and Timmerhaus (1959). As given in Table 1, The CAPEX is estimated based on the bare equipment cost (BEC) of the units. These values are adjusted to the scale, and the base year of 2019 using six tenth rule and cost indices respectively. OPEX is evaluated based on the direct calculations from the process models and is called process-related cost (PRC). PRC consists of the cost of raw materials, utilities, and waste management. Raw material includes syngas, NH<sub>3</sub>, NH<sub>4</sub>OH, and deionized water. The cost of utilities is evaluated from the energy balances and average utility price in the Netherlands (2019). The cost of waste management is estimated using the avg. carbon releasing tax in Europe (5 Euro/ ton CO<sub>2</sub> release, 2019). For the depreciation cost, a straight-line depreciation method is used with 10% of the purchase cost as salvage value after a plant life of 15 years. The annual tax paid is calculated from the annual sales and the tax rate which is taken as 25%.

To measure the economic viability, the relative competitive percentage (RCP) is measured as follows;

$$RCP = 100 * \frac{(MP - MSP)}{MP} \quad (2)$$

Where MP is the market price of ethanol in 2019 and is 780 €/kg. And Minimum Selling Price (MSP) is evaluated for a payback time of 3 years.



**Table 1 Calculation of CAPEX (Peters, 1959)**

	Items	Factors on FCI (%)
Direct cost (DC)	Bare Equipment cost	21.8
	Installation	7.6
	Process piping	11.1
	instrumentation	4
	insulation	1.4
	electrical	4.7
	buildings	8.4
	yard improvements	3.1
	service facilities	7.2
	land	1.7
Indirect cost (IDC)	Engineering	7.5
	construction	9.7
	contractor's fee	4
	contingency	7.8
Fixed Capital Investment (FCI)	Sum up above	100
Start-up related cost	working capital	OPEX/12
	start-up cost	9
CAPEX	FCI + working capital + start-up cost	

**Table 2 Calculation of OPEX (Peters,1959)**

	Items	Calculation
Facility-dependent costs	Depreciation	$(DC-PC*0.1)/10$
	maintenance	0.15 purchase cost
	insurance	0.01 TFC
	local taxes	0.01 TFC
	plant overhead	0.05 Revenue

Process related cost	Utility	
	Raw material cost	
	Waste management	
	(CO <sub>2</sub> tax)	
Labour-related cost	labour	0.07 (PRC+FDC)
	Laboratory charges	0.1 labour
OPEX	Sum up above	

2.3. Parametric studies

A sensitivity study was carried out for various conversions and the corresponding recycling ratios. The purge was fixed at 10%, and the rest 90% are recycled in all cases. In each case, the RCPs have been calculated.

3. Results

Figure 2 shows the relation between the single pass conversion of CO, and the recycle ratio (recycle flowrate/(fresh syngas+ recycle flowrate)). In all cases, the split ratio of recycle flow rate is fixed at 0.9. As we see, a lower conversion would lead to high recycle flow rates.

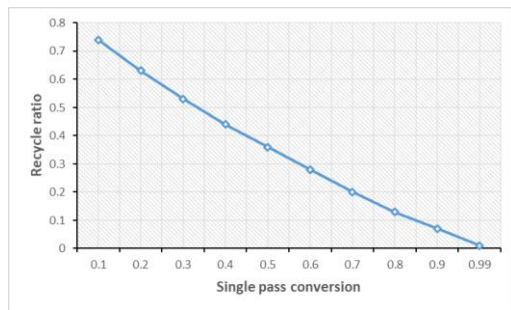


Figure 2 Single pass conversion vs Recycle ratio

unreacted CO. None of the cases showed positive viability compared to the existing market, and therefore further optimization of the process, and support (subsidies) may be required to bring ethanol from syngas fermentation to the market.

The effect of recycling on the economic feasibility of the process is shown in figure 3. The economic feasibility is represented in RCP as indicated in the previous section. It shows the positive or negative economic viability of ethanol as compared to the present market. Two scenarios are compared in this figure. Firstly, the RCP of the once-through process (no recycling) and secondly RCP of the process with recycling of

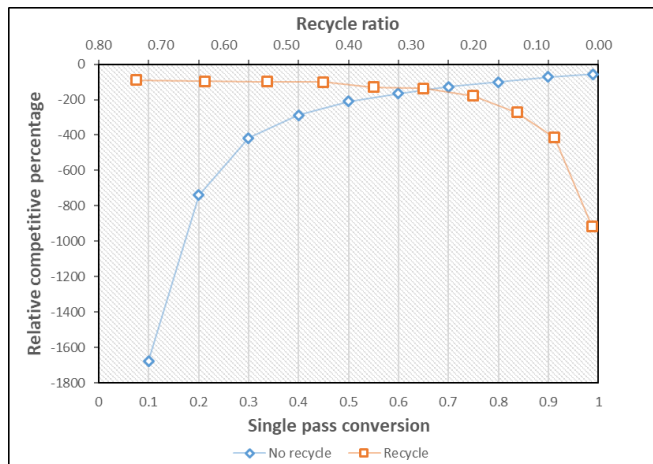


Figure 3 RCP (relative competitive percentage) for once through process and process with recycling in the production of ethanol from BOF gas via syngas fermentation.

Comparing the scenarios, recycling has improved the economic viability in multiple folds at lower conversions. It is due to the improved gas utilization, and production rates of ethanol. This is the trend up to around 60% of single pass conversion, and then the additional production cost of recycling is dominating. Therefore, at higher conversion rates, scenarios without recycling have better economic viability.

These results show a clear tradeoff between additional recycling costs and productivity improvement due to recycling. Therefore, to recycle the unreacted CO in improving the economic viability, the economic hot spots due to recycling must be identified, These units or variables must be optimized to minimize the recycling cost. Similarly, the limiting mechanism of improved production performance must be identified and must be tuned further to fully benefit from recycling CO back to the fermenter.

### Acknowledgements

This publication is part of the project Microsync (with project number P16-10 project 7 of the research programme Perspectief Novel Approaches for Microbial Syngas Conversion to Chemical Building Blocks 2017 TTW. Project title: 'Process design and sustainability assessment' which is (partly) financed by the Dutch Research Council (NWO).

### References

- Benalcázar, Eduardo Almeida, Henk Noorman, Rubens Maciel Filho, and John Posada. 2017. Production of bulk chemicals from lignocellulosic biomass via thermochemical conversion and syngas fermentation: a comparative techno-economic and environmental assessment of different site-specific supply chain configurations. *Biofuels*, Bioprod. Bioref. 11:861–886.
- Heijnen J.J et.al. 1992. A blackbox mathematical model to calculate auto- and heterotrophic biomass yields based on Gibbs energy dissipation. *Biotechnology and Bioengineering*. 40. 10:1139- 1154.
- Peters, M. S. & Peters, J. I. 1959. Plant design and economics for chemical engineers. *Eng. Economics*.
2019. DACE price booklet. 33<sup>rd</sup> edition.
2019. Carbon taxes in Europe. Tax foundation. *Taxfoundation.org*

# Modeling, simulation and techno-economic analysis of an integrated biorefinery based on halophytes

Tutku Taşçı Çilak,<sup>a</sup> Sanketkumar Raval,<sup>a</sup> Sylvia Fasse,<sup>a</sup> Mette H. Thomsen,<sup>b</sup>  
Axel Gottschalk,<sup>a</sup>

<sup>a</sup>*Institute of Process Engineering, Bremerhaven University of Applied Sciences, An der  
Karlstadt 8, 27568, Bremerhaven, Germany*

<sup>b</sup>*Department for Energy Technology, Aalborg University, Esbjerg Campus, Niels Bohrs  
Vej 8, 6700 Esbjerg, Denmark*

## Abstract

Many existing biorefineries use limited feedstocks and technologies to produce biofuels. However, an integrated biorefinery provides a variety of feedstocks and highly marketable final products leading to better economic viability. For this reason, this paper aims to perform a techno-economic analysis of the whole integrated biorefinery with modeling of the process in a flowsheet simulation tool, here Aspen Plus<sup>®</sup>. After having mass and energy balances, the assessment of technical aspects, and the economics of the integrated biorefinery process is the step forward in the development of a sustainable and biobased industrial process. In that matter, capital expenditures (CAPEX), operational expenditures (OPEX) and production costs are estimated for the integrated biorefinery, processing lignified halophyte biomass.

**Keywords:** Biorefinery, second-generation feedstock, value-added compounds, flowsheet simulation, cost-estimation.

## 1. Introduction

To cover the rising demand for food and energy is a strengthening challenge in today's world. Additionally, there is a strong need for sustainable production processes. Therefore, the concept of circular economy is more and more applied in biorefineries. There are numerous options to design a biorefinery based on feedstocks used, processes applied and products yielded. In second-generation feedstock biorefineries, wastes from agriculture and lignocellulosic biomass are processed. Thus, the decline in agriculturally usable land can be circumvented (Kamm et al. 2000). But the lignocellulosic and recalcitrant structure of the biomass with lignin, cellulose, and hemicellulose requires more process steps than first-generation feedstocks. For the production of biofuels, the thus higher production cost can make the biorefinery not competitive with fossil sources. Integrated biorefineries are advanced biorefineries including multiple processes and products for different industries to maximize the biomass valorization (Patel and Shah 2021). Hence, the incorporation of the extraction of value-added compounds in integrated biorefineries can improve economic performance (Kamm et al. 2000).

In this paper, the biomass category in focus is halophytes. These are salt-tolerant plants, which, in addition to their capability of growing without freshwater supply, are also yielding a range of bioactive compounds. These secondary metabolites are synthesized as part of the halophyte's survival mechanism in saline environments. Those phenolics

act as natural antioxidants and are of interest for their application in a variety of foods, cosmetics, and pharmaceuticals. This is additionally based on their antimicrobial activity against several pathogenic microorganisms (Jdey et al. 2017). Several halophyte species are traditional medical and nutritional plants, with the edible tips being part of gourmet foods (Antunes et al. 2021). But the dominant part of the plant is lignocellulosic biomass, which is to be valorized in circular economy biorefineries (Brown et al. 2014).

The halophyte species in the scope of this article is *Salicornia ramosissima*, which is already available on the market and can be harvested in saline areas across Europe. Further *S. ramosissima* biomass has elevated amounts of hydroxycinnamic acids and is therefore of interest for industrial valorization (Antunes et al. 2021). The lignified biomass contains extractable phenolic compounds such as hydroxycinnamic acids (HCAs) and polyphenolics, being of interest for the cosmetic and pharmaceutical industry.

Consequently, extraction methods are established also including the utilization of residual biomass for biogas and biochar production. Hence, it is a combination of different processing steps such as biomass fractionation and the production of biogas and biochar after the recovery of value-added compounds.

In the course of the process design of a whole integrated biorefinery, it is important to perform a techno-economic analysis to evaluate its technical feasibility and economic performance. Additionally, the choice of the technologies to be implemented, the operating conditions of the process units, and their sizing can be optimized through techno-economic assessment. (Bangalore Ashok et al. 2022).

## 2. Conceptual design of the integrated biorefinery

The conceptual process design of the integrated biorefinery with its individual process steps is shown in Figure 1.

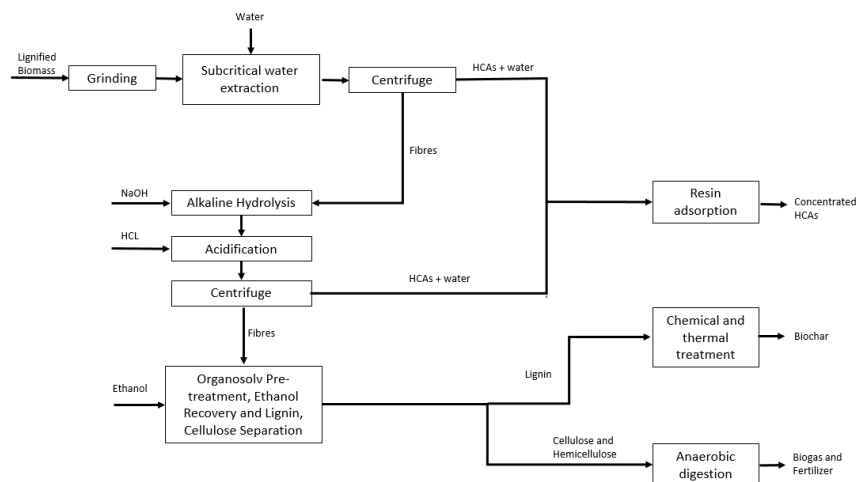


Figure 1: Conceptual process flowsheet of the integrated halophytes biorefinery

As a raw material for the process, lignified biomass was assumed to be sundried in the field, therefore grinding is the first step as pretreatment to reduce the particle size. It is followed by subcritical water extraction (SWE), a non-conventional extraction process (Taşçı Çilak et al. 2022). Downstream the SWE process, the remaining fibers are sent to alkaline hydrolysis to extract remaining HCAs. Alkaline hydrolysis at elevated

temperatures is used as post-treatment of SWE. After alkaline hydrolysis, extracts are acidified by using hydrochloric acid (HCl aq.) to lower the pH value. Consequently, both extracts from acidification and SWE are combined and by ion exchange resin to concentrate HCAs and remove sugars, salts, and all water-soluble contaminants. To realize maximum benefits, residual lignified fibers are subjected to organosolv pretreatment to generate streams that are rich in cellulose and lignin. In the organosolv process, lignin is separated from cellulose and hemicellulose by using 50/50 v/v ethanol and water solution. After that, the cellulose rich stream is valorized in anaerobic digestion to generate biogas while the lignin stream is sent to a biochar production unit.

The process flowsheet modelling and simulation of the integrated biorefinery is carried out in the commercial computer-aided process engineering (CAPE) tool Aspen Plus<sup>®</sup> V11 to enable mass and energy balance calculations.

### 3. Methods

#### 3.1. Process design and assumptions

In this study, the main focus is on the extraction of value-added compounds from the lignified biomass of the halophyte species, *S. ramosissima*. The respective approach offers the opportunity to expand the product portfolio by adding biogas and biochar as products along with high-value extracts. As a starting point to set the assumption regarding the capacity of the biorefinery, a production scenario of 1000 tons/year of lignified biomass is considered as the biorefinery capacity with 8000 h of operating hours. Furthermore, Germany is chosen for the biorefinery location. In the simulation tool (Aspen Plus<sup>®</sup>), the composition of the biomass is defined based on dry weight as 27 wt% ash, 5 wt% proteins, 3 wt% lipids, 9.5 wt% lignin, 11.5 wt% cellulose, 17 wt% hemicellulose, and 27 wt% extractives. Also, the moisture content in the feed is defined to be 15 wt%. In the process concept, recovery of lignin is assumed as 64 wt% whereas cellulose retention is 70 wt%. Also, the recovery of the targeted HCAs is estimated as 85 wt%. Assumptions, which are taken in this article, are inspired from the overall framework of the AQUACOMBINE project (AQUACOMBINE). In the flowsheet simulation model, the thermodynamic property packages ‘Solid’ and ‘NRTL’ have been applied. The physical properties such as boiling point, molecular weight and heat capacities of components being not available in the Aspen Plus<sup>®</sup> database (e.g. lignin, cellulose) have been entered by hand. Conventional compounds are selected from the AspenPlus<sup>®</sup> database. Moreover, the models that are used to define the unit operation of the process are given in Table 1.

Table 1: List of equipment and its model used in Aspen Plus<sup>®</sup>

Separation process for HCAs		Organosolv process and biochar generation	
Equipment	Model in Aspen Plus <sup>®</sup>	Equipment	Model in Aspen Plus <sup>®</sup>
Crusher	Gyratory Crusher	Mixer	Mixer
Reactor	Rstoic	Reactor	RStoic
Separator	Flash2	Separator	Flash2
Heat exchangers	Heater	Heat exchangers	Heater
Resin Adsorbtion	User Model	Washing	Swash
Centrifuge	Decanter	Separator	Sep
Pump	Pump	Distillation	RadFrac
		Precipitator	RStoic

### 3.2. Cost Estimation Method

In the simulation of the conceptual process, as shown in Figure 1, the various types of equipment such as heat exchangers, crushers, pumps, reactors, and vessels were used. The costs of each equipment have been calculated based on parameters like temperature, pressure, mass flow rate, and utilities which were compiled from the simulation. In addition to that, material of construction, efforts for installation, etc. were considered in the equipment costs. Furthermore, the whole economic assessment has been done with reference to the year 2019, using the chemical engineering plant costs index (CEPCI), a currency conversion rate of 0.87 USD/EUR and the estimated lifetime of the plant being 20 years. The cost estimation method based on the individual factor method of Guthrie (1969, 1974) is followed (Seider et al. 2009) to evaluate the total capital investment (TCI). Initially, the purchase costs of equipment are determined using empirical equations. Further, the respective cost estimation method is based on the bare module factors which are based on the type of equipment (e.g. pumps: 3.30, shell and tube heat exchanger: 3.17, crusher: 1.39). The sum of all bare module costs of equipment, process machinery, spares, costs of catalysts, etc. gives the total bare module investments (TBM). The direct permanent investment (DPI) includes the sum of the costs for site preparation, services facility and allocated cost for utility plant and related facilities as they are shown in Table 2.

Table 2: Assumptions for calculating total capital investment (Seider et al. 2009)

Item	Assumption
Total bare module investment TBM	Contains costs related to purchase and installation of equipment
Cost of site preparation	5 % of TBM
Cost of service facilities	1.5% of TBM
Allocated cost for utility plant and related facilities	14% of TBM
Total of direct permanent investment DPI	Contains costs of TBM, site preparation, service facility, utility plants and related facility
Cost of contingencies and contractor's fee	18% of DPI
Total depreciable capital TDC	Contains costs of DPI, contingencies and contractor's fees
Cost of land	2% of TDC
Cost of royalties	2% of TDC
Cost of plant start up	10% of TDC
Total permanent investment TPI	Contains costs of TDC, land, royalties and plant start up
Total permanent investment TPI corrected	120% of TPI
Working capital	18% of TPI corrected
Total capital investment TCI	Sum of costs of TPI and working capital

Along with the further costs such as costs of contingencies and contractor's fees, costs of land, royalties, plant start-up, and working capital are calculated based on the percentages of TDC. Finally, the TCI is calculated as the sum of total permanent investment corrected and working capital (Seider et al. 2009).

## 4. Results

Based on the simulation results and selected operational parameters, the sizing of equipment has been performed. The cost estimation has been performed and the TCI has been calculated as 6869 k€ which covers the investment required for the considered biorefinery concept. The breakdown of the TCI is depicted in Table 3. The TBM is 2993 k€ whereas the total permanent investment corrected is 5821 k€. The addition of working capital (1048 k€) into the total permanent investment corrected provides the TCI.

Table 3: Break down of TCI

Costs	k€
Total bare module investment TBM	<b>2993</b>
Cost of site preparation	150
Cost of service facilities	45
Allocated cost for utility plant and related facilities	419
Total of direct permanent investment DPI	<b>3606</b>
Cost of contingencies and contractor's fee	649
Total depreciable capital TDC	<b>4255</b>
Cost of land	85
Cost of royalties	85
Cost of plant start up	426
Total permanent investment TPI	<b>4851</b>
Total permanent investment TPI corrected	<b>5821</b>
Working capital	<b>1048</b>
Total capital investment TCI	<b>6869</b>

Along with the estimation of total capital investment, the annual production costs are calculated and amount 4210 k€/year which is shown in Table 4.

Table 4: Break down of total production costs

Costs	k€/year
Raw material costs (excl. Biomass costs)	0.24
Utility costs	880
Direct wages and benefits (DW&B)	1300
Labour related operations (O)	1773
Maintenance cost (M)	343
Operation overhead	383
Property tax and insurance	85
Depreciation, D	167
Licensing fees	87
Cost of manufacturing	2869
General expenses	461
<b>Total production cost</b>	<b>4210</b>

The total production costs have been performed as the sum of variable cost items such as operational expenditures (OPEX) including raw material (e.g. water, ethanol, etc.) and utility costs (e.g. steam (2 bara, 120°C), cooling water (15°C), electricity, etc.) and other cost items as fixed operating expenses and general expenses. Fixed operating expenses include direct wages and benefits (DW&B) with consideration of 5 shifts, including illness and holidays, costs of labour related operations, maintenance costs, costs of operation overhead, cost of property tax and insurance (2% of total depreciable capital (TDC)), licensing fees (2% of sales), and cost of manufacturing. Moreover, the depreciation (D) is taken into consideration which is 4% of total depreciable capital (TDC). Also, selling expenses, research and development costs, management incentive compensation have been included in general expenses.

## 5. Conclusion

A techno-economic analysis on the integrated biorefinery processing lignified halophyte biomass was presented with the conceptual process design. The corresponding energy and mass balances have been calculated by the flowsheet simulation tool Aspen Plus®.



The estimated TCI and production costs amount to 6869 k€ and 4210 k€/year respectively. The selling prices of the high-value products are not included in this study because they are depending on the target sectors such as food, cosmetics and pharmaceutical industries. Hence, in this study total capital cost and production cost were aimed to be evaluated without focusing on revenue from the products. Nevertheless, in the context of the economics of the conceptual process, many factors such as the capacity of the biorefinery, production rate, type of feedstock, and technologies have an impact. In that regard, the sensitivity analysis and energy integration are required to reduce the cost in upcoming work. Therefore, the presented results in this study are preliminary, as they will be optimized and improved in the future.

## 6. Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No 862834. Any results of this project reflect only this consortium's view and the European Research Executive Agency is not responsible for any use that may be made of the information it contains.



## References

- M. Antunes; C. Gago; A. Guerreiro; A. Sousa et al., 2021, Nutritional Characterization and Storage Ability of *Salicornia ramosissima* and *Sarcocornia perennis* for Fresh Vegetable Salads, *Horticulturae*, 7, 1, p. 6. <https://doi.org/10.3390/horticulturae7010006>
- R. P. Bangalore Ashok; P. Oinas; S. Forssell, 2022, Techno-economic evaluation of a biorefinery to produce  $\gamma$ -valerolactone (GVL), 2-methyltetrahydrofuran (2-MTHF) and 5-hydroxymethylfurfural (5-HMF) from spruce, *Renewable Energy*, 190, pp. 396–407. <https://doi.org/10.1016/j.renene.2022.03.128>
- J. J. Brown; I. Cybulska; T. Chaturvedi; M. H. Thomsen et al., 2014, Halophytes for the Production of Liquid Biofuels, *Sabkha Ecosystems*, Springer, Dordrecht, pp. 67–72. DOI: 10.1007/978-94-007-7411-7\_4
- A. Jdey; H. Falleh; S. Ben Jannet; K. Mkadmini Hammi et al., 2017, Anti-aging activities of extracts from Tunisian medicinal halophytes and their aromatic constituents, *EXCLI Journal*, 16, pp. 755–769. DOI: 10.17179/excli2017-244
- B. Kamm; P. R. Gruber; M. Kamm, 2000, Biorefineries-Industrial Processes and Products, *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, pp. 1–38. [https://doi.org/10.1002/14356007.104\\_101](https://doi.org/10.1002/14356007.104_101)
- A. Patel; A. R. Shah, 2021, Integrated lignocellulosic biorefinery: Gateway for production of second generation ethanol and value added products, *Journal of Bioresources and Bioproducts*, 6, 2, pp. 108–128. <https://doi.org/10.1016/j.jobab.2021.02.001>
- W. Seider; J.D. Seader; D. R. Lewin; S. Widagdo, 2009, Product and process design principles Synthesis, analysis and evaluation. 3. ed. Hoboken, NJ: Wiley.
- T. Taşçı Çilak; S. Fasse; S. Raval et al., 2022, Extraction Methods of Value-Added Compounds from Biomasses, *European Biomass Conference and Exhibition Proceedings*, pp. 512–528 DOI: 10.5071/30thEUBCE2022-3CV.6.26
- AQUACOMBINE, 2019-2023, EU H2020 IA project, Grant agreement ID: 862834, <https://doi.org/10.3030/862834>

# Optimal Control for Deriving Policies for Global Sustainability

Urmila Diwekar<sup>1,2</sup>, Apoorva Nisal<sup>2</sup>, Yogendra Shastri<sup>3</sup>, and Heriberto Cabezas<sup>4</sup>

1. Vishwamitra Research Institute, Crystal Lake, IL, USA
2. The University of Illinois at Chicago, Chicago, IL, USA.
3. Indian Institute of Technology, Mumbai, India.
4. University of Miskolc, Miskolc, Hungary

## Abstract

Inordinate consumption of natural resources by humans over the past century and unsustainable growth practices have necessitated a need for enforcing global policies to sustain the ecosystem and prevent irreversible changes. This study utilizes the Generalized Global Sustainability model (GGSM), which focuses on sustainability for the Food-Energy-Water (FEW) Nexus. GGSM is a 15-compartment model with components for the food web, microeconomic framework, energy, industry and water sectors, and humans. It was validated based on historical data for global sectors and can predict population, global and regional water stress, GHG emissions, and the gross domestic product (GDP) for the next century. GGSM shows that an increased per capita consumption scenario is unsustainable. In this study, an optimal-control theory-based approach is devised to address the unsustainable scenario through policy interventions to evaluate sustainability by employing multiple global indicators and controlling them.

**Keywords:** GGSM, global sustainability, Fisher information, optimal control

## 1. Introduction

Global models depicting the state of the world help in the study of the complex interconnected dynamics of sustainable growth and provide a scientific basis for the decision-making process behind global policies. Some of the noteworthy global models are given below.

- WORLD 3 Model (Meadows et al., 1972[1,2])
- GUMBO Model (Boumans et al., 2002[3])
- Puliato et al. (2008)[4]: “Predator-Prey” model for population and gross domestic product (GDP)
- HANDY Model (Motesarra et al., 2014[5])
- EARTH 3 Model (Randers et al., 2017[6])
- HARMONEY Model (King et al., 2020 [7])

However, these models have the following deficiencies.

- These models are mostly empirical models
- They are not focused on FEW nexus
- They are mostly descriptive and not Prescriptive.

To circumvent these deficiencies, we developed a new global model called Generalized Global Sustainability Model (GGSM). This model consists of differential-algebraic equations focused on FEW nexus, both descriptive and prescriptive.

## 2. The GGSM Model

In this section, the generalized global sustainability model and its various features will be described. The current model is conceptually similar to the previous model (Kotecha et al., 2013 [8]). This work's exception is that it is a global model and focuses on the food-energy-water (FEW) nexus. GGSM was created to represent a simplified global ecosystem with enough detail to allow the pursuit of further study. The 15-compartment model consists of a simplified ecological food web set in a macroeconomic framework and a rudimentary legal system. The different compartments in the model, shown in Fig 1, are three primary producers (P1, P2, and P3), three herbivores (H1, H2, and H3), two carnivores (C1 and C2), human households (HH), industry (IS), energy producer (EP), fuel source, water reservoir, resource pool (RP) and an inaccessible resource pool (IRP).

The food web is modeled by Lotka–Volterra type expressions, whereas the economy is represented by a price-setting model wherein firms and human households attempt to maximize their economic well-being. The model development is done using global-scale data for stocks and flows of food, energy, and water, which were used to parameterize this model. Appropriate proportions for some of the ecological compartments, like herbivores and carnivores, are used to model those compartments. The modeling of the human compartment was carried out using historical data for the global mortality rate. Historical data were used to parameterize the model.

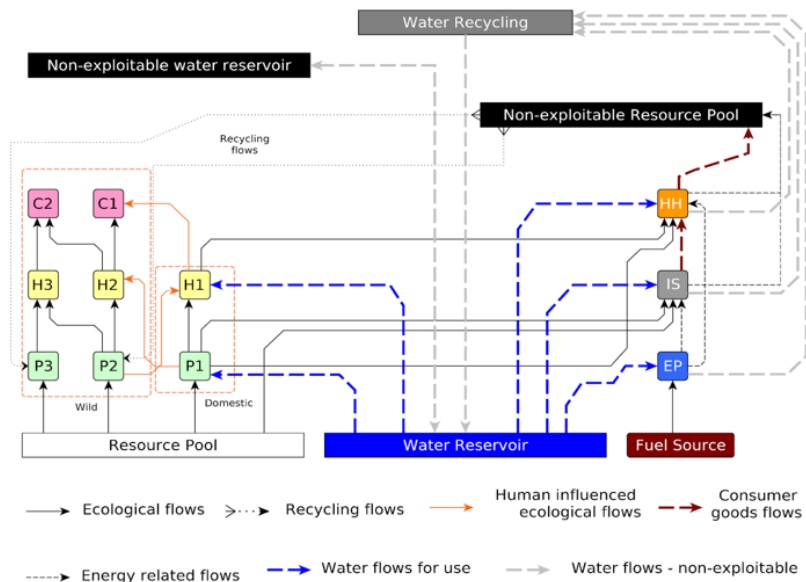


Fig 1. Generalized Global Sustainability Model (GGSM). Compartments: primary producers (P1, P2, P3), herbivores (H1, H2, H3), carnivores (C1, C2), human households (HH), industry sector (IS), energy producer (EP), fuel source, water reservoir, resource pool, inaccessible resource pool.

We used data for key variables like the human population, GDP growth, and greenhouse gases like CO2 and NOX emissions [9-12] to validate the model, as shown in Figure 2. The validated model becomes the foundation for the next part of the study and will be referred to in the following section as the "Base Case ."The model was then used to make long-term forecasts and study global sustainability over time.

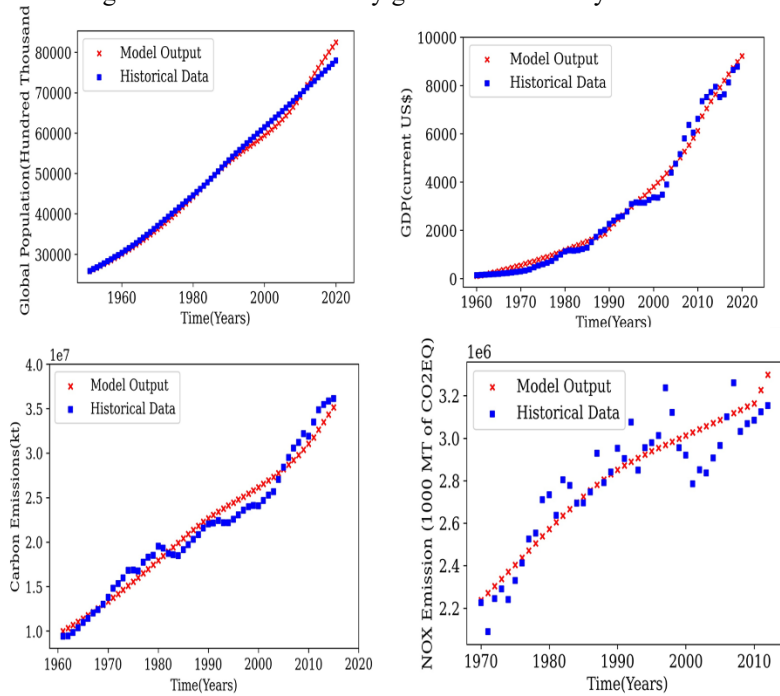


Figure 2: Results of Model Validation for Key Outputs

### 3. Scenario Analysis

The base case scenario represents the current state of the world. The other scenarios we considered are given below. We consider the forecast up to the year 2150.

- Scenario 1: Base Case: Business at the current level
- Scenario 2: Population Explosion: Human population grows at a rapid rate and then tapers off.
- Scenario 3: Consumption Increase: Consumption levels increase at a higher rate
- Scenario 4: Population Explosion with Consumption Increase: Both the population and consumption levels increase together

The ecological compartment results of scenarios 1 and 2 show that the planet can support the increase in population if human consumption does not increase significantly. However, for scenarios 3 and 4, the compartments start dying, as shown in Figure 3 for primary producer P1.

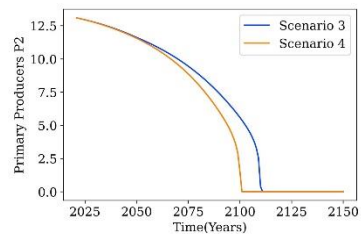


Figure 3: Scenarios 3 & 4, for P1

We know that these scenarios are the likely scenarios of our future. This means our planet will become unsustainable within this time frame. We are using optimal control theory to derive optimal techno-socioeconomic policies to make the planet sustainable in the above time frame.

#### 4. Deriving Techno-Socio-Economic Policies

GGSM showed that an increased per capita consumption scenario is unsustainable. In this study, an optimal-control theory-based approach is devised to address the unsustainable scenario through policy interventions to evaluate sustainability by employing multiple global indicators and controlling them given below.

- Fisher Information (FI): Measurement of the stability of the system
  - Fisher Information (FI) has been used as a measure of sustainability (Doshi et al. 2015 [13])
  - Thus, the objective for the optimal control problem is to minimize the variance of FI between the unsustainable system and an ideal sustainable system or the base case
- Ecological Buffer
  - Equivalent land demand of the population or land required to support the annual average consumption and the waste production of an individual
- Green Net Product(GNP)
  - Estimation of sustainability from a social welfare standpoint
  - Sum of the real value of consumption and real value of net investments
- Gross Domestic Product (GDP)
  - Economic performance evaluation
- Green House Gases (GHG)
  - CO<sub>2</sub> and NO<sub>x</sub> emissions to regulate total CO<sub>2</sub>EQ contribution by humans
- Global Water Stress
  - A measure of the stress on global water systems

The optimal control model assesses these multiple objectives by minimizing the variance in the Fisher Information. One significant result from this study is that optimizing for the Fisher Information-based objective is adequate to attain sustainability and manage the other objectives under consideration, as can be seen in Figure 4. Thus, forgoing a multi-objective problem framework. The results show that cross-dimensional policy interventions (Figure 5) such as increased vegetarianism (a), increased industrial efficiency (b), and increased penalty on industrial discharge (c) are shown to have a positive impact on scale. It should be noted that this increase in discharge fees for industries does not affect the GDP, as shown in Figure 4.

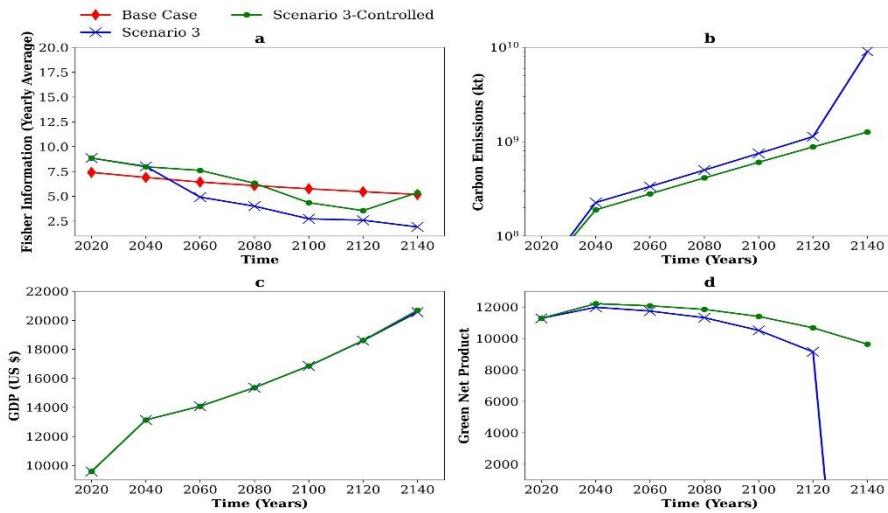


Figure 4: Behavior of Various Objectives when Fisher Information Variance Minimized

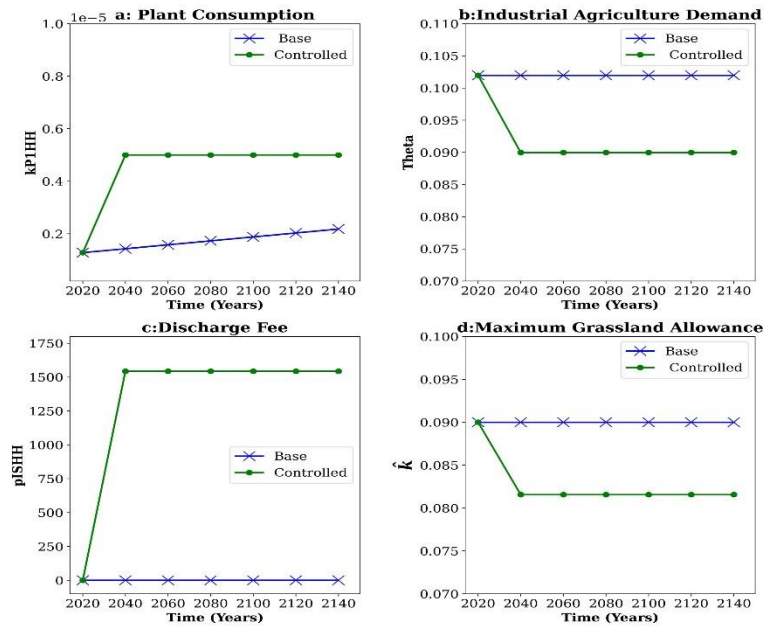


Figure 5: Optimal Profiles of Some of the Decision Variables

## 5. Summary

The Generalized Global Sustainability Model (GGSM), including ecological and economic dimensions to study the Food, Energy, and Water (FEW) Nexus is developed. Irrational human consumption would lead to an unsustainable future. Water availability is adequate, but regional water stress can be concerning. **Fisher Information**-based

optimization is enough to manage other sustainability objectives. GGSM can provide **techno-socioeconomic policy solutions** to make sustainable changes globally

## References

1. The World3 Model: Classic World Simulation | Insight Maker. The World3 Model: Classic World Simulation | Insight Maker; 2021.
  2. Meadows D, Randers J, Meadows D. Limits to growth: The 30-year update. Chelsea Green Publishing; 2004.
  3. Boumans R, Costanza R, Farley J, Wilson MA, Portela R, Rotmans J, et al. Modeling the dynamics of the integrated earth system and the value of global ecosystem services using the GUMBO model. *Ecological economics*. 2002; 41(3):529–560.
  4. Puliafito SE, Puliafito JL, Grand MC. Modeling population dynamics and economic growth as competing species: An application to CO<sub>2</sub> global emissions. *Ecological Economics*. 2008; 65(3):602–615.
  5. Motesharrei S, Rivas J, Kalnay E. Human and nature dynamics (HANDY): Modeling inequality and use of resources in the collapse or sustainability of societies. *Ecological Economics*. 2014; 101:90–102.
  6. Randers J, Rockström J, Stoknes PE, Goluke U, Collste D, Cornell SE, et al. Achieving the 17 Sustainable Development Goals within 9 planetary boundaries. *Global Sustainability*. 2019; 2.
  7. King CW. An integrated biophysical and economic modeling framework for long-term sustainability analysis: The HARMONEY model. *Ecological Economics*. 2020; 169:106464.
  8. Kotecha P, Diwekar U, Cabezas H. Model-based approach to study the impact of biofuels on the sustainability of an ecological system. *Clean Technologies and Environmental Policy*. 2013; 15(1):21–33.
  9. The World Bank. CO<sub>2</sub> Emissions (kt) Data; 2020. Available from <https://data.worldbank.org/indicator/EN.ATM.CO2E.KT>.
  10. The World Bank. GDP (current USD) Data; 2020. Available from: <https://data.worldbank.org/indicator/NY.GDP.MKTP.CD>.
  11. The World Bank. Nitrous Oxide Emissions (thousand metric tons of CO<sub>2</sub> equivalent) Data; 2020. Available from: <https://data.worldbank.org/indicator/EN.ATM.NOXE.KT.CE>.
  12. United Nations D. World Population Prospects Highlights, 2019 revision Highlights, 2019 revision. United Nations; 2019.
  13. Doshi R, Diwekar U, Benavides PT, Yenkie KM, Cabezas H. Maximizing sustainability of ecosystem *Technologies and Environmental Policy*. 2015; 17(6):1573–1583.
- Computer Code and All the data for the GGSM model is available from: <https://zenodo.org/record/6331602>.

# Application of CAPE Tools into Prospective Life Cycle Assessment: A Case Study in Feedstock Recycling of Waste Plastics

Yasunori Kikuchi,<sup>a,b,c\*</sup> Yuki Nomura,<sup>c</sup> Takuma Nakamura,<sup>c</sup> Shoma Fujii,<sup>a</sup>  
Aya Heiho,<sup>b</sup> Yuichiro Kanematsu<sup>b</sup>

<sup>a</sup> *Institute for Future Initiatives, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8654, Japan*

<sup>b</sup> *Presidential Endowed Chair for “Platinum Society”, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*

<sup>c</sup> *Department of Chemical System Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*  
*ykikuchi@ifl.u-tokyo.ac.jp*

## Abstract

In this study, we are tackling systems design with assessments for emerging technologies. Because of the data limitation on the systems and processes adopting emerging technologies, computer-aided process engineering (CAPE) tools such as process design heuristics, process simulation, optimization, parametric analysis for characterizing sensitivity and alternative generation. In this paper, we examine the feedstock recycling of waste plastics for raw materials in refineries. As for the pretreatment process of waste plastics, catalytic pyrolysis is applied. By modifying the condition of catalysts and reaction conditions, thermal decomposition of waste plastics into chemical feedstock was experimentally demonstrated to obtain low-molecular-weight hydrocarbons. This study converts the experimental results into process simulation models to conduct prospective life cycle assessment (LCA). For such purpose, the mathematical models on catalytic reactions and oil refinery were developed. The results showed that the performance of plastics pyrolysis could be linked with the improvements of environmental performances.

**Keywords:** catalytic pyrolysis, LCA, process modeling, process simulation.

## 1. Introduction

When carbon neutrality (CN) is achieved, assuming that no fossil carbon is used, carbon sources in the chemical industry will be largely limited to those from biomass, recycling, and CO<sub>2</sub> capture from the air (Meng et al., 2022). The current availabilities of these carbon sources are not sufficient to meet the demand for chemical raw materials due to their individual issues such as the limitation of the expansion of agriculture and forestry, the quality of waste-derived carbon, and the low technology readiness level (TRL). While the availability of biomass-derived carbon is limited (Energy Transition Commission, 2021), it should be utilized as carbon sources considering the effect of de-fossilization if waste biomass is included (Kikuchi et al., 2022) and the economic spill-over effects on agriculture and forestry (Kikuchi et al., 2020). Waste plastics as resources can be pyrolyzed to become hydrocarbons similar in composition to those used in conventional refineries (Koller et al., 2022). Although the qualities of waste plastics can be largely



different from the collection routes and the purpose of use of plastic products, there is a possibility that they can be converted into hydrocarbons inputted to refineries through appropriate treatment.

In this study, we are tackling systems design with assessments for emerging technologies. Because of the data limitation on the systems and processes adopting emerging technologies, computer-aided process engineering (CAPE) tools such as process design heuristics, process simulation, optimization, parametric analysis for characterizing sensitivity and alternative generation. In this paper, we examine the feedstock recycling of waste plastics for raw materials in refineries. As for the pretreatment process of waste plastics, catalytic pyrolysis is applied. By modifying the condition of catalysts and reaction conditions, thermal decomposition of waste plastics into chemical feedstock was experimentally demonstrated to obtain low-molecular-weight hydrocarbons. This study converts the experimental results into process simulation models to conduct prospective life cycle assessment (LCA). For such purpose, the mathematical models on catalytic reactions and oil refinery were developed. For catalytic cracking, the yield of the reaction was set and the inventory per unit of plastics was estimated by process simulation using Aspen HYSYS™. The yields and complexity factors for each unit operation in the refinery were taken from previous inventory data, and the energy consumption of the entire refinery was distributed to each process by the throughput and complexity factors. Naphtha reforming was assumed to be equivalent to that in the refinery, and inventory data for naphtha cracking were taken from a LCA database.

## 2. Materials and methods

### 2.1. Application of CAPE tools into prospective LCA

Figure 1 shows the description of systems assessments applying CAPE tools for prospective LCA. In management activity and resource provider, data estimation and interpretations are assigned to CAPE tools considering the conditions in prospective LCA.

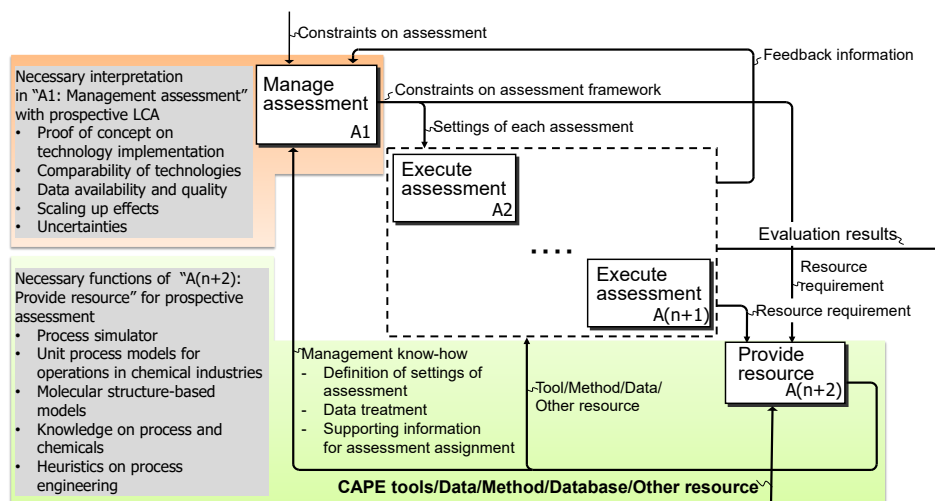


Figure 1 Description of multiple assessment activities with the necessary conditions for prospective LCA.  $n$  is the number of assessment methods. (Modified from previous studies (; Kikuchi, 2014; Kikuchi et al., 2010; Kikuchi and Hirao, 2009))

Life Cycle Thinking (LCT) is the process of considering the impact of a product throughout its entire life cycle, from the stage of raw material acquisition to its final disposal, when evaluating a product. LCA is a method for quantifying environmental impacts. There are two types of LCA: attributable LCA, which aims to visualize the environmental attributes of a product during its current life cycle, and consequential LCA, which aims to visualize the changes that occur when changes are made to the current life cycle (Ekvall, 2019). In the introduction of technological systems, where significant changes from the current system may occur, an analysis based on consequential LCT is necessary. The introduction of biomass-derived/waste-derived production technologies in chemical production and related life cycles, which are mainly based on the production of products derived from fossil resources, will bring about obvious changes and requires the design of the right type and scale of technical systems based on an assessment that takes into account the consequential impacts induced by these changes.

Conventional LCA can be carried out by collecting data on the target technology or system from actual production sites and combining other data extracted from LCA databases and other sources. On the other hand, in order to consider the development and introduction of CN and other sustainability-oriented technologies and systems, it is necessary to obtain information on non-existing life cycles and to carry out LCA. The applicability of LCA methods for the future of such technologies and systems is discussed (van der Giesen, 2020). LCA methods such as Consequential LCA, Dynamic LCA, Anticipatory LCA, Prospective LCA, and Ex-Ante LCA can be used for LCA of technologies under development, in other words technologies with low TRL, and systems after their introduction on a larger scale than at present, and for LCA of technologies that do not yet exist.

For the generation of life cycle inventory data, CAPE tools could be applicable. The previous review (Parvatker and Eckelman, 2019) demonstrated that the generation methods of inventory data should be selected in terms of the accuracy and the data/time requirements considering assessment objectives. For estimating inventory data, the applicable CAPE tools include process simulation, process input and output calculations, stoichiometry, and molecular structure-based models such as quantitative structure activity relationship (QSAR) models. Even if no stoichiometric/synthesis information is available, the inventory data of proxy chemical, which has already registered in LCA databases, can be selected with the understandings of chemistry of the technologies. While these CAPE tools have been developed mainly for chemical process systems, the inventory data of mechanical process systems such as metal processing and assembling might be estimated in the same manner.

### *2.2. Case study: Feedstock recycling of waste plastics*

Figure 2 shows the life cycle boundary of plastics including catalytic pyrolysis as feedstock recycling of waste plastics for raw materials in refineries. Catalytic pyrolysis can produce hydrocarbons similar in compositions to those used in conventional oil refineries. The profiles of obtained hydrocarbons have various patterns based on the temperature of pyrolysis and compositions of resins in waste plastics. There is a distribution of carbon chains in the pyrolysis products, and it is not easy to extract only specific hydrocarbons. Therefore, if the pyrolysis products are fed directly into the refinery flow, they may be fed together with crude oil into an atmospheric distillation column. The catalytic cracking process can be followed by a separation process to

separate the pyrolysis products into groups of certain carbon chains. This allows the products to be fed directly into certain unit operations of the refinery, thereby omitting some unit operations of the refinery.

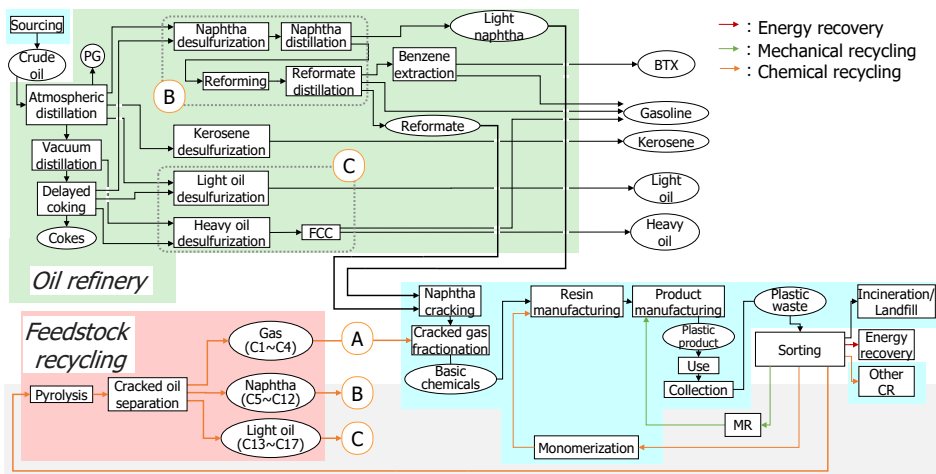


Figure 2 Life cycle boundary of plastics including catalytic pyrolysis as feedstock recycling of waste plastics for raw materials in refineries, where CR means chemical recycling, MR is material recycling, PG is purged gas, and FCC is fluidized-bed catalytic cracking.

The pyrolysis products were separated into three groups based on their carbon chains. Product A contains hydrocarbons (C1 to C4) which could be directly used as raw materials for basic chemical products through fractionation processes. This is similar to the stream after naphtha cracking in conventional petrochemical process systems. Product B is similar to naphtha, containing hydrocarbons (C5 to C12). This is fed to the unit operations around naphtha production in oil refinery. Product C contains heavy content and must be fed to FCC for further cracking. The unit operations and flows inside of oil refinery were extracted from the reports on existing refinery (Yoshitome et al., 2022).

Process simulation was applied as a CAPE tool to estimate inventory data. For catalytic cracking, the temperature dependence of pyrolysis products was collected from literature values, and the production yield from waste plastics was calculated (Keller et al., 2022). For the utility of catalytic cracking, inventory data for existing pyrolysis were used (Kikuchi et al., 2022). For inventory data in the separation system after catalytic cracking, the data were estimated using Aspen HYSYS<sup>TM</sup>. For unit operations in the refinery, the inventory of the existing refinery was calculated from the complexity factors and the utility of the entire refinery. For unit operations where the flow rate increases due to the merging of waste plastic-derived components, the increase in energy input due to the increased flow rate was calculated using a linear approximation. For each unit operation, a facility capacity was set.

### 3. Results and discussion

Figure 3 shows the example of process simulation results; greenhouse gas (GHG) emission for oil refinery and petrochemical process for a conventional oil refinery in Japan, the throughput of which is 18,115 kL-crude oil per day. The functional unit is set as the operation of this oil refinery plant. The results of this prospective LCA

demonstrated that the catalytic pyrolysis of waste plastics could reduce the GHG emissions attributable to oil refinery by reducing refined crude oil. Based on the compositions of pyrolysis products, for example, the flowrates in oil refinery were changed, which resulted in the changes of the product ratios in oil refinery. If the pyrolysis products contain hydrocarbons similar to naphtha or ones from naphtha cracker, the carbon source from waste plastics can be easily circulated as raw materials of chemicals.

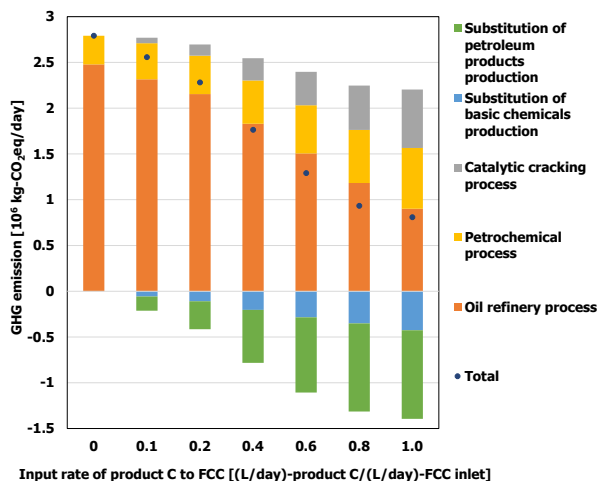


Figure 3 Example of process simulation results; GHG emission for oil refinery and petrochemical process for a conventional oil refinery in Japan, the throughput of which is 18,115 kL-crude oil per day. The inputted waste plastics were capped by the capacity of FCC. The background process inventories were extracted from Japanese LCA database (IDEA v3.2).

Converting waste plastics to oil through pyrolysis allows the reuse of existing refinery equipment, which may reduce the barriers to the introduction of new technologies. On the other hand, the composition of hydrocarbons obtained by catalytic cracking of waste plastics is different from that contained in crude oil. Therefore, when pyrolysis products derived from waste plastics are fed into refineries, the composition of the products obtained as petroleum derivative products will change. Although it is possible to adjust the composition of the product system by changing the temperature and catalyst in catalytic cracking, it is necessary to modify the refinery to meet future demand for petroleum products. These issues need to be analyzed through LCAs that take into account the consequences of other social scenarios, such as reduced gasoline consumption due to the electrification of automobiles and reduced demand for fossil fuels due to changes in the composition of power sources.

#### 4. Conclusions

In this study, a mathematical model was constructed for catalytic pyrolysis of waste plastics and oil refinery receiving the pyrolysis products as raw materials as feedstock recycling of plastics. As a result, we confirmed the changes in GHG emissions from the production of various petroleum-based products and products due to the introduction of catalytic pyrolysis. The effect of conversion of waste plastics to chemical feedstock aimed at reducing fossil resource consumption cannot be fully assessed only by evaluating the GHG emissions resulting from the substitution of fuel oil and chemical feedstock. Changes in demands of energy and chemicals in society should be taken into account. As a sustainable or recyclable carbon sources, waste plastics should be able to become a

stable source which can be gathered from industries and municipalities. However, the current waste plastics are contaminated with other materials such as papers, metals, sands, garbage, and so on. Additionally, the resin types are also mixed in waste plastics. The pretreatment process, i.e., sorters, is important for the quality of raw materials for catalytic pyrolysis. The prospective design of life cycle should be conducted with social systems and supported by LCA with CAPE tools.

### Acknowledgement

This is based on results obtained from a project, JPNP20012, subsidized by the New Energy and Industrial Technology Development Organization (NEDO). This work was supported by MEXT/JSPS KAKENHI Grant Number JP20K20016, 21K17919, and JST COI-NEXT JPMJPF2003. Activities of the Presidential Endowed Chair for “Platinum Society” at the University of Tokyo are supported by Mitsui Fudosan Corporation, Sekisui House, Ltd., the East Japan Railway Company, and Toyota Tsusho Corporation.

### References

- T. Ekvall, 2019, Attributional and Consequential Life Cycle Assessment, Sustainability Assessment at the 21st century, <https://www.intechopen.com/chapters/69212>
- Energy Transition Commission, 2021, Bioresources within a Net-Zero Emissions Economy: Making a Sustainable Approach Possible, <https://www.energy-transitions.org/publications/bioresources-within-a-net-zero-economy/>
- van der Giesen, C., Cucurachi, S., Guinée, J., Kramer, G.J., Tukker, A., 2020, A critical view on the current application of LCA for new technologies and recommendations for improved practice, *J. Clean. Prod.*, 259, 120904
- F. Keller, R.L. Voss, R.P. Lee, B. Meyer, 2022, Life cycle assessment of global warming potential of feedstock recycling technologies: Case study of waste gasification and pyrolysis in an integrated inventory model for waste treatment and chemical production in Germany, *Resour. Conserv. Recycl.* 179, 106106
- Y. Kikuchi, M. Hirao. 2009, Hierarchical Activity Model for Risk-Based Decision Making Integrating Life Cycle and Plant-Specific Risk Assessments, *J Ind Ecol*, 13(6) 945-964.
- Y. Kikuchi, K. Mayumi, M. Hirao. 2010. Integration of CAPE and LCA Tools in Environmentally-Conscious Process Design: A Case Study on Biomass-Derived Resin, *Computer-Aided Chemical Engineering*, 27, 1051-1056.
- Y. Kikuchi. 2014. Activity and Data Models for Process Assessment Considering Sustainability, *Kagaku Kogaku Ronbunshu*, 40(3) 211-223.
- Y. Kikuchi, M. Nakai, Y. Kanematsu, K. Oosawa, T. Okubo, Y. Oshita, Y. Fukushima, 2020, Application of technology assessments into co-learning for regional transformation: A case study of biomass energy systems in Tanegashima, *Sustain. Sci.*, 15, 1473-1494
- Y. Kikuchi, N. Torizaki, L. Tähkämö, A. Enström, S. Kuusisto, 2022, Life cycle greenhouse gas emissions of biomass- and waste-derived hydrocarbons considering uncertainties in available feedstocks, *Proc. Saf. Environ. Protec.*, 166, 693-703
- F. Meng, A. Wagner, A.B. Kremer, D. Kanazawa, J.J. Leung, P. Goult, M. Guan, S. Herrmann, E. Speelman, P. Sauter, S. Lingeswaran, M.M. Stuchtey, K. Hansen, E. Masanet, A.C. Serrenho, N. Ishii, Y. Kikuchi, J.M. Cullen, 2022, Planet compatible pathways for transitioning the chemical industry, *ChemRxiv*, <https://doi.org/10.26434/chemrxiv-2022-hx17h-v2>
- National Institute of Advanced Industrial Science and Technology, LCA-database IDEA v3.2. (2022)
- A.G. Parvatker, M.J. Eckelman, Comparative evaluation of chemical life cycle inventory generation methods and implications for life cycle assessment results, *ACS Sustain. Chem. Eng.*, 7 (2019), pp. 350-367,
- T. Yoshitome, K. Saito, K. Tamura, 2022, Modeling of refinery equipment configuration and CO<sub>2</sub> emissions, *PETROTECH*, 45 (1), 21-28

# Designing roadmaps for transitioning to value chains with net-zero emissions: Case of the chemical industry

Amrita Sen,<sup>a</sup> Vyom Thakker,<sup>a</sup> George Stephanopoulos,<sup>b,c</sup> Bhavik R Bakshi,<sup>a</sup>

<sup>a</sup> *William G Lowrie Department of Chemical and Biomolecular Engineering, The Ohio State University, Columbus OH 43201, USA*

<sup>b</sup> *The Global KAITEKI Center, Arizona State University, Tempe, Arizona 85287, USA*

<sup>c</sup> *Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

## Abstract

The global economy must restrict net greenhouse gas emissions to the environment to zero by 2050. However, in the absence of rigorous methods, the precise path to net-zero remains ambiguous, despite the overabundance of candidate solutions. Our formulation builds on the existing net-zero roadmapping literature by including technologies at various readiness levels within the life-cycle network superstructure. A case study for chemicals and plastics value chains demonstrates the framework. We model the current global plastics and chemicals industry and the promising low carbon solutions, based on electrification, biomass, carbon capture and circularization. Our framework allows the evolution of background emissions and maturities of technological solutions. We solve a multi-objective optimization problem at discrete time steps to roadmap for projected chemicals production till 2050. Our results show that net-zero value chains are only possible with a combination of energy efficiency and material recycling. Lower TRL technologies enable this transition, but only on upfront investments into their development. Our framework incorporates the realistic limitations of adopting emerging technologies with low TRL values into the decision-making process. The product specific insights gained may be used to guide future strategies.

**Keywords:** net-zero, circular economy, multi-objective optimization, life cycle

## 1. Introduction

The irreversible repercussions of incumbent climate change threaten the future of humankind. This has led policymakers across the world to seek to steer the global economy to operate with net-zero emissions. While most current research efforts in this direction are geared towards delivering cleaner energy with minimum emissions to industrial and residential systems, “net-zero” cannot be reached unless material value-chains are reinvented.<sup>1</sup> The direct emissions from the chemicals sector account for a humble 7% of the global emissions. However, as chemical energy carriers like hydrogen and methanol become increasingly important in other residential sectors, eliminating emissions from this hard-to-decarbonize sector becomes critical.

Many new technologies and innovations have been developed in the recent past to address this transition. However, in the absence of modelling techniques, choosing the most

sustainable, economic and robust solutions is impossible. In addition to numerous life-cycle studies evaluating the emissions abatement potential of emerging technologies in isolation, research has compared the net-zero potential of technology combinations in product systems such as methanol.<sup>2</sup> Existing research has also evaluated the precise combination of technologies for the global chemicals industry to get to net-zero, while investigating possible cost drivers.<sup>3</sup> This steady state analysis, however, does not evaluate nascent innovations or the temporal evolution of technologies. Even though research, such as that by Zibunas et al. account for the capital investments required for technologies to become viable, the gap in developing roadmaps to reach net-zero chemical systems, remains unbridged.<sup>4</sup> The evolution of innovations through time is a crucial consideration in the design of roadmaps and remains unaddressed in current literature.

In this paper, we design a roadmap to guide the transition of the global chemical industry towards net-zero greenhouse gas emissions. Life-cycle assessment methods and the Sustainable Circular Economy (SCE) framework are used to model emerging technologies.<sup>5</sup> The temporal evolution of emerging technologies is utilized to design roadmaps i.e., value chains of chemicals and plastics with the least greenhouse gas emissions or generate the greatest possible value. The linear evolution model that we assume for innovations can be easily replaced by deterministic models and eventually surrogated with simpler functions in the optimization problem.

## 2. Methodology

### 2.1. Model of Chemical Industry

The global chemical industry is modelled as a network of individual processes. We use bottom-up linear steady state models to represent individual processes in the network. These processes are assumed to scale linearly to meet the global demand of chemicals and plastics. This assumption allows the scale-up of linear life-cycle models to meet demands at the economy scale. This version of the study does not model fractional yield and assumes stoichiometric conversion of reactants to products instead. Since the global demand of chemicals and plastics is used to solve for the optimal resource use, the calculated resource use via stoichiometric calculations essentially serves as a lower limit on resource use, which is the worst case scenario. Chemical value chains are modelled including refinery processes to convert fossil feedstock to olefins and aromatics, production of various chemical intermediates, solvents, monomers, fertilizers, and polymerization reactions to produce rubbers, plastics, and plasticizers. The thermal energy and electricity needs of these processes are modelled using global averages of industrial heating and electricity mixes. Our model covers all high-volume chemicals and plastics, and most minor flows in the industry.<sup>9,10</sup> Publicly available data are used to populate the network. These models are represented as a matrix of linear equations, similar to the LCA framework. Processes at the refinery stage are populated with globally averaged data over multiple refineries from industrial reports. Averaged composition of inputs to and outputs from refineries are used for the calculations. Further downstream, conversion and polymerization reactions are modelled using chemical reaction data. The model contains multiple production pathways and final demands are assumed to be fulfilled by appropriate contributions from these pathways. Energy demands of the processes are calculated from specific energy consumption data from IEA.<sup>11</sup> A detailed version of the approach can be found in the authors' past publication.<sup>9</sup>

### 2.2. Modelling emerging technologies

We supplement the baseline network of conventional technologies by models of emerging technologies to obtain a superstructure network with alternatives at all life-cycle stages.

At present, the sector of chemicals production is inherently dependent on carbonaceous fossil feedstock. These feedstocks can be bypassed by more circular pathways such as the utilization of carbon dioxide captured from other emission sources. In this work, we model the conversion of carbon dioxide by electrochemical and thermochemical means to ethylene and methanol, with and without the intermediate formation of syngas. We also model the electrolytic production of hydrogen via water splitting (green hydrogen).

Biomass feedstock may replace most hydrocarbons produced by refinery operations. Our model includes the production of syngas from woody biomass, such as pine and eucalyptus, as well as the catalytic biomass-to-aromatics processes discovered at the Catalysis Center of Energy Innovation, University of Delaware, referred to henceforth as the CCEI process. Since the emerging technologies modelled thus far only produce methanol as a platform chemical, we also include the acid catalyzed transformation of methanol to aromatics to allow production of aromatics derived products.

Mixed plastic waste generated from residential and industrial sectors may be landfilled or recycled to maintain some fraction of its original value. Our superstructure network includes the mechanical (primary) recycling of PET bottles, the chemical (secondary) recycling of PET viz enzymatic hydrolysis and methanolysis reactions, pyrolysis (tertiary recycling) of mixed plastic waste to produce pyrolysis gas as well as syngas, and the transformation of waste to energy (tertiary recycling) via incineration. We also model the recovery of monomers from polyethylene by pyrolysis to produce a mixture of hydrocarbons. Recycling is included in the superstructure network as circular flows which intrinsically determine the emissions associated with circular flows, instead of estimating them with stationary emission factors, using the SCE framework.<sup>5</sup>

Refinery operations, boilers and furnaces are responsible for the majority of the emissions from the chemical industry. We model steam methane reforming of natural gas, partial oxidation of coal and oil, steam cracking of natural gas and naphtha, and catalytic reforming of naphtha to be amenable to post combustion capture using Monoethanol amine (MEA), since flue gases from these processes have relatively high concentrations of carbon dioxide (5-15%).<sup>6</sup> The carbon dioxide thus captured is considered available for utilization i.e., conversion to value added products.

### 2.3 Cost Models

We model the costs of each module as the difference between product and resource costs. This formulation allows the estimation of emerging technological pathways with unknown costs, in terms of costs of input and output flows. We implicitly assume that emerging technologies replace conventionally made products at the same price point. This assumption requires investments into research and development, policy intervention and sharing of environmental burden across stakeholders.

### 2.4 Road-mapping Framework

$$\begin{aligned}
 \min_{s_t} Z(s_t) &:= -T_t = -C^{Module} s_t \\
 \min_{s_t} Z(s_t) &:= h_{GWP} = \phi_{GWP} r_t \\
 \text{s. t. } &As_t \geq f_t \\
 &r_t \leq r_t^U \\
 &0 \leq s_t \leq s_t^U y_t \\
 &\mathcal{F}(s_t \geq 0) \\
 \text{where } &f_t = f_0 + f_0 \gamma t \\
 &r_t = B s_t \\
 &r_t^U = r_0^U + r_0^U \beta T \\
 &t \in T; T = \{0, 1, 2, \dots, 10\}
 \end{aligned}$$



$$\begin{aligned}
 y_t &= \{y_{i,t}\}; i \leq m; y_{i,t} \in \{0,1\} \\
 y_{i,(t+1)} &= 1 \text{ if } t \geq (11 - TRL_{i,t}) \forall i \in I \\
 TRL_{i,t+1} &= 1 + TRL_{i,t}
 \end{aligned}$$

We perform a multi-objective optimization on our superstructure network, minimizing greenhouse gas emissions or maximizing the value generated by chemicals production. The epsilon constraint method is used to draw Pareto fronts through the optimal points. The underlying superstructure network is subjected to fulfillment of societal needs of chemicals. The global chemicals capacity  $f_t$  is used to impose a lower limit on production, which evolves linearly through subsequent time steps at current growth rates.  $\gamma$  is the vector of growth rates. An upper limit  $r_t^U$  is imposed on the availability of some resources like biomass, which also evolve linearly over time. The growth rate of biomass availability,  $\beta$ , is estimated using the current availability of 35 EJ of bioenergy and the assumption that, of the 150 EJ biomass required for sustainable bio-energy production by 2050, 90 EJ will be available for chemicals production by 2050.<sup>7</sup> Renewable electricity is also assumed to evolve linearly from generating at the most, 30% of the supply at the beginning, to 85% at the last time step.<sup>8</sup> Our formulation restricts the inclusion of technologies within the superstructure, depending upon their Technology Readiness Levels (TRL). The parameter  $y_{i,t}$  determines whether a technology is available at time constraining the scale up of technologies until they are mature enough. All immature technologies ( $TRL < 10$ ) are assumed to reach the next TRL at each time step.  $y_{i,t}$  is updated accordingly, A technology is considered to be mature and ready to be adopted once its TRL reaches 10. Finally, governing equations are imposed as inequality constraints, ensuring no non-negative product accumulation.

### 3. Results and Discussions

This optimization problem, formulated as a linear program with 136 variables and 259 constraints, yields globally optimal solutions within reasonable time ( $\leq 1$  min) on an Intel Core i7-1165G7 CPU 2.8 GHz with 12 GB RAM. We visualize the trade-offs between emissions and cost as Pareto curves at each time step in Fig 1. Individual Pareto fronts indicate the lowering of value addition as emissions are progressively decreased. Any point on a Pareto front represents a value chain, all of which are equivalent solutions and improving one of the objectives worsens the other. However, as more technologies mature with each time step, the Pareto curves shift, allowing improvement in both objectives. The technologies introduced at each time step are indicated on the figure. We observe greater value generation even as emissions decrease, as products of the same value can now be generated with consumption of less valuable resources and inputs. The introduction of mechanical recycling shows greater value creation with the retention of plastics within the economy, even though emissions do not decrease significantly. The first net-zero solution is obtained on the introduction of biomass-based processes, which allow the sequestration of carbon during growth phase of the biomass. The shift of Pareto fronts between time steps 3 and 4 is win-lose as thermochemical carbon dioxide conversion pathways can lower emissions but are unable to create greater value. A huge jump in value addition results when plastic waste pyrolysis is introduced in the superstructure, allowing value creation in the form of chemical products and diesel, from waste. Finally, as electrochemical processes mature enough to be adopted, the resource

cost decreases significantly, along with emissions. We plot the adoption of individual technologies at each time step to better analyze how the value chains change with time.

Figure 2 shows the amounts of key hydrocarbons produced by technologies. Benzene is initially produced by a combination of steam cracking and toluene dehydrogenation and disproportionation. As soon as the recovery of monomers from polyethylene matures, it becomes economically viable to produce benzene primarily by the chemical recycling of polyethylene. The use of toluene as feedstock for benzene production is completely replaced. Methanol, initially produced by syngas from steam methane reforming, first switches to being produced thermochemically from carbon dioxide, and eventually to production from syngas. To track the generation of syngas, we look at hydrogen production. Initially, renewables support the production of green hydrogen. However, as more biomass becomes available, the system adopts syngas production from woody biomass. Finally, as the pyrolysis of mixed plastic wastes matures, the pyrolysis of plastics is adopted to supplement biomass. We also see the re-adoption of methane reforming when carbon capture becomes viable. Lastly, the production of xylene from catalytic reforming is phased out and replaced by production from methanol and biobased syngas when readiness levels permit.

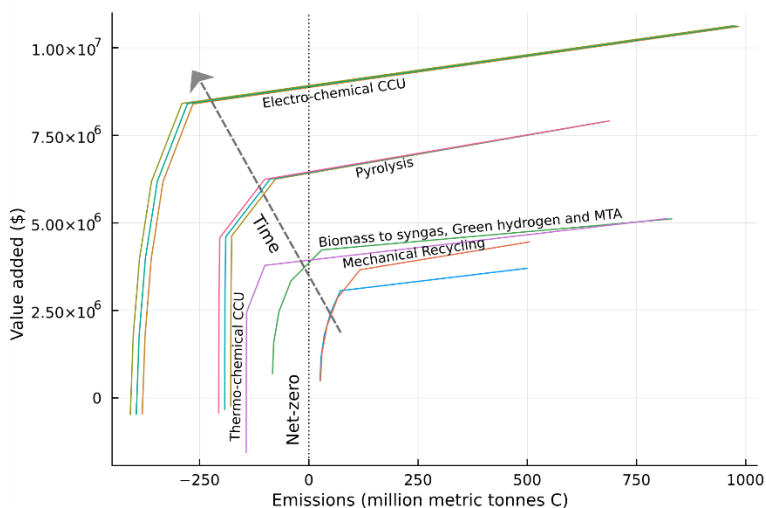


Figure 1: Improvement of both emissions and value addition objectives with adoption of new technologies over time

#### 4. Conclusions and Future Work

Our results reconfirm the need to combine all low carbon solutions to cross the net-zero obstacle and get win-win solutions. Solutions like electrochemical conversion of carbon dioxide to hydrocarbons and bio-based chemicals emerge as robust solutions. Refinery operations such as catalytic reforming are replaced unless they can be retrofit with carbon capture technologies. Additionally, unless constraints are levied on landfilling for example, fossil based CCU solutions are preferred to the recycling of plastics. Ultimately, the ability of nascent technologies to ensure net-zero emissions is contingent on upfront investments into their research and development as well as policy interventions to ensure they are economically competitive with their fossil derived counterparts. The preliminary

findings from our work indicate that emissions abatement, as well as cost, are highly sensitive to the constraints imposed on the network. Therefore, utilizing predictions of renewables and resource availabilities, assessing climate change scenarios to ensure resilience to background emissions, and utilizing historical data to predict technological advancements are crucial. This work will be extended in the near future, to include integrated assessment models, utilize research and development costs as well as patent literature to stochastically predict technology evolution.

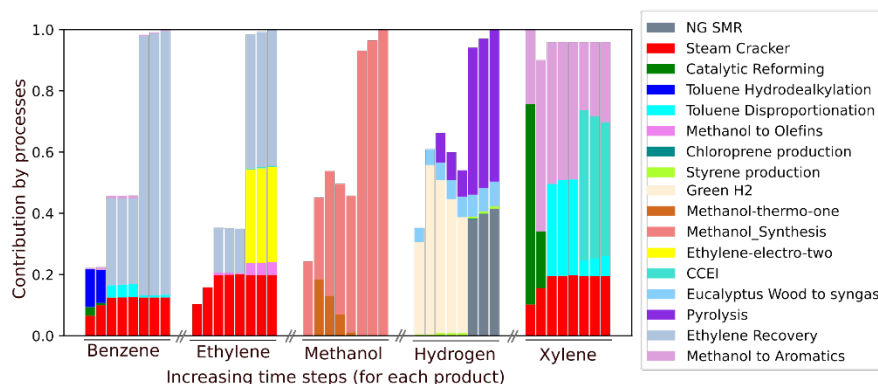


Figure 2: Contribution of processes to key hydrocarbon production. Bio-based and electrochemical processes are readily adopted as soon as they mature.

## References

- <sup>1</sup>Stephanopoulos, G., Bakshi, B.R. and Basile, G., 2022. Reinventing the Chemicals/Materials Company: Transitioning to a Sustainable Circular Enterprise. In *Computer Aided Chemical Engineering* (Vol. 49, pp. 67-72). Elsevier.
- <sup>2</sup>Gabrielli, P., Gazzani, M. and Mazzotti, M., 2020. The role of carbon capture and utilization, carbon capture and storage, and biomass to enable a net-zero-CO<sub>2</sub> emissions chemical industry. *Industrial & Engineering Chemistry Research*, 59(15), pp.7033-7045.
- <sup>3</sup>Meys, R., Kätelhön, A., Bachmann, M., Winter, B., Zibunas, C., Suh, S. and Bardow, A., 2021. Achieving net-zero greenhouse gas emission plastics by a circular carbon economy. *Science*, 374(6563), pp.71-76.
- <sup>4</sup>Zibunas, C., Meys, R., Kätelhön, A. and Bardow, A., 2022. Cost-optimal pathways towards net-zero chemicals and plastics based on a circular carbon economy. *Computers & Chemical Engineering*, 162, p.107798.
- <sup>5</sup>Thakker, V. and Bakshi, B.R., 2021. Toward sustainable circular economies: A computational framework for assessment and design. *Journal of Cleaner Production*, 295, p.126353.
- <sup>6</sup>Van Straelen, J., Geuzebroek, F., Goodchild, N., Protopapas, G. and Mahony, L., 2010. CO<sub>2</sub> capture for refineries, a practical approach. *International Journal of Greenhouse Gas Control*, 4(2), pp.316-320.
- <sup>7</sup>IEA (2021), What does net-zero emissions by 2050 mean for bioenergy and land use?, IEA, Paris <https://www.iea.org/articles/what-does-net-zero-emissions-by-2050-mean-for-bioenergy-and-land-use>, License: CC BY 4.0
- <sup>8</sup>Global Energy Perspective, 2022, Mckinsey & Company
- <sup>9</sup>Sen, A., Stephanopoulos, G. and Bakshi, B.R., 2022. Mapping Anthropogenic Carbon Mobilization Through Chemical Process and Manufacturing Industries. In *Computer Aided Chemical Engineering* (Vol. 49, pp. 553-558). Elsevier.
- <sup>10</sup>Sen, A., Stephanopoulos, G. and Bakshi, B.R., 2022. Mapping the Chemicals and Materials industry. In preparation
- <sup>11</sup>IEA, 2013. Energy and GHG Reductions in the Chemical Industry via Catalytic Processes.

# Data Envelopment Analysis of Ammonia and Methanol as Sustainable Energy Storage Vectors

Sebastiano C. D'Angelo,<sup>a</sup> Michael Bregy,<sup>a</sup> Philipp Steiner,<sup>a</sup> Raul Calvo-Serrano,<sup>a</sup> Gonzalo Guillén-Gosálbez<sup>a</sup>

<sup>a</sup>*Institute for Chemical and Bioengineering, Department of Chemistry and Applied Biosciences, ETH Zürich, Vladimir-Prelog-Weg 1, 8093 Zürich, Switzerland*  
[gonzalo.guillen.gosalbez@chem.ethz.ch](mailto:gonzalo.guillen.gosalbez@chem.ethz.ch)

## Abstract

The increasing concern about global warming presses for a transition to more sustainable energy sources such as solar and wind power. To tackle the associated intermittency, chemical energy storage is witnessing a surge in interest from investors and policymakers. The efficient development of such technologies requires comprehensive analyses encompassing: (1) power-to-chemicals (P2C), (2) storage, and (3) chemicals-to-power (C2P). This work aims to compare ammonia and methanol as chemical energy vectors with other energy storage technologies. Data envelopment analysis (DEA) was used to compare the assessed technologies in terms of economic and environmental life cycle assessment (LCA) indicators, ranking the different technologies and estimating their required improvement targets. Overall, this work sheds light on the potential role of chemicals in energy storage applications, analyzing their advantages and disadvantages, and proposing improvement strategies to facilitate their development and future implementation.

**Keywords:** ammonia, methanol, energy storage, life cycle assessment, data envelopment analysis.

## 1. Introduction

A sustainable transition of the power sector to a decarbonized economy calls for an increased penetration of renewables, primarily solar and wind power. However, the intermittency of such technologies constitutes a challenge: new energy storage solutions are urgently needed, featuring high scalability and long-duration support. Highly suitable vectors for such a purpose are platform chemicals able to effectively store hydrogen (H<sub>2</sub>) in high energy density forms, such as ammonia (NH<sub>3</sub>) and methanol (MeOH) (Davies et al., 2020). Accordingly, a large set of technologies to convert power to chemical storage vectors (P2C) and then back to power (C2P) has been recently investigated, giving rise to a plethora of routes with wide variations in economic and environmental performance and inherent trade-offs. Systematic approaches to quantify and assess such trade-offs could, thus, support informed decisions in developing more sustainable routes.

In this work, acknowledging the scarce literature using similar case studies (Rostami et al., 2022), we apply for the first time Data Envelopment Analysis (DEA) to a large set of routes for power-to-chemical-to-power (P2C2P) involving NH<sub>3</sub> and MeOH. The goal is to assess their efficiency by considering multiple criteria simultaneously and to identify margins of improvement for suboptimal scenarios. DEA originally stems from the field of economics (Charnes et al., 1979) and has been only recently applied to sustainability problems (Vázquez-Rowe et al., 2010). Unlike other multi-criteria decision-making tools

(Tzeng and Huang, 2011), DEA does not rely on user-defined weight preferences to quantitatively assess and classify sets of alternatives based on multiple decision criteria simultaneously. Here we shall use DEA to gauge the efficiency of a chemical storage route based on all the considered economic and environmental decision criteria, further ranking all the overall efficient alternatives through the use of super-efficiencies, and estimating the required improvement targets for the sub-efficient routes.

## 2. Methodology

### 2.1. Generation of scenarios

For the P2C section, the Haber-Bosch process was simulated in Aspen Plus<sup>®</sup> to produce NH<sub>3</sub>, using H<sub>2</sub> from different sources and nitrogen from an air separation unit (D'Angelo et al., 2021). For MeOH, the single-step hydrogenation of CO<sub>2</sub> was assumed (González-Garay et al., 2019). Several H<sub>2</sub> routes were considered, including steam methane reforming (H<sub>2</sub>-SMR), H<sub>2</sub>-SMR using carbon capture and storage (CCS) (H<sub>2</sub>-SMR-CCS), biomass gasification with CCS (H<sub>2</sub>-G), and water electrolysis powered by a set of different sources: grid mix (H<sub>2</sub>-CE), nuclear (H<sub>2</sub>-N), wind (H<sub>2</sub>-W), solar (H<sub>2</sub>-S), bioenergy with CCS (H<sub>2</sub>-B), and hydropower (H<sub>2</sub>-H) (D'Angelo et al., 2021). CO<sub>2</sub> from carbon capture from combined cycle power plants (CC) or from direct air capture (DAC) powered by the same sources as for electrolytic H<sub>2</sub> was considered. For the C2P section, for NH<sub>3</sub>, a power plant was simulated (Sánchez et al., 2021) considering three different downstream treatments for the flue gas: H<sub>2</sub> recovery with nitrogen-argon separation (HR-NAS), H<sub>2</sub> recovery with subsequent flue gas release (HR-FGR), and H<sub>2</sub> combustion with nitrogen-argon separation (HC-NAS). For MeOH, the C2P section considered direct combustion (MC), MeOH decomposition (MD), and steam reforming (MR) (Deng and Hynes, 2012), with CCS applied for the flue gas in all cases. The results of all the combinations of feedstocks and C2P alternatives were compared with an adiabatic compressed air energy storage with a modified Kalina cycle (CAES) (Liu and Wang, 2016) and pumped hydro energy storage (PHES) (Dones et al., 2007).

### 2.2. Environmental assessment

The life cycle assessment (LCA) was performed following the ISO 14040 (ISO, 2014). A cradle-to-gate study based on an attributional approach was adopted, considering the whole life cycle until the P2C step. The selected functional unit was 1 kWh of delivered electrical energy from the various storage routes.

In the second LCA phase, the life cycle inventories (LCIs) are modeled. The foreground system includes all the steps from P2C to C2P, according to the defined scenarios. At the same time, the underlying energy and raw materials suppliers belong to the background system, here modeled with Ecoinvent v3.5, accessed through SimaPro v9.2. All the inventories assumed global (GLO) regionalization where available and Rest-of-the-World (RoW) otherwise. For the NH<sub>3</sub> and MeOH P2C steps, the average global grid power mix was considered for the processes, excluding H<sub>2</sub> and CO<sub>2</sub> generation. Salt caverns were assumed for intermediate H<sub>2</sub> storage (D'Angelo et al., 2021), and the oxygen by-product from water splitting was considered vented. In the third LCA phase, ReCiPe 2016 with the hierarchical approach (Huijbregts et al., 2017) was selected, with a focus on the midpoint indicators, namely: global warming (GW), stratospheric ozone depletion (SOD), ionizing radiation (IR), fine particulate matter formation (FPMF), human health ozone formation (OFHH), terrestrial ecosystem ozone formation (OFTE), terrestrial acidification (TA), freshwater eutrophication (FWE), marine eutrophication (ME), water consumption (WC), land use (LU), mineral resource scarcity (MRS), fossil resource scarcity (FRS), marine ecotoxicity (MECOTOX), freshwater ecotoxicity

(FWECOTOX), terrestrial ecotoxicity (TECOTOX), human carcinogenic toxicity (HCT), human non-carcinogenic toxicity (HNCT). Finally, in the last step of LCA, the results are interpreted, and recommendations are drawn. Here, we applied DEA to provide additional insights into the environmental efficiency of the generated scenarios.

### 2.3. Mathematical models of DEA

DEA provides relative efficiencies of the different alternatives (n) as the maximum ratio between the weighted sum of outputs (j, criteria to maximize) to the weighted sum of inputs (i, criteria to minimize). To this end, it solves a linear programming (LP) model, which, when reformulated into its dual form, also provides improvement targets for the different sub-efficient alternatives. In this work, we consider the input-oriented variable return to scale dual model M1, presented below as described by Cooper et al. (2007):

$$\begin{aligned}
 \text{M1} \quad & \min_{\theta_{n'}, S_i^-, S_j^+, \lambda_{n'n}} \theta_{n'} - \varepsilon (\sum_i S_i^- + \sum_j S_j^+) & (1) \\
 \text{s.t.} \quad & \sum_n \lambda_{n'n} x_{ni} + S_i^- = \theta_{n'} x_{n'i}, \forall i \\
 & \sum_n \lambda_{n'n} y_{nj} - S_j^+ = y_{n'j}, \forall j \\
 & \sum_n \lambda_{n'n} = 1 \\
 & \lambda_{n'n}, S_i^-, S_j^+ \geq 0, \forall n, i, j
 \end{aligned}$$

where the model variables include  $\theta_{n'}$ , the efficiency value of alternative  $n'$ ,  $S_i^-$  and  $S_j^+$ , the input and output, respectively, slack variables, and  $\lambda_{n'n}$ , the calculated efficiency projection weights of alternative  $n'$  on the efficient alternatives  $n$ . The model parameters include  $x_{ni}$ , the input  $i$  value for alternative  $n$ ,  $y_{nj}$ , the output  $j$  value for alternative  $n$ , and  $\varepsilon$ , a non-archimedean infinitesimal value.

Model M1 is solved for each of the considered alternatives  $n'$ , obtaining for each one its respective efficiency and projection weights. In essence, the best technologies are those with an efficiency score equal to one, while the rest could match the best-performing units by attaining some improvement targets provided by DEA. More precisely, it is possible to estimate the relative input improvement targets  $\tau_{n'i}$  for the different sub-efficient alternatives  $n'$  and input criteria  $i$  as the relative difference between the current input value  $x_{n'i}$  and the value projected onto the efficient frontier  $\sum_n \lambda_{n'n} x_{ni}$ . The relative input improvement target can be calculated as follows:

$$\tau_{n'i} = 1 - \frac{\sum_n \lambda_{n'n} x_{ni}}{x_{n'i}} \quad (2)$$

In order to rank the alternatives defined as efficient in model M1, it is possible to use super-efficiency models to calculate the super-efficiency scores  $\hat{\theta}_{n'}$  for the already efficient alternatives. These super-efficiency scores can be calculated by model M2, which only differs from model M1 in that the projection constraints consider only the other alternatives  $n$  and not the alternative being evaluated  $n'$ , as seen in the equations below. As in model M1, model M2 is solved for each of the efficient alternatives  $n'$ , obtaining the super-efficiency score  $\hat{\theta}_{n'}$ .

$$\begin{aligned}
 & \sum_{n \neq n'} \lambda_{n'n} x_{ni} + S_i^- = \hat{\theta}_{n'} x_{n'i}, \forall i \\
 & \sum_{n \neq n'} \lambda_{n'n} y_{nj} - S_j^+ = y_{n'j}, \forall j
 \end{aligned} \quad (3)$$

In essence, DEA is applied to identify the best technologies, which are further ranked according to their super-efficiency values, while providing valuable insight via improvement targets on how to make sub-optimal technologies efficient.

### 3. Case study: an application to chemical energy vectors route selection

The DEA method has been applied to 218 energy storage alternatives, encompassing 189 MeOH routes, 27 NH<sub>3</sub> routes, and 2 conventional technologies (compressed air energy storage, CAES, and pumped hydro energy storage, PHES). These alternatives have been analyzed considering their economic performance (i.e., relative cost of storage, RCS) and environmental performance. The latter has been evaluated through 18 LCA mid-point indicators in the ReCiPe2016 methodology (Huijbregts et al., 2017).

In total, these 19 economic and environmental indicators represent the input decision criteria (i.e., indicators to be minimized) for the DEA methodology. The electricity output of the system, 1 kWh for all alternatives, is the only output indicator considered. In order to ensure the numerical stability and robustness of the DEA results, all indicators have been linearly normalized between 0 and 1. The DEA model was implemented in GAMS 39.3.0 and solved with CPLEX 22.1.0.0 on an Intel® Core™ i7-4790 processor operating at 3.6 GHz. On average, each DEA model instance was solved with a zero optimality gap, taking between 0.03 and 0.14 CPU seconds. Models M1 and M2 consist of 240 variables and 22 constraints.

In addition to the considered economic and environmental indicators, the technical performance of each route was estimated through its total energy efficiency  $\eta_E = \frac{1}{E_{In} + T_{In}}$ , which considers the amount of electricity ( $E_{In}$ ) and thermal energy ( $T_{In}$ ) required through the entire route to obtain 1 kWh after releasing the stored energy. This indicator is not considered in the DEA, instead it is compared with the super-efficiency scores to determine whether super-efficient alternatives also present high energy efficiencies.

### 4. Results and discussion

Out of a total of 218 processes, only 52 (24% of them) were found to be efficient. Inspecting the efficient cases, no correlation was found between energy efficiency and super-efficiency scores (Figure 1). Thus, energy efficiency is not sufficient to establish the optimality level of an energy storage process from an economic and environmental perspective. The ten most super-efficient scenarios consist of eight NH<sub>3</sub>-based scenarios, CAES, and PHES, with the NH<sub>3</sub> | H<sub>2</sub>-B+HC-NAS scenario performing best, followed by CAES. Although not shown in Figure 1, efficient MeOH routes represent the majority of the remaining 42 alternatives, albeit their super-efficiency scores are strikingly smaller than the NH<sub>3</sub> and conventional storage alternatives.

Moving to the inefficient scenarios, the distribution of the relative improvement targets along the different indicators (Figure 2A) highlights that the worst-performing indicators, especially environmental metrics such as FWE, TA, and FPMF, require sensible improvements (39% on average). Moreover, the average improvement targets of the 12 metrics with the worst results lie within a narrow range (28-39%), and the 13 worst-performing indicators, on average, require substantial improvements (75-86% improvement margin). Conversely, the three best-performing indicators (GW, WC, LU) require an improvement below 10%, on average, and lower than 25%, even when accounting for extreme cases. Notably, the RCS is the fourth metric requiring the lowest improvement on average (12%) and in the worst case (37%). Such a result highlights how most technologies present relatively similar operational costs and greenhouse gas emissions in their life cycle, as well as in other generally considered environmental indicators (e.g., water and land use). This finding reflects how particular emphasis should be placed on often disregarded environmental metrics (e.g., FWE, TA) rather than solely GHG emissions or economic indicators when aiming to design fully sustainable energy storage strategies.

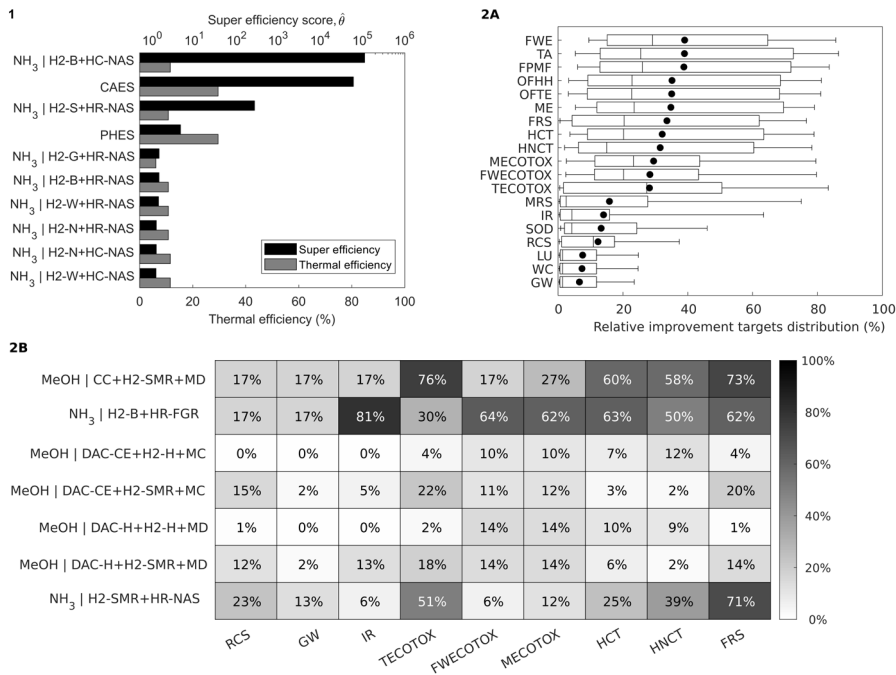


Figure 1. Super-efficiency (upper x-axis) and thermal efficiency (lower x-axis) scores for 10 alternatives with the highest super-efficiency scores.

Figure 2A. Relative improvement target distributions for all 19 input categories and their means (dots). Boxplots represent the 5<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup>, and 95<sup>th</sup> percentiles.

Figure 2B. Relative improvement targets for sub-efficient alternatives with the best performance (not considering the efficient alternatives) in one or more input indicators.

Finally, Figure 2B presents the disaggregated relative improvement targets for the sub-efficient scenarios with the best performance in one or more of the considered 19 indicators. In contrast with previous results, MeOH-based technologies represent the majority of this set of sub-efficient technologies, with some of these alternatives presenting very high improvement targets in toxicity (e.g., TECOTOX) and resource scarcity (e.g., FRS) related indicators. This highlights how seemingly promising scenarios may still perform suboptimally, with only a few alternatives performing efficiently in these specific metrics. Among other scenarios, NH<sub>3</sub> | H2-B+HR-FGR would require significant improvements in toxicity-related categories due to the use of fertilizers in biomass cultivation. On the other hand, scenario MeOH | DAC-CE+H2-H+MC presents relatively low improvement targets (in some cases performing as other efficient alternatives), indicating how grid-powered DAC CO<sub>2</sub> capture and sequestration may be seen as a promising energy storage strategy based on methanol.

## 5. Conclusions

With the increasing number of developed alternative pathways involving energy storage, there is a spiking need for decision-support tools to systematically assess the current state of the art and optimally drive research efforts in this field. In this work, we apply a systematic methodology based on DEA to analyze a large representative set of P2C2P routes involving NH<sub>3</sub> and MeOH as energy storage vectors, considering both economic



and environmental criteria. The approach here considered has been able to, given a set of economic and environmental metrics and without the use of user-defined weights, identify and rank the overall best-performing energy storage alternatives and provide necessary improvement targets for inefficient alternatives. Our results highlight how energy efficiency is not correlated with the overall economic and environmental performance of the different routes. Also, NH<sub>3</sub>-based routes seem to be the most promising alternatives, along with CAES and PHES. Overall, this work aims to shed light on the potential of chemical energy storage by evaluating and comparing multiple alternative routes, highlighting the benefits that decision-support tools such as DEA can bring to the study and development of more sustainable energy systems.

### Acknowledgments

This publication was created as part of NCCR Catalysis (grant 180544), a National Centre of Competence in Research funded by the Swiss National Science Foundation. SCD and GGG are authors associated with NCCR Catalysis.

### References

- A. Charnes, W.W. Cooper, E. Rhodes, 1979, Measuring the efficiency of decision-making units, *Eur. J. Oper. Res.*, 3, 339.
- W.W. Cooper, L.M. Seiford, K. Tone, 2007, *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, 2nd ed, Springer Science & Business Media, New York, NY.
- S.C. D'Angelo, S. Cobo, V. Tulus, A. Nabera, A.J. Martín, J. Pérez-Ramírez, G. Guillén-Gosálbez, 2021, Planetary Boundaries Analysis of Low-Carbon Ammonia Production Routes, *ACS Sustain. Chem. Eng.*, 9, 9740–9749.
- J. Davies, F. Dolci, D. Klassek-Bajorek, R. Ortiz Cebolla, E. Weidner Ronnefeld, 2020, Current Status of Chemical Energy Storage Technologies, Joint Research Centre (European Commission).
- S. Deng, R. Hynes, 2012, Advanced combined cycle systems based on methanol indirect combustion, *J. Eng. Gas Turbines Power*, 134.
- R. Dones, C. Bauer, R. Bolliger, B. Burger, T. Heck, A. Röder, M. Faist, R. Frischknecht, N. Jungbluth, M. Tuchschnid, 2007, Life Cycle Inventories of Energy Systems: Results for Current Systems in Switzerland and other UCTE Countries, Dübendorf.
- A. González-Garay, M.S. Frei, A. Al-Qahtani, C. Mondelli, G. Guillén-Gosálbez, J. Pérez-Ramírez, 2019, Plant-to-planet analysis of CO<sub>2</sub>-based methanol processes, *Energy Environ. Sci.*, 12, 3425–3436.
- M.A.J. Huijbregts, Z.J.N. Steinmann, P.M.F. Elshout, G. Stam, F. Verones, M. Vieira, M. Zijp, A. Hollander, R. van Zelm, 2017, ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level, *Int. J. Life Cycle Assess.*, 22, 138–147.
- International Organization for Standardization, 2014, ISO 14040:2006 - Environmental management - Life Cycle Assessment: Principles and framework, ISO.
- J.L. Liu, J.H. Wang, 2016, A comparative research of two adiabatic compressed air energy storage systems, *Energy Convers. Manag.*, 108, 566–578.
- F. Rostami, Z. Kis, R. Koppelaar, L. Jiménez, C. Pozo, 2022, Comparative sustainability study of energy storage technologies using data envelopment analysis, *Energy Storage Mater.*, 48, 412–438.
- A. Sánchez, E. Castellano, M. Martín, P. Vega, 2021, Evaluating ammonia as green fuel for power generation: A thermo-chemical perspective, *Appl. Energy*, 293, 116956.
- G.-H. Tzeng, J.-J. Huang, 2011, *Multiple Attribute Decision Making: Methods and Applications*, 1st ed, Chapman and Hall & CRC, New York, NY.
- I. Vázquez-Rowe, D. Iribarren, M.T. Moreira, G. Feijoo, 2010, Combined application of life cycle assessment and data envelopment analysis as a methodological approach for the assessment of fisheries, *Int. J. Life Cycle Assess.*, 15, 272–283.

# Simulation of thermochemical recycling of waste plastic to produce dimethyl ether

Mohammaed Awwad <sup>a</sup> and Umer Zahid <sup>a, b</sup>

<sup>a</sup> *Chemical Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia.*

<sup>b</sup> *Interdisciplinary Research Center for Membranes & Water Security, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia.*

## Abstract

In this study, the thermochemical plastic recycling process namely gasification has been investigated to process 10,000 kg/h of plastic waste. This study rigorously modeled and simulated the thermochemical processes that can recycle the plastic wastes efficiently to useful chemical such as di-methyl ether (DME). Currently, the cost of recycled plastic is not competitive compared to the fossil-based virgin plastics, which leads to the lack of industrial willingness for recycling. This study explored the technical feasibility to understand the processing of recycled plastics to useful chemical conversion. The results show that the gasification can provide a competitive solution with flexibility in terms of waste feed composition to produce DME at a large scale.

**Keywords:** Waste recycling; techno-economics; clean fuel; process design.

## 1. Introduction

Plastic production has readily invaded the global market by the middle of the previous century due to its wide range of applications in construction, textiles packaging, transportation, healthcare, and electronics sectors. Plastics have replaced the traditional materials such as wood, steel, and ceramics due to its versatile properties, i.e., low density, durability, low price, and corrosion resistance. Polyethylene (PE) and polypropylene (PP) represent nearly the half of production of plastics, on the other hand, polyvinyl chloride (PVC), polystyrene (PS), polyurethane (PUR), and polyethylene terephthalate (PET) are also produced significantly as shown in figure 1 (Lopez et al., 2018). Currently, the plastic production rate is increasing approximately at 4% per year, reaching an annual production of more than 353 million metric tons in 2019 (OECD, 2022). Owing to this mega-scale utilization of plastics in everyday domestic and industrial products poses a dramatic problem of plastic waste management that needs firm decisions be taken. According to Organization for Economic Co-operation and Development (OECD), only 9% of global plastic waste is recycled, while 49% of waste plastic was discarded, mainly in landfills, 19 % wastes were incinerated, and 22% of the waste remained uncollected and mismanaged (OECD, 2022). Due to plastics low degradability, landfilling of plastics represents a significant problem to the planet's ecosystem. Besides, plastic incineration has several environmental issues, such as formation of dioxins, fly ash, sulfur, nitrous oxides, and other hazardous chemicals. At the same time, the world is facing a real challenge from climate change, where plastics contribute around 3.4 % of greenhouse gas emissions globally (OECD, 2022).

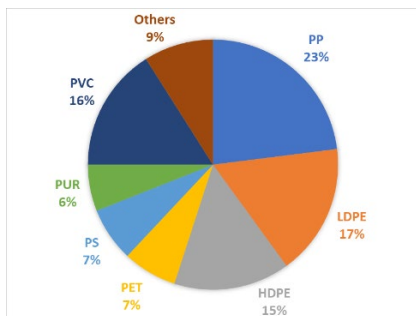


Figure 1: Global plastics production by type.

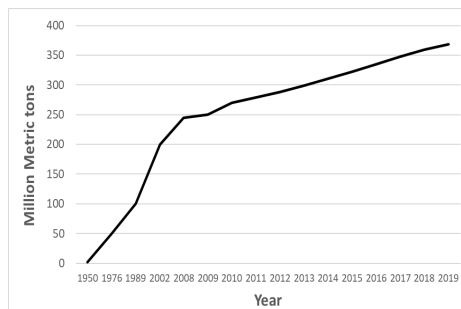
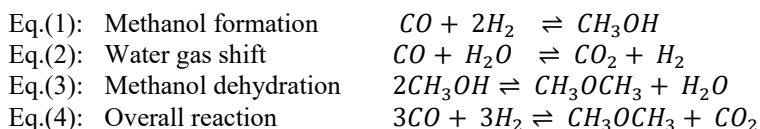


Figure 2: Global plastic production.

Some developed countries have started a waste plastic management plan; however, the required efforts are not enough, and mandate local, regional, and global actions be taken. Currently, different waste plastic recycling methods are already under investigation. Mechanical recycling is an efficient method that reconstruct the plastic into simpler shape while preserving the chemical structure. Although mechanical recycling of plastics is the most common plastic waste management method, it has some limitations. The feedstock needs first to be purified before proceeding further to achieve a high product quality; also, recycled product's market feasibility is still not proved (Ragaert et al., 2017). Recently, several countries are evaluating alternative fossil fuel options, and waste plastic is one such opportunity that is seen as lost resource that could be useful economically to produce valuable chemicals. Thermochemical plastic recycling is a technology where the heat is used to break a larger molecule to a smaller one that already made the plastic. Hence, chemical recycling is a potential route to enable a sustainable life cycle of plastic, and it is more flexible with feedstock than mechanical recycling [1,5]. Both pyrolysis and gasification processes fall under the category of thermochemical recycling. Gasification is a process where the waste plastic is heated at a high temperature with a limited amount of  $O_2$ . On the other hand, pyrolysis is a process where the plastic is being heated at a relatively lower temperature in the absence of  $O_2$ . The main advantage of gasification over pyrolysis is less sensitivity to different feed compositions or mixtures of plastics waste. Gasification of plastic mainly produces a gas stream that is rich in  $H_2$ ,  $CO_2$ ,  $CO$ , and  $CH_4$ . This syngas stream can then be employed either to produce energy or can serve as a raw material to produce a wide-range of important chemicals such as methanol, ammonia, di-methyl ether (DME) and others. DME has gained significant attention due to its clean combustion properties and easy handling as a fuel (Sansaniwal et al., 2017). DME can be synthesized from syngas in two ways, indirect and direct method. In the indirect method, methanol is first produced from syngas followed by the dehydration process. The direct method employs a bifunctional catalyst where methanol synthesis reaction, water gas shift reaction, and dehydration reaction co-occur as shown below in Eq. (1) – (4):



Direct synthesis of DME is a potential route as it can overcome the strong thermodynamic limitations of syngas conversion to methanol by shifting the equilibrium. This study is

focused on the DME production from the gasification of waste plastic. Some previous works studied the plastic gasification. Shehzad et al., (2016) investigated the municipal solid waste (MSW) performance by varying gasification operational conditions. They reported that a higher gasification temperature has a significant impact on syngas composition. The results showed that gasification process has lower efficiency when utilizing MSW for electricity generation application, compared to the incineration; however, since chemicals could be produced from syngas, further economic analysis to make a suitable judgment was recommended. Weiland et al., (2021) studied the gasification of complex plastics and found that high temperature gasification processes that are usually assisted with high purity oxygen can generate syngas with heating value in the range of 10 – 13 MJ/Nm<sup>3</sup>. Similarly, many studies explored the DME synthesis through a one-step route. For example, Chen et al., (2012) studied the DME direct route synthesis from the syngas by developing a bifunctional Cu–ZnO–Al<sub>2</sub>O<sub>3</sub>/ZSM5 catalyst. They found that the DME yield decreases as the CO<sub>2</sub> concentration in the syngas increases. Guo et al., (2022) studied the DME one-step synthesis from the syngas produced by the biomass. This study was mainly focused on the catalyst development and reported that a hybrid catalyst comprised of CuO–ZnO–Al<sub>2</sub>O<sub>3</sub> and HZSM-5 has a synergetic effect on the DME synthesis. Most of the previous works shed light on the waste plastic gasification experimentally and one-step direct DME synthesis independently. Besides, there are no remarkable studies that evaluated the DME production from waste plastic gasification. Therefore, the aim of this study is to design and simulate a commercial scale DME production unit considering waste plastic as a feedstock.

## 2. Process and Simulation details

The process contains three main sections: waste plastic gasification, acid gas removal (AGR) unit, and DME production and purification, as shown in figure 3.

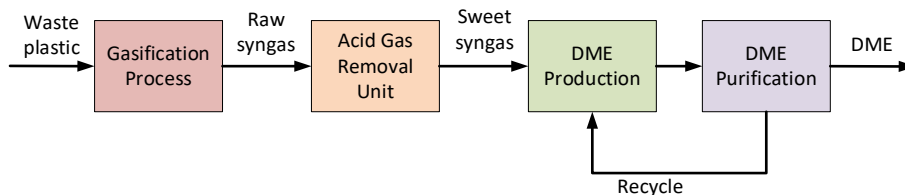


Figure 3: Simplified block diagram of plastic waste to DME production process.

### 2.1. Gasification Process

The waste plastic is first shredded before feeding it to the gasification process. The gasification reactor operates at 887°C and 1 bar in the presence of high purity oxygen. The gasifier effluent produces syngas rich in H<sub>2</sub>, CO<sub>2</sub>, CO, and CH<sub>4</sub> and slag which comprise of ash and unconverted carbon. One operational trouble with plastic gasification is the sticky nature of plastic material, however, this problem can be minimized by operating gasification at higher temperatures. A multi-stage cooler system is used to cool the reactor effluent before separating the syngas from the slag. Waste plastic is simulated as a non-conventional component in Aspen Plus by defining its ultimate and proximate analysis. Peng-Robinson thermodynamics property was utilized to model the gasification process. The base assumes the plastic waste feed rate of 10000 kg/h, while the oxidant is assumed to be 98 mol.% pure oxygen. The gasifier has been modeled using a RGIBBS subroutine that works on the principle of minimization of Gibbs free energy.

### 2.2. Acid gas removal (AGR) unit

A cleaning unit is an important process needed to reduce the CO<sub>2</sub> concentration before directly utilizing the raw syngas. In general, absorption with amines is the most convenient method to decrease the CO<sub>2</sub> level. A typical AGR unit consists of an absorber where the acid gases are absorbed by the amine, and a regenerator column where the rich amine is regenerated by stripping of the absorbed acid gases from the amine solvent. In this study, mono-ethanol-amine (MEA) has been utilized for the removal of acid gases. The AGR process has been modeled using AMINES thermodynamic package. The absorber is modeled using a RADFRAC block that removes 98% of the CO<sub>2</sub> in the syngas stream coming from the gasifier. The rich amine solvent is then sent to the regenerator (also modeled using RADFRAC) where the MEA solvent is regenerated, cooled, and recirculated to the absorber column. The calculator block in the Aspen plus is utilized to calculate the make-up solvent.

### 2.3. DME synthesis and purification unit

Sweet syngas from the AGR unit is first compressed in a multi-stage compressor stage with intercooling to reach the favorable reaction conditions (250°C and 60 bar). The compressed syngas is then fed to the DME synthesis reactor where a bifunctional catalyst yield DME via a one-step process explained earlier. The reactor effluent mainly contains unreacted syngas, DME, and methanol. The reactor effluent is sent to a series of separation units for the separation of product from the unreacted syngas and by-products. First, a flash drum is employed to separate most of the gas phase which mainly consists of unreacted syngas. The liquid phase leaving at the bottom of the flash drum is then sent to a series of distillation column where CO<sub>2</sub> is removed in the first column, product DME is recovered in the second column, and by-product methanol is separated in the third column which is recycled back to the DME synthesis reactor. DME synthesis and purification section has been modeled using Soave-Redlich-Kwong (SRK) equation of state. DME synthesis reactor is modeled using RPLUG by employing the Langmuir–Hinshelwood–Hougen–Watson (LHHW) kinetics. Figure 4 shows the process flow diagram of the overall plastic waste to DME process.

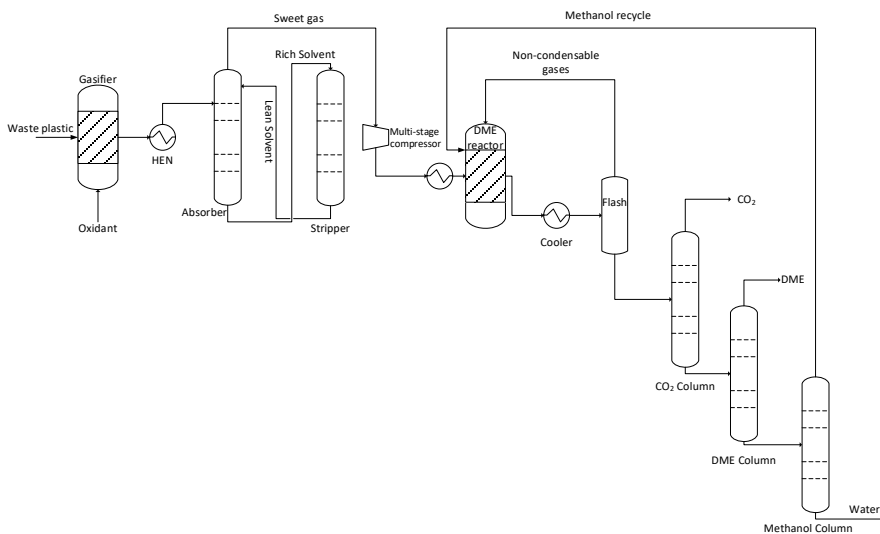


Figure 4: Process flow diagram of waste plastic to DME production process.

### 3. Results and discussions

#### 3.1. Model validation

Model Validation is a crucial step in the design process that should not be overlooked before analyzing or extending the model to different operational conditions. Therefore, model validation is performed to ascertain how rigorous the designed model is. Figure 5 shows the validation of the DME synthesis reactor compared to the experimental data reported in the literature. The results show an amicable agreement between the simulation results and experimental data. In addition, composition of DME synthesis reactor effluent stream has been compared using RGIBBS and REQUIL reactors, and the results show a fair agreement with RPLUG reactor results as shown in Figure 6.

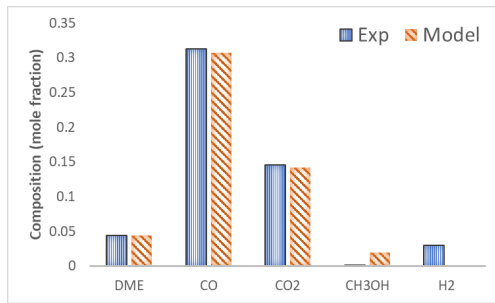


Figure 5: Validation of DME synthesis reactor.

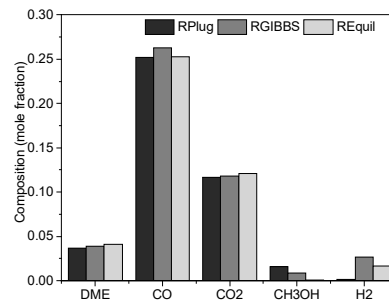


Figure 6: Comparison of different reactor models for DME synthesis.

#### 3.2. Energy analysis

A detailed energy analysis has been performed to explore the energy consumption in various sections of the process. The results for the main sections of the plastic waste to DME process are presented in Figure 7 and 8. The results show that the DME synthesis and purification section consumes the most energy in the process amounting to around 52% of the complete process. This is obvious because of intensive separation system in the DME purification section. AGR unit and gasification section contribute around 34.4% and 13.6% of the energy consumption respectively.

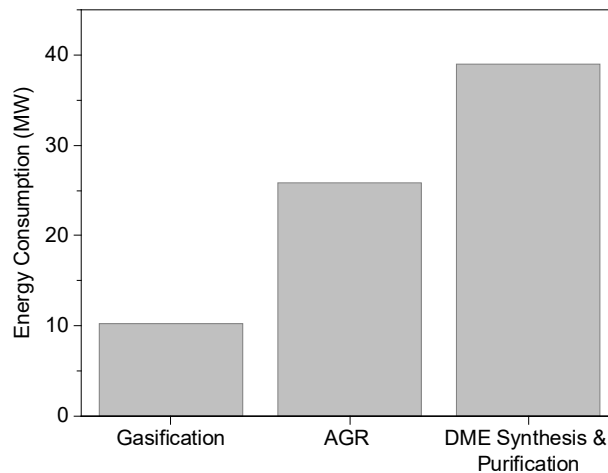


Figure 7: Energy consumption in various sections of the plastic waste to DME process.

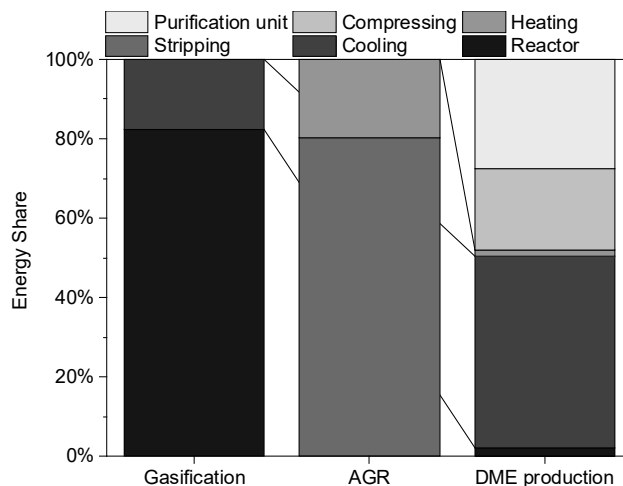


Figure 8: Energy share in various section of the plastic waste to DME process.

#### 4. Conclusion

Plastic wastes are produced in significant amounts every year and can be utilized to produce useful clean chemicals such as DME. In this work, a conceptual design has been proposed for the production of high purity DME from the waste plastic gasification. This study showed the technical feasibility of DME production from waste plastic. The results showed that the proposed DME production process requires around 2.88 MJ per kg of DME production.

#### Acknowledgement

The author(s) would like to acknowledge the support provided by the Deanship of Scientific Research at King Fahd University of Petroleum & Minerals (KFUPM).

#### References

- Chen, W.-H., Lin, B.-J., Lee, H.-M., Huang, M.-H., 2012. One-step synthesis of dimethyl ether from the gas mixture containing CO<sub>2</sub> with high space velocity. *Appl. Energy* 98, 92–101.
- Guo, X., Liu, F., Hua, Y., Xue, H., Yu, J., Mao, D., Rempel, G.L., Ng, F.T.T., 2022. One-step synthesis of dimethyl ether from biomass-derived syngas on CuO-ZnO-Al<sub>2</sub>O<sub>3</sub>/HZSM-5 hybrid catalyst: Combination method, synergistic effect, water-gas shift reaction and catalytic performance. *Catal. Today*.
- Lopez, G., Artetxe, M., Amutio, M., Alvarez, J., Bilbao, J., Olazar, M., 2018. Recent advances in the gasification of waste plastics. A critical overview. *Renew. Sustain. Energy Rev.* 82, 576–596.
- OECD, 2022. Global Plastics Outlook.
- Ragaert, K., Delva, L., Van Geem, K., 2017. Mechanical and chemical recycling of solid plastic waste. *Waste Manag.* 69, 24–58.
- Sansaniwal, S.K., Pal, K., Rosen, M.A., Tyagi, S.K., 2017. Recent advances in the development of biomass gasification technology: A comprehensive review. *Renew. Sustain. energy Rev.* 72, 363–384.
- Shehzad, A., Bashir, M.J.K., Sethupathi, S., 2016. System analysis for synthesis gas (syngas) production in Pakistan from municipal solid waste gasification using a circulating fluidized bed gasifier. *Renew. Sustain. Energy Rev.* 60, 1302–1311.
- Weiland, F., Lundin, L., Celebi, M., van der Vlist, K., Moradian, F., 2021. Aspects of chemical recycling of complex plastic waste via the gasification route. *Waste Manag.* 126, 65–77.

# Sustainable recycling of refrigerants: Analysis of alternatives

J. Moreno<sup>a</sup>, I. Karpov<sup>a</sup>, A. Ahmed<sup>a</sup>, J. Foglia<sup>a</sup>, H. Wu<sup>a</sup>, A. Gadekar<sup>a</sup>, P.L.J. Swinkels<sup>a</sup>, R. Gani<sup>b,c</sup>

<sup>a</sup>*Chemical Engineering, Delft University of Technology, Van der Maasweg 9, 2629HZ, Delft, The Netherlands*

<sup>b</sup>*PSE for SPEED Company, Ordtup Jagtvej 42D, 2920 Charlottenlund, Denmark*

<sup>c</sup>*Sustainable Energy and Environment Thrust, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China*

Corresponding author: [P.L.J.Swinkels@tudelft.nl](mailto:P.L.J.Swinkels@tudelft.nl)

## Abstract

The use of refrigerants has been continuously increasing in thermal control systems employed in residential premises, offices, storage, process operations and many more. Widely used refrigerants such as R-410A have become a target for regulations to prohibit their use because of environmental issues. R-410A is a blend of R-32 (Difluoromethane) and R-125 (Pentafluoroethane). Because of high global warming potential (GWP) of R-125, it has become a target for removal from R-410A and reuse. An additional issue is the recovery of the refrigerant from current and decommissioned thermal control units. This paper reports the results of an investigation related to the separation of refrigeration blend compounds, recycle and reuse of the refrigerant mixture compounds and alternative blend compositions, with emphasis on environmental, health and physical hazards as well as optimal refrigeration cycle operation. Two alternatives for feasibility of operating a continuous separation-recycling process are studied.

**Keywords:** Refrigerant blends, R-32, R-125, R-410A, recycling, global warming potential, safety-health-environmental hazards.

## 1. Introduction

Hydrofluorocarbons (HFCs) are known for their excellent performance as refrigeration fluids. Their durability, inertness and thermal stability have made them the selection of choice in air-conditioning (AC) units. Nevertheless, their stability makes their presence persistent when released to the atmosphere, with a lifetime between 15 and 29 years (Xu, 2013). Furthermore, their capacity to absorb infrared light make them potent greenhouse gases (Castro et al. 2021).

R-410A is an HFC-based refrigerant, initially invented and patented by Allied Signal (now Honeywell) in 1991, that became popular as a substitute for R-22 (Chlorodifluoromethane) after the Montreal Protocol (UNEP, 2021) took place. The commercial R-410A consists of 50 wt% of R-32 and R-125 whose individual global warming potential (GWP) values are 677 and 3170, respectively, and giving the blend GWP value of 1923. Note that the mixture of R-32 and R-125 forms an azeotrope at 91 mol% R-32 (Castro et al. 2021). The objective of this paper is checking the feasibility of



removal of R-125 from R-410A as well as recycling and reuse of the separated compounds. Monjur et al. (2022) have reported the separation of R-32 and R-125 by a continuous extractive distillation process with ionic liquids as the solvent. The feasibility of this process is compared with a pressure-swing distillation process, which also generates various R-410A blend alternatives. The performance of the generated alternatives is compared in terms of coefficient of performance (COP) of the refrigerant within the refrigeration cycle. Also, the effect of hazardous properties (health, physical and environmental) is considered from the points of view of safe and optimal refrigeration cycle operation.

Because of limited space, calculation details, hazardous properties of chemicals, etc., are not given in this manuscript. They can be obtained from the corresponding author.

## 2. Data-based Analysis of Opportunities for Recycling of F-gases

According to EU F-Gas legislation, all HFC refrigerants from refrigeration systems and air-conditioning systems must be recovered (European Parliament and Council of European Union, 2014) and replaced. If recovered material is too contaminated for further recycling, it should be decomposed. According to the European Environmental Agency, by 2020, the refrigerant HFC-23 account for nearly 90% of F-gases destroyed in the EU. Considering limitations of resources as well as undesirable impacts, the recovery and reuse of refrigerants is an option worth considering (Castro et al. 2021), provided they satisfy the safety, health and environmental concerns as well as operability of the refrigerant cycles. For R-410A and its components (R-32, R-125) requirements mentioned below are applied:

- Vapor phase contaminants: Air and other non-condensables must be < 1.5% by volume at 25.0 °C
- Liquid phase contaminants: Water must be <10 ppm by weight; all other volatile impurities must be < 0.5 wt%; high boiling residue must be < 0.01 wt%

Note that the above requirements are not an issue as R-410A is considered to be a blend of only R-32 and R-125.

### 2.1. R-410A recycling capacity limitations

The main use for R-410A as a refrigerant is in air-conditioning and heat pumps. In 2013 around 86% of all R-410A produced was used for HVAC systems only. Therefore, this study has selected R-410A as a representative example for the study of recycling of refrigeration fluids.

The global production of R-410A increased from 193 kMT in 2012 to 238 kMT in 2016, and the estimated production in 2022 will be 292 kMT. China is the leading producer of R-410A, with about 45% of the global production in 2016, followed by North America, with a production share of around 30% (360 Research Reports, 2020). The share of R-410A produced in EU in the years 2009 and 2013 was around 12% (Makhnatch & Khodabandeh, 2019).

Analysis of the technical data on HVAC systems indicate a life cycle of an AC unit is in the range of 10-15 years. Therefore, in China, the USA and the EU, the total maximum capacity of R-410A available for recycling in the years 2022-2027 is expected to be around 12, 8, and 3 MT/h, respectively (assuming 8000 h/yr). To understand the amount of bulk material (AC units) to be recycled, the assumption made is that R-410A can be extracted only from small-scale AC units (less than 150 kg). Based on the available information on commercial AC units, the average unit weights of 70 kg and as a mass fraction of R-410A of 4 wt%. Therefore, to obtain 1 ton of refrigerant, recovery from 26

tons or 370 units of decommissioned cooling devices should be done. Considering an average recovery time from one unit of around 20 min, and costing USD 16 per hour as a median manufacturing operator salary in USA, together with the average mass of refrigerant mentioned above, recovery costs for R-410A in USA would be 2200 \$/MT.

## 2.2. R-410A separation

In the case of R-410A, separation of the pure compounds is complex due to an azeotrope for a R-32 purity of 91% mol. Monjur et al. (2022) proposed separation via extractive distillation with an ionic liquid in a continuous process.

An alternative to solvent-based extractive distillation is pressure-swing distillation if the azeotrope composition changes when the pressure is changed. Figure 1 shows the process flowsheet for separation of R-32 from R-125, along with the changes in the Txy diagrams. In the first column, operation at 17 bara yields a distillate with a near azeotropic composition of 90 mol% R-32 (80 wt%) at a temperature of 25 °C, allowing the use of cooling water in the condenser. Then, pressure is lowered to 1 bara, shifting the system to the right side of the azeotrope and resulting in a bottom product which is 99.5 mol% R-32 (99 wt%). Details regarding the process and the comparative study can be obtained from the authors. Operational expenses derived from consumption of utilities are 18 times higher for distillation at pressure. Since distillation at pressure is more energy intensive than extractive distillation, this result was expected. Nonetheless, when considering the cost of replenishing the ionic liquid BMIN PF<sub>6</sub>, processing costs for extractive distillation increase from 0.004 \$/hr to 2.4 \$/hr, while remaining at 0.07 \$/kg for the proposed process. An additional advantage of not using an ionic liquid is that, as seen in section 3.2, the coefficient of performance of the recycled R-32 could be affected considering the low vapour pressures which characterize ionic liquids, even at small concentrations.

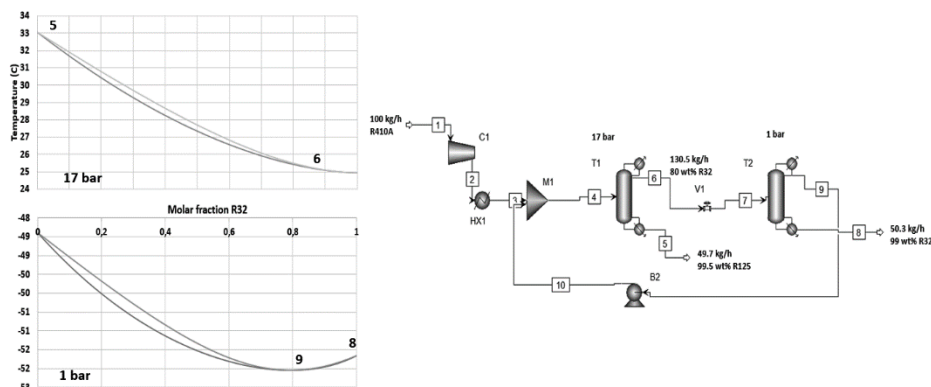


Figure 1: Flowsheet of pressure-swing distillation for separation of R-410A compounds, with Txy diagrams

Note, that the discussion in Section 2.1 indicates that reclaiming enough refrigerant to sustain a continuous process in a processing plant is not feasible, unless a periodic operational policy of collection followed by short periods of operation is applied. The design of the continuous process is sufficiently flexible to allow changes in production rates.

### 3. Evaluation of Alternatives: Recycle of F-gases

The high-pressure azeotropic distillation method provides opportunities for separation of R-32 and R-125 with individual purities of up to 95 and 99.5 wt%, respectively. In addition to the pure component product streams, the pressure swing distillation process allows intermediate product fractioning resulting in R-32 / R-125 mixtures of different compositions, which are investigated below

#### 3.1. R-32 / R-125 mixture as a blend

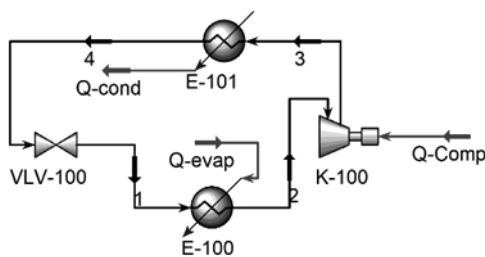
Recently, several options were proposed as a substitute for phasing-out of R-410A. Among the options are:

- Non-flammable (A1 rating) blends: R-448A, R-449A,B,C, R-458A, and R-407C,F with high GWP rating (>1000)
- A2L safety class blends such as R-447A (GWP = 572), R-447B (GWP = 714), R-452B (GWP = 674), with a fraction of R-32 and R-125 in the range of 89.5 – 95.1 and 4.9 – 10.5 wt% respectively

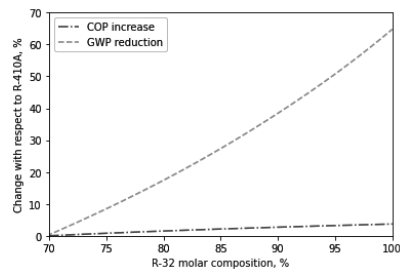
A1-rated refrigerants are available ranging from light commercial systems (0.15 – 5kW) to industrial solutions (up to 10MW). In contrast, A2L refrigerants due to safety concerns have limited use in low-power capacity units (Danfoss A/S, 2022). The full list of patented and approved blends can be obtained from the authors. Tailor-made blends can also be generated (Udomwong et al. 2021).

#### 3.2. Evaluation of performance of R-32 / R-125 blend systems

To investigate the effect of the R-125 composition in the refrigerant mix on the mixture's GWP and coefficient of performance COP a refrigerant cycle (shown in Fig 2a) has been simulated with Aspen HYSYS using the Peng-Robinson equation of state as the thermodynamic model for phase equilibrium. The refrigerant cycle has four operations: compressor, evaporator, condenser, and valve. The ProREFD software (Udomwong et al. 2021) and ChemSub (Syeda et al. 2022) software also allows a detailed cycle simulation, hazardous effects analysis and refrigerant substitution.



(2a)



(2b)

Figure 2: Refrigeration cycle and effect of R-32 composition on COP of refrigeration cycle

At first, the standard composition of R-410A, which is 70 mol% of R-32, was simulated resulting in a GWP of 1924 and a COP of 7.5. It is shown that increasing the composition of R-32 results in a lower GWP but does not decrease the COP (see Fig 2b). A total

replacement of R-125 with R-32 towards a pure compound product, results in a GWP reduction of 65% and a COP increase of 4%. Nevertheless, there is an inherent safety risk associated to R-32 as it is classified with a safety level of A2L, meaning that it does not pose a toxicity risk at concentrations of less than or equal to 400 parts per million, but it is a mildly flammable gas. Its flammability limits are between 14.4% to 29.3%, which poses a risk of explosion if leakage occurred in the AC system (Osuagwu, 2022). Conversely, R-410A is classified as an A1 refrigerant (highest rank in safety classification), meaning that it is non-toxic, non-flammable and relatively safe to store, use and distribute (Osuagwu, 2022).

### 3.3. Evaluation of second life of recycled R-125

R-125 has recently found limited application in other sectors rather than thermal control systems, even though 99% of R-125 produced worldwide by the reporting companies is used as a blend component for commercial refrigerants. It could also be used as a fire suppression agent because it is odorless, colorless, electrically non-conductive, non-corrosive, and leaves no residue. It was also reported to act as a foam-blowing agent (Yelisetty & Visco, 2009) for making polyurethane, as the polyurethane foams are made by reacting a di-isocyanate molecule with a polyol in the presence of a blowing agent. R-125 is to substitute chlorofluorocarbons and hydrochlorofluorocarbons, because of the ozone-depleting potential of the latter ones. It only works as an interim solution because of R125's GWP concerns. According to Karecki (2000), R-125 can also be used as a substitute for 1,1-difluoroethane for a Dielectric Etch Application in an Inductively Coupled Plasma Etch Tool. A comparison of the emissions from R-125/R-152a processes and a reference C<sub>2</sub>F<sub>6</sub>-based process showed that emissions reductions of 68 to 76% were attainable, in large part due to much higher feed gas conversion efficiency (Karecki, 2000).

## 4. Conclusions

It is found that, while the separation by ionic liquid based extractive distillation is theoretically feasible, its implementation is not practical as the amount of R-410A available for recovery would not sustain a continuous process. For a processing capacity of 100 kg/h of R-410A, the number of average-size units to be recovered per day is 900. This is equivalent to 270,000 average-size AC units per year, or 4% of the R-410A units that are estimated to go out of service in China in 2022. Considering the vast area for which the availability of refrigerant is calculated, the process capacity is deemed infeasible, while pressure swing distillation is found to give 34 times lower operating costs. Increasing the pressure shifts the azeotrope to 91 mol% of R-32 at 17 bara (from 80 mol% R-32, at 1 bara), making separation of both components possible by shifting between both sides of the azeotrope in two distillation columns operating at 1 and 17 bara. However, both continuous operations need to be operated in a periodic (or batch) mode with periods of collection and short separation operation. Additionally, the presence of ionic liquid in recycled R-32, even at small concentrations, presents a health hazard for humans. Regarding the use of R-32 as a stand-alone refrigerant, coefficient of Performance (COP) value is found to increase by up to 4% with respect to R-410A, while displaying a 65% reduction in GWP. However, the flammability and health hazard concerns for R-32 also need to be considered and a stricter monitoring of the safety issues would need to be implemented. As the ionic liquid does not vaporize, the batch operation mode for ionic liquid based extractive distillation may need new batch operations designs, unlike the organic solvent-based batch extraction operations.

## References

- P. J. Castro, J. M. Araújo, G. Martinho, A. B. Pereira, (2021). Waste Management Strategies to Mitigate the Effects of Fluorinated Greenhouse Gases on Climate Change. *Applied Sciences*, 4367.
- Danfoss A/S. (2022). *Refrigerant options now and in the future*. Danfoss A/S. Retrieved from <https://assets.danfoss.com/documents/211728/AD224586434178en-001001.pdf>
- Difference Between R410A and R32*. (2022). Retrieved from Acontech PLT: <https://acontech.com.my/difference-between-r410a-and-r32/>
- European Environment Agency. (2020). *Fluorinated greenhouse gases 2020*. Luxembourg: European Environment Agency.
- European Parliament and Council of European Union. (2014). Regulation (Eu) No 517/2014 of the European Parliament and of the Council of 16 April 2014 on fluorinated greenhouse gases and repealing Regulation (Ec) No 842/2006—European Environment Agency [Policy Document]. *Official Journal of the European Union, L 150*, 195–230.
- 360 Research Reports (2020). *Global R410A Market Research Report 2020*.
- S. M. Karecki, (2000). *Development of novel alternative chemistry processes for dielectric etch applications*. PhD Thesis, Massachusetts Institute of Technology publications.
- P. Makhnatch, R. Khodabandeh, (2019). Future Refrigerant Mix Estimates as a Result of the European Union Regulation on Fluorinated Gases. *International Institute of Refrigeration*, 2318–2325.
- M. S. Monjur, A. Iftakher, M. M. F. Hasan, (2022). Separation Process Synthesis for High-GWP Refrigerant Mixtures: Extractive Distillation using Ionic Liquids. *Industrial and Engineering Chemistry Research*, 4390-4406.
- S. Osuagwu, (2022). *R410a vs R32 – A Comparative Study of Features & Benefits*. Retrieved from Rx Mechanic: <https://rxmechanic.com/r410a-vs-r32/>
- V. Panato, D. F. Pico, E. P. Filho, (2022). Experimental evaluation of R32, R452B and R454B as alternative refrigerants for R410A in a refrigeration system. *International Journal of Refrigeration*, 135, 221-230.
- S. R. Syeda, E. A. Khan, N. Kuprasertwong, O. Padungwatanaroj, R. Gani, (2022), A Model-Data Driven Chemical Analysis System for Products and Associated Processes, *Computer Aided Chemical Engineering*, 49, 181-186

- K Udomwong, A Robin, N Kuprasertwong, O Padungwatanaroj, AK Tula, ...R Gani, (2021), ProREFD: Tool for Automated Computer-Aided Refrigerant Design, Analysis, and Verification, *Computer Aided Chemical Engineering* 50, 457-462
- R. Velgara, (2003, August). What You Need to Know When Recovering R-410A. *Yellow Jacket*, 1(3).
- Y. Xu, D. Z. (2013). The role of HFCs in mitigating 21st century climate change. *Atmos. Chem. Phys.*
- S. S. Yelisetty, D. P. Visco, (2009). Solubility of HFC32, HFC125, HFC152a, and HFC143a in Three Polyols. *Journal of Chemical & Engineering Data*, 781-785.
- UNEP (2021). *About Montreal Protocol*. Retrieved from <https://www.unep.org/>



# Thermodynamic study and simulation of the process of separation of the IPA+water mixture by heterogeneous azeotropic distillation

Elvira Spatolisano<sup>a</sup>, Camilla Barbieri<sup>a</sup>, Laura A. Pellegrini<sup>a</sup>, Stefania Moioli<sup>a,\*</sup>

*<sup>a</sup>GASP, Group on Advanced Separation Processes & GAS Processing, Dipartimento di Chimica, Materiali e Ingegneria Chimica "Giulio Natta", Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133 Milano, Italy  
[stefania.moioli@polimi.it](mailto:stefania.moioli@polimi.it)*

## Abstract

The market of 2-Propanol is expected to reach a value of 4.8 billion dollars by 2032, considering that one possible use of this alcohol is as substitute or as additive to fossil fuels, with the aim of employing biofuels to increase the use of renewable energies. To dehydrate alcoholic mixtures containing azeotropes advanced processes are needed, as a heterogeneous azeotropic distillation process. To select the thermodynamic method best representing the binary mixture IPA+water for the simulation in ASPEN Plus<sup>®</sup>, new experimental data of Vapor-Liquid Equilibrium (VLE) have been collected at the Process Thermodynamics laboratory (PT lab) of Politecnico di Milano at conditions for which no information was found in the literature and used with other literature points for the thermodynamic description of the system IPA+water. The selected model has then been employed for the simulation of the novel azeotropic distillation process, that has been optimized.

**Keywords:** biofuels, isopropanol, azeotrope, distillation, PT lab.

## 1. Introduction

The industrial relevance of 2-Propanol (IsoPropyl Alcohol, IPA) is outstanding. It can be not only used as substitute or as additive to fossil fuels, but also as solvent and cleaning agent in a high number of industries as the medical (Guang et al., 2019) and the semi-conductor ones (Arifin and Chien, 2008). Its market, in 2021 of about 3.2 billion \$, is estimated to grow reaching about 1% of the consumption of organic chemicals by 2032 (Fact.MR, 2022).

In several IPA production processes as the indirect hydrogenation or the fermentation of cellulosic materials, the raw product stream contains water and needs to be dehydrated to obtain pure IPA (De Guido et al., 2021). The process of water removal cannot be performed by employing one standard distillation column because the IPA+water mixture forms an azeotrope (at 101.3 kPa of about 87.5% (Moioli et al., 2021)), and the system, as other alcohol+water systems (CRC Handbook of Chemistry and Physics), is strongly non ideal. Advanced processes are therefore needed, as the one proposed by GASP and studied elsewhere (Spatolisano and Pellegrini, 2022).

To select the thermodynamic model for carrying out the simulation of the dehydration process, experimental data of Vapor-Liquid Equilibrium (VLE) are needed. The experimental activity carried out by GASP at the Process Thermodynamics laboratory (PT lab), part of the Process Design & Process Thermodynamics laboratory (PD&PT lab)



of Politecnico di Milano, strives to fill the gaps of experimental data that the literature presents and it is a further development of the research started in 2021 (Moioli et al., 2021). In particular in this work VLE points of the 2-Propanol+water mixture have been collected at the pressure of 800 mbar. The new equilibrium data, together with other points at different pressures from the literature, have been used for the thermodynamic description of the system IPA+water. Then, the selected model has been employed for the simulation of a novel heterogeneous azeotropic distillation process of the mixture of interest. The process has been optimized taking into account several variables as the solvent flowrate, the number of theoretical stages of the column and the reflux flowrate.

## 2. Experimental set up and protocol

### 2.1. Chemicals

Demineralized water (ISO 3696 Q3, ASTM D 1193 TYPE 4) was used and isopropanol (CAS 67- 63 - 0) was supplied by Sigma-Aldrich, with a purity of at least 99.8% wt., molecular weight equal to 60.10 g/mol and specific gravity < 0.7840. Nitrogen (>99.99%, mole), coming from a tank storing N<sub>2</sub> in the liquid phase, was used to create the vacuum condition inside the VLLE unit (Fischer Labodest VLLE 602) up to the desired pressure and as a carrier for the gas chromatograph.

### 2.2. Apparatus

The Fischer® Labodest® VLLE 602 unit has been employed to collect the VLE data at PT lab. The condensed vapor and the liquid samples at the equilibrium conditions have been drawn from this unit and analyzed with the Agilent 7820A Gas Chromatograph equipped with a Thermal Conductivity Detector (TCD).

The details of the experimental protocol are reported in Moioli et al. (2021). In this work the mixture has been prepared under the hood in a container provided with a cap (400 mL) and, after mixing, part (100 mL) of it has been fed to the feed tank of the VLLE unit. After sampling, the unit is emptied and then refilled by another portion of the already prepared mixture. In order to ensure the repeatability of the test and to check the accuracy of the results obtained, each test (complete procedure) was repeated at least three times.

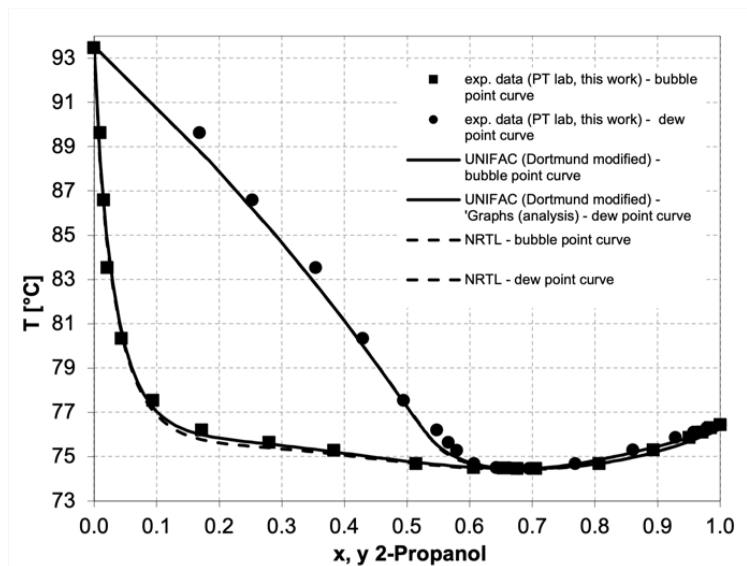
## 3. Results and Discussion

### 3.1. Experimental results and thermodynamic model

Figure 1 reports the experimental points collected at PT lab for a pressure of 800 mbar. The minimum temperature azeotrope occurs at about 68% mol of 2-Propanol. The bubble point curve becomes very close to the dew point curve on the right side of the azeotrope and this underlines the strong non-ideality of this mixture. However this behavior is expected also on the basis of the experimental data already collected at PT lab (Moioli et al., 2021) and from the literature at different pressures.

To describe the thermodynamic behavior of the system, the performance of the NRTL model (Renon et al., 1968) and the one of the UNIFAC-Dortmund modified (Gmehling et al., 1993) model have been tested, by considering 48 sets of isobaric or isothermal VLE data from NIST databank (detailed in Barbieri, 2022) and 8 sets of data collected at the PT lab at different pressures. As from Figure 1 and Table 1, reporting the percent Absolute Average Deviation (%AAD) of the two considered models, the UNIFAC-Dortmund modified (Gmehling et al., 1993) model best represents the system.

The UNIFAC (Dortmund modified) has therefore been selected as the thermodynamic package to be used for the simulation of the novel azeotropic distillation process in ASPEN Plus®.



**Figure 1:** VLE experimental data for the IPA and water system collected at PT lab at 800 mbar and results obtained with the NRTL model with ASPEN Plus® V11.0 default parameters and the UNIFAC Dortmund modified model.

	Isobaric VLE data			Isothermal VLE data		
	AAD% T	AAD% Y <sub>2-Propanol</sub>	AAD% Y <sub>water</sub>	AAD% P	AAD% Y <sub>2-Propanol</sub>	AAD% Y <sub>water</sub>
<b>NRTL</b>	0.17395	7.77114	4.93413	18.722	50.383	11.214
<b>UNIFAC</b>	0.15138	6.63100	4.47493	10.113	19.061	7.256

**Table 1:** %AAD of the temperature for isobaric data or the pressure for isothermal data and the mole fraction of IPA and water in the vapor phase estimated with NRTL and UNIFAC (Dortmund modified) for all the considered sources of data.

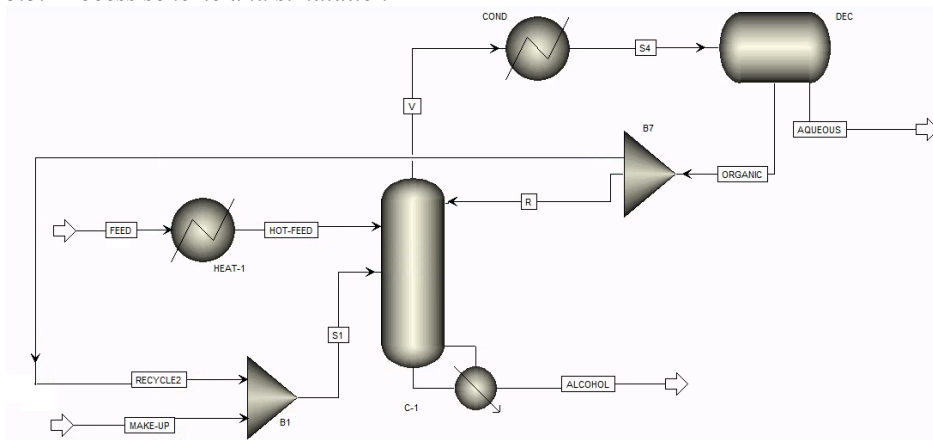
### 3.2. Solvent selection

The literature offers two main alternatives in order to carry out the separation of the IPA+water mixture in its pure components: the extractive distillation (Arifin and Chien, 2008) and the heterogeneous azeotropic distillation (Arifin and Chien, 2007; Cho and Jeon, 2006).

In the extractive distillation the entrainer alters the separation factor of the mixture without introducing azeotropic behaviors with all the components of the system to be separated. On the other hand, in the case of the heterogeneous azeotropic distillation, the solvent is added in order to introduce a new lower boiling azeotrope which for its part can easily be separated. The solvent selection is crucial and, if properly done, can significantly affect the process performances. Among the high number of possibilities, disclosed in literature (Gmehling et al., 1998), the methyl butyl ether has been selected in this work. In fact it obeys to the criterion disclosed in a recent article by GASP (Spatolisano and

Pellegrini, 2022), where it is demonstrated that the best ideal entrainer, thus the one that minimizes the vapor flowrate and, ultimately, the reboiler duty, is the one that forms a binary azeotrope with water, as the methyl butyl ether does. In addition, its boiling point is lower than the boiling point of the IPA, which can be recovered as the bottom product. The other big advantage is that if some traces of methyl butyl ether are present in the IPA product, they are considered acceptable for the use of 2-Propanol as bio-fuel.

### 3.3. Process scheme and simulation



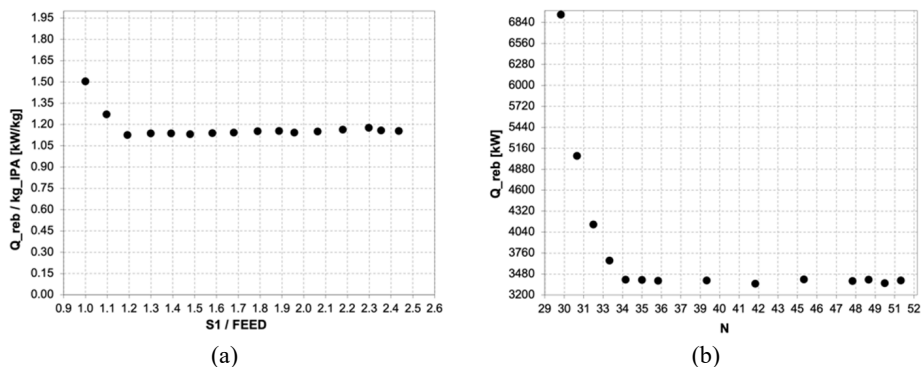
**Figure 2:** Scheme of the heterogeneous azeotropic distillation process with binary azeotrope for the 2-Propanol+water mixture in ASPEN Plus® V11.0.

The proposed flow sheet for the IPA-water separation exploiting methyl butyl ether as solvent is reported in Figure 2. The IPA-water mixture (FEED) and the solvent (S1) are fed into the distillation column at stage 25 and 20 respectively (the stages are counted from top to bottom with the first tray as the first stage and the reboiler as the last stage). The feed, an equimolar mixture of IPA and water, is available at atmospheric pressure and 30°C and has a flowrate of 100 kmol/hr. In the base-case configuration the column has 50 stages and the solvent flow rate is 250 kmol/h. The column separates IPA as bottom product, whose water content has to be lower than 3000 ppm in weight, and the binary water-solvent heterogeneous azeotrope, from the top, which can be separated into two liquid phases. The organic phase, exiting from the decanter, very rich in methyl butyl ether, is refluxed back to the column. A solvent make up stream is provided in order to counterbalance unavoidable solvent losses through the product streams.

The base-case configuration has been optimized in ASPEN Plus® V11.0 taking into account several variables as the solvent flow rate, the number of theoretical stages of the column and the reflux ratio (here defined as the ratio between the reflux flow rate and the organic flow rate exiting the decanter). The objective function, to be minimized, was the reboiler heat duty, as it is the main cost item of the process.

The initial solvent flow rate of 250 kmol/h has been decreased during each simulation (a delta of 10 kmol/h among each simulation has been considered) and within each run of the simulator the IPA bottom purity has been checked. If the bottom product purity was not enough, the reflux ratio would have been increased. This means that the solvent flow

rate and the reflux ratio have been optimized simultaneously. The results of the simulation are reported in Figure 3a).



**Figure 3:** Duty of the reboiler as a function of a) the ratio between the solvent flowrate and the feed flow rate [100 kmol/h] and b) the number of theoretical stages, with a solvent flowrate equal to the optimum value of 120 [kmol/h].

The trend is almost constant (some little fluctuation are due to numerical issues) up to a solvent/feed ratio equal to 1.2. Then the reboiler heat duty tends to increase as the ratio further decreases. This means that the optimum solvent/feed ratio is 1.2, which corresponds to a solvent flowrate of 120 kmol/h. The optimum reflux ratio is equal to 0.56325.

Optimized the solvent flowrate and the reflux ratio, the theoretical number of stages was reduced with respect to the initial value of 50. Indeed the thermal profile and the composition profile of the vapor and liquid mole fractions of the IPA inside the column turned to be almost flat in a high number of stages inside the column, meaning that a lot of stages are not making any contribution to the separation process.

The results of the optimization are reported in the Figure 3b). Again the trend is almost flat up to a number of stages equal to 34 and then it tends to increase as the  $N$  decreases further. Therefore 34 is the optimum number of stages.

#### 4. Conclusions

The isobaric vapor-liquid equilibrium for the system IPA+water at 800 mbar, conditions not already reported in the literature, has been experimentally investigated at the PT lab of Politecnico di Milano. The collected experimental points and the data from the literature have been used to select the thermodynamic model for the simulation and optimization in ASPEN Plus<sup>®</sup> of a novel azeotropic distillation process based on a binary water-solvent azeotrope. The future developments of this work could focus on the collection of more experimental data of phase equilibria in deeper vacuum conditions and on the economic evaluation of the simulated process of azeotropic distillation.

#### References

C. Guang, X. Shi, Z. Zhang, C. Wang, C. Wang, J. Gao, 2019, Comparison of heterogeneous azeotropic and pressure-swung distillations for separating the diisopropylether/isopropanol/water mixtures, ChERD, 143, 249-260.

- S. Arifin, I. Chien, 2008, Design and control of an Isopropyl Alcohol dehydration process via extractive distillation using dimethyl sulfoxide as an entrainer, *Ind. Eng. Chem. Res.*, 47, 790-803. *Fact.MR*, 2022. <https://www.factmr.com/report/836/isopropanol-market>.
- G. De Guido, C. Monticelli, E. Spatolisano, L. A. Pellegrini, 2021 Separation of the mixture 2-Propanol + Water by Heterogeneous Azeotropic Distillation with Isooctane as an Entrainer, *Energies* 14, 17 (2021): 5471.
- S. Moiola, G. De Guido, M. Gilardi, L.A. Pellegrini, D. Bonalumi, G.G. Lozza, 2021, Isobaric Vapor–Liquid Equilibrium Data for the Isopropanol–Water System, *J. Chem. Eng. Data*, 66, 4148-4158.
- CRC Handbook of Chemistry and Physics, 44<sup>th</sup> ed., 2143-2184.
- E. Spatolisano, L.A. Pellegrini, Dehydration of IPA-H<sub>2</sub>O mixture: review of fundamentals and proposal of novel energy-efficient separation schemes, *Chemical Engineering Science*, (Article in press).
- H. Renon, J. M. Prausnitz, 1968, Local compositions in Thermodynamic excess functions for liquid mixtures, *AIChE Journal*, 14, 135-144.
- J. Gmehling, J. Li, M. Schiller, 1993, A modified UNIFAC model. 2. Present parameter matrix and results for different thermodynamic properties, *Ind. Eng. Chem. Res.*, 32, 178-193.
- C. Barbieri, Raccolta e analisi di dati sperimentali di equilibri di fase di sistemi non ideali e studio del processo di distillazione azeotropia, Master Degree Thesis, Politecnico di Milano, 2022.
- S. Arifin, I-L. Chien, 2007, Combined preconcentrator/recovery column design for isopropyl alcohol dehydration process, *Ind. Eng. Chem. Res.*, 46, 2535-2543.
- J. Cho, J. Jeon, 2006, Optimization study on the azeotropic distillation process for isopropyl alcohol dehydration, *Korean J. Chem Eng.*, 23, 1, 1-7.
- J. Gmehling, C. Möllmann, 1998, Synthesis of distillation processes using thermodynamic models and Dortmund Data bank, *Ind. Eng. Chem. Res.*, 37, 3112-3123.

# Sustainability assessment of poly(butylene succinate) production and End-of-Life options from wheat straw

Sofia-Maria Ioannidou<sup>a</sup>, Dimitrios Ladakis<sup>a</sup>, Ricardo Rebolledo-Leiva<sup>b</sup>, Maria Teresa Moreira<sup>b</sup>, Ioannis K. Kookos<sup>c</sup>, Apostolis Koutinas<sup>a,\*</sup>

<sup>a</sup>*Department of Food Science and Human Nutrition, Agricultural University of Athens, Iera Odos 75, 118 55 Athens, Greece*

<sup>b</sup>*Department of Chemical Engineering, School of Engineering, Universidade de Santiago de Compostela, D Rúa Lope Gómez de Marzoa, s/n, 15782, Santiago de Compostela, Spain*

<sup>c</sup>*Department of Chemical Engineering, University of Patras, 26504, Rio, Greece,*

\* [akoutinas@aua.gr](mailto:akoutinas@aua.gr)

## Abstract

This study presents sustainability assessment of poly(butylene succinate) (PBS) production from wheat straw and post-consumer alternative End-of-Life (EoL) recirculation routes of the used biopolymer, under both techno-economic and environmental perspective. The processing steps for the whole life cycle of PBS production have been designed and simulated using the UniSim Design software (Honeywell). Techno-economic assessment employing discounted cash flow analysis has been implemented to compare the profitability of PBS production against the appropriate reference product, by estimating important techno-economic indicators. The EoL options for PBS recirculation have been evaluated, targeting the recycling of the material back into the production process, with emphasis on chemical recycling. The environmental impact of the production process and the EoL recirculation route have been evaluated via Life Cycle Assessment. A holistic evaluation of sustainability that takes into account the environmental and techno-economic pillars and combines them into a final life cycle cost has been presented.

**Keywords:** Biopolymer, Lignocellulosic residues, Life Cycle Costing, Sustainability assessment, Poly(butylene succinate)

## 1. Introduction

Global plastic production reached ca. 370 million t in 2019 with an annual increase of 2.5% [1]. The transition towards the bio-economy era necessitates the production of bio-based and biodegradable polymers to reduce the fossil resources used for the current production of the majority of plastics. Poly(butylene succinate) (PBS) is among the most widely used biopolymers, having an annual worldwide production capacity of 86,500 t [2], and constitute a biodegradable polyester produced via polycondensation of succinic acid (SA) and 1,4-butanediol (BDO). It has similar properties to polystyrene (PS), polypropylene (PP) and polyethylene terephthalate (PET) [3], fact that make it ideal candidate for the replacement of such polymers.

The development and implementation of the biorefinery concept are of great importance in order to reinforce the vision toward a sustainable economy based on bioresources. The goal within the integrated biorefinery concept is to start with a biomass feedstock and produce multiple end-products in a systematic and technologically feasible way so as to improve the overall financial performance. The combination of two principles: exploiting the residual biomass use and reducing the associated emissions are approached in the biorefinery concept [4]. Wheat is the most extensively cultivated crop worldwide, playing a vital crop in food supply chains [5]. The wheat residues are a valuable natural resource and their efficient management is essential for sustainable crop production. This study aims to assess potential valorisation of wheat residues, including major processing steps such as pre-extraction, pre-treatment, fermentation, downstream separation and purification (DSP) and polymerization, for the evaluation of the production process of PBS under the life cycle costing perspective.

## 2. Methodology

Process design has been carried out with the design software UniSim (Honeywell). The proposed biorefinery is based on the utilisation of wheat straw, as residue of wheat production. The chemical analysis of raw material used is based on the study of Ballesteros et al. (2006), while the pretreatment process is designed according to Al-Zuhair et al. (2013) [6-7] (Figure 1). Steam explosion and subsequent hydrolysis are utilized for the pretreatment of wheat straw and the production of the sugar rich hydrolysate. More specifically, the straw is mixed with saturated steam at 180°C for 10 min to break the fibers of the lignocellulosic material, and then the pressure is drastically reduced to atmospheric level. After cooling to 45°C, the enzymatic hydrolysis takes place where the lignocellulosic material are degraded into corresponding monomers with the aid of suitable enzymes.

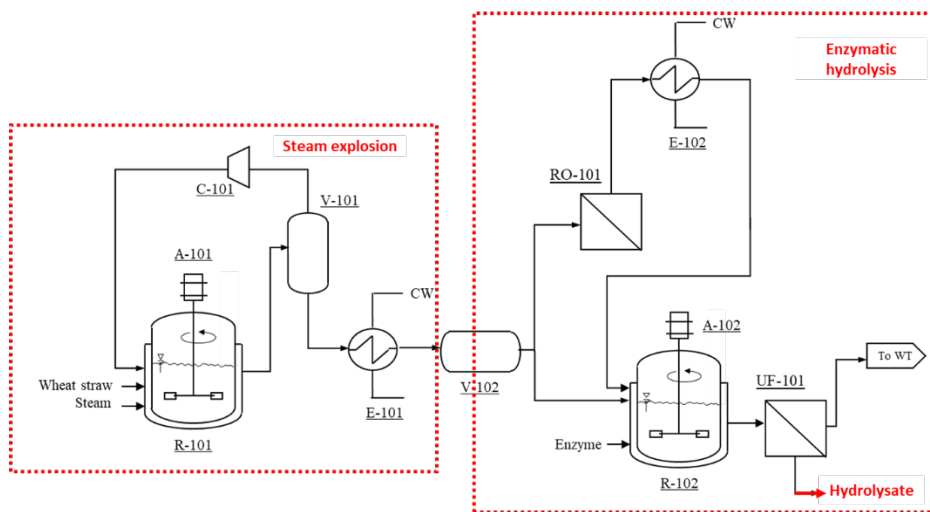


Figure 1: Process flow diagram for wheat straw pretreatment.

The sugar-rich hydrolysate produced from the pretreatment stage is then utilized for the fermentative production of PBS' monomers, e.g. SA and 1,4-BDO. Figure 2 presents the different stages of all production process. The bioconversion stage for both monomers

includes a heat-sterilisation in a continuous process consisting of two heat exchangers and a holding tube. The fermentation medium is sterilized at 140°C and then cooled to the fermentation temperature before addition into the bioreactors. In case of SA, The selected parameters for the fermentation are selected from Ma et al. (2011) and are equal to 101.0 g<sub>SA</sub>·L<sup>-1</sup> concentration, 0.78 g<sub>SA</sub>·g<sup>-1</sup><sub>sugars</sub> yield and 1.18 g<sub>SA</sub>·L<sup>-1</sup>h<sup>-1</sup> productivity [8]. CO<sub>2</sub> supply is used during fermentation due to metabolic requirements in the reductive TCA cycle to produce SA. Likewise, in case of 1,4-BDO the fermentation efficiency is reported by Burgard et al. (2016) using a genetically engineered *E. coli* strain. The final concentration of BDO is 125 g<sub>BDO</sub>·L<sup>-1</sup> with a yield of 0.40 g<sub>BDO</sub>·g<sup>-1</sup><sub>sugars</sub> and a productivity of 3.5 g<sub>BDO</sub>·L<sup>-1</sup>h<sup>-1</sup>, while microaerobic conditions are used during fermentation (0.02 vvm) [9].

The DSP of SA is designed according to Alexandri et al. (2019) [10]. The fermented broth is centrifuged (CF<sub>PBS</sub>-201) to remove the bacterial biomass, fed into the activated carbon columns (V<sub>PBS</sub>-202) for decolorisation and impurity removal and processed by cation-exchange resin columns (V<sub>PBS</sub>-203) in order to transform organic acid salts into their corresponding organic acids. The acidified liquid is concentrated using the MVR-forced circulation evaporator system (EV<sub>PBS</sub>-201) and the concentrated liquid is subsequently treated via crystallisation in continuous crystallizers (CR<sub>PBS</sub>-201-202) at 4°C. The wet succinic acid crystals are dried in a spray dryer (DR<sub>PBS</sub>-201), while the remaining liquid is recycled at the evaporation stage. The SA crystal purity achieved is higher than 99.5%, while the overall succinic acid recovery yield in the DSP is ca. 95% (w/w). In the DSP stage of 1,4-BDO, product is purified to 99.7% purity with recovery yield of 92%. The microbial biomass is initially removed via centrifugation (CF<sub>PBS</sub>-202). The bacterial mass free liquid stream is processed through a series of cation- (V<sub>PBS</sub>-206) and anion-exchange (V<sub>PBS</sub>-207) resin columns to remove the minerals and organic acid salts that are present in the fermentation broth. The outlet liquid stream is subsequently concentrated using a MVR-forced circulation evaporator system (EV<sub>PBS</sub>-202). BDO is purified via distillation (T<sub>PBS</sub>-201) at atmospheric pressure and 180°C in order to separate the water and GBL.

The final stage of the production process is the polymerization stage to form the polymer PBS. The unit operations and the process conditions for the polymerization of PBS were taken from Kamikawa et al. (2013) [11]. This process is divided into three sections, the preparation of raw materials, esterification and polymerization. BDO and SA are initially mixed in a mixing tank (VPBS-301) at a molar ratio of 1.3:1 and 80°C using low pressure steam (LPS). The liquid outflow enters the esterification reactor (R<sub>PBS</sub>-301), after it is heated to 180°C. The reaction takes place at 230°C and 1 bar in order for an appropriate rate of reaction to be achieved and the ester of succinic acid and BDO is formed after 3 h. The polymerization of ester is a polycondensation reaction in the presence of titanium tetrabutoxide as catalyst with a concentration of 2000 ppm (with respect to succinic acid) (R<sub>PBS</sub>-302). The temperature of the reaction is 240°C and the vacuum applied is 2 torr. After 16.5h, the final product from the polycondensation reactor contains PBS with molecular weight of 70,000 Da.



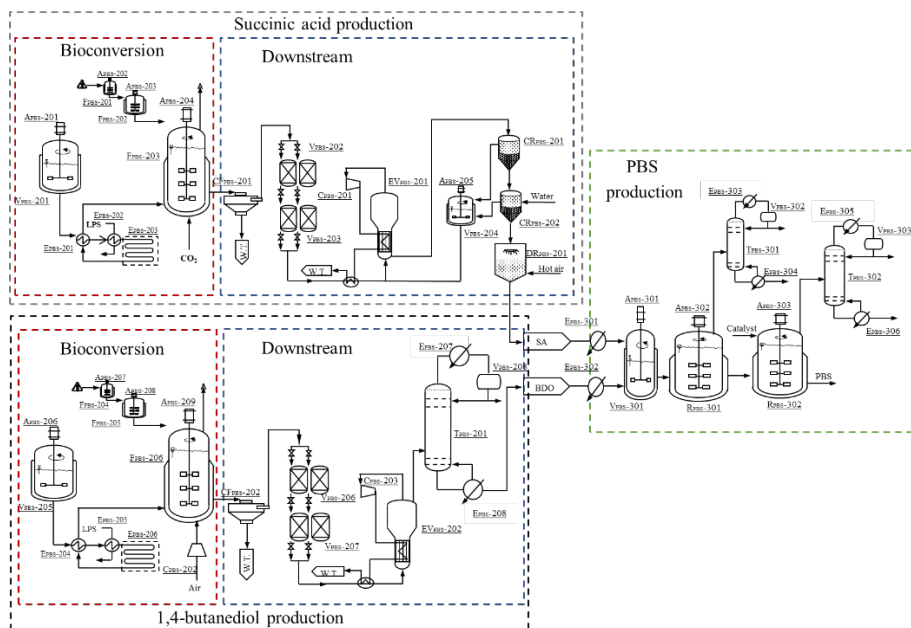


Figure 2: Process flow diagram for PBS production process.

The chemical recycling of PBS is finally performed to evaluate the potential recirculation of the monomer back to the process. The lipase-catalyzed degradation of PBS into cyclic 2-mer is designed according to Okajima et al. (2003) [18]. The reaction is carried out using a mixture of PBS and toluene having a solid to liquid ratio 1:1 (w/v) and 300% immobilized lipase CA (with respect to PBS) at 85 °C for 24 h. The repolymerization of the oligomer product obtained by the degradation reaction is also evaluated with employment of toluene under azeotropic dehydration conditions.

Techno-economic evaluation is based on preliminary economic analysis (accuracy up to  $\pm 30\%$ ). Process design data were used for the estimation of fixed capital investment (FCI) and other important techno-economic indicators. Equipment sizing was based on established methodologies, while the estimation of free-on-board purchased equipment costs estimation was based on established textbooks and literature-cited reports [12-15]. The techno-economic indicators used in this study are the Cost of Manufacture (COM), Minimum Selling Price (MSP), Optimum Plant Capacity (OPC) leading to minimum COM, Discounted Payback Period (DPP) and Minimum Feedstock Requirements (MFR). MSP and DPP are estimated after a discounted cash flow (DCF) is performed. Life Cycle Assessment (LCA) is carried out using the GaBi software and the the ReCiPe 1.08 LCA methodology. The system boundaries for the analysis is “cradle to gate” and the functional unit is 1 kg of PBS. Finally, LCC is performed by estimating twelve different environmental categories followed by their conversion into monetized environmental externalities. The principal stages for the implementation of methodology are described by Bickel and Friedrich (2005) [16]. The monetization of the estimated impacts is based on the environmental prices reported by De Bruyn et al. (2018) for EU28 countries using a dollar to euro exchange rate of 0.856 [17].

### 3. Results and Discussion

Table 1 presents the results for the techno-economic evaluation of the production process of PBS, including monomers production and polymerization stages, without considering the cost of sugars production, as the wheat straw pretreatment is under development.

The FCI and COM per kg PBS are estimated at different plant capacities (10-120 kt/y) to identify the one where these indicators reach a constant value is reached (OPC 58.63 kt PBS per year). The MSP of PBS is lower than its current market price \$4/kg<sub>PBS</sub> [19] with a decrease percentage of 45%. As regards the LCA, some representative results for the Global Warming Potential and the Abiotic Depletion Fossil potential, the most typical environmental indicators, are 2.31 kg CO<sub>2</sub>-eq/kg<sub>PBS</sub> and 15.78 MJ/kg<sub>PBS</sub> when only the production process is taken into consideration. The results for both techno-economic and environmental assessment are expected to be varied when the pretreatment stage is also added in the estimations.

Finally, regarding chemical recycling, it was found that PBS having an Mw of 99000 was quantitatively degraded by lipase to produce cyclic oligomers mainly consisting of the cyclic diester of BS units.

*Table 1. Techno-economic indicators for PBS and PLA production at the optimum plant capacity*

Indicator	PBS
OPC (kt/year)	58,630
FCI (\$/kg)	2.40
COM (\$/kg)	1.85
NPV (M\$)	649.8
MSP (\$/kg)	2.18

### Acknowledgements

This research has been supported by the project “Enhancing diversity in Mediterranean cereal farming systems (CerealMed)” funded by the General Secretariat for Research and Innovation of the Ministry of Development and Investments under the PRIMA Programme. PRIMA is an Art.185 initiative supported and co-funded under Horizon 2020, the European Union’s Programme for Research and Innovation.

### References

- [1] Plastics Europe, 2019. The Circular Economy for Plastics - A European Overview. PlasticsEurope. [https://www.plasticseurope.org/application/files/9715/7129/9584/FINAL\\_web\\_version\\_Plastics\\_the\\_facts2019\\_14102019.pdf](https://www.plasticseurope.org/application/files/9715/7129/9584/FINAL_web_version_Plastics_the_facts2019_14102019.pdf), (accessed February 2021).
- [2] European Bioplastics, 2019. Bioplastics facts and figures. Inst. Bioplastics Biocomposite, Nova-Institute. [https://docs.european-bioplastics.org/conference/Report\\_Bioplastics\\_Market\\_Data\\_2020\\_short\\_version.pdf](https://docs.european-bioplastics.org/conference/Report_Bioplastics_Market_Data_2020_short_version.pdf), (accessed February 2021).
- [3] H.I. Moussa, S.B. Young, Y. Gerand (2012). Polybutylene succinate life cycle assessment variations and variables, in: AIChE Annual Meeting, Conference Proceedings.

- [4] F. Cherubini (2010). The biorefinery concept: using biomass instead of oil for producing energy and chemicals. *Energy Convers. Manag.* 51 (7), 1412–1421.
- [5] R. Rebolledo-Leiva, M.T. Moreira, S. González-García (2022). Environmental assessment of the production of itaconic acid from wheat straw under a biorefinery approach. *Bioresour. Technol.* 345, 126481.
- [6] L. Ballesteros, M.J. Negro, J. M. Oliva, A. Cabañas, P. Manzanares, M. Ballesteros (2006). Ethanol production from steam-explosion pretreated wheat straw. In Twenty-seventh symposium on biotechnology for fuels and chemicals (pp. 496-508). Humana Press.
- [7] S. Al-Zuhair, M. Al-Hosany, Y. Zooba, A. Al-Hammadi, S. Al-Kaabi (2013). Development of a membrane bioreactor for enzymatic hydrolysis of cellulose. *Renew. Energy*, 56, 85-89
- [8] J.F. Ma, M. Jiang, K.Q. Chen, B. Xu, S.W. Liu, P. Wei, H.J. Ying, H.N. Chang, P.K. Ouyang (2011). Strategies for efficient repetitive production of succinate using metabolically engineered *Escherichia coli*. *Bioprocess Biosyst. Eng.* 34 (4), 411–418
- [9] A. Burgard, M.J. Burk, R. Osterhout, S. Van Dien, H. Yim, (2016). Development of a commercial scale process for production of 1,4-butanediol from sugar. *Curr. Opin. Biotechnol.* 42, 118–125.
- [10] M. Alexandri, R. Schneider, H. Papapostolou, D. Ladakis, A. Koutinas, J. Venus (2019). Restructuring the conventional sugar beet industry into a novel biorefinery: fractionation and bioconversion of sugar beet pulp into succinic acid and value-added coproducts. *ACS Sustain. Chem. Eng.* 7 (7), 6569–6579.
- [11] M. Kamikawa, T. Matsuo, K. Oka, T. Kondo, Y. Sase, M. Tanto (2013). U.S. Patent No. 8,604,156. Washington
- [12] M.S. Peters, K.D. Timmerhaus, R.E. West (2003). Plant Design and Economics for Chemical Engineers 5th edition, McGraw-Hill.
- [13] R. Turton, R.C. Bailie, W.B. Whiting, J.A. Shaeiwitz, D. Bhattacharyya (2018). Analysis, Synthesis, and Design of Chemical Processes, Fifth Edition. Prentice Hall, New Jersey.
- [14] E. Dheskali, K. Michailidi, A.M. de Castro, A.A. Koutinas, I.K. Kookos, (2017). Optimal design of upstream processes in biotransformation technologies. *Bioresour. Technol.* 224, 509-514.
- [15] D. Humbird, R. Davis, L. Tao, C. Kinchin, D. Hsu, A. Aden, P. Schoen, J. Lukas, B. Olthof, M. Worley, D. Sexton, D. Dudgeon (2011). Process Design and Economics for Biochemical Conversion of Lignocellulosic Biomass to Ethanol: Dilute-Acid Pretreatment and Enzymatic Hydrolysis of Corn Stover, National Renewable Energy Laboratory.
- [16] P. Bickel, R. Friedrich (2005). ExternE Externalities of Energy: methodology 2005 update, Reproduction. Office for Official Publications of the European Communities, Luxembourg.
- [17] S. De Bruyn, M. Bijleveld, L. de Graaff, E. Schep, A. Schroten, R. Vergeer, S. Ahdour (2018). Environmental Prices Handbook EU28 Version - Methods and numbers for valuation of environmental impacts. CE Delft.
- [18] S. Okajima, R. Kondo, K. Toshima, S. Matsumura (2003). Lipase-catalyzed transformation of poly (butylene adipate) and poly (butylene succinate) into repolymerizable cyclic oligomers. *Biomacromolecules.* 4(6), 1514-1519.
- [19] E4tech, Re-Cord, Wur. (2015). From the Sugar Platform to biofuels and biochemicals. Final Rep. Eur. Comm. Dir. Energy.

# Computer simulation of a plastic waste gasification-solid oxide fuel cell power generation system using Aspen Plus

Mahmudul Hassan Riyad,<sup>a</sup> Sadman Fakid,<sup>a</sup> Nahid Sanzida<sup>a,\*</sup>

<sup>a</sup>*Department of Chemical Engineering, Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh*

\**nahidsanzida@che.buet.ac.bd*

## Abstract

This paper presents a simulation study of a plastic-waste-based combined heat and power (CHP) generation system. The process was modeled using the Aspen Plus V10 simulation software. The proposed power generation system is divided into two sections. The first is the plastic gasification system containing a pyrolyzer and a gasifier. The second part is the electrical power generation process with a pre-reformer, a solid oxide fuel cell (SOFC), and a gas turbine. Sensitivity analysis is done on the gasification process to find the optimum condition for gasification. It was observed that the combinations of relatively moderate temperature and steam to PE mass ratio are beneficial to the performance of the overall gasification system in terms of producing more syngas with greater energy/ exergy efficiency. This study still in the simulation stage aims to pave a way for establishing the plastic-waste-based CHP system as an alternative source of energy.

**Keywords:** Gasification, waste plastic, power generation, solid oxide fuel cell.

## 1. Introduction

Plastic manufacturing began commercially in the 1950s and has grown rapidly to the current worldwide annual production of 368 million tons in 2019 (Errera, 2022). According to the Geneva Environment Network, plastic is one of the major environmental threats (Geneva Environment Network, 2023). Plastic has had a severe impact on the natural ecology, causing problems for humans and wildlife, such as hurting plant life and putting animals in danger (Kumar *et al.*, 2021). Plastic contamination in the oceans is a growing problem all around the world.

Furthermore, harmful chemicals are released when plastic is burned or buried in soil (Verma *et al.*, 2016). Recycling is probably the best way to cut down on plastic waste. Plastic wastes are used in thermo-chemical processes such as combustion, hydrogenation, pyrolysis, and gasification due to their high energy content. These processes have been shown to be effective waste treatment alternatives. They not only eliminate waste, but also convert it into fuel. Separation of feedstock is not required here, unlike in mechanical recycling. Different feed materials, for example, can be combined and transformed to fuel in the same gasifier. Chemical recycling is a potential approach for the treatment of plastic waste for these reasons. This technology can be used to convert a variety of plastic waste materials, such as polyethylene terephthalate (PET), polystyrene (PS), polymethylmethacrylate (PMMA), polypropylene (PP), polyethylene (PE), low and high-

density polyethylene (LDPE and HDPE), and polyvinyl chloride (PVC) (McKeown and Jones, 2020).

On the other hand, there is a growing demand for electricity in present world. Fossil fuels are still the primary source of energy for electricity generation around the world. Using fossil fuels has resulted in global warming and pollution due to the release of greenhouse gases, acid gases, and aerosol particles (Wang et al., 2019). It is crucial to use alternative fuel to generate power in sustainable approach. A recent study by Zhao et al., 2021, on biomass-based combined heat and power (CHP) system consisting of solid oxide fuel cell and gas turbine fed by syngas from a gasifier is reported to have achieved high efficiency for the generation of electrical power. Since plastic waste gasification gives similar products as biomass gasification, a plastic waste based CHP system can be envisioned, which may also be able to generate electrical power.

The aim of this work is to simulate a plastic waste-based small-scale power generation system where all the conversion steps are optimized within the limits of the simulation. This work uses Aspen Plus as the simulation software and waste PE (polyethylene) as the fuel for the simulation. It provides some preliminary qualitative and quantitative data on the gasification process's overall behavior, as well as the sensitivity of process parameters. Moreover, in the plastic waste-based CHP system, along with the optimization of the gasification process, consequent electrical power generation from SOFC is also simulated in Aspen Plus. The overall energy and exergy efficiency of the process is calculated and evaluated in comparison with already existing data.

## 2. Model Development

### 2.1. Modeling the gasification process

Plastic wastes considered in this work consist of polyethylene (PE). The equilibrium model utilized in this study to investigate the simulation of PE gasification is based on the previous work by Mitta *et al.*, 2006. The simplified PE gasification model's process flow sheet is shown in Figure 1. At the very outset, the non-conventional fuel polyethylene was defined, and the ultimate and proximate analyses were provided as input to the model based on previous literature (He *et al.*, 2009). The polymer NRTL/Redlich-Kwong equation of state with Henry's law "POLYNRTL" and "POLYSRK" were chosen as parameter models to calculate the thermo-physical properties of the components. The simulation began with unit processes of drying and separation which was used on the polyethylene sample to rid it of any moisture. The dried fuel is first pyrolyzed by adding external heat to facilitate breaking it down into simpler components suitable for normal gasification. Before being transferred to the gasifier, the gasifying mediums, air and steam, are preheated and combined.

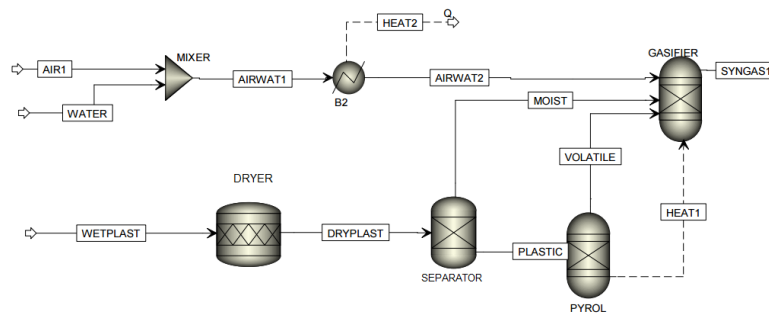


Figure 1: Gasification process flowsheet

2.2. SOFC Stack and Gas Turbine modelling

The SOFC stack model is developed based on the previous works by Doherty *et al.*, 2010. The process flow sheet is shown in Figure 2. Since ion cross-over across the electrolyte cannot be modeled in Aspen Plus, the overall oxidation of H<sub>2</sub> was studied rather than the cell half-reactions; and because only H<sub>2</sub> is electrochemically reacted, it is presumed that CO is shifted to CO<sub>2</sub> and CH<sub>4</sub> is reformed to H<sub>2</sub> (Panopoulos *et al.*, 2006). A fraction of processed air and unreacted CH<sub>4</sub> at SOFC output react in a burner. The high temperature flue gas is then introduced into a gas turbine to produce additional electrical power. In addition, coupling the syngas and air compressors to the gas turbine on the same shaft reduces energy consumption. The description of the simulation blocks in Figures 1 and 2 is given in Table 1.

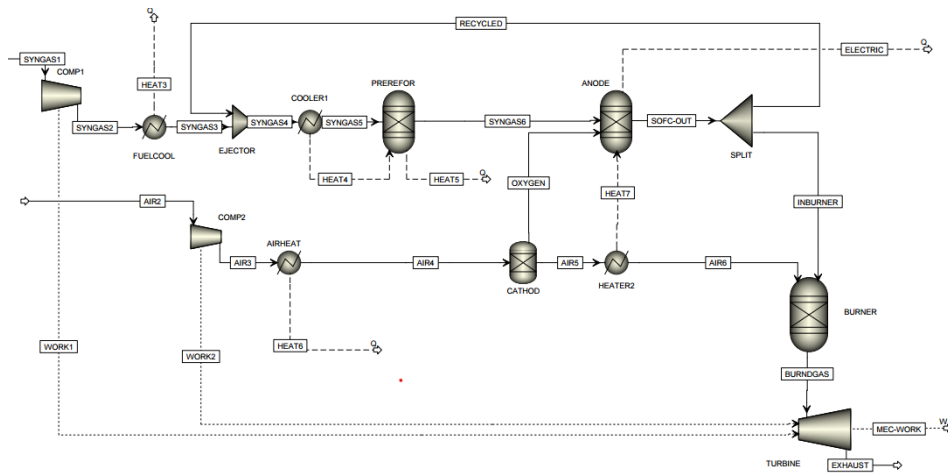


Figure 2: SOFC and Gas Turbine Process Flowsheet

Table 1: Description of Aspen Plus flowsheet unit operation blocks presented in Figure 1 and 2.

Simulation Model	Block ID	Unit	Operating Conditions
RStoic	DRYER	Dryer	Temperature: 110 °C, Pressure: 1atm
Sep	SEPARATOR	Separator	
	CATHODE	SOFC Cathode	
Ryield	PYRO	Pyrolyzer	Temperature: 500 °C, Pressure: 1 atm
RGibbs	GASIFIER	Plastic Gasifier	Temperature: 850–1100 °C (Reactions were specified with equilibrium restricted), Pressure: 1 atm
	PREREFOR	Pre-reformer	Temperature: 571.4 °C (Doherty <i>et al.</i> , 2010), Pressure: 3 atm
	ANODE	SOFC Anode	Temperature: 850 °C (Doherty <i>et al.</i> , 2010), Pressure: 3 atm
	BURNER	Combustor	Temperature: 910 °C (Zhao <i>et al.</i> , 2021), Pressure: 3 atm
Compr	COMP1	Compressor	Isentropic, Pressure ratio: 3 (Doherty <i>et al.</i> , 2010) Mechanical Efficiency: 75%
	COMP2	Compressor	Isentropic, Pressure ratio: 3 (Doherty <i>et al.</i> , 2010) Mechanical Efficiency: 75%
	TURBINE	Gas Turbine	Isentropic, Pressure ratio: 0.33 Mechanical Efficiency: 75%

### 3. Results and Discussion

#### 3.1. Sensitivity analysis of the gasification process

The simulation predicted equilibrium temperature for the gasification process to be 805°C. The effect of O<sub>2</sub> flow rate on the PE gasification process has been investigated in the range of 20 kg/h to 220 kg/h with a constant PE feed rate of 100 kg/h. As demonstrated in Figure 3, an O<sub>2</sub> flow rate of 99.78 kg/h maximizes CO and minimizes CO<sub>2</sub> production. It corresponds to the optimum O<sub>2</sub> to PE mass ratio of nearly 1. Similarly, the effect of steam flow rate is investigated in the range of 50 kg/h to 500 kg/h keeping the PE feed rate constant as before. Figure 4 shows that a steam flow rate of 50.04 kg/h maximizes CO and minimizes CO<sub>2</sub> production. It corresponds to the optimum steam to PE mass ratio of 0.5. Figure 5 shows that operating pressure does not affect the syngas composition much. For this reason, the operating pressure is selected to be 1 atmosphere.

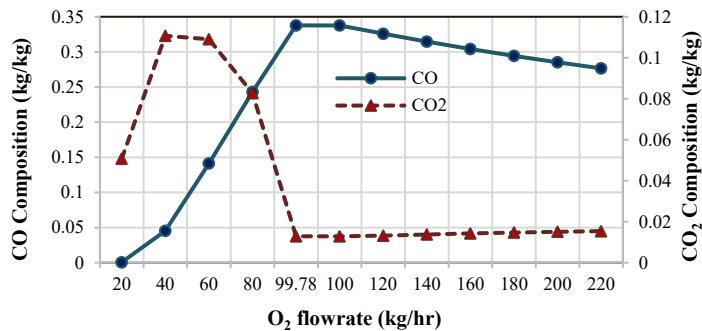


Figure 3: Effect of O<sub>2</sub> flow rate on syngas composition

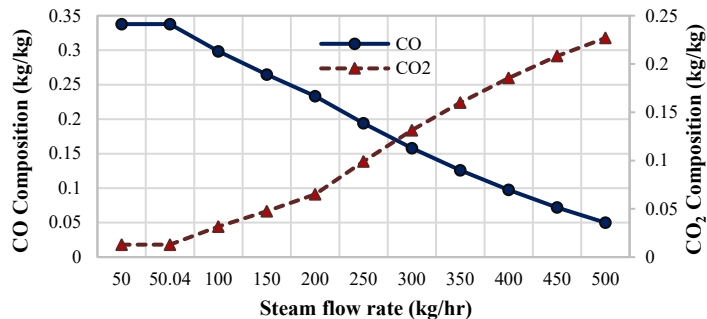


Figure 4: Effect of steam flow rate on syngas composition

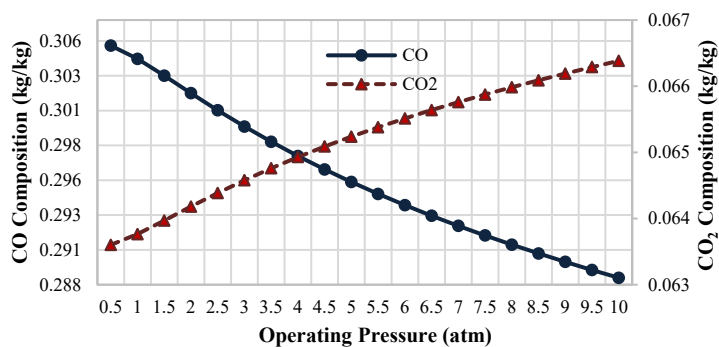


Figure 5: Effect of operating pressure on gasification products

### 3.2. Energy and Exergy analysis of the power generation system

In this study, to ensure maximum power generation, the power generation system is simulated at the optimum gasifier operating temperature of 805°C and the optimum steam to PE mass ratio of 0.5. The Sankey diagram of the process shown in Figure 6 demonstrates the energy flow for evaluating the plant's performance. In the figure, the widths of the different colored bars represent the quantity of energy in kilowatts. Using this system, plastic waste input equivalent to 2269 kW of higher heating value (HHV) results in an electrical power output of 721 kW. The electricity output from SOFC is 582 kW (for details please see Suther *et al.*, 2007) and the gas turbine can generate 139 kW. The energy efficiency is calculated to be 31.4% and the exergy efficiency of the process is 34.2%. These compare favorably to many gasification CHP systems studied in previous literature where the exergy ranged from 30-45% (Colpan *et al.*, 2010; Jia *et al.*, 2015; Mojaver *et al.*, 2019; Karimi *et al.*, 2020).

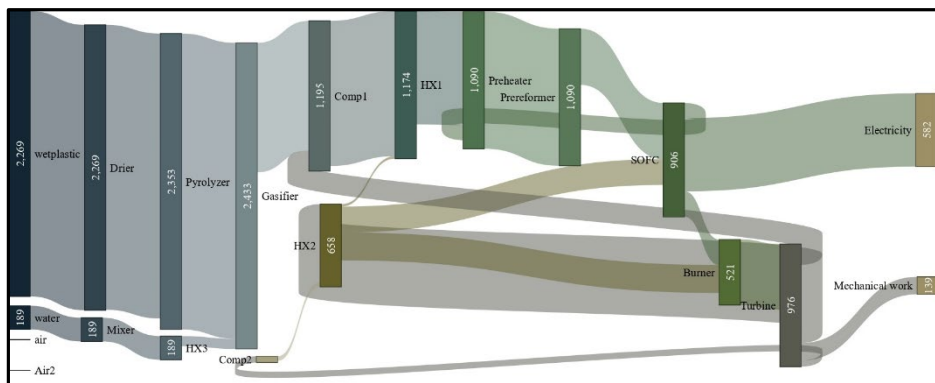


Figure 6: Sankey Diagram of the process

## 4. Conclusions

In this paper, a small-scale CHP system based on the concept of gasifying waste plastic to be used as fuel for power generation through solid oxide fuel cells (SOFC) has been developed using Aspen Plus. To determine the ideal operating conditions for the gasification system, effects of temperature, pressure, and O<sub>2</sub>-steam flow ratio with the feed PE have been examined. The optimal operating conditions for gasification are determined to be 805 °C and 1 atm. The optimum O<sub>2</sub> to PE mass ratio is nearly 1, and the optimum steam to PE mass ratio is 0.5. In addition, the energy and exergy assessments of the complete power generation system are performed. This power generation system generates 721 kW of electrical power with an overall energy efficiency of 31.4 percent. The overall exergy efficiency is 34.2 percent. The results compare favorably to contemporary works and offer solutions to the prevailing plastic-waste recycling problem and alternative fuel crisis. Even though the results from this work are heavily dependent on the assumptions made, and it was within many limits of the simulation, it would help to establish plastic waste-based CHP system as a viable topic for research.

## References

- C. O. Colpan, F. Hamdullahpur, I. Dincer, and Y. Yoo, 2010, Effect of Gasification Agent on the Performance of Solid Oxide Fuel Cell and Biomass Gasification Systems, *International Journal of Hydrogen Energy*, 35(10), Pages 5001–5009. doi: <https://doi.org/10.1016/j.ijhydene.2009.08.083>.



- W. Doherty, A. Reynolds, and D. Kennedy, 2010, Computer Simulation of a Biomass Gasification-Solid Oxide Fuel Cell Power System using Aspen Plus, *Energy*, 35(12), Pages 4545–4555. doi: 10.1016/j.energy.2010.04.051.
- R. Errera, 2022, Surprising Facts about Plastic (Production, Use, and Recycling - Global Data 2023), Available at: <https://www.tonerbuzz.com/blog/facts-about-plastic/#:~:text=Global%20Plastic%20Industry%20Facts%20and%20Statistics&text=The%20world%20produced%200.3%20percent,of%20118%20million%20metric%20tons>. Accessed 30 January 2023.
- Geneva Environment Network, 2023, Plastics and the Environment, Available at: <https://www.genevaenvironmentnetwork.org/resources/updates/plastics-and-the-environment/>. Accessed 30 January 2023.
- M. He, B. Xiao, Z. Hu, S. Liu, X. Guo, and S. Luo, 2009, Syngas Production from Catalytic Gasification of Waste Polyethylene: Influence of Temperature on Gas Yield and Composition, *International Journal of Hydrogen Energy*, 34(3), Pages 1342–1348. doi: <https://doi.org/10.1016/j.ijhydene.2008.12.023>.
- J. Jia, A. Abudula, L. Wei, B. Sun, and Y. Shi, 2015, Thermodynamic Modeling of an Integrated Biomass Gasification and Solid Oxide Fuel Cell System', *Renewable Energy*, 81, Pages 400–410. doi: <https://doi.org/10.1016/j.renene.2015.03.030>.
- M. H. Karimi, N. Chitgar, M. A. Emadi, P. Ahmadi, and M. A. Rosen, 2020, Performance Assessment and Optimization of a Biomass-based Solid Oxide Fuel Cell and Micro Gas Turbine System Integrated with an Organic Rankine Cycle, *International Journal of Hydrogen Energy*, 45(11), Pages 6262–6277. doi: <https://doi.org/10.1016/j.ijhydene.2019.12.143>.
- R. Kumar, A. Verma, A. Shome, R. Sinha, S. Sinha, P.K. Jha, R. Kumar, P. Kumar, Shubham, S. Das, P. Sharma, P.V. Vara Prasad, 2021, Impacts of Plastic Pollution on Ecosystem Services, Sustainable Development Goals, and Need to Focus on Circular Economy and Policy Interventions, *Sustainability*, 13, Page 9963. <https://doi.org/10.3390/su13179963>.
- P. McKeown, and M. D. Jones, 2020, The Chemical Recycling of PLA: A Review, *Sustainable Chemistry*, 1(1), Pages 1–22. doi: 10.3390/suschem1010001.
- N. R. Mitta, S. Ferrer-Nadal, A. M. Lazovic, J. F. Parales, E. Velo, and L. Puigjaner, 2006, Modelling and Simulation of a Tyre Gasification Plant for Synthesis Gas Production, in W. Marquardt, and C. Pantelides, (eds), *Computer Aided Chemical Engineering*. Elsevier 21, Pages 1771–1776. doi: [https://doi.org/10.1016/S1570-7946\(06\)80304-4](https://doi.org/10.1016/S1570-7946(06)80304-4).
- P. Mojaver, S. Khalilarya, and A. Chitsaz, 2019, Multi-Objective optimization using Response Surface Methodology and Exergy Analysis of a Novel Integrated Biomass Gasification, Solid Oxide Fuel Cell and High-Temperature Sodium Heat Pipe System, *Applied Thermal Engineering*, 156, Pages 627–639. doi: <https://doi.org/10.1016/j.applthermaleng.2019.04.104>.
- K. D. Panopoulos, L.E. Fryda, J. Karl, S. Poulou, and E. Kakaras, 2006, High Temperature Solid Oxide Fuel Cell Integrated with Novel Allothermal Biomass Gasification: Part I: Modelling and Feasibility Study, *Journal of Power Sources* 159 (1), Pages 570-585. <https://doi.org/10.1016/j.jpowsour.2005.12.024>
- T. Suther, A. Fung, and M. Koksai, 2007, Simulation of a Solid Oxide Fuel Cell-Gas Turbine System Using AspenPlus™, in *Proceedings of 3<sup>rd</sup> International Energy, Exergy and Environment Symposium*, Portugal 2007, Pages 1–18.
- R. Verma, K.S. Vinoda, M. Papireddy, A.N.S. Gowda, 2016, Toxic Pollutants from Plastic Waste- A Review, *Procedia Environmental Sciences*, 35, Pages 701–708. doi: 10.1016/j.proenv.2016.07.069.
- H. Wang, Z. Lei, X. Zhang, B. Zhou, and J. Peng, 2019, A Review of Deep Learning for Renewable Energy Forecasting, *Energy Conversion and Management*, 198, Page 111799. doi: <https://doi.org/10.1016/j.enconman.2019.111799>.
- Z. Zhao, Y. A. Situmorang, P. An, J. Yang, X. Hao, J. Rizkiana, A. Abudula, and G. Guan, 2021, A Biomass-based Small-Scale Power Generation System with Energy/Exergy Recuperation, *Energy Conversion and Management*, 227, Page 113623. doi: <https://doi.org/10.1016/j.enconman.2020.113623>.

# Enhancement of hydrochar from coal discards and sewage sludge (co)-HTC for transition to a low-carbon economy

Gentil Kahilu Mwengula,<sup>a</sup> Orevaoghene Eterigho-Ikelegbe,<sup>b</sup> Jean Mulopo,<sup>a</sup>

Samson Bada<sup>a</sup>

*DSI-NRF SARCHI Clean Coal Technology Research Group, School of Chemical and Metallurgical Engineering, Faculty of Engineering and the Built Environment, University of the Witwatersrand, Wits 2050, Johannesburg, South Africa.*

*<sup>b</sup>Sustainable Energy and Environment Research Group, School of Chemical Engineering, University of Witwatersrand PO Box 3, Johannesburg, Wits 2050, South Africa.*

\*Correspondence and requests for materials should be addressed to G.K (email:

[1902723@students.wits.ac.za](mailto:1902723@students.wits.ac.za))

## Abstract

Climate change scenarios highlight the significance of modifying production patterns in order to transition to a low-carbon economy, as well as the value of recycling innovations in this regard. This study examines the co-hydrothermal treatment (Co-HTC) of coal discard (CD) and sewage sludge (SS) in order to enhance the carbon content of the produced hydrochars for a sustainable circular and low-carbon economy. Using a coal-sewage sludge blend ratio of 5:1, the optimal hydrothermal carbonization (HTC) and Co-HTC parameters for producing hydrochars with the highest fixed carbon (FC) and lowest ash content (A) were 150°C, 27bar, 92.13 minutes and 208°C, 22.5 bar, 331 minutes, respectively. The physicochemical parameters of the optimized hydrochar were evaluated and compared to those of the raw materials. According to the results, HTC and Co-HTC raised the calorific value of CT and CS to 19.33 MJ/kg, 25.79 MJ/kg, and 24.31 MJ/kg, respectively.

**Keywords:** hydrothermal carbonization; coal tailings; coal slurry; sewage sludge, hydrochar, low-carbon economy.

## 1. Introduction

South Africa (SA) is one of the world's greatest coal producers and uses coal as its primary energy source (Bohlmann et al., 2018). According to the 2001 National Coal discard and slurry inventory compiled by the South African Department of Energy, about 65 million tons of coal wastes are produced annually. These coal by-products are disposed of in either tailings heaps or slurry dams (Van de Venter et al., 2021). Due to the solubilization of harmful pollutants from coal discard (CD) and the potential of spontaneous combustion, coal waste disposal is seen as a significant threat to the country's environmental waste management (Onifade and Genc 2020). Beneficiation procedures, such as physico-chemical processes and regeneration techniques, have evolved over time,

however they are regarded as inefficient, not ecologically friendly, difficult, and costly (Petlovanyi et al., 2018). While the mining industry contributes to global CO<sub>2</sub> emissions, it may also be critical to lowering global CO<sub>2</sub> emissions, particularly in developing economies, not only by supporting responsible mineral extraction and processing, but also by ensuring the adoption of a circular economy approach based on mineral waste reuse and beneficiation as shown in Figure 1 for the cases of coal discards and sewage sludge. Conversely, South African wastewater treatment plants produce a substantial amount of sewage sludge (SS) (Grobelak et al., 2019). Current SS management procedures, including on-site land disposal and waste piling, are deemed unsustainable (Shaddel et al., 2019). Therefore, creative approaches to managing the CD and SS are thought essential. Hydrothermal carbonization technology (HTC) was used in this study to improve the physicochemical properties of CT, CS, and a blend of the two in order to develop suitable precursors for activated carbon, and other carbonaceous materials (value added products). The Co-HTC treatment of CD-SS mixture has the potential to enhance the carbon life cycle in coal waste while also offering an ecologically closed loop for WWTP effluent treatment, which is the cornerstone of this work (Figure 1).

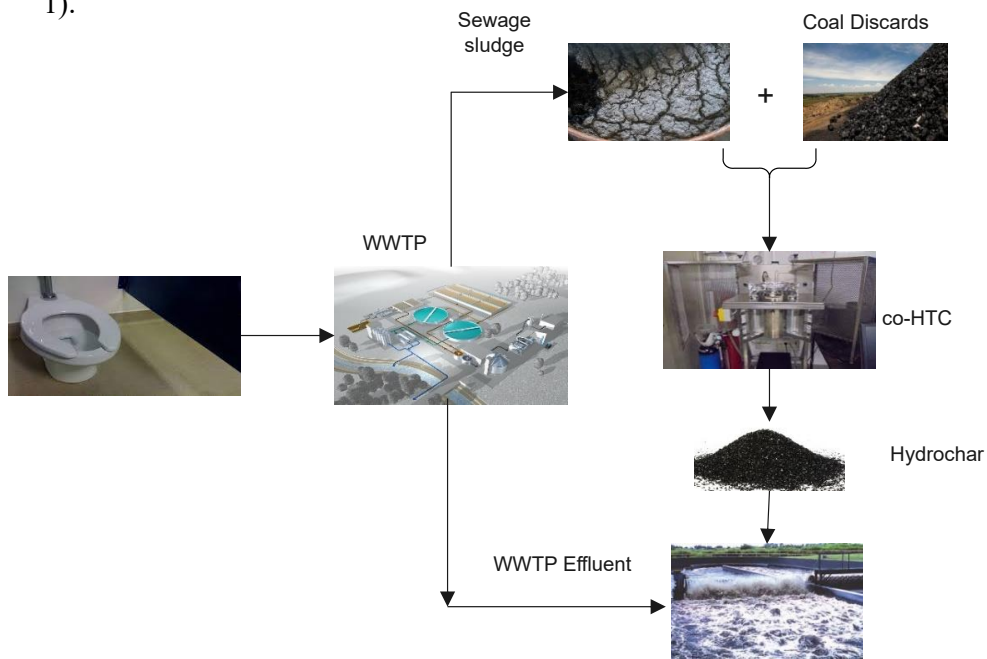


Figure 1: Co-treatment of coal waste and sewage sludge to improve the carbon life cycle in coal waste while providing an ecologically closed loop for WWTP effluent treatment.

## **2. Materials and methods**

### **2.1. Materials**

The coal samples utilized in this study are coal tailing (CT) and coal slurry (CS) obtained from a coal beneficiation plant in Mpumalanga, South Africa. The sewage sludge (SS) obtained from a wastewater treatment plant (Ekurhuleni Water Care Company (ERWAT)).

### **2.2. Experimental Methods**

The HTC and Co-HTC tests were conducted in the high-pressure tubular reactor using –1mm coal samples. HTC were performed, individually, on coal tailing (CT) and coal slurry (CS), while Co-HTC were performed on the coal-sewage sludge mixtures. 25g of the feedstock (coals, mixtures of coal and sewage sludge) was mixed with water as input solvent in a solid to liquid ratio of 1:4 in the reactor. The HTC and Co-HTC process parameters (temperature, pressure, residence time and CT+CS:SS ratio) were varied according to the design expert matrix. For experimental designs, the central composite design (CCD) and the custom design (CD) were utilized to know the number of runs that will be required to optimize the HTC and Co-HTC process respectively. Based on the CCD, 20 experimental runs were performed for the HTC of CT and CS, six center points, height factorial points and six axial points and was repeated five times to ensure reproducibility of the responses and to determine accurate experimental error. However, 40 experimental runs were performed according to the CD matrix of the Co-HTC process. The CD matrix predicted 20 edge\* interior points and 20 vertex\* interior points to accurately determine the process optimum conditions. The obtained solid products are identified as hydrochar coal tailing (HCT), hydrochar coal slurry (HCS) and hydrochar blended coal and sewage sludge (HCB).

## **3. Results and discussions**

The impact of operating conditions on the FC of the produced hydrochars (HC) is illustrated by the 3D plot of the response surface (Fig 2-3). The slopes observed on the surface plots of responses indicate the impact of reaction temperature on the FC of the produced hydrochars. The increase of temperature above 150°C decreases the FC of the produced HCT and HCS. However, the increase of temperature up to 208°C and (CT+CS): SS ratio up to 3:1 increases the FC of the produced HCB.

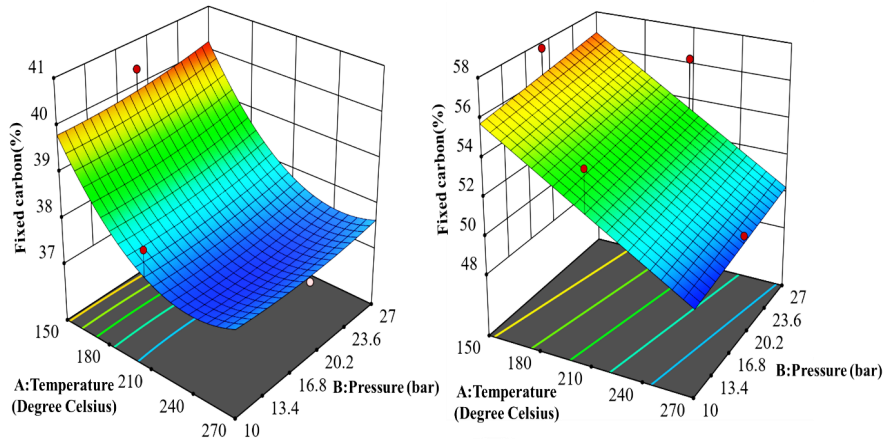


Figure 2: Surface plot showing the influence of temperature and pressure on the Fixed Carbon of Coal tailing's HC (left) and Coal slurry's HC (right).

The results indicate that 150°C, 27bar and 92.13 minutes were the optimum HTC conditions, while 208°C, 22.5 bar, 331 minutes and 20.02:4.98(CT+CS: SS) were optimum Co-HTC conditions. The process water (PW) analysis during the HTC and Co-HTC processes identified several chemical components that were dissolved in the liquid phase. For economic and environmental reasons, the PW generated by the HTC and Co-HTC processes was recycled back into the processes in this work.

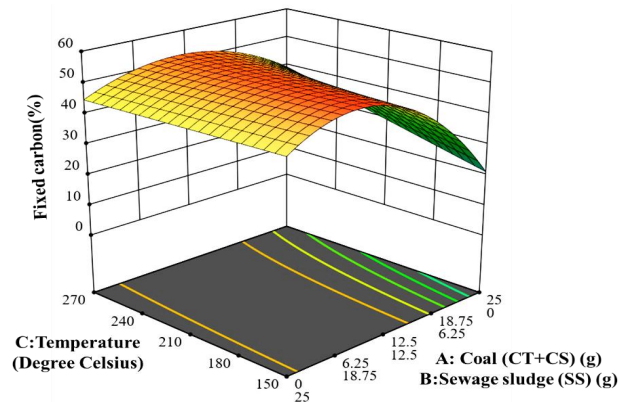


Figure 3: Surface plot showing the influence of temperature, pressure, time and the (CT+CS): SS ratio on the Fixed Carbon of produced hydrochar.

Additionally, the HTC and Co-HTC processes have shown the ability to upgrade the fixed carbon (FC) contents of coal discards and/or sewage sludge by releasing mineral matter (ash content), oxygen, and sulphur content from their molecular structures, which improved their porosity characteristics (Aliakbari *et al.*, 2018; Saba *et al.*, 2020). The data obtained further show that the optimal HTC and Co-HTC conditions favored the

hydrolysis of aliphatic components, the dissolution of mineral materials, and the prevention of carboxylate gas release. Table 1 shows the effect of the HTC and co-HTC processes under optimum settings on the physicochemical properties of the raw materials. Additionally, the HTC and Co-HTC processes raised the calorific value of raw CT, CS, and CB from 16.59MJ/kg, 24.40MJ/kg, and 17.68MJ/kg to 19.33MJ/kg, 25.79MJ/kg, and 24.99MJ/kg, respectively. Due to comparable interactions (acidic environment caused by the decomposition of SS) occurring between the feedstocks during the Co-HTC process, the CV increase of produced HCB (41.34%) was greater than that of HCT and HCS (16.51% and 5.69%). Table 2 compares the properties of the produced hydrochars to various previous study findings. The data confirmed that the optimal HTC and co-HTC conditions established in this study could be efficient for prospective carbonization of comparable feedstocks.

Table 1. Physicochemical properties of the raw materials and produced hydrochars

Analyses	CT	HCT	CS	HCS	SS	CB	HCB
<i>Proximate Analysis (wt%, adb)</i>							
Inherent Moisture	4.17	1.07	3.94	1.73	8.12	2.71	1.72
Ash content	38.64	37.96	23.22	22.78	36.07	37.75	19.56
Volatile Matter	21.44	21.04	21.28	18.38	47.07	21.92	20.31
Fixed Carbon	<b>35.75</b>	<b>39.94</b>	<b>50.98</b>	<b>57.11</b>	<b>8.75</b>	<b>37.62</b>	<b>58.41</b>
<i>Ultimate Analysis (wt. %, adb)</i>							
Carbon	<b>42.82</b>	<b>49.80</b>	<b>61.85</b>	<b>66.90</b>	<b>29.7</b>	<b>45.64</b>	<b>67.04</b>
Hydrogen	3.01	2.87	3.56	2.98	4.88	3.05	2.78
Nitrogen	1.14	1.72	1.39	1.99	4.15	1.82	2.43
Oxygen	8.79	5.83	4.78	2.75	15.22	8.2	6.08
<i>Total Sulfur (wt. %, adb)</i>	1.43	0.75	1.26	0.87	1.86	0.83	0.39
<i>Calorific value (MJ/kg)</i>	16.59	19.33	24.4	25.79	13.2	17.68	24.31
<i>Surface area (m<sup>2</sup>/g)</i>	6.06	11.88	6.37	14.35		6.17	20.35
<i>CO<sub>2</sub> emission (mole)</i>		0.16		0.1			0.6
<i>Mass yield (%)</i>	-	<b>97.66</b>	-	<b>94.74</b>	-	-	<b>88.38</b>

CT: coal Tailing. CS: Coal Slurry. HCT: Hydrochar coal tailing. HCS: Hydrochar coal slurry. Solid/liquid ratio: ¼. adb: air dried basis. Oxygen% = 100 – (Moiture + Ash + Total Carbon + Hydrogen + Nitrogen + Sulphur)

Table 2. Comparison of the produced hydrochars from this work with literature values

Material	HTC conditions		Proximate analysis (wt.%)			Ultimate analysis (wt.%)					Source
	Temperature (°C)	Time (Min)	Ash*	Fixed carbon*	Volatile matter*	C	H	N	O*	S	
Clarion coal (Cc)	-	-	11.7	54.0	34.3	63.8	4.1	1.5	26	4.6	Saba et al., 2017
HC	260	30	7.3	57.3	35.4	69.1	4.3	1.7	21.3	3.5	
Miscanthus (Ms)	-	-	2.7	12.8	84.5	48.7	5.7	0.2	45.4	0	
HC	260	30	0.7	47.8	51.5	69.6	4.4	0.3	25.6	0	
Blend (Cc + Ms)	200	90	4.2	33.1	62.7	58.2	5.3	0.8	34.4	1.2	
HC	260	90	3.5	53.4	43.1	68.7	4.1	1.1	24.4	1.6	
Bituminous Coal (CW1)	-	-	64.4	15.5	20.1	20.5	1.9	0.5	4.2	9.6	Reza et al., 2022
HC	230	30	62.8	18.7	18.5	22.5	1.5	0.5	5.4	8.4	
Bituminous Coal (CW)			67.4	14.2	18.4	18.6	1.8	0.6	4.1	8.5	
HC	230	30	65.4	17.6	17.0	20.2	1.6	0.7	5.2	7.9	
Food waste (FW)			9.6	18.8	71.6	39.3	6.0	1.5	44.0	0	
CW+FW	230	30	28.6	31.2	40.2	49.5	4.5	1.9	14.8	1.4	
Sub-bituminous coal (SbC)	-	-	25.5	40.5	34	75.1	5.1	1.1	18.4	0.33	Saqib et al., 2018
Coal			11.74	55.48	32.78	64.07	4.24	1.23	18.07	0.65	Zhao et al., 2021
HC (Coal+SbC)	340	60	7.16	65.12	27.72	74.81	3.5	1.88	12.15	0.56	
Coal discard (CD)	-	-	41.95	35.83	20.17	48.90	2.67	1.15	1.93	1.34	Setsepu et al., 2021
Searsia lancea Trees Grown (Ssl)	-	-	3.89	18.41	69.40	45.12	6.35	0.44	35.78	<0.10	
HC (CD + Ssl)	280	30	1.49	50.33	46.12	-	-	-	-	-	
Coal tailing (CT)	-	-	40.32	37.31	22.37	42.82	3.01	1.14	12.96	1.43	<b>This work</b>

HC	150	95	38.36	40.37	21.27	49.80	2.87	1.72	6.9	0.75
Coal slurry (CS)			24.17	53.02	22.69	61.85	3.56	1.39	8.72	1.26
HC	150	95	20.12	57.69	22.19	66.90	2.98	1.99	4.48	0.87
Sewage sludge (SS)	-	-	39.25	9.52	51.23	29.7	4.88	4.15	23.34	1.86
CT+CS+SS blend	-	-	38.8	38.67	22.53	45.64	3.05	1.82	10.91	0.83
HC	208.10	360	19.71	58.82	21.46	67.04	2.78	2.43	7.8	0.39

Ash: Ash content. FC: Fixed carbon. VM: Volatile matter. \*: dried basis (moisture free). HC: Hydrochar.  $Oxygen\% = 100 - (Moisture + Ash + Total\ Carbon + Hydrogen + Nitrogen + Sulphur)$

#### 4. Conclusion

The conversion of carbon-based wastes into energy and chemical products is a potential approach for waste treatment. In order to improve the hydrothermal carbonization (HTC) of coal discards from coal beneficiation operations and/or sewage sludge, this work optimized a range of process parameters to enable the synthesis of hydrochar with improved physicochemical properties. The best hydrothermal carbonization (HTC) and Co-HTC parameters for synthesizing hydrochars with the highest fixed carbon (FC) and lowest ash content (A) using a coal-sewage sludge blend ratio of 5:1 were 150°C, 27bar, 92.13 minutes and 208°C, 22.5 bar, 331 minutes, respectively. The enhanced hydrochar's physicochemical properties were assessed and compared to those of the raw materials. HTC and Co-HTC increased the calorific value of CT and CS to 19.33 MJ/kg, 25.79 MJ/kg, and 24.31 MJ/kg, respectively.

#### References

1. Aliakbari, Z., Younesi, H., Ghoreyshi, A. A., Bahramifar, N., & Heidari, A. (2018). Production and Characterization of Sewage-Sludge Based Activated Carbons Under Different Post-Activation Conditions. *Waste Biomass Valor*, 9(3), 451–463. <https://doi.org/10.1007/s12649-016-9823-7>
2. Bohlmann, J. A., & Inglesi-Lotz, R. (2018). Analysing the South African residential sector's energy profile. *Renewable and Sustainable Energy Reviews*, 96, 240–252. <https://doi.org/10.1016/j.rser.2018.07.052>
3. Grobelak, A., Grosser, A., Kacprzak, M., & Kamizela, T. (2019). Sewage sludge processing and management in small and medium-sized municipal wastewater treatment plant-new technical solution. *Journal of Environmental Management*, 234, 90–96. <https://doi.org/10.1016/j.jenvman.2018.12.111>
4. Mazumder, S., Saha, P., & Reza, M. T. (2022). Co-hydrothermal carbonization of coal waste and food waste: fuel characteristics. *Biomass Conv. Bioref.*, 12(1), 3–13. <https://doi.org/10.1007/s13399-020-00771-5>
5. Onifade, M., Genc, B., & Bada, S. (2020). Spontaneous combustion liability between coal seams: A thermogravimetric study. *International Journal of Mining Science and Technology*, 30(5), 691–698. <https://doi.org/10.1016/j.ijmst.2020.03.006>



6. Petlovanyi, M. V., Lozynskyi, V. H., Saik, P. B., & Sai, K. S. (2018). Modern experience of low-coal seams underground mining in Ukraine. *International Journal of Mining Science and Technology*, 28(6), 917–923. <https://doi.org/10.1016/j.ijmst.2018.05.014>
7. Saba, A., Saha, P., & Reza, M. T. (2017). Co-Hydrothermal Carbonization of coal-biomass blend: Influence of temperature on solid fuel properties. *Fuel Processing Technology*, 167, 711–720. <https://doi.org/10.1016/j.fuproc.2017.08.016>
8. Saqib, N. U., Baroutian, S., & Sarmah, A. K. (2018). Physicochemical, structural and combustion characterization of food waste hydrochar obtained by hydrothermal carbonization. *Bioresource Technology*, 266, 357–363. <https://doi.org/10.1016/j.biortech.2018.06.112>
9. Setsepu, R. L., Abdulsalam, J., Weiersbye, I. M., & Bada, S. O. (2021). Hydrothermal Carbonization of *Searsia lancea* Trees Grown on Mine Drainage: Processing Variables and Product Composition. *ACS Omega*, 6(31), 20292–20302. <https://doi.org/10.1021/acsomega.1c02173>
10. Shaddel, S., Bakhtiary-Davijany, H., Kabbe, C., Dadgar, F., & Østerhus, S. (2019). Sustainable Sewage Sludge Management: From Current Practices to Emerging Nutrient Recovery Technologies. *Sustainability*, 11(12), 3435. <https://doi.org/10.3390/su11123435>
11. Sultana, A. I., & Reza, M. T. (2022). Investigation of hydrothermal carbonization and chemical activation process conditions on hydrogen storage in loblolly pine-derived superactivated hydrochars. *International Journal of Hydrogen Energy*, 47(62), 26422–26434. <https://doi.org/10.1016/j.ijhydene.2022.04.128>
12. Van de Venter, F. F. *Producing briquettes for domestic household fuel applications from coal tailings (Doctoral dissertation)*. (2021). (n.d.).
13. Wang, L., Chang, Y., & Li, A. (2019). Hydrothermal carbonization for energy-efficient processing of sewage sludge: A review. *Renewable and Sustainable Energy Reviews*, 108, 423–440. <https://doi.org/10.1016/j.rser.2019.04.011>
14. Zhao, P., Lin, C., Zhang, J., Huang, N., Cui, X., Tian, H., & Guo, Q. (2021). Moisture re-adsorption characteristics of hydrochar generated from the Co-hydrothermal carbonization of PVC and alkali coal. *Fuel Processing Technology*, 213, 106636. <https://doi.org/10.1016/j.fuproc.2020.106636>

# Fully Electrified Conversion of Low-Quality Plastic Waste to Polymer Precursors

Kristiano Prifti<sup>a,b</sup>, Andrea Galeazzi<sup>a,b</sup>, Adrián Pacheco-López<sup>a,c</sup>, Antonio Espuña<sup>c</sup>, Flavio Manenti<sup>a,b\*</sup>

<sup>a</sup>*Politecnico di Milano, CMIC Department "Giulio Natta", Piazza Leonardo da Vinci, 32, Milan 20133, Italy*

<sup>b</sup>*Consorzio Interuniversitario Nazionale per la Scienza e Tecnologia dei Materiali, Via Giusti 9, 50121 Firenze, Italy*

<sup>c</sup>*Department of Chemical Engineering, Universitat Politècnica de Catalunya, Escola d'Enginyeria de Barcelona Est, C/ Eduard Maristany 16, Barcelona 08019, Spain*

[\\*flavio.manenti@polimi.it](mailto:*flavio.manenti@polimi.it)

## Abstract

A new process design for plastic waste gasification and conversion to methanol is described in its main elements. An electrification approach is used to maximize the CO<sub>2</sub> credit generation of the plant. The overall process yields are described for three different feedstocks made of HDPE, PP, and PS. Electrical consumption is broken down by unit contributions to identify process bottlenecks. Preliminary profitability analysis is done with and without accounting for the possibility of selling credits CO<sub>2</sub> showing that the process is unprofitable unless CO<sub>2</sub> credits are accounted for. Depending on the feedstock, CO<sub>2</sub> credits between 89.5 and 103.2 €/t are shown estimated to reach a payback time of 3.5 years without accounting for labor and taxes.

**Keywords:** process simulation, circular economy, electrification, plastic waste, sustainable development, chemical recycling

## 1. Introduction

The issue of developing a fully circular supply chain for plastic materials is a crucial topic for waste management and the carbon economy (Pacheco-López et al., 2021). Chemical recycling of plastic waste has become a hot topic in recent years with pyrolysis being the main focus (Qureshi et al., 2020). Mechanical recycling has reached commercial maturity with the current bottleneck being feedstock purity and property degradation over multiple cycles (Schyns and Shaver, 2021). Gasification is a supplementary route for contaminated feedstocks that have become unsuitable for mechanical recycling and are available in large amounts. Continuous gasification processes are derived from biomass and asphalts treatment technologies to manufacture syngas which could be used to synthesize polymer monomeric precursors. The circular vision seeks to reach polymeric precursors that can be used to make polymers of virgin quality, which in turn can be recycled over multiple loops after usage, thus drastically extending the lifespan of the raw material in the market. This work develops and analyses a plant configuration that converts plastic waste through mild gasification (Kristiano Prifti et al., 2021) and fully electrifies the downstream of the gasification unit to make methanol with no CO<sub>2</sub> emissions assuming enough renewable electrical energy is available. The process was designed with the intent of maximizing CO<sub>2</sub> credit generation and carbon emission abatement. The boundaries are a mix of PE,

PS, PP waste, demi-water, and renewable electricity as an input to reach methanol of commercial purity as a final product. The process is divided into five conceptual blocks: gasification, light hydrocarbon bi-reforming, syngas conditioning, methanol synthesis, and methanol purification. The model was developed in Aspen HYSYS V11 using a thermodynamic equilibrium approach for the gasification and reforming reactors and kinetic models in methanol synthesis (Bisotti et al., 2021). Syngas conditioning is carried out with a commercial PEM electrolyzer (Carmo et al., 2013) which generates the hydrogen for CO<sub>2</sub> hydrogenation to methanol and the oxygen necessary for the gasification consuming 55 kWh/kg hydrogen. The environmental impact is evaluated with a simple CO<sub>2</sub> equivalent method while the economic performance was estimated using the CORO framework (Prifti et al., 2022) to compute both capital and operative expenditures. A sensitivity analysis on feed composition is carried out to create a surrogate model (Galeazzi et al., 2022). Which is intended to be introduced in the ontological framework developed by Pacheco-López et al (2021) to further expand such ontology in both applicability and flexibility, as well as evaluate the role this process might have in the value chain of plastic waste in the future.

## 2. Methods

The process was designed to start with a feedstock specification of 1000 kg/h of plastic waste mixtures of different compositions. Three feedstocks were analyzed:

- 48% HDPE, 32% PP, 20% PS: H/C ratio of 1.55, Mass Fraction of C: 88.59%
- 54% HDPE, 36% PP, 10% PS: H/C ratio of 1.73, Mass Fraction of C: 87.40%
- 60% HDPE, 40% PP: H/C ratio of 2, Mass Fraction of C: 85.71%

Changing the overall mass fraction of carbon greatly impacts the performance of the gasification and oxygen demand of the system. The PEM electrolyzer is assumed to have a consumption of 55 kWh/kg of hydrogen, and the specification of the module is given on hydrogen demand. The PEM produces all the hydrogen necessary for the plant to fully hydrogenate CO<sub>2</sub>. Oxygen, if produced in excess is sold as a secondary product, while, if not produced in sufficient amount, is supplemented with an outside source (cylinders or Air Separation Unit). The PEM electrolyzer is a bottleneck for pressure in the upstream of the methanol loop. Pressure is set at 20 bar, which corresponds with the highest pressure supported by commercial solutions for a PEM electrolyzer considered in this study.

The steam methane reformer is electrified and follows the design commercialized by Topsoe A/S (Wismann et al., 2019). Thermal losses toward the environment are neglected and the outlet temperature of the syngas can reach 950°C due to the compact design and direct catalyst heating. Moreover, the unit uses Ni-CeO<sub>x</sub>-MgAl<sub>2</sub>O<sub>4</sub> catalyst instead of the conventional Nickel catalyst, which is active also for dry reforming reactions and to further help with the conversion of CO<sub>2</sub> (Ye et al., 2019).

The gasification unit is an entrained flow gasification reactor (Bader et al., 2018) working with pure oxygen coming from the electrolyzer and steam. A controller keeps the temperature at the outlet of the gasification at 850±1°C by regulating the oxygen demand and thus, adjusting the equivalence ratio. The dynamic variation of the equivalence ratio is made to keep the level of tars in the gasification chamber as constant as possible (the assumption is that tar degradation depends more on temperature than the composition of the feedstock, even if the higher aromatic content in PS affects tar concentration to some extent).

The methanol synthesis loop is done with two shell and tube reactors cooled with boiling feed water and with intermediate sequestration of the product. The kinetics model used

for the reactor is based on Graaf kinetics (Bisotti et al., 2021) which was chosen due to its reliable performance at pressures close to 60 bar (Bisotti et al., 2022).

The distillation section works with a triple-column scheme. The scheme could be simplified since the kinetic scheme does not account for the formation of Dimethyl Ether, higher alcohols, and formaldehyde, but it was intentionally kept like the industrial designs to allow for a more accurate CapEx estimation.

The resulting process layout is a more detailed and improved design with regard to the conventional process layout as previously implemented in Prifti et al. (2021).

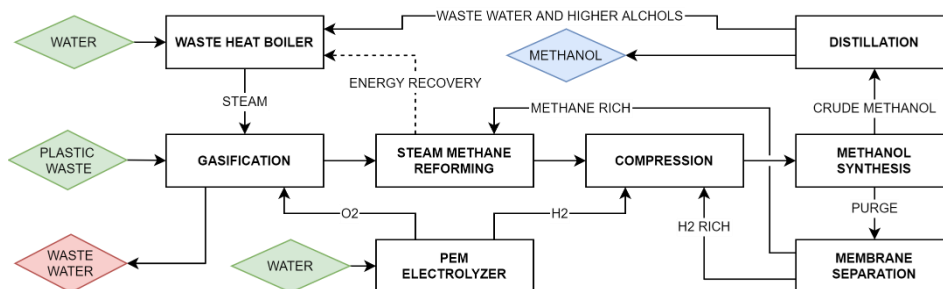


Figure 1: Block flow diagram of the process layout

### 3. Results

The overall input-output plant performance is reported in Table 1. As can be seen from the mass flow rates, the bottleneck of this process is the carbon content in the feedstock. This causes methanol production to rise as the H/C ratio of the feedstock decreases. This behavior can be explained by considering that, from the point of view of the process, there is an unlimited supply of hydrogen thanks to the electrolyzer's presence.

Table 1: Process input and output streams

		48PE32PP20PS	54PE36PP10PS	60PE40PP
		Hydrogen/Carbon ratio		
		1.55	1.73	2
INPUT	Electricity [kWh]	9650	9110	8349
	Plastic Waste [kg/h]	1000	1000	1000
	Demi water [kg/h]	1607	1468	1268
OUTPUT	Oxygen [kg/h]	249	139	-18
	Off gasses [kg/h]	26	26	28
	Wastewater [kg/h]	19	19	19
	Methanol [kg/h]	2316	2285	2238

Electricity shows the highest variability when changing the feedstock composition. Looking at the detailed breakdown of electricity consumption in Figure 2, the PEM electrolyzer is the main consumer of electricity as well as the most sensitive to feed changes. The PEM consumption from the most hydrogenated feedstock to the least hydrogenated one increases by 20%, while compression duty increases by 8% and the reformer one decreases by 8%. The resulting overall electricity consumption is 15.6% higher for the process with the most hydrogenated feedstock.

While the additional electricity consumption translates to higher operative costs, the plant has higher methanol and oxygen production, which could compensate for these higher operative costs. To understand the relative profitability of each feedstock an analysis of the operative margin was made and is reported in Table 2. First, the more hydrogenated the feedstock, the more profitable the process, even with lower methanol yield; second, assuming 8000 hours of operation a year, the plant can generate earnings between 940 k€/y in the best-case scenario, and 260 k€/y in the worst-case scenario, before taxes and without accounting for labor and maintenance. Using the framework described in Prifti et al. (2022) the plant CapEx was estimated at 15.7 M€. With these assumptions, the payback time would range from 16.7 to 60.4 years, being unfeasible in both cases.

Table 2: Process operative margin sensitivity made with EU October 2022 price assumptions.

	€/unit	[(€/h) 48PE32PP20PS]	[(€/h) 54PE36PP10PS]	[(€/h) 60PE40PP]
Methanol [kg/h]	0.58	1343.3	1325.3	1298.0
Electricity [kWh]	0.14	-1351.0	-1275.4	-1168.9
Oxygen [kg/h]	0.2	49.9	27.9	-3.6
Off gasses [kg/h]	0	0	0	0
Waste Water [kg/h]	0	0	0	0
Demi Water [kg/h]	0.006	-9.6	-8.8	-7.6
Net		32.5	68.9	117.9

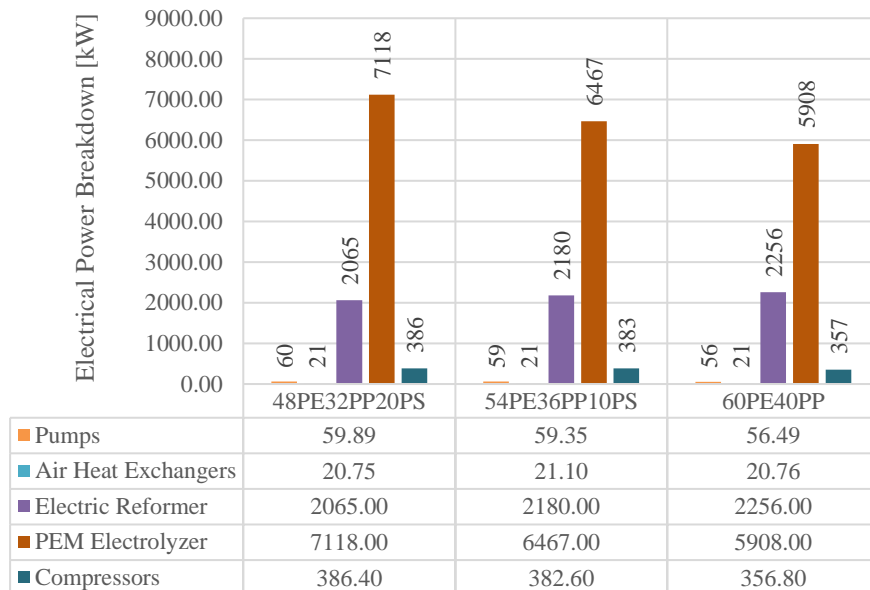


Figure 2: Electrical consumption breakdown by unit operation

Among the revenues, CO<sub>2</sub> credits have been neglected until now. Currently, methanol production has an average footprint of 110 gCO<sub>2</sub>/MJ or converted in mass, 2.21 gCO<sub>2</sub>/gCH<sub>3</sub>OH (Hamelinck and Bunse, 2022). Since the plant does not emit CO<sub>2</sub> and is assumed to consume only renewable energy, for each kg of methanol, 2.21 kg of CO<sub>2</sub> credits are generated which translates into revenues.

As reported in Table 3, including CO<sub>2</sub> credits value at 67 €/t (Europe, October 2022), it considerably increases the process profitability but does not change the trend seen in the sensitivity analysis with the more hydrogenated feedstock still being the most profitable. In conclusion, the process feasibility relies heavily on CO<sub>2</sub> credits and thus the business model depends on subsidies, either local or regional. Therefore, the identification of an appropriate location with political stability and willingness to address the climate challenge is of key importance for the feasibility of the process.

Table 3: Process economics accounting for the CO<sub>2</sub> footprint

	48PE 32PP 20PS	54PE 36PP 10PS	60PE 40PP
CO <sub>2</sub> Credits from Methanol [kg/h]	5118.36	5049.85	4945.98
CO <sub>2</sub> credit revenues [€/h]	342.93	338.34	331.38
October 2022 CO <sub>2</sub> credit value of 67 €/ton			
Annual revenues from CO <sub>2</sub> credits sales [€/y]	2,743,440.96	2,706,719.60	2,651,045.28
Annual revenues by process sales [€/y]	259,984.00	551,616.00	943,744.00
<b>Total revenues [€/y]</b>	<b>3,003,424.96</b>	<b>3,258,335.60</b>	<b>3,594,789.28</b>
<b>Payback time [y]</b>	<b>5.23</b>	<b>4.82</b>	<b>4.37</b>
CO <sub>2</sub> credit value to have PBT of 3.5y [€/t]	103.2	97.4	89.5

Two other aspects that might come into play in the future are the revenues from plastic waste treatment that are intentionally underestimated to account for pre-treatment costs not discussed in the paper, and further subsidies on methanol produced from non-fossil sources. However, both terms require a legislative push towards chemical recycling against incineration and landfill, as well as economic commitments to the circular economy development model from governments.

#### 4. Conclusions

The paper described a possible pathway to methanol synthesis from plastic waste through gasification, maximizing the generation of carbon credits by electrifying key unit operations and exploiting a PEM electrolyzer that generates hydrogen and oxygen to use on site. A sensitivity study on the process feed was carried out and used to obtain mass and energy balances using a detailed simulation. The results of the preliminary economic analysis show the process as unfeasible unless CO<sub>2</sub> credit sales are accounted for. The process becomes economically appealing for the highest price of CO<sub>2</sub> credits. To reach a payback time of 3.5 years, carbon credit values between 89.5 and 103.2 €/t are necessary, even though at current CO<sub>2</sub> credit values of 67 €/t, acceptable payback times lower than 6 years are achieved for every feedstock analyzed.

#### 5. Acknowledgments

Adrian Pacheco-Lopez acknowledges the financial support received from the "Ministerio de Ciencia e Innovación" (grant ref. PRE2018-087135) within the research project CEPI (PID2020-116051RB-I00 granted by the Spanish "MINECO" and the European Regional Development Fund), and the Erasmus+ programme of the EU 2021-27 (KA131 call).

## References

- Bader, A., Hartwich, M., Richter, A., Meyer, B., 2018. Numerical and experimental study of heavy oil gasification in an entrained-flow reactor and the impact of the burner concept. *Fuel Processing Technology* 169, 58–70. <https://doi.org/10.1016/j.fuproc.2017.09.003>
- Bisotti, F., Fedeli, M., Prifti, K., Galeazzi, A., Dell'Angelo, A., Barbieri, M., Pirola, C., Bozzano, G., Manenti, F., 2021. Century of Technology Trends in Methanol Synthesis: Any Need for Kinetics Refitting? *Industrial and Engineering Chemistry Research* 60, 16032–16053. <https://doi.org/10.1021/acs.iecr.1c02877>
- Bisotti, F., Fedeli, M., Prifti, K., Galeazzi, A., Dell'Angelo, A., Manenti, F., 2022. Impact of Kinetic Models on Methanol Synthesis Reactor Predictions: In Silico Assessment and Comparison with Industrial Data. *Industrial and Engineering Chemistry Research* 61, 2206–2226. <https://doi.org/10.1021/acs.iecr.1c04476>
- Carmo, M., Fritz, D.L., Mergel, J., Stolten, D., 2013. A comprehensive review on PEM water electrolysis. *International Journal of Hydrogen Energy* 38, 4901–4934. <https://doi.org/10.1016/j.ijhydene.2013.01.151>
- Galeazzi, A., Prifti, K., Gallo, F., Manenti, F., 2022. A Methodology for The Optimal Surrogate Modelling of Digital Twins Using Machine Learning. *Computer Aided Chemical Engineering* 51, 1543–1548. <https://doi.org/10.1016/B978-0-323-95879-0.50258-7>
- Hamelinck, C., Bunse, M., 2022. Carbon footprint of Methanol [WWW Document]. Methanol Institute. URL <https://www.methanol.org/> (accessed 10.19.22).
- Pacheco-López, A., Somoza-Tornos, A., Graells, M., Espuña, A., 2021. Synthesis and assessment of waste-to-resource routes for circular economy. *Computers & Chemical Engineering* 153, 107439. <https://doi.org/10.1016/j.compchemeng.2021.107439>
- Pacheco-López, A., Somoza-Tornos, A., Muñoz, E., Capón-García, E., Graells, M., Espuña, A., 2020. Synthesis and Assessment of Waste-to-resource Routes for Circular Economy, in: Pierucci, S., Manenti, F., Bozzano, G.L., Manca, D. (Eds.), *Computer Aided Chemical Engineering, 30 European Symposium on Computer Aided Process Engineering*. Elsevier, pp. 1933–1938. <https://doi.org/10.1016/B978-0-12-823377-1.50323-2>
- Prifti, K., Galeazzi, A., Barbieri, M., Manenti, F., 2022. A Capex Opex Simultaneous Robust Optimizer: Process Simulation-based Generalized Framework for Reliable Economic Estimations. *Computer Aided Chemical Engineering* 51, 1321–1326. <https://doi.org/10.1016/B978-0-323-95879-0.50221-6>
- Prifti, Kristiano, Galeazzi, A., Margarita, I., Papale, A., Miele, S., Bargiacchi, E., Barbieri, M., Petea, M., Manenti, F., 2021. CONVERTING END-OF-LIFE PLASTIC WASTE INTO METHANOL: THE GASIFORMING PROCESS AS NEW, EFFICIENT AND CIRCULAR PATHWAY. *Environmental Engineering and Management Journal* 20, 1629–1636. <https://doi.org/10.30638/eemj.2021.151>
- Prifti, K., Galeazzi, A., Margarita, I., Papale, A., Miele, S., Bargiacchi, E., Barbieri, M., Petea, M., Manenti, F., 2021. Converting end-of-life plastic waste into methanol: The gasiforming™ process as new, efficient and circular pathway. *Environmental Engineering and Management Journal* 20, 1629–1636. <https://doi.org/10.30638/eemj.2021.151>
- Qureshi, M.S., Oasmaa, A., Pihkola, H., Deviatkin, I., Tenhunen, A., Mannila, J., Minkinen, H., Pohjakallio, M., Laine-Ylijoki, J., 2020. Pyrolysis of plastic waste: Opportunities and challenges. *Journal of Analytical and Applied Pyrolysis* 152, 104804. <https://doi.org/10.1016/j.jaap.2020.104804>
- Schyns, Z.O.G., Shaver, M.P., 2021. Mechanical Recycling of Packaging Plastics: A Review. *Macromol Rapid Commun* 42, e2000415. <https://doi.org/10.1002/marc.202000415>
- Wismann, S.T., Engbæk, J.S., Vendelbo, S.B., Bendixen, F.B., Eriksen, W.L., Aasberg-Petersen, K., Frandsen, C., Chorkendorff, I., Mortensen, P.M., 2019. Electrified methane reforming: A compact approach to greener industrial hydrogen production. *Science* 364, 756–759. <https://doi.org/10.1126/science.aaw8775>
- Ye, R.-P., Ding, J., Gong, W., Argyle, M.D., Zhong, Q., Wang, Y., Russell, C.K., Xu, Z., Russell, A.G., Li, Q., Fan, M., Yao, Y.-G., 2019. CO<sub>2</sub> hydrogenation to high-value products via heterogeneous catalysis. *NatCom* 10, 5698. <https://doi.org/10.1038/s41467-019-13638-9>

# Deep neural network-based multi-objective optimization of NO<sub>x</sub> emission and profit by recovering lignocellulosic biomass

Kim Y.,<sup>a,b§</sup> Park J.,<sup>a,b§</sup> Lim J.,<sup>a,b</sup> Joo C.,<sup>a,b</sup> Cho H.,<sup>a</sup> Kim J.<sup>a,b\*</sup>

<sup>a</sup>*Green Materials and Processes R&D Group, Korea Institute of Industrial Technology, 55, Janga-ro, Ulsan, 44413, Korea*

<sup>b</sup>*Department of Chemical and Biomolecular Engineering, Yonsei University, 50, Yonsei-ro, Seoul, 03722, Korea*

<sup>§</sup>*Kim Y. and Park J. contributed equally to this work as first authors.*

<sup>\*</sup>*kjh31@kitech.re.kr*

## Abstract

In the pulp and paper industry, the external energy and pulping chemical consumption have been reduced by recovering the lignocellulosic biomass (LB) produced during the pulping process. However, this involves the inevitable emission of thermal NO<sub>x</sub> owing to the high pyrolysis reaction temperature. Therefore, it is necessary to simultaneously optimize energy and pulping chemical recovery and minimize NO<sub>x</sub> emissions. Hence, this study focuses on the multi-objective optimization of maximizing the net profit from energy and pulping chemical recovery while minimizing NO<sub>x</sub> emissions in recovering LB. For multi-objective optimization, a deep neural network (DNN)-based optimization model for the net profit and NO<sub>x</sub> emissions was developed with the 1,071 simulation data points according to the operating conditions. Consequently, Pareto-optimal solutions with profits between 5,241,520 and 1,329,558 \$/y and NO<sub>x</sub> emissions between 87.95 and 78.27 ppm were obtained. The proposed Pareto-optimal front can offer comprehensive solutions to decision-makers.

**Keywords:** Lignocellulosic biomass, Profit, NO<sub>x</sub> emission, Multi-objective optimization, Deep neural network

## 1. Introduction

The pulp and paper industry is energy-intensive, involving high operating costs because of the large amounts of chemicals used during the pulping process. Therefore, lignocellulosic biomass (LB) is recovered to reduce energy and pulping chemical consumption. LB is produced during the pulping process and is composed of lignin, inorganic compounds, and 80–85% water. The energy consumption in the pulp and paper industry is reduced by reusing lignin in LB as a primary energy source. Further, the consumption of pulping chemicals is reduced by recovering the pulping chemicals converted from the inorganic components in the LB via pyrolysis.

The recovery of energy and pulping chemicals is increased by decreasing the moisture content in LB via evaporation, followed by efficient combustion with multi-level injected air through the biomass boiler. Currently, to increase energy and pulping chemical recovery with LB reuse, research is actively conducted to reduce energy consumption by modifying the configuration of the evaporation process (Kim et al., 2021), or increase



energy and chemical recovery by retrofitting the internal structure of the biomass boiler to improve combustion performance (Pérez et al., 2016).

However, increasing the efficiency of LB combustion leads to higher emission of thermal NO<sub>x</sub> owing to the high temperature during the pyrolysis reaction. Therefore, research has been conducted on reducing thermal NO<sub>x</sub> by combining an additional system with a biomass boiler (Tu et al., 2018).

Despite the considerable contributions of previous studies, some challenges remain to be solved. First, previous studies have focused on retrofitting the conventional evaporation process and biomass boiler (Kim et al., 2021; Pérez et al., 2016; Tu et al., 2018). The presented solutions have limitations in that on-site application to the existing pulp and paper industry is difficult due to structural changes or additional equipment installation. Second, previous studies have not simultaneously considered minimizing the energy consumption of the evaporation process, maximizing energy and pulping chemical recovery, and minimizing NO<sub>x</sub> emissions in the biomass boiler (Kim et al., 2021; Pérez et al., 2016; Tu et al., 2018). Removal of higher amounts of moisture in LB during evaporation leads to increased energy recovery through LB reuse; however, it also results in increased energy consumption during evaporation. In addition, more efficient LB combustion in a biomass boiler results in increased energy and pulping chemical recovery; however, it also leads to higher thermal NO<sub>x</sub> emissions. Therefore, it is necessary to simultaneously focus on minimizing the energy consumption during the evaporation process, maximizing energy and pulping chemical recovery, and minimizing NO<sub>x</sub> emissions in the biomass boiler.

To overcome these challenges, this study focused on the deep neural network (DNN)-based multi-objective optimization of NO<sub>x</sub> emission, and net profit from energy and pulping chemicals recovery by reusing LB. This study aimed to maximize net profit and minimize NO<sub>x</sub> emissions in an applicable on-site way by simultaneously considering the evaporation process and biomass boiler operating conditions. To this end, simulation models of the evaporation process and biomass boiler were developed to derive NO<sub>x</sub> emissions and net profit according to the operating conditions. Then, a DNN-based optimization model for the net profit and NO<sub>x</sub> emissions was developed with the 1,071 simulation data points to derive the optimal operating conditions.

## **2. Methodology**

### *2.1. Model development*

#### *2.1.1. Process model*

The simulation conducted in this study involved an evaporation process, followed by a biomass boiler. The evaporation process was simulated using Aspen Plus V11.0, ELECNRTL method, and industrial data. Lignocellulosic biomass (LB) with a concentration of 17% entered an intermediate evaporator and flowed partially forward and backward. Using this configuration, LB was efficiently concentrated to 69–85%. The higher the target LB concentration, the more energy sources are required for the evaporation process to remove more moisture. A schematic of the evaporation process is shown in Fig. 1.

The biomass boiler was simulated using Aspen Plus V11.0, Peng Robinson–Boston Mathias method, and industrial data. Tertiary-level air and concentrated LB were directed into the biomass boiler. Subsequently, dehydration, pyrolysis, gasification, reduction, reoxidation, and combustion reactions occurred. First, the concentrated LB was sprayed in the furnace in the biomass boiler, and the moisture in the LB was completely evaporated via dehydration. Second, the dried LB in the solid state dropped toward the bottom of the furnace and was combusted with secondary air. Lignin in the dried LB was

decomposed into fixed carbon and volatiles via pyrolysis. Thermal NO<sub>x</sub> was also produced owing to the high pyrolysis reaction temperature. The volatiles (including thermal NO<sub>x</sub>) moved to the upper part of the furnace, and the fixed carbon and remaining inorganics in the LB were stacked in the form of a char bed at the furnace bottom. Third, the fixed carbon in the char bed was combusted with primary air and the energy required for the reduction reaction was produced. In the char bed, Na<sub>2</sub>SO<sub>4</sub> in the inorganic components was converted to Na<sub>2</sub>S via gasification with the remaining fixed carbon. Partial Na<sub>2</sub>S was converted to Na<sub>2</sub>SO<sub>4</sub> via reoxidation with the O<sub>2</sub> present in the primary air. Furthermore, a part of the fixed carbon was converted to CO<sub>2</sub> via gasification, and the converted CO<sub>2</sub> reacted with Na (the inorganic component in LB), forming Na<sub>2</sub>CO<sub>3</sub>. After the reaction, inert components such as Na<sub>2</sub>S, Na<sub>2</sub>SO<sub>4</sub>, and Na<sub>2</sub>CO<sub>3</sub> exited the furnace bottom in the form of smelt. Subsequently, Na<sub>2</sub>CO<sub>3</sub> in the smelt was converted to NaOH by additional treatment, and Na<sub>2</sub>S from the smelt and NaOH were recovered as pulping chemicals. Finally, all the volatiles moved to the upper part of the furnace and combusted with the tertiary air. In this combustion reaction, polluting gases such as H<sub>2</sub>S, SO<sub>2</sub>, CO, and NO<sub>2</sub> were reduced. The final flue gas exited the furnace, and steam was produced in the upper region (economizer, boiler bank, and superheaters) of the biomass boiler. A schematic of the biomass boiler is shown in Fig. 2.

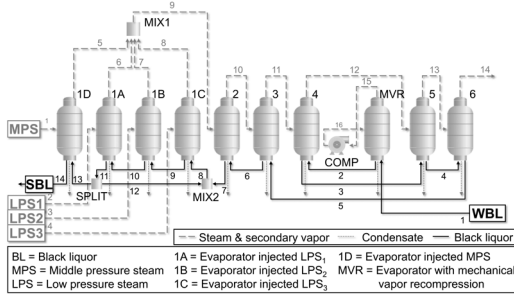


Fig. 1. Schematic of the evaporation process

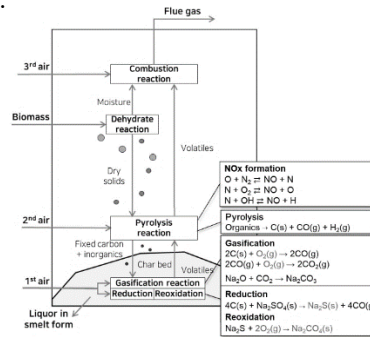


Fig. 2. Schematic of the biomass boiler

The main operating parameters that affect the energy, pulping chemical recovery, and NO<sub>x</sub> emissions are the LB concentration and tertiary-level injected air distribution. The net profit by reusing the LB according to the operating conditions was calculated using Eqs. (1)–(4). The net profit was calculated by subtracting the energy consumption cost from the sum of the profit by recovering the energy (e.g., steam) and pulping chemicals, as shown in Eq. (1). The profit from steam recovery was calculated by multiplying the mass flow and specific enthalpy of the flue gas with the steam price, as shown in Eq. (2). The profit obtained by recovering pulping chemicals was calculated by multiplying the mass flow with the prices of Na<sub>2</sub>S and NaOH, the main components of the pulping chemical, as shown in Eq. (3). Here, NaOH could be recovered only through additional treatment of Na<sub>2</sub>CO<sub>3</sub>; therefore, the NaOH profit was additionally considered in terms of Na<sub>2</sub>CO<sub>3</sub> conversion and NaOH yield during the treatment. The steam consumption cost was calculated by multiplying the price of each type of steam consumed in the evaporation process with the consumed steam mass flow, as shown in Eq. (4).

$$Net\ profit = P^S + P^{PC} - C^S, \quad (1)$$

$$P^S = M^{flue\ gas} \times E^{flue\ gas} \times Price^S, \quad (2)$$

$$P^{PC} = Price^{Na_2S} \times M^{Na_2S} + Price^{NaOH} \times \alpha \times M^{Na_2CO_3}, \quad (3)$$

$$C^S = \sum_{i=1} C_i^S \times M_i^S. \quad (4)$$

### 2.1.2. DNN-based optimization model

Using the limited datasets derived from the process model makes it easy to fall into the local optimum problem. However, the DNN model enables global optimization by efficiently generating numerous datasets. Therefore, this study developed a DNN model for robust multi-objective optimization to maximize net profit while minimizing NO<sub>x</sub> emissions. First, the initial net profit and NO<sub>x</sub> emission datasets were generated through the process model. Subsequently, the DNN model was trained using these datasets. Among these datasets, 60% were used for training, 20% for validation, and 20% for testing. The input variables of the developed DNN model were the concentration of the LB exiting the evaporation process and tertiary-level injected air distribution ratio. The output variables of the model were net profit and NO<sub>x</sub> emissions, and the model consisted of 3 hidden layers and 20 hidden nodes. ReLu and R<sup>2</sup> were used as the activation function and evaluation index of the developed DNN model, respectively.

### 2.2. Multi-objective optimization

This section describes the multi-objective optimization methodology. There were several constraints in solving the multi-objective optimization problem of minimizing NO<sub>x</sub> while maximizing the net profit. First, the NO<sub>x</sub> emissions should not be higher than those of the existing operation. Second, the net profit should not be lower than the existing values. Third, the concentration of LB exiting the evaporation process was 69–85%. Finally, the sum of the distributed air was 100%. Multi-objective optimization was conducted using the datasets from the process and DNN models. Finally, this study proposed Pareto-optimal solutions that satisfy the specified constraints.

## 3. Results and discussions

### 3.1. Model validation

Conventionally, LB is concentrated to 70% during evaporation, and combusted with the distributed tertiary-level air at ratios of 24.89%, 49.06%, and 26.04% in the biomass boiler. The simulation results for the evaporation process and biomass boiler were validated using industrial data. The performance of the models showed little error compared to the industrial data ( $\leq 0.24\%$  and  $\leq 4.55\%$ , respectively). Consequently, the operating conditions of the evaporation process consumed 368,157 t/y of steam, 1,629,187 t/y of steam and 591,542 and 1,789,136 t/y of Na<sub>2</sub>CO<sub>3</sub> and Na<sub>2</sub>S, respectively, for converting to pulping chemicals were recovered by the existing biomass boiler. Therefore, the net profit was 1,329,641 \$/y in the existing operation condition: 1,521,825 \$/y of energy consumption cost, and 5,518,458 and 6,635,017 \$/y by recovering energy and pulping chemicals, respectively. Furthermore, the NO<sub>x</sub> emissions from the existing operation were 87.95 ppm. The DNN model was trained using 1,091 data samples for profit and NO<sub>x</sub> emissions according to the operating conditions obtained by the process model. Consequently, the developed DNN model was confirmed to have high accuracy, with an R<sup>2</sup> of 0.982 for profit and 0.981 for NO<sub>x</sub> emission.

### 3.2. Optimization of NO<sub>x</sub> emissions and net profit

This section first describes the changes in energy consumption, energy and pulping chemical recovery, and NO<sub>x</sub> emissions according to the operating conditions: the concentration of LB exiting the evaporator and tertiary-level injected air distribution ratio. Then, representative solutions are presented along with the results of multi-objective optimization for maximizing the net profit while minimizing NO<sub>x</sub> emissions.

The concentration of LB exiting the evaporator was changed to 69–85%, and the resulting steam consumption is shown in Fig. 3. The evaporation process consumed steam of 367,462 t/y at 69%, 368,157 t/y at 70%, and 387,810 t/y at 85%. The steam consumption

of the evaporation process increased as the LB concentration increased, because the higher the target LB concentration, the more energy sources were required to remove the higher amount of moisture.

The effects of the LB concentration and tertiary-level injected air distribution ratio on steam, pulping chemical recovery, and  $NO_x$  emissions are shown in Fig. 4 (a). The flue gas enthalpy, which is directly related to steam recovery, increased when the LB concentration and secondary air flow ratio increased. The higher the LB concentration, the less energy that is lost to remove moisture from the LB. In addition, as the secondary air flow ratio increases, more volatiles can be introduced into the flue gas by the activated pyrolysis reaction.

The  $Na_2S$  recovery for pulping chemicals increases as the secondary air flow ratio increases, and the primary air flow rate approaches 14.74%, as shown in Fig. 4 (b). As the secondary air flow ratio increases, the pyrolysis reaction actively occurs so that more substances for  $Na_2S$  conversion (C, CO) are decomposed from the LB. As the primary air flow ratio increases, more energy can be generated during gasification. However, it has an optimum point of 14.74% because it also promotes the reoxidation reaction of  $Na_2S$ .

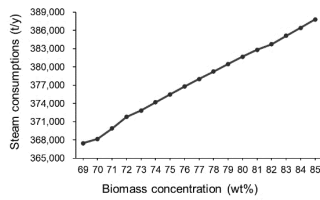


Fig. 3. Steam consumption according to LB concentration

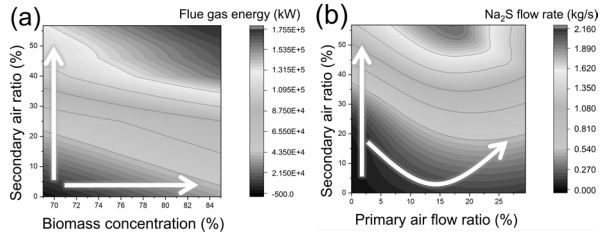


Fig. 4. Steam and pulping chemical recovery according to operating conditions: (a) steam; (b)  $Na_2S$

The  $Na_2CO_3$  recovery for the pulping chemicals increases as the primary air flow ratio and biomass concentration increases, as shown in Fig. 5 (a). As the primary air flow ratio increases,  $CO_2$  generation increases because of the activated gasification reaction. The  $Na_2CO_3$  conversion is promoted by increased  $CO_2$  generation. In addition, as the biomass concentration increases, the energy used for the  $Na_2CO_3$  reaction increases by decreasing the energy lost to moisture removal from LB.

Finally, the  $NO_x$  emissions decrease as the secondary air flow ratio decreases and the tertiary air flow ratio increases, as shown in Fig. 5 (b). As the secondary air flow ratio increases, the pyrolysis reaction becomes more active; accordingly, the amount of thermal  $NO_x$  production increases with increasing temperature of the pyrolysis reaction. As the tertiary air flow ratio increases,  $NO_x$  emissions decrease owing to the activated  $NO_x$ -reducing combustion reactions.

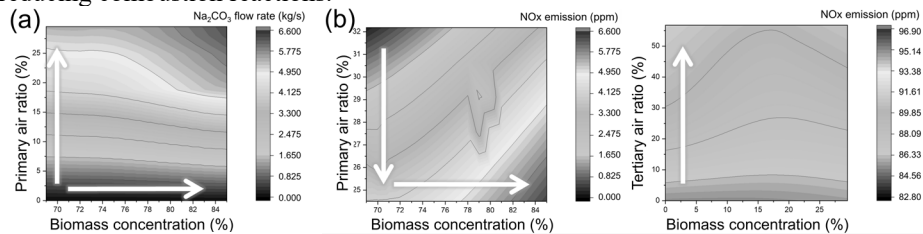


Fig. 5. Pulping chemical recovery and  $NO_x$  emissions according to operating conditions: (a)  $Na_2CO_3$ ; (b)  $NO_x$  emissions

$NO_x$  emissions were derived. Fig. 6 shows the Pareto-optimal front with representative solutions. Representative solution 1 has  $NO_x$  emissions equal to the existing operating conditions but can increase profits by up to 294.23%. Representative solution 3 has the

same net profit as before but can reduce NOx emissions by up to 11.01%. Finally, Solution 2 is the average of Solutions 1 and 3, which can increase the net profit by 147.41% and reduce NOx emissions by 5.50%. In addition to these three solutions, decision-makers can select solutions that fit the situation from the proposed Pareto-optimal front.

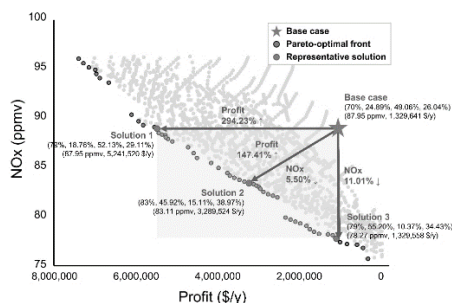


Fig. 6. Pareto-optimal front with representative solutions

#### 4. Conclusions

This study focuses on multi-objective optimization of maximizing net profit and minimizing NOx emissions while recovering LB. This study makes two major contributions to the existing literature. First, the solutions proposed in this study do not include any retrofitting of the conventional evaporation process and the biomass boiler, and are thus applicable on-site. Second, robust multi-objective optimization was realized by simultaneously considering the net profit and NOx emissions in the overall processes in the pulp and paper industry using the DNN model. The derived representative solutions increased the net profit by up to 294.23% and decreased the NOx emissions by up to 11.01%. In addition to representative solutions, decision-makers can select solutions that fit the situation from the proposed Pareto-optimal front. Therefore, we believe this study can offer comprehensive solutions to on-site decision-makers.

#### Acknowledgments

This work was supported by the Korea Institute of Industrial Technology within the framework of the “Development and application of carbon-neutral engineering platform based on carbon emission database and prediction model” project [Grant Nos. KM-23-0098] and “Development of complex parameter smart analysis modules for color customizing” project [Grant Nos. EH-23-0005].

#### References

- Kim Y, Lim J, Cho H, Kim J. Novel mechanical vapor recompression-assisted evaporation process for improving energy efficiency in pulp and paper industry, *Int J Energy Res*, 46, 3, 3409-3427
- Pérez MG, Vakkilainen E, Hyppänen T, Fouling growth modeling of kraft recovery boiler fume ash deposits with dynamic meshes and a mechanistic sticking approach, *Fuel*, 185, 872–85.
- Tu Y, Zhou A, Xu M, Yang W, Siah KB, Subbaiah P, NOx reduction in a 40 t/h biomass fired grate boiler using internal flue gas recirculation technology, *Appl Energy*, 220, 962–973.
- Lim J, Kim J, Optimizing ash deposit removal system to maximize biomass recycling as renewable energy for CO<sub>2</sub> reduction, *Renew Energy*, 190, 1006–1017.
- Lim J, Cho H, Kwon H, Park H, Kim J, Reinforcement learning-based optimal operation of ash deposit removal system to improve recycling efficiency of biomass for CO<sub>2</sub> reduction, *J. Clean. Prod.*, 370, 133605.

Park J, Kim Y, Lim J, Cho H, Kim J, Optimal operation of the evaporator and combustion air distribution system in a pulp mill to maximize biomass recycling and energy efficiency, J. Clean. Prod., 367, 133048.



# Decarbonization of gray ammonia production through the Allam power cycle: a techno-economic assessment and optimization

Andrea Isella, Mattia Forina, Mario Spagna, Davide Manca\*

*PSE-Lab, Process Systems Engineering Laboratory, Dipartimento di Chimica, Materiali e Ingegneria Chimica “Giulio Natta”, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy*

*\*davide.manca@polimi.itw*

## Abstract

Ammonia is one of the leading products of the chemical industry worldwide. In addition, its manufacturing is significantly both energy- and emissions-intensive. As ammonia production is heavily based on fossil fuels, its synthesis integration with the Allam power cycle is a promising technology that can lead to substantial decarbonization of the whole process and an important abatement of the electrical energy costs (and indirect emissions). The Allam cycle is a system for high-efficiency power generation with supercritical CO<sub>2</sub> working fluid. Its integration into conventional ammonia plants allows for simultaneously producing electricity, water, and pipeline-ready carbon dioxide. An original, improved layout of a gray ammonia plant combined with the Allam power cycle is simulated to run a techno-economic assessment of the process and detect the optimal operating conditions that improve its economic and environmental sustainability.

**Keywords:** Haber-Bosch; carbon dioxide; sustainability; climate change mitigation.

## 1. Introduction

The rising energy transition era calls for the decarbonization of hard-to-abate sectors, with the chemical industry at the forefront. Specifically, in 2020 only, ammonia synthesis totaled more than 440 Mt<sub>CO<sub>2</sub>eq</sub>/y, being the most carbon-intensive process of the whole sector (Isella and Manca, 2022). Considering a conventional ammonia plant, the core of its direct CO<sub>2</sub> emissions is the so-called “front-end”. Such a term, which refers to the reforming, shift, and syngas purification sections as a whole, is opposed to the “back-end”, which refers to the ammonia converters and the separation section that allows the final product recovery. This is because conventional front-ends heavily rely on fossil fuels (primarily natural gas) to meet both the hydrogen and energy demands of the process. In answer to this, the integration of a recently introduced power cycle known as the “Allam cycle” in the front-end of a conventional (“gray”) ammonia plant is here proposed and performed through process simulations and techno-economic evaluations. Indeed, since the main peculiarity of the Allam cycle lies in exploiting supercritical carbon dioxide as a working fluid for electricity production (Allam *et al.*, 2015), the present work investigates how ammonia production may be decarbonized by diverting the whole CO<sub>2</sub> produced in the front-end of the production site to the Allam cycle. Such a strategy would lead to producing a sort of “blue” ammonia (to borrow the hydrogen color codes) and



generating electric power (which would fulfill first the electricity demand of the whole plant and then export the surplus).

## 2. Process simulation

The simulations of this study were performed by UniSim® Design R491. Specifically, three distinct sections constitute the plant: the Front-end, the Air separation unit, and the Allam cycle. First, these sections were singularly modeled and finally joined into the whole process flow diagram reported in Figure 1.

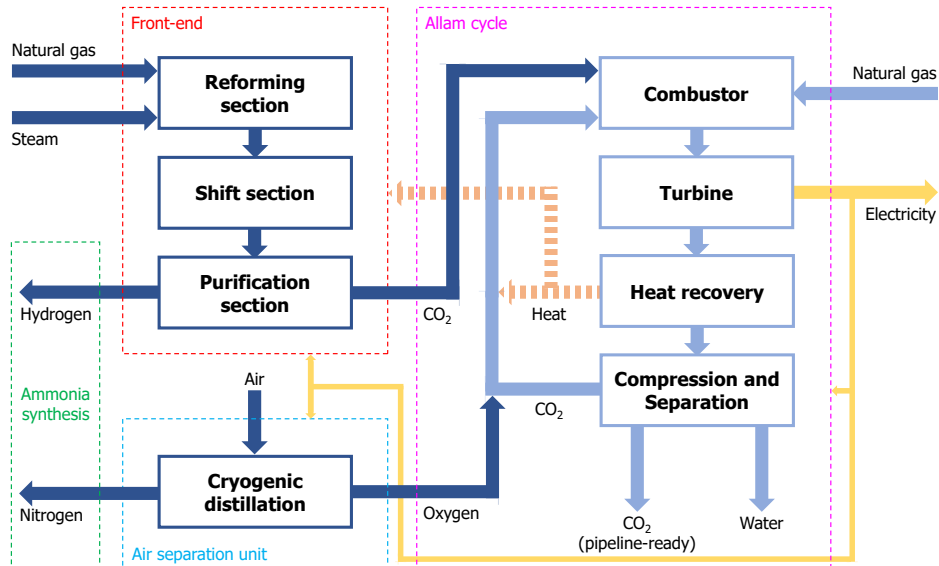


Figure 1: Block flow diagram of the simulated ammonia plant with the Allam cycle integration.

As evident from Figure 1, the main goal of the proposed integration is to fully abate the gaseous  $\text{CO}_2$  emissions generated in the Front-end by routing all the tail gases (separated in the purification section from the hydrogen product) to the Allam cycle power plant. Such a  $\text{CO}_2$ -rich stream feeds the combustor of the Allam cycle. Specifically, this unit burns natural gas and oxygen to increase further the temperature above  $1100\text{ }^\circ\text{C}$  and produce even more  $\text{CO}_2$  that feeds a turbine for electricity production. Next, a series of heat exchangers cool the exhaust stream and recover a significant amount of heat. One of them, the recuperator, has the crucial function of pre-heating the recycle stream before it returns to the combustor. After water removal, the  $\text{CO}_2$  mass flow is compressed through a train of centrifugal compressors that make the expanded  $\text{CO}_2$  supercritical again: a minor fraction is removed as pipeline-ready carbon dioxide, and the remaining part is recycled back to the combustor. It is worth noting that the Allam cycle contributes to an increase in the total amount of carbon dioxide produced by the whole process. However, unlike the gaseous emissions from the ammonia plant, the  $\text{CO}_2$  product is collected in such a state that easily allows exporting it for industrial uses (e.g., urea synthesis, enhanced oil recovery, food market). Indeed, the usage of carbon dioxide in the industrial sector records remarkable extents: around  $230\text{ Mt}_{\text{CO}_2}/\text{y}$  are currently used for the said purposes (IEA, 2022). Even water is among the products of the Allam cycle. Concerning the current integration layout, it can be used in the ammonia section either as a utility for the on-site steam generation required by the reformers or as cooling water. Finally, an air

separation unit (ASU) that concurrently produces the nitrogen required for ammonia synthesis meets the oxygen demand of the Allam cycle combustor. Indeed, a nitrogen-to-hydrogen ratio equal to 3 is mandatory at the inlet of the ammonia converters. The following subsections show each simulated section of the integrated plant.

### 2.1. Front-end

The Front-end model primarily comes from Soltani *et al.* (2014). It is a simplified model of a typical steam methane reforming (SMR) process, comprehensive of the furnace, the reformer and shift reactors, and the purification section.

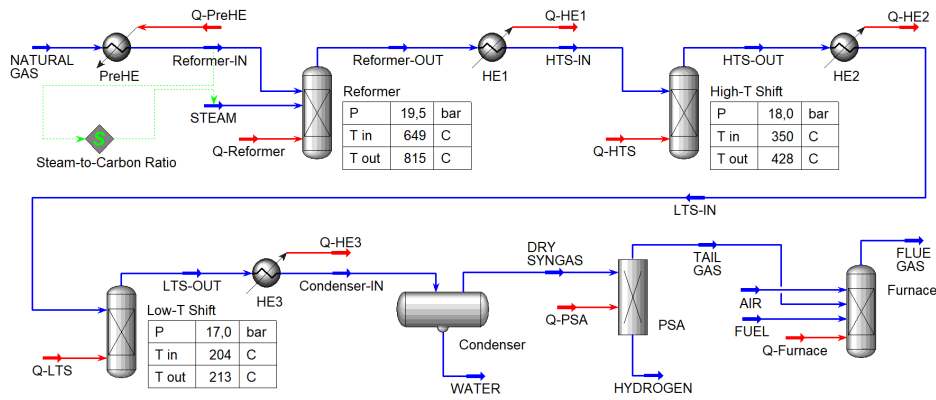


Figure 2: The simulated Front-end process flow diagram.

Precisely, all the reactors were simulated as “Equilibrium reactors” (*i.e.* reactors whose outlet streams achieve both chemical and physical equilibria according to the specified reaction set). In contrast, the furnace was modeled as a “Gibbs reactor” (*i.e.* a reactor that simply minimizes the Gibbs free energy of its products, suited to avoid complicated and inadequately accurate stoichiometric reactor models). Note that the furnace appears at the end of the flowsheet because it downstream evaluates the amount of fuel it needs. In this respect, the flue gases from the furnace virtually enter the reformer to exchange their enthalpy with the reacting flow.

### 2.2. Air separation unit

The Air separation model comes from Ebrahimi *et al.* (2015). Expressly, it consists of a conventional double-column cryogenic distillation process.

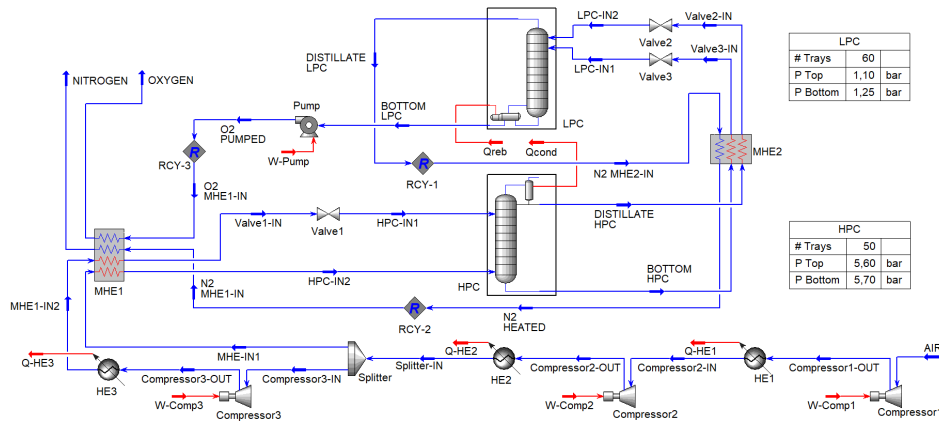


Figure 3: The simulated Air separation process flow diagram.

Concerning product purities, the nitrogen stream should preferably have purity above 99 mol% to avoid oxygen poisoning the catalyst of the ammonia synthesis loop. In contrast, the oxygen stream headed to the Allam cycle should be over 95 mol%, *i.e.* the near-optimal oxygen purity value for oxy-fuel combustion (Ye *et al.*, 2019).

### 2.3. Allam cycle

The Allam cycle model mainly relies on Allam *et al.* (2015) and Wang *et al.* (2021). The inlet tail gas and oxygen streams come from the Front-end and the Air separation unit sections, respectively.

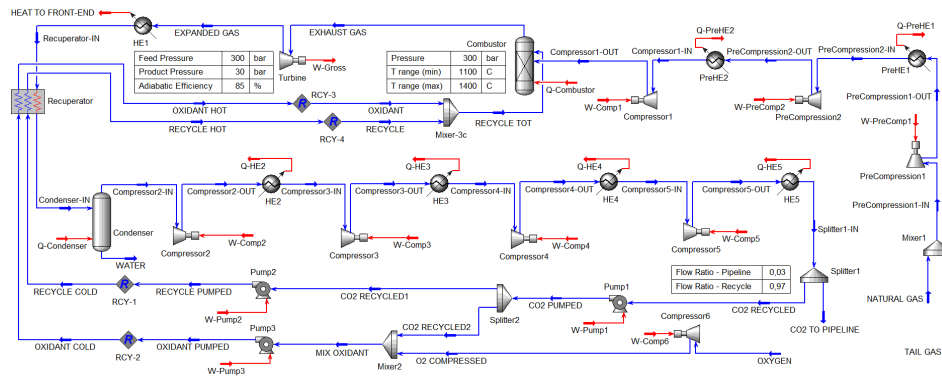


Figure 4: The simulated Allam cycle process flow diagram.

As happened for the furnace of the Front-end, even the combustor was modeled as a Gibbs reactor. Specifically, its temperature operability range spans from 1100 to 1400 °C. Also, note that an additional heat exchanger (missing in the original layout of the Allam cycle patent) was added to recover heat beneficial to the Front-end reforming section, as proposed by Wang *et al.* (2021). The CO<sub>2</sub> purge, ready for the pipeline, equals 3% of the processed flow, while the remaining 97% is recycled back to the combustor.

All the simulations implement the Peng-Robinson equation of state for thermodynamic computations. Finally, the three sections were combined in a single flowsheet to assess the performance of the retrofitted ammonia synthesis plant.

### 3. Results and discussion

Both CapEx (Capital Expenditure) and OpEx (Operational Expenditure) were evaluated for the economic assessment of the whole plant. The main hypotheses for such calculations consist in considering 8000 working hours in a year and a 10-year depreciation time for the equipment in each section of the plant. Regarding the Front-end, according to the rather simplified model, the CapEx and the OpEx terms rely on the indicative values of 390 USD/t<sub>H2</sub> and 970 USD/t<sub>H2</sub> (IEA, 2019), respectively. Conversely, concerning the Air separation unit, the CapEx estimation mainly derives from Guthrie’s formulae (for columns and compressors) and the equipment costs reported by Young *et al.* (2021) for the two multi-pass heat exchangers provided in the process flow diagram. Finally, the costs of the Allam cycle originate from an indicative value of 685.7 USD/kW<sub>e,output</sub> (IEAGHG, 2015) for the CapEx, while market values of utilities (*i.e.* natural gas, 0.08 USD/kWh in the 1<sup>st</sup> semester of 2022) and products (*i.e.* electricity, 0.25 USD/kWh in the 1<sup>st</sup> semester of 2022, and ready-to-pipeline carbon dioxide, 10 USD/t<sub>CO2</sub>) for the OpEx. Four ammonia plant capacities were considered: 750, 1000, 2000, and 3000 t<sub>NH3</sub>/d. In this regard, a “0.8 power law” replaced the conventional “0.6 power law” for the economies of scale to run a conservatively safer economic assessment. Indeed, both the combustor and the turbine of the Allam cycle are rather peculiar units (especially as they have to deal with CO<sub>2</sub>-rich streams), thus likely less liable to such a typical behavior of the conventional units of chemical plants.

Table 1: Main results of the techno-economic assessment. Key process variables, CapEx and OpEx, and the percentage cost increases ( $\Delta\%$ ) of the performed Allam cycle integration as a function of the ammonia plant capacity. A combustor temperature equal to 1100 °C is considered.

Plant capacity [t <sub>NH3</sub> /d]	CO <sub>2</sub> to pipeline [t <sub>CO2</sub> /d]	Water output [t <sub>H2O</sub> /d]	Net power [MWe]	Max pipe diameter [m]	CapEx [MUSD]	OpEx [MUSD/y]	$\Delta\%$ CapEx [-]	$\Delta\%$ OpEx [-]
750	1426	664	147	0.58	510	82.6	+82.4%	-34.5%
1000	1902	885	205	0.67	627	100.5	+85.9%	-28.4%
2000	3803	1770	438	0.95	1073	208.7	+88.8%	+5.4%
3000	5704	2655	670	1.16	1497	351.8	+87.4%	+37.7%

Table 1 shows the main process variables and the results of the CapEx and OpEx assessment performed on the integrated facilities. Specifically, increasing the operative temperature of the combustor does not show any beneficial effect on both the investment and operative costs of the plant. Thus, the table reports only the results corresponding to the simulated lowest feasible operative temperature at the combustor (*i.e.* 1100 °C). Nevertheless, as evident from Table 1, the integration of the Allam cycle to gray ammonia plants becomes increasingly unprofitable as the ammonia plant capacity increases. Indeed, although higher ammonia plant capacities allow producing considerable amounts of electric power, they also lead to extremely high flowrates in the Allam cycle loop (note that the CO<sub>2</sub>-to-pipeline stream is just 3% of the total processed flow). In other words, this means the power cycle needs high piping diameters to process all the flowing mass within the loop. The maximum internal diameter of a standard stainless-steel tube (material needed to resist the corrosivity of the CO<sub>2</sub>-rich stream: indeed, in presence of water, severe corrosion of the infrastructure may result due to the formation of carbonic

acid) is around 0.8 m: thus, both 2000 and 3000  $t_{\text{NH}_3}/\text{d}$  plants would inevitably require at least a parallel configuration. As previously mentioned, the water produced by the Allam cycle could be exploited to generate steam as required by the reforming section: specifically, at each ammonia plant capacity scenario, it would reduce about 40 wt% of its demand. Finally, the oxygen demand of the Allam cycle is another crucial factor making the integration disadvantageous at high capacities. Indeed, an air separation unit producing about 1900  $t_{\text{O}_2}/\text{d}$  was simulated to deal with the nitrogen demand of the 750 and 1000  $t_{\text{NH}_3}/\text{d}$  ammonia plants but, in the case of 2000 and 3000  $t_{\text{NH}_3}/\text{d}$  facilities, such oxygen productivity just covers 25% and 50% of the final demand of the power cycle, respectively.

#### 4. Conclusions

The integration of the Allam power plant into an ammonia synthesis process proved to be a feasible strategy to decarbonize ammonia production. The results of the presented techno-economic assessment showed that such a solution primarily suits low/medium-capacity ammonia plants. Indeed, concerning the 750 and 1000  $t_{\text{NH}_3}/\text{d}$  capacities, substantial reductions in OpEx complement the inevitable increases in CapEx. On the other hand, medium/high-capacities would suffer greatly from the extremely high flowrates circulating in the Allam cycle, although this would make it capable to export much more electricity. In addition, such higher capacities would call for further measures such as parallel configurations and additional raw materials and utilities entering the system. However, if one disregards any economic considerations, it is rather important to observe that every integrated layout was able to simultaneously prevent all the gaseous emissions originating from the front-end section of the ammonia plant and cover the electricity demand of the whole facility. Therefore, although occasionally proving less profitable than the original configurations, the retrofitting of conventional ammonia plants with the Allam cycle turned out to be an available-to-date strategy capable of fully abating both direct (*i.e.*  $\text{CO}_2$  released on-site by the front-end) and indirect (*i.e.*  $\text{CO}_2$  generated by the power stations for the production of electricity) emissions at all plant capacities. Finally, the significant amount of ready-to-pipeline  $\text{CO}_2$  produced by the shown configurations may represent an attractive feed for several industrial applications. Indeed, although over 35  $\text{Gt}_{\text{CO}_2}/\text{y}$  are currently emitted into the atmosphere every year since the last decade (Our World in Data, 2022), shortages in industrial-grade carbon dioxide have been particularly recurrent in the past few months in several European countries, including Italy, Germany, and the United Kingdom (The Guardian, 2022). Moreover, by doing so, completely abating carbon dioxide emissions in the atmosphere allows to avoid paying any “carbon tax”, which are expected to substantially raise in the next years.

#### References

- Allam, R.J., Brown Jr., G.W., & Palmer, M.R. (2015). System and method for high efficiency power generation using a carbon dioxide circulating working fluid (U.S. Patent N° 8,959,887).
- Ebrahimi, A., Meratizaman, M., Reyhani, H.A., Pourali, O., & Amidpour, M. (2015). Energetic, exergetic and economic assessment of oxygen production from two columns cryogenic air separation unit. *Energy*, 90, 1298-1316.
- IEA (2019). The Future of Hydrogen. International Energy Agency.
- IEA (2022). CO2 Capture and Utilisation. International Energy Agency.
- IEAGHG (2015). Oxy-Combustion Turbine Power Plants. International Energy Agency GHG.
- Isella, A., & Manca, D. (2022). GHG Emissions by (Petro)Chemical Processes and Decarbonization Priorities—A Review. *Energies*, 15(20), 7560.

- Our World in Data (2022). Annual CO<sub>2</sub> emissions. Available at: <https://ourworldindata.org/explorers/co2> (accessed on November 30<sup>th</sup>, 2022)
- Soltani, R., Rosen, M., & Dincer, I. (2014). Assessment of CO<sub>2</sub> capture options from various points in steam methane reforming for hydrogen production. *International Journal of Hydrogen Energy*, 39, 20266-20275.
- The Guardian (2022). UK food and drink sector sounds alarm over CO<sub>2</sub> shortage as plant halts output. Available at: <https://www.theguardian.com/business/2022/aug/25/co2-producers-meet-food-needs-halt-production-energy-prices> (accessed on November 30<sup>th</sup>, 2022).
- Wang, S., Fernandes, D., Xu, Q., & Chen, D. (2021). New Conceptual Design of an Integrated Allam-Cycle Power Complex Coupling Air Separation Unit and Ammonia Plant. *Industrial & Engineering Chemistry Research*, 60(49), 18007-18017.
- Ye, H., Zheng, J., & Li, Y. (2019). Feasibility analysis and simulation of argon recovery in low oxygen-purity cryogenic air separation process with low energy consumption. *Cryogenics*, 97, 109-121.
- Young, A.F., Villardi, H.G.D., Araujo, L.S., Raptopoulos, L.S.C., & Dutra, M.S. (2021). Detailed Design and Economic Evaluation of a Cryogenic Air Separation Unit with Recent Literature Solutions. *Industrial & Engineering Chemistry Research*, 60(41), 14830-14844.



# Into the Valley of Death Rode the Green Transition

Robert Pujan<sup>a,b</sup> and Heinz A. Preisig<sup>b</sup>

<sup>a</sup>*BNT Chemicals GmbH, PC-Straße 1, 06749 Bitterfeld-Wolfen, Germany*

<sup>b</sup>*NTNU Norwegian University of Science and Technology, Høgskoleringen 5,  
7491 Trondheim, Norway*

*robert.pujan@bnt-chemicals.de*

## Abstract

Although the concept of Valley of Death (VoD) has been widely documented, also in the context of the slow uptake of the Green Transition (GT), the available literature rarely provides any contemplations from more than one point of view. By reviewing works from different viewpoints like economics, policies, research, and industries, this contribution discusses the phenomenon of VoD in the GT more comprehensively and outlines the extent of the issues at hand. The study argues that VoDs of individual biorefinery technologies are not solved independently but only by overcoming the systemic barriers that stranded the GT in its overarching VoD. In fact, the GT efforts may be the first venture that has to address such a multitude of technological, economic, social, political, and market aspects on its way to maturity and commercialization.

**Keywords:** biorefinery, assessment, sustainability, policy, industry

## 1. Motivation

Despite the interest and support for sustainable process industries, shared by many actors from the public, politics, industries and research, being higher than ever, the Green Transition (GT) is leaping behind considerably. With most biorefinery concepts still being stuck between the experimental and demonstration phases, unable to cross this so-called Valley of Death (VoD), the deployment of sustainable technologies on a full-industrial scale has been slow at best. The concept of VoD refers to a point in process or technology development where the sum of available research grants and existing commercialization funding is at its minimum. This is usually experienced after the technology readiness level (TRL) 4 has been achieved - the lab-scale validation, and before TRL 7 can be met, that is, the technology's prototype implementation in a relevant, operational environment. As depicted in Figure 1, research activities push new concepts until TRL 4, mainly aided by public funding. From TRL 6 on, those concepts have to be pulled by the industries since private investments become the predominant funding. TRL 5–7 are therefore the bridge between basic knowledge generation and industrial application. If, before this phase, the technology has not reached

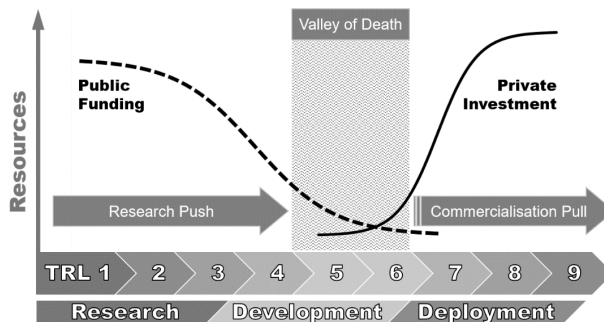


Figure 1: The Valley of Death in technology development



sufficient legitimacy and reliability for commercialization pulls, it strands in a calm belt of next to no push or pull; the VoD.

Since development processes are not linear and often even require the major scientific, technological and social disruptions of so-called paradigm shifts (Kuhn, 1970), the existence of VoDs is nothing new. Accordingly, the barriers involved are not only technical ones; those of economic, sustainable and socio-political nature continue to exist, although many technical issues may be long resolved. Identifying and understanding these barriers and how they can be overcome is thus essential for supporting the GT.

In an unfortunate plot twist concerning the GT, most of the VoDs observed with individual technologies can not be resolved independently - they are mere projections of the much larger, overarching VoD that the GT is currently stranded in. Analogously, the aforementioned barriers are rooted in significantly larger and systemic causalities, thereby necessitating the discussions on them to be lifted to a broader scale as well. The GT is probably the first genuinely global technical transformation humankind has ever aimed for; a new Industrial Revolution and an absolute at that. It is thus of no surprise, that the GT faces such a multitude of interrelated, global-scale barriers of scientific, technological, economic, social and political natures. However, the available literature on the concept of VoD and barriers of the GT rarely addresses more than one discipline and point of view. This contribution, therefore, bases its discussion on reviewing prior work from different sources like economics, policies, research and industries, thereby painting a comprehensive picture of the root causes and the extent of the issue at hand.

## **2. The Green Valley of Death**

During the Industrial Revolution in the 18<sup>th</sup> and 19<sup>th</sup> centuries, hand-production workshops have largely been replaced with mechanized factories, thereby prompting such an unprecedented, radical technological shift that society was fundamentally altered. Considering the GT's scope and expected global impact, disruptive industrial shifts similar to those aforementioned are more likely than the comparably confined societal impacts of recent technological innovations. However, the GT is still often handled as a slow, iterative process of incremental changes and regime reconfigurations, which may hinder renewal and profound structural change if left unchallenged in policies and governance (Bauer, 2018). This notion is fortified by the current policies' emphasis on primarily supporting new technologies' developments, which seem to implicitly assume default commercialization of them as soon as sufficiently developed (Hansen & Coenen, 2017). Apparently, many actors still try to enforce traditional transformation measures on the GT instead of enabling much-needed paradigm shifts. Meanwhile, the private sector proves those notions wrong by being unwilling to finance biorefineries alone due to residual GT-associated risks (Philip, 2018). Even after over 20 years of research and development (R&D), the GT still struggles to move from a formative R&D phase towards one of commercial growth. Interrelated barriers that involve co-evolving changes in technologies, markets, policies and industrial strategies as well as capabilities still limit interests and abilities to mobilise resources around the emerging technologies, thereby embedding the GT in a typical case of systemic lock-in (Hansen & Coenen, 2017).

### *2.1. Industrial Constraints*

Although innovation fosters economic growth, the GT depends on industries entering into uncertain environments, where they may be well advised to proceed and invest cautiously. Correspondingly, Nilsen (2017) describes that despite a few corporative strategy changes in recent years, many mature companies like Statoil ASA (since renamed Equinor ASA) remain highly dependent on the petroleum path as the main competitive advantage in the

years to come. Industrial incumbents have lower incentives to prioritise radical innovations since efficiency and profit come from established, predictable and reliable technologies, and the organizational routines surrounding them. It also implies bureaucratic inertia and resistance to change. The Technology Innovation System (TIS) analyses by Hansen & Coenen (2017) and Frishammar et al. (2019) identified, among others, high capital intensity, a lack of strategic value-chain partnerships and market knowledge, as well as overall system weaknesses rooted in uncertainty, complexity, equivocality and ambiguity as the main limiters in the technological adoption of the GT. A limited industrial interest may also be inferred from present business models certainly having to drastically adapt in the GT. Contrary to traditional industrial characteristics like the economics of scale and competition based on market prices, businesses arising from the GT may be characterized by distributed, small-scale plants, economics dominated by local supply/demand, and competition in terms of sustainability. Growing new sectors like this, however, requires operational support via appropriate policies due to long formative periods in which investments can be substantial with little to no return and new businesses struggling to develop any legitimacy (Peck et al., 2009). Industrial initiatives require investment motivation and assistance by long-term, reliable policies.

### *2.2. Market Barriers*

Instead of achieving the multi-product biorefinery configuration, already implemented biorefineries are mostly working as stand-alone facilities without any integral, well-designed process for the maximum use of an extensive feedstock portfolio or a range of different products (Cardona-Alzate et al., 2020). Consequentially, they are at the mercy of market fluctuations in the absence of operational support. Although renewable, the capitalism-inherent intention of unlimited growth cannot be met within the narrow constraints of biomass. Policies that aim on evolving the GT concurrent to the petroleum industries, increasing the industries' overall size instead of substituting fossil with sustainable resources, will thus fail due to many overlapping resource and market competitions. These competitions are, besides a lack of coordination and policy instruments in niche markets, already today a predominant cause for weak resource mobilization and market formation (Hansen & Coenen, 2017). There is a strong need to support the creation of bio-product markets through, for example, public procurement, price guarantees and biomass/bio-product production incentives (Fevolden et al., 2017). The implementation of carbon taxations and fossil fuel subsidy reforms would especially aid the creation of new markets, however, policies explicitly facilitating market adoption and commercialisation are weak or altogether lacking (Hellsmark et al., 2016). As a result, current market structures neither secure a sufficient supply of sustainable feedstocks nor generate ample demand for bio-products, thereby creating one of the conditions that deter investors (Philip, 2018). Supply chain analyses are therefore critical to the success of the entire GT. Their reliability may, however, be questioned: Seidenberger et al. (2010) and Schueler et al. (2016) demonstrate significant variability in published global biomass potentials due to several weaknesses and a lack of short-term estimates (i.e. 2030).

### *2.3. R&D Bottlenecks*

Even though segueing to barriers within R&D may be appropriate at this point, most of the discussion can be found in Pujan & Preisig (2022), therefore keeping this paper's considerations on the topic to a minimum.

The conservative notion of forcing bio-products into the shapes of fossil products seems to be a predominant factor that is hindering innovations in R&D. However, despite early propositions by, for example, Marquardt et al. (2010) to use the GT as a unique opportunity for redesigning industrial value chains, current literature shows the majority

of R&D trends still aiming at bio-feedstock conversions to products very similar to their fossil counterparts. In this regard, knowledge of biorefinery integrations into existing technical systems may remain critical for further developments, however, Bauer et al. (2017) reported a consensus in the literature that R&D may be necessary but investing more resources in it would certainly not enable biorefineries to cross the VoD.

#### *2.4. Erratic Policies*

Since the Industrial Revolution has been a century-long effort, one may conclude that the GT's slow uptake is par for the course. However, scientific innovation methods have tremendously improved since the 18<sup>th</sup> century and are accelerated by the rise of computing. This evolution is, however, barely visible since the GT faces barriers unknown to the Industrial Revolution. The latter was mainly driven by economic growth, disregarding environmental and social ramifications, and did not encounter fierce competition from deprecated industries. Additionally, the early years of the Industrial Revolution were confined to national scales, whereas the GT was always a global venture. Industrial boundaries often do not coincide with national borders and may vary between different industrial sectors. Accordingly, without supra-national policies that are stable in the long term, private investments in disruptive GT technologies remain too risky for consideration. Many EU and national policies still show insufficient levels of coherence and private sector incumbents question the coordination between EU, national and regional policy levels, especially in terms of biomass use (Peck et al., 2009). Due to the rush and hype regarding biofuels and bioenergy in recent years, many supra-/national policies are incentivizing the energetic utilisation of biomass, while almost neglecting material exploitation. Since the private sector must necessarily respond to market forces, a broad uptake of new businesses is generally dependent on significant political support with a consistent future vision. Regional developments, national R&D and global issues necessitate sufficiently flexible policies that are coherent across boundaries and minimum in duplications (Philip, 2018). Policies are of particular importance where significant positive externalities of the new system, like environmental and social benefits, are yet not priced by current market structures (Peck et al., 2009). However, although current political discussions centre on sustainability narratives, the resulting policies remain focused on efficiency, productivity and industrial competition instead of environmental benefits (Bauer, 2018). Standard policies like these do not yield the same effect in the GT as in other sectors, as demonstrated by Maes & Passel (2019). By investigating the policy frameworks for past technological transitions, Verbong et al. (2008) conclude that close adherence to industrial policies is problematic and that the innovation learning effects were not sufficiently guarded by past policy instruments. More than any other sector, the GT requires a combination of technology-push and demand-pull policies to stimulate sufficient levels of innovation and progress (Frishammar et al., 2019). Public-private partnerships, governmental loan guarantees and other innovative funding instruments like Green Bonds are necessary to de-risk private investments in the GT (Philip, 2018). This asks for large capital sums and long-term commitments since any biorefinery would need external support until the industry achieves critical mass and supply-demand stability. Relocating funding from fossil subsidies and introducing carbon taxations can be a huge source for these endeavours. However, their widespread realisations are still due.

#### *2.5. Legitimacy Thresholds*

Conducting the GT calls for eliminating more barriers than mere technological ones. It requires transitions in consumer behaviour and expectations, institutional norms, standards and regulations, as well as technological, economic and organizational innovations throughout supply and value chains – it is such a politically contested

transformation where public perceptions are key (Bauer, 2018). Understanding public perceptions has been shown by Marciano et al. (2014) to be critical for managing the costs of emerging biorefineries. Peck et al. (2009) outlines how an understanding of the processes of building cognitive and socio-political legitimacy is relevant to the GT and how early legitimacy enhancements can reduce the duration and difficulties of formative phases in the evolution of innovative businesses. The socio-economic aspect influences most of the barriers to the critical adoption of the GT, which necessitates policy intervention (Ubando et al., 2020). Unfortunately, the public is largely oblivious to the environmental and social costs of fossil fuel subsidies, as well as the state-of-the-art in the GT. Accordingly, public and political opinions are erratic, and actors may even be confused by the aims and activities of the GT. An informant in the study of Frishammar et al. (2019) reported public backlash to the company when buying palm oil, even though it was certified – strangely, the public is less discerning with regard to crude oil. The impact and pace of climate change are just not obvious enough and too complex to grasp for most of the public. Similarly, the GT's complexity and its multitude of technological pathways are confusing even to scientists working in the field (Peck et al., 2009). Researchers still debate whether the bioeconomy is itself inherently sustainable or rather some weak ecological modernization and thus a threat to sustainability (Pfau et al., 2014). Truly sustainable solutions require much wider perspectives than presently considered. There are some mainstream claims that the GT's early adopters will mostly be interested in the green image, thus, technologies would not have to be refined or reasonably priced. Although those early adopters may provide initial finances, relying on them would certainly not enable an absolute GT. Furthermore, since the sustainability concept is often not objectively defined in current assessment methods (Cardona-Alzate et al., 2020), reaching public and political legitimacy in terms of sustainability may be a lost game.

Although R&D is proposing several technologies for the GT, hardly any evidence is provided to ensure widespread economic legitimacy. Evidence has to prove the feasibility and reliability of the GT technologies in development that is systematic scale-up studies, extensive simulation runs and reliable assessment analyses. After failed R&D hypes in recent decades, like, for example, microalgae, public funding measurements and investors understandably became very hesitant.

### **3. Concluding Remarks**

As reasoned before by others as well, the GT can be successful in terms of economics, sustainability and society only if developed under certain conditions (Devaney & Henchion, 2018). However, as discussed here, the lack of these conditions as well as a multitude of interrelated, global-scale barriers of scientific, technological, economic, social and political natures keep the GT stranded in a VoD. Conflicting narratives dominate public opinion, policy-makers are acting too reluctant, R&D fails to establish legitimacy for GT technologies, and industries are enclosed in established routines and market constraints. Many barriers are rooted in notions that still try to undertake the GT without drastic shifts. Often, outdated measures as well as the profile of traditional fossil industries are enforced on the GT, thereby ignoring the largely different nature of this industrial revolution. More than 20 years of mainstream attention to the GT, yet, the nurture of “monkey see, monkey do” remains a dominating approach, trying to reproduce traditional chemical industries and old societal systems in the GT. It is about time to encourage that old monkey to take more risks and teach it some innovations.

## Acknowledgements

This paper originated from ongoing work in the research projects Bio4Fuels (Norwegian Centre for Environment-friendly Energy Research (FME), project 257622) and MarketPlace (Horizon 2020, project 760173).

## References

- F. Bauer, 2018, Narratives of biorefinery innovation for the bioeconomy: Conflict, consensus or confusion?. *Environ. Innov. Soc. Trans.*, 28, 96-107
- F. Bauer, L. Coenen, T. Hansen, K. McCormick & Y.V. Palgan, 2017, Technological innovation systems for biorefineries: a review of the literature, *Biofpr.*, 11(3), 534-548
- C.A. Cardona-Alzate, S. Serna-Loaiza & M. Ortiz-Sanchez, 2020, Sustainable Biorefineries: What was learned from the design, analysis and implementation, *J. Sustain. Dev. Energy Water Environ. Syst.*, 8(1), 88-117
- L. Devaney & M. Henchion, 2018, Consensus, caveats and conditions: International learnings for bioeconomy development, *J. Clean. Prod.*, 174, 1400-1411
- A.M. Fevolden, L. Coenen, T. Hansen & A. Klitkou, 2017, The role of trials and demonstration projects in the development of a sustainable bioeconomy. *Sustainability*, 9(3), 419
- J. Frishammar, P. Söderholm, H. Hellsmark & J. Mossberg, 2019, A knowledge-based perspective on system weaknesses in technological innovation systems, *Sci. Public Policy*, 46(1), 55-70
- T. Hansen & L. Coenen, 2017, Unpacking resource mobilisation by incumbents for biorefineries: The role of micro-level factors for technological innovation system weaknesses, *Technol. Anal. Strateg. Manag.*, 29(5), 500-513.
- H. Hellsmark, J. Mossberg, P. Söderholm & J. Frishammar, 2016, Innovation system strengths and weaknesses in progressing sustainable technology: The case of Swedish biorefinery development, *J. Clean. Prod.*, 131, 702-715
- T.S. Kuhn, 1970, *The structure of scientific revolutions* (Vol. 111), University of Chicago Press
- D. Maes & S. Van Passel, 2019, Effective bioeconomy policies for the uptake of innovative technologies under resource constraints, *Biomass and Bioenergy*, 120, 91-106
- J.A. Marciano, R.J. Lilieholm, M.F. Teisl, J.E. Leahy & B. Neupane, 2014, Factors affecting public support for forest-based biorefineries: A comparison of mill towns and the general public in Maine, USA. *Energy Policy*, 75, 301-311
- W. Marquardt, A. Harwardt, M. Hechinger, K. Kraemer, J. Viell & A. Voll, 2010, *The Biorenewables Opportunity – Toward Next Generation Process and Product Systems*, *AIChE Journal*, 56(9), 2228-2235
- T. Nilsen, 2017, Innovation from the inside out: Contrasting fossil and renewable energy pathways at Statoil, *Energy Res. Soc. Sci.*, 28, 50-57
- P. Peck, S.J. Bennett, R. Bissett-Amess, J. Lenhart & H. Mozaffarian, 2009, Examining understanding, acceptance, and support for the biorefinery concept among EU policy-makers, *Biofpr.*, 3(3), 361-383
- J. Philp, 2018, The bioeconomy, the challenge of the century for policy makers, *New Biotechnol.*, 40, 11-19
- R. Pujan & H.A. Preisig, 2022, Biorefinery modelling is in tatters, and here is why. *Comp. Aid. Chem. Eng.*, Vol. 51, 295-300
- V. Schueler, S. Fuss, J.C. Steckel, U. Weddige & T. Beringer, 2016, Productivity ranges of sustainable biomass potentials from non-agricultural land, *Env. Res. Letters*, 11(7), 074026
- T. Seidenberger, D. Thrän, R. Offermann, U. Seyfert, M. Buchhorn & J. Zeddies, 2010, *Global Biomass Potentials. Investigation and assessment of data, remote sensing in biomass potential research, and country-specific energy crop potentials, energy [r] evolution - A sustainable world energy outlook*, 3, 166-168
- A.T. Ubando, C.B. Felix & W.-H. Chen, 2020, Biorefineries in circular bioeconomy: A comprehensive review, *Bioresour. Technol.*, 299, 122585

# A novel industrial biotechnology approach to valorize fatty acids to bioplastics: scope for scale-up and process efficiency using an integrated approach

Pantelis Vasilakis<sup>a</sup>, Kostas Pyrgakis<sup>a</sup>, Melina Psycha<sup>a</sup>, Antonino Biundo<sup>b</sup>,  
Antonis Kokossis<sup>a,\*</sup>

<sup>a</sup>Chemical Engineering National Technical University of Athens, Athens, Greece

<sup>b</sup>REWOW srl, Via Ciasca 9, 70124, Bari, Italy

[akokossis@mail.ntua.gr](mailto:akokossis@mail.ntua.gr)

## Abstract

Waste Cooking Oils (WCO) are considered a harmful waste, though also constitute a valuable carbon source if treated correctly. This paper illustrates a novel WCO valorization chemistry to produce added-valued bioplastics. The core process is based on bio-based catalysis, in which *E.Coli* is fermented to produce oleate hydratase enzyme, which catalyzes FFAs conversion into the key building block of 10-Hydroxystearic-Acid (10-HA). The process is further integrated with upstream saponification-acidification stages for the pre-treatment of mixed WCOs, while downstream chemical processes (hydrogenation, esterification, polymerization) are upgrading the 10-HA building block into bio-based poly-10-HAME polymer. As upstream and downstream stages involve different scales, the paper addresses the scope to scale-up using a systems approach that builds efficiency and optimizes interactions. Overall, 3 tn of WCO yield in 1.5 tn of poly-10-HAME, CAPEX and OPEX are estimated to 141 M€ and 65 M€/yr, while emissions are estimated to 2.4 kg CO<sub>2</sub><sub>eq</sub>/kg biopolymer.

**Keywords:** WCO, FFA, Bio-catalysis, Fermentation, *E.Coli*.

## 1. Scope of work

Waste Cooking Oils (WCO) constitute a valuable carbon source composed of unsaturated triglycerides (TGs) collected as wastes after frying with vegetable oils. They are usually disposed in landfills and/or along with wastewater resulting in huge environmental impacts; nevertheless, there are still options for recycling. The traditional solutions include the well-known transesterification chemistry that returns 1G biodiesel. Biofuel production is questioned in terms of social (edible sources), economic and environmental aspects (Lopes et al, 2019). Other emerging technologies use alternative renewable feedstocks (Pateromichelakis et al, 2022), or target to other bio-based products, like alcohols for vehicle and/or shipping fuels (Pyrgakis et al, 2016). Instead, a novel green solution based on a biocatalytic chemistry is proposed enabling the valorization of WCOs towards high-value bioplastics. The chemistry has been validated at laboratory scale (TRL4) and, in this work, it has been studied at modelling level to build knowledge and efficiencies for scaling up purposes.

The chemistry is based on the catalysis of the transformation of FFAs into 10-HA by means of oleate hydratase enzyme, which is produced by *Escherichia Coli*. The unsaturated fats of WCO are processed to prepare the required FFAs feedstock for the parallel enzymatic conversion to 10-HA, while, finally, 10-HA is upgraded to prepare the

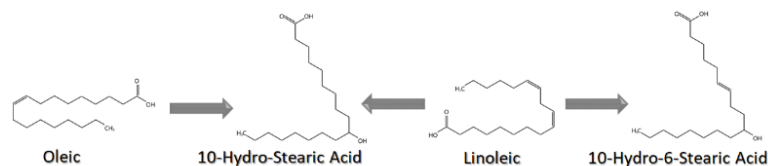


Figure 1 Reactions of FFAs transformation to biopolymer building blocks

10-HAME monomer for biopolymer production. Specifically, saponification is decomposing triglycerides (WCOs) into soaps, while acidification of soaps returns the exploitable FFAs co-substrate. *E. Coli* is fermented using glucose substrate and nutrients as cultivation media for growth and the production of oleate hydratase enzyme. The latter catalyzes the transformation of oleic and linoleic acids (FFAs) into 10-Hydroxyl-stearic acid and 10-Hydroxy-linoleic acid (10-HAs) following the reaction paths of Figure 1. Next, both 10-HAs and unconverted FFAs are hydrogenated generating their saturated fractions, which are next esterified forming the 10-HAME precursor and biodiesel (FAME). The latter closes the gap of carbon losses of the organic feedstock, which are not catalyzed to 10-HAs. The two fractions (10-HAME and FAME) are split by distillation returning drop in biodiesel and pure 10-HAME for polymerization purposes.

Based on experimental data for the core and surrounding chemistries, an integrated biorefinery flowsheet was developed from scratch in ASPEN Plus for analysis and simulation purposes.

## 2. Challenges in WCO valorisation

The studied chemistry was never addressed before at a modeling level, or as a complete biorefinery flow diagram. Experimental data concerning the biocatalysis cultivation media and its substrates and conversion efficiencies of FFAs are available for analysis. In this scope, challenges rise in terms of transferring the new-to-nature chemistry into a complete biorefinery flowsheet from raw (WCOs) and auxiliary (i.e., glucose) substrates to end-biopolymer products. Moreover, this plant utilizes 5 different chemical paths (saponification, acidification, fermentation, hydrogenation, polymerization) and challenging separation unit operations (e.g., spray drying). As far as concerns biocatalysis, the chemistry brings new challenges in modelling hybrid systems utilizing a fermentation-driven pathway and enzymatic conversion across the same system. Biocatalysis combines growth with glucose and production of primary metabolites (enzyme), which are exploited as catalyzing means of other reactions converting the FFAs co-substrate into desired products. These conditions increase the size and complexity of coupling different kinetic models that essentially require Monod-based kinetics for the *E. Coli* fermentation, and Michaelis-Menden-based kinetics for the enzymatic reactions. In addition, fermentation is affected by  $O_2$  inhibition and temperature bounds that define fermentation and enzymatic efficiencies. Finally, the modeling procedure was challenged by lack of properties for special nutrients/auxiliaries, like the biochemical inducer IPTG.

## 3. Methodological steps

This work followed a 6-step approach to transfer chemistries at laboratory scale into a complete process biorefinery flowsheet for simulation, scaling up and analysis purposes of new-to-nature biotechnology concepts. The approach includes (1) estimation of biocatalysis kinetics; (2) process flowsheeting from scratch; (3) simulation and scaling up; (4) energy integration; (5) LCA for estimation of environmental impacts; and (6) techno-economic analysis of bio-based polymers.

### 3.1. Development of biocatalysis kinetics

An engineered strain of *E.coli* is incubated to achieve desirable growth rates for target scales. In parallel, the raw feedstock is prepared by means of saponification (KOH) and acidification (H<sub>2</sub>SO<sub>4</sub>) resulting treated FFAs; the conventional reaction systems are modeled by RSTOIC in ASPEN. Centrifuges are employed to remove water from organic phases. Next, experimental data including biocatalysis yields were used for kinetics modeling. Conditions such as growth, temperature and O<sub>2</sub> inhibition are considered resulting in an extended version of Monod kinetic model, which integrates the following terms for each limiting condition (Wang et al, 2014; Cruz et al, 2018):

$$\text{Growth: } r_c = k^{cell} \cdot \frac{C_c \cdot C_S}{C_S + C_M^{Mon}} \Leftrightarrow \frac{dC_R}{dt} = k^{cell} \cdot \frac{C_c \cdot C_S}{C_S + C_M^{Mon}} \cdot \frac{C_R - C_{R0}}{C_c - C_{c0}} \quad (1)$$

$$\text{Temperature: } r_c = r_c^{max} \cdot \frac{T_{max} - T}{T_{max} - T_{opt}} \cdot \left( \frac{T}{T_{opt}} \right)^{\frac{T_{opt}}{T_{max} - T_{opt}}} \quad (2)$$

$$\text{O}_2 \text{ inhibition: } k_{obs} = k^{cell} \cdot \left( 1 - \frac{C_{O_2}}{C_{O_2}^*} \right)^n \quad (3)$$

$$\text{Enzymatic conversion: } r_{10HA} = k^{enz} \cdot \frac{C_{E0} \cdot C_{FFA}}{C_{FFA} + C_M^{enz}} \quad (4)$$

where  $C_i$  and  $C_{i0}$  are the concentrations of component  $i$  at the end of each batch experiment and the beginning, respectively;  $C_{O_2}^*$  the limiting concentration of inhibition;  $r_c$  the reaction rate of cells growth;  $k$  the maximum growth rate;  $C_M$  the Monod constant;  $k_{obs}$  the observed growth rate;  $n$  the exponential factor of inhibition; and  $T_{max}$ ,  $T_{opt}$  and  $T$  the maximum, the optimum and the actual temperatures in cultivation medium. The indexes C, S, R and O<sub>2</sub> account for cells, Substrate (glucose), product (enzymes) and oxygen components that participate and affect the biocatalysis system.

### 3.2. Process flowsheeting of biocatalysis chemistry

The overall biorefinery diagram is represented in Figure 2. At preliminary steps of the plant, the WCOs feedstock is chemically pre-treated. The cell factories and the processed feedstock are mixed in conventional reactor systems generating 10-HAs (with and without hydroxyl groups) that are saturated by hydrogenation. The saturated FFAs are esterified with methanol yielding in the 10-HAME biopolymer precursor and FAME. Finally, 10-HAME is polymerized resulting in the end-product of poly-10-HAME.

The above process diagram of Figure 2 was designed in detail from scratch including all process operations. The flowsheeting procedure was followed to deliver all pre-treatment, bioconversion and formulation chemistries; including intermediate separation stages for cost-effective removal of by-products, co-products and water-organic streams resulting high purity poly-10-HAME. In Aspen, the ‘‘General’’ type and the ‘‘LHHW’’ class of reactions have been used to integrate the complex kinetics. Based on the calculated Monod parameters, the optimal cells and glucose concentrations that result in maximum production rate and conversion have been defined. Accordingly, based on the tuned fermentation and enzyme production rates, the optimal feed rate of FFAs and the retention time for maximum production of 10-HA are estimated. Accordingly, the reactor volume is specified and the RCSTR model (Aspen plus) was used for the biocatalysis system.

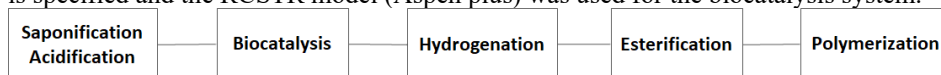


Figure 2 Biorefining stages



Based on the finalized process flowsheet, techno-economic and LCA analyses have been conducted to address the economic and environmental sustainability of biopolymers production in front of competitive synthetic and fossil-based polymers.

### 3.2.1. Scaling up and analysis

The developed process flowsheet was modeled in Aspen for the simulation of mass and energy balances at higher scales (3 tn/hr). Thereafter, energy integration has been performed to identify the maximum feasible energy savings at higher scales. Moreover, the biorefinery input-outputs (energy, materials and wastes) fed the LCA analysis for the estimation of mid- and end-point indicators; including the global warming potential (CO<sub>2</sub> equivalent) for biopolymers production. Finally, a techno-economic analysis focusing on capital and operating expenses was performed to identify profitability margins and potentials for market entry of alternative biopolymers.

## 4. Results

The process flowsheet that is presented in Figure 3 consists of 4 sections for (i) WCO pre-treatment, (ii) biocatalysis, (iii) FFAs upgrade and (iv) biopolymers production.

### 4.1. Saponification-Acidification stage (WCO pre-treatment)

Saponification (RSTOIC model) takes place at 80 °C using KOH (5 M) to break the bonds of triglycerides of WCO resulting in glycerol and potassium oleic, linoleic, palmitic and stearic acid at conversion rates of 78.3%. In the basis of 3 tn of WCO, 2349.7 kg of potassium-FFAs are produced, which are then acidified at 80 °C in presence of H<sub>2</sub>SO<sub>4</sub> returning the treated FFAs for biocatalysis. A vapor-liquid separation stage is employed to recover volatile compounds in acidification effluents, while a downstream decanter is implemented to remove the water-glycerol content finalizing the biocatalysis feedstock.

### 4.2. Biocatalysis stage

The biocatalyst was prepared by inoculation to reach 3 tn of biomass ready for use in fermentation (37 °C). The biocatalysis step results in 1827.1 kg (77.7% conversion) of 10-Hydroxystearic-Acid. Based on experimental data, the kinetic parameters are accordingly estimated as follows:  $k^{cell} = 0.022388 h^{-1}$ ,  $C_M^{Mon} = 74.6269 mmol/L$ ,  $T_{max} = 315.2 K$ ,  $T_{opt} = 311.9 K$ ,  $C_{O_2}^* = 0.12 mg/L$ ,  $n = 1$ ,  $k^{enz} = 1.0288 mol/L$ ,  $C_M^{enz} = 0.025 mol/L$ . The fermentation exit stream is treated with ethyl-acetate solvent (5% v/v) to facilitate mixing and extraction of the biocatalysis phase (desired products) and removal of the excess water content by centrifugation. Next, spray drying was implemented to recover an ethyl-acetate rich phase (43%), the unsaturated FFAs (47%), and the remaining water (10%).

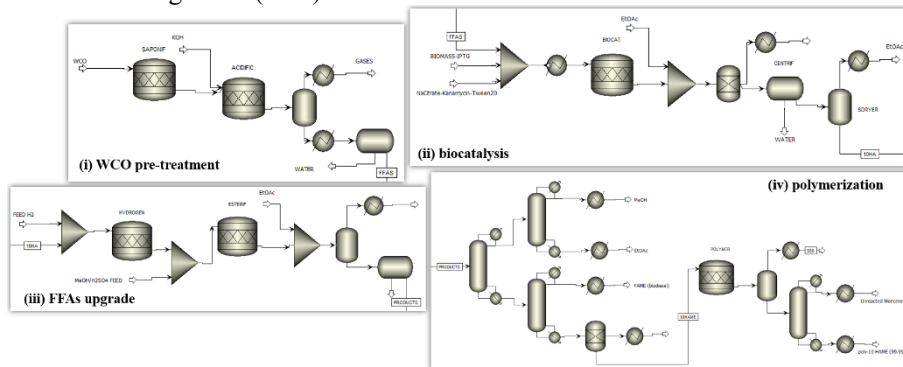


Figure 3 Biorefinery process flowsheet for biocatalysis implementation

#### 4.3. Hydrogeneration and Esterification stage (FFAs upgrade)

The recovered organic phase was hydrogenated (fixed bed reactor) with the use of a nickel-based catalyst to saturate double carbon bonds with hydrogen reaching nearly 99.9% saturation. The esterification stage is next implemented to treat the saturated acids with methanol. The key component, 10-Hydroxystearic-Acid, is esterified into 10-Hydroxystearic-Acid-Methyl-Ester (10-HAME) by 99.4% (1816.72 kg), while the methyl-esters (FAME) of all remaining FFAs are produced. The effluent methyl-esters stream is mixed with ethyl-acetate solvent to entrain the organic phase from water, which are recovered through a 2 liquid phase separation.

#### 4.4. Polymerization stage

Downstream esterification, a 3-distillation system is employed to recover the 4 key fractions from the ethyl-acetate entrainment: (i) unreacted methanol, (ii) FAME, (iii) 10-HAME and (iv) ethyl-acetate. The distillation separations have been designed to perform sharp splits at high purities ( $\geq 99\%$ ). Methanol and ethyl-acetate can be recovered back to esterification and liquid-liquid separation stages, while FAME (1440.5 kg) constitutes the biodiesel co-product. The 10-HAME (1816.54 kg) is finally driven to the polymerization stage producing the poly-10-Hydroxystearic-Acid-Methyl-Ester (1453.19 kg), that is equivalent to an overall yield of 48.44%.

#### 4.5. Evaluation: Energy, techno-economic and environmental analyses

Energy integration is applied to minimize operating and capital costs; the biorefining concept will fall short to compete with conventional processes or address sustainability targets [Kokossis et al, 2015; Pyrgakis et al, 2019]. Prior to energy integration, hot and cold utilities consumption have been at 7.4 MW and 14.1 MW, respectively. After integration the total utilities ware estimated to 2.24 MW (hot) and 8.9 MW (cold), which is equivalent to 70% and 37% of energy savings in heating and cooling demands. The optimal distribution over the utilities levels was accordingly estimated to 2.05 MW (LP steam), 0.19 (VHP steam) and 8.9 MW cooling water, as presented in the Grand Composite Curve of Figure 4.a.

Moreover, an LCA approach was followed using SimaPro to estimate the mid- and end-point indicators to produce biopolymers based on the input-output flows estimated by the biorefinery model. The Ecoinvent databases were used for the environmental indicators, while the ReCiPe methods was followed for the calculations. Accordingly, the contribution of the difference flows to the respective indicators are summarized in Figure 4.b. The emissions of CO<sub>2</sub> equivalent for the production of biopolymer were estimated

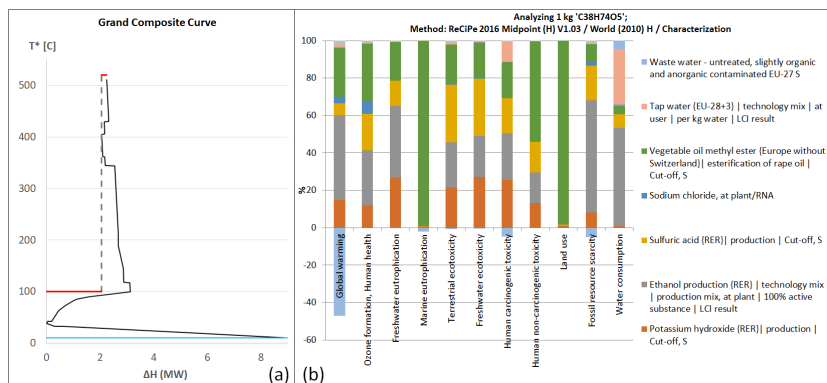


Figure 4 Energy Integration (a) and LCA (b) results

to 2.4 kg CO<sub>2</sub><sub>eq</sub>/kg biopolymer, while the emissions for conventional fossil-based polymers range between 2-3.5 kg CO<sub>2</sub><sub>eq</sub>/kg polymer.

Finally, the Aspen Process Economic Analyzer (APEA) was used to approximate the capital and operating expenses for production of 17,600 tn/yr FAME and 11,200 tn/yr biopolymers; accordingly, a mass allocation of 61-39% can be considered. The CAPEX was estimated to 141 M€, while OPEX was 65 M€/yr. As a result, for a depreciation period of 20 years (and the mass allocation), a production cost of 1.5 €/kg of biodiesel and ≈1 €/kg of biopolymer are estimated, while respective market costs are identified to 0.91 €/kg (biodiesel) and 1.8 €/kg (PVC equivalent cost). This is enough to consider that if another cost allocation strategy was accounted (i.e. by 34-64%), then an average price of 1.66 €/kg of biopolymer and 0.84€/kg of biopolymer could be achieved, resulting in improved prices (by 8%) compared to existing ones.

## 5. Conclusions

The work presented a new chemistry driven at industrial scales for production of biopolymers and biodiesel (co-product). The chemistry is attributed by complex kinetics in terms of cellular growth, enzymes production and enzymatic conversion of FFAs into valuable building blocks for biopolymers. Biodiesel was also considered to maximize utilization of residual carbon content of WCOs. The chemistry was presented at biorefining scales and achieved reductions at both costs and environmental impacts compared to existing solutions and products. Future work could include the investigation of alternatives auxiliaries (KOH, H<sub>2</sub>SO<sub>4</sub>) with lower cost and environmental impacts, as well as alternatives for downstream separations (e.g., membranes).

## Acknowledgments

Financial support from the project Industrial Biotechnology Innovation and Synthetic Biology Accelerator (PREP-IBISBA) is gratefully acknowledged.

## References

- Cruz, M., Costa, E., Fonseca Almeida, M., da Conceição Alvim-Ferraz, M., Maia Dias, J., 2018. Recovery Of By-Products From The Olive Oil Production And The Vegetable Oil Refining For Biodiesel Production. *Detritus In Press*, 1.
- Kokossis, A.C., Tsakalova, M., Pyrgakis, K., 2015, Design of integrated biorefineries, *Computers & Chemical Engineering*, 81, 40-56
- Lopes, M., Miranda, S.M., Belo I., 2019. Microbial valorization of waste cooking oils for valuable compounds production – a review. *Critical Reviews in Environmental Science and Technology*, 50 (24), 2583-2616
- Pyrgakis, K.A., Kokossis, A.C., 2019, A Total Site Synthesis approach for the selection, integration and planning of multiple-feedstock biorefineries, *Computers & Chemical Engineering*, 122, 326-355
- Pateromichelakis, A., Psycha, M., Pyrgakis, K., Maréchal, F., Kokossis, A., 2022, The Use of GVL for Holistic Valorization of Biomass, *Computers & Chemical Engineering*, 164, 107849
- Pyrgakis, K.A., de Vrije, T., Budde, M.A.W., Kyriakou, K., López-Contreras, A.M., Kokossis, A.C., 2016, A process integration approach for the production of biological iso-propanol, butanol and ethanol using gas stripping and adsorption as recovery methods, *Biochemical Engineering Journal*, 116, 176-194
- Wang, H., Wang, F., Wang, W., Yao, X., Wei, D., Cheng, H., Deng, Z., 2014. Improving the expression of recombinant proteins in *E. coli* BL21 (DE3) under acetate stress: an alkaline pH shift approach. *PLoS One*. 9(11):e112777

# On the re-deployment of lignocellulosic biorefineries with solvent-free designs and process-to-process integration

Iraklis Sarris, Kostas Pyrgakis, Antonis Kokossis\*

*School of Chemical Engineering, National Technical University of Athens, Iroon Polytechniou 9, GR-15780, Greece*

*\*Corresponding author: [akokossis@mail.ntua.gr](mailto:akokossis@mail.ntua.gr)*

## Abstract

This work follows a systems engineering approach to re-deploy a real-life lignocellulosic biorefinery for lower operating and capital costs, featuring benefits as solvent demands are reduced, and a better scope to integrate with first generation biorefineries. The redeployment focused on hot spots with high materials and energy use that affect cost-effective production of pulp, lignin, and bio-based products. The collaborative work investigates alternative designs for process, feedstocks and products portfolio and combines process-to-process integration to ensure optimal valorization of materials and energy. The reference real-life biorefinery case is modeled in ASPEN and involves high volumes of solvents for delignification, so re-deployment addressed a delignification-free case with parallel integration of downstream ethanol fermentation resulting in 22.5% lignin and 26% ethanol yields. The revised mass and energy balances of alternative scenarios were estimated and new insights for CAPEX (€ 39.6M), OPEX (€ 212M/yr) and energy costs (74% lower hot utilities) are obtained.

**Keywords:** biorefinery, delignification, bioethanol, techno-economic, energy integration.

## 1. Introduction

To date, there are 93 European lignocellulosic biorefineries (BBIC, 2022). Given the extended experience in 2G biorefining and the current social, political, and economic environments, the design contexts need to be revisited to face critical challenges related to cost and production efficiencies and serve replicable solutions for broader commercialization of the biorefining concept. State of the Art 2G biorefineries undertake the fractionation of lignocellulosic biomass into its main ingredients: C5, C6 sugars and lignin. All fractions can be integrated in bioeconomy markets and downstream chemistries for fuels, power, chemicals, and food additives production (Pyrgakis et al., 2016; Kokossis et al., 2015). This work is focusing on a real-life 2G biorefinery (CIMV, 2022), whose operation is based on processing wheat straw and poplar residues with organic acids (organosolv technology) returning C5, C6 sugars fraction and a functional biolignin based on the CIMV technologies (Delmas et al., 2011; Delmas et al., 2013; Nadja et al., 2014). Still, alternative organosolv technologies have been also recently appeared using biobased solvents and new fractions recovery concepts (Pateromichelakis et al., 2022).

Biolignin™ is a phenolic polymer with linear structure and low molecular weight. Moreover, the cellulosic polysaccharide is composed of a linear chain of  $\beta$ -1,4 linked d-glucose units, while the pentose sugar syrup consists of xylose monomers and oligomers (Mesa et al., 2016). The CIMV process takes advantage of the organosolv fractionation

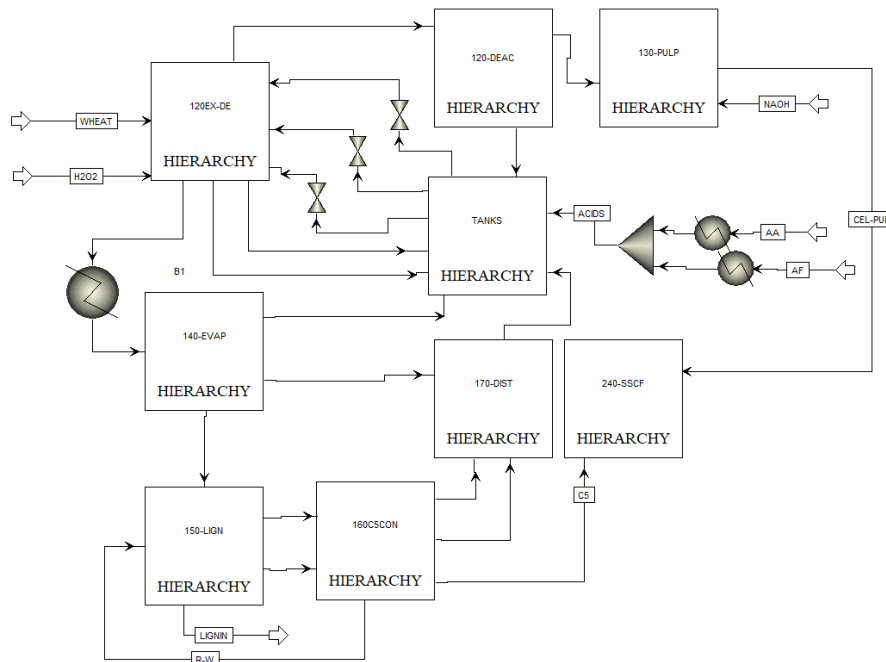


Figure 1 CIMV biorefinery flow diagram

concept utilizing a mix of organic acids as solvents of the three lignocellulosic components. The technology is applicable to both forestry and woody feedstocks and the overall biorefinery flow diagram is presented in Figure 1.

The biorefinery of CIMV consists of the main Biomass fractionation section (120EX-DE, in Figure 1), where C6 sugars and C5-soluble lignin are recovered from the raw lignocellulosic biomass. The C6 sugars are next delignified with a H<sub>2</sub>O<sub>2</sub>-organic acids mixture (Khama et al., 2005). After deacidification of the cellulosic fraction (120-DEAC), that is finally pulped (130-PULP) resulting the C6 sugars of the reference biorefinery. The C5-lignin mixture (from 120EX-DE) is dried through a thin film evaporation system (140-EVAP), while the lignin content is precipitated with water (150-LIGN), driving the residual C5 sugars towards the concentration section (160C5CON) where C5 sugars are collected. Any residual organic acids are recovered by distillation (170-DIST) and stored in the biorefinery acids tanks (TANKS) to be recycled across biorefinery.

Across the reference biorefinery, there are 3 main challenges identified that affect cost-effective valorization of biomass:

- i. the high auxiliaries' costs for delignification (water and H<sub>2</sub>O<sub>2</sub> solvents)
- ii. the co-production of added-value chemicals to face technoeconomic bottlenecks
- iii. the use of alternative feedstocks to face pricing seasonality and build flexibility

## 2. Problem description

As shown in Figure 2, the delignification process uses a 1:1 H<sub>2</sub>O<sub>2</sub>-water mixture to extract 2.6% of lignin trapped within pulp fraction returning 550 kg/hr of soluble Lignin from 21.000 kg/hr of the pulp feedstock. The high water and H<sub>2</sub>O<sub>2</sub> demands ( $\approx 4.500$  kg.hr) in combination with the high energy needs for downstream recovery of auxiliary organic acids and water crucially affect the sustainability of this section and profitability

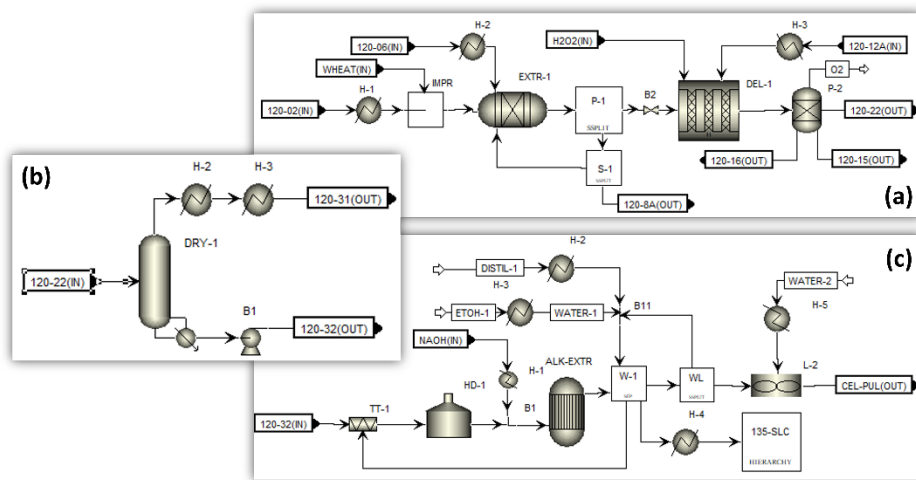


Figure 2 The delignification (a), deacidification (b) and alkaline extraction (c) sections of the whole biorefinery, raising concerns of re-engineering, or even removing, this section. Such decision will incur significant effects on the quality of delivered C6 sugars as end-product and might require its valorisation through alternative, still mature, technologies to other products. Modelling of the design alternatives will reveal vital answers on the trade-offs and redevelopment of the biorefining concept towards the exploitation of the C6 biomass fraction. Under these conditions, the replacement of the delignification section with direct hydrolysis and co-fermentation of C5 and C6 sugars for bioethanol production was addressed; otherwise, pulp could attribute commercial value as lower-grade C6 sugars.

The case of co-producing biofuels is a radical swift on the reference biorefining concept in technological and commercial contexts by changing the products portfolio. Bioethanol production can be provided by regular glucose (C6) fermentation, while simultaneous valorization of the xylan and glucan substrates are also considered to boost ethanol production (Hassan et al., 2019; CIMV, 2022; Pyrgakis and Kokossis, 2019).

The implementation of alternative biomass feedstocks featuring alternative C5-C6 sugars fractions are capable of significantly affecting bioethanol and functional lignin productivities. Changes in feedstocks will infer a series of changes in mass and energy balances of the high complexity biorefinery raising questions about sustainability margins on the different lignocellulosic feedstocks. Biorefining is naturally affected by seasonality of the different woody and forestry feedstocks – with subsequent effects on integration potential and energy savings – thus, flexibility in operation is indispensable to face pricing volatilities and seasonality. As a result, besides wheat straw, corn stover is also examined to investigate new production potentials and sustainability margins. All above challenges map a set of alternative design options for re-designing the reference biorefinery and address trade-offs in terms of capital, operating and integrated energy costs.

### 3. Methodology

Dealing with above challenges, this work employs process systems engineering know-how and tools to identify better solutions than the reference biorefinery scenario. The whole and interconnected sections of the reference biorefinery has been modeled in Aspen in contexts of a previous work (Mountraki et al., 2016). This work elaborates the existing process flowsheet by re-engineering the existing sections of Figure 2 providing a

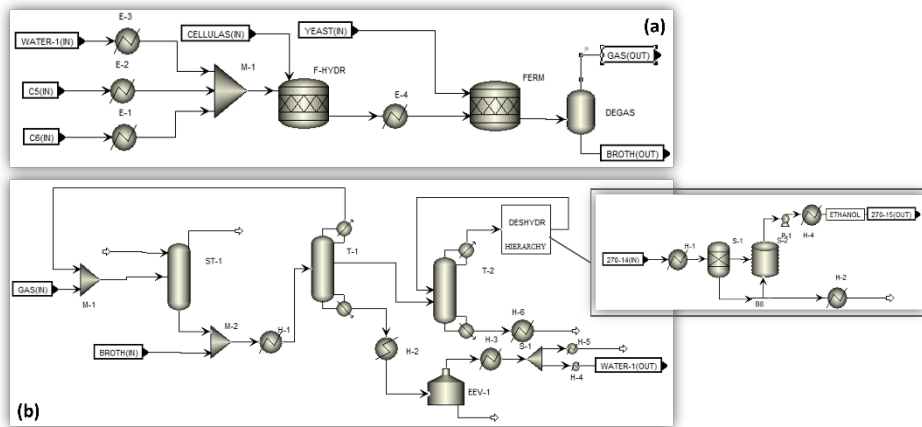


Figure 3 Ethanol production chemistry: Fermentation (a), products recovery (b)

set of alternative biorefining scenarios for analysis.

The case of removing the challenging delignification section, at cost of low-grade C6 sugars production has been examined. Such decision inherently affected mass and energy balances; let alone given the high complexity of the interconnected sections. The reference biorefinery has been appropriately amended and converged to the new balances. As an alternative, bioethanol production has been integrated along with hydrolysis, fermentation and downstream separations systems targeting to the production of high purity ethanol. The biorefinery section simultaneously valorizes the C5 (xylose) and C6 (pulp) fractions and the appropriate chemistry has been integrated (Figure 3). The analysis focused on all alternative combinations of using the delignification and ethanol chemistries resulting in 4 main scenarios with/without delignification, and with/without ethanol production. In addition, the 4 alternative biorefinery designs have been also tested in presence of two alternative feedstocks (corn stover and the reference-wheat straw) to adapt flows, unit operations and productivities to each carbon source. In this scope, 8 alternative biorefinery scenarios, and the corresponding simulations, have been developed and performed establishing a valuable portfolio of test cases for analysis.

The analysis involves energy integration and techno-economics to estimate the new sustainability margins of alternative scenarios. The energy analysis of multiple-process and multiple-product biorefineries is essentially challenged by the complex and variable integration routes (Pyrgakis and Kokossis, 2019). Furthermore, the investigation of alternative chemistries (e.g., bioethanol production) rises questions on the new integration patterns that need to be systematically addressed for all potential biorefinery scenarios. The guidelines of Pyrgakis and Kokossis (2019) have been followed to handle the energy integration alternatives, while the Aspen Process Economic Analyzer was used to estimate the new capital and operating expenses for each case.

#### 4. Results

The process systems engineering analysis of the lignocellulosic biorefinery results a set of 8 alternatives scenarios to compare with the reference biorefinery in terms of capital and operating expenses. To this end, the design options account for (i) the use, or not, of the pulp delignification with sequent costs in the quality and quantity of end-pulp product;

Table 1: Production yields of the 8 scenarios

Scenario	Delignif. section	Ethanol production	Feedstock	Mass flows [tn/hr]			
				C5	C6	lignin	ethanol
Reference	YES	NO	Wheat straw	5.82	15.79	8.96	
2	YES	YES	Wheat straw	0.57	1.48	8.96	9.44
3	NO	NO	Wheat straw	5.67	14.57	8.43	
4	NO	YES	Wheat straw	0.58	1.55	8.43	9.74
5	YES	NO	Corn stover	6.00	8.63	4.13	
6	YES	YES	Corn stover	0.56	0.75	4.13	11.63
7	NO	NO	Corn stover	5.31	7.88	3.75	
8	NO	YES	Corn stover	0.56	0.94	3.75	12.00

Table 2: Energy and economic results of the addressed scenarios

Design strategy	Hot [MW]	Cold [MW]	CAPEX [k€]	OPEX [k€/yr]
Del. (YES) - Eth. (NO)	133 (62%)	148 (54%)	31.8	215.6
Del. (YES) - Eth. (YES)	256 (60%)	64 (55%)	42.4	221.0
Del. (NO) - Eth. (NO)	126 (58%)	141 (50%)	28.9	207.0
Del. (NO) - Eth. (YES)	148 (74%)	57 (53%)	39.6	212.4

(ii) the integration, or not, of bioethanol production replacing profits from biomass C5 and C6 fractions with biofuels, at cost of additional equipment and (iii) the changes in biorefinery economic performance at use of different feedstocks. Accordingly, the production yields of end-products (Table 1) considered in each scenario have been calculated preparing the economic analysis for comparison. The 8 scenarios have been also addressed in terms of alternative energy integration patterns of the revised energy balances across biorefinery resulting in significant savings and cost reductions, as shown in Table 2; the value in parenthesis reflects the savings compared to the non-energy integrated biorefineries justifying advanced economics for all scenarios. In addition, Table 2 also summarizes the capital and operating expenses of each biorefinery scenario.

## 5. Conclusions

As indicated by the presented data, there has been a significant hot utilities reduction (up to 74%), after the successful bypassing of the delignification section and its consequent blocks. In addition, by the design of the downstream ethanol production unit, the insignificant amount of delignified lignin (0.53 tn/hr) and the sugars fractions are replaced by produced ethanol (9.74 tn/hr). With CIMV's starting point referring to a case scenario containing Delignification which was afterward deleted, the "energy needs" reduction was an expected change. Though the ethanol production integration has an apparent increase in CAPEX by 33%, there is a slight effect (3%) on the operating expenses (including the energy integrated costs) favoring payback of the new investment. In this context, the delignification section was observed to have a strong effect of operating costs by € 8.7M/yr, while the also reduces the efficiency of the ethanol production (ethanol production drops by 300kg/hr). As a conclusion, the most promising results are observed in the scenario of replacing the delignification section with ethanol production resulting



in 22.5% lignin and 26% ethanol yields, which are the highest among all cases; CAPEX and OPEX indexes are accordingly estimated to € 39.6M and € 212M/yr. Feedstock seasonality is capable of affecting the biorefinery operations and production capacities by 2-55%, while corn stover appears better productivities and, thus, process economics than wheat straw.

## Acknowledgements

Financial support from the Marie Curie European Research Program RENESENG-II is gratefully acknowledged.

## References

- Hassan, S.S., Williams, G.A., Jaiswal, A.K., 2019. Lignocellulosic biorefineries in Europe: Current State and prospects. *Trends in Biotechnology*, 37, 3, 231–234.
- Pyrgakis, K.A. and Kokossis, A.C., 2019. Systematic synthesis and integration of multiple-effect distillation into overall processes: The case of biorefineries. *AIChE Journal*, 65, 7, e16631).
- Mesa, L., Albornas, Y., Morales, M., Corsano, G., González, E., 2016, Integration of organosolv process for biomass pretreatment in a Biorefinery. *Biomass Fractionation Technologies for a Lignocellulosic Feedstock Based Biorefinery*, 229-254.
- CIMV, France: <http://www.cimv.fr/> (Accessed: December 2, 2022).
- Delmas, G.-H., Benjelloun-Mlayah, B., Le Bigot, Y.s, Delmas, M., 2011. Functionality of wheat straw lignin extracted in organic acid media. *Journal of Applied Polymer Science*, 121, 1, 491-501.
- Nadja, C., Séverinem C., Benjelloun-Mlayah, B., Jean-Stephane, C., Delmas, M., 2014. Esterification of organosolv lignin under supercritical conditions. *Industrial Crops and Products*, 58, 287-297.
- Delmas, G.-H., Benjelloun-Mlayah, B., Le Bigot, Y.s, Delmas, M., 2013. Biolignin™ based epoxy resins. *Journal of Applied Polymer Science*, 127, 3), 1863-1872.
- BBIC-Bio Based Industries Consortum, Mapping European Biorefineries: <https://biconsortium.eu/news/mapping-european-biorefineries> (Accessed: December 2, 2022)
- Khama, L., Le Bigot, Y., Delmas, M., Avignon, G., 2005. *Industrial Crops and Products*, 21, 9-15.
- Mountraki, A.D., Mlayah, B.B., Kokossis, A.C., 2016. A Study on the Endogenous Symbiosis of First and Second Generation Biorefineries: Towards a Systematic Methodology. *Computer Aided Chemical Engineering*, 38, 2181-2186.
- Pateromichelakis, A., Psychas, M., Pyrgakis, K., Maréchal, F., Kokossis, A., 2022, The Use of GVL for Holistic Valorization of Biomass, *Computers & Chemical Engineering*, 164, 107849.
- Pyrgakis, K.A., Kokossis, A.C., 2019, A Total Site Synthesis approach for the selection, integration and planning of multiple-feedstock biorefineries, *Computers & Chemical Engineering*, 122, 326-355.
- Pyrgakis, K.A., de Vrije, T., Budde, M.A.W., Kyriakou, K., López-Contreras, A.M., Kokossis, A.C., 2016, A process integration approach for the production of biological iso-propanol, butanol and ethanol using gas stripping and adsorption as recovery methods, *Biochemical Engineering Journal*, 116, 176-194.
- Kokossis, A.C., Tsakalova, M., Pyrgakis, K., 2015, Design of integrated biorefineries, *Computers & Chemical Engineering*, 81, 40-56.

# Digital-twin development for a novel vibrating membrane aiming at fractionating fermentation broths

Oscar A. Prado-Rubio<sup>a,b</sup>, Wai Fung Hui<sup>b</sup>, Mads Stevnsborg<sup>b</sup>, Manuel Pinelo<sup>b</sup>,  
Jakob Kjøbsted Huusom<sup>b</sup>

<sup>a</sup>*Departamento de Ingeniería Química, Universidad Nacional de Colombia – 170003 Manizales, Colombia*

<sup>b</sup>*Department of Chemical and Biochemical Engineering, Technical University of Denmark (DTU), DK-2800 Lyngby, Denmark*  
*oapr@kt.dtu.dk*

## Abstract

In the near future, tools brought on by digitalization will play a more significant role in technology development and operation for debottlenecking challenges in the biochemical industry. In that regard, a novel vibrating membrane filtration system at pilot scale is currently under investigation to fractionate bio-succinic acid fermentation broth. Remarkably sustained flux has been obtained that is 3.6 times higher than the best results found in literature. Experimental data was used to develop a digital object that combine a robust communication system and hybrid model to predict hydrodynamics and fouling rate. The digital object is used to forecast membrane performance 25 min ahead, predicting the flux with high accuracy (4% error) and the onset of the irreversible fouling formation during step-up experiments. These tools will become the core of a digital twin for the adaptive operation of the ultrafiltration system under uncertainty.

**Keywords:** Digital-twin, Vibrating membrane ultrafiltration, Fractionation fermentation

## 1. Introduction

Extensive research in debottlenecking bioprocesses gains traction as manufacturers commit to more sustainable development practices (Ruales-Salcedo et al. 2022). Digitalization initiatives including advances in process monitoring, data storage, computing power, modelling approaches, and simulation tools join to improve bioprocess development. The concept of integrating digital and physical objects is not revolutionary and has existed for many years in areas such as aeronautics. However, in chemical- and biochemical processes lack of advanced process analytical tools to measure critical compounds and computationally expensive models hindered harnessing digitalization in these fields. Digital shadow and twin have been defined as a system of an interconnected physical process and in-silico/digital replica which have continuous information exchange. In a digital shadow, the virtual plant representation can assess process performance (past and present) to evaluate future scenarios, while the digital twin is capable to determine and implement open or closed loop actions given by an optimal objective function (Gargalo et al. 2021). The heart of a digital-twin is data-driven and hybrid modelling that has a great potential to exploit a priori system understanding

complemented by nonparametric models. This approach is particularly relevant in emerging technologies where there is a lack of understanding of the underlying phenomena, thus, a data-driven or hybrid approaches could provide a more accurate system description (Azevedo et al. 2019).

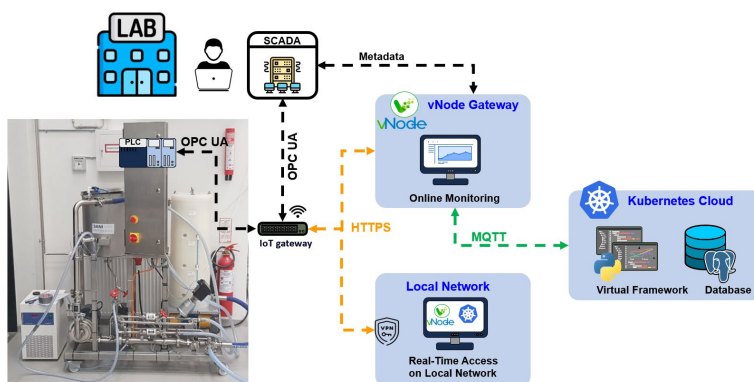
This research is focused on optimizing downstream process for the fractionation of succinic acid fermentation broth. Succinic acid bioproduction is interesting since it is one of the most important building blocks in the biorefinery concept (Ruales-Salcedo et al. 2022). Despite that succinic acid bioproduction is already operated at a commercial scale, there is still has room for improvements to reach an even larger scale. In this context a novel vibrating membrane ultrafiltration system at pilot scale is currently under investigation as the first stage of a possible downstream processing. This step aims to separate cells, cell debris and proteins from the main dissolved organic acids (i.e., succinic, formic, lactic, and acetic). This separation is particularly challenging and conventional technologies like centrifugation reports around 90% recovery of cells and only 50% of proteins (Wang et al. 2013). Membrane ultrafiltration has been proposed as an alternative to obtain better separation performance with the possibility of integrating a membrane bioreactor. Despite the promising components recovery (i.e., 100% cells and 83% proteins), there is up to 75% throughput loss due to membrane fouling. A novel vibrating membrane technology is a possible alternative to conventional membrane ultrafiltration where fouling is mitigated by inducing a relatively high shear rate at the membrane surface (Prochaska et al. 2018). In that way, operation at higher throughput is attempted while reducing energy consumption. However, determining and understanding the maximum membrane throughput (so called critical flux), requires extensive experiments due to time-variant fouling nature plus potential inlet disturbances. Therefore, the design and optimal operation of the system will benefit by developing a digital twin aimed to predict the hydrodynamics and irreversible fouling onset. This research aims to present the challenges on building a digital twin for the vibrating membrane system including communication architecture, experimental campaign, digital object development and evaluation. The main outcome of this research is to build a tool which can estimate the performance of the novel membrane system operation and employ the model to optimize the energy demanding operation improving the sustainability of the process.

## 2. Methodology

### 2.1. Digitalization infrastructure

Last year, the Department of Chemical- and Biochemical Engineering at the Technical University of Denmark (DTU) heavily retrofitted the existing digital infrastructure aiming to centralize data collection, process monitoring and model validation activities across the portfolio of pilot scale laboratory equipment (Jones et al. 2022). A representation of the communication infrastructure is depicted in Figure 1. Process data and logged at-line tags from each unit are now forwarded to an IoT gateway software, vNode that facilitates online bidirectional communication to the process equipment through a secure connection to the DTU network. A Kubernetes cluster has been implemented in parallel to store historical data and contain PSE and modelling tools. Kubernetes is a container management architecture that manages a collection of independent internally connected server entities. One of the entities are designated as the “master” node, while the

remaining entities are categorized as “worker” nodes (Gupta et al. 2021). The master-node orchestrates the available hardware across the collection of entities and designates resources depending on the demand of active applications. Local virtualization makes it possible to simulate several server entities and orchestrations on a local machine to test application deployment dataflow and database schemas before large scale implementation. vNode ensures that the OPC-UA signal received from the units are routed to the Kubernetes cluster during regular operation. Additionally, a relational SQL database standardized data collection for the process units and contextualizes data across different experimental runs. The cluster also acts as a deployment domain for containerized model application that can be used for prediction, control, and monitoring in real-time. Easy access to the process real-time data and historian improves the capabilities of utilizing data-driven methods and real-time optimization for these systems. The vibrating membrane module is being integrated with this new infrastructure.



**Figure 1.** Schematic of the connectivity to the digital infrastructure at DTU Chemical and Biochemical Engineering. The OPC-UA signal from the PLC cabinet is sent to the IoT gateway software to be joint with metadata from offline measurements thus send to Kubernetes cluster where the tag values are stored and online accessible for monitoring or modelling tasks.

## 2.2. Experimental campaign

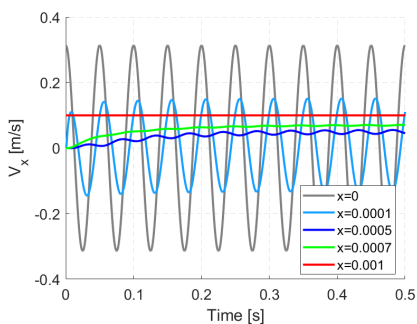
The equipment is a 7.5 m<sup>2</sup> Vibro Pilot membrane separation unit from SANI Membranes. For the separation, a polyethersulfone membrane (PES) is employed with a Molecular Weight Cut-Off of 30 KDa, which is enough to retain cells and most of the expected protein. Due to the experiment’s scale, a considerable amount of fermentation broth is required. To circumvent this limitation, a simulated fermentation broth is used for tuning purposes. The simulated fermentation broth composition is extracted from literature for continuous operation including biomass (8 g/l yeast cells), proteins (0.25 g/L BSA) and organic acids (25 g/l succinic, 0.5 g/l acetic, 0.5 g/l formic acids). The experimental strategy to find the operating point is the well-known step-up method at constant transmembrane pressure. The feed channel pressure is increased stepwise, while the system pressures, flowrates and temperatures are monitored every 15 s. Because of the fouling, the membrane flux is expected to decline with time. According to the resistance in series model, the flux decline with time corresponds to the fouling rate which determine when the membrane should be cleaned. The influence of the crossflow velocity and fermentation broth composition are currently under investigation.

### 2.3. Developing the digital object

A multipurpose digital twin for design and control, consists of multiple digital objects that can combine different modelling approaches (i.e., first principles, data-driven and hybrid)(Grisales Díaz and Prado-Rubio 2022; Prado-Rubio, Grisales Diaz, and Kjøbsted Huusom 2022). Herein, a hybrid model is proposed with the potential to incorporate forthcoming process understanding provided by a more dedicated first principles models (e.g., CFD) and experiments. As initial step, the digital object approach uses a parallel implementation of a hydrodynamic and fouling rate models. The system hydrodynamics is modelled using the unsteady laminar flow model near an oscillating plate. This model is used to predict velocity profiles and shear rate in the zone adjacent to the membrane surface. This information is relevant to compare model prediction with literature. A MIMO ARX model with forgetting factor is then tuned using the online recursive least squared method, where the model structure and forgetting factors are defined by a sensitivity analysis using MAPE, FIT and NRMSE as performance indexes. This model is calibrated to predict the membrane flux and transmembrane pressure based on the pressure and recovery setpoints plus the crossflow rate. Thus, the model is used to forecast 25 min in the future the flux decline and its trend which represent the dynamic fouling rate.

## 3. Results

To date, five campaigns have been conducted including 10-15 experiments each. The results of the experimental campaigns showed remarkable performance where the vibrating membrane module achieved a maximum sustainable operation of around 53 LMH at TMP of 2 bars, and up to 100%, 89.4% and ~25% rejection of yeast, BSA and organic acids, respectively. Previous research has shown that ultrafiltration membrane with analogous properties and operating conditions achieved 14.74 LMH at laboratory scale (Wang et al. 2013), which represent a 3.6 fold improvement at a more larger scale with the potential to be further optimized. The best result from the campaigns is herein used for the modelling tuning and forecasting analysis. The hydrodynamic modelling results for the conventional operation show that the velocity profile decreases, as expected, linearly towards the membrane surface where the shear rate is zero.

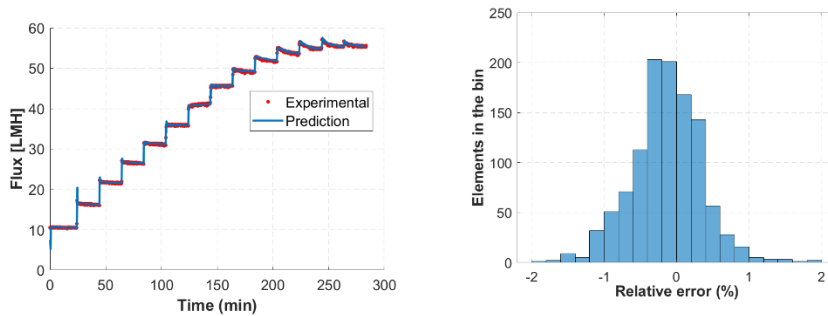


**Figure 2.** Velocity profiles for the conventional crossflow ultrafiltration and vibrating membrane operation at 20 Hz and an oscillation amplitude of 2.5 mm.

Figure 2 shows the velocity profiles at the boundary layer. The oscillating membrane at 20 Hz and with 2.5 mm amplitude generates a relative velocity at the membrane surface up to 0.31 m/s, which is 1.5 to 3 times higher than the crossflow. This corresponds to an average surface shear rate of 2239.5 1/s to mitigate fouling. Previous research with vibrating microfiltration of yeast and BSA at 20 Hz and 1.375 mm amplitude, generated a shear rate of 1230 1/s (Prip Beier & Jonsson, 2009). Under those conditions, a sustained flux of 30 LMH was obtained. The current results show how almost doubling shear rate has substantially enhanced flux.

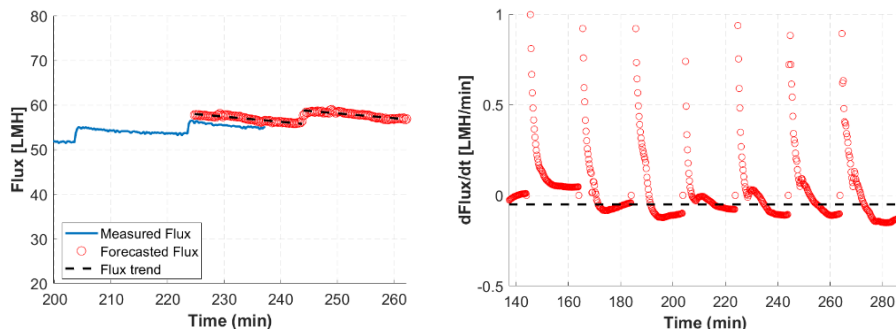
The increased shear rate appeared to be more efficient in removing the pore blocking and cake-controlled fouling expected from fermentation broths.

The adaptive online system identification model was able to capture the time variant flux nature with high accuracy (Model quality: NRMSE = 0.0092, MAPE = 0.57% and FIT = 97.7%). Experimental data and model predictions are shown in Figure 3, where 97% of data have relative error lower than 1%.



**Figure 3.** Online system identification predictive power during 14 experiments. Flux prediction and histogram of the relative model error

The model is also used in real time to forecast 100 samples into the future (25 min) and estimate the flux decline and fouling rate. A zoom-in is depicted in Figure 4. Note that there is a small deviation of ~4% in the flux forecast. Despite that the step-up disturbances during the experiment are mild (i.e., 0.2 bar at the time), there is a tremendous effect on the fouling rate which tends to stabilize after 10 min. This result is important since it defines that the experimental tuning should last at least 20 min to stabilize the fouling when vibration is on. Now, defining an affordable fouling rate of -0.05 LMH/min (corresponding to a 50% of max flux decline in 8 h operation), it can be predicted that experiments beyond 190 min (exp 10) the system consistently overcomes the threshold. Every step further leads to a higher fouling rate where the last represent 2.4 times the desired target. All these points then represent super-critical operation and should be avoided. This tool can then predict the onset of the irreversible fouling formation during experiments which will be tested in the near future when the full digitalization infrastructure is implemented.



**Figure 4.** Moving horizon flux and fouling rate forecasting 25 min ahead using the last estimated ARX model

## Conclusions

The department of Chemical and Biochemical Engineering at DTU is designing a robust pilot plant infrastructure capable of coping with the current and forthcoming challenges of biomanufacturing using a digitalization strategy. Herein, the development of a digital twin for a novel vibrating membrane is elaborated upon, discussing communication, experimental campaigns, and modelling approaches. This technology has shown remarkable performance during the fractionation of succinic acid fermentation broth by increasing the membrane flux 3.6 times compared to the best reported conventional ultrafiltration. Employing an online system identification, it is possible to predict the flux behaviour with high accuracy (MAPE = 0.57%). This was used to forecast the onset of irreversible fouling formation and detecting 4 super-critical experiments. The developed digital object plus the digitalization infrastructure will become the core of the implementation of digital twin for process optimization and adaptive system operation under uncertainty.

## Acknowledgements

The authors of this work would like to thank the Novo Nordisk Foundation for their support of this project through grant NNF19SA0035474.

## References

- Azevedo, Cristiana Rodrigues et al. 2019. "Hybrid Semiparametric Modeling: A Modular Process Systems Engineering Approach for the Integration of Available Knowledge Sources." In *Systems Engineering in the Fourth Industrial Revolution*, eds. Ron S. Kenett, Robert S. Swarz, and Avigdor Zonnenshain. Wiley, 345–73.
- Gargalo, Carina L. et al. 2021. "Towards the Development of Digital Twins for the Bio-Manufacturing Industry." *Advances in biochemical engineering/biotechnology* 176.
- Grisales Diaz, Victor Hugo, and Oscar Andrés Prado-Rubio. 2022. "Flux Estimation of Dynamic Ultrafiltration for Wastewater Treatment Using Combinations of Machine Learning Methods." In *Euromembrane 2022*, Sorrento, Italy.
- Gupta, Manu, Konte Sanjana, Kontham Akhilesh, and Mandepudi Nobel Chowdary. 2021. "Deployment of Multi-Tier Application on Cloud and Continuous Monitoring Using Kubernetes." In *2021 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques, ICEECCOT 2021 - Proceedings*.
- Jones, Mark Nicholas et al. 2022. "Pilot Plant 4.0: A Review of Digitalization Efforts of the Chemical and Biochemical Engineering Department at the Technical University of Denmark (DTU)." *Computer Aided Chemical Engineering* 49: 1525–30.
- Prado-Rubio, Oscar Andrés, Victor H.; Grisales Diaz, and Jakob Kjøbsted Huusom. 2022. "A Platform for Online System Identification of Dynamic Ultrafiltration Systems." In *Computer Aided Chemical Engineering*, , 1075–80.
- Prochaska, K., J. Antczak, M. Regel-Rosocka, and M. Szczygielda. 2018. "Removal of Succinic Acid from Fermentation Broth by Multistage Process (Membrane Separation and Reactive Extraction)." *Separation and Purification Technology* 192.
- Ruales-Salcedo, Angela V. et al. 2022. "Production of High-Added Value Compounds from Biomass." *Biofuels and Biorefining: Volume 1: Current Technologies for Biomass Conversion*: 381–445.
- Wang, Caixia et al. 2013. "Clarification of Succinic Acid Fermentation Broth by Ultrafiltration in Succinic Acid Bio-Refinery." *Journal of Chemical Technology & Biotechnology* 88(3): 444–48. <https://onlinelibrary.wiley.com/doi/full/10.1002/jctb.3834> (November 30, 2022).

# Post-Combustion Carbon Capture and Utilization by Sodium Hydroxide Aqueous Solution for Bicarbonate Microalgae Cultivation

Jialin Liu,<sup>a\*</sup> Yoong Kit Leong,<sup>a</sup> Jo-Shu Chang<sup>a</sup>

<sup>a</sup>*Research Center for Smart Sustainable Circular Economy, Tunghai University,  
No. 1727, Sec.4, Taiwan Boulevard, Taichung, Taiwan  
[jialin@thu.edu.tw](mailto:jialin@thu.edu.tw)*

## Abstract

The increasing amounts of CO<sub>2</sub> in the atmosphere, which is caused by the burning of fossil fuels, contribute greatly to global warming. In a standard post-combustion carbon capture (PCC) process, the energy penalty attributes to regenerate the CO<sub>2</sub> lean solvents. Microalgae have been established as the most promising candidate to the CO<sub>2</sub> capture and utilization (CCU) technology due to a high growth rate, oil productivity, and photosynthetic efficiency. However, the gaseous form of CO<sub>2</sub> is utilized by microalgae only to a very limited degree due to its low water solubility and high outgassing rate. In addition, the flue gases from a coal-fired power plant, which contains 10% to 16% of CO<sub>2</sub>, cannot be supplied as the carbon source for the microalgae, directly. The bicarbonate-based microalgae cultivation may offer a promising alternative that relieves the burden of regeneration energy and improves the low carbon utilization efficiency of the gaseous form of CO<sub>2</sub> cultivation. In the present work, the rate-based model of Aspen Plus is adopted to simulate a flue gas in the flow rate of 1800 m<sup>3</sup>/h with 14% CO<sub>2</sub> that is absorbed by aqueous NaOH solution. The absorbent is recycled from the mother liquor where most of bicarbonate is precipitated, and the fresh makeup with 30 wt% NaOH is applied. The column height to achieve 90% of CO<sub>2</sub> removal rate is investigated by varying the recycle rate of absorbent and the solid fraction in the bicarbonate product. In addition, the carbon utilization efficiency (CUE) by gaseous CO<sub>2</sub> and bicarbonate solution for the microalgae cultivation is compared. The results show that the CUE of the bicarbonate-based cultivation is far superior to that of the counterpart.

**Keywords:** carbon capture and utilization, chemical absorption, rate-based model, bicarbonate-based microalgae cultivation.

## 1. Introduction

In 2015, the United Nations General Assembly (UNGA) set out 17 sustainable development goals (SDGs) to be achieved by 2030. Microalgae have an increasing interest in various sectors in the SDGs, such as SDG-6 “clean water and sanitation”, SDG-7 “affordable and clean energy”, and SDG-13 “climate action”. The microalgae can be cultivated in wastewater or seawater that provides a way for water reclamation (Wang et al., 2015). Moreover, microalgae could complete the entire growth cycle via photosynthesis reactions that convert light energy into renewable energy. The CO<sub>2</sub> emitted from the power generation and heavy industries is the largest contributor to the global warming problem. Vuppaladadiyam et al. (2018) reviewed the technical feasibility



of combined carbon fixation and microalgae cultivation for carbon reuse. They concluded that the rate of carbon fixation by microalgae tends to be too slow to capture CO<sub>2</sub> from flue gases. Unless the rate of CO<sub>2</sub> uptake can be drastically improved, culturing algae may not be suited for the post-combustion carbon capture (PCC) from the power generation and heavy industries.

Amine scrubbing has been used to separate acid gas, such as H<sub>2</sub>S, SO<sub>2</sub>, and CO<sub>2</sub>, from the flue or tail gases since 1930 (Bottoms, 1930). It is a robust technology and is ready to be tested and used on a larger scale for CO<sub>2</sub> capture from coal-fired power plants. However, the high regeneration energy of the amine-based solvents is the most critical challenge for extensively applying the PCC technology to thermal power plants. It is well known that the improvements of process configuration and solvent development may reduce the regeneration energy that is around 4.0 GJ/t-CO<sub>2</sub> for a standard PCC process configuration using 30 wt% monoethanolamine (MEA) aqueous solution (Rochelle, 2009). Solvent development has historically focused on reducing the enthalpy of absorption, increasing the absorption capacity, and enhancing the reaction kinetics by blending and synthesizing new amines. Recently, Benquet et al. (2021) reported that the second generation of amine-based solvent, named CO<sub>2</sub> Enhanced Separation and Recovery (CESAR1), was tested at large scale with a real flue gas at the pilot plants in Technology Centre Mongstad (TCM). The CESAR1 solvent, which is a blend of 27 wt% amino-methyl-propanol (AMP) and 13 wt% piperazine (PZ), that poses the lower regeneration energy (3.0 GJ/t-CO<sub>2</sub>), higher stability, and lower corrosivity compared to those of traditional 30 wt% MEA solvent. The energy reduction of 25% can be achieved by blending amine-based solvents in the standard PCC process configuration. On the other hand, Le Moulec et al. (2014) reviewed available literature as well as the patent database and summarized 20 elementary modifications to save regeneration energy for the PCC. These methods can be categorized into three groups: absorption enhancement, heat integration, and heat pump. According to the survey by Le Moulec et al. (2014), up to 39.0% of energy savings could be achieved by modifying the PCC process configuration.

Recently, Zhu et al. (2022) suggested the integration of the bicarbonate-based microalgae cultivation with the carbon capture that significantly reduces the production costs of cultivation and the regeneration energy in the PCC. The authors estimated that the cost of bubble CO<sub>2</sub> gas for microalgae production ranges from \$1.47 to \$7.33 kg<sup>-1</sup> and the carbon utilization efficiency (CUE) is in the range of 1.0% to 5.0%. In contrast, the carbon cost of bicarbonate-based *Spirulina* production is around \$0.359 kg<sup>-1</sup> due to the high CUE around 90%, particularly. Kim et al. (2017) reported that only 3.59% of the gaseous CO<sub>2</sub> was used for biomass synthesis, bicarbonate was effectively incorporated into the biomass with 91.40% of CUE. The CUE of bicarbonate-based cultivation is significantly superior to that of the gaseous CO<sub>2</sub> because CO<sub>2</sub> supplied in a gaseous form easily escapes from the culture medium, rather than being captured and utilized by the cell.

Wang et al. (2022) developed a PCC process by chemical absorption using aqueous sodium glycinate solutions (SGS) to produce high-value NaHCO<sub>3</sub> nanomaterials. CO<sub>2</sub> was captured from a split flue gas stream emitted from the 600 MWe coal-fired power plant in the Wolverine Clean Energy Venture project. The operating expenditure (OPEX) and capital expenditure (CAPEX) were calculated as \$27.53 and \$7.96 per metric ton of CO<sub>2</sub>, respectively. In this study, aqueous NaOH solvents are applied to capture CO<sub>2</sub> to form the precipitation of NaHCO<sub>3</sub>, which can be diluted for bicarbonate-based microalgae cultivation. Subsequently, the PCC process design is briefed in Section 2. Section 3 compares the experimental results of cultivating a variant strain of *Chlorella vulgaris* by gaseous CO<sub>2</sub> and bicarbonate, respectively. Conclusions are presented in Section 4.

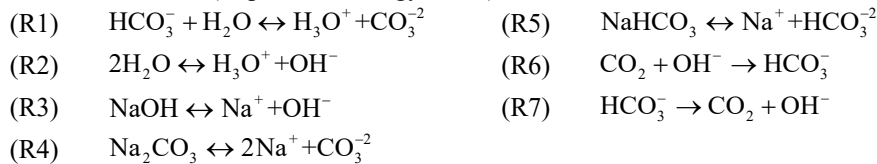
## 2. Process Design for Bicarbonate Precipitation

The rate-based model of Aspen Plus is applied to design an absorber to achieve 90% of CO<sub>2</sub> removal rate from flue gases by sodium hydroxide solutions. The Redlich-Kwong equation of state and the electrolyte-NRTL method were used to compute the properties of the vapor and liquid phases, respectively. In addition, all the ionic reactions and solid dissociation can be assumed in equilibrium reactions, as listed in (R1) – (R5), where the equilibrium constants are calculated from the Gibbs free energy changes. In the rate-based model, the mass transfer rate of CO<sub>2</sub> diffused into the liquid phase needs to be considered; therefore, the rate-controlled reactions of CO<sub>2</sub> with OH<sup>-</sup> are concerned as following (R6) and (R7). The kinetic reaction rate is described by the reduced power law:

$$r = k \exp(-E/RT) \prod_{j=1}^N (C_j)^{a_j}$$

where the concentration basis is Molarity,  $N$  is the number

of components and  $a_j$  is the stoichiometric coefficient of component  $j$ ; the rate constants of  $k$  and  $E$  are listed in Table 1 (Aspen Technology, 2014). The parameter settings for the rate-based model, such as Film resistance and Flow model options can be found in the technical document (Aspen Technology, 2014).



$r$ (kmol·m <sup>-3</sup> ·s <sup>-1</sup> )	$k$	$E$ (J·kmol <sup>-1</sup> )
(R6)	$4.32 \times 10^{13}$	55.4
(R7)	$2.83 \times 10^{17}$	123.2

The flow rate of the flue gas from a coal-fired boiler was 1800 m<sup>3</sup>/h, containing 14% CO<sub>2</sub> and the saturated water vapor at 40 °C. Assuming that SO<sub>x</sub> in the flue gas was removed before entering the CO<sub>2</sub> absorber, the reaction of forming sodium sulfite can be neglected. In this study, the fresh makeup of sodium hydroxide solution was set to 1500 kg/h, which contains 30 wt% NaOH and the balanced water. The column diameter was determined by calculating 80% of the maximum flood capacity, so that the diameter of 0.85 m was applied. The column height was adjusted to fulfill the CO<sub>2</sub> removal rate at 90%. Figure 1(a) shows the process flow diagram for the PCC process where the flue gas and lean solvent enter the absorber from the bottom and top, respectively. The outlet temperatures of the absorber are higher than those of inlets, because the CO<sub>2</sub> absorption is an exothermic reaction. The temperature of the settling tank was maintained at 20 °C for the NaHCO<sub>3</sub> precipitates, in which the recycled flow rate and the weight fraction of NaHCO<sub>3</sub> drawn in the outlet flow are two adjustable variables. Figure 1(b) shows the packed height that needs to maintain the CO<sub>2</sub> removal rate at 90% under varying the recycle-to-outlet ratio (Recycle/Outlet) and the solid fraction (SF) of precipitates in the outlet flow. The results show that the packed height needed is under 2 m once the SF is larger than 0.7, whereas the operating condition of the Recycle/Outlet ranges from 0.5 to 3.0 that is a robust design for operations. In contrast, the feasible operating range of Recycle/Outlet limits from 0.8 to 2, once the SF is operated at 0.5, as shown in Figure 1(b), where the packed height is between 2 to 2.5 m. Figure 1(a) show the simulation results of the ratio of Recycle/Output at 1.8 and the SF operated at 0.7. The outlet flow contains 28 wt% of NaHCO<sub>3</sub>, whereas 6.6 wt% of bicarbonate in the recycle flow can be dissolved after

mixing with the makeup flow, by which the lean temperature after mixing was risen due to the exothermic reactions of mixing the recycle flow and the fresh makeup.

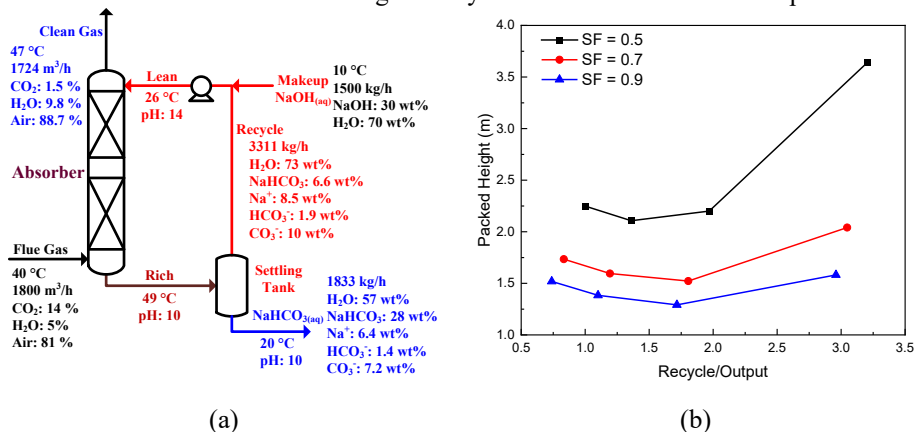


Figure 1. The simulation results of the makeup flowrate being 0.7 times flue gas, (a) process flow diagram and (b) the packed heights under varying the designed parameters. In this study, the makeup-to-feed flow ratio was set as 0.7 (kg/kg) for the proof of concept, which microalgae could be cultivated by captured CO<sub>2</sub> through chemical absorption. The CAPEX, OPEX, and levelized cost of carbon (LCOC) will be investigated under varying the makeup-to-feed ratio, whereas the robust operations are also taken into consideration for designing the PCC process in the future work.

### 3. Carbon fixation by Bicarbonate-Based Microalgae Cultivation

The microalga used in this study (*Chlorella vulgaris* SU-1) is a variant strain of *Chlorella vulgaris*, which is similar to that of the study by Wang et al. (2015). BG-11 medium was used as preculture medium with continuously supplied carbon dioxide (2%) at a flow rate of 0.2 L/min for the baseline cases. On the other hand, 5 g/L NaHCO<sub>3</sub> were supplemented as the carbon source at the beginning of cultivation and the 4<sup>th</sup>, 8<sup>th</sup>, and 12<sup>th</sup> day after sampling. The microalga was cultivated in a 1-L glass photobioreactor illuminated by external light sources (white fluorescent lamps) mounted on both sides of the photobioreactor at a light intensity of 200 W/m<sup>2</sup>, and the room temperature was controlled at 25 ± 2 °C. A volume of 10 mL microalgae suspension was collected every day from each photobioreactor for nutrient analysis. Meanwhile, the pH value for each photobioreactor was adjusted to 7, 8, and 9, respectively, by adding hydrochloric acid on a daily basis for the bicarbonate-based cultivation. On the other hand, the pH value for cultivating microalgae by gaseous CO<sub>2</sub> was not controlled. The daily samples were centrifuged at 7000 rpm for 2 min. Dry cell concentrations were obtained based on the optical density (OD) values measured at 685 nm with a spectrophotometer (model U-2001, Hitachi, Tokyo, Japan). Figure 2(a) show the calibration curve of OD<sub>685</sub>, in which the R<sup>2</sup> is 0.98, that was used to calculate the biomass concentration by the measured OD values. Figure 2(b) show the measured pH values during the experimental period where the pH values in the culture medium by the gaseous CO<sub>2</sub> were not controlled and the averaged pH value is around 8.5. The pH values of the other counterparts were measured and controlled to the corresponding values on the daily basis.

The growth curves are shown in Figure 3(a) where the biomass concentration ( $X$ , g/L) was calculated by the measured OD<sub>685</sub> values and the calibration curve in Figure 2(a). Figure 3(a) shows that the growth curves of the bicarbonate-based cultivation were

inferior to the gaseous CO<sub>2</sub> cases; the final biomass concentration under the gaseous CO<sub>2</sub> condition was over four times as high as those by the bicarbonate-based cultivation. In addition, the figure shows that the pH values maintained at 8 may favor the bicarbonate-based cultivation for this algae strain by contrasting to pH values controlled at 7 and 9. The biofixation of CO<sub>2</sub> can be calculated through the biomass concentration (Adamczyk et al., 2016):  $CO_2 \text{ fixation} = C \times X \times (M_{CO_2} / M_C)$  where  $C$ ,  $M_{CO_2}$  and  $M_C$  are the carbon content in the biomass and the molecular weights of CO<sub>2</sub> and C, respectively. In this study, the carbon content of 47.1 wt% was applied for the strain of *Chlorella vulgaris* (Silva et al., 2016). Figure 3(b) shows the carbon utilization efficiency for each photobioreactor, in which the efficiency for the gaseous CO<sub>2</sub> cases is not over 6%, because the CO<sub>2</sub> gas was supplied at the steady flow of 0.2 L/min. Over 90% of CO<sub>2</sub> were slipped by directly exposing the gaseous form of carbon source that is impractical on the PCC perspective. On the other hand, the average CUE of the bicarbonate-based cultivation was around 40%, as shown in Figure 3(b); in some cases of the pH controlled at 8, the efficiency could be over 60%. In this study, the CUE of the bicarbonate-based cultivation is not as promising as the work of Kim et al. (2017) reported, which is over 90%, because the final biomass concentration was too low for the cases by bicarbonate cultivation. Kim et al. (2017) reported that the final biomass yields by gaseous CO<sub>2</sub> and bicarbonate were comparable.

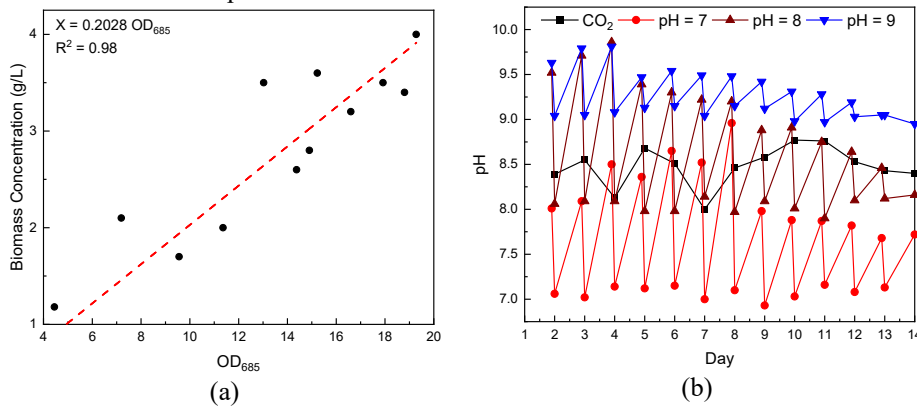


Figure 2. Cultivation of *Chlorella vulgaris* SU-1 with pH controlled, (a) calibration curve of OD<sub>685</sub> to biomass concentration, (b) measured pH values.

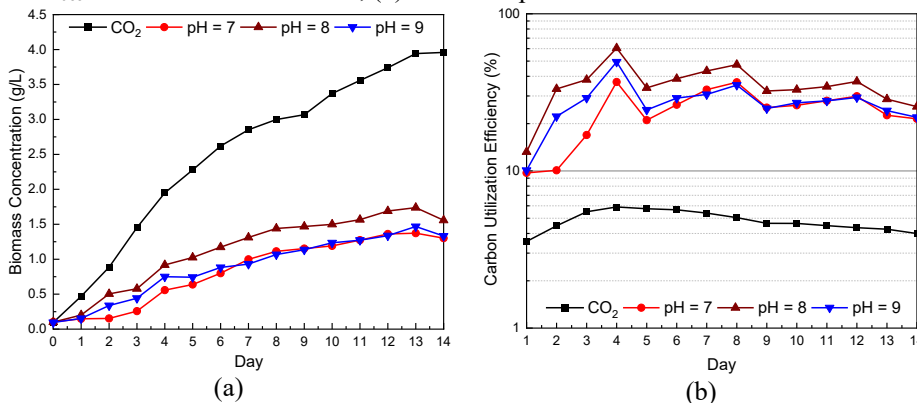


Figure 3. Growth curves of *Chlorella vulgaris* under different carbon sources, (a) biomass concentration (g/L) and (b) carbon utilization efficiency (%).

In this preliminary study, the biomass yield by bicarbonate is inferior to that by the gaseous form of carbon source. The alkaline culture medium may suppress the *Chlorella vulgaris* strain growth, although the pH values of the bicarbonate-based photobioreactors were maintained on a daily basis. However, the CUE of the bicarbonate-based cultivation is still superior to the counterpart. In the future work, the amount and frequency of the supplemented  $\text{NaHCO}_3$  will be adjusted to enhance the productivity. More importantly, adopting a strain of microalgae that is resilient to the alkaline culture medium may improve the final biomass concentration, whereas increases the CUE.

#### 4. Conclusions

Microalgae cultivation to recycle material and energy represents a significant circular economic approach. However, the flue gases from a power plant cannot be directly supplied to cultivate microalgae due to the high concentration of  $\text{CO}_2$  in the flue gases. On the other hand, the energy penalty for regenerating absorbents, which is around 15%–20% reduction in overall electricity, is a major obstacle to implement the PCC for a coal-fired power plant. This study demonstrated that the penalty-free PCC through bicarbonate-based microalgae cultivation is a feasible solution for the circular economy. However, the results show that the biomass yield of bicarbonate-based cultivation is inferior to that of using gaseous  $\text{CO}_2$  as the carbon source. The limitation will be investigated by cultivating the microalgae that can be found in hypersaline environments. In addition, the simulation results show that the robust design for the absorber can be achieved, for which the designed column height is insensitive to the recycle flow rate, by removing sodium bicarbonate from the settling tank as possible.

#### References

- M. Adamczyk, J. Lasek, A. Skawińska, 2016,  $\text{CO}_2$  biofixation and growth kinetics of *Chlorella vulgaris* and *Nannochloropsis gaditana*, *Appl. Biochem. Biotechnol.*, 179, 1248-1261.
- Aspen Technology, 2014, Rate-based model of the  $\text{CO}_2$  capture process by NaOH using Aspen Plus.
- C. Benquet, A. B. N. Knarvik, E. Gjernes, O. A. Hvidsten, E. Romslo Kleppe, S. Akhter, 2021, First process results and operational experience with CESAR1 solvent at TCM with high capture rates (ALIGN-CCUS Project), In *Proceedings of the 15th Greenhouse Gas Control Technologies Conference*, 15-18.
- R. R. Bottoms, 1930, Separating acid gases, U.S. Patent 1783901.
- G. Y. Kim, J. Heo, H. S. Kim, J. I. Han, 2017, Bicarbonate-based cultivation of *Dunaliella salina* for enhancing carbon utilization efficiency, *Bioresour. Technol.*, 237, 72-77.
- Y. Le Moullec, T. Neveux, A. A. Azki, A. Chikukwa, K. A. Hoff, 2014, Process modifications for solvent-based post-combustion  $\text{CO}_2$  capture, *Int. J. GreenH. Gas Control*, 31, 96-112.
- G. T. Rochelle, 2009, Amine scrubbing for  $\text{CO}_2$  capture, *Science*, 325, 1652–1654.
- C. M. Silva, A. F. Ferreira, A. P. Dias, M. Costa, 2016, A comparison between microalgae virtual biorefinery arrangements for bio-oil production based on lab-scale results, *J. Clean. Prod.*, 130, 58-67.
- A. K. Vuppaladadiyam, J. G. Yao, N. Florin, A. George, X. Wang, L. Labeeuw, Y. Jiang, R. W. Davis, A. Abbas, P. Ralph, P. S. Fennell, M. Zhao, 2018, Impact of flue gas compounds on microalgae and mechanisms for carbon assimilation and utilization, *ChemSusChem*, 11, 334-355.
- R. Wang, H. E. Ashkanani, B. Li, B. I. Morsi, 2022, Development of an innovative process for post-combustion  $\text{CO}_2$  capture to produce high-value  $\text{NaHCO}_3$  nanomaterials, *Int. J. Greenh. Gas Control.*, 120, 103761.

- Y. Wang, W. Guo, H. W. Yen, S. H. Ho, Y. C. Lo, C. L. Cheng, N. Ren, J. S. Chang, 2015, Cultivation of *Chlorella vulgaris* JSC-6 with swine wastewater for simultaneous nutrient/COD removal and carbohydrate production, *Bioresour. Technol.*, 198, 619-625.
- C. Zhu, S. Chen, Y. Ji, U. Schwaneberg, Z. Chi, 2022, Progress toward a bicarbonate-based microalgae production system, *Trends Biotechnol.*, 40, 180-193.



# Comparing different modelling approaches for metabolic network dynamic simulation under uncertainty

Oliver Pennington,<sup>a</sup> Dongda Zhang,<sup>a</sup>

<sup>a</sup> *Department of Chemical Engineering, University of Manchester, Oxford Road, Manchester, M1 3AL, UK.*

[Dongda.zhang@manchester.ac.uk](mailto:Dongda.zhang@manchester.ac.uk)

## Abstract

Modelling strategies in metabolic engineering cover a broad spectrum to simulate microbial cell behaviour and product formation. Understanding differences and limitations of various approaches is key for accurately simulating metabolic network activities and designing high-efficiency strains. One of the main challenges lies within the simulation of dynamic environments, where steady-state assumptions are invalid. However, modelling a dynamic operation is essential for developing batch processes, which are immensely common for fermentation and pharmaceutical industries. This study is to propose and compare different dynamic modelling approaches for metabolic network simulation. The overall aim is to identify the most cost-effective and robust approach to model a metabolic network under non-steady-state conditions.

**Keywords:** Dynamic metabolic flux analysis, kinetic modelling, metabolic network simulation, parameter estimation, dynamic modelling.

## 1. Introduction

Microbial fermentation modelling lies at the epicenter of biochemical engineering, as such systems can produce renewable biofuels, bioplastics, and many other high value bioproducts, and mammalian cell lines account for a large proportion of therapeutic recombinant proteins. The UK bioeconomy alone is worth around £220 billion (as of 2018), and is expected to double by 2030, making this a lucrative research field (Harrington, 2018). Such growth requires overcoming many challenges, including low yields in reactor scale-up, deficient metabolic and secretory phenotype for protein production, accumulation of by-products ( increasing separation costs and product loss), and finally significant variation between batches, leading to quality-control challenges.

Understanding the metabolic pathways is essential to conquering such challenges for two reasons: 1) to develop industrially desired microbial strains for large scale fermentation; and 2) modify mammalian cells to overexpress recombinant proteins and improve process performance. Several modelling techniques can be employed to infer the behaviour of the metabolic network. These include two main approaches: 1) flux-based methods, which can be applied to both large or small-scale networks to either steady-state or dynamic systems; and 2) mechanistic kinetic models, which try to establish a mechanistic understanding of the behaviour of each reaction pathway. However, there are challenges associated with the various modelling approaches. For example, flux-based



methodologies often utilize a fundamental steady-state assumption, as reviewed by Ahn et al. (2012), that is invalid for batch and fed-batch operation and cannot account for intracellular metabolite accumulation. Although existing dynamic MFA approaches can encapsulate the changing extracellular environment, most methodologies do not account for any intracellular accumulation, constant or dynamic (Antoniewicz, 2013). Regarding kinetic modelling, finding an appropriate model structure can be time-consuming and challenging – many structures also exhibit stiffness, causing trouble with parameter estimation. This study aims to propose a new dynamic metabolic flux analysis framework and then compare with other existing methods to determine the most reliable and robust (least susceptible to simulation deviations for a given input deviation) modelling approach, while considering minimal computational expense.

## 2. Methodology

### 2.1. Kinetic modelling

The first approach to be considered is the kinetic modelling, in which each reaction flux obeys the Michaelis-Menten model structure as shown in Eq. 1, as described in the study by (Goffaux et al., 2017):

$$v = v_{f,max} \prod_f \alpha_f \frac{R_f}{K_{R,f} + R_f} - v_{r,max} \prod_r \beta_r \frac{P_r}{K_{P,r} + P_r} \quad (1)$$

where  $\alpha_f$  and  $\beta_r$  represent stoichiometric coefficients associated with the forward and reverse reactions, respectively, which have their corresponding maximum rates,  $v_{f,max}$  and  $v_{r,max}$ , respectively. Equilibrium constants,  $K_{R,f}$  and  $K_{P,r}$ , are associated with every reactant,  $R_f$ , and product,  $P_r$ , respectively. Although this structure follows basic assumptions, it is utilized to avoid high computational expense to remain competitive with other modelling approaches. Using a mass balance across each metabolite, the reaction fluxes determine the rate of change of each metabolite – integrating this generates concentration profiles that can be compared to recorded data. A sum of least squares objective function can be used to minimise the difference between the actual and simulated concentrations,  $C_{j,meas_t}$  and  $C(t)_j$ , respectively, at the times measurements were taken, as shown in Eq. 2:

$$\min \sum_{t,meas} \sum_{j,meas} \left( \frac{C_{j,meas_t} - C_j(t)}{C_{j,meas_t}} \right)^2 \quad (2)$$

Parameters of the Michaelis-Menten model structure,  $v_{f,max}$ ,  $v_{r,max}$ ,  $K_{R,f}$ , and  $K_{P,r}$ , are constant with time, but vary for each reaction.

### 2.2. Dynamic metabolic flux analysis (DMFA)

For the first DMFA methodology, a traditional approach is undertaken in which it is assumed no changes in intracellular metabolite concentrations occur, unless proven otherwise by measurements. 2<sup>nd</sup> degree polynomial expressions can be used to fit measured intracellular and extracellular metabolite concentrations:  $C_j(t) = a \cdot t^2 + b \cdot t + c$ , where  $a$ ,  $b$ , and  $c$  are parameters to be fitted. No higher order polynomials were utilized to avoid overfitting. Such polynomials construct mass balances, as shown in Eq. 3; an equality constraint. For extracellular metabolites, a similar concept is used to minimise the difference between the simulated and predicted rates of concentration at each timestep  $t$ , shown as Eq. 4. In this case, the reaction fluxes are the problem variables.

$$\frac{dC_{j,int}(t)}{dt} = \sum_i \alpha_i v_i \quad (3)$$

$$\min \left( \frac{X(t) \cdot \mu_t - \frac{dX(t)}{dt}}{\frac{dX}{dt}} \right)^2 + \sum_{j, meas, ex} \left( \frac{10^3 \cdot X(t) \cdot \sum_i \alpha_{i,j} \cdot v_{i,t} - \frac{dC_{j,ex}(t)}{dt}}{\frac{dC_{j,ex}}{dt}} \right)^2 \quad (4)$$

For the second DMFA methodology, the cell is assumed to behave as efficiently as possible. This is done by minimising the magnitude of each reaction flux at each timestep  $t$ , shown by Eq. 5, where  $\omega_i$  is the weight associated with flux  $v_i$ . As with the previous MFA approach, the same polynomials are utilized to determine the intracellular concentration profiles. For extracellular metabolites, the mass balance forms equality constraints to avoid over-complication of the objective function, presented via Eq. 6.

$$\min \sum_i \omega_i v_{i,t}^2 \quad (5)$$

$$\frac{dC_{j,ex}(t)}{dt} = 10^3 \cdot X(t) \cdot \sum_i \alpha_{i,j} \cdot v_{i,t} \quad (6)$$

Finally, we proposed a novel approach to DMFA where accumulation is considered. The accumulation is captured by an accumulation variable  $A$ . This coincides with the mass balance of each intracellular metabolite,  $\frac{dC_{j,int,t}}{dt} = \sum_i \alpha_{i,j} v_{i,t} = A_{j,t}$ . The aim is to recognize that the dynamic system is unlikely to be at complete steady-state but that the cell will ideally not excessively produce or consume any metabolites. Hence, it is appropriate to penalise accumulation in the objective function Eq. 7.

$$\min \left( \frac{X(t) \cdot \mu_t - \frac{dX(t)}{dt}}{\frac{dX}{dt}} \right)^2 + \sum_{j, ex} \left( \frac{10^3 \cdot X(t) \cdot \sum_i \alpha_{i,j} \cdot v_{i,t} - \frac{dC_{j,ex}(t)}{dt}}{\frac{dC_{j,ex}}{dt}} \right)^2 + w_{acc} \sum_{j, in} \left( \frac{A_{j,t}}{\max(|A_{j,t-1}|, 10^{-20})} \right)^2 \quad (7)$$

Metabolite concentrations are updated linearly, which is valid given a small time increment,  $\Delta t$ , between MFA calculations,  $C_{j,in,t} = C_{j,in,t-1} + A_{j,t-1} \cdot \Delta t$ .

### 2.3. Regularizing flux inconsistency

In the DMFA methodologies discussed, there is potential for consecutive fluxes to display severe inconsistency, which is unlikely to be a true representation of the system. To avoid such behaviour, a penalty function is implemented to penalise deviations from the flux at the previous timestep. This penalty has a corresponding weight, which is determined iteratively, to ensure it does not overrule the main objective function.

## 3. Case Study

A metabolic network of Myeloid-Derived Suppressor Cells (MDSCs) studied in a previous paper (Goffaux et al., 2017) was used as a case study given the availability of essential metabolites data. The metabolic network consists of 46 reactions, including biomass generation, and 44 metabolites. The data available contains measurements, taken every 8 hours for 96 hours, of 7 extracellular metabolites, biomass concentration, and 12 intracellular metabolites. The remaining 25 metabolites have an initial concentration measurement only. The metabolic reaction network is shown as Fig. 1. More details about the case study can be found in (Goffaux et al., 2017). Four modelling methodologies are

to be considered and compared: a class kinetic modelling approach and three variations of dynamic metabolic flux analysis (DMFA).

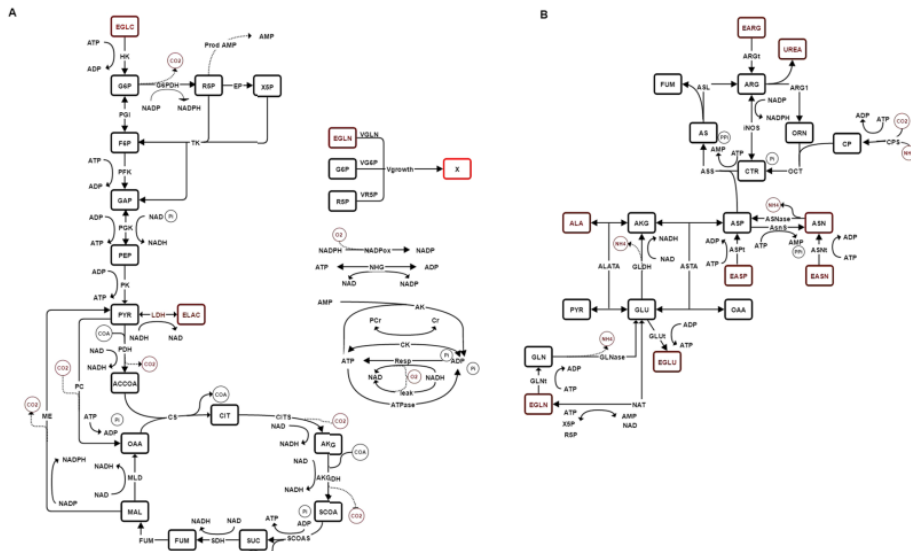


Figure 1: Metabolic network to be simulated (Goffaux et al., 2017), consisting of: (A) Central carbon metabolism and bioenergetics; (B) Urea cycle and amino acid catabolism.

#### 4. Results and Discussion

When comparing the fitting results of the four approaches, the kinetic model is found to performed far worse (mean error 37.8%) than the other methodologies and the three DMFA approaches were almost inseparable (mean error 16.6%), due to the added flexibility provided by the data-driven elements.

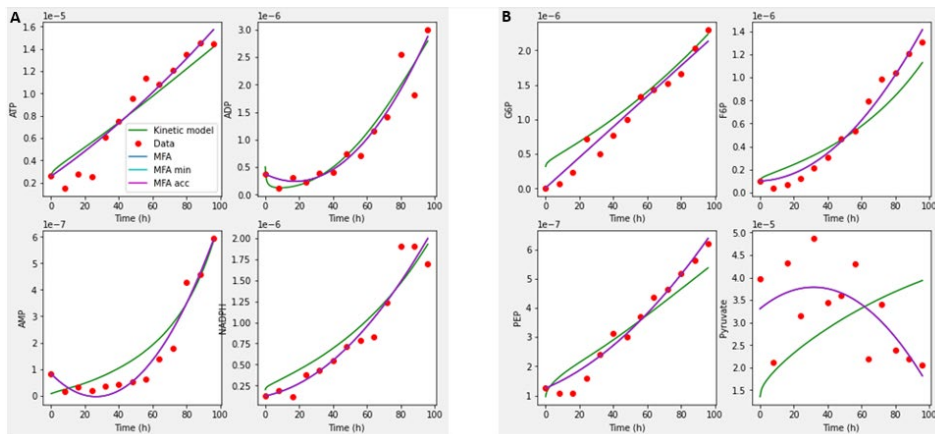


Figure 2: Comparison of experimental data (red) against the simulations of various modelling methodologies, DMFA (blue), DMFA with minimisation (cyan), and DMFA with accumulation (magenta), for both extracellular (A) and intracellular (B) metabolites.

For much of the extracellular metabolites, good fits were obtained – especially in instances of low-level noise in the data. For intracellular metabolites, much of the same was generally observed. However, some datasets were extremely noisy, such as pyruvate, which led to poor fitting – especially for the kinetic model, as shown in Fig. 2B. The excessive noise causes great uncertainty in the initial pyruvate concentration. This is a potential cause of the stiff region simulated by the kinetic model across the initial timesteps (up to a time of 0.3 hours). The stiffness was alleviated by allowing the initial concentration of pyruvate to vary within a sensible region. In addition, the penalty function was successfully implemented, greatly reducing the inconsistency in several reaction fluxes without hindering the performance of the main objective function. When comparing the different methodologies, the DMFA and DMFA with flux minimisation showed the most agreement between any two methodologies. A few cases in which the kinetic modelling and DMFA with accumulation showed better agreement than the traditional DMFA approaches could highlight a metabolite that is in fact violating steady-state or a reaction that requires some form of accumulation to stabilise it. When comparing to DMFA studies and flux observations of similar networks (Tang et al., 2009 and Chan et al., 2003), it is apparent that DMFA and DMFA with flux minimisation simulate more realistic fluxes for both central carbon metabolism, and the urea cycle with amino acid catabolism. DMFA with accumulation and kinetic modelling have a tendency to allow large accumulation of some metabolites as opposed to a larger flux consuming said compound. Since the DMFA approaches used the maximum flux parameters found in the kinetic modelling approach for the flux constraints, it can be concluded that is likely that these parameters were not well estimated since the DMFA and DMFA with minimisation were unable to converge to a feasible solution with such tight constraints. Several constraints, especially in the cyclical structures, had to be relaxed, and literature agrees that the kinetic model and DMFA with accumulation did in fact underestimate said fluxes (Tang et al., 2009) (Chan et al., 2003). DMFA with accumulation was able to overcome these tight constraints by allowing excessive accumulation, despite being unrealistic. Literature containing intracellular metabolite concentrations were immensely scarce, but studies on mammalian cells (Yishitoshi et al., 2017) indicate that the kinetic model was generally predicting sensible concentrations, where data was available.

Finally, the correlation matrix of the kinetic parameters in the kinetic model is also estimated and presented in Fig. 3. About 5% of the correlation values exceed 0.7 and the parameter pairs involved include 25% of the model parameters, meaning the kinetic model has an appropriate structure in most cases, but could benefit from fine-tuning. A high concentration of excessive correlation is observed in parameters associated with biomass growth, indicating it either does not obey the mechanism suggested, or the mechanism involves less relevant metabolites.

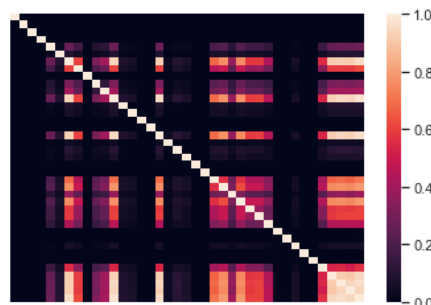


Figure 3: Correlation matrix of the kinetic model parameters as a heat-map.

## 5. Conclusion

Kinetic modelling can account for intracellular metabolite accumulation and is able to generate reasonable fits, although less accurate than DMFA. If such fits are improved, it could show a good understanding of the mechanisms within the metabolic pathway. However, it is time-consuming to construct a model and computationally expensive to solve. High stiffness meant that gradient-based optimisation failed, and stochastic search optimisation (Particle-Swarm) was utilized, thus not guaranteeing a local optimum being found. Literature also indicates that the fluxes are being underestimated, and this, along with the high stiffness, cases of high parameter correlation and highest average error, indicates that the correct structure was not found, which is a great challenge in kinetic modelling. The DMFA approaches yielded excellent fits to the experimental data and had a much lower computational cost due to the convex nature of each optimisation approach; allowing a global optimum to be found with gradient-based optimisation. However, there is always a high risk of overfitting when using purely data-oriented models for the measured metabolites and so it is essential to not over-complicate the model structure.

Penalizing inconsistent fluxes was effective for DMFA and DMFA with flux minimization. For DMFA with accumulation, the weights of the penalty functions were more sensitive and challenging to determine due to the presence of three terms in the objective function: 1) the main objective; 2) the accumulation penalty; and 3) the flux inconsistency penalty. All approaches can yield good results to simulate experimental data, but it is essential to study the metabolic network prior to experimental work and identify potential bottleneck metabolites (those leading into a cyclical structure) and try to measure these metabolites to help conclude the appropriate flux within these cycles. Traditional DMFA approaches appear to give more realistic flux simulations, which is intuitive as cells aspire to maintain a balance and not over-consume/produce any metabolites. DMFA with accumulation is the most robust approach due to its ability to overcome changes to the input via its allowance for accumulation.

As a result, based on the current research, an overall approach to modelling metabolic networks is suggested; begin with DMFA or DMFA with flux minimisation. If good results are unattainable, add accumulation terms as required. Apply kinetic modelling last; use the DMFA results to inform the kinetic model to reduce computational expense.

## References

- W. S. Ahn et al, 2012, Towards dynamic metabolic flux analysis in CHO cell cultures, *Biotechnology Journal*, Volume 7, Issue 1, Pages 61-74
- M. R. Antoniewicz, 2013, Dynamic metabolic flux analysis—tools for probing transient states of metabolic networks, *Current Opinion in Biotechnology*, Volume 24, Issue 6, Pages 973-978
- C. Chan et al, 2003, Metabolic Flux Analysis of Cultured Hepatocytes Exposed to Plasma, *Biotechnology and Bioengineering*, Volume 81, Issue 1, Pages 33-49
- G. Goffaux et al, 2017, A Dynamic Metabolic Flux Analysis of Myeloid-Derived Suppressor Cells Confirms Immunosuppression-Related Metabolic Plasticity, *Scientific Reports*, Volume 7, Issue 9850
- R. Harrington, HM Government, 2018, Growing the bioeconomy: a national strategy to 2030.
- Y. J. Tang et al, 2009, Metabolic Flux Analysis of *Shewanella* spp. Reveals Evolutionary Robustness in Central Carbon Metabolism, *Biotechnology and Bioengineering*, Volume 102, Issue 4, Pages 1161-1169
- E. Y. Yoshitoshi-Uebayashi et al, 2017, Modelling urea-cycle disorder citrullinemia type 1 with disease-specific iPSCs, *Biochemical and Biophysical Research Communications*, Volume 3, Issue 468, Pages 613-619

# Integrating hybrid modelling and transfer learning for new bioprocess predictive modelling

Sam Kay,<sup>a</sup> Harry Kay,<sup>a</sup> Alexander W. Rogers<sup>a</sup>, Dongda Zhang,<sup>a\*</sup>

<sup>a</sup>*Department of Chemical Engineering, University of Manchester, Oxford Road, Manchester, M1 3AL, UK.*

\*[dongda.zhang@manchester.ac.uk](mailto:dongda.zhang@manchester.ac.uk)

## Abstract

Hybrid modelling combines data-driven and mechanistic modelling, providing a cost effective solution to modelling complex (bio)chemical processes when the underlying mechanisms are not fully understood. While more data-efficient than pure data-driven methods, sufficient experimental data must still be collected through time-consuming experiments for hybrid model construction when applied to simulate a newly investigated process. To address this challenge and accelerate design of new systems, transfer learning has been recently implemented to transfer knowledge from a well-understood source domain to a less-studied target domain. Therefore, in this study, hybrid modelling and transfer learning are integrated for the first time to predict a novel dynamic process for microalgal lutein synthesis through photo-production. Specifically, a hybrid model was built to extract primary knowledge about a well-known lutein production strain, *Desmodesmus* sp., for which considerable process data and a high-fidelity kinetic model are available from the literature. Then transfer learning was used to update the hybrid model using limited process data for a recently isolated microalgal strain, *Chlorella sorokiniana*. Once built, the hybrid transfer model was tested by predicting biomass growth and lutein production of the new strain under various operating conditions. Furthermore, bootstrapping was adopted to estimate and propagate the uncertainty of the hybrid transfer models when predicting each batch process. The performance and efficiency of the hybrid transfer learning model were thoroughly demonstrated through experimental verification. It was found that hybrid transfer learning more efficiently modularised primary knowledge than previous pure data-driven transfer learning techniques, enabling more selective updates of source-specific knowledge to preserve generalisable knowledge. This study, therefore, provides a novel contribution to the predictive modelling and knowledge transfer of complex bioprocesses.

**Keywords:** Machine learning, hybrid modelling, transfer learning, microalgae photo production, small data problems.

## 1. Introduction

Small data problems are prevalent throughout bioprocess industries, where collecting samples of experimental data is both time consuming and resource intensive. Datasets classified as small data lack sufficient data points, so the system domain is sparsely explored hence, it is not possible to parameterise a model which fully represents the process. Recently, the limitations of small data have been discussed in the literature, and several techniques have been proposed to tackle this such as artificial data generation (Abdul Lateh et al., 2017). Another such approach is transfer learning, where modelling knowledge of another system can be applied to the new system in a controlled manner. Literature reports significant advantages in data efficiency using transfer learning to

impart knowledge from a previously studied (larger) dataset to another (similar) system characterised by a lack of data (Barman et al., 2019).

Traditionally, modelling is categorised into data driven and mechanistic approaches (i.e., models derived from statistics and those derived from first principles), however, in recent years there has been a rise in the use of hybrid modelling throughout the literature (Hodgson, 2005). Hybrid modelling combines the two traditional approaches, supplementing the current mechanistic knowledge with a surrogate statistics-based model able to capture additional complex non-linearities within the system. The inclusion of physical knowledge reduces the data required to construct an accurate model which, for systems where data is scarce, can present large benefits. In the bioprocess industries, it is common for first-principle models to be restricted in complexity due to limited mechanistic knowledge of an organisms metabolic reactions and the difficulties in deriving models that represent the immense number of possible metabolic pathways present within the biological system (Hodgson, 2005). Hybrid models, therefore, offer a unique solution to overcome the challenge of limited mechanistic knowledge and small data by integrating data driven techniques to represent the extremely non-linear system complexities that may be left unaccounted for in current first-principle models.

## 2. Problem statement

The focus of this study is directed at determining the possible effectiveness of combining hybrid modelling of a well understood biological system and transfer learning of the learned knowledge to a less studied target domain. The aim is to develop a framework for accelerating the design of new systems from scarce data, typical of that experienced in laboratory conditions for such newly researched systems. The source model is built for the microalgal strain, *Desmodesmus* sp., and the target model is constructed for a recently isolated strain, *Chlorella sorokiniana*. Sufficient data is available for the prior whereas little data is available for the latter. The biomass concentration in each case is recorded alongside the substrate concentration (nitrate) and the concentration of product (lutein) is monitored through the full experiment timespan. In the source domain (*Desmodesmus* sp.), a nitrate feed stream is introduced at the same hour for every experiment, the concentration of which is altered between experiments. This substrate feed stream is not present in the target domain. Furthermore, for both strains, light intensity is a controlled variable modified between experiments. More detailed introduction to the experimental setup and data collection of the case study can be found in (Rogers et al., 2022).

For a successful transfer of knowledge, the two systems under study must share similar underlying mechanisms such that information extracted from one system can be relevant to the other (Weiss et al., 2016). In this study, it is believed that the two microalgal strains share similar kinetics and metabolic pathways such that some level of mechanistic and statistical knowledge can be used to inform the target strain, reducing the data requirements of the hybrid model. The greatest advantages of transfer learning are observed when the target domain is severely lacking in data, which, in this study manifests itself as a lack of experiments available for model construction. The prerequisite to designing a successful transfer model is first to develop a robust, accurate source model from which information can be shared. Without a solid foundation, any information transferred to the target domain will be non-generalisable and will likely limit the performance of the target model.

Although the use of transfer learning and hybrid modelling has been well explored in recent studies, the combination of the two has not been explored. In this work, the data

driven component of the hybrid models will be managed by an artificial neural network (ANN). Bootstrapping techniques are used to estimate the uncertainty associated with the batch predictions. Once a robust source model is designed, the optimised kinetic parameters and ANN parameters (weights and bias') are implemented in their respective objective functions to provide guidance to the target model during its construction.

### 3. Methodology

First, the kinetic model structure will be introduced. For the production of lutein from the *Desmodesmus* sp. strain, (del Rio-Chanona et al., 2017) proposes an informed kinetic model using modified Monod kinetics. The equations proposed are used in this study to represent the source domain and are as follows:

$$\frac{dX}{dt} = u_0(X, N, I) \cdot \frac{N}{N + K_N} \cdot X - \mu_d \cdot X \quad (1a)$$

$$\frac{dN}{dt} = -Y_N \cdot u_0(X, N, I) \cdot \frac{N}{N + K_N} \cdot X + F_{N0} \quad (1b)$$

$$\frac{dP}{dt} = k_0(X, N, I) \cdot \frac{N}{N + K_P} \cdot X - k_d \cdot P \cdot X \quad (1c)$$

Where  $\frac{dX}{dt}$ ,  $\frac{dN}{dt}$  and  $\frac{dP}{dt}$  represent the rate of change of biomass, substrate (nitrate) and product (lutein) concentrations respectively. In equation 1a,  $u_0$  is the specific growth rate,  $I$  is incident light intensity,  $K_N$  is the nitrate half velocity constant and  $\mu_d$  is the specific decay rate. In equation 1b,  $Y_N$  is the nitrate yield coefficient and  $F_{N0}$  is the mass flow rate of nitrate influent. In equation 1c,  $k_0$  is the product synthesis rate parameter,  $K_P$  is the nitrate half velocity constant for synthesis of lutein and  $k_d$  is the product consumption rate constant. Derivation and justification of the model composition is explained thoroughly in (del Rio-Chanona et al., 2017).

For this methodology,  $K_N$ ,  $\mu_d$ ,  $Y_N$ ,  $K_P$  and  $k_d$  are estimated as time constant parameters where as  $u_0$  and  $k_0$  are assumed time varying. Using experimental data for the *Desmodesmus* sp. system, the model parameters are estimated with finite element discretisation and optimised using interior point algorithms in python. In the parameter estimation problem, a least squares error function is implemented to optimise the constant and time-varying terms; for an input of  $n$  experiments and  $N$  timesteps,

$$obj = \operatorname{argmin} \left( \frac{1}{n} \sum_{j=1}^n \sum_{i=1}^N (\theta_{j,i} - \theta'_{j,i})^2 \right); \theta_{j,i} = \{X, N, P\} \quad (2)$$

where,  $\theta$  and  $\theta'$  are the set containing the measured and estimated states respectively. Similarly, a least squares approach is employed in the ANN loss function.  $u_0$  and  $k_0$  are complex functions of  $X$ ,  $N$  and  $I$  so, to reduce the non-linearity of the parameter optimisation problem,  $u_0$  and  $k_0$  are solved as independent parameters at each timestep by minimising the fitting error, then an ANN is implemented to alter the terms achieving a better model fit by better approximating the inherent complexities in the system. Let us address the structure and intent behind the ANN component of the hybrid model. The objective of the ANN is to perform alterations to the time varying parameters ( $u_0$  and  $k_0$ ) to the aim of providing accurate estimations of biomass, nitrate, and lutein concentrations. In addition to this, the uncertainty of the predictions must be determined to ensure confidence in the models' performance hence, bootstrapping is employed to this effect. Two-fold validation was employed to assess the predictive abilities of the model and optimisation of the model hyperparameters was carried out. The structure of the ANN is shown in Figure 1.



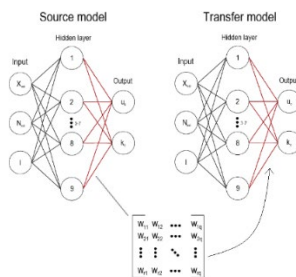


Figure 1: ANN component of the hybrid source and transfer models, detailing the structures, inputs and outputs.

To provide a meaningful comparison of results, the following metrics are defined for the mean relative percentage error (MRPE) and mean relative percentage uncertainty (MRPU) respectively:

$$MRPE\% = \frac{x' - x}{\frac{1}{2}(x' + x)} \cdot 100 \quad (3a)$$

$$MRPU\% = \frac{\sigma'}{\frac{1}{2}(x' + x)} \cdot 100 \quad (3b)$$

Where  $x'$  and  $x$  represent the prediction and experimental datapoint respectively, and  $\sigma'$  represents the uncertainty estimate obtained from bootstrapping. It is important to consider the relative error and uncertainty due to the proximity of some datapoints to zero values. If simply the mean average percentage error was taken, when the substrate concentration decreased close to zero, the denominator would then tend to zero yielding unrealistic percentage errors and uncertainties.

Herein, we will shift focus to the transfer of knowledge from the source domain to the target domain and the methodology that entails. It is known that the advantages offered by transfer learning are at their greatest when data is severely limited so, in this study only two experiments are used for estimation of model parameters and training of the target ANN. The target domain mechanistic structure is held almost identical to equations 1a, 1b and 1c, differing only by the  $F_{N0}$  term in equation 1b which is absent for the *Chlorella sorokiniana* strain due to no addition of nitrate solution in the experimental datasets. The transfer of information between mechanistic structures is achieved via implementation of penalty terms for specific constant and time varying parameters estimated by the transfer model which penalises deviation from the source model values. The exchange of statistical knowledge was executed via the direct implementation of the ANN architecture and optimal hyperparameters determined for the source model. Furthermore, an additional penalty was enforced in the ANN loss function to discourage divergence between the weights and bias' calculated through backpropagation in the transfer ANN and the parameters of the finalised source ANN.

#### 4. Results and Discussion

The source kinetic model was fitted using all available datasets. Single-fold cross validation on all 8 experiments of the *Desmodesmus* sp. dataset was used to determine the optimal hyperparameters of the source model ANN. This was achieved using the average MRPE per experiment alongside the determined training loss values to inform this optimisation procedure. Once the parameters were obtained, the source model was tested on the entire dataset to evaluate its performance with respect to both predictive

accuracy and uncertainty estimation. The results in Figure 2 detail the model predictions of biomass concentrations (g/L), nitrate concentration (g/L) and lutein concentration (mg/L) for experiments 6 and 8.

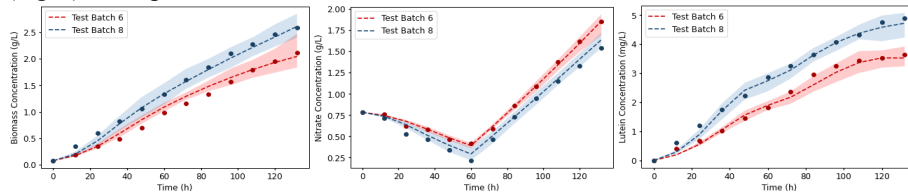


Figure 2 - Hybrid model predictions and corresponding uncertainty limits for test batches 6 and 8. From Figure 2, it can be seen that the hybrid model captures the underlying mechanisms to a considerable extent for the evolution of biomass, substrate and product concentration through time, providing reasonable uncertainty estimations which cover the residuals between the predictions and experimental datapoints. It was therefore determined that the source model was robust in its construction and, due to its success in capturing the fundamental physics driving the system, it is deemed appropriate for use as the foundation to support knowledge transfer to the novel *Chlorella sorokiniana*. system. The transfer model was constructed using only two experiments, severely limiting the information available to the model for informed predictions. The kinetic model for the transfer domain was parameterised using experiments 3 and 6 however a modified objective function was implemented such that the constant parameters  $K_N$  and  $K_P$  were penalised the further they deviated from the source domain values. A similar approach was taken for the time varying parameter,  $u_0$  however, the source model values were averaged at each timestep to obtain a single vector of values which the transfer kinetic model parameters could be compared against. Similarly, the ANN component of the hybrid model was then trained on experiments 3 and 6 and optimised using a MSE loss function adjusted to discourage significant variation of the weights and bias from the source model parameters. The predictions of the transfer model on testing experiments 5 and 7 are shown in Figure 3.

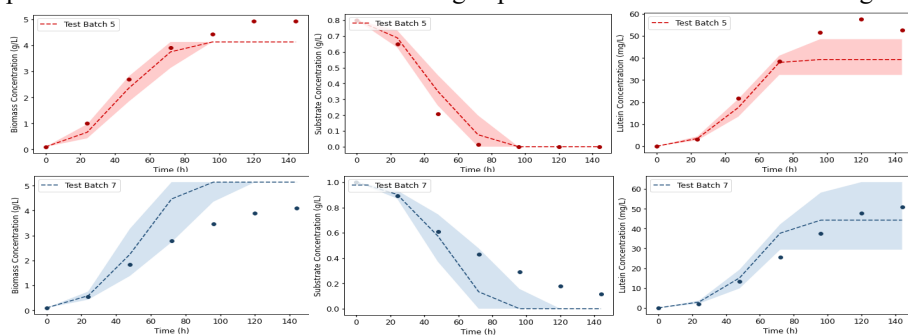


Figure 3: Transfer model predictions and corresponding uncertainty limits for test batches 5 and 7.

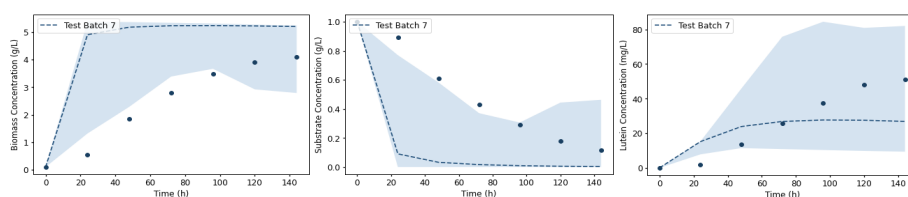


Figure 4: Benchmark predictions and corresponding uncertainty limits for test batch 7.

From Figure 3, it is observed that, with very small data, the transfer model successfully reproduced the trends within the process, while offering adequate estimations of state parameters as well as representative uncertainty bounds. To assess the potential of the transfer model, it is necessary to construct a benchmark model built under the same conditions of extreme data scarcity. The mechanistic structure of the benchmark and its corresponding ANN were once again parameterised however, no external influence from the source model was introduced. The results from this analysis are shown in Figure 4.

From Figure 4, it can be observed that the trends captured by the benchmark are generic and lacking in knowledge specific to the provided dataset. When compared to the transfer model, it is clear that the introduction of prior knowledge from the source domain provides notable advantages. It is also evident that the information held within the small training set is insubstantial and is not representative of the *Chlorella sorokiniana*. domain and, as such, knowledge transfer provides a surrogate for this missing information.

## 5. Conclusion

In conclusion, transfer learning offers a unique solution when used in conjunction with hybrid modelling to overcome the necessity of large datasets to achieve satisfactory model performances. The controlled introduction of prior mechanistic and statistical knowledge into a novel, less studied domain establishes a robust framework from which accelerated design can be accomplished. The combination of both data-driven and kinetic modelling simultaneously reduces the data intensity of the transfer learning process while extracting primary knowledge from which some degree of transferable information is contained. Furthermore, the use of hybrid modeling allows for the extension of the framework to systems in which mechanistic understanding is bounded, as increased reliance on the data-driven component can be used to account for this. Overall, this work demonstrates the novel combination of hybrid modelling and transfer learning and exhibits the potential entailed within the use of this framework for situations in which data availability is limited.

## References

- Abdul Lateh, M., Muda, A. K., Izzah Mohd Yusof, Z., Azilah Muda, N., & Sanusi Azmi, M. (2017). Handling a Small Dataset Problem in Prediction Model by employ Artificial Data Generation Approach: A Review. *Journal of Physics: Conference Series*, 892(1), 012016.
- Barman, R., Deshpande, S., Agarwal, S., & Inamdar, U. (2019). *Transfer Learning for Small Dataset*.
- del Rio-Chanona, E. A., Ahmed, N. rashid, Zhang, D., Lu, Y., & Jing, K. (2017). Kinetic modeling and process analysis for *Desmodesmus sp. lutein* photo-production. *AIChE Journal*, 63(7), 2546–2554.
- Hodgson, B. J. (2005). *Hybrid modelling of bioprocesses*.
- Rogers, A. W., Vega-Ramon, F., Yan, J., del Rio-Chanona, E. A., Jing, K., & Zhang, D. (2022). A transfer learning approach for predictive modeling of bioprocesses using small data. *Biotechnology and Bioengineering*, 119(2), 411–422.
- Weiss, K., Khoshgoftaar, T. M., & Wang, D. D. (2016). A survey of transfer learning. *Journal of Big Data*, 3(1), 1–40.

# Machine learning-supported cybergenetic modeling, optimization and control for synthetic microbial communities

Sebastián Espinel-Ríos,<sup>a</sup> Katja Bettenbrock,<sup>a</sup> Steffen Klamt,<sup>a</sup> José L. Avalos,<sup>b</sup>  
Rolf Findeisen<sup>c</sup>

<sup>a</sup>*Analysis and Redesign of Biological Networks, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany*

<sup>b</sup>*Department of Chemical and Biological Engineering, Princeton University, 08544 Princeton, United States*

<sup>c</sup>*Control and Cyber-Physical Systems Laboratory, Technical University of Darmstadt, Landgraf-Georg-Straße 4, 64283 Darmstadt, Germany*

\*Correspondence: [rolf.findeisen@iat.tu-darmstadt.de](mailto:rolf.findeisen@iat.tu-darmstadt.de). †This work was supported by the International Max Planck Research School for Advanced Methods in Process and Systems Engineering.

## Abstract

Synthetic microbial communities are promising production strategies that can circumvent, via division of labor, many challenges associated with monocultures in biotechnology. Here, we consider microbial communities as lumped metabolic pathways where their members catalyze different metabolic submodules. We outline a machine learning-supported cybergenetic strategy for manipulating the reaction rates of microbial consortia via dynamic regulation of key biomass population levels. To do so, we show a quasi-unstructured modeling approach for synthetic microbial communities with external regulation of intracellular growth regulatory components. Then, we formulate an optimal control problem to find the optimal initial conditions and dynamic input trajectories. We use model predictive control to address system uncertainty, which can be coupled to an observer based on moving horizon estimation. Using a two-member community with optogenetic control as a simulation example, we found the optimal initial biomass concentrations and light intensity trajectories to maximize naringenin production.

**Keywords:** synthetic microbial communities, machine learning-supported optimization, model predictive control, estimation, optogenetics.

## 1. Introduction

Monocultures dominate biotechnological production, i.e., processes catalyzed by single-cell types. While metabolic engineering may increase the production efficiency of cells, it often involves intense genetic modifications. As a result, the engineered cells may suffer from resource burden and disturbances in the energy and redox balances, thus portraying poor growth and low volumetric productivity rates. Synthetic microbial communities (SMCs) are possible solutions to these challenges as metabolic tasks are distributed among a designed consortium of microorganisms, hence division of labor is achieved (Roell et al., 2019). In this work, we outline a (machine-learning supported) modeling, optimization and control framework for maximizing the production efficiency of SMCs.

We visualize SMCs as lumped metabolic pathways consisting of several metabolic submodules, each catalyzed by a different member population. We consider that the

reaction rates of the submodules in the SMCs can be dynamically adjusted by online modulation of the biomass population levels. Being able to do this in a structured and flexible manner can unlock a powerful optimization degree of freedom in biotechnology. To enable online modulation of the population levels, we propose to introduce tunable genetic systems into (some of) the community members for regulating the gene expression of key intracellular growth regulatory elements via external inputs.

The remainder of this paper is structured as follows. In Section 2, we describe a general modeling approach for SMCs in batch processes. We lump up most of the intracellular bioprocesses and focus on the dynamics of the external species and the externally manipulated intracellular growth regulatory components. With this, we aim to decrease the model complexity as much as possible, while still capturing the most relevant process dynamics. We outline the use of machine learning for improving the prediction capacity of the derived models by learning uncertain or unknown parts of the dynamic equations. In Section 3, we formulate a model-based optimal control problem to find the optimal initial conditions and dynamic input trajectories. In Section 4, system uncertainty is addressed via model predictive control (MPC), optionally coupled to a state observer based on moving horizon estimation (MHE) (Findeisen and Allgöwer, 2002; Rawlings et al., 2020). Finally, in Section 5, a two-member consortium for naringenin production (Lalwani et al., 2021) is used as a simulation example.

## 2. Modeling of synthetic microbial communities with cybergeneic inputs

Let  $B \in \mathbb{R}^{n_B}$  refer to the biomass concentrations of the community members. We comprise in  $p \in \mathbb{R}^{n_p}$  all the intracellular concentrations of growth regulatory components present in the SMC and in  $z \in \mathbb{R}^{n_z}$  the concentrations of extracellular metabolites. The reaction rates of the model  $V \in \mathbb{R}^{n_V}$  are divided into rates of production  $V_p \in \mathbb{R}^{n_p}$  and degradation/dilution  $V_d \in \mathbb{R}^{n_p}$  of intracellular growth regulatory components, rates of biomass growth  $V_B \in \mathbb{R}^{n_B}$  and rates of metabolic conversion of extracellular metabolites  $V_z \in \mathbb{R}^{n_z}$ . Hence,  $V := [V_p^T, V_d^T, V_B^T, V_z^T]^T$ . Without loss of generality, we focus on batch processes, thus the dynamics of the SMC follows

$$\dot{p}(t) = S_p V(t), \quad p(t_0) = p_0, \quad (1)$$

$$\dot{B}(t) = S_B V(t), \quad B(t_0) = B_0, \quad (2)$$

$$\dot{z}(t) = S_z V(t), \quad z(t_0) = z_0, \quad (3)$$

where  $S_p \in \mathbb{R}^{n_p \times n_V}$ ,  $S_B \in \mathbb{R}^{n_B \times n_V}$  and  $S_z \in \mathbb{R}^{n_z \times n_V}$  are matrices that map the reaction rates to the corresponding differential equations.

The control inputs are collected in  $u \in \mathbb{R}^{n_u}$  and the model parameters in  $\theta \in \mathbb{R}^{n_\theta}$ . For simplicity, we assume that a given control input can affect the expression of at most one growth regulatory component. In addition, we consider that there is no more than one tunable growth regulatory element per community member. Hence,

$$V_{B_i} = \begin{cases} B_i \mu_{\text{reg}_i}(p_i, z, \theta) & \text{for regulation via } p_i, \forall i \in \mathbb{B} \\ B_i \mu_{\text{unr}_i}(z, \theta) & \text{otherwise,} \end{cases} \quad (4)$$

$$V_{p_i} = \eta_i(u_i, \mu_{\text{reg}_i}, \theta), \quad \forall i \in \mathbb{P}, \quad (5)$$

$$V_{d_i} = d_i(\mu_{\text{reg}_i}, p_i, \theta), \quad \forall i \in \mathbb{P}. \quad (6)$$

$$V_{z_i} = B_j q_i(\mu_j, \theta), \quad \forall i \in \mathbb{Q}(j). \quad (7)$$

where  $\mathbb{B}$  is the set of community members,  $\mathbb{P}$  be the set of growth regulatory elements and  $\mathbb{Q}(j)$  is the set of conversion rates catalyzed by a member  $B_j$ . Furthermore,  $\mu_{\text{reg}_i}: \mathbb{R} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_\theta} \rightarrow \mathbb{R}$ ,  $\mu_{\text{unr}_i}: \mathbb{R}^{n_z} \times \mathbb{R}^{n_\theta} \rightarrow \mathbb{R}$ ,  $\eta_i: \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{n_\theta} \rightarrow \mathbb{R}$ ,  $d_i: \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{n_\theta} \rightarrow \mathbb{R}$  and  $q_i: \mathbb{R} \times \mathbb{R}^{n_\theta} \rightarrow \mathbb{R}$  are appropriate kinetic functions.

Given the quasi-unstructured nature of the model, the kinetic functions in Eqs. (4)-(7) are often phenomenological relations. Although the latter generally offer good fitting properties, they usually lack good predictability (Zhang et al., 2019). In some cases, one might not even have a mechanistic or phenomenological relation at hand due to the complexity of a given bioprocess. Therefore, *if necessary*, we can learn the uncertain parts of the model using machine learning. In Section 5 we show an example of a *hybrid* model.

### 3. Optimal cybergenetic control of synthetic microbial communities

Let us define  $x := [p^T, B^T, z^T]^T$ ,  $x \in \mathbb{R}^{n_x}$ . We formulate a model-based optimal control problem to maximize the production efficiency of a SMC (e.g., productivity, profit, etc.)

$$\max_{u(\cdot), x_0} \int_{t_0}^{t_f} l(x(t), u(t)) dt + e(x(t_f)), \quad (8)$$

$$\begin{aligned} \text{s. t.} \quad & \text{Eqs. (1) - (7),} \\ & x(t_0) = x_0, \quad (9) \\ & 0 \leq g(x(t), u(t)). \quad (10) \end{aligned}$$

In the previous formulation,  $u(\cdot)$  is the input function,  $l: \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}$  is the stage cost and  $e: \mathbb{R}^{n_x} \rightarrow \mathbb{R}$  is the terminal cost. The system state and input constraints are captured by  $g: \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_g}$ . The process time spans from the initial time  $t_0$  to the final time  $t_f$ . Note that we consider both  $u(\cdot)$  and the initial state concentration  $x_0$  as degrees of freedom (DOF). We also refer to this formulation as the *open-loop* optimization as no online corrective actions are taken if the system is out of “specs” or if disturbances occur.

### 4. Predictive cybergenetic control of microbial communities with estimation

We compensate for possible system uncertainty such as model-plant mismatch and disturbances using online feedback or closed-loop control in the form of MPC (Findeisen and Allgöwer, 2002; Rawlings et al., 2020). For the batch case we use a shrinking horizon, i.e., the control problem is solved iteratively over a time horizon  $[t_k, t_f]$ , where  $t_{k+1} = t_k + \Delta t$  and  $\Delta t$  is the sampling time size. Considering that  $\tilde{x}_k$  is the measured state at the sampling time  $t_k$ , the MPC problem reads

$$\max_{u(\cdot)} \int_{t_k}^{t_f} l(x(t), u(t)) dt + e(x(t_f)), \quad (11)$$

$$\begin{aligned} \text{s. t.} \quad & \text{Eqs. (1) - (7), (10),} \\ & x(t_k) = \tilde{x}_k. \quad (12) \end{aligned}$$

When not all the states can be measured, we can *reconstruct* them via MHE. We collect the measurements in  $y \in \mathbb{R}^{n_y}$  and Eqs. (1)-(3) in the vector function  $f(x, u, \theta)$ . Also, let  $h: \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\theta} \rightarrow \mathbb{R}^{n_y}$  be the measurement equations and  $w \in \mathbb{R}^{n_x}$  the state noise. We denote with  $(\cdot)_i$  and  $(\tilde{\cdot})_i$  the optimization and measured variables at time  $t_i$ , respectively. We indicate with  $(\bar{\cdot})$  the prior information of a variable. We assume that the input signal is piecewise constant over the sampling times. Given a time-discrete model, the MHE problem at time  $t_k$  reads (Rawlings et al., 2020)

$$\min_{x_{k-N}, \theta_k, w_{k-N}, \dots, w_k} \left\| \begin{bmatrix} x_{k-N} - \bar{x}_{k-N} \\ \theta_k - \bar{\theta} \end{bmatrix} \right\|_P^2 + \sum_{i=k-N}^k \|y_i - \bar{y}_i\|_R^2 + \|w_i\|_W^2 \quad (13)$$

$$\text{s. t.} \quad x_{i+1} = x_i + \int_{t_i}^{t_i + \Delta t} f(x, u, \theta_k) dt + w_i, \quad (14)$$

$$\begin{aligned} & \text{Eqs. (4) - (7), (10)} \\ & y(t_i) = h(x_i, u_i, \theta_k), \quad (15) \\ & i \in [k - N, k], k, N \in \mathbb{N}, \end{aligned}$$

where  $\|a\|_A^2 := a^T A a$ ,  $N$  is the length of the estimation window and  $P \in \mathbb{R}^{n_x + n_\theta, n_x + n_\theta}$ ,

$R \in \mathbb{R}^{n_y \times n_y}$  and  $W \in \mathbb{R}^{n_x \times n_x}$  are weighting matrices. In general, the decision variables of the MHE are the states at the beginning of the estimation window  $x_{k-N}$ , the parameters and the state noise; with these we can reconstruct the states via Eq. (14). The estimated state vector  $\hat{x}$  can be used to update the MPC. Parameters are assumed to be constant in the MHE horizon, but they can be updated at each iteration to enable an *adaptive* MPC.

## 5. Microbial community with optogenetic control: naringenin biosynthesis

A summary of the application example is presented in Fig. 1. The SMC consists of an engineered *Escherichia coli* that produces tyrosine, and an engineered *Saccharomyces cerevisiae* that consumes tyrosine to synthesize naringenin. *E. coli* expresses a toxin constitutively, while it expresses an antitoxin, MazE, in an inducible manner using blue light intensity  $I_b$  as control input. MazE counteracts the effect and halts the expression of the toxin, thereby enabling different growth levels (Lalwani et al., 2021). This prevents *E. coli* from outgrowing *S. cerevisiae* and turning the SMC into a monoculture.

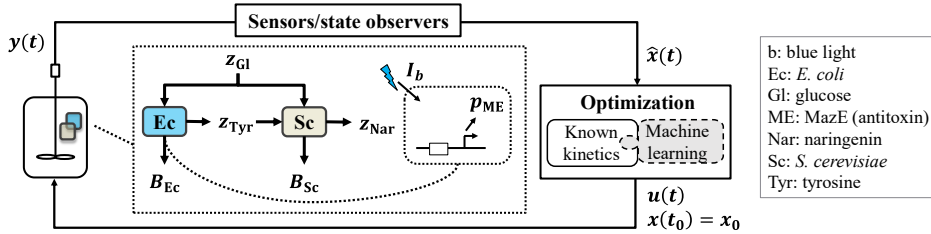


Figure 1. Cybergnetic strategy for naringenin production by a two-member SMC.

We consider that the dynamics of the SMC follows

$$\dot{p}_{ME} = V_{p_{ME}} - (k_d + \mu_{Ec})p_{ME}, V_{p_{ME}} = \left( \eta_{ME0} + \eta_{ME,m} \frac{I_b^{n_1}}{I_b^{n_1} + k_{ME}^{n_1}} \right) \mu_{Ec}, \quad (16)$$

$$\dot{B}_{Ec} = \mu_{Ec} B_{Ec}, \mu_{Ec} = \mu_{m,Ec} \left( \frac{z_{Gl}}{z_{Gl} + k_{Gl,Ec}} \right) \left( \frac{p_{ME}^{n_2}}{p_{ME}^{n_2} + k_p^{n_2}} \right), \quad (17)$$

$$\dot{B}_{Sc} = \mu_{Sc} B_{Sc}, \mu_{Sc} = \mu_{m,Sc} \left( \frac{z_{Gl}}{z_{Gl} + k_{Gl,Sc}} \right), \quad (18)$$

$$\dot{z}_{Gl} = -(Y_{Gl/Ec} \mu_{Ec}) B_{Ec} - (Y_{Gl/Sc} \mu_{Sc}) B_{Sc}, \quad (19)$$

$$\dot{z}_{Tyr} = (Y_{Tyr/Ec} \mu_{Ec}) B_{Ec} - \left( Y_{Tyr/Sc} \mu_{Sc} \frac{z_{Tyr}}{z_{Tyr} + k_{Tyr}} \right) B_{Sc}, \quad (20)$$

$$\dot{z}_{Nar} = \left( Y_{Nar/Tyr} Y_{Tyr/Sc} \mu_{Sc} \frac{z_{Tyr}}{z_{Tyr} + k_{Tyr}} \right) B_{Sc}, \quad (21)$$

with  $x = [B_{Ec}, B_{Sc}, z_{Gl}, p_{ME}, z_{Tyr}, z_{Nar}]^T$ ,  $x(t_0) = x_0$ . We refer to this model as the *nominal* or *real plant* model. All the optimizations and the machine learning-related methods presented hereafter were performed with HILO-MPC (Pohlodek et al., 2022).

The open-loop optimization results are presented in Fig. 2. We maximized  $z_{Nar}(t_f)$  using the optimal control problem in Section 3 based on the nominal model. We selected the parameter values of the nominal model such that biologically sound process dynamics under aerobic conditions could be simulated. Considering three DOF,  $B_{Ec}(t_0)$ ,  $B_{Sc}(t_0)$  and  $I_b(\cdot)$ , we obtained 12.0 mg/L naringenin. We also analyzed the case where *E. coli* does not possess the toxin/antitoxin system, i.e.,  $\dot{p}_{ME} = 0$  and  $\mu_{Ec} = \mu_{m,Ec} (z_{Gl}/(z_{Gl} + k_{Gl,Ec}))$ . In this second scenario, with  $B_{Ec}(t_0)$  and  $B_{Sc}(t_0)$  as DOF, we obtained 9.3 mg/L naringenin, 23 % less than with the previous optimization. Furthermore, the optimizer predicts a very low  $B_{Ec}(t_0)/B_{Sc}(t_0)$  to reduce the risk of *E. coli* outgrowing *S. cerevisiae*.

In contrast, with dynamic growth regulation of *E. coli*, one can *safely* start the process at a higher *E. coli* concentration, which increases the tyrosine available for conversion to naringenin by *S. cerevisiae*, hence the better performance.

In the previous cases, the predictions of the optimization matched the plant simulations as we did not consider model-plant mismatch. However, often we do not know the exact plant model. In a third open-loop scenario, we assumed that  $V_{p_{ME}}$  in Eq. (17) and the  $p_{ME}$ -dependent factor in the  $\mu_{EC}$ -kinetics (Eq. (18)) were unknown. A feedforward fully connected neural network was trained to learn these unknowns. We simulated eight batches with the nominal model subject to randomized input values and added 2.5 % Gaussian white noise to the *measurements*. We considered  $p_{ME}$  and  $I_b$  as features and the two unknown parts of the model as labels. From the eight batches, 20 % of the data was used for validation. Three additional independent batches were used to test the quality of the hybrid model (Fig. 3). Despite some model-plant mismatch, the hybrid model was able to predict overall trends. Nevertheless, using the hybrid model in the open-loop optimization with  $B_{EC}(t_0)$ ,  $B_{Sc}(t_0)$  and  $I_b(\cdot)$  as DOF rendered 8.5 mg/L naringenin (cf. Fig. 2), 29 % less than with the open-loop optimization based on the nominal model. The hybrid model-based optimization overestimated *E. coli* growth and tyrosine synthesis, thus negatively impacting on naringenin production.

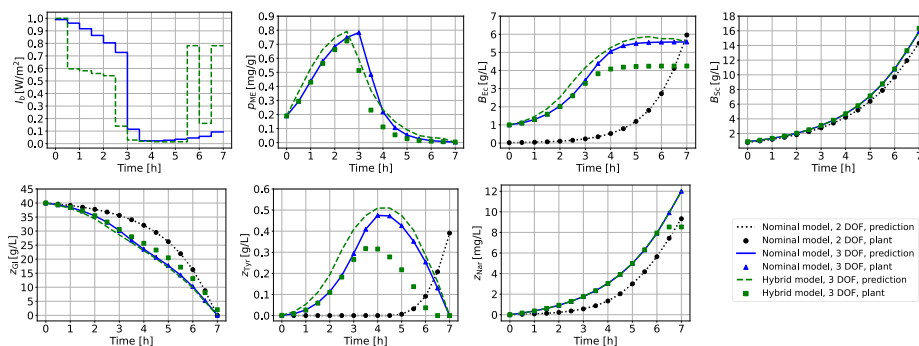


Figure 2. Open-loop optimization results of the SMC for naringenin production. Nominal model:  $\mu_{m,EC} = 0.825 \text{ h}^{-1}$ ,  $\mu_{m,Sc} = 0.418 \text{ h}^{-1}$ ,  $\eta_{ME_0} = 5.71 \times 10^{-5} \text{ g/g}$ ,  $\eta_{ME,m} = 2.86 \times 10^{-3} \text{ g/g}$ ,  $n_1 = 2.49$ ,  $n_2 = 1.5$ ,  $k_d = 1.39 \text{ h}^{-1}$ ,  $k_{GI,EC} = 6.8 \times 10^{-5} \text{ g/L}$ ,  $k_{GI,Sc} = 0.06 \text{ g/L}$ ,  $k_{ME} = 0.138 \text{ W/m}^2$ ,  $k_p = 5 \times 10^{-4} \text{ g/g}$ ,  $k_{TYr} = 1 \times 10^{-6} \text{ g/L}$ ,  $Y_{GI/EC} = 2.2 \text{ g/g}$ ,  $Y_{GI/Sc} = 2.0 \text{ g/g}$ ,  $Y_{Nar/TYr} = 0.016 \text{ g/g}$ ,  $Y_{TYr/EC} = 0.164 \text{ g/g}$ ,  $Y_{TYr/Sc} = 0.05 \text{ g/g}$ . The relevant product yields were inferred *at our best* from Lalwani et al. (2021). Similarly, *realistic* optogenetic-related parameters were inferred from Olson et al. (2014).

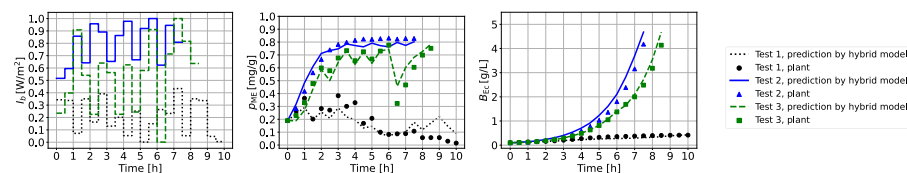


Figure 3. Testing of the hybrid model. We only show the dynamics of  $p_{ME}$  and  $B_{EC}$ . Artificial neural network configuration: two hidden layers with ten neurons each, sigmoid activation function.

The hybrid model could be improved by increasing the amount of data used for training. However, experiments in biotechnology are generally expensive and time-consuming, which limits the available data. Therefore, in Fig. 4 we use MPC to address the model uncertainty of the hybrid model. First, we assumed that all the states can be measured online. This led to 11.5 mg/L naringenin, significantly compensating for the uncertainty



of the hybrid model by adjusting the dynamic input trajectory. Despite the good results, online monitoring of intracellular components is challenging. Therefore, we also analyzed the case where  $p_{ME}$  cannot be measured online. A special case of MHE is when all the measurements from  $t_0$  are considered; this is called full information estimation (FIE). We estimated  $p_{ME}$  via FIE, where we minimized the first two terms in Eq. (13) and did not consider the parameters as decision variables. Note that from the estimated state vector, we only passed  $\hat{p}_{ME}$  to the MPC as we *trusted* the measurements. This resulted in 10.8 mg/L naringenin, representing a 21 % enhancement compared to the open-loop scenario with the hybrid model and only 6 % less efficient than the MPC with full state measurement. As expected, the FIE estimates improved as the estimation window grew due to its *memory* effect, i.e., more historical measurements available for the estimation.

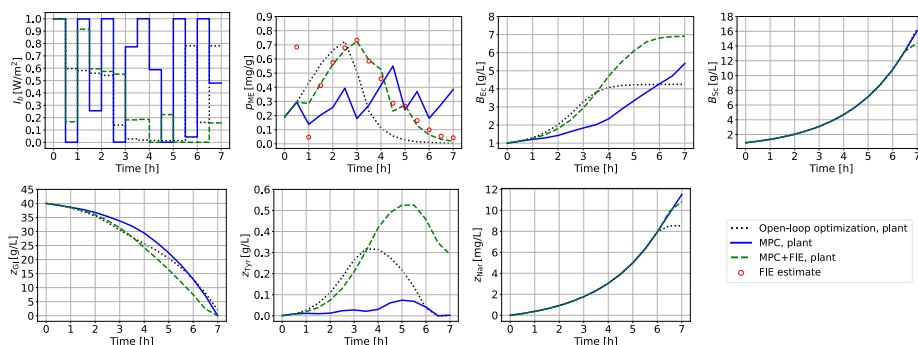


Figure 4. Machine learning-supported feedback control of the SMC for naringenin production.  $B_{EC}(t_0)$ ,  $B_{SC}(t_0)$  and the first step of  $I_b$  were found at the first MPC iteration, then the MPC only had  $I_b$  as a DOF. All the measurements had 2.5 % Gaussian white noise.

## 6. Conclusion

We outlined a (machine learning-supported) cybergenetic modeling, optimization and control strategy for exploiting biotechnological production by SMCs. We also presented a state observer for inferring unmeasured states. We demonstrated the potential of our framework with a two-member SMC for naringenin production with optogenetic control as a simulation example. Although we focused on batch setups, the framework can in principle be extrapolated to other modes of operation. Future work deals with the experimental implementation of the framework and additional hybrid modeling and control approaches, e.g., based on Gaussian processes, Bayesian neural networks, physics-informed machine learning and reinforcement learning.

## References

- Findeisen, R., Allgöwer, F., 2002. An introduction to nonlinear model predictive control, in: 21st Benelux Meeting on Systems and Control. pp. 119–141.
- Lalwani, M.A., Kawabe, H., Mays, R.L., Hoffman, S.M., Avalos, J.L., 2021. Optogenetic control of microbial consortia populations for chemical production. *ACS Synth. Biol.* 10, 2015–2029.
- Olson, E.J., Hartsough, L.A., Landry, B.P., Shroff, R., Tabor, J.J., 2014. Characterizing bacterial gene circuit dynamics with optically programmed gene expression signals. *Nat. Methods* 11, 449–455.
- Pohlodek, J., Morabito, B., Schlauch, C., Zometa, P., Findeisen, R., 2022. Flexible development and evaluation of machine-learning-supported optimal control and estimation methods via HILO-MPC. arXiv:2203.13671.
- Rawlings, J., Mayne, D., Diehl, M., 2020. Model predictive control: theory, computation and design, 2nd edition. ed. Nob Hill Publishing, LLC, Santa Barbara.
- Roell, G.W., Zha, J., Carr, R.R., Koffas, M.A., Fong, S.S., Tang, Y.J., 2019. Engineering microbial consortia by division of labor. *Microb. Cell Factories* 18, 35.
- Zhang, D., Del Rio-Chanona, E.A., Petsagkourakis, P., Wagner, J., 2019. Hybrid physics-based and data-driven modeling for bioprocess online simulation and optimization. *Biotechnol. Bioeng.* 116, 2919–2930.

# Model-based optimization of a recombinant *E. coli* fed-batch process for best time-space yield and conversion efficiency

Julian Kager<sup>a</sup>, Nora Horst<sup>b</sup>, Christoph Herwig<sup>b</sup> and Ioannis V.Skiadas<sup>a</sup>

<sup>a</sup>*Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 228A, 2800 Kgs. Lyngby, Denmark*

<sup>b</sup>*Institute of Chemical, Environmental and Bioscience Engineering, TU Wien, 1040 Vienna, Austria*  
*jukager@kt.dtu.kt*

## Abstract

Within this contribution we use a mechanistic model to optimize the feed addition and harvest time point of the induction phase of a recombinantly produced protein in *E. coli*. The model includes the metabolic load of the forced product formation and its effects on the production rate as well as the biomass conversion yield. The model was fitted to three experiments with different constant substrate feed rates during the production phase. Under usage of the calibrated model, optimal substrate addition and harvest time points with the objective to improve the time-space yield were determined. In addition, a constant feed rate was compared to an increasing feed rate with the aim to maintain a constant biomass specific substrate uptake. Different optimal combinations were identified and compared considering time and resource efficiency. The work shows how recombinant protein production and the inevitable metabolic load can be incorporated into a simple mechanistic model. Once calibrated to a limited number of experiments the model is suited for the design of a resource and time optimal induction phase. Future experiments and uncertainty analysis will show how well the model predicted the real optimal operating points.

**Keywords:** bioprocessing, optimization, production phase modelling, simulation, recombinant protein production

## 1. Introduction

Fed-batch processes using genetically modified *E. coli* bacteria are widely used to produce high value recombinant proteins (Selas Castineiras et al., 2018). State of the art production processes are divided into three main phases, including a batch and fed-batch phase to produce a high amount of active biomass followed by the production phase (Yee and Blanch, 1992). The cell behavior during the first two phases is well understood and optimized to reach high biomass concentrations within short times (Lee, 1996) and to avoid by product formation (Abadli et al., 2021). Due to the metabolic load of the heterologous production of the target protein, cell metabolism changes in course of the induction phase (Neubauer et al., 2003). Therefore, modelling, optimization and control of the induction phase is a challenging task and harvest time points and feed addition are often static parameters derived from historic experiments, which often leads to suboptimal operation and low time space yields. Models can help to improve the exploration of the design space by simulating all different combinations and identifying optimal operation ranges. Möller et al. (2019) for examples used a model to explore the

design space of a CHO cell process to identify optimal feed addition. Bano et al. (2019) extended this concept and recalculated the design space in real-time as soon novel information was available to the model. For recombinant protein production in *E. coli* strong influences on product formation from the feed rate was reported for different products (Mears et al., 2017; Huang et al., 2012; Chen et al., 2012), but due to the lack of proper, mechanistic models, successful examples for insilico exploration and optimization of the production phase are still missing. Although constant substrate feeds are often used during the production phase, recent works show possibilities and potentials to control the specific substrate uptake rate (Kager et al., 2022b) or even the product formation rate (Kager et al., 2022a) by manipulating the feed rate of the limiting substrate. Although the latter being very promising, the control of the specific substrate uptake is easier realizable and was already shown in different applications (Kager et al., 2020; Spadiut et al., 2013; Picon et al., 2005) and can be therefore regarded as a good alternative to constant substrate addition rates.

Within this contribution we show how a previously developed mechanistic model describing the production phase of a recombinantly produced protein can be used to determine optimal feed addition in combination with harvest time. Optimal constant feed addition will be compared to optimal set-points of the cell specific substrate uptake for their time and resource efficiency. After describing the three calibration experiments, the model and its fitting to the experiments are presented. The results show the design space of the analyzed process and the dynamics of the optimal process followed by the conclusion.

## 2. Methods

### 2.1. Experimental data

A modified K12 *E. coli* strain with a rhamnose-inducible expression system (rhaBAD promoter), producing a single chain antibody fragment with a transporter sequence to be transported into the periplasm was used. Fermentations were conducted in a DASGIP multibioreactor system (Eppendorf, Germany). Composition of the used minimal media was based on Wilms et al. (2001). Temperature was kept at 35°C, stirrer speed at 1400 rpm and aeration at 1.4 vvm for the whole process. The pH was controlled at 7.0 ± 0.2 with addition of 12.5 % NH<sub>4</sub>OH solution and the dissolved oxygen (DO<sub>2</sub>) was kept over 25% by supplementing pure oxygen to the air. The reactors with an initial glycerol concentration were inoculated with 25 mL of an overnight shake flask culture (30°C and 170 rpm). The 20 g/L glycerol in the batch phase was consumed after approx. 12 h and a fed-batch phase with an exponential feed ramp with a  $\mu$  of 0.14 according to the equation given in Lee (1996) was applied to reach high biomass concentrations of 40 gL<sup>-1</sup> after another 12 h runtime. Recombinant protein production was induced by a one-point addition of sterile filtrated rhamnose solution (1.5 g L-rhamnose in 10 mL H<sub>2</sub>O) and the glycerol feed rate was set to different constant values (*high* = 14.5 mLh<sup>-1</sup>; *medium* = 8.5 mLh<sup>-1</sup>; *low* = 2.2 mLh<sup>-1</sup>).

Periodical offline measurements were taken and analyzed. Biomass dry content was determined by centrifugation at 4500 g, 10 min, 4°C including one washing step before drying at 105°C for min. 72h. For product concentration, homogenized, centrifuged and gel filtrated (PD MiniTrap g-25) cell suspension was applied on a protein G affinity column (HiTrap ProtG (GE Healthcare; USA) with a flow rate of 2 mlmin<sup>-1</sup> (20 mM phosphate buffer) at 25°C and elution with a change of the pH from 7.4 to 2.5). Acetate and glycerol concentrations were quantified from the supernatant by enzymatic,

photometric principle in a robotic system (BioHT, Roche, Germany) and were always under the detection limit during the induction phase.

## 2.2. Model equations

The system differential equations for the ideally stirred tank reactor in fed-batch mode are:

$$\begin{aligned}
 \frac{dc_X}{dt} &= \mu c_X - \frac{F_{in}}{V_R} c_X \\
 \frac{dc_S}{dt} &= -q_S c_X + \frac{F_{in}}{V_R} (c_{S,in} - c_S) \\
 \frac{dc_P}{dt} &= q_P c_X - \frac{F_{in}}{V_R} c_P \\
 \frac{dV_R}{dt} &= F_{in}; \quad \frac{dS_{met}}{dt} = q_S
 \end{aligned} \tag{1}$$

describing the concentration changes over time of the: biomass  $c_X$ , glycerol  $c_S$ , product  $c_P$ , metabolized substrate  $S_{met}$  and the reactor volume  $V_R$ . Glycerol is added to the system by the substrate inflow  $F_{in}$  with a concentration  $c_{S,in}$ . The biological conversion rates, namely the growth rate  $\mu$ , the substrate consumption rate  $q_S$  and the product formation rate  $q_P$  are described by the following reaction kinetics:

$$\begin{aligned}
 q_S &= \frac{q_{S,max} c_S}{c_S + K_S} \\
 \mu &= Y_{X/S} (q_S - m_S) \\
 Y_{X/S} &= \frac{Y_{X/S,max} \exp(-S_{met} K_{YX/S})}{S_{met}} \\
 q_P &= q_{P,max} \frac{q_S}{q_S + K_{SqS}} \frac{\frac{S_{met}}{K_{IqP}} + S_{met} + K_{SqP}}{S_{met}}
 \end{aligned} \tag{2}$$

The substrate uptake rate  $q_S$  is described by a Monod kinetic with  $q_{S,max}$  being the maximum uptake rate in function of the glycerol concentration  $c_S$  and the half saturation constant  $K_S$ . Biomass growth  $\mu$  is derived from consumed substrate  $q_S$  by substrate conversion yield  $Y_{X/S}$  reduced by the substrate needed for cell maintenance  $m_S$ . Due to the metabolic load  $Y_{X/S}$  decreases by an asymptotic decay ( $K_{YX/S}$ ) in function of the metabolized substrate ( $S_{met}$ ). The product formation  $q_P$  is composed by a Monod term describing the dependency of  $q_{P,max}$  on the substrate uptake  $q_S$  with the half saturation constant  $K_{SqS}$  and a Haldane term to describe the underlying start up and decline phase with  $S_{met}$  as a trigger,  $K_{SqP}$  as the delay coefficient,  $K_{IqP}$  as the decay coefficient and  $k$  as the Haldane exponent determining the shape of the decay.

Table 1: Model Parameter

Parameter	Value	Unit	Description
$c_{S,in}$	850	$gL^{-1}$	feed concentration
$q_{S,max}^*$	1.0	$gg^{-1}h^{-1}$	max. substrate uptake
$K_S^*$	0.0050	$gL^{-1}$	half saturation constant
$m_S$	0.02	$h^{-1}$	maintenance constant
$Y_{X/S,max}^*$	0.47	$gg^{-1}$	max. growth yield
$K_{YX/S}$	0.34	$gg^{-1}$	growth yield decay
$q_{P,max}$	0.0066	$gg^{-1}h^{-1}$	max. production rate
$K_{SqS}$	0.082	$gg^{-1}h^{-1}$	$q_P$ affinity to $q_S$
$k$	4.6	–	Haldane exponent
$K_{IqP}$	3.8	$(gg^{-1})k^{-1}$	product decay
$K_{SqP}$	0.096	$gg^{-1}$	product formation delay

The model was fitted to three experimental data sets, with three different constant feed rates, displayed in figure 1. The model parameters are given in table 1. As fitting criterion the weighted sum of squared errors between measurements of biomass, product and sugar concentrations and the respective model simulations of all three experiments was minimized by a local optimizer (Fmincon:MATLAB). All calculations were done in MATLAB (R2022b, Mathworks, USA) and the differential equations were solved by ode15s. To keep the substrate uptake constant the following equation was used to derive the corresponding feed rate based on the current biomass concentration and aimed  $q_S$  setpoint ( $q_S \ll q_{S,max}$ ).

$$F_{in} = \frac{q_S c_X V_R}{c_{S,in} - c_S} \quad (3)$$

### 3. Results and Discussion

#### 3.1. Process & Model dynamics

In figure 1 the results of the model simulation of the three calibration experiments are shown in comparison to the measured product concentrations. Overall, with a normalized root mean square error (NRMSE) of 11,5 % the model has an acceptable accuracy considering the big feed rate range ( $2.2 - 14.5 \text{ mLh}^{-1}$ ) of the calibration experiments. From the model and the measurements, it can be seen that higher feed rates lead to a faster product formation but also to a faster stagnation. The lowest feed rate leads to highest product titers. Compared to the other processes it takes however longer to reach those concentrations, with a duration of up to 100 h.

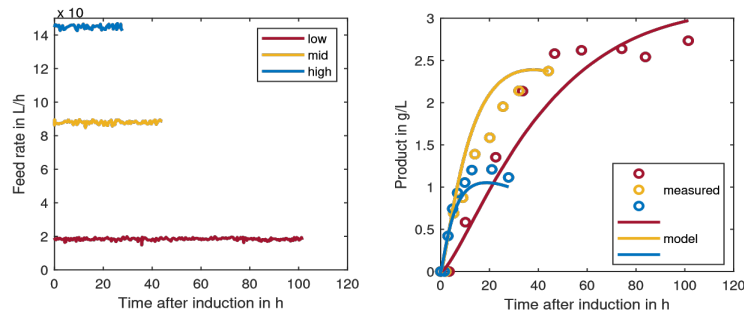


Figure 1: Applied constant feed rates of the three model calibration experiments and alignment of model output to measured product concentrations with NRMSE of 10,7% for low, 12.7 % for mid and 11.2 % for high experiment.

#### 3.2. Design space exploration

Although the fitted model displayed in figure 1, is not perfect it can be used to explore the design space and to determine potentially optimal operating regions. In order to identify optimal combinations of feed rates,  $q_S$  setpoints and harvest time-points the time space yields (TSY) were calculated considering the duration of the induction phase in addition to batch (12 h), fed-batch (12 h) and setup times (7 h). The resulting design spaces for the constant feed rate (a) and the controlled  $q_S$  (b) are displayed in figure 2. According to the model the optimal TSY is reached by applying constant feed rates around 5 mL/h and harvesting after 23 h, which is between the *low* and the *mid* feed rate experiments. It seems that for this process, high constant feed rates are highly unfavorable

with very low TSY's. A similar pattern can be seen for the  $q_S$  controlled induction phases (figure 2 b). Also, here smaller setpoints ( $< 0.1 \text{ g g}^{-1}\text{h}^{-1}$ ) and harvest times between 18 - 40 hours, lead to best TSY's. In contrast to the constant feed rate two optimal regions are identified. A setpoint of  $0.1 \text{ g g}^{-1}\text{h}^{-1}$  in combination with a harvest time of 19.2 h and a smaller setpoint of  $0.055 \text{ g g}^{-1}\text{h}^{-1}$  with a harvest timepoint at 30.2 hours.

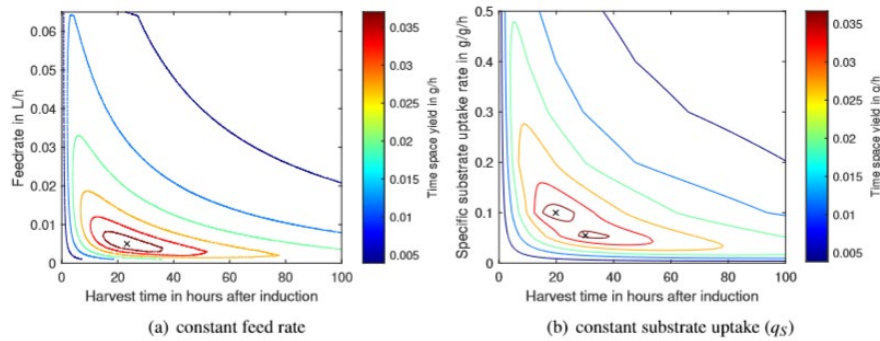


Figure 2: Design spaces for optimal time space yield for induction phases with different constant feed rates (a) and specific substrate uptake ( $q_S$ ) setpoints (b) in dependence on the harvest time-point.

In figure 3 the trajectories of the three potential, optimal conditions are displayed with some performance criteria summarized in table 2. It can be seen that the different optimal conditions differ in their feed profiles, which are all in the lower range (medium feed = 8.5 mL/h). The constant feed rate results in a steadily decreasing  $q_S$  compared to the two optimal constant  $q_S$  setpoints. The product dynamics reveal a similar behavior as seen in the calibration experiments. Hereby, higher feed rates lead to a faster product increase but the lowest setpoint to the overall highest product titer (see also table 2). Although, the constant feed rate and the higher  $q_S$  setpoint have slightly higher TSY's, due to faster product formation dynamics, the low  $q_S$  setpoint reaches the highest product titer with overall lowest consumed glycerol and is therefore the most resource efficient.

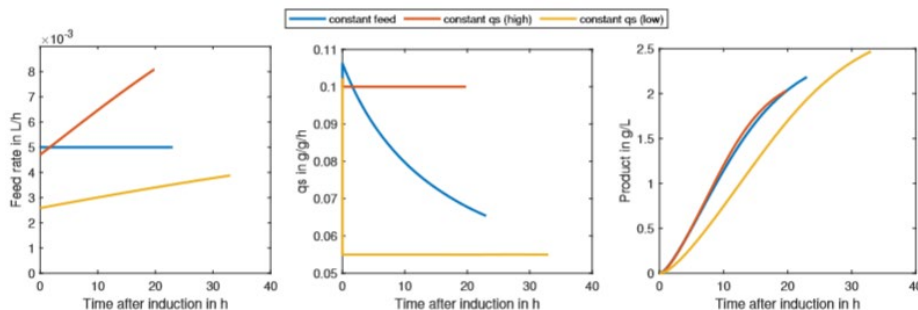


Figure 3: Model simulations of identified optimal constant feed and two  $q_S$  setpoints

Table 2: Output of the three potentially good performing processes

Feeding	Harvest (h)	TSY (g/h)	Product (g/L)	Biomass (g/L)	Glycerol (g)
$F_{in} = 0.005 \text{ L/h}$	23.0	0.0389	2.19	58.3	97.8
$q_S = 0.10 \text{ g/g/h}$	19.8	0.0386	2.03	61.0	108.3
$q_S = 0.055 \text{ g/g/h}$	30.2	0.0372	2.47	54.2	91.4

#### 4. Conclusion

In this contribution a mechanistic model was fitted to the production phase of three experiments with different feed addition. Based on the obtained model the design space could be explored and promising regions identified. Based on this methodology potentially optimal experiments can be planned and different control strategies can be bench-marked with only few available experiments. With a proper uncertainty analysis of the preliminary model, a more probability-based approach could be followed in future to predict the most probable design space rather than optima which are highly uncertain due to model-plant mismatches. To show the validity of the obtained results, verification experiments and a proper model uncertainty analysis are necessary. Future verification experiments are also the basis to further improve the underlying model.

#### Acknowledgements

This work obtained funding from the Novo Nordisk Foundation, start package grant tenure track Assistant Professor on bioprocess digitalization and control (Nr. 111279).

#### References

- M. Abadli, L. Dewasme, S. Tebbani, D. Dumur, A. V. Wouwer, 2021. An experimental assessment of robust control and estimation of acetate concentration in *Escherichia coli* b121 (de3) fed-batch cultures. *Biochemical Engineering Journal*, 108103.
- G. Bano, P. Facco, M. Ierapetritou, F. Bezzo, M. Barolo, Aug. 2019. Design space maintenance by online model adaptation in pharmaceutical manufacturing. *Computers & Chemical Engineering* 127, 254–271.
- X. Chen, L. Liu, J. Li, J. Liu, G. Du, J. Chen, Feb. 2012. Optimization of glucose feeding approaches for enhanced glucosamine and N-acetylglucosamine production by an engineered *Escherichia coli*. *Journal of Industrial Microbiology & Biotechnology* 39 (2), 359–365.
- C. Huang, Jr., H. Lin, X. Yang, 2012. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *Journal of Industrial Microbiology & Biotechnology* 39 (3), 383–399.
- J. Kager, J. Bartlechner, C. Herwig, S. Jakubek, Mar. 2022a. Direct control of recombinant protein production rates in *E. coli* fed-batch processes by nonlinear feedback linearization. *Chemical Engineering Research and Design*, S0263876222001460.
- J. Kager, N. Horst, J. Bartlechner, C. Herwig, S. Jakubek, Jan. 2022b. Generic model based control of different specific rates in recombinant *E. coli* fed-batch processes. *IFAC-PapersOnLine* 55 (7), 756–761.
- J. Kager, A. Tuveri, S. Ulonska, P. Kroll, C. Herwig, Mar. 2020. Experimental verification and comparison of model predictive, PID and model inversion control in a *Penicillium chrysogenum* fed-batch process. *Process Biochemistry* 90, 1–11.
- S. Y. Lee, 1996. High cell-density culture of *Escherichia coli*. *Trends in biotechnology* 14 (3), 98–105.
- L. Mears, S. M. Stocks, G. Sin, K. V. Gernaey, Mar. 2017. A review of control strategies for manipulating the feed rate in fed-batch fermentation processes. *Journal of Biotechnology* 245, 34–46.
- J. Moller, K. B. Kuchemüller, T. Steinmetz, K. S. Koopmann, R. Pörtner, May 2019. Model-assisted Design of Experiments as a concept for knowledge-based bioprocess development. *Bioprocess and Biosystems Engineering* 42 (5), 867–882.
- P. Neubauer, H. Lin, B. Mathiszik, 2003. Metabolic load of recombinant protein production: inhibition of cellular capacities for glucose uptake and respiration after induction of a heterologous gene in *Escherichia coli*. *Biotechnology and bioengineering* 83 (1), 53–64.
- A. Picon, M. J. Teixeira de Mattos, P. W. Postma, Apr. 2005. Reducing the glucose uptake rate in *Escherichia coli* affects growth rate but not protein production. *Biotechnol Bioeng* 90 (2), 191–200.
- T. Selas Castineiras, S. G. Williams, A. G. Hitchcock, D. C. Smith, 2018. *E. coli* strain engineering for the production of advanced biopharmaceutical products. *FEMS microbiology letters* 365 (15), fny162.
- O. Spadiut, S. Rittmann, C. Dietzsch, C. Herwig, Jan. 2013. Dynamic process conditions in bioprocess development. *Engineering in Life Sciences* 13 (1), 88–101.
- B. Wilms, A. Hauck, M. Reuss, C. Syldatk, R. Mattes, M. Siemann, J. Altenbuchner, 2001. High-cell-density fermentation for production of L-n-carbamoylase using an expression system based on the *Escherichia coli* rhabd promoter. *Biotechnology and bioengineering* 73 (2), 95–103.
- L. Yee, H. Blanch, 1992. Recombinant protein expression in high cell density fed-batch cultures of *Escherichia coli*. *Biotechnology* 10 (12), 1550–1556.

# Closed-loop optimization of high-throughput robotic platforms for reproducible bioprocess development

Federico M. Mione<sup>a</sup>, Judit Aizpuru<sup>b</sup>, Martin F. Luna<sup>a</sup>, Pablo Rodriguez Bahamon<sup>b</sup>, Jong Woo Kim<sup>b</sup>, Ernesto C. Martinez<sup>a,b</sup>, M. Nicolas Cruz B.<sup>b</sup>

<sup>a</sup> *INGAR (CONICET-UTN), Avellaneda 3657, S3002GJC Santa Fe, Argentina*

<sup>b</sup> *KIWI-biolab, Bioprocess Engineering, TU Berlin, Ackerstrasse 76, 13355 Berlin, Germany*  
*mariano.n.cruzbournazou@tu-berlin.de*

## Abstract

To confidently accelerate bioprocess development using high-throughput robotic platforms, parallel experiments should be reproducible whereas the generated data must be FAIR. To this aim, higher levels of automation for planning, redesigning, executing, and monitoring of related experimental-computational workflows are needed. In this work, an automation hierarchy for closed-loop optimization based on Bayesian update, belief propagation and Thompson sampling is proposed. For online redesign of parallel dynamic experiments, asynchronous *rolling out beliefs about beliefs* based on simulated data is used for counterfactual assessment of the effect of alternative (re)design decisions on the distributions of model parameters. A computational implementation of the Bayesian automation hierarchy using Apache Airflow is described. Simulation results obtained using an *E. coli* case study are presented to argue that parallel dynamic experiments made using the proposed Bayesian hierarchy are reproducible whereas experimental-computational workflows and data are FAIR by design.

**Keywords:** automation hierarchy, closed-loop experimentation, belief propagation, FAIR principles, reproducibility crisis.

## 1. Problem statement

Bioprocess development requires to explore efficiently huge design spaces for specific bioproducts and involves efficient strain selection, bioprocess optimization, scale-up, and optimal control strategies for reproducible industrial production. The shift towards high-throughput technologies and automation in bioprocess development (Cruz Bournazou et al., 2017; Bromig et al., 2022) demands a more autonomous approach for generating informative data sets to support model-based experimental design in the face of uncertainty. For reproducible bioprocess development and decision-making provenance, experimental data sets used for predicting the dynamic behaviour of genetically modified microorganism and the physiological effect of different combinations of process parameters on the productivity levels after scale-up must be obtained by following FAIR principles (Wilkinson et al., 2016; Mione et al., 2022). A significant impact is expected from using these principles in model-based autonomous experimentation for strain selection and the optimization of bioprocess parameters based on FAIR data, experimental protocols, and computational workflows. In this work, an automation hierarchy for integrating Bayesian inference with experimental (re)design and execution control using probabilistic models and belief propagation is proposed.



## 2. Bayesian automation hierarchy

First and foremost, automating the execution of an experiment requires that the specific objective for data gathering is clearly stated, the *a priori* available knowledge be explicit modeled, and the initial uncertainty quantified. Without any loss of generality, it is assumed here that a causal probabilistic model (see Martinez et al, 2021 for details) of the microorganism dynamic behavior is going to be used to guide experimental off-line design, online redesign, execution monitoring and control. The aim of the asynchronous parallel experiments would be to gather FAIR data to increase the model predictive power so that biomass production is maximized. Alternatively, the objective of the experiment would be strain screening for physiological robustness after scaling up. Gathering data for model selection would also be the aim of an experiment. In this work, it is assumed that the *a priori* uncertainty is summarized in the means and variances corresponding to the (normal) distributions of the parameters for the causal dynamic model.

A schema of the proposed automation hierarchy in the chosen Bayesian setting is given in Fig. 1. The top layer deals with off-line design where the *a priori* optimal sequence of control actions for a given objective is calculated alongside with the expectations regarding observations based on prior distributions for model parameters. In the middle layer, as new data is available, the *posterior* distributions are revised using variational Bayesian inference. Then, the optimal sequence of redesign decisions is re-estimated and future data expectations, which are needed to monitor the experiment execution, are updated accordingly. In the bottom layer, a predictive control strategy is used to safely implement (re)design decisions by properly handling both physiological and platform-related constraints using (stochastic) model predictive control.

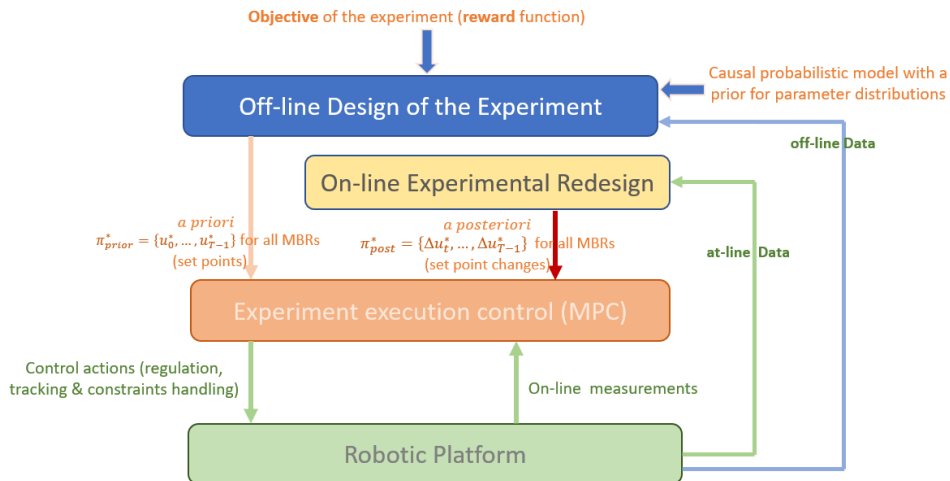


Figure 1. Bayesian hierarchy for closed-loop optimization of robotic platforms

At their three layers, the unifying approach is that by considering the counterfactual consequences of alternative (re)design decisions in generating alternative sampled data, belief propagation (Friston et al., 2021; Martinez et al., 2021) is used to update posterior distributions about the evolution of hidden states and the estimated sequence of optimal redesign decisions for each MiniBioReactor (MBR) in the platform based on simulated sampled data and Thompson sampling of the posterior distributions for model parameters. This recursive form of propagating beliefs over the shrinking horizon of a dynamic

experiment effectively implements a deep tree search over alternative sequences of optimized (re)design decisions and their corresponding expectations regarding state evolution and sampled data at different times for dynamic parallel experiments.

For asynchronous *rolling out beliefs about beliefs* in each MBR, the same recursive procedure in Fig. 2 is used for both off-line experimental design and on-line redesign as new observations from all MBRs are available to update the distributions  $f(\theta)$  for model parameters. The internal loop implements a forward pass where, for a given MBR, a *prior* density is increasingly converted into a *posterior* density for model parameters upon simulated data in a rollout where a stagewise greedy design procedure based on Thompson sampling (Russo et al., 2018) is recursively implemented. At each redesign iteration  $k$ , a decision is chosen to implement in the  $i$ th MBR by solving an optimization problem for the expected look-ahead reward based on the  $r_t(\hat{y}_{t+1}^i)$  computed from simulated data  $\hat{y}_{t+1}^i$  for alternative design decisions in a bounded set  $\Omega$  and accounting for physiological and platform constraints  $C(x_{i,t}, u_{i,t})$ :

$$u_{i,t}^* = \operatorname{argmax}_{u_i \in \Omega} r_t(y_{t+1}^i), \text{ subject to: } C(x_{i,t}, u_{i,t}) \geq 0 \quad (1)$$

*Inputs:*  $T, K, x_{i,0}$ , prior  $f(\theta)$ ,  $n_{MBR}$ , state evolution and observation functions  $f, g$

▷ **For**  $t = 1$  to  $T - 1$

*Infer* current state  $\hat{x}_{i,t}$  using  $u_{i,t-1}^*$ ,  $\hat{x}_{i,t-1}$  and Thompson *Sampling* of prior  $f(\theta)$

▷ **For**  $k = 1$  to  $K$

$\hat{f}(\theta_k) = f(\theta)$

**While**  $t < T$  (Forward Pass)

▷ **For**  $i = 1, 2, \dots, n_{MBR}$  **do**

Thompson *Sampling* of the prior  $\hat{f}(\theta_k)$ :  $\Theta_k = \mathbf{TS}[\hat{f}(\theta_k)]$

$u_{i,t}^k = \operatorname{argmax}_{u_i \in \Omega} (r_t(\hat{y}_{t+1}^i) | \Theta_k, \hat{x}_{i,t})$ , S. to:  $f, g$  &  $C(x_{i,t}, u_{i,t}) \geq 0$

*Simulate* redesign using  $u_{i,t}^k$  and predict  $r_{t+1}(\widehat{y}_{t+1}^k | u_{i,t}^k, \hat{x}_{i,t})$

*Update* prior:  $\hat{f}(\theta_k) \leftarrow \hat{f}(\theta_k | u_{i,t}^k, \hat{x}_{i,t})$  using  $(u_{i,t}^k, \hat{x}_{i,t})$

*Accumulate* reward:  $\mathcal{R}_k = \mathcal{R}_k + r_{t+1}(\widehat{y}_{t+1}^k | u_{i,t}^k)$

▷ **End for**

**End while**

Define the redesign policy:  $\pi_t^k = \{u_{i,t}^k, \dots, u_{i,T-1}^k\}$  with its corresponding  $\mathcal{R}_k$

▷ **End for**

*Rank* policies  $\pi_{i,t}^k, k = 1, \dots, K$ , using  $\mathcal{R}_k$

*Select* the best policy  $\pi_{i,t}^* = \{u_{i,t}^*, \dots, u_{i,T-1}^*\}$  with the highest  $\mathcal{R}_k$

*Redesign the experiment* in the MBR  $i$  using only  $u_{i,t}^*$  and measure  $y_{i,t+1}$

*Update* prior:  $f(\theta) \leftarrow f(\theta | u_{i,t}^*, y_{i,t+1})$  using data  $(u_{i,t}^*, y_{i,t+1})$  from all MBRs

▷ **End for**

*Outputs:*  $\pi^* = \{u_{i,1}^*, \dots, u_{i,T-1}^*\}, (u_{i,t}^*, y_{i,t+1}) \ i = 1, \dots, n_{MBR}, f(\theta)$

Figure 2. Asynchronous rolling out beliefs about beliefs for online redesign in an MBR.

where  $r$  is a *reward* function for observations and their underlying hidden states which accounts for the specific objective of the experiment. The predicted observation  $\hat{y}_{t+1}$  and the estimated optimal action  $u_{i,t}^*$  are then used to update the posterior distribution for model parameters  $\hat{f}(\theta_k)$  to be used at  $t+1$ . Thompson sampling is again applied and using

Eq. (1) the action  $u_{i,t}^*$  is calculated and the corresponding simulated observation  $\hat{y}_{t+2}$  is computed, and so on and so forth. This forward pass ends when  $u_{i,T-1}^*$  is computed and the *optimal* redesign sequence  $\pi_{i,t}^* = \{u_{i,t}^*, \dots, u_{i,T-1}^*\}$  is completely defined for a given MBR. For each iteration of the outer loop, a different sequence  $\pi_k$  is computed. The generated sequences are then ranked based on their corresponding cumulative rewards  $\mathcal{R}_k$  over the planning horizon  $[t, \dots, T]$ . From the top ranked redesign policy  $\pi_{i,t}^*$ , only the first action  $u_{i,t}^*$  is used for redesigning the dynamic experiment in the *i*th MBR and, at the next sampling time  $t+1$ , the observation  $y_{i,t+1}$  is obtained. Using sampled data  $(u_{i,t}^*, y_{i,t+1})$  from all MBRs, the joint posterior distribution  $f(\boldsymbol{\theta})$  for the parameters of the probabilistic model is updated using variational Bayesian inference (based on the Laplace approximation) and the master (external) loop begins re-estimating the optimal redesign sequence  $\pi_{i,t+1}^*$  for each MBR from  $t+1$  until the end of the experiment at time  $T$  using belief propagation. It is worth noting that online redesign in each MBR is only coupled with others' MBRs redesign by the global update of the posterior distribution  $f(\boldsymbol{\theta})$ .

### 3. Implementation in Apache Airflow

The FAIR principles (Wilkinson et al., 2016), describe a set of requirements for data management and stewardship to make research data **F**indable, **A**ccessible, **I**nteroperable, and **R**eusable. The FAIR principles are a set of guidelines that aim to maximize the value and usefulness of experimental data and highlight the importance of making experimental protocols and workflows digital objects findable and reusable by others. To this aim, the representation of the automation hierarchy in Fig. 1 and the (re)design algorithm in Fig. 2 using directed acyclic graphs (DAGs) foster their implementation in the Apache Airflow ecosystem to enforce FAIR principles by design in high-throughput experimentation facilities. For details on making experimental-computational workflows FAIR in Apache Airflow (Harenslak and de Ruiter, 2021), the interested reader is referred to (Mione, et al., 2022) for details.

In Fig. 3, an abstract DAG is used to represent the hierarchy in Fig. 1 as a pipeline for implementing computational tasks such as update of parameter distributions as soon as new sampled data are available from all MBRs, followed by asynchronous experimental redesign using the algorithm in Fig. 2 and then predictive control of redesign changes to the set-points for controlled variables to guarantee feasibility. It is worth noting the importance of execution control nodes since due Thompson sampling changes to the set-points of control loops may violate physiological constraints. For FAIRness and reproducibility, is key to save all seeds used for random number generation in posterior sampling at each iteration  $k$  in the belief propagation algorithm for each MBR. As a result, the parameter values used for solving optimization problems in Eq. (1) will be known.

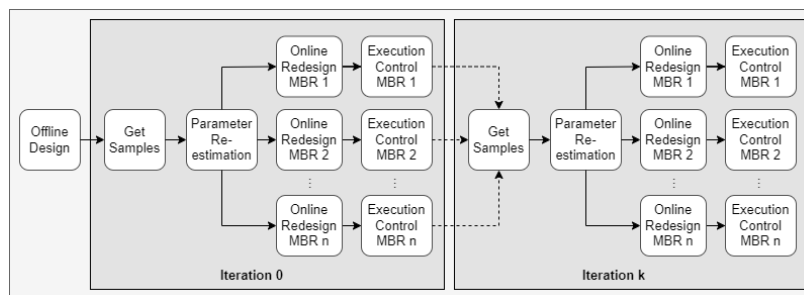


Figure 3. The Bayesian hierarchy with belief propagation as an Airflow pipeline.

#### 4. Case study

The proposed method was tested using artificial data generated based on the model proposed in Nickel et al. (2017), consisting of six differential equations and eighteen parameters. Eight MBRs for *Escherichia coli* fed-batch cultivations were run over a six-hour period. Samples are taken every thirty minutes and biomass ( $X$ ), glucose ( $Glc$ ) and acetate ( $Ac$ ) concentrations are measured at-line, whereas the dissolved oxygen tension ( $DOT$ ) in each MBR is measured continuously using an online sensor. Based on new data, the sequence of glucose pulses for each MBR are computed to implement model-based optimized feeding strategies (see Fig. 2) so that the experiment is redesigned online to maximize the information content. The aim of the experimental design is to find *a posteriori* an optimal feeding profile that maximizes the amount of biomass obtained at the end of the run and prevents the  $DOT$  to drop below 20% throughout. For details regarding all the assumptions, models, software packages, parameters and hyper-parameters used to implement the automation hierarchy, the reader is referred to the repository:

<https://git.tu-berlin.de/bvt-htbd/kiwi/tf2/experimental-computational-workflows/-/tree/escape>

Once the parallel eight experiments have been completed, the probabilistic model is the updated using the complete data set. The Bayesian nature of the resulting probabilistic model allows for a chance-constrained formulation of the optimization problem

$$u^* = \operatorname{argmax}_{u \in \mathcal{U}} E[m_{X,end}] \quad (2.1)$$

$$\text{Subject to } \Pr(DOT_{min} > 20) > 95\% \quad (2.2)$$

where  $m_{X,end}$  is the biomass amount obtained (in g/L) at the end of a testing experiment and  $DOT$  is in percentage of saturation. The decision variables are the duration of the batch phase and the feeding profile. Some results obtained using the *posterior* parameter distributions are shown in Fig. 4 to highlight the effectiveness of the proposed approach.

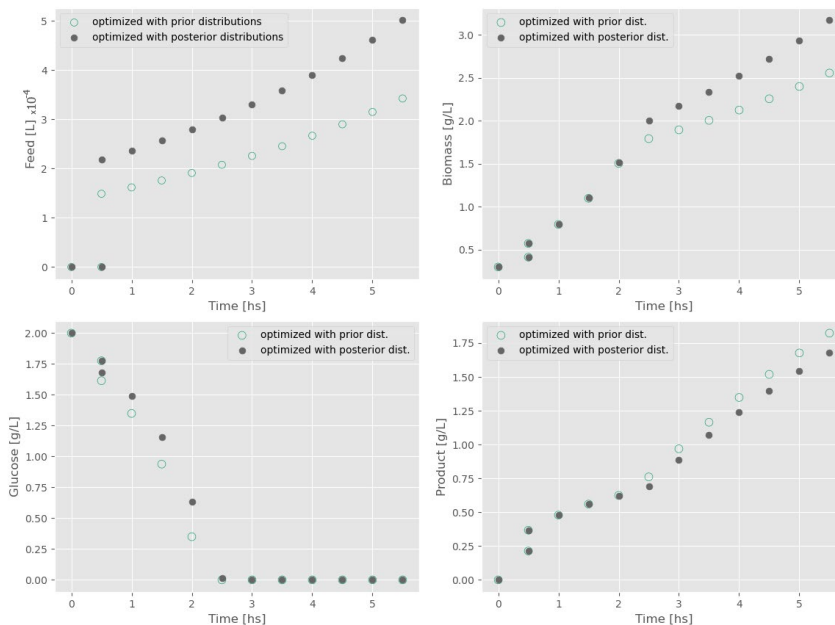


Figure 4. Feed rate optimization with prior and posterior distributions compared.

As it is shown, the *a priori* optimized feeding profile is more cautious than the one with the probabilistic model using the *a posteriori* distributions of model parameters. For the later, a higher amount of biomass is obtained which is aligned with the experiment objective which bias data gathering in the operating region where biomass is maximized whereas the dissolved oxygen tension is safely kept above the threshold value of 20% during the optimized run. The effect of data bias is clear in Fig. 5 which describes the evolution of the distributions of some sensitive model parameters as the iteration count  $k$  increases. The parameter distributions for Iter. #1 are the priors before the experiment.

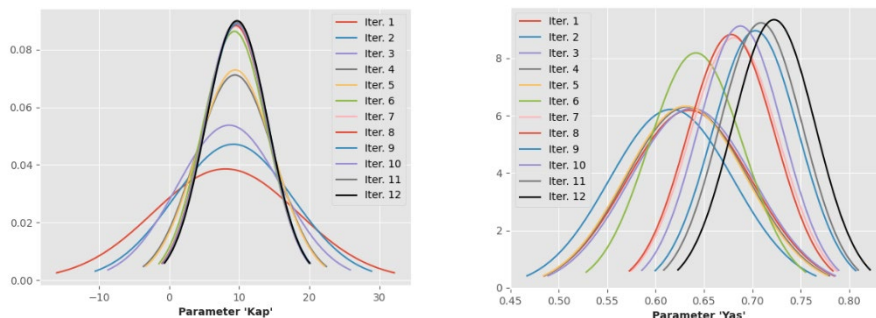


Figure 5. Evolution of two distributions of model parameters during the experiment.

## 5. Concluding remarks

A new automation hierarchy for closing the loop in high-throughput robotic platforms was proposed. The implementation of the hierarchy in Apache Airflow makes possible to reproduce the computational experiments and make sampled data FAIR. Finally, the asynchronous belief propagation algorithm in Fig. 2 is best suited for parallel processing.

## References

- L. Bromig, N. von den Eichen, D. Weuster-Botz, 2022, Control of parallelized bioreactors I: dynamic scheduling software for efficient bioprocess management in high-throughput systems, *Bioprocess and Biosystems Engineering*, 45,1927–1937.
- M. N. Cruz Bournazou, T. Barz, D. B. Nickel, D. C. Lopez Cárdenas, F. Glauche, A. Knepper, P. Neubauer, 2017, Online optimal experimental re-design in robotic parallel fed-batch cultivation facilities. *Biotechnol. Bioeng.*, 114, 610–619.
- K. Friston, L. Da Costa, D. Hafner, C. Hesp, T. Parr, 2021, Sophisticated inference, *Neural Computation*, 33, 713–763.
- B. Harenslak, J. de Ruiter, 2021, *Data Pipelines with Apache Airflow*. Manning Publications.
- E. C. Martinez, J. W. Kim, T. Barz, M. Nicolas Cruz B., 2021, Probabilistic modeling for optimization of bioreactors using reinforcement learning with active inference, *Proc. of ESCAPE-31 2021*, 50 (Part A), 419-424.
- F. M.Mione, A. N. Silva, M. F.Luna, M. N. Cruz B., E. C.Martinez, 2022, Managing Experimental-Computational Workflows in Robotic Platforms using Directed Acyclic Graphs, *Computer Aided Chemical Engineering*, 49, 1495-1500.
- D. B. Nickel, M. N. Cruz-Bournazou, T. Wilms, P. Neubauer, & A. Knepper, 2017, Online bioprocess data generation, analysis, and optimization for parallel fed-batch fermentations in milliliter scale, *Engineering in the Life Sciences*, 17, 11, 1195-1201.
- D. J. Russo, B. Van Roy, A. Kazerouni, Ia. Osband, Z. Wen, 2018, A Tutorial on Thompson Sampling, *Foundations and Trends in Machine Learning*, 11, 1, 1-96.
- M. D. Wilkinson, et al., 2016, The FAIR guiding principles for scientific data management and stewardship. *Nature* 3:160018 DOI 10.1038/sdata.2016.18.

# Dynamic microbial-community metabolic modeling for yogurt fermentation based on the metagenome of starter culture

Sizhe Qiu<sup>a,b</sup>, Zhijie Yang<sup>b</sup>, Hong Zeng<sup>b</sup>, Bei Wang<sup>b\*</sup>, Aidong Yang<sup>a\*</sup>

<sup>a</sup> *Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, United Kingdom*

<sup>b</sup> *Beijing Advanced Innovation Center for Food Nutrition and Human Health, Beijing Technology & Business University (BTBU), Beijing 100048, China*

*\*Corresponding author*

*wangbei@th.btbu.edu.cn (B. Wang)*

*aidong.yang@eng.ox.ac.uk (A. Yang)*

## Abstract

Genome-scale metabolic models (GSMMs) and flux balance analysis (FBA) have been extensively used to model and design bacterial fermentation. However, models that can simulate co-culture fermentation dynamically with quantitative accuracy are relatively scarce. To investigate metabolic interactions in yogurt starter culture of *Streptococcus thermophilus* (ST) and *Lactobacillus delbrueckii subsp. bulgaricus* (LB), this study built a dynamic community-level metabolic model based on metagenomic data. We first assessed the quality of the model by comparing the predicted growth curve with reference experimental data, and then used it to simulate the impact of different initial ST:LB ratios and search for the optimal composition of starter culture for efficient acidification for yogurt fermentation. The modeling pipeline presented in this work provides a basis for an efficient tool for computer-aided process design of the production of fermented dairy products, contributing to the development of precision fermentation in the food industry.

**Keywords:** Metagenomics; Fermentation; Flux balance analysis; Genome scale metabolic model; Proteome allocation.

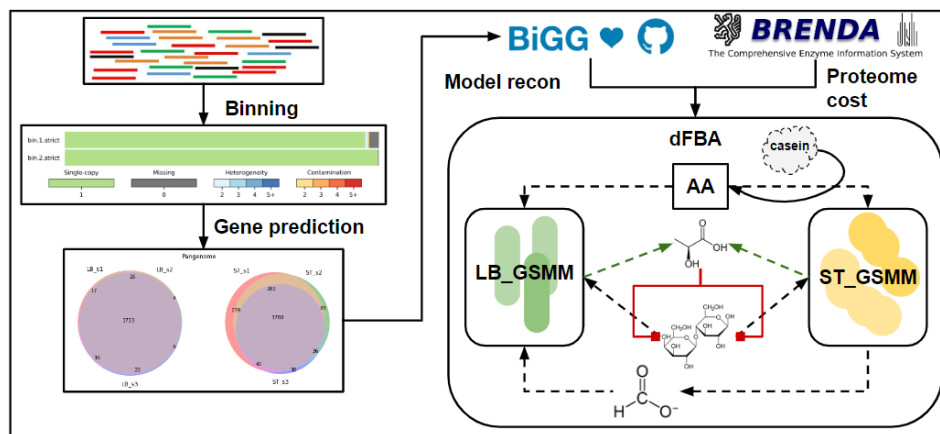
## 1. Introduction

Yogurt is an important fermented dairy food, traditionally made by lactic acid bacteria (LAB). In industrial production of yogurt, the control of the fermentation process, in terms of acidification and flavor/probiotic compound production, largely depends on the composition of the starter culture. To investigate and engineer the microbial community of yogurt starter culture, a community-level metabolic model of LABs used in the production of yogurt is needed. Currently, several genome scale metabolic models (GSMMs) for dairy-origin LABs have already been reconstructed, and a dynamic co-culture metabolic model for cheese starter culture using those GSMMs has been built (Özcan et al., 2021). However, there is still a lack of metagenome-scale metabolic models (different from GSMMs of certain strains) that can simulate the growth and metabolism of the actual (as opposed to an assumed) LAB community in yogurt fermentation.

Furthermore, existing models could not simulate unique inter-species interactions in yogurt fermentation. In this work, we built a dynamic community-level metabolic model with reconstructed GSMMs of major species identified in the starter culture, i.e., *Streptococcus thermophilus* (ST) and *Lactobacillus delbrueckii subsp. bulgaricus* (LB), using the information obtained from metagenomic analysis. We then showed how the model can be used to design the control strategy of the yogurt fermentation process.

## 2. Methodology

Building the dynamic community-level metabolic model of the yogurt starter culture comprised two steps: 1. from metagenome to annotated protein coding genes; 2. from coding genes to GSMMs (**Figure 1**). With GSMMs, dynamic flux balance analysis (dynamic FBA, or dFBA) was implemented to simulate bacterial growth and metabolism.



**Figure 1.** the workflow diagram of metabolic modeling from metagenomic analysis to GSMM reconstruction and dFBA.

### 2.1. Metagenome assembly, binning, and protein coding gene identification

Three parallel DNA samples of the commercial yogurt starter culture were sequenced by Illumina PE150. Reads were filtered for high quality ones, and then assembled by MEGAHIT.

MetaBAT2, MaxBin2 and CONCOCT were used to obtain individual genomes from the assembled metagenome. Bins were then processed using refinement and reassembly modules in metaWRAP. The final outputs are individual genomes of high quality, assessed by CheckM. GTDB-Tk performed taxonomic classification for refined individual genomes to identify major species.

Protein-coding gene prediction for each individual genome was done by Prodigal, generating the nucleotide sequence and translated protein sequence of each gene. Then, functional annotation provided information on cellular pathways, carbohydrate active enzymes and biosynthetic gene clusters.

### 2.2. Reconstruction of genome-scale metabolic models

Non-redundant protein sequences were filtered from concatenated protein sequences of different samples and used as the input for automatic GSMM reconstruction by CarveMe.

In addition to protein sequences, other inputs were chemically defined media (CDM) for gap filling and the universal bacterial template. The CDM was adopted from the co-culture metabolic model of cheese starter (Özcan et al., 2021). The universal bacterial template used for reconstruction was specialized for gram-positive bacteria.

The growth function (normalized to be the formation of 1gDW biomass) was adjusted based on biomass composition reported in previous studies and existing GSMMs of the same species. The non-growth associated ATP maintenance was set to be 0 to avoid infeasible solutions. Based on functional annotation of coding genes and existing GSMMs, reactions included erroneously were removed and missing reactions were added. In addition, to model the proteolysis activity in the co-culture system, a self-defined reaction decomposing casein protein into amino acids was added to the GSMM:



Stoichiometric coefficients in the proteolysis reaction were approximated from average amino acid distribution in the casein protein.

### 2.3. *Dynamic flux balance analysis and proteome allocation constraint*

To simulate the growth of bacteria and the production of target metabolites in time, dFBA was adopted (Mahadevan et al., 2002). The intracellular metabolic fluxes,  $v_j$ , were computed by parsimonious FBA, maximizing the growth rate while minimizing the total sum of fluxes, based on the assumption that the cell minimizes the use of enzyme catalyzed reactions due to limited recourse (Eqs. 1, 2), subject to mass conservation (Eq. 3), in which  $S_j$  represented the stoichiometric matrix. Inter-species exchanges were not modeled directly, the concentration change of metabolites and biomass in the extracellular space was modeled by differential equations to account for biomass accumulation (Eq. 4) and exchange fluxes of metabolites  $M$  from major species (Eq. 5).

$$\text{Maximize } v_{j,Growth} \quad \forall j \in \text{major species} \quad (1)$$

$$\text{Minimize } \sum_{i=0}^{\#reactions} v_{j,i} \quad (2)$$

$$S_j v_j = 0 \quad (3)$$

$$\frac{d[Biomass_j]}{dt} = v_{j,Growth} [Biomass_j] \quad (4)$$

$$\frac{d[M]}{dt} = \sum_{j=0}^{\#species} v_{j,EX\_M\_e} [Biomass_j] \quad (5)$$

Furthermore, proteome allocation was implemented to constrain reaction fluxes of central metabolism to model the flow of carbon source uptake (Zeng et al., 2020). Proteome was divided into sectors of inflexible housekeeping (Q), anabolism (A), transportation (T), catabolism (C), and the free sector. The upper bound of all flexible sectors combined was assumed to be 50% of the total proteome (Eqs. 6, 7). The proteome cost on each reaction was computed as the flux divided by the multiplicative product of enzyme activity,  $a_i$ , and saturation degree,  $\sigma_i$  (Eq. 8). The saturation degrees,  $\sigma_i$ , were assumed to be 0.5, except



for the lactose transporter, which was considered as fully saturated by abundant lactose,  $\sigma_{LT} = 1$ . The activity of the ribosome,  $a_{ribosome}$ , for the anabolism sector was adopted from literature (Regueira et al., 2020). Other enzyme activities were obtained from BRENDA Enzyme Database. For the transportation of amino acids, the flux was constrained by michaelis-menten equation ( $v_{max} \frac{[S]}{[S]+K_m}$ ), average  $v_{max}$  and michaelis-menten constant  $K_m$  were set based on parameters obtained from BRENDA and SABIO-RK.

$$\frac{P_{tot}}{1 \text{ gDW}} = p_Q (\sim 50\% \frac{P_{tot}}{1 \text{ gDW}}) + p_C + p_A + p_T + p_{Free} \quad (6)$$

$$0 < p_C + p_A + p_T \leq 50\% \frac{P_{tot}}{1 \text{ gDW}} \quad (7)$$

$$p_C = \sum \frac{v_i}{\sigma_i a_i}; p_A = \frac{v_{Growth}}{a_{ribosome}}; p_T = \frac{v_{LT}}{\sigma_{LT} a_{LT}} + \frac{v_{lacT}}{\sigma_{lacT} a_{lacT}} + \frac{v_{act}}{\sigma_{act} a_{act}} \quad (8)$$

The activity of lactose uptake incorporated inhibition by undissociated lactate (LacH), the final product of glycolysis in anaerobic conditions. The exponential decay equation (Eq. 9) to model the inhibition of lactose transporter activity was adopted from Bouguettoucha et al. (2011). The concentration of undissociated lactate was computed with Henderson-Hasselbalch equation (Eq. 10), and pH was approximated as a linear function of lactic acid concentration,  $pH = C_1[Lac] + C_2$ .

$$a_{LT} = a_{LT}^0 e^{(-k_{LacH}[LacH])} \quad (9)$$

$$[LacH] = \frac{[Lac]}{10^{pH-pK_a}} \quad (10)$$

### 3. Results

#### 3.1. Major species identified and status of reconstructed GSMMs

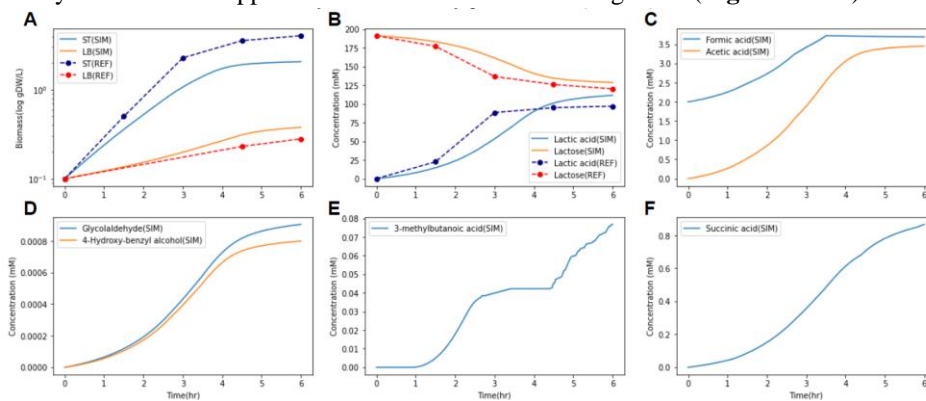
The taxonomic classification revealed that *Streptococcus thermophilus* (ST) and *Lactobacillus delbrueckii subsp. bulgaricus* (LB) were two major species in the starter culture. After binning and refinement, individual genomes of ST and LB showed good completeness and low level of contamination. There were 2499 and 1801 non-redundant protein coding genes predicted from ST and LB's genomes. In reconstructed GSMMs, reactions assigned with gene-protein-reaction relations were around 60% in both models. About 25% gap-filled reactions were added into the model for a complete metabolic network. The rest were boundary reactions for metabolite exchange.

#### 3.2. Simulation of growth and metabolism of ST/LB co-culture during yogurt fermentation

To assess the predictive power of the dynamic community-level metabolic model, dynamic simulation was performed using initial conditions of the fermentation system from reference data (Oliveira et al., 2012). For the biomass of ST and LB, the initiation

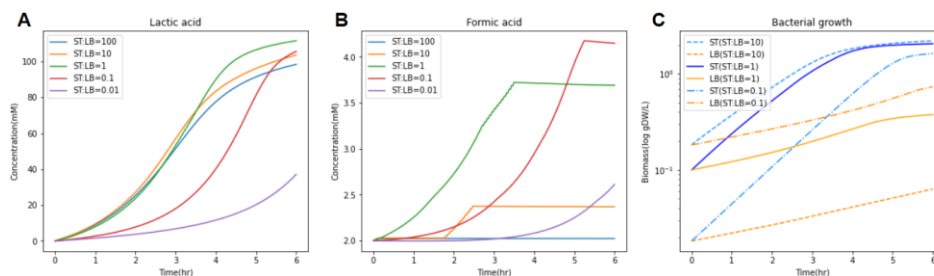
of exponential phase and transition to stationary phase were adequately captured by the model (**Figure 2A**). Besides, our model also predicted the consumption of lactose and production of lactic acid with fairly good accuracy (**Figure 2B**).

In addition to lactic acid, acetic and formic acids are also important fermentative products of glycolysis. The predicted concentration curve of acetic acid showed a similar trend with that of lactic acid (**Figure 2C**). The accumulation of formic acid produced by ST initiated the exponential growth of LB, as it activated LB's synthesis of purine, whose concentration was low in milk. In return, the activation of growth and metabolism of LB with strong proteolytic activity enhanced the growth of ST by supplementing amino acids and small peptides, which was also limited in milk (**Figure 2AC**). Apart from end products of glycolysis, the model also predicted the secretion of glycolaldehyde, 4-hydroxy-benzyl alcohol, 3-methylbutanoic acid and succinic acid, among which 3-methylbutanoic acid appeared unassociated with cellular growth (**Figure 2DEF**).



**Figure 2.** Model predictions of growth and metabolism of ST/LB co-culture during yogurt fermentation (solid line: simulation results, dashed line: reference experimental data; simulation results were obtained with a ST to LB ratio of 1, both 0.1g/L).

**3.3. Prediction of the impact of different ST:LB ratio on the fermentation behavior**  
 Perturbations on the initial ratio of ST and LB in yogurt starter culture was conducted to investigate its impact on the fermentation behavior and search for the optimal starter culture composition. The simulated concentration profile of lactic acid indicated that the acidification rate (the rate of lactic acid accumulation) by starter cultures of different ST:LB ratios ranked as follows: ST:LB = 1 > ST:LB = 10 > ST:LB = 100 > ST:LB = 0.1 > ST:LB = 0.01 (**Figure 3A**). When the acidification rate is selected as the sole objective to maximize, an optimal ST:LB ratio is predicted to be between 1 and 10. The concentration profiles of formic acid and growth curves for different ST:LB ratios showed that LB's exponential growth would be activated when formic acid accumulated to a certain level (**Figure 3BC**). The supplement of nitrogen source (amino acids and small peptides) by LB was also necessary for ST. When ST:LB increased from 1 to 10, the proteolytic activity became weaker due to the lower biomass of LB, and thus the growth of ST was inhibited, leading to reduced acidification achieved in 6 hours.



**Figure 3.** Simulation of yogurt fermentation by starter cultures of different ST:LB ratios (from 100 to 0.01, fixed total starting biomass of 0.2 g/L).

#### 4. Conclusion

Overall, this work presented the metabolic reconstruction of the classical yogurt starter culture, i.e., ST and LB, and provided a dynamic community-level GSMM modeling approach. The proposed model is able to capture the bacterial growth, the acidification of yogurt via lactic acid and predict active fluxes of flavor compounds like acetic acid, methyl butanoic acid and succinic acid during the yogurt fermentation process. Although several limitations remain to be overcome (e.g., the lack of calibration with experimentally determined biomass composition and ATP requirements), the model was shown to have the potential to offer an efficient tool to guide engineering decisions in the food industry, which could be used to address issues such as the optimal initial biomass ratio of ST and LB to maximize the rate of acidification and possibly other process targets.

#### References

- E. Özcan, M. Seven, B. Şirin, T. Çakır, E. Nikerel, B. Teusink, E. T. Öner, 2021, Dynamic co-culture metabolic models reveal the fermentation dynamics, metabolic capacities and interplays of cheese starter cultures, *Biotechnology and Bioengineering*, Vol. 118, Issue 1, Pages 223–237
- R. Mahadevan, J. S. Edwards, F. J. Doyle III, 2002, Dynamic Flux Balance Analysis of Diauxic Growth in *Escherichia coli*, *Biophysical Journal*, Vol. 83, Issue 3, Pages 1331–1340
- H. Zeng, A. Yang, 2020, Bridging substrate intake kinetics and bacterial growth phenotypes with flux balance analysis incorporating proteome allocation, *Scientific Reports*, Vol. 10, Issue 1, No. 4283
- A. Regueira, J. L. Rombouts, S. A. Wahl, M. Mauricio-Iglesias, J. M. Lema, R. Kleerebezem, 2020, Resource allocation explains lactic acid production in mixed-culture anaerobic fermentations, *Biotechnology and Bioengineering*, Vol. 118, Issue 2, Pages 745–758
- A. Bouguettoucha, B. Balanec, A. Amrane, 2011, Unstructured Models for Lactic Acid Fermentation – A Review, *Food Technology and Biotechnology*, Vol. 49, Issue 1, Pages 3–12
- R. P. S. Oliveira, B. R. Torres, P. Perego, M. N. Oliveira, A. Converti, 2012, Co-metabolic models of *Streptococcus thermophilus* in co-culture with *Lactobacillus bulgaricus* or *Lactobacillus acidophilus*, *Biochemical Engineering Journal*, Vol. 62, Pages 62–69

# Multilevel Discrete Modelling of Microbial Cell Cultures using Population Balances

Menwer Attarakih<sup>a</sup>, Hans-Jörg Bart<sup>b</sup>, Khaled Rawajfeh<sup>a</sup>

<sup>a</sup>The University of Jordan, Queen Rania Str., Amman, 11942, Jordan

<sup>b</sup>The University of Kaiserslautern, Gottlieb-Daimler Str., 67663, Germany  
[m.attarakih@ju.edu.jo](mailto:m.attarakih@ju.edu.jo)

## Abstract

Microbial cell population balance models are used because of culture heterogeneity due to complex intercellular phenomena governed by cell growth, substrate utilization, and product formation. To solve these models, an efficient multilevel system approach based on conservation of cell distribution moments is introduced. This conservation is propagated along the characteristic curves of these hyperbolic conservation laws. This is assured by the careful design of numerical discrete schemes that pass the conserved discrete information from the initial state to its final extracellular state. The discrete model on each layer is derived with a guarantee to satisfy the conservation of cell total number and the mean cell biomass. Many reference solutions based on the Chebyshev-QMOM, the maximum entropy method, and the Lax-Wendroff finite difference scheme are used to validate the developed model. When compared to these methods, the model exhibits extreme accuracy and rapid convergence at a low CPU time cost.

**Keywords:** Microbial cells; Population Balances; Multilevel; Discrete modelling.

## 1. Introduction

The continuous population balance equation (PBE) with particle nucleation, growth, and breakage is a nonlinear integral-partial differential equation. For microbial cell cultures, the PBE presents a framework to describe the complex phenomena manifested by cell growth, division, substrate utilization, and product formation. With no general analytical solution, this calls for numerical modeling to understand the cell culture behavior and to predict, optimize, and control cell growth in biochemical reactors (Mantzaris et al., 2001; Ramkrishna et al., 2014; Jerono et al., 2021).

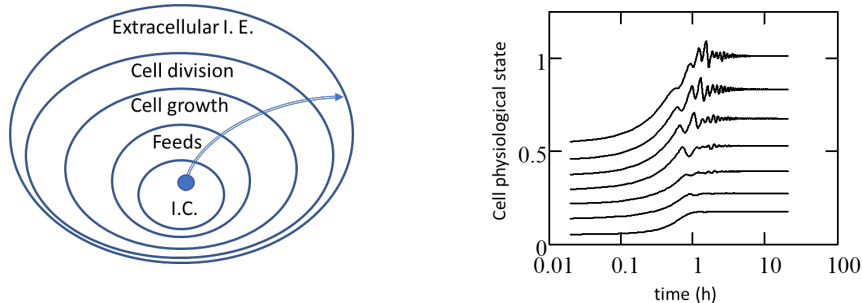


Figure (1): Left-Multilevel system discretization approach for microbial cell dynamics. Right-system characteristic curves as predicted by ChQMOM using 7 quadrature nodes.

Since the heterogeneity of microbial cultures results from cell size and its intercellular concentrations of DNA, we confine ourselves to the mono-variate cell probability density

function  $f(x, t)$  with cell mass ( $x$ ) as its internal coordinate. These cells are embedded in a continuous extracellular environment with a single substrate ( $S$ ) (Durr & Buk, 2020; Jerono et al., 2021). Our numerical modeling method for this system starts by introducing a multilevel discrete approach that is evolved in five evolutionary layers, as shown in Fig. 1. This layer structure reflects the hyperbolic nature of the discrete conservation and population balance laws. The first layer in Fig. (1-Left) shows the initial conditions and the external feeds, which are considered an essential part of the internal inconsistency problem arising from the transformation of the continuous PBE into its discrete form. This is to prevent the propagation of internal inconsistency along the characteristic lines (Fig. 1-Right) associated with the growth, division, and extracellular information exchange layers, which if ignored would increase the computational cost. To explain this, given a discrete initial state vectors with cell boundaries  $x$  and centers  $\chi$  which satisfy  $x_i < \chi_i < x_{i+1}$  and initial cell population number concentration  $N_i(0)$ ,  $i = 1, 2, \dots, M$ . In this discrete framework, the integral properties such as total cell number and mass will never be the same as those calculated from its continuous probability density function except for only one integral property (Ramkrishna, 2000). For example,  $N_i(0) = \int f(x,0)dx$  but  $\sum x_i N_i(0) + \varepsilon = \int x f(x,0)dx$  where  $\varepsilon$  is called the error of moving the state from its continuous to its discrete states. Unfortunately, this error is of cumulative nature and is propagated from the initial state layer to the higher layers with increasing amplitude. Its elimination needs infinitely large number of discrete states of uniform grid:  $x_i = ix_{\min}$ . However, this error can be eliminated for two selected integral properties starting from the initial state layers and moving forward to the growth and division layers. As the total number and mass are the two most important cell transport properties, we derive our discrete PBEs that satisfies these properties in each layer within this multilevel approach as shown in Fig.(1-Left).

## 2. Mathematical model

The heterogeneity of the cell population is adequately modelled using the population balance equation (Ramkrishna, 2014). In this equation, the cell density function  $f(x,t)$  accounts for variances in the cell population in a homogeneous physical space (CSTR):

$$\frac{\partial f(x, t)}{\partial t} + \frac{\partial [G(x, t)f(x, t)]}{\partial x} = -D(t)f(x, t) + R\{f(x, t), S\} \quad (1)$$

where  $x$  is the physiological state of single cell (biomass components) which depends on the intercellular growth rate  $G(x,S)$ , where  $S$  is the single substrate concentration in the extracellular continuous medium, and  $D$  is the dilution rate.  $R$  is a volumetric source which represents the net number of cells per unit biomass volume due to birth of cells with the physiological state  $x$  because of binary division of all mother cells of size  $y$ :

$$R\{f(x, t), S\} = -\Gamma(x, S)f(x, t) + 2 \int_x^{x_{\max}} \Gamma(y, S) \beta(x, y) f(y, t) dy \quad (2)$$

In this equation  $\Gamma(x,S)$  is the cell division rate and  $\beta(x,y)$  is the cell partition probability function for a daughter cell of physiological state  $x$  given a mother cell of state  $y$ . This function should satisfy the mass and number conservation conditions as well as the symmetrical properties (Ramkrishna, 2000). The cell PBEs (1) and (2) are coupled with the dynamic abiotic environment through the substrate mass balance as shown by Eq.(3). In this equation the source term represents the net flow of substrate through reactor boundary along with the total loss of substrate as utilized by the microbial cells.

$$\frac{dS}{dt} = D(S_f - S) - \int_{x_{\min}}^{x_{\max}} q(x, S) f(x, t) dx \quad (3)$$

$$q(x, S) = \frac{1}{Y} \left( \frac{\mu_{\max} S}{K_s + S} \right) x \quad (4)$$

In these equations  $q(x, S)$  is single cell consumption rate of a substrate ( $S$ ),  $Y$  is a constant biomass product yield,  $\mu_{\max}$  is the maximum growth rate and  $K_s$  is the half saturation substrate concentration. The growth rate is given by  $G(x, S) = Yq(x, S)$ , and  $\Gamma(x, S) = \gamma(x)G(x, S)$  where  $\gamma(x)$  is expressed in terms of Gaussian probability density function (Mantzaris et al., 2001). The initial conditions are  $f(x, 0) = f_0(x)$  and  $S(0) = S_0$  while  $S_f$  is the substrate feed concentration and the boundary condition is  $f(x_{\min}, t) = f(x_{\max}, t) = 0$ .

### 3. Multilevel model discretization

A multilevel model discretization approach that is consistent with that shown in Fig.(1-Left) is used here to ensure the consistency of the discrete model when compared to the continuous one as given by Eqs.(1-4).

#### 3.1. Level 1: Initial conditions and feed discretization

The cell internal property is discretized into  $M$  contiguous numerical cells with geometric progression  $x_i = x_{\min}[r]^{i/M}$ , with  $r > 1$  and cell centers  $\chi_i = (x_i + x_{i+1})/2$ ,  $i = 0, 1, \dots, M$ .

$$N_i^0 = \alpha_0 N_i(0) + \beta_0 N_{i+1}(0), i = 0, 1, \dots \quad (5)$$

where  $\alpha_0$  and  $\beta_0$  are solutions of two linear equations to conserve the zero and first moments of the cell number concentration functions  $f_0(x, 0)$  and  $f_1(x, 0)$  in which 0 refers to the initial condition and 1 refers to the inlet feed distribution. This is a crucial step in this multilevel discrete framework to avoid the propagation of any discrete error along the characteristic lines of the system of Eqs. (1-4).

#### 3.2. Level 2: Cell growth

$$\frac{dN_i(t)}{dt} = \frac{2}{x_i} \frac{1}{r^2 - 1} (G(\chi_i, S)N_i(t) - rG(\chi_{i-1}, S)N_{i-1}(t)), i > 0 \quad (6)$$

At this level, the discrete cell growth needs to guarantee the conservation of total number and mass concentrations using a new first order upwind scheme as given by Eq.(6). This conservation law can handle smooth and discontinuous numerical fluxes with suitable flux-vector splitting depending on the sign of  $G(\chi, S)$  where  $r = [x_{\max}/x_{\min}]^{(1/M)}$ .

#### 3.3. Level 3: Cell division

$$\frac{dN_i(t)}{dt} = -\eta_i \Gamma(\chi_i, S)N_i(t) + \sum_{j=0}^M \sigma_j \Gamma(\chi_j, S)\pi_{i,j}N_j(t), i = 0, 1, \dots, M \quad (7)$$

This equation is the discrete counterpart of Eqs.(1) and (2) without growth term where the loss and formation terms in Eq.(7) are modified by placing the two discrete functions ( $\eta$  and  $\sigma$ ) to conserve exactly the zero and first moments of the cell population. This ensures passing the conserved information to higher levels without errors induced by the inconsistency problem. The two functions  $\sigma$  and  $\eta$  tends to unity as the grid is refined showing the consistency of these functions with the mean-value theorem of integrals over

$\beta(x,y)$ . The discrete function  $\pi_{i,j}$  calculates the total  $i^{\text{th}}$  fraction of daughter cells at any given mother of size  $y_j$ .

### 3.4. Level 4: Extracellular dynamics

$$\frac{dS(t)}{dt} = D(t)(S_f - S) - \sum_{j=0}^M q(\chi_j, S) N_j(t), i = 0, 1, \dots, M \quad (8)$$

This level provides a gate to pass the conserved discrete information from level 4 to the cell continuous environment and then returns any information from the continuous phase to the discrete one. The final discrete solution is then recovered at any time step as  $n(\chi_i, t) = N_i(t)/\Delta x_i, \forall i, i = 0, M$ .

## 4. Results and discussion

The multilevel discrete model that is given by Eqs.(5-8) provides a complete set of ODEs supplemented by consecutive equations for cell division rate, cell partition probability function, and cell intercellular growth rate.

Table (1): Parameters for test problem 2 as reported by Mantzaris et al. (2001).

$\mu_{\max}$	$K_s$	$Y$	$D$	$S_f$	$S_0$	Gaussian I.C.	
1.0/h	0.2 g/L	0.5 g/g	0.5/h	10 g/L	2 g/L	$\mu_0$	$\sigma_0$
Gaussian division probability function				$\mu_1 = 0.575$	$\sigma_1 = 0.125$	0.2875	0.0675

The initial condition and the cell division probability functions are taken as Gaussian densities with mean and standard deviations as shown in Table (1). The cell partition probability function is taken as a beta distribution:  $\beta(x, y) = c(q)y^{-1}(x/y)^{q-1}(1 - x/y)^{q-1}$  where  $c(q)$  is normalization factor and  $q = 40$  (Mantzaris et al., 2001).

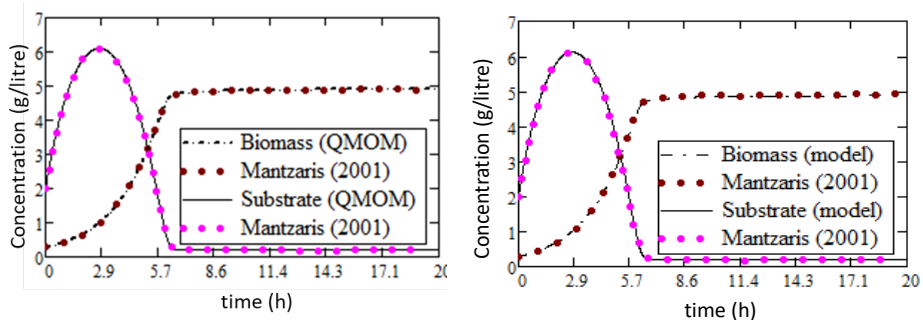


Figure (2): Dynamics of cell growth and division in a CSTR using Chebyshev-QMOM, the Lax–Wendroff scheme (Mantzaris et al., 2001, Toro, 1999) and the present multilevel discrete model.

### 4.1. Reference solutions

To validate the multilevel discrete model in terms of the conserved moments; namely, the zero (total cell number) and first moments (total cell mass) of the discrete cell distribution, two standard methods are used: The first one is the Chebyshev-QMOM (Upadhyay, 2012) and the conservative finite difference scheme FDS (Mantzaris et al., 2001). The Chebyshev-QMOM (ChQMOM) is superior to the standard QMOM which is based on the Product Difference Algorithm (PDA), and the Direct Quadrature Method Of Moments

(DQMOM) where ill-conditioned linear system of equations is encountered when the number of the quadrature nodes increases.

4.1.1. Chebyshev-QMOM reference solution

The system of Eqs.(1-5) can be written in compact form in terms of the moments of  $f(x,t)$  where the unclosed integrals are approximated using the Chebyshev-QMOM. The cell internal property integration domain is  $\Omega = [x_{min}, x_{max}] = [0, 1]$  where the resulting system of ODEs was solved using the explicit Adams-Bashforth method with error tolerance =  $10^{-3}$  and number of quadrature nodes  $N_q$  equals to seven.

4.1.2. Mantzaris et al. (2002) reference solution

The second reference solution for Eqs.(1-5) is regenerated from the work of Mantzaris et al. (2001) using conservative discretization finite difference schemes for a system of hyperbolic conservation laws. The results from the Lax-Wendroff scheme are generated with the same numerical parameters as shown in Table (1). The results are displayed in Fig.(2) where the number of numerical cells used by this scheme is 500.

4.1.3. Comparison of the present model predictions with reference solutions

Firstly, we compared the substrate and biomass dynamics as predicted by the ChQMOM and the reference solution of Mantzaris et al. (2001). The results are found indistinguishable using seven quadrature nodes which is obvious from the results shown in the left panel of Fig.(2). This result is not surprising since the ChQMOM is well-known for its high accuracy in predicting distribution moments (Attarakih et al., 2012, 2014). Secondly, and as gaining trust in the solution of the present model, the multilevel discrete model (Eqs.(5-8)) is solved in time using the explicit Adams-Bashforth method with error tolerance =  $10^{-3}$  and number of grid points  $M=50$ . To ensure the correct propagation of information along the hyperbolic laws characteristic curves (Fig.(2-Right)), the Gaussian I.C. with parameters given in Table (1), is discretized according to the scheme shown in Eq.(5). This solution conserves exactly the total number and mass concentrations.

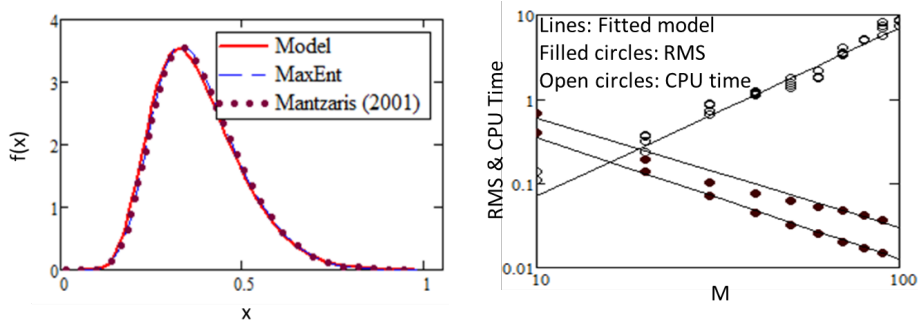


Figure (3): Left-Converged cell distribution. Right- CPU time (s) denoted by open circles and RMS (Root Mean Square) error in zero and 1<sup>st</sup> moment of cell distribution (Filled circles).

Starting from the first layer (see Fig.(1-Left)) and for  $M = 50$ , the factors  $\alpha_0$  and  $\beta_0$  are found to be 0.99 and 0.01 respectively. This means 99.0 percent of the discrete population is assigned to the  $i^{\text{th}}$  numerical cell center while only 1.0 percent is assigned to the next larger numerical cell center. This ensures the correct prediction of the population moments even if low number of numerical cells is used (e.g. 15). Based on this, the cell growth layer conserves exactly zero and first moments of the cell distribution which are then propagated correctly to the cell division layer in spite of the numerical diffusion associated with the used first-order upwind schemes (Toro, 1999). In the cell division layer, and as implied by Eq.(7), the computation of the two correction functions ( $\eta, \sigma$ )



are computed offline and are found to depend on the cell cell partition distribution function ( $\beta(x,y)$ ) and the numerical grid structure. Once these correction vectors are available, the cell division is advanced in time and is then propagated to the extracellular dynamic layer without loss of any global information because of discretization. As a result of this system conservation approach, the biomass and substrate utilization are correctly predicted by our model as compared to the two reference solutions mentioned in sections (4.1.1 and 2). This is clear by referring to the right panel of Fig.(2). In a second level of validation, our model is compared with the FDS of Mantzaris et al (2001) and the Maximum Entropy (MaxEnt) method of Attarakih and Bart (2014) to predict the cell population distribution. Our present model converges rapidly to the FDS and the MaxEntM as the number of numerical cells increases. In Fig.(3)-Left, we used  $M = 50$  as compared to the FDS solution using 500 numerical cells. On the other hand, five moments are used in the MaxEnt method to reconstruct the distribution at final time of 20 hrs. To quantify the convergence characteristics of the present model, the RMS error is calculated using the zero and first moments of the distribution as function of  $M$ . Fig.(3)-Right shows rapid decrease in the RMS as function of  $M$  with approximate order of  $O(M^{-0.4})$ . Concerning the computational cost of the present discrete model, and as can be seen from Fig.(3)-Right, the mean CPU time (four runs are used per  $M$  value) is increasing as  $O(M^2)$ .

## 5. Summary and conclusions

To conclude, the present multilevel discrete model is yet simple, with equations that are elegant in structure, efficient in implementation, and that handle the cell culture dynamics with high accuracy. These equations conserve the propagated information along the characteristic lines of the hyperbolic conservation laws through a hierarchical structure. When compared to the QMOM and the FDS, the present model converges rapidly as the number of numerical cells is increased. With more than four quadrature nodes, the present model computational cost is lower than that of the FDS or the ChQMOM.

## Acknowledgments

This work was supported by the German Science Foundation (DFG).

## References

- M. Attarakih and H.-J. Bart, 2014, Solution of the population balance equation using the Differential Maximum Entropy Method (DMaxEntM): An application to liquid extraction columns. *Chemical Engineering Science*, 108, 123-133.
- M. Attarakih, H. B. Jildeh, M. Mickler and H.-J. Bart, 2012, The OPOSPM as a nonlinear autocorrelation population balance model for dynamic simulation of liquid extraction columns, *Computer Aided Chemical Engineering*, 31, 1216-1220.
- R. Durr and A. Buck, A., 2020, Approximate Moment Methods for Population Balance Equations in Particulate and Bioengineering Processes. *Processes*, 8, 1-18.
- P. Jerono, A. Schaum & T. Meurer, 2021, Moment-based Kalman filter design for cell population balance models in batch fermentation processes. *IFAC-PapersOnLine*, 54, 19-24.
- N. V. Mantzaris, P. Daoutidis, & F. Sieni, 2001, Numerical solution of multi-variable cell population balance models: I. Finite difference methods. *Comp. Chem. Eng.*, 25, 1411-1440.
- D. Ramkrishna and M. R. Singh, 2014, Population balance modeling: Current status and future Prospects. *Annual Review of Chemical and Biomolecular Engineering*, 5, 123-146.
- D. Ramkrishna, 2000, *Population balances: Theory and applications to particulate systems in engineering*, Academic Press, San Diego.
- E. F. Toro, 1999, *Riemann solvers and numerical methods for fluid dynamics*, Springer, Berlin.
- R. R. Upadhyay, 2012, Evaluation of the use of the Chebyshev algorithm with the quadrature method of moments for simulating aerosol dynamics. *Journal of Aerosol Science*, 44, 11-23.

# Binary Classification of the Endocrine Disrupting Chemicals by Artificial Neural Networks

Zahir Aghayev,<sup>a,b</sup> George F. Walker,<sup>a,b</sup> Funda Iseri,<sup>c,d</sup> Moustafa Ali,<sup>c,d</sup> Adam T. Szafran,<sup>e</sup> Fabio Stossi,<sup>e,f</sup> Michael A. Mancini,<sup>e,f</sup> Efstratios N. Pistikopoulos,<sup>c,d</sup> Burcu Beykal<sup>a,b</sup>

<sup>a</sup>*Department of Chemical and Biomolecular Engineering, University of Connecticut, Storrs, CT 06269, USA*

<sup>b</sup>*Center for Clean Energy Engineering, University of Connecticut, Storrs, CT 06269, USA*

<sup>c</sup>*Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77843, USA*

<sup>d</sup>*Texas A&M Energy Institute, Texas A&M University, College Station, TX 77843, USA*

<sup>e</sup>*Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX 77030, USA*

<sup>f</sup>*GCC Center for Advanced Microscopy and Image Informatics, Baylor College of Medicine, Houston, TX 77030, USA*

*burcu.beykal@uconn.edu*

## Abstract

We develop a machine learning framework that integrates high content/high throughput image analysis and artificial neural networks (ANNs) to model the separation between chemical compounds based on their estrogenic receptor activity. Natural and man-made chemicals have the potential to disrupt the endocrine system by interfering with hormone actions in people and wildlife. Although numerous studies have revealed new knowledge on the mechanism through which these compounds interfere with various hormone receptors, it is still a very challenging task to comprehensively evaluate the endocrine disrupting potential of all existing chemicals and their mixtures by pure *in vitro* or *in vivo* approaches. Machine learning offers a unique advantage in the rapid evaluation of chemical toxicity through learning the underlying patterns in the experimental biological activity data. Motivated by this, we train and test ANN classifiers for modeling the activity of estrogen receptor- $\alpha$  agonists and antagonists at the single-cell level by using high throughput/high content microscopy descriptors. Our framework preprocesses the experimental data by cleaning, scaling, and feature engineering where only the middle 50% of the values from each sample with detectable receptor-DNA binding is considered in the dataset. Principal component analysis is also used to minimize the effects of experimental noise in modeling where these projected features are used in classification model building. The results show that our ANN-based nonlinear data-driven framework classifies the benchmark agonist and antagonist chemicals with 98.41% accuracy.

**Keywords:** machine learning, artificial neural networks, estrogenic potential, high throughput analysis, nonlinear classification.

## 1. Introduction

There is a wide array of naturally occurring and synthetic chemical substances that interfere with the endocrine system and cause detrimental developmental, reproductive, neurological, and immunological consequences in humans which are referred to as

endocrine disrupting chemicals (EDCs). The interaction of these substances with the estrogen receptor (ER) ligand binding domain induces the activation or blockage of the biological response, known as an agonist or antagonist chemical, respectively. Thus, rapid and accurate endocrine disrupting potential identification is a crucial task in environmental risk assessment.

By observing the mechanistic endpoints in the ER signaling pathway, it is possible to assess the estrogenic potential of various substances using cell-based or cell-free assays. While these approaches can be used to accurately assess a small number of substances, evaluating tens of thousands of chemicals and their mixtures becomes quickly infeasible. On the other hand, the availability of larger data sets from high throughput experimentation, increased computing capabilities, and cutting-edge machine learning models such as deep neural networks that can be trained to capture highly nonlinear behaviors in datasets make it possible to mathematically model and predict the expected biological activity of any chemical compound with prompt function evaluations.

Previous studies by Chierici et al. (2019) and Collins and Barton-Maclaren (2022) investigated a variety of machine learning models, including Support Vector Machines (SVMs), Random Forest (RF), ANNs, for predicting the ER potential of chemicals in binary and/or multi-class classification settings. However, the predictive performance of these studies has suffered from relatively low predictive balanced accuracies in certain blind testing sets with wide performance margins ranging from 59% to 93%.

In our earlier studies, we also presented a framework to categorize chemicals according to their ER activities using population-averaged high content imaging data. By analyzing different subsets of feature groups using correlation analysis, we formulated linear (Logistic Regression) and nonlinear classification models (RF) that achieved an accuracy of more than 90% on the benchmark chemicals (Mukherjee et al., 2020). Subsequently, we integrated RF and SVM algorithms with the single-cell ER data, where characteristic shape features of the non-central gamma distribution that were used to fit the cell populations (alpha, beta, and gamma) were used as descriptors in our predictive models (Ganesh et al., 2021). Most recently, we demonstrated a predictive modeling pipeline for the endocrine disrupting potential identification that directly uses the single-cell level information, along with principal component analysis, and SVM and RF classifiers (Aghayev et al., 2022). We were able to show that the SVM model classifies unseen testing chemicals with more than 96% accuracy and data preprocessing plays a key role in achieving favorable performance from the trained classifiers.

Different than our previous studies, here, we are proposing an ANN-based predictive modeling framework that will further improve the classification accuracy and sensitivity of the separation between agonist and antagonist chemicals using single-cell level data. By further incorporating a feature selection routine based on the Fowlkes-Mallows (FM) Index in the analysis pipeline, we analyze smaller feature subsets of the high-dimensional experimental data and quantify the predictive performance of the final models using standard evaluation metrics.

## 2. Methods and Materials

### 2.1. Experimental Methodology

Sixty reference EDCs with known ER $\alpha$  activities are analyzed with an engineered biosensor to collect the image-based high content data. The studied compound library contains 32 agonists with different activity levels (very weak, weak, moderate, strong), 15 antagonists, and 13 inactive chemicals. To record the estrogenic activity of these

reference chemicals, biosensor cells were treated with varying concentrations of the chemicals for two hours, and high-throughput microscopy and image analysis pipelines detailed in Ashcroft et al. (2011) are followed. This analysis yielded 70 size, shape, and intensity attributes that relate to the various mechanistic endpoints on the ER signaling pathway with measurements recorded for every single treated cell (Szafran et al., 2017). Further details of the experimental data generation, processing, and quantification are provided in Aghayev et al., 2022.

## 2.2. Computational Methodology

### 2.2.1. Data Preprocessing

Once the multidimensional imaging data were generated and standardized using robust z-score, we conducted percentile analysis over cells that have visible nuclear spots (receptor-DNA binding, array positive cells). The goal of this step is to expand the dataset with engineering features that will potentially enable a better separation than the raw descriptors. We previously explored different percentile ranges and their effects on nonlinear classification algorithms (Aghayev et al., 2022). Based on our findings, only single-cell features that fall into the middle percentile (between 25<sup>th</sup> and 75<sup>th</sup> percentiles) were preserved and the dataset size was decreased from 165,525 to 22,286 observations.

The produced dataset was then put through missing data check and rigorous cleaning steps, where cell data representing the inactive compounds and control treatments were eliminated from the analysis. As a result, we end up with a preprocessed dataset size of 9420 observations  $\times$  36 features prior to the feature selection phase.

### 2.2.2. Feature Selection via Clustering Analysis

We used clustering analysis to identify promising subsets of the original features that will facilitate the separation of agonist and antagonist compounds. To this end, we created 11 different feature groups that capture different compartments and characteristics of the measured cells (e.g., cytoplasm, nucleus, nucleoplasm, shape, etc.), including the full set of features. We then performed hierarchical clustering on the mean cell response of each active compound and selected feature subsets using Euclidean distance and complete linkage. Each clustering dendrogram is then compared to a synthetic “ideal” cluster, where this cluster shows perfect separation between agonistic and antagonistic compounds. The cluster similarity of the subset features and the ideal dendrogram is quantified using the FM index (Onel et al., 2019). FM index evaluates the similarity between two dendrograms using the geometric mean of precision and recall (Eq. 1).

$$FM = \sqrt{\frac{TP}{TP + FP} \cdot \frac{TP}{TP + FN}} \quad (1)$$

A higher FM index value (1-perfect match, 0-no match) indicates greater similarity between the analyzed dendrograms which in this case shows that the studied set of features will either yield a similar clustering to the ideal case or not. This way we can make deductions about whether a feature subset can yield a perfect classifier and have superior performance compared to the full feature set. Additionally, we tested the correlation between the dissimilarity matrices of the clusters using the Mantel test across 10,000 permutations and computed Pearson’s product-moment correlation coefficient to assess whether any observed relationships were due to random arrangement.

### 2.2.3. Principal Component Analysis (PCA)

After identifying promising feature subsets, these datasets were standardized to conduct PCA. As a result, the selected subsets of feature groups were projected to a new feature

space - new orthogonal features that are the linear combinations of the original features – principal components (PCs) with descending explained variance. Fig 1 demonstrates the scree plot of % explained variance and the top 10 PCs for the dataset with all features. Based on the location of the elbow of the scree plot, the first 3 PCs are enough to retain around 91.07% of the variation in the original datasets.

Following the feature selection and dimensionality reduction steps, the transformed datasets are split for the training and testing of the ANNs. To train the model effectively, we considered a compound-wise split procedure by considering their ER activities. During this process, we kept the testing set unseen to assess the predicted performance of the model without any bias.

#### 2.2.4. The Deep Neural Network Model

The proposed neural network model is a feed-forward neural network with 3 nodes in the input layer and 2 nodes in the output layer. Every single partition of PCs is propagated through 3 fully connected hidden layers with 8, 16, and 20 nodes, respectively. For all the layers Rectified Linear Unit (ReLU) activation functions are used (Eq. 2).

$$g(z) = \begin{cases} 0 & \text{if } z < 0 \\ z & \text{otherwise} \end{cases} \quad (2)$$

To estimate the network parameters, we use the Adaptive Moment Estimation (Adam) optimizer which inherits characteristics of both the Adagrad (Adaptive Gradient Descent) and RMSprop (Root Mean Square Propagation) algorithms (Kingma and Ba, 2014). Regarding the loss function selection, as we are trying to categorize the compounds according to their agonistic and antagonistic behaviors (binary classification), we used binary cross-entropy. Finally, after training the model, its predictive capability is measured using accuracy, sensitivity, specificity, and balanced accuracy criteria by classifying the testing set that was reserved unseen to the model.

### 3. Results and Discussion

The results obtained from the feature selection phase are provided in Table 1. After thorough statistical analysis of the listed 11 feature groups, we found that the FM index value for 4 of them (highlighted with grey shade) is equal to 0.89, which means that they have high similarity to the ideal clustering of agonist and antagonist compounds. These 4 feature subsets, including all features, also have highly correlated distances matrices with respect to the ideal cluster and their p-values are below 0.05. This indicates that the observed correlations are not by random arrangement and the results are statistically significant. The other feature subsets are observed to have lower FM indices, and in one case p-value greater than 0.05, indicating that their dendrograms are more different than the ideal case with less distinct separation of agonist and antagonist compounds into two clusters. This shows that the underlying dataset might not yield a good classification model as the information carried on these features are not descriptive enough to have a

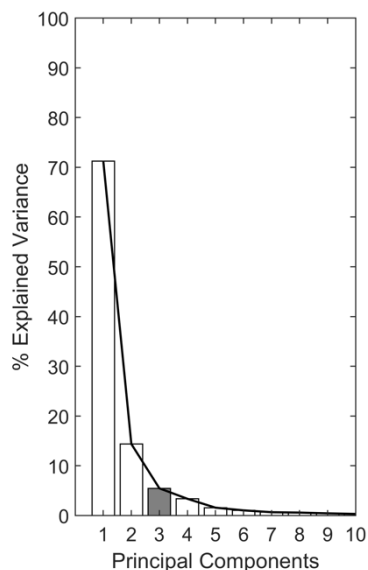


Fig 1. Scree plot of % explained variance and the top 10 PCs for the dataset with all features

generalized classifier. Hence, we only take the most promising feature subsets with the highest FM index to perform classification analysis. Although we have reduced dimensionality through this procedure, we still use PCA to decorrelate the experimental features and use the PCs to train the neural network. The training and testing results of the deep neural network are summarized in Table 2.

Table 1 The FM indices, Pearson's product-moment correlation coefficient ( $r$ ), and the p-value resulting from the Mantel test for the studied feature subsets.

Feature Subset	FM Index	Mantel Test $r$	p-value
Array	0.80	0.93	0.0001
-Array	0.89	0.94	0.0001
Cell	0.89	0.92	0.0001
Cytoplasm	0.57	0.37	0.0001
Nucleus	0.89	0.94	0.0001
Nucleoplasm	0.58	0.93	0.0001
Intensity	0.74	1.00	0.0001
-Intensity	0.64	0.23	0.0104
Secondary	0.53	0.16	0.0534
Shape	0.59	0.31	0.0001
All	0.89	1.00	0.0001

According to the classification results, the dataset with all features outperforms the other 3 feature groups with 98.41% accuracy. This dataset has the highest accuracy, sensitivity, and balanced accuracy values, as well as a very high specificity value. However, datasets with only nucleus features and without array features have nearly perfect specificity values, which represent the fraction of correctly predicted antagonist compounds. Interestingly, although the dataset with only nucleus features mapped nearly all the antagonist compounds correctly (only 1 cell misclassified out of 1791 cells), its capability is very limited for agonist compounds (67.55%). We observe the opposite trend for the dataset with only cell features, where a deep neural network constructed with the projected set of these features yields a classifier with 97.54% sensitivity and 67.67% specificity in the testing set. Overall, the results show that the neural network classifier constructed from the dataset with all features outperforms other models and provides consistently very high performance scores in both training and testing steps.

Table 2 Performance metrics of deep neural network model for the 4 selected feature groups.

Feature Group	Set	Accuracy	Sensitivity	Specificity	Balanced Accuracy
W/o Array Features	Testing	92.42%	84.28%	<b>100.00%</b>	92.14%
	Training	98.64%	99.05%	98.27%	98.66%
Cell Features	Testing	82.07%	97.54%	67.67%	82.61%
	Training	94.35%	99.40%	89.77%	94.59%
Nucleus Features	Testing	84.33%	67.55%	99.94%	83.75%
	Training	98.09%	96.01%	99.97%	97.99%
All Features	Testing	<b>98.41%</b>	<b>97.66%</b>	99.11%	<b>98.39%</b>
	Training	98.96%	100.00%	98.02%	99.01%

#### 4. Conclusion

We provided a classification framework based on multidimensional imaging data and deep neural networks to assess the endocrine disrupting potential of estrogen receptor- $\alpha$  (ER $\alpha$ ) agonists and antagonists. To overcome the curse of dimensionality, both feature selection and dimensionality reduction techniques were incorporated. Feature selection is performed using hierarchical clustering, FM index, and Mantel test techniques, where the similarity between different feature subgroups was quantified with the ideal grouping of agonist and antagonist compounds. Four feature subsets, including the entire set of features, were identified as the most promising ones that will enable the agonist/antagonist separation. Principal component analysis is also performed to further reduce the dimensionality of the studied datasets and to remove any correlations between the features. The results show that the dataset with all features succeeded in predicting the estrogenic activity of unknown chemicals with 98.41% accuracy, which was followed by the dataset without array features (92.42%). Additionally, we observed an interesting trend in the sensitivity and specificity values of the other two datasets; while the dataset with only cell features could predict the ER activity of agonist compounds with very high accuracy (97.54%), the dataset with only nucleus features classified 1790 out of 1791 cells correctly (99.94%) for antagonist class. Overall, we demonstrated that highly accurate neural network models can be constructed to predict the biological activity of chemical compounds. This research was funded by the U.S. National Institutes of Health (NIH) grant P42 ES027704.

#### References

- Z. Aghayev, A.T. Szafran, A. Tran, H.S. Ganesh, F. Stossi, L. Zhou, M.A. Mancini, E.N. Pistikopoulos, B. Beykal, 2022, Machine learning methods for endocrine disrupting potential identification based on single-cell data, *Chemical Engineering Science* (Under Review).
- F.J. Ashcroft, J.Y. Newberg, E.D. Jones, I. Mikic, M.A. Mancini, 2011, High content imaging-based assay to classify estrogen receptor- $\alpha$  ligands based on defined mechanistic outcomes, *Gene*, 477(1-2), 42-52.
- M. Chierici, M. Giuliani, N. Bussola, G. Jurman, C. Furlanello, 2018, Machine learning models for predicting endocrine disruption potential of environmental chemicals, *Journal of Environmental Science and Health, Part C*, 36(4), 237- 251.
- S.P. Collins, T.S. Barton-Maclaren, 2022, Novel machine learning models to predict endocrine disruption activity for high-throughput chemical screening. *Frontiers in Toxicology*, 4:981928.
- H.S. Ganesh, B. Beykal, A.T. Szafran, F. Stossi, L. Zhou, M.A. Mancini, E.N. Pistikopoulos, 2021, Predicting the estrogen receptor activity of environmental chemicals by single-cell image analysis and data-driven modeling, *Computer Aided Chemical Engineering*, 50, 481-486.
- D.P. Kingma, J. Ba, 2014, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980.
- R. Mukherjee, B. Beykal, A.T. Szafran, M. Onel, F. Stossi, M.G. Mancini, D. Llyod, F.A. Wright, L. Zhou, M.A. Mancini, E.N. Pistikopoulos, 2020, Classification of estrogenic compounds by coupling high content analysis and machine learning algorithms, *PLoS Computational Biology*, 16(9), e1008191.
- M. Onel, B. Beykal, K. Ferguson, W.A. Chiu, T.J. McDonald, L. Zhou, J.S. House, F.A. Wright, D.A. Sheen, I. Rusyn, E.N. Pistikopoulos, 2019, Grouping of complex substances using analytical chemistry data: A framework for quantitative evaluation and visualization, *PloS one*, 14(10), e0223517.
- A.T. Szafran, F. Stossi, M.G. Mancini, C.L. Walker, M.A. Mancini, 2017, Characterizing properties of non-estrogenic substituted bisphenol analogs using high throughput microscopy and image analysis, *PloS one*, 12(7), e0180141.

# Digital Twins in Pilot Scale Fermentation: Non-Linear State Estimation for Improving Induction Timing

Mads Stevnsborg<sup>a</sup>, Kurt Selle<sup>b</sup>, Ryan Barton<sup>b</sup>, Oscar A. Prado-Rubio<sup>a,c</sup>, Carina Gargalo<sup>a</sup>, Krist V. Gernaey<sup>a</sup>, Gary Gilleskie<sup>b</sup>, Jakob K. Huusom<sup>a,\*</sup>

<sup>a</sup>*PROSYS, Dept. of Chemical and Biochemical Engineering, Technical University of Denmark, Søtofts Plads, Building 228A, 2800 Kgs. Lyngby, Denmark.*

<sup>b</sup>*Golden LEAF, BTEC, NC State University, 27695 Raleigh, NC, USA*

<sup>c</sup>*Departamento de Ingeniería Química, Universidad Nacional de Colombia – 170003 Manizales, Colombia*

\**jkh@kt.dtu.dk,*

## Abstract

In this work, a model is developed and implemented for GFP<sub>UV</sub> production with aerobic fed-batch fermentation of *E. coli* BL21 (DE3). The model parameters are estimated using historical process data and minimizing the prediction and measurement error. The model implements an extended Kalman filter for non-linear state estimation of biomass, glucose, and dissolved oxygen concentration. The filter includes an existing cascade feed-back loop for dissolved oxygen control which improves the predictive accuracy of the filter. The estimator is used during fermentation to predict the induction point based on a threshold glucose concentration which is otherwise determined exclusively with at-line measurements. The validation examples presented in this work show great agreement between the estimated and measured glucose concentrations, making it a useful tool for predicting the time until induction without requiring high-frequency at-line sampling.

**Keywords:** Digital Twin, Biobased Manufacturing, Operator Support, Scheduling

## 1. Introduction

Digitalisation in biomanufacturing aims to increase production efficiency and reduce time-to-market by tightly integrating physical and digital entities to enable smart manufacturing and faster characterization of fermentation kinetics through improved data-utilization [1,2]. Despite many biobased manufacturers are improving their digital infrastructures and data collection paradigms, few have addressed specific value generating activities associated with improved data access and -generation [2]. A challenge in biobased manufacturing is the generally low degree of automation compared to the petrochemical industry. The complex and non-linear nature of microbial and metabolic systems make it difficult and computationally expensive to predict the fermentation progression [1]. Understanding how to operate these processes and mitigate disturbances during operation often require operators with significant hands-on experience to ensure a steady operation thus making manufacturers heavily dependent on maintaining a stable of skilled operators (and clear procedures) to meet production targets [1]. Retaining the operator knowhow could help to reduce the time and cost of educating new employees and simultaneously guide the new operators during operation [3].



This work presents a protocol employing a standard host organism, *Escherichia coli* BL21 (DE3), commonly used to produce recombinant proteins in the biopharmaceutical industry [4]. The protocol is used in a course curriculum where students with little experience conduct a pilot-scale exercises. Here the operators are required to sample the fermentation broth and measure the glucose concentration and optical density at-line to determine when to induce. After induction, the operators initiate a fed-batch phase. Induction shifts the metabolic pathways to favour the formation of the recombinant protein, and mistimed induction has previously been shown to significantly affect the yield and productivity [5]. Tools that support the operators during the experimental run is therefore especially valuable to ensure a steady and predictable operation. Since the protocol is employed extensively as part of the course curriculum, there exists a large historian of process- and at-line lab data pertaining to the production of green fluorescent protein by *E. coli* aerobic fermentation. This data will be used to build a non-linear state-estimator to predict the induction timing based on concentration thresholds dictated by the protocol.

## 2. Materials & Methods

### 2.1 Fermentation Protocol

The presented protocol serves as a practical exercise in a short course curriculum on topics in advanced fermentation offered at the Golden LEAF Biomanufacturing Training and Education Center (BTEC), located at North Carolina State University. Recombinant green fluorescent protein (GFP<sub>UV</sub>) exhibiting a high ultraviolet (UV) reflectivity is produced from a batch- & fed-batch aerobic fermentation of *E. coli* BL21 (DE3). Initially a seed culture is inoculated with a glucose- and yeast extract rich media and fermented in aerobic batch conditions at 50% dissolved oxygen (DO) saturation. The glucose is added in excess, resulting in the production and accumulation of Acetic acid during the lag and log phase. After running the batch condition for 5-6 hours, the glucose concentration has decreased to approximately 2 g/L. The operator then induces the broth with isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG), shifting the metabolic pathways to favour the formation of the recombinant GFP<sub>UV</sub>. The fermentation then enters a fed-batch phase, where the operator controls the feed flow rate of a 2 L glucose and yeast extract rich media until the supply runs out. As the organism enters the stationary phase and the feeding media is depleted, the accumulated Acetic acid is also consumed. Approximately 24 hours after inoculation the fermentation is terminated, and the product is harvested.

### 2.2 Equipment and Data Handling

BTEC employs 5 identical 30L Sartorius® bioreactors to run several instances of the protocol in parallel. Each bioreactor is augmented with high frequency sensors measuring pressure, temperature, DO, pH, and optical density (OD). The OD sensor measures near-infrared (NIR) wavelengths and correlates the results to OD<sub>600</sub> absorbance. The absorbance at 600 nm corresponds to the dry cell weight of *E. coli* and functions as a soft sensor determination of the cell density. Additionally, the operator periodically takes samples from the broth and perform at-line determination of the glucose concentration and OD<sub>600</sub> through spectrophotometric methods. The DO is controlled by a closed loop cascade that manipulates agitation and aeration. The cascade control strategy primarily manipulate agitation until it reaches a maximum value. The inner control loop will then start to manipulate air sparging flowrate to reach the DO setpoint.

Data collection, storage and querying is orchestrated by AVEVA PI hosted on a central database server located at BTEC. A human-machine interface (HMI) is developed using AVEVA's DeltaV software and hosted on the same central server. The operator can easily input at-line measurement results through the HMI ensuring that the low-frequency measurements are stored in the same concise database and historian as the high-frequency sensors. Online accessibility and tight integration between the digital- and physical entities allow for easy online monitoring and control of the bioreactors. This presents itself as a digitally mature system with a great potential for augmenting a digital twin as depicted in Figure 1. This can be used to support operators with control suggestions during run-time or as virtual laboratories to practice operation strategies before the exercise.

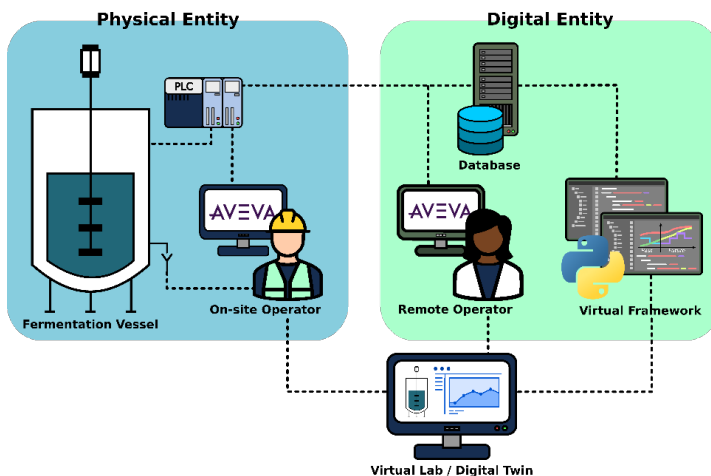


Figure 1: A schematic representation of the digital infrastructure at Golden LEAF BTEC. Data flows from the process unit to a central database and historian hosted with AVEVA PI and human machine interface (HMI) built in DeltaV. The HMI can be accessed remotely for control and monitoring when located on the BTEC network. The Virtual Laboratory / Digital Twin can be used to perform high-level model development and operator support during run-times.

A total of 53 fermentation runs have been extracted from the PI Vision database, conducted between 2020 and 2022. The measurements from these experimental runs are used in the objective function presented in Table 1 and include sensor- and at-line OD, DO sensor and at-line glucose.

### 2.3 Model & Nonlinear State-Estimator

This study employs an acetate overflow model to describe the metabolic behaviour for GFP<sub>UV</sub> formation in *E. coli* BL21 (DE3) this model is derived from a modification to the implementation originally presented by Christina et al. in 2003 [6]. The model by Christina et al has served as a foundation for recent studies in similar systems of *E. coli* BL21 (DE3) [5, 7]. The protocol only requires operators to collect endpoint titer data, which complicates parameter estimation for kinetic GFP production. Therefore, it is not currently possible to include product formation as a modelled state in the dynamic model. Since this study has focus on the pre-induction kinetics, the formation of GFP can be assumed to be reasonably small such that it should not skew the estimated yield coefficients significantly by excluding it from the mass balance constraints. Instead, the aim is to develop an accurate glucose concentration estimator to assist operators in

determining when the glucose concentration is expected to reach the induction target given by the fermentation protocol. The number of parameters in the model explodes with increasing complexity. As a result, the final model is restricted to three rate expressions [4,5,7]. Parameter bounds are determined from literature review and heuristics to reduce the solution space [4,7]. Regardless, the problem is currently a concave optimization problem, and solutions will thus depend heavily on initial conditions. Good initial guesses are therefore necessary to extract dynamics that have realistic behaviour. The yield coefficients ( $Y_{ij}$ ) are restricted to satisfy a mass balance for each rate expression. In the equations presented in Table 1. the species Glucose (S), Acetate (A), Oxygen (O), Carbon Dioxide (CO<sub>2</sub>), Biomass (X) and Yeast Extract (YE) are modelled with Monod expressions, including the parameters, maximum rate ( $q_{S,max}$ ,  $q_{O,max}$ ,  $q_{A,max}$ ), half-times and inhibitors (K and K<sub>I</sub>).

Table 1: stoichiometric equations, rate expressions, objective function and parameters used in the presented model.

<b>STOICHIOMETRY</b>		
Glucose respiration	$S + Y_{OS1}O + Y_{YS1}YE \xrightarrow{\mu_1 X} Y_{XS1}X + Y_{CS1}CO_2$	$q_{S,max}$ 2.25
Glucose respiration	$S + Y_{OS2}O + Y_{YS2}YE \xrightarrow{\mu_2 X} Y_{XS2}X + Y_{CS2}CO_2 + Y_{AS}A$	$q_{S,max}$ 1.12
Glucose respiration	$A + Y_{OA}O + Y_{YA}YE \xrightarrow{\mu_3 X} Y_{XA}X + Y_{CA}CO_2$	$q_{O,max}$ 0.98
		$K_S$ 1.00
		$K_{SO}$ $3.97 \cdot 10^{-3}$
		$K_A$ $5.33 \cdot 10^{-2}$
		$K_{AO}$ $1.01 \cdot 10^{-4}$
		$K_{IS}$ 0.19
		$K_{IA}$ $2.48 \cdot 10^{-3}$
		$Y_{XS1}$ 0.48
		$Y_{XS2}$ 0.67
		$Y_{XA}$ 0.41
		$Y_{AS}$ 0.29
		$Y_{OS1}$ 1.72
		$Y_{OS2}$ 0.95
		$Y_{OA}$ 1.36
		$Y_{CS1}$ 0.27
		$Y_{CS2}$ 0.42
		$Y_{CA}$ 0.35
		$Y_{YS1}$ 0.14
		$Y_{YS2}$ 0.16
		$Y_{YA}$ 0.15
<b>RATE EXPRESSIONS</b>		
Critical substrate uptake rate	$q_{S,crit} = \frac{q_{O,max}}{Y_{OS}} \cdot \left( \frac{K_{IA}}{K_{IA}+A} \right)$	
Substrate uptake rate	$q_S = q_{S,max} \cdot \left( \frac{S}{K_S+S} \right) \cdot \left( \frac{O}{K_O+O} \right)$	
Acetate uptake rate	$q_A = q_{A,max} \cdot \left( \frac{A}{K_A+A} \right) \cdot \left( \frac{O}{K_{AO}+O} \right) \cdot \left( \frac{K_{IS}}{K_{IS}+S} \right)$	
<b>REACTION RATES</b>		
	$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} \min(q_S, q_{S,crit}) \\ \max(0, q_S - q_{S,crit}) \\ \max(0, q_A) \end{bmatrix}$	
<b>PARAMETER FITTING</b>		
Objective function	$obj = \min_{\theta}  y_{DO,OD,S} - \hat{y}_{O,X,S}(\theta) _Q$ s.t.	
Mass balance	$mb = \sum Y \cdot \mu$	
Redox balance	$rb = \sum Y \cdot \mu \cdot \epsilon$	

The objective function is defined to minimize the error between sensor measurements and the simulated results. This includes online optical density and dissolved oxygen sensor values combined with at-line determination of OD<sub>600</sub> and glucose concentrations. The weights are assigned to improve prediction accuracy for glucose and balance the contribution from at-line and online OD measurements. The optimization scheme employs a Nelder-Mead algorithm that penalises a scaled error between the measured- and simulated state, restricted by mass- and redox-balances. From the parameterized model, a non-linear state estimator is constructed. For this purpose, the Extended Kalman Filter (EKF) is formulated.

### 3. Results & Discussion

A preliminary assessment of the datasets shows high variance in total GFP titer ( $2.57 \pm 1.05$  g/L), product yield on the substrate ( $0.0743 \pm 0.0011$  g/g), glucose concentration at

the induction point ( $2.57 \pm 1.05$  g/L), and induction time after inoculation ( $5.91 \pm 0.27$ ). These results support the hypothesis that inexperienced operators naturally infer high variance to a process despite following the same protocol. To validate the performance of the parameterized model, two batch runs are assessed that were not previously included in the parameter estimation scheme. Here an  $R^2$ -value for prediction of glucose and biomass were found to be 0.93 and 0.90 respectively, confirming a good accuracy between the model and the system.

Assuming that the most recent sensor measurements for the two cases are collected at  $\sim 5.25$  hours and  $\sim 5.80$  hours, it is possible to predict the evolution using the EKF. The measurements and mixture details at inoculation are passed to the EKF, which computes states and predict measurement trajectories. A qualitative assessment of the predictions presented in Figure 3 show a reasonable agreement between the predictions and sensor measurements, especially for determination of glucose concentration. The glucose predictions are then used to determine an estimate for timing the induction.

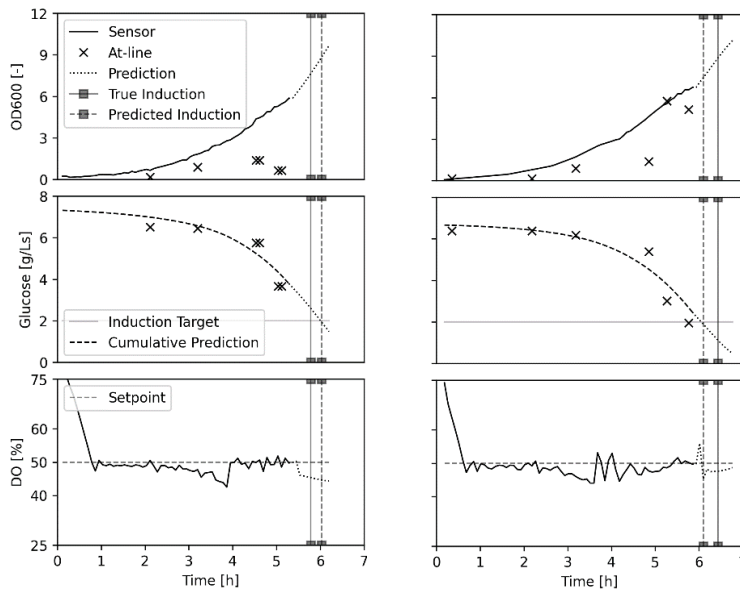


Figure 2: Measurements and predictions for two fermentation runs (left #1 & right #9) over time, beginning at inoculation of the fermentor. (left) Achieves induction close to the desired condition. (right) Overshoots the induction point by approximately 15 min reducing the concentration of glucose to 0.94 g/L.

As show that in Figure 2 the batch runs are expected to reach induction conditions approximately 6 hours after inoculation. In this way, operators would be able to receive the prediction in advance and adequately prepare for induction at the desired conditions without relying solely on at-line glucose measurements. Alternatively, it could assist operators in determining sections in the fermentation run where high-frequency at-line sampling has an added value. As evident by the case depicted on the right-hand side of Figure 2, operators miss the induction point by approximately 20 min. This could be caused by too infrequent at-line sampling during the phase with steep consumption of glucose leading to the operators being poorly prepared to perform the induction. From the

results presented by Lopes et al. [5], this deviation should be identifiable in the yield and productivity by comparing the two batch runs at the end of the fermentation. However, assessing the final GFP<sub>UV</sub> titer at harvest after 24 hours shows that both fermentation runs achieve a similar titer and yield on substrate despite the run on the right-hand side initiating induction 20 min past the suggested glucose concentration listed in the protocol. Since the post-induction dynamics change and include feeding of the glucose and yeast extract solution, there are elements potentially explaining how the delayed induction run bounces back, which the current model cannot account for. Variations in later operating points could justify this observation, but another cause could be that the selected induction condition is not optimized to maximise the GFP<sub>UV</sub> yield or productivity. A more thorough study of final GFP<sub>UV</sub> maximisation and the inclusion of fed-batch fermentation behaviour post induction could address this caveat.

## Conclusion

Existing historical data from pilot scale experiments were contextualised and used to estimate parameters in a suggested mechanistic model for *E. coli* BL21 based on state-of-the-art acetate overflow models. The system was validated against process data and used to develop an extended Kalman filter that predicts the evolution of biomass growth, glucose, and oxygen consumption with reasonable accuracy. In contrast to shake flask experiments investigating the same system, delayed induction did not appear to have a significant effect on the final fermentation performance. Since the agitation and aeration are controlled by the closed loop cascade with a DO setpoint at 50% for all fermentations, the root cause of the observed variation must be found elsewhere. The current implementation only account for the pre-induction dynamics, and variance in the glucose feeding strategy after induction might have greater influence on the performance than previously anticipated.

## Acknowledgements

Nordisk Foundation in the frame of the ‘Accelerated Innovation in Manufacturing Biologies’ (AIMBio) project (Grant number NNF19SA0035474).

## References

- F. Böhner et al., 2021, Challenges in Optimization and Control of Biobased Process Systems: An Industrial-Academic Perspective, *Industrial and Engineering Chemistry Research*, Vol. 60, 14985-15003
- I. A. Udugama et al., 2021, Towards Digitalization in Bio-Manufacturing Operations: A Survey on Application of Big Data and Digital Twin Concepts in Denmark: *Frontiers in Chemical Engineering*, Vol. 3, 727152
- M. Muldbak et al., 2022, Digital Twin of a pilot-scale bio-production setup, *Computer Aided Chemical Engineering*, Vol. 49, 1417-1422
- M. Abadil et al., 2021, An experimental assessment of robust control and estimation of acetate concentration in *Escherichia coli* BL21(DE3) fed-batch cultures, *Biochemical Engineering Journal*, 174, 108103
- C. Lopes et al., 2019, Improving the cost effectiveness of enhanced green fluorescent protein production using recombinant *Escherichia coli* BL21 (DE3): Decreasing the expression inducer concentration, *Biotechnology and Applied Biochemistry*, Vol. 66, 527-536
- I. Cristina et al., 2003, Model-based strategies process for computer-aided operation of a recombinant *E. coli* fermentation
- C. Retamal et al., 2018, Parameter estimation of a dynamic model of *Escherichia coli* fed-batch cultures, *Biochemical Engineering Journal*, 135, 22-35

# Optimal dosing of thyroid hormones in hypothyroid patients with an individualized compartmental model

Davide Manca\*, Federico Appiani, Giovanni Colombo

*PSE-Lab, Process Systems Engineering Laboratory, Dipartimento di Chimica, Materiali e Ingegneria Chimica "Giulio Natta", Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy*

*\*davide.manca@polimi.it*

## Abstract

Recent years have seen a growing interest in studying new methods to predict the correct dose of levothyroxine (LT<sub>4</sub>) that is used as a synthetic drug for hypothyroidism treatment and supplements/substitutes the T<sub>4</sub> hormone secreted by the healthy thyroid. The paper proposes a quantitative procedure for the optimal administration of LT<sub>4</sub> through a model-based approach dependent on some individualized parameters of the patient to free the endocrinologist from the current rough guesses of LT<sub>4</sub> dosage. A dedicated algorithm receives as input data the patient features along with the major indicator of thyroid status, *i.e.* the serum concentration of TSH, the thyroid-stimulating hormone. Then the procedure calculates the Residual Thyroid Function (RTF), which quantifies the thyroid efficiency status, and eventually the optimal LT<sub>4</sub> dose. As a complement to the quantitative models available in the scientific literature, we also take into account the age of the patient to determine the optimal dose that can avoid any suboptimal prescriptions regarding hypothyroidism diagnosis for both young and elderly individuals.

**Keywords:** LT<sub>4</sub>; TSH; individualized medicine; anthropometric parameters.

## 1. Introduction

The thyroid gland is the largest organ in our body capable of hormonal secretion, located in the lower part of the neck in front of the trachea, below Adam's apple. It is divided into two conical lobes linked by an isthmus. The physiological functional sites of the gland are the follicles. They comprise a central nucleus, a so-called colloid, which is surrounded by cuboidal cells. These cells exchange the necessary biological elements and products through the cardiovascular system that highly perfuses the thyroid gland. The thyroid is unique among all body organs thanks to its ability to store and concentrate iodine above forty times the normal plasma concentration. Iodine is the essential substance involved in the complex mechanism of hormone production. Elemental iodine is introduced into the body with nourishment, and it is later exploited by the thyroid to produce triiodothyronine (T<sub>3</sub>) and tetraiodothyronine (T<sub>4</sub>), the two main hormones secreted by the gland. The role of this organ is of vital importance since it can modulate processes such as thermoregulation, body growth, cellular metabolism, breathing, and heart rate. The two thyroid-secreted hormones are synthesized in the follicle in a complex series of metabolic interactions, mainly involving thyroglobulin protein, which is produced in the thyrocytes and then delivered into the colloid. Here, it interacts with iodine to form T<sub>3</sub> and T<sub>4</sub> precursors, containing three and four iodine atoms per molecule respectively, which are

metabolized by specific enzymes to obtain the actual hormones eventually released into the bloodstream. A specific pituitary-secreted molecule called thyroid-stimulating hormone (TSH), which in turn is stimulated by the hypothalamic hormone thyrotropin-releasing hormone (TRH), promotes the cascade of events related to hormonal synthesis. From these intricate relations, it is evident the tight interconnections that link the hypothalamus, pituitary gland, and thyroid, in what is known as the H-P-T axis. These organs enable a peculiar control mechanism. TSH can stimulate the thyroid gland by increasing its  $T_3$  and  $T_4$  production, but the thyroid hormones, especially  $T_3$ , inhibit the TSH release by acting on the gene expression of TRH, which consequently hinders the TSH secretion (Werner *et al.*, 2005). Thyroid diseases can be traced back to the failure of critical links in the chain of events that have been described so far. These disorders can be divided into hypothyroidism and hyperthyroidism. Hypothyroidism is the most common thyroid disorder, where the patient experiences a low level of  $T_4$  and a high level of TSH, caused either by pathologies directly linked with thyroid failure, referred to as primary hypothyroidism, or pathologies associated with hypothalamus or pituitary gland dysfunction, better known as central hypothyroidism. This disease has an estimated incidence of 4 to 5 individuals every 1000 per year in women and 0.6 to 0.9 individuals every 1000 per year in men. Prevalence exceeds 5% of the global population (Flynn *et al.*, 2004).

Researchers in the last six decades put constant efforts into establishing a mathematical model capable of mimicking the thyroid behavior. More recently, there has been an attempt to use the available models to predict the drug dosage for specific thyroid pathologies. The implementation of such models is motivated by two main controversies regarding the diagnosis of hypothyroidism and the prescription criterion of levothyroxine ( $LT_4$ ), which is the active principle used to cure hypothyroidism and is the levorotatory enantiomer of the actual  $T_4$  hormone in the form of sodium salt. The first issue concerns the evaluation methods of the hypothyroidism status, which are primarily based on the analysis of the TSH values, whose allowed range must fall between 0.45 and 4.5 mIU/l even though that range is mired in debates (Werner *et al.*, 2005, Chapter 46). Based on data belonging to patients under levothyroxine treatment, it is rather common to register miscalculations of the  $LT_4$  dose, which is a major health risk. This issue should be considered of utmost importance. Health studies showed that 22% of analyzed subjects were improperly treated and that 18% of them were receiving higher  $LT_4$  doses than necessary, with consequent negative side effects, namely dyslipidemia, cardiovascular disorder risks, and, particularly, iatrogenic thyrotoxicosis (Werner *et al.*, 2005, Chapter 46). In addition,  $LT_4$  has been classified as a narrow therapeutic index drug, given the utter care that must be put into the titration of the dose, considering that even small changes can significantly affect the existing equilibrium of the H-P-T axis (Werner *et al.*, 2005, Chapter 13A). Current goal of the recently developed thyroid models is including individualized parameters in the attempt to set new standards for hypothyroidism diagnosis and  $LT_4$  prediction. Among the available models, p-Thyrosim stands out for its capability of simulating the thyroid biological mechanism as a function of the patient's parameters such as gender, height, weight, residual thyroid function (*i.e.* quantification of the hypothyroidism status), and  $LT_4$  administered dose (Cruz-Loya *et al.*, 2022).

## 2. Methods

Eisenberg and coauthors (2008) outlined the base structure of a multi-compartmental PBPK model to describe the physiological feedback control of the H-P-T axis. The reference model explains the H-P-T axis components by means of homogenous

compartments. There are three main submodels: thyroid hormones (TH), Brain (B), and TH input (THI). The interconnection of each of them represents the interaction of the H-P-T axis that allows evaluating the trend of both TSH and the thyroid hormones in plasma. The TH submodel includes a set of ordinary differential equations that govern the thyroid hormones' fluxes from the plasma to the tissues and vice versa, occurring in the whole body. Then, the B submodel describes the endogenous circadian TSH secretion by lumping the interaction between the pituitary gland and the hypothalamus in a single equation, which is mainly a function of the  $T_3$  level in the brain. Finally, the THI submodel contains the equations concerning the exogenous administration of  $LT_4$  employing four alternative paths: intravenous, intramuscular, subcutaneous, or oral administration (Eisenberg *et al.*, 2008). This model saw a relevant development described in Cruz-Loya *et al.*, (2022) where the authors succeeded in implementing individual parameters such as body weight, height, and gender and proposed a suitable formulation to describe the circadian oscillation of TSH that affects the thyroid secretory pattern.

In case of patients with thyroid problems, it is common practice to consider the TSH concentration measured with a first blood analysis. If there is a positive match for hypothyroidism, the endocrinologist proceeds with an initial dose of  $LT_4$ , which can be determined in a variety of ways. The preferred  $LT_4$  dose is  $1.6 \mu\text{g}/\text{kg}$  per day but there is also an extremely simplified approach based on a one-size-fits-all of  $100 \mu\text{g}/\text{day}$ . Afterwards, the patient has to take the remedy every morning before breakfast for at least six weeks after which a new blood sample provides the updated value of TSH that allows establishing if the thyroid levels belong to the recommended ranges. If they are not, the dose is empirically adjusted, and the procedure is iterated until normal hormone levels are reached (Werner *et al.*, 2005, Chapter 46). It is rather common to undergo a few iterations before an optimal prescription is achieved, which makes the whole procedure demanding for the patient and subject to the biased evaluation of the endocrinologist. Therefore, an optimized, individualized, and unbiased protocol is highly recommended.

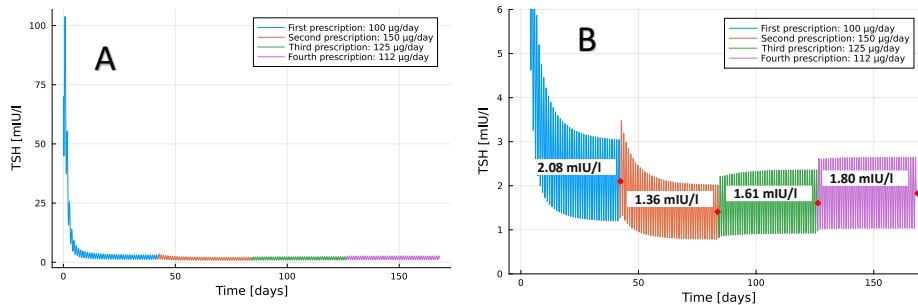


Figure 1: Panel (A) simulated endocrinologist trial and error approach in  $LT_4$  dose administration for a 70-year-old hypothyroid female patient of 78 kg, 1.63 m, and TSH starting level of 71.44 mIU/l. Panel (B) zoomed trend. Red diamonds report the asymptotic TSH values after the iterative adjustments of  $LT_4$  prescriptions.

Panel A of Figure 1 reproduces a hypothetical situation based on the traditional approach to the  $LT_4$  dosing scheme where the endocrinologist wants the TSH to be far from the upper limit to prevent the persistence of hypothyroidal symptoms. Based on this assumption, the optimal value of TSH is set below the average of the euthyroid range (0.45-4.5 mIU/l), to a value of 1.9 mIU/l. A 70-year-old female, weighing 78 kg and 1.63 m tall undergoes a blood sample to measure her TSH concentration for the first time. TSH is 71.44 mIU/l, and the endocrinologist (neglecting any individualized features) prescribes a theoretical starting dose of  $100 \mu\text{g}/\text{day}$  (Werner *et al.*, 2005, Chapter 46).



This produces a six-week asymptotic TSH value of 2.08 mIU/l, which is higher than the expected value and might make the symptoms persist. The dose is then increased to 150 µg/day and, after six weeks, TSH reaches 1.36 mIU/l, which is now too low to be acceptable and bearable for prolonged treatments. The endocrinologist then fine-tunes the dose to 125 µg/day after which TSH reaches 1.61 mIU/l. Since the LT<sub>4</sub> has a narrow therapeutic window, the endocrinologist chooses to further change the dose to 112 µg/day, which eventually corresponds to a TSH of 1.80 mIU/l which is a rather good value reasonably near to the expected one (1.9 mIU/l).

An alternative approach consists of the automated optimal LT<sub>4</sub> dosing by exploiting the power of quantitative deterministic models such as p-Thyrosim. With respect to the above iterative search for a suitable prescription by the human interaction between the endocrinologist and the patient, now the goal is to minimize the blood samples (and therefore the time to find the optimal dosage as each dose adjustment calls for 50-60 day more settling times) by getting the starting dose right the first time. To this purpose, we implemented an optimizing procedure that can return the optimal LT<sub>4</sub> tablet dose for hypothyroid patients, using p-Thyrosim and accounting for the patient's individual features such as weight, height, gender, age, and the TSH value from a recent blood sample. First, the proposed procedure evaluates the RTF (Residual Thyroid Function) value by zeroing the following function:

$$f(RTF) = TSH^{exp} - TSH(RTF) \quad (1)$$

Where the  $TSH^{exp}$  and  $TSH$  [mIU/l] are the TSH blood sample and the p-Thyrosim asymptotic prediction of the in-silico individualized patient  $TSH$  plasma dynamics, respectively. Once RTF is computed, the procedure zeros Equation (2), by determining the optimal LT<sub>4</sub> dose that would yield to euthyroid values of both  $T_4$  and  $TSH$ :

$$f(LT_4) = \left| \frac{TSH^{sp} - TSH(LT_4)}{4} \right| + \left| \frac{T_4^{sp} - T_4(LT_4)}{60} \right| \quad (2)$$

where  $TSH^{sp}$  [mIU/l] and  $T_4^{sp}$  [µg/l] are the optimal expected values (*i.e.* setpoints) for the corresponding euthyroid values. Initially,  $TSH^{sp}$  was set at 1.80 mIU/l for any hypothyroid patient as Cruz-Loya *et al.*, (2022) suggest a minor dependency from age as far as euthyroid conditions are concerned. However, given the high precision that must be kept in drug dosage, we also accounted for the dependency of  $TSH^{sp}$  from age. We derived an age-dependent trend for the upper limit of TSH euthyroid from Raverot *et al.*, (2020) that took to Equation (3).

$$TSH_{M,F}^{sp}(age) = \frac{TSH_{M,F}^{ul}(age) - TSH^{ll}}{2.7} \quad (3)$$

Where  $TSH_{M,F}^{ul}$  and  $TSH^{ll}$  are the upper and lower limits of the TSH range respectively for either a male or female euthyroid patient.  $TSH_{M,F}^{ul}(age)$  was evaluated by extrapolating the values from the TSH upper limit trends, both for males and females, from the work of Walsh (2022). Therefore, the data of Figure 2 were regressed with a nonlinear procedure. This methodology generates also the TSH upper limit formula from which it is possible to calculate, for every combination of age and gender, the associated upper limit and eventually derive the correct  $TSH^{sp}$ :

$$TSH_{Male}^{ul}(age) = 4.642520589559373 - 0.02039878317500956 \cdot age + 0.0002614981700062368 \cdot age^2 \quad (4)$$

$$TSH_{Female}^{ul}(age) = 4.1777986229928485 + 0.00463599130382207 \cdot age + 0.0001309278501772748 \cdot age^2 \quad (5)$$

Instead, the  $T_4^{sp}$  plasma values have been considered equal for both genders and fixed at 80 µg/l as the literature does not report any significant variability with age (Werner *et al.*, 2005, Chapter 11A).

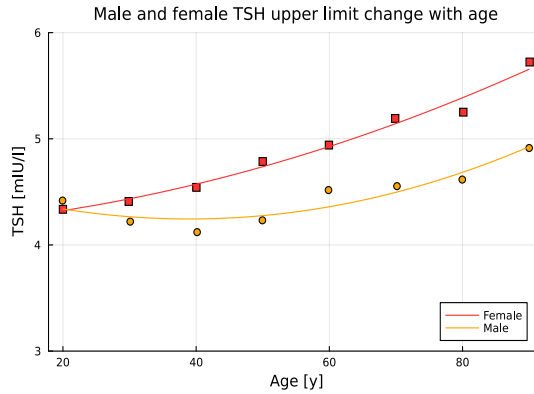


Figure 2: TSH upper limit trend for males and females at different ages derived by a nonlinear regression procedure.

### 3. Results and discussion

The proposed automated algorithm predicts the necessary  $LT_4$  tablet dose (commercially available) to be administered to a hypothyroid patient by accounting for their individual features and endogenous TSH (before the treatment). The procedure, based on the p-Thyrosim model (Cruz-Loya *et al.*, 2022) can investigate how the optimal dose changes as a function of the patient's age. The dependency of  $TSH^{ul}$  with age, seen in Equation (3), affects  $TSH^{sp}$ . If the TSH upper limit is somehow increased, the  $LT_4$  dose must decrease according to Equation (2). Thus, the implementation of the automated procedure of the  $TSH^{sp}$  age dependency ensures the achievement of different dosages for the same hypothyroid patient at different ages, allowing a better approach to individualized treatment.

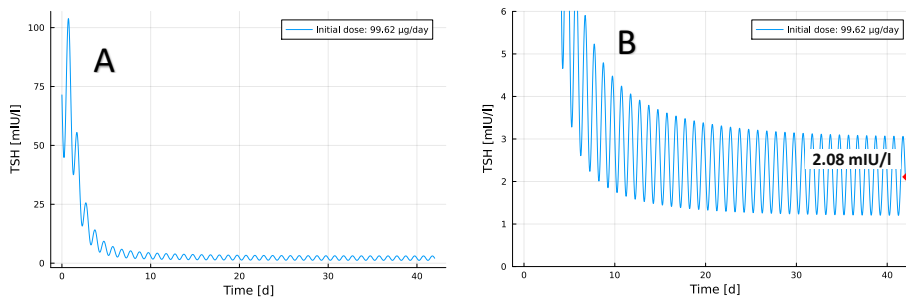


Figure 3: (A) Simulated TSH plasma value for a 70-year-old hypothyroid female patient with a weight of 78 kg, 1.63 m, and TSH starting level of 71.44 mIU/l under  $LT_4$  treatment calculated through the double-step algorithm routine. (B) Simulated TSH plasma value for the same patient highlighting with a red marker the achievement of the correct TSH setpoint thanks to the theoretical optimal  $LT_4$  dose (99.62 µg/day) evaluated by the optimizing procedure.

With reference to Figure 3, we tested our double-step algorithm for the same female hypothyroid patient in Figure 1. The main differences between the graphs must be searched in the diverse  $LT_4$  dose treatment schemes and in the age-dependent approach related to the individualized  $TSH^{sp}$  value. Figure 1 shows how the iterative human-based procedure resulted in a total of 160 days and 4 adjustments before identifying the correct  $LT_4$  dose centered on the endocrinologist's perspective. Conversely, Figure 3 shows the direct assessment of the optimal dose, which allows for reaching the expected TSH

optimal value by a one-shot prescription. Furthermore, the TSH setpoint chosen in the dosing scheme of Figure 1 does not account for the  $TSH^{ul}$  dependency with age as the patient is assigned a constant setpoint at any stage of their life. Thus, the proposed optimizing procedure assigns the TSH setpoint as a function of both the patient's age and gender. For the considered 70-year-old female patient the correct TSH level should be 2.08 mIU/l, which is achieved by an  $LT_4$  dose of 99.62  $\mu\text{g}/\text{day}$ , found by the optimization algorithm. That dose is not available in commercial tablets and should be rounded to the nearest available formulation of 100  $\mu\text{g}/\text{day}$  which would bring an asymptotic TSH value of 2.07 mIU/l, almost equal to the optimal one. It is worth observing that the scenario reported in Figure 1 calls for the involvement of an earnest physician. In real life, the endocrinologist might not look for those precise adjustments and might be satisfied with any reasonable dosages. Nonetheless, our automated and unbiased procedure can overcome the endocrinologist's sensitivity in the dose prescription and displays a lower time to reach the optimal dose based on a one-shot and more accurate treatment that takes into account the individual features of the patient.

#### 4. Conclusions

The findings reported in the present paper were made possible by the decades-long studies published in the scientific literature. Our efforts concerned the application of compartmental models simulating thyroid secretion to replicate real-life scenarios, in the attempt to increment the precision of the first  $LT_4$  tablet dose for hypothyroid patients. We developed an automated approach through a double-step procedure to support the endocrinologist's decisions in terms of optimal  $LT_4$  dosage by exploiting the individualization features and unbiased calculations of p-Thyrosim. The attempt of implementing age as a further individualized parameter for hypothyroid patients allowed for outlining the dependency of TSH at different stages of life and improving the evaluation of the  $LT_4$  optimal dose.

#### References

- Cruz-Loya, M., Chu, B., Jonklaas, J., Schneider, D., & DiStefano, J. (2022). Optimized Replacement T4 and T4+T3 Dosing in Male and Female Hypothyroid Patients With Different BMIs Using a Personalized Mechanistic Model of Thyroid Hormone Regulation Dynamics. *Frontiers in Endocrinology*, 13.
- Eisenberg, M., Samuels, M., & DiStefano J. (2008). Extensions, validation, and clinical applications of a feedback control system simulator of the hypothalamo-pituitary-thyroid axis. *Thyroid*, 18(10), 1071–1085.
- Flynn, R. W. v, MacDonald, T. M., Morris, A. D., Jung, R. T., & Leese, G. P. (2004). The thyroid epidemiology, audit, and research study: thyroid dysfunction in the general population. *The Journal of Clinical Endocrinology & Metabolism*, 89(8), 3879–3884.
- Hollowell, J. G., Staehling, N. W., Flanders, W. D., Hannon, W. H., Gunter, E. W., Spencer, C. A., & Braverman, L. E. (2002). Serum TSH, T4, and thyroid antibodies in the United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III). *The Journal of Clinical Endocrinology & Metabolism*, 87(2), 489–499.
- Raverot, V., Bonjour, M., Abeillon du Payrat, J., Perrin, P., Roucher-Boulez, F., Lasolle, H., Subtil, F., & Borson-Chazot, F. (2020). Age- and sex-specific TSH upper-limit reference intervals in the general French population: there is a need to adjust our actual practices. *Journal of Clinical Medicine*, 9(3), 792.
- Walsh, J. P. (2022). Thyroid Function across the Lifespan: Do Age-Related Changes Matter? *Endocrinology and Metabolism*, 37(2), 208.
- Werner, S. C., Ingbar, S. H., Braverman, L. E., & Utiger, R. D. (2005). *Werner & Ingbar's the thyroid a fundamental and clinical text* (9th ed.). Lippincott Williams & Wilkins.

# Technoeconomic process modelling of waste remediation within a supply chain network

Robert Milton, Jagroop Pandhal and Solomon F. Brown

*Department of Chemical and Biological Engineering, The University of Sheffield,  
Sheffield, S1 3JD, United Kingdom  
r.a.milton@sheffield.ac.uk*

## Abstract

Technoeconomic modelling links engineering performance to economic outcomes. In this paper we present a method for embedding technoeconomic unit process models within a geographically distributed supply chain simulator. This enables the simultaneous optimization of how and where processing is performed. This is particularly suited to waste remediation, where the costs of transporting hazardous materials may outweigh those of performing some stages of treatment at the waste source. The approach is applied to a microalgal-bacterial remediation of hydrocarbon waste (sump oil mainly) produced by the MOD Sites (military bases) across the UK. Through this treatment, valuable bioproducts are recovered from algae which consumes hydrocarbon waste and CO<sub>2</sub> as it grows.

**Keywords:** Technoeconomic analysis, supply chain modelling, waste remediation.

## 1. Introduction

Technoeconomic modelling has long been a basic step in process design, linking engineering performance to economic outcomes (Seider et al 2019). The operation and optimisation of individual processes within wider, distributed supply networks is typically considered with simplified models of the underlying processes, and may be independent of the scheduling of the operations contained therein (Garcia et al 2015). In this study, we address this by presenting software for embedding technoeconomic unit process models within a geographically distributed supply chain simulator. This enables the simultaneous optimization of how and where processing is performed. This is particularly suited to waste remediation, where the costs of transporting hazardous materials may outweigh those of performing some stages of treatment at the waste source.

As an example, the approach is applied to a microalgal-bacterial remediation of hydrocarbon waste. Microalgal-bacterial remediation is a multi-stage process to cultivate, extract and exploit the bioproducts produced by microalgae grown in hydrocarbon effluent. Through this treatment, valuable bioproducts are recovered from algae which consumes hydrocarbon waste and CO<sub>2</sub> as it grows (Zahra et al 2020). This offsets the costs, both economic and environmental, of our global abundance in hydrocarbon effluents, used oils and lubricants (Fregie et al, 2019).

Unit operations are represented by kinetic models for each stage that is linked to economic performance via cashflow modelling of revenue streams, resource costs (OpEx) and plant costs (CapEx). The unit processes (stages) are implemented as agents in an agent-based supply chain model. Incorporating transport costs, processing is then optimized across

the entire network of waste producing and processing sites. This enables optimal choices between local and centralized processing.

## 2. Architecture

The basic problem architecture, embedding a hierarchy of agents (i.e. agents containing lower-level agents) in a geographically distributed supply chain network is depicted in Fig. 1.

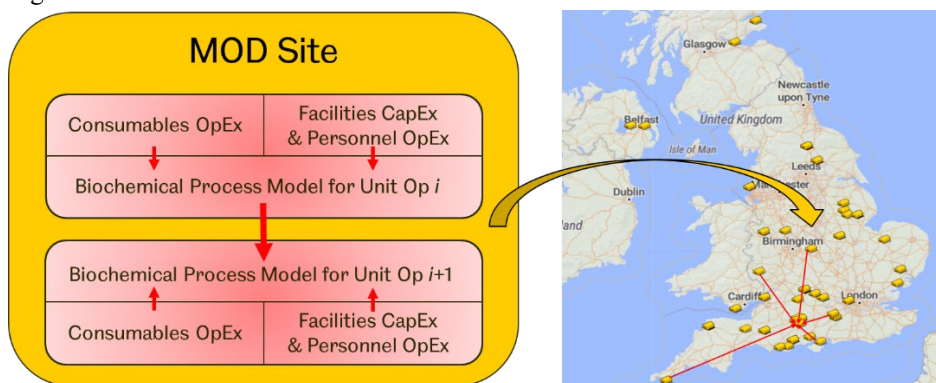


Fig.1. Embedding the layers of architecture – MOD Sites containing Unit Ops – in a supply chain network

### 2.1. Implementation

The Agent Based Model is implemented in AnyLogic (Borshchev 2014) extended with our own Java code). At every level, the representation in AnyLogic/Java is as generic as possible, the details of implementation being stored in a separate Excel spreadsheet for each layer of the agent hierarchy. Each spreadsheet ultimately lists every agent at that level of the hierarchy, providing for each a description specifying every parameter required to construct the agent. Upon model startup, the MOD Sites (military bases) are located by the postcode embedded in the (Excel) description of the site. This automatically exposes the road distance between any two sites, as a feature of AnyLogic.

### 2.2. MOD Site

The description of an MOD Site goes far beyond geographical location, populating all of the parameters in Fig. 2. These specify, amongst other matters, the frequency and amount of material to be processed at this site. Material is based on a single limiting consumable (e.g. hydrocarbon waste) by which to scale reagent amounts and energy requirements. The limiting consumable may be produced at another site, transported and received here, or be hydrocarbon waste produced onsite (the limiting consumable amount). Once processed, products 0-4 may be dispatched to destinations 0-4, which are MOD Sites specified in the description with product routing information. These matters are handled on the right of Fig. 2, by the receiver and dispatcher agents, which calculate scheduling (based on the trucks listed in this MOD Site's description), mileage, and thereby Haulage OpEx including fuel, depreciation per km and personnel.

The type of an MOD Site is used to populate the sequence of Unit Ops to the right of Fig. 2. For example, if the MOD Site is described as "Empty" it does nothing but produce and.

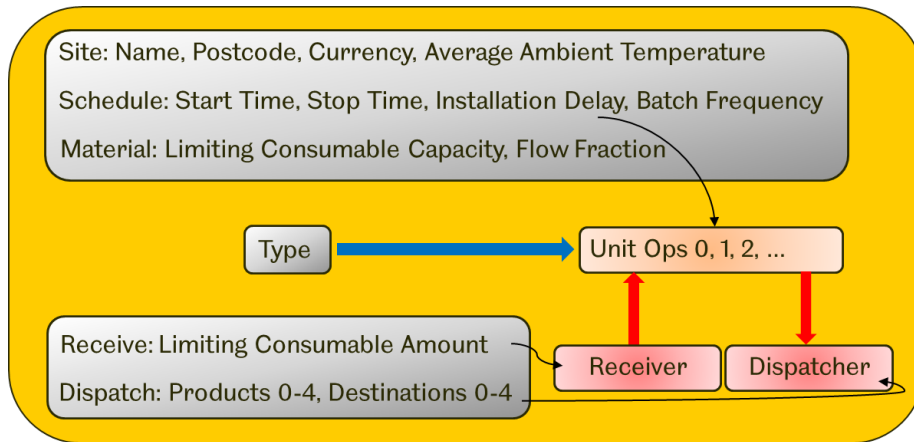


Fig. 2. A generic MOD Site, whose Unit Ops sequence is determined by the Site Type. Parameters in grey feed information according to black arrows. Material flows according to the red arrows.

dispatch hydrocarbon waste, whereas “Cultivation-Centrifuge” processes waste into biomass then dries it

### 2.3. Unit Op

In this implementation a Unit Op is an agent encapsulating the kinetics of a physio-chemical process together with its economics. Based on a generic template (base class or agent), there are dedicated sub-types of Unit Op for each stage of processing -- in this case study cultivation of biomass, dewatering and drying. Further Unit Ops could be added for downstream processing such as cell disruption, and bio-oil extraction and separation. Each Unit Op may operate in batch or flow mode, or a combination of the two. After each regular batch duration the reactor and the outlet are emptied into the product store, cleaned, and refreshed by recovering (from the product store), purchasing and/or receiving the consumables required for the next batch duration. The OpEx record is amended accordingly at this time.

The batch intervals are fixed by the Unit Op’s start time, stop time, installation delay and batch frequency, which are inherited from its parent MOD Site. This information is simultaneously used to calculate CapEx cashflows, including any amortization or depreciation in the description. The product store is at any time available to supply the next Unit Op and/or the haulage dispatch agent at this MOD Site.

There remains the problem of sizing the facilities for any Unit Op, which dictates CapEx expenditure, and heat requirements via volume and surface area. For the initial Unit Op at any MOD Site, the size is determined by the limiting consumable annual capacity scaled by batch frequency. This is used to calculate a product capacity for this Unit Op. This product capacity is used to calculate the limiting consumable annual capacity for the subsequent Unit Op, enabling us to size it, and calculate its product capacity. This reasoning is repeated along the sequence of Unit Ops at any MOD Site.

### 2.4. Economy

At the bottom of the agent hierarchy there is a simple economy layer, encapsulating cashflow schedules, amortization and depreciation, discounting, inflation and currency

LIMITING_CONSUMABLE			yieldCapacities		RATE_PRODUCT		RATE_CONSTANT	
BIOMASS_2_PERCENT			BATCH	1.00	BIOMASS_15_PERCENT		17.8049	
			FLOW	1.00				
			RECOVERY	0.00	BATCH_DURATION		BATCH_LAG	
			LIMITING_CONSUMABLE_SCALE_FACTOR	0.00	0.1151		0.0575	
RECOVERY			yieldFractions		kineticParameters			
LIMITING_CONSUMABLE_SCALE_FACTOR					INERT		KS	
0.00%	0%	0%	OTHER_ELECTRICITY	0.00%	OTHER_ELECTRICITY	100%	1.0000E-12	
0.00%	0%	0%	LIGHT	0.00%	LIGHT	100%	1.0000E-12	
0.00%	0%	0%	HEAT	0.00%	HEAT	100%	1.0000E-12	
0.00%	0%	0%	STEAM	0.00%	STEAM	100%	1.0000E-12	
98.00%	0%	0%	WATER	86.67%	WATER	100%	1.0000E-12	
98.00%	0%	0%	BIOMASS_2_PERCENT	-100.00%	BIOMASS_2_PERCENT	100%	1.0000E-12	
98.00%	0%	0%	BIOMASS_6_PERCENT	0.00%	BIOMASS_6_PERCENT	100%	1.0000E-12	
98.00%	0%	0%	BIOMASS_15_PERCENT	13.33%	BIOMASS_15_PERCENT	100%	1.0000E-12	
98.00%	0%	0%	BIOMASS_20_PERCENT	0.00%	BIOMASS_20_PERCENT	100%	1.0000E-12	
98.00%	0%	0%	BIOMASS_25_PERCENT	0.00%	BIOMASS_25_PERCENT	100%	1.0000E-12	
98.00%	0%	0%	BIOMASS_92_PERCENT	0.00%	BIOMASS_92_PERCENT	100%	1.0000E-12	
0.00%	0%	0%	CHITOSAN	0.00%	CHITOSAN	100%	1.0000E-12	
0.00%	0%	0%	GENERAL_WASTE	0.00%	GENERAL_WASTE	100%	1.0000E-12	
0.00%	0%	0%	SEWERAGE	0.00%	SEWERAGE	100%	1.0000E-12	
			FACILITIES		facilities			
				27	11	FIXED_AMOUNT	BATCH_FACTOR	FLOW_FACTOR
			VOLUME	0	0	1.0000	1.0000	1.0000
			SURFACE_AREA	1	1	1.0000	0.0000	0.0000
			LAND	100,000	100,000	1.0000	1.0000	0.6667
			BUILDINGS	100,001	100,001	1.0000	1.0000	0.6667
			GBP_SUPERVISOR	200,000	200,000	0.0000	0.0000	0.0000
			GBP_TECHNICIAN	200,001	200,001	1.0000	0.0000	0.0000
			CENTRIFUGE	5,000	5,000	0.0000	17.3810	0.0000
			SPIRAL_PLATE	5,100	5,100	0.0000	0.0000	0.0000
			PRESSURE_FILTER	5,200	5,200	0.0000	0.0000	0.0000
			MEMBRANE_FILTER	5,300	5,300	0.0000	0.0000	0.0000
			VACUUM_FILTER	5,400	5,400	0.0000	0.0000	0.0000
			SPRAY_DRYER	5,500	5,500	0.0000	0.0000	0.0000
			DRUM_DRYER	5,600	5,600	0.0000	0.0000	0.0000

Fig 3. A fragment of the Excel description for a dewatering or drying Unit Op, showing some key inputs. Start and stop times are inherited from the parent MOD Site, not the Unit Op description.

Exchange rates. Consumables are stored in a price list that governs purchase, inflation, insurance and other consumables rates. Waste and sewerage are consumables that we sell for a negative price. Likewise, facilities and personnel are stored in a price list, which governs all economic matters including purchase, salvage, lifetime, inflation, contingencies, maintenance, Lang factors and scale factors. These default to industry benchmarks from (Fasaei et al 2018, Seider et al 2019) and market values circa 2021.

### 3. Results

The agent based supply chain model has been applied to preliminary studies of microalgal-bacterial remediation of hydrocarbon waste. Biomass cultivation is assumed to follow Monod kinetics (Monod, 1949), calibrated from laboratory work in a phycoflow reactor. The techno-economic inputs for dewatering and drying are based on literature review (Fasaei et al 2018). The cultivation and dewatering Unit Ops are flexibly distributed amongst 41 military bases across the UK. Scheduling consists of dispatching available products to the receivers of their destinations at the end of each batch duration. Transit times incorporating a loading/unloading delay are taken into account automatically, but must currently be scheduled manually in setting the start and stop times for each MOD Site. Optimal scheduling has not been automated.

A typical simulation result is shown in Fig. 5. This is summary information for the entire network, a complete breakdown of the time evolution of reactants and products, and all CapEx and OpEx cashflows is available down to the Unit Op Level (Fig. 6). This information has been used to show, for example, that centralized processing saves some £5.5m over six years in this network, as economies of scale outweigh transport costs. Surprisingly, drying locally cultivated biomass centrally saves £1.8m over local drying. So economies of scale dominate transport costs, even though wet biomass is very bulky.

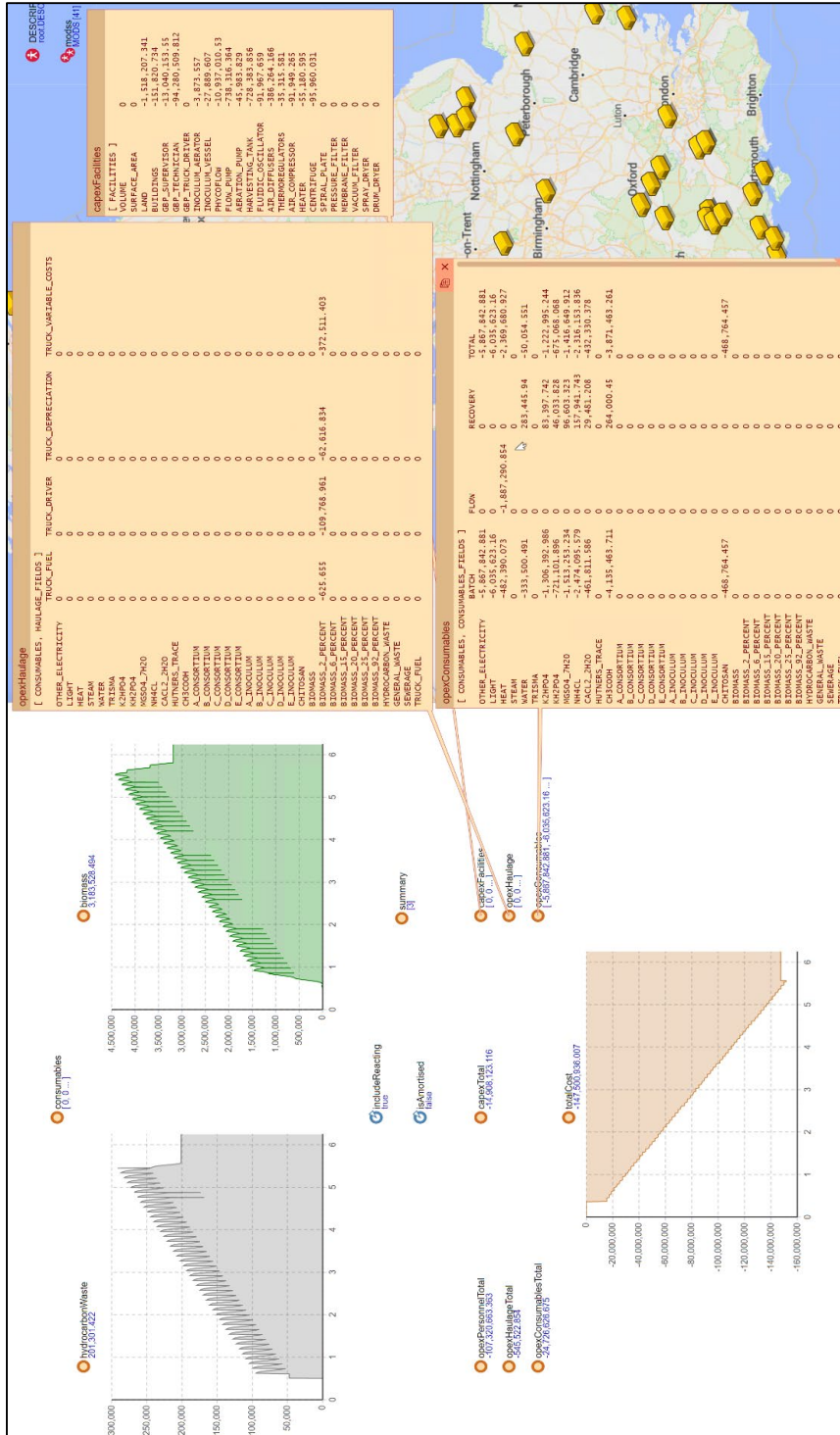


Fig. 5. Accumulated results for a six year nationwide project, showing hydrocarbon waste processed and biomass produced, above expenditure.



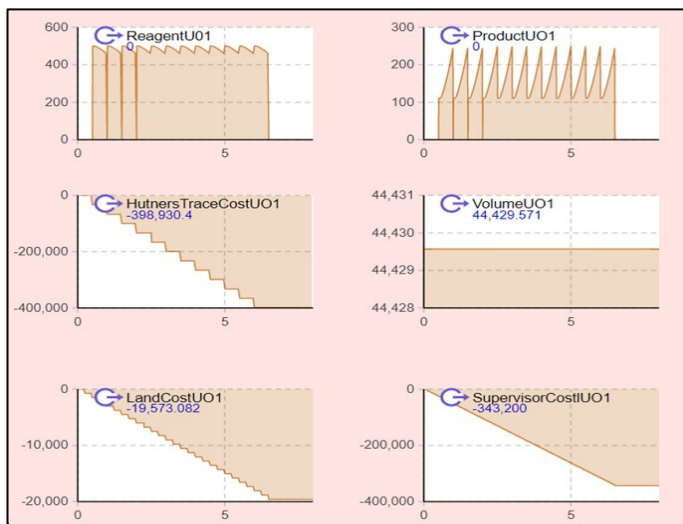


Fig 6. Drill-down results from a cultivation Unit Op.

#### 4. Conclusion

In this paper we have presented a techno-economic process model distributed within a supply chain network. The model is designed as a set of building blocks (Unit Ops) which may be distributed within the supply chain as desired. The many key parameters are neatly organized in a spreadsheet interface of descriptions, specifying location, timing, kinetics and economics. The software is highly extensible, and is almost agnostic to the nature of the Unit Ops. Adding new Unit Ops can be done in a few hours. The runtime for the 41 sites and 82 Unit ops considered is a few seconds. The software architecture is very satisfactory in these regards.

The major challenges only partially addressed concern optimal site capacity and transit scheduling. These could perhaps be optimized automatically in future work. Furthermore, the key outputs – namely material consumed, material produced and economic cost/benefit – should ideally be considered in some optimization scheme in future. Ultimately, it would be wise to analyse the sensitivity of any results to the techno-economic inputs, the future values of which may be quite uncertain.

#### References

- W.D. Seider, J.D. Seader, D. R. Lewin and S. Widagdo. *Product and Process Design Principles: Synthesis, Analysis and Evaluation*, 4<sup>th</sup> edition. John Wiley & Sons Inc, 2019. ISBN 1119588006.
- Garcia, D. J., & You, F. (2015), "Supply chain design and optimization: Challenges and opportunities", *Computers & Chemical Engineering*, vol. 81, pp. 153-170.
- Z. Zahra, D. H. Choo, H. Lee, and A. Parveen, "Cyanobacteria: Review of Current Potentials and Applications", *Environments*, vol. 7, no. 2, p. 13, Feb. 2020.
- S. B. Eregie and S. F. Jamal-Ally, "Comparison of biodegradation of lubricant wastes by *Scenedesmus vacuolatus* vs a microalgal consortium", *Bioremediat. J.*, vol. 23, no. 4, pp. 277–301, Oct. 2019.
- A. Borshchev, *The Big Book of Simulation Modeling: Multimethod Modeling with Anylogic 6*, Anylogic North America, 2014, ISBN 0989573176.
- J. Monod, "The growth of bacterial cultures", *Annual Review of Microbiology*, vol 3, p.371, 1949. doi:10.1146/annurev.mi.03.100149.002103
- F. Fasaei, J.H. Bitter, P.M. Slegers, A.J.B. van Boxtel, " Techno-economic evaluation of microalgae harvesting and dewatering systems", *Algal Research*, vol 31, pp. 347–362, 2018.

# Spatio-temporal analysis of milk safety under climate change

Lydia Katsini<sup>1</sup>, Satyajeet S. Bhonsale<sup>1</sup>, Styliani Roufou<sup>2</sup>, Sholeem Griffin<sup>2</sup>,  
Vasilis ValDRAMIDIS<sup>3</sup>, Simen Akkermans<sup>1</sup>, Monika Polanska<sup>1</sup>,  
Jan F.M. Van Impe<sup>1</sup>

<sup>1</sup>*KU Leuven, Chemical Engineering Department, BioTeC+, Chemical & Biochemical Process Technology & Control, Gebroeders De Smetstraat 1, Ghent 9000, Belgium*

<sup>2</sup>*University of Malta, Department of Food Sciences and Nutrition, MSD 2080, Malta*

<sup>3</sup>*National and Kapodistrian University of Athens, Department of Chemistry, Panepistimioupolis Zografou, Athens, 157 84 Greece*

[jan.vanimpe@kuleuven.be](mailto:jan.vanimpe@kuleuven.be)

## Abstract

Bio-industrial processes can be optimized considering climate change using climate change impact assessments. These are based on models that describe the effect of climate on the affected sector. In this case, we focus on milk safety, thus we study Somatic Cell Counts (SCC), which is an indicator of mastitis, and Total Bacterial Counts (TBC), an indicator of the microbiological state of raw milk. The objective is to identify SSC and TBC patterns by looking into their spatial variability. For this purpose, traditional data mining methods are adopted to cover both space and time using data from 53 farms (2014-2019). According to the results, both SCC and TBC indicate spatial correlation. This means that apart from the already established temporal dependency, there is also a spatial dependency. In conclusion, collecting temporal data only from one farm may be misleading. Climate change-proof process design is essential for the dairy sector.

**Keywords:** milk, climate change, data mining, spatio-temporal analysis

## 1. Introduction

As climate change is expected to impact all aspects of life on the planet, it ranks high in both public and private decision making. The main tool needed in such decision making is climate change impact assessments, enabling a climate change-proof process design. It is based on impact models that describe the effect of multiple climate factors on affected sectors (Ahmad et al., 2001), here the dairy sector.

By utilizing these impact models, future food safety risks can be estimated under climate change (Katsini et al., 2021). For example, Gunn et al. (2019) estimated milk yield losses. By looking into raw milk quality, the already established processing standards can be revisited and re-optimized by taking into account the climate change impact on raw milk. A major consideration for process design is milk safety, i.e., the reduction of microflora in milk. Thus, the analysis involves the Somatic Cell Count (SCC), which is an indicator of the inflammation of the breast tissue of cattle, known as mastitis. In such a case, the milk is unsafe and not fit for human consumption. Moreover, the Total Bacterial Count (TBC) is studied, which is an indicator of the microbiological quality as it expresses the plate counts of the total amount of microflora in raw milk. TBC and SCC are recorded on

a regular basis as they form regulatory thresholds for raw milk according to the EU legislation (EC 2073/2005).

The seasonal patterns of SCC and TBC have already been identified (Stürmer et al., 2018), nevertheless microbial dispersal takes place both in space and time. Thus, the objective of this work is to identify spatial patterns of milk contamination by analyzing the SCC and TBC temporal and spatial variability using data-driven modelling approaches.

## 2. Materials and Methods

### 2.1. Data set and data pre-treatment

Data from 682 dairy farms, located in Spain (A Coruña, Galicia), collected over the timespan 2014–2019 comprise the dairy data set. The first step of data pre-treatment was to screen the farms based on data completeness, e.g., cover all the years. After this screening, 53 farms were selected. Secondly, as the sampling frequencies of different farms were inconsistent, data was resampled to obtain the same sampling frequency for all farms, i.e., every 4 days. Existing outliers were removed and the gaps arising from the previous step were filled in using the median of the adjacent samples (10% interpolated data). Finally, the temporal trajectories for each farm were smoothed using a moving window weighted average. The address of each farm is represented in terms of longitude and latitude coordinates.

### 2.2. Methodology

The farms are clustered using non-linear unsupervised learning methods in terms of their temporal variability. The resulting clusters of farms are compared with their real topology in terms of coordinates. Farms that belong to the same cluster and are located close to each other indicate spatial correlation. SCC and TBC data are treated independently, which means that the above is done first for SCC and then repeated for TBC. On the one hand data-driven modelling techniques can be valuable to find patterns in data, however they remain black-box approaches, which means that the results are not always interpretable.

#### 2.2.1. *t-Distributed Stochastic Neighbor Embedding (t-SNE)*

The first method applied in this paper is the non-linear t-distributed stochastic neighbor embedding (t-SNE), which was introduced by van der Maaten et al. (2008). t-SNE performs dimensionality reduction by resulting in a two-dimensional space, that enables the visualization of data in clusters. The method aims at keeping the probabilities of pair distances of the data in the original space as similar as possible to the corresponding probabilities in the reduced space. The clustering is done by visualizing data that have similarities in the original space as points close to each other in the reduced space. In this paper, this method is used, firstly, to identify the number of clusters of farms for SCC and TBC. Secondly, t-SNE results are exploited to assess any possible spatial correlations among pairs of farms. The results were obtained from the `tsne` function of MATLAB using perplexity equal to 4 and the distance metric as the Euclidean.

#### 2.2.2. *Self-Organizing Map (SOM)*

The second data mining methodology exploited is self-organizing maps (SOMs), which is a type of artificial neural network (Kohonen 1982). It is an unsupervised learning method that performs dimensionality reduction by maintaining the topological and metric aspects of the input space. The high-dimensional input space is compressed in a two-dimensional space, which comprises neurons. The SOM algorithm assigns each data point to one of the neurons. Augustijn et al. (2013) explain in detail how SOMs can serve

spatio-temporal analysis. In this paper, the SOM has the same number of neurons as the number of the resulting clusters from t-SNE for SCC and TBC. The selforgmap function of MATLAB is used with two grids. The grid is chosen to be hexagonal and the SOM batch algorithm is used to train the network for 1000 epochs.

### 3. Results and Discussion

#### 3.1. Clusters based on t-SNE

The geographical coordinates act as the reference to assess the clustering of the farms as resulting from the t-SNE technique. t-SNE results are illustrated by the topology of the farms in the two-dimensional space. In the following figures, each dot represents one farm. The farms that are located close to each other for both the t-SNE reduced space and the geographical coordinates are connected by black lines. The dots representing those farms are accompanied by a label including the corresponding code of each farm to facilitate the comparison of their proximity in terms of geographical coordinates.

##### 3.1.1. Results for SCC

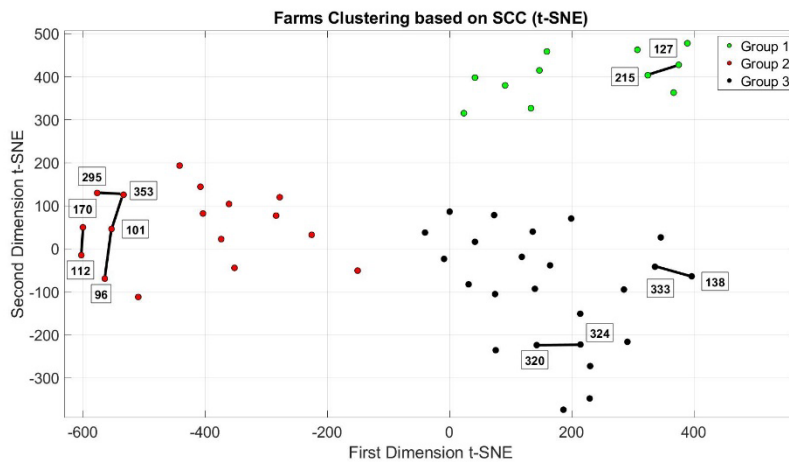


Figure 1. The t-SNE reduced space based on SCC.

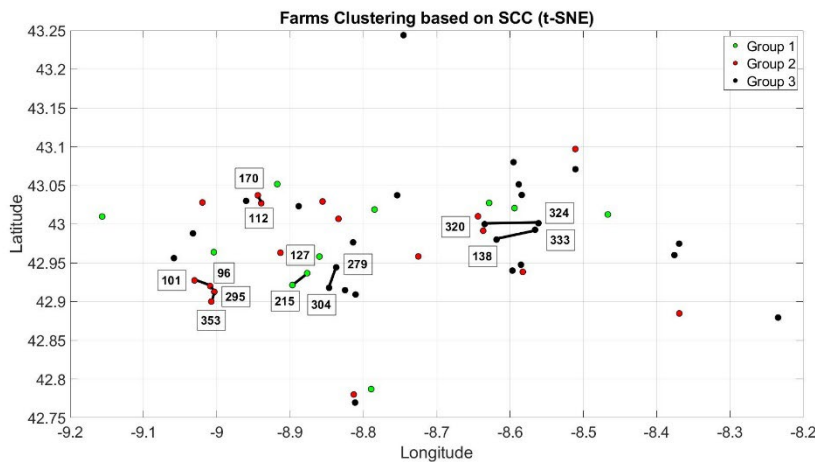


Figure 2. Clustering of farms based on SCC using t-SNE.

In the case of SCC, three clusters are identified in the reduced t-SNE space. By assessing the topology of each farm in the reduced space (Figure 1) with regard to their map coordinates (Figure 2), four pairs of farms come up, i.e., farm 127 & 215, 333 & 138, 324 & 320, and 112 & 170, as well as a cluster of four farms, i.e., farm 96, 101, 353 & 295. This indicates spatial correlation among those farms in terms of SCC temporal variability.

### 3.1.2. Results for TBC

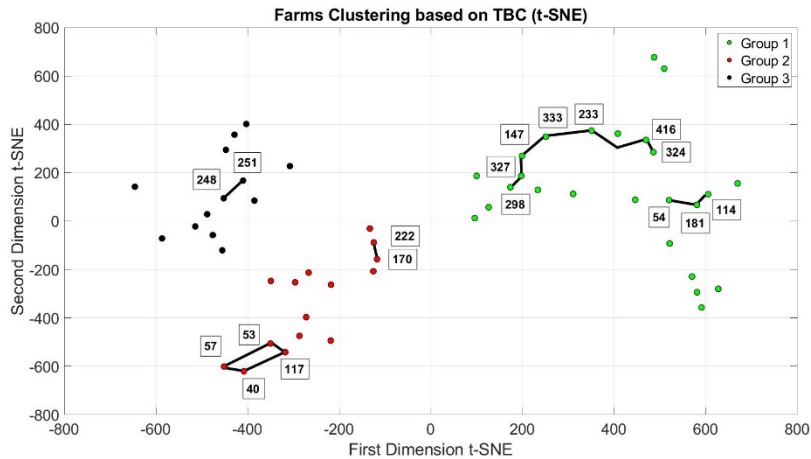


Figure 3. The t-SNE reduced space based on TBC.

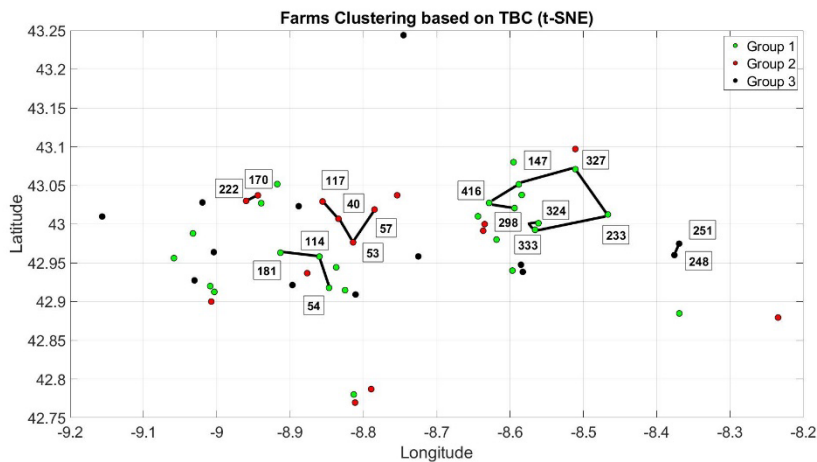


Figure 4. Clustering of farms based on TBC using t-SNE.

The t-SNE clustering based on TBC results in three clusters as well (Figure 3). Following the same rationale as in the case of SCC (Figure 3 & 4), two pairs of farms can be identified, i.e., farm 248 & 251 and 222 & 170. Furthermore, one group of three farms, i.e., farm 181, 114 & 54, along with one group of four farms, i.e., farm 40, 53, 57 & 117, are found. Lastly, the group formed by farm 298, 327, 147, 333, 233, 295, 416 & 324, are located close to each other (Figure 4), however in the t-SNE reduced space (Figure 3) they form a line instead of a group of points. This indicates spatial correlation among those farms in terms of TBC temporal variability as well.

### 3.2. Clusters based on SOM

As mentioned in the previous section, the SOM for each variable has the same number of neurons as the clusters resulting from the t-SNE analysis. In the case of SCC, three clusters of farms are formed in the t-SNE reduced space, thus the SOM is formed by three neurons which represent three groups of farms. However, for TBC, the clusters are less clear in the t-SNE reduced space. Thus, two SOM structures were assessed for TBC, one with two and one with three neurons or groups of farms. The results from the SOM with three neurons are presented here as it provides more insights. The clustering results are presented in terms of the geographical coordinates of the farms, i.e., each farm is depicted as a dot and colored based on the neuron it is assigned to after training the neural network.

#### 3.2.1. Results for SCC

Results from the SOM trained based on SCC data with three neurons indicate clustering of farms with spatial characteristics (Figure 5). Groups 1 and 3 mostly consist farms that are located in the western part of the map, while Group 2 is in the eastern one. Furthermore, all the farms with an indication of spatial correlation following the t-SNE technique (Figures 1 & 2) are clustered accordingly based on the SOM methodology.

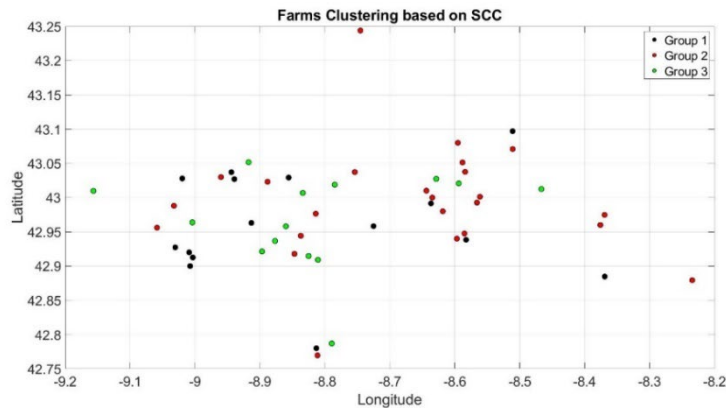


Figure 5. Clustering of farms based on SCC using SOM.

#### 3.2.2. Results for TBC

In the case of the SOM trained based on the TBC data with three neurons, the geographical distinction is less clear compared with the SOM for SCC (Figure 6). Group 1 mostly consists of western while Group 3 of eastern farms and Group 2 shows no such geographical characteristic. Nevertheless, Group 3 appears to represent farms that are located around the longitude -8.6. Interestingly, some of these farms appear close to each other in the t-SNE reduced space (Figures 3 & 4) as well, i.e., the group of farms that forms a line in Figure 3.

The agreement between the t-SNE and SOM methodologies both based on SCC and TBC for those sets of farms makes a stronger argument regarding the hypothesis of this study, i.e., whether there is a spatial correlation among dairy farms in terms of SCC and TBC. Even though the results do not show a clear geographical clustering, the indication of its existence is presented, as resulting from these two methodologies.

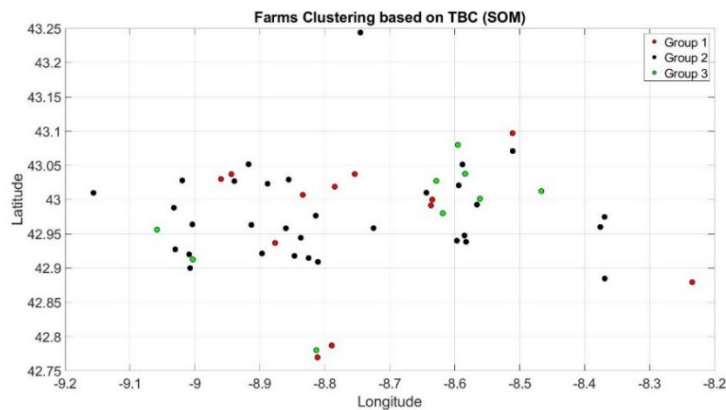


Figure 6. Clustering of farms based on TBC using SOM.

#### 4. Conclusion

Processes such as milk pasteurization, can be optimized considering climate change using climate change impact assessments. As pasteurization aims to reduce the microbial load, such an impact assessment is based on a model that describes the effect of multiple climatic factors on milk contamination, which is expressed with SCC and TBC. This study aimed at identifying spatial correlations among dairy farms in terms of SCC and TBC by applying data mining techniques. According to the results, both SCC and TBC indicate spatial correlation among dairy farms. This means that apart from the temporal dependency between SCC and TBC and climate, there is also a spatial dependency among farms. Therefore, collecting data only from one farm may be misleading. In this way, predicting food safety aspects of milk under climate change is less uncertain. Assessing the application potential of such models to different locations may be realized by evaluating if predictive models developed from Spanish data can explain the spatio-temporal characteristics of raw milk produced in another region.

#### Acknowledgement

This work was supported by the PROTECT project funded by the European Union's Horizon 2020 Research and Innovation Programme [MSCA grant 813329]. SA was supported by the Research Foundation – Flanders (FWO), under grant 1224623N.

#### References

- Q. K. Ahmad & R. A. Warrick, 2001. Methods and tools. In: J. J. McCarthy, O. F. Canziani, N. A. Leary, D. J. Dokken and K. S. White. (eds) *Climate change 2001: impacts, adaptation, and vulnerability*. Cambridge University Press, 105
- E.W. Augustijn & R. Zurita-Milla, 2013, Self-organizing maps as an approach to exploring spatiotemporal diffusion patterns, *International Journal of Health Geographics*, 12, 60
- K. M. Gunn, et al., 2019, Projected heat stress challenges and abatement opportunities for US milk production. *PLoS one*, 14,3, e0214665
- L. Katsini, et al., 2022, Quantitative methods to predict the effect of climate change on microbial food safety: A needs analysis. *Trends in Food Science & Technology*, 126, 113
- T. Kohonen, 1982. Self-organized formation of topologically correct feature maps. *Biological cybernetics*, 43,1, 59
- M. Stürmer, et al., 2018. Relationship between climatic variables and the variation in bulk tank milk composition using canonical correlation analysis. *International Journal of Biometeorology*, 62,9, 1663
- L. Van der Maarten & G. Hinton, 2008. Visualizing data using t-SNE. *Journal of Machine learning research* 9, 2579-2605.

# Ensemble Kalman Filter for estimation of intracellular nucleotide sugars from extracellular metabolites in monoclonal antibodies

Luxi Yu <sup>a</sup>, Ehecatl Antonio del Rio Chanona <sup>a</sup>, Cleo Kontoravdi <sup>a</sup>

<sup>a</sup> *Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, London, SW7 2AZ, United Kingdom*

## Abstract

The emergence of Quality by Design (QbD) and Process Analytical Technology (PAT) paradigm supported by the FDA imposes a strong motivation for digital transformation in biopharmaceutical industry. The inherent complexity of bioprocess dynamics, batch-to-batch variability resulting from raw materials and process operations, as well as the need for accelerating product manufacturing, makes dynamic soft sensors such as Kalman Filters highly desirable for process development, monitoring, and control. In this work, we develop an Ensemble Kalman Filter framework in the context of monoclonal antibody bioprocessing, where the noise on physical sensors is mitigated for extracellular metabolite states by integrating the process' dynamic mechanistic model and sensor measurements. More importantly, the framework accurately estimates the nucleotide sugar concentrations, an intracellular state of the cell that is not routinely measured in industry due to experimental complexity. The proposed EnKF soft sensor retrieves this knowledge through state inference, providing valuable insights for monitoring and control of key quality attributes such as glycan distribution.

**Keywords:** Antibody Bioprocessing; Metabolites; Nucleotide Sugar Donors; Ensemble Kalman Filter

## 1. Introduction

Glycoproteins such as monoclonal antibodies represent the largest group of biologically derived medicines with substantial market growth in the foreseeable years. Nucleotide sugar donors (NSD) are activated forms of monosaccharides that act as direct co-substrates for synthesizing glycosylated proteins, with their intracellular levels providing insights on the cell state. Real-time NSD information in bioprocesses can support decision-making on extracellular nutrient feeding towards controlling and monitoring key product quality attributes such as glycan distribution. However, NSD measurements are challenging to obtain experimentally and are therefore not routinely performed in industry. As NSDs are synthesised metabolically, several predictive models have been built to link extracellular metabolites with intracellular NSD concentrations. Whilst these models can be highly accurate in predicting NSD concentrations (Kotidis et al., 2019, Jedrzejewski et al., 2014, Sou et al., 2017, Villiger et al., 2016), they are usually system-specific and require extensive re-parameterization for changes in experimental conditions, cell line or product.

As an alternative to the system-specific predictive models, Kalman Filters (KF) employ a mechanistic process model to predict system transition from a previous to the current



state while integrating the information from hard sensor measurements. In addition to unmeasured state inference, the idea of KF is to produce better estimations of the measured states compared to either measurements or model predictions alone (Narayanan et al., 2020). Here, a Monte-Carlo based KF for non-linear systems, Ensemble Kalman Filter (EnKF), is proposed to estimate NSD concentrations using a mechanistic model and corrected by easy-to-obtain extracellular metabolite measurements. The EnKF is able to filter noise in metabolite measurements and subsequently enables a more accurate estimation of NSD concentrations. The combination of sensor observations and first principles modeling reduces the reliance on highly specific mechanistic models and makes obtaining accurate NSD knowledge possible without directly measuring them, offering great promise for applying the QbD paradigm to ensure product quality during manufacturing.

## 2. Material and Methods

### 2.1. Experiment Description

A fed-batch CHO cell culture process producing an IgG antibody was used to evaluate the performance of the EnKF-based soft sensor, where the culture was supplemented with glucose and amino acid nutrients on even days of the entire 12-day cell culture period. Daily measurements were taken for all of the metabolites and NSDs. The experimental set up and associated dataset can be found in full in Jędrzejewski (2015).

### 2.2. Mechanistic Model Description

The mechanistic model used in this study consists of two parts, the cell culture model and the NSD metabolic model. Briefly, the cell culture model describes the cell growth and death, extracellular metabolism, and antibody production, whereas the NSD metabolic model was formulated to account for the specific cell growth rate, intracellular metabolite reactions and the transport of NSDs into the Golgi. A detailed model description can be found in Kotidis *et al.* (2019).

### 2.3. Simulation of True Values

Here, it is assumed that the true value for each of the extracellular metabolites and intracellular NSDs can be simulated by adding a white Gaussian process noise to the mechanistic model. Although true values are not available in real-life scenarios, the simulated true states are only used to examine the effectiveness of the EnKF-based soft sensor and their values are not used in executing the algorithm.

### 2.4. Data Augmentation of Experimental Observations

Due to nonlinearities in the growth rate equation, the availability of additional, artificial measurements allows implementation of a more stable and robust EnKF estimation. For this purpose, we have augmented the experimental dataset with artificial measurements. The latter were generated by adding sensor noise, in the form of variance calculated from the experimental triplicate to the true state trajectories described in 2.3.

### 2.5. Ensemble Kalman Filter

The EnKF is a Monte-Carlo based KF, where an ensemble of state vectors is randomly sampled from a large cloud of states that represents a specific probability density function. The ensemble is then integrated forward in time through the true non-linear process model and updated when a new observation becomes available (Evensen, 1994). The update step in EnKF introduces additional approximations when non-Gaussianity or non-linearity are involved, where the ensemble ‘shift’ towards the new observation through linear Gaussian state-space model (Katzfuss et al., 2016). Regardless, EnKF has been successfully applied to many complex non-linear and non-

Gaussian cases, although the deviation from Gaussian behaviour becomes more important when the ensemble predictions are substantially different from the observations (Aanonsen et al., 2009).

Figure 1 presents the implementation of the EnKF in this work. The algorithm starts by sampling an ensemble of states for each metabolite and NSD, with a mean equal to the estimated initial conditions and a covariance representing the uncertainty of this first-estimate of the initial state. The number of draws for the ensemble is determined by  $2L+1$  where  $L$  is the number of states, a rule of thumb that is commonly used in Unscented Kalman Filter (UKF) (Hommels et al., 2009). Since UKF is an approximation of EnKF and this rule of thumb is used to propagate the approximate distribution on the UKF, we decided to apply the same heuristic to approximate the distribution via Monte Carlo by the EnKF. The selected EnKF ensemble is then propagated through the non-linear process model  $f$ , with a white Gaussian noise  $w$  added at every time step. Here, the effect of the error propagation through correlated states must be taken into account when designing the covariance matrix of the process noise  $w$ . This is to ensure a reliable evolution of the ensemble variance, as a sensible estimate of the model variance is of vital importance for the success of EnKF (Evensen, 1994). In practice, the design of process noise covariance matrix is a trial-and-error procedure based on which matrix yields the most accurate state estimation for that specific filter structure (Alag and Gilyard, 1990).

The mean at time step  $k$ ,  $\hat{x}_k \in \mathbb{R}^{n_x}$ , of the ensemble is then calculated following the prediction step and the state covariance,  $P_k \in \mathbb{R}^{n_x \times n_x}$ , can be estimated from the spread of the ensemble  $X_k$ . It should be noted that the state covariance  $P_k$  is not an explicitly expressed matrix for implementation of EnKF, but rather an intermediate state for calculating the Kalman gain in the update step (Houtekamer and Mitchell, 2005). In this work, the EnKF algorithm has been designed to only go through the update step if the measurement is available at that time point. Following the prediction step, the Kalman gain  $K$  in the update step can be interpreted as a ratio of how much the algorithm trusts the model prediction versus the measurement. The Kalman gain  $K$  is then used as a weighting factor to scale the residual between sensor measurement and model prediction, which ultimately output the filtered state  $x$  for the next prediction step.

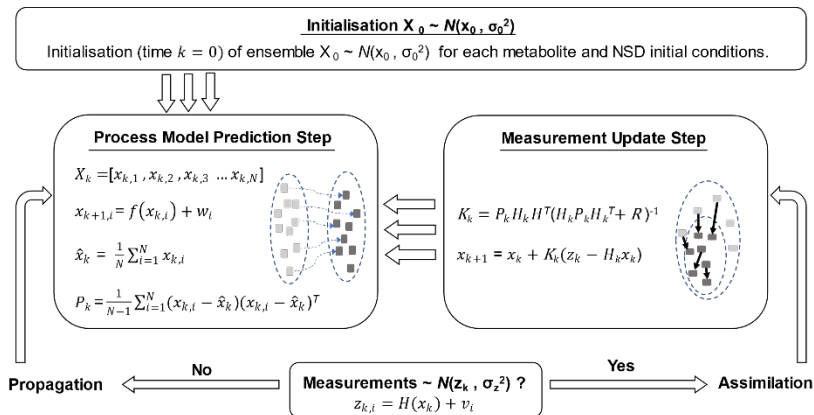


Figure 1. The Ensemble Kalman Filter algorithm, adapted from Pellenq and Boulet (2004). **Notations:**  $X$ , sampled ensemble;  $f$ , process model;  $w$ , white Gaussian noise of process;  $x$ , process states;  $P$ , state covariance;  $K$ , Kalman gain;  $H$ , measurement function;  $R$ , measurement noise covariance;  $z$ , measurement;  $v$ , white Gaussian noise of measurement

### 3. Results & Discussion

#### 3.1. Metabolites correction

As shown in Figure 2, the EnKF showcases an excellent performance on predicting the viable cell density, antibody titre, extracellular metabolites glucose and ammonia concentrations when plotted against the true trajectory of the states. As the Kalman gain  $K$  is involved in the update step as a scaling factor, the filtered results will lie in-between the measurements and the model prediction. The mechanistic model predicts all four states accurately from the start of the cell culture period until the lag phase, and then gradually deviates from the true value. This results in an overestimation for viable cell density and antibody production titre, underestimation of glucose levels, and slight underestimation and then overestimation of the ammonia levels in the exponential and stationary phase, respectively. In contrast, the EnKF estimation is closer to the true value for the entire cell culture period and converges with time propagation. This phenomenon is particularly obvious for glucose, for which the EnKF lies closer to the model prediction than the true value from Day 4 to Day 6, but gradually converges towards the end of the culture period and matches almost exactly to the true value trajectory.

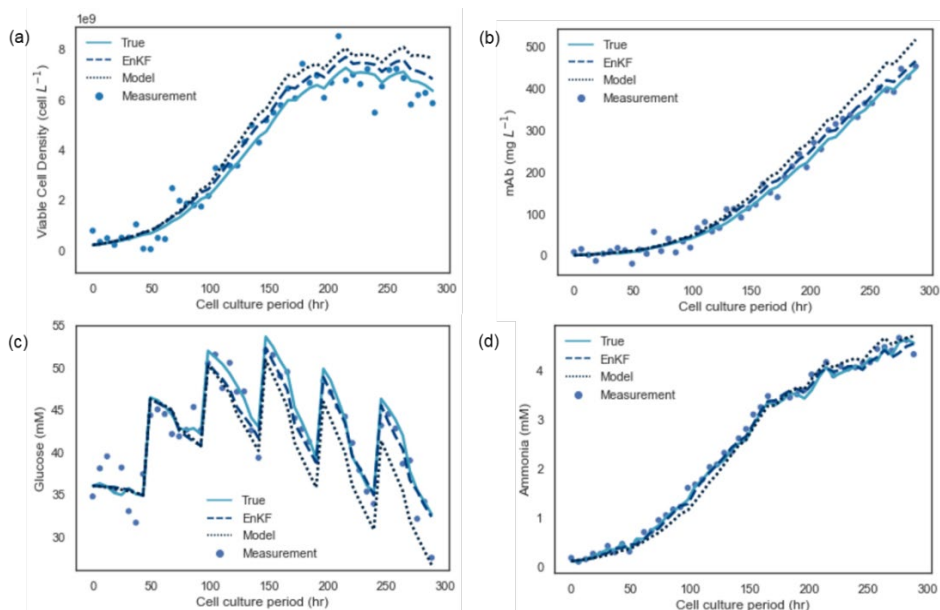


Figure 2. Comparison of true trajectory (solid line), states estimated by EnKF propagation (dashed line), mechanistic model prediction (dotted line) and noisy measurements (scatter) for (a) viable cell density  $X_v$  (b) antibody titre  $mAb$  (c) glucose concentration and (d) ammonia concentration

#### 3.2. NSD estimator

The estimation of unmeasured NSD states, UDP-Galactose (UDPGal) and UDP-glucose (UDPGlc) concentrations, is compared with their respective true value trajectories, as illustrated in Figure 3 (a) and (b). UDPGlc is the high-energy donor form of glucose, which synthesizes glycogen and other glucose-containing molecules such as dolichol P-glucose, the key substrate for initiating the glycosylation process (Varki, 2017). On the other hand, UDPGal serves as the main precursor for galactosylation, a highly desirable characteristic for enhanced antibody-dependent cellular cytotoxicity (Kotidis et al.,

2019). Therefore, real-time estimation of UDPGal levels offers valuable knowledge on developing process engineering strategies for optimizing the product glycan distribution. For example, extracellular galactose feeding has been consistently reported to increase the UDPGal pool, which then promotes galactosylation on the recombinant product (Blondeel and Aucoin, 2018). Here, although sensor measurements for UDPGal and UDPGlc concentrations are not directly available for implementation of the EnKF, the extracellular data appear to be correlated with intracellular NSD levels. NSD concentrations are therefore implicitly corrected from the available extracellular metabolite measurements through the EnKF update step and propagated forward in time with the EnKF prediction step. Overall, the EnKF outperforms the mechanistic model, especially towards the end of the cell culture period where the EnKF converges towards the true value. The computed time-resolved mean squared error (MSE) in Figure 3 (b) and (c) shows that pure mechanistic model is more prone to system instability and model inaccuracies. The MSE for both NSD concentrations is consistently lower for EnKF. For UDPGal concentration estimation, the MSE reduces past the exponential phase, while the MSE for UDPGlc concentration remains low throughout the culture period, in contrast with that for the mechanistic model, which increase rapidly in late-stage culture.

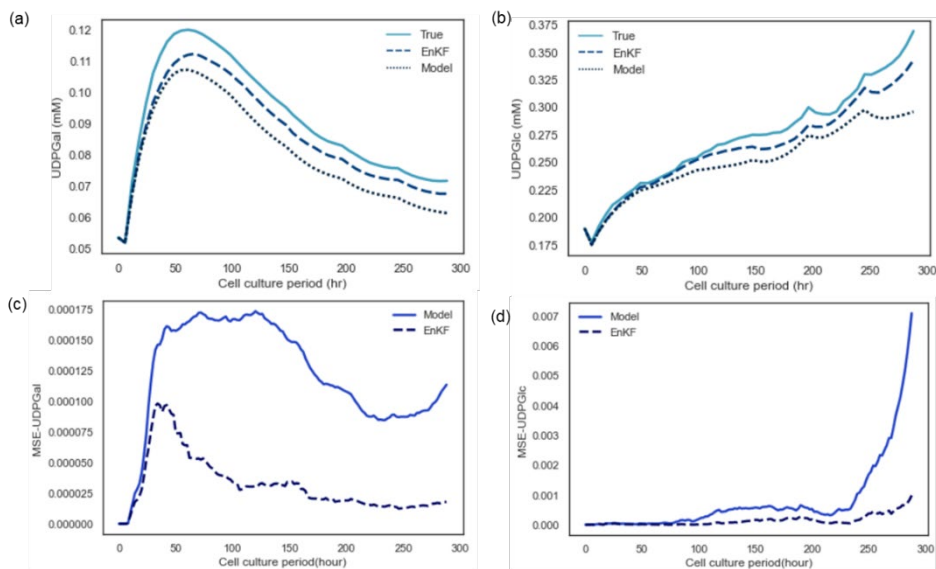


Figure 3. Comparison of true trajectory (solid line), states estimated by EnKF propagation (dashed line), mechanistic model prediction (dotted line) for (a) UDPGal concentration (b) UDPGlc concentration. Model prediction and EnKF estimation MSE computed against the true trajectory for (c) UDPGal concentration and (d) UDPGlc concentration

#### 4. Conclusion & Outlook

Herein, we present the implementation of EnKF for estimation of intracellular NSD concentrations. The proposed EnKF framework performs well with respect to reducing sensor noise and estimating unmeasured NSD levels accurately, which therefore support in-process decision-making and product quality control based on easy-to-obtain extracellular measurements. Moreover, the incorporation of time-discrete sensor

observations infuses a degree of data-driven characteristic to the mechanistic process model while retaining the system dynamic knowledge, thereby potentially providing greater flexibility in transferring the framework to other process conditions, cell line or product.

## References

- AANONSEN, S. I., NÆVDAL, G., OLIVER, D. S., REYNOLDS, A. C. & VALLÈS, B. 2009. The Ensemble Kalman Filter in Reservoir Engineering—a Review. *SPE Journal*, 14, 393-412.
- ALAG, G. S. & GILYARD, G. B. 1990. A proposed Kalman Filter Algorithm for Estimation of Unmeasured Output Variables for an F100 Turbofan Engine. *NASA Technical Memorandum 4234*.
- BLONDEEL, E. J. M. & AUCOIN, M. G. 2018. Supplementing glycosylation: A review of applying nucleotide-sugar precursors to growth medium to affect therapeutic recombinant protein glycoform distributions. *Biotechnol Adv*, 36, 1505-1523.
- EVENSEN, G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99, 10143-10162.
- HOMMELS, A., MURAKAMI, A. & NISHIMURA, S. 2009. Comparison of the Ensemble Kalman filter with the Unscented Kalman filter: application to the construction of a road embankment. *GEO International*.
- HOUTEKAMER, P. L. & MITCHELL, H. L. 2005. Ensemble Kalman filtering. *Quarterly Journal of the Royal Meteorological Society*, 131, 3269-3289.
- JEDRZEJEWSKI, P. M. 2015. A platform for the optimisation of metabolic pathways for glycosylation to achieve a narrow and targeted glycoform distribution. *PhD Thesis*, Imperial College London.
- JEDRZEJEWSKI, P. M., DEL VAL, I. J., CONSTANTINOU, A., DELL, A., HASLAM, S. M., POLIZZI, K. M. & KONTORAVDI, C. 2014. Towards controlling the glycoform: a model framework linking extracellular metabolites to antibody glycosylation. *Int J Mol Sci*, 15, 4492-522.
- KATZFUSS, M., STROUD, J. R. & WIKLE, C. K. 2016. Understanding the Ensemble Kalman Filter. *The American Statistician*, 70, 350-357.
- KOTIDIS, P., JEDRZEJEWSKI, P., SOU, S. N., SELICK, C., POLIZZI, K., DEL VAL, I. J. & KONTORAVDI, C. 2019. Model-based optimization of antibody galactosylation in CHO cell culture. *Biotechnol Bioeng*, 116, 1612-1626.
- NARAYANAN, H., LUNA, M. F., VON STOSCH, M., CRUZ BOURNAZOU, M. N., POLOTTI, G., MORBIDELLI, M., BUTTE, A. & SOKOLOV, M. 2020. Bioprocessing in the Digital Age: The Role of Process Models. *Biotechnol J*, 15, e1900172.
- PELLENQ, J. & BOULET, G. 2004. A methodology to test the pertinence of remote-sensing data assimilation into vegetation models for water and energy exchange at the land surface. *Agronomie, EDP Sciences*, 2004, 24 (4), pp.197-204.
- SOU, S. N., JEDRZEJEWSKI, P. M., LEE, K., SELICK, C., POLIZZI, K. M. & KONTORAVDI, C. 2017. Model-based investigation of intracellular processes determining antibody Fc-glycosylation under mild hypothermia. *Biotechnol Bioeng*, 114, 1570-1582.
- VARKI, A. 2017. *Essentials of glycobiology*, Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press.
- VILLIGER, T. K., SCIBONA, E., STETTLER, M., BROLY, H., MORBIDELLI, M. & SOOS, M. 2016. Controlling the time evolution of mAb N-linked glycosylation - Part II: Model-based predictions. *Biotechnology Progress*, 32, 1135-1148.

# Towards a unified multi-scale strategy for bio-manufacturing process development

Thomas Bisgaard,<sup>a</sup> Nima Nazemzadah,<sup>b</sup> Eduardo Krebs Kleingesinds,<sup>a</sup> Negin Yousefi,<sup>a</sup> Christian Beenfeldt,<sup>b</sup> Seyed Soheil Mansouri,<sup>a,\*</sup>

<sup>a</sup>*PROSYS, Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søtofts Plads, Buidling 228A, 2800 Kongens Lyngby, Denmark*

<sup>b</sup>*Knowledge Hub Zealand, Holbækvej 141 B, DK-4400, Kalundborg*

*seso@kt.dtu.dk*

## Abstract

Biomanufacturing has been increasingly receiving attention in the past few decades due to the need for sustainable production of various consumables, global energy crisis, and climate change. The biomanufacturing sector or the so called biosolution has found application in the production of various products such as fine and bulk chemical, pharmaceuticals, food and feed products, and biofuels. However, this industrial sector is facing a lot of challenges mainly due to scalability issues and lack of fundamental understanding from the process development point of view, which need to be addressed. Hence, a unified systematic multi-scale approach is required to address some of the current challenges by providing profound knowledge of the fundamental phenomena at various scales for a more efficient process scale-up. In this study, a conceptual systematic framework is proposed that incorporated the fundamental knowledge from various disciplines that potentially could tackle some of the aforementioned challenges within the biosolutions framework.

**Keywords:** multi-scale, modelling, scale-up, bioprocess, microalgae.

## 1. Introduction

Biosolutions is a multidisciplinary commercial area that targets the use of biological systems such as enzymes, microorganisms, bio-based chemicals, and bio-based materials. The ongoing “green transition” and the increasing demand for nutritious and sustainable foods for the growing global population are key drivers of biosolutions. Novel proteins satisfy both aspects and remain a field of high interest. Industries involved in the production of commodities or foods (such as biofuels and novel proteins) in large quantities suffer from low profit margin thereby making the economic /environmental feasibility very sensitive to for example (i) scale of production and thus the scale-up process, (ii) raw material prices, and (iii) energy prices. Therefore, a robust scale-up methodology and experimental infrastructure are of utmost importance. Biosolutions often lean towards large scales of production and low value raw materials (e.g., local raw materials, valorization of “waste” streams such as CO<sub>2</sub> and process water) to obtain a feasible business case. Biosolutions Zealand is an international lighthouse project that aims to develop state-of-the-art biosolution technologies with commercial focus, while strengthening its sector in Denmark with a focus on Kalundborg Symbiosis. The

symbiosis in Kalundborg, Denmark, known as one of the largest symbiosis self-sufficient networks, is comprised of various production facilities for water, energy, food, feed, and biopharmaceutical products and has been an originally-evolving environmental collaboration for more than five decades. The network aims to tackle sustainability challenges within economic, social, and environmental scopes by exploiting the potential within a sustainable collaboration framework. A multi-scale strategy can have a crucial role in developing a fundamental understanding of process development activities in such a complex framework. This could potentially lead to optimized process operation, less lab and pilot experiments, lower unnecessary product losses, and more importantly lower time required from process development to meeting market demands. As the energy crisis is rapidly growing as a global challenge, such a strategy will lead to optimum energy utilization, potentially lower global warming potential and harmful emission within a production facility or a production network such as Kalundborg Symbiosis. In this work, we sketch a framework that provides a suitable chronological methodology and toolbox to bring a potential product to the market by virgin raw materials or by upgrading waste stream(s) microorganisms. The challenges and perspectives are highlighted and discussed for each step. Finally, a case study within the Biosolutions Zealand project is presented.

## 2. Framework for Biosolutions

It is well-known that bioreactors comprise a multi-scale complexity, in the sense that the links between hydrodynamic (bulk scale), local fermentation media conditions experienced by the individual cells (heterogeneities), and the cellular reactions and networks, are far from established. Therefore, experimentation is currently considered as an inevitable yet costly and time-consuming activity in production scale-up. Ideally, scale-up can be performed successfully and robustly *in silico* (simulation) using minimum amount of experimental work. This framework (Figure 1) summarizes the promising tools for the process of taking a target product to the market (a biosolution), and incorporates the scale-up elements presented by (Wang et al. 2014; Wang et al. 2020).

### 2.1. Step 1. System Identification

An idea for a biosolution either arises from a targeted desired product or a targeted technology with the aim of identifying valuable product(s) for a given symbiotic environment. A targeted technology could be microalgae photobioreactor (as the case study presented later). Hence, in this step, all  $P$  product candidates,  $S$  material side-streams must be systematically identified. Additionally, all energy sources must be identified, however, this is only used for Step 6. Due to advances in strain optimization using either genetic engineering (i.e., genetically modified organisms, GMO's) and high throughput cultivation for natural random mutations, we recommend also to list the  $N$  potential candidate production host microorganisms. The biosolution identification problem now comprises a  $P \times S \times N$  multidimensional search space as a single-product example.

High throughput screening (HT) has become common practice in most biobased manufacturing. It is carried out in micro-liter scale at controlled conditions, with typical commercial configurations up to 96 simultaneous cultivations (e.g., 96 well-plate). The target is to screen among the many candidates using optical or spectroscopic measurement techniques to significantly reduce the search space. Parameters such as mass transfer, strain robustness towards concentration changes, etc., are commonly not

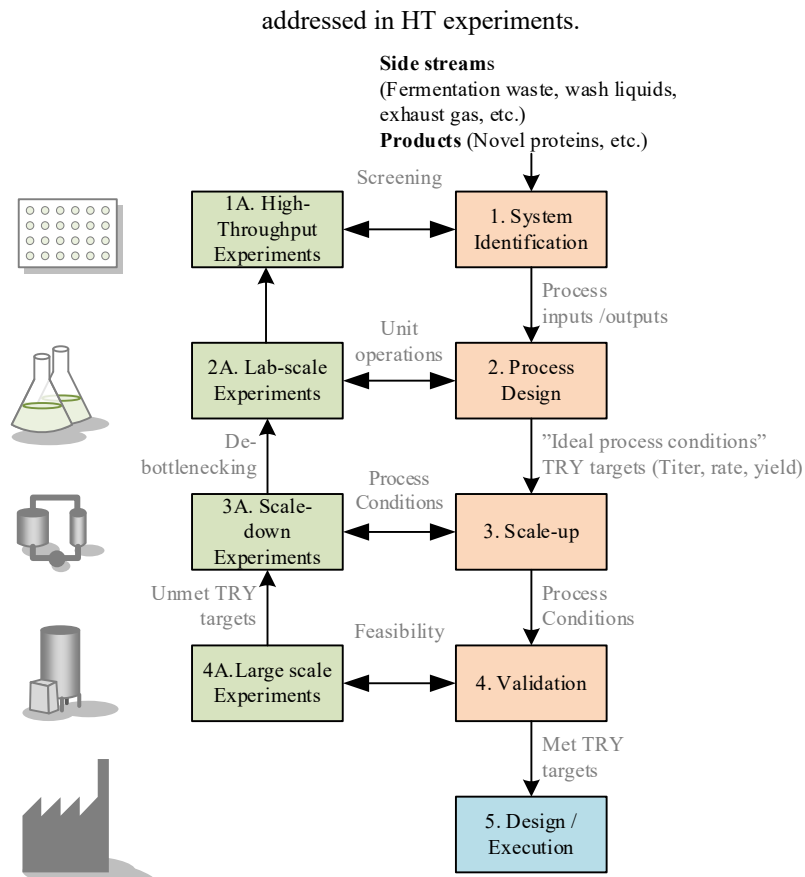


Figure 1. Biosolutions Framework.

## 2.2. Step 2. Process Design

The conceptual is a classical discipline, that aims to identify process flowsheet, mass balance, overall plant layout, techno-economic and sustainability feasibility, and a coarse price estimate. Another important aspect of the conceptual design includes assessment of performance indicators (PI's), which covers as a minimum the techno-economic feasibility and sustainability. The process design task will define targets for the process and hence the following tasks. The targets are titer (gram product per volume), rate (gram product per volume per time), yield (gram product per gram raw material), together with recovery rates at each separation step. The targets define the techno-economic feasible design space. Titer, rate, and yield are commonly known as TRY.

Laboratory experimentation may be used to identify feasible upstream and/or downstream unit operations. For upstream, the experiments could cover shake flask and bench-scale bioreactors. The aim of shake flask studies is to map the optimal ("ideal") conditions of the microorganisms. Such studies typically involve data-driven optimization such as one factor at the time or factorial design (e.g., RSM, Response surface methodology), complemented by statistical tools such as ANOVA. When the optimal conditions are identified, kinetic models must be fitted to the experimental data. Simple stoichiometric models should be developed to obtain the ideal cultivation trajectory. The conditions



determined at this step will be a proper initial estimate of the trajectory that must be further refined during the scale-up procedure using model-based scale-up techniques. Integration of Step 1 and Step 1 are possible with modern facilities, e.g., Gamble et al. 2021.

### 2.3. Step 3. Scale-up

The main task in performing scale-up is to understand the metabolic–hydrodynamic coupling during fermentation and achieve your final goal. Computational Fluid Dynamics (CFD) Euler-Lagrangian particles appear to provide a useful tool for this, in which the lifeline of single cells can be analyzed quantitatively (Lapin et al. 2004; Haringa 2022), thereby facilitating targeted experimentation, e.g., down-scale studies (Lapin et al. 2004; Noorman 2011). Hence, CFD can be coupled with a segregated description of the cell population (Pigou & Morchain 2015). Wang et al. (Wang et al. 2020) highlight the four key challenges in model development for scale-up studies:

- I. Lack of metabolic data representing the variable conditions inside bioreactors.
- II. Kinetic models do not capture the nonlinear structure of cellular kinetics.
- III. Inappropriately designed experiments for parameter identification and estimation.
- IV. No minimization of uncertainty via data reconciliation.

Metabolomics is one of the promising tools to support key challenges outlined in points I, II, and IV. It has been demonstrated that metabolic activity depends on the culture media (Pigou & Morchain 2015; Noorman 2014; Xu et al. 1999).

### 2.4. Step 4. Validation

Validation is performed to address the uncertainty associated with scale-up. Access to an experimental scale-up facility comprises a huge challenge for companies (in particular small to SME's, small and medium-sized enterprises) due to high investment cost and due to the gap of knowledge that lies across various scales in process development. To avoid unnecessary capital investment, several alternatives exist and continue to emerge, such as rental facility, e.g., mobile scale-up facility (containerized) (Fahr et al. 2022). On another note, these validation scale-up strategies should be coupled with model-based scale-up solutions to support the validation phase quite efficiently. In the biosolutions project, a combinatorial strategy using model-based scale-up and mobile facilities are explored to provide a understanding in depth of the various process steps and to avoid large investment on a fixed facility. A more generalized solution can be developed to support as many production facilities as possible that are capable of mimicking the larger scale process steps. The intention is to develop a mobile facility that is easier to adapt to a variety of processes by using a modularity concept. The modularity concept in this project lies on the different process steps such as upstream and downstream, waste handling and utility management. However, on the early-step design of the mobile unit, several considerations have been made in terms of capacity and compatibility for replacing unit operations at various steps.

### 2.5. Step 5. Design /Execution

Upon completion of conceptual design (Step 4A through 4C), the following phases focus on the translation of process to physical equipment, via specifications, etc., during basic design (BD). Following BD is the detailed design (DD) in which technical drawings,

equipment and pipe sizing are made to initiate the execution (construction, C&Q, and validation). Process design changes should be minimized at this step since degrees of freedom are lowered with project progression.

### 3. Case Study: Microalgae

Microalgae can be a source of high valued biomolecules such as proteins and lipids of interest of food and /or pharmaceutical production, while being relatively robust to the process conditions of side streams (Andrade et al. 2018). Hence, they are subject of bioremediation and valorization applications.

Five industrial side streams from a Danish company was investigated initially to illustrate the proposed methodology. At present *Chlorella vulgaris* was selected as the microorganism and photobioreactor was selected as the technology. The following aim is to valorize the side stream by selecting valuable compounds produced by microalgae after cultivation.

The *Chlorella vulgaris* was cultivated in the five different side streams at three different concentrations as a preliminary example of HT screening (diluted versions of the side stream) and under sterile and non-sterile conditions (Figure 2A). As a control experiment, cultivation was also carried out in a chemically defined media (CDM). The experiments were conducted under controlled environment (150 rpm, 25 C, 40  $\mu$ mol intensity, 12h/12h cycle).

The results suggested that *Chlorella vulgaris* presents a favorable growth pattern using one of the side-streams rather than CDM (Figure 2B) which is supposed to be tested at larger scales that allow extended process monitoring and control with a subset of conditions from the HTS.

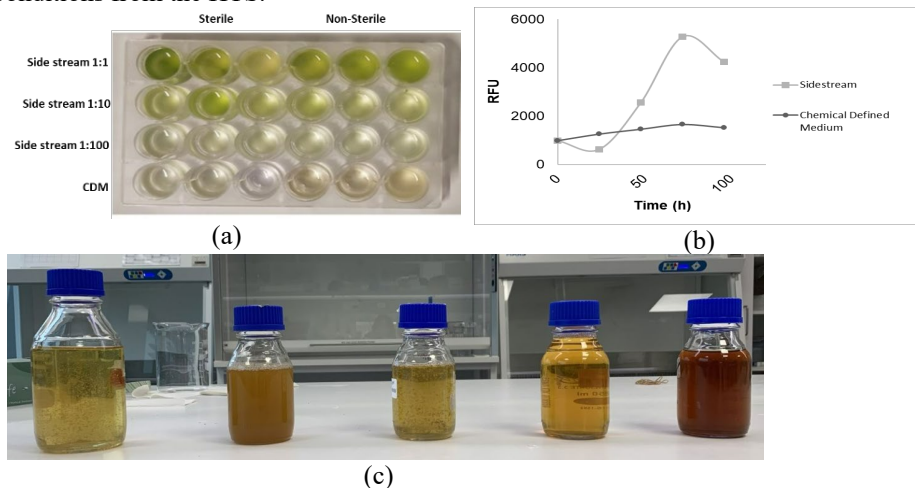


Figure 2. Step 1A HT Screening example (a) 24 well-plate of cultivation at different side streams (b) Absorbance of 24 well-plate experiment that is used to correlate the growth rate. (c) Industrial side streams.

At the time of writing, the necessary biomass is produced in shake flasks to quantify the TRY. Furthermore, the biomass will be used in supercritical CO<sub>2</sub> extraction lab-scale test to quantify the recovery. Ultimately, the goal of this microalgae study is to develop a

process that can be validated in containerized pilot scale at the company's facility. I.e., up to Step 4.

### Acknowledgement

This work is conducted in Biosolutions Zealand project under Danmarks Erhvervsfremmebestyrelse, funded by the European Regional fund.

### References

- Andrade, L. M., Andrade, C. J., Dias, M., Nascimento, C., & Mendes, M. A. (2018). Chlorella and spirulina microalgae as sources of functional foods. *Nutraceuticals, and Food Supplements*, 6(1), 45-58.
- EM-DK, Danish Ministry of Industry, Business, and Financial Affairs, <https://em.dk/ministeriet/arbejdsomraader/erhvervspolitik/biosolutions/>. Accessed November 2022.
- Fahr, S., Peña-Benavides, S.A., Thiel, L., Sengoba, C., Karacasulu, K., Ihling, N., Sosa-Hernández, J.E., Gilleskie, G., Woodley, J.M., Parra-Saldivar, R., Mansouri, S.S., Roh, K. (2022), *Industrial and Engineering Chemistry Research*, 61(35), 13191-13204.
- Gamble, C., Bryant, D., Carrier, D., Bixby, E., Dang, J., Marshall, J., Doughty, D., Colwell, L., Berndt, M., Roberts, J., Frumkin, F. From bioRxiv 28-Jan-2023, <https://doi.org/10.1101/2021.08.06.453272>, Machine Learning Optimization of Photosynthetic Microbe Cultivation and Recombinant Protein Production, 2021.
- Haringa, C. (2022). An analysis of organism lifelines in an industrial bioreactor using Lattice-Boltzmann CFD. *Engineering in Life Sciences*, 1–16.
- Lapin, A., Müller, D., & Reuss, M. (2004). Dynamic behavior of microbial populations in stirred bioreactors simulated with Euler-Lagrange methods: Traveling along the lifelines of single cells. *Industrial and Engineering Chemistry Research*, 43(16), 4647–4656.
- Noorman, H. (2011). An industrial perspective on bioreactor scale-down: What we can learn from combined large-scale bioprocess and model fluid studies. *Biotechnology Journal*, 6(8), 934–943.
- Pigou, M., & Morchain, J. (2015). Investigating the interactions between physical and biological heterogeneities in bioreactors using compartment, population balance and metabolic models. *Chemical Engineering Science*, 126, 267–282.
- Fahr, S., Peña-Benavides, S.A., Thiel, L., Sengoba, C., Karacasulu, K., Ihling, N., Sosa-Hernández, J.E., Gilleskie, G., Woodley, J.M., Parra-Saldivar, R., Mansouri, S.S., Roh, K. Mobile On Demand COVID-19 Vaccine Production Units for Developing Countries, *Industrial & Engineering Chemistry Research*, 61, 35, 2022, 13191-13204.
- Wang, G., Chu, J., Noorman, H., Xia, J., Tang, W., Zhuang, Y., & Zhang, S. (2014). Prelude to rational scale-up of penicillin production: A scale-down study. *Applied Microbiology and Biotechnology*, 98(6), 2359–2369.
- Wang, G., Haringa, C., Noorman, H., Chu, J., & Zhuang, Y. (2020, August 1). Developing a Computational Framework To Advance Bioprocess Scale-Up. *Trends in Biotechnology*, 38(8), 846–856.
- Xu, B., Jahic, M., & Enfors, S. O. (1999). Modeling of overflow metabolism in batch and fed-batch cultures of *Escherichia coli*. *Biotechnology Progress*, 15(1), 81–90.

# Development of a framework for simulation of biotechnological processes

Priscila Marques da Paz<sup>a</sup>, Caroline Satye Martins Nakama<sup>b</sup> and Galo Antonio Carrillo Le Roux<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, Polytechnic School, University of São Paulo. Av. Prof. Luicano Gualberto, trav. 3, 380, São Paulo, 05508-010, Brazil*

<sup>b</sup>*Department of Chemical Engineering, Norwegian University of Science and Technology, Høgskoleringen 1, Trondheim, 7034, Norway*

## Abstract

The development of Process Systems Engineering tools that can integrate experimental information with models is fundamental to obtain higher yields and increase productivity of a bioprocess. This is not a simple task, and there is a need for tools that facilitate the collaboration in a multidisciplinary group. In this context, developing computational frameworks for simulation and parameter estimation of bioprocesses that could be used comprehensively by researchers with different backgrounds is the goal of this work. The structure of such frameworks is carefully designed based on an ontology to describe bioprocesses. A helpful tool is the Unified Modeling Language and its use for software generalization is demonstrated along with a case study, showing that the software could serve as a model or example of good practice of software development for guiding simulations and parameter estimations of bioprocess in a structured way.

**Keywords:** bioprocesses, object-oriented programming, prototype, systems biology, unified modelling language.

## 1. Introduction

Biotechnological processes promote microorganisms' synthesizing ability to produce complex molecules that are difficult to access by traditional chemical synthesis, besides to promoting a sustainable and eco-friendly alternative. However, to make the production of a target compound economically viable, a bioprocess needs to be designed and developed to adapt an engineered microbial producer strain (Hemmerich et al., 2021). Process Systems Engineering (PSE) tools together with an experimental effort are necessary to improve the economic potential and sustainability of a new bioprocess. Besides, the required dataset size increases steadily over the years, mainly due to the technical improvements in the data measurement techniques. Often, the extensive experimental data can only be processed and interpreted through the development of quantitative methods. Combining PSE and biotechnological abilities is not a simple task, and there is a need for tools that facilitate the work and collaboration in a multidisciplinary group, which is the mark of the systems biology area (Meyer and Saez-Rodriguez, 2021). Systems biology is the scientific field that studies how the properties of biological systems emerge from the interaction of multiple components. This approach has become essential to cope with the increase in information (Bassalo and Gill, 2016; Meyer and Saez-Rodriguez, 2021). An advanced systems biology model would ideally be based on

physical laws and verified by comparing how its predictions fit experimental data. However, this is not always possible. Indeed, knowledge of the fundamental physical principles that give rise to biochemical processes is often incomplete, even for simple systems, and the available information which describes a biological system requires simplifying assumptions to allow model building. Thus, terms based first principles can be studied and incorporated into models, making them more and more complete through the years (Meyer and Saez-Rodriguez, 2021). This is the reason why systems biology is considered a multidisciplinary topic, which indicates that rarely a single person will have a deep knowledge in all fields (Zuo and Zhao, 2018).

In this context, developing computational frameworks become desirable to guide applications in bioprocess modeling. It is recommended the structure be carefully designed, since researchers with different backgrounds are able not only to use the software, but also to contribute with new models as studies in biotechnology are updated. In this way, architecture software becomes as important as developing it. Even in simple cases, it is recommended that the entire system be modeled before starting its implementation, since systems often tend to increase in size, complexity and scope. Developing the architecture is common in the computer science field, and demonstrates the software's flexibility to new editions. For the reasons given, it is interesting to bring these concepts to biotechnology to have a tool along these lines. To guarantee the (re)usability of the system, ontologies are an option to support the sharing of information structures (Guedes, 2018). Ontology is an explicit specification of a conceptualization, and typically involves classes, their relationships and axioms to describe the intended semantics (Marquardt et al., 2010). It is important that every party involved, from developers to users, reach a common understanding of the abstractions and semantics, since the software is intended to be distributed across multiple locals and researchers.

The knowledge represented by ontology allows instantiations of bioprocess components, which makes the information easy and readily shared since these components are standardized and organized. One of the tools used for the design of ontologies is the Unified Modeling Language (UML), which provides the basis for developing software architecture in an organized and generalized way and avoids inconsistencies between requirements specification documents and implementation (Arora et al., 2020; Yurin and Dorodnykh, 2020). In this work, methods and techniques which can describe bioprocesses are implemented in Julia, covering options of growth models, as well as different operation modes. The use of UML for software generalization is demonstrated with a case study, showing that the software could serve as a model or example of good practice of software development for guiding simulations and parameter estimations of bioprocess in a structured way.

## 2. Methodology

### 2.1. Software architecture: ontology for bioprocesses

Ontology has a philosophical origin, where there is interest in the conceptualization of the world of what exists and the essence of things. The basic element of the ontology is the class, which represents a collection of components that share common characteristics. It is organized by hierarchy, in which each inherited property is assigned to subclasses. Components belonging to the same class are called instances, and the characteristics or parameters of the class are the attributes. Each attribute can be identified by its name and has one or more values specific to the class it belongs to. More details can be described in the visual tool that represents ontology: UML (Marquardt et al., 2010; Zhang et al.,

2013). A flowchart that summarizes the methodology used in this paper is shown in Figure 2.1.

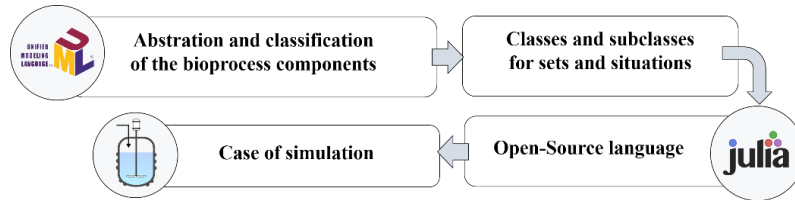


Figure 2.1. Overview of methodology for software architecture

For this work, UML diagrams were implemented using the StarUML software (version 4.5.1), and two types of diagrams were chosen: use case (Figure 3.1.b) and class diagrams (Figure 3.1.a). The first is the most general form of UML and presents an external version of the functionality that the system should offer to users, without details about how such functionality will be implemented. It presents a simple and easy-to-understand language, so that users can readily get an idea of how the system will behave. This diagram is of great help in identifying and understanding the system requirements, helping to specify, visualize and document the system features, functions and services desired by the user. It consists of two main items: actors and use cases. The actors, represented by “skinny dolls”, can be the users, other systems or even some special hardware that will use the system in some way. The functionalities that the system will make available to the actors are known as use cases, which is represented by ellipses containing a succinct text describing the action. Interactions between them are represented by lines (Guedes, 2018). The second diagram (class diagram) includes classes, interfaces and relationships. Classes are the model for creating objects, providing specification of attributes and operations that an instance of the class can complete. These operations are treated as functions or procedures, i.e., how actions will happen. Each class can implement more than one interface and each interface can be implemented by several different classes. Relationships have a similar concept to that in the use case diagram, but with more details and more types of relationship (Guedes, 2018).

## 2.2. Bioprocess modelling: starting equations

Bioprocess are represented by the class diagram, abstraction and classification of its components are done based on equations in the most generic way possible, which allows additions of more complex concepts related to the component to be made, either in the class itself or through interactions with other created ones. The first class created was *OperationMode* and the abstract characteristics for it came from the general mass balance (Doran, 2013):

$$\frac{dM}{dt} = M_i - M_o + R_G - R_C \quad (2.1)$$

where the derivative  $dM/dt$  is the mass accumulated within the system and it is zero if the system is in steady-state,  $M_i$  is the mass entering through system boundaries, i.e., the mass flow rate of a component entering the reactor,  $M_o$  is the mass flow rate leaving the system,  $R_G$  is the mass rate of generation of a component by reaction, and  $R_C$  is the mass rate of its consumption. Inherited classes are created for: continuous mode, which considers the entire equation, batch and fed-batch mode with same characteristics, but differing by the first having  $M_i$  and  $M_o$  equal to zero, while in the second only  $M_o$  is zero.

Considering that, when a bioprocess is carried out by microorganisms,  $R_G$  is given by:

$$R_G = \mu_X X \quad (2.2)$$

where  $\mu_X$  is the specific growth rate. A similar situation can be found for  $R_C$  with a specific death constant.

$$R_C = k_d X \quad (2.3)$$

Continuing with the creation of classes, it becomes noticeable that a class for microbial growth and death should be created. Cell growth and product formation reflect general kinetics and stoichiometry of intracellular reactions. So, the new class *CellGrowth* will interact with the  $R_G$  attribute of the *OperationMode* class, as *CellDeath* with  $R_C$ . *CellGrowth* will enable that different types of growth can be chosen for the system, since rates like  $\mu_X$  can be described in several ways. One of the most commonly used method is relating  $\mu_X$  to substrate  $S$ , taking the form of saturation kinetics. For example, when it is assumed that a single substrate  $S$  is limiting for  $\mu_X$ , while changes in other nutrient concentrations have no effect, this can be described by the Monod equation or alternative ones, such as Blackman, Tessier, Moseer and Contois equations. On the other hand, when substrate or product concentrations are high and inhibitory substances are present in the medium,  $\mu_X$  depends on the inhibitor concentration. Then, these situations can be described by others equations. A similar situation occurs with the *CellDeath* class, which must consider the mechanisms for decreasing cell mass, for example, when temperature changes,  $K_d$  can vary according to the Arrhenius equation (Shuler et al., 2002).

Growth kinetics has stoichiometrically related parameters, such as cell maintenance and yield coefficients, and it is also influenced by environmental conditions. There is a set of factors to be considered and consecutively, new classes arise and interactions with other classes can be performed. For example, in aerobic processes, oxygen becomes a limiting factor for growing and therefore, the oxygen transfer (OTR) from the gas phase to the liquid phase and the oxygen consumption rate (OUR) should be considered as subclasses belonging to a new *Oxygen* class, which can describe the model's dissolved oxygen. Thus, the architecture provides situations and paths for assembling models (Shuler et al., 2002).

### 2.3. Case study for bioprocess

The intended use of the system and the mentioned classes, both projected so far in the UML diagrams, were used to guide the software programming. The language chosen for the implementation of the code was Julia by its high-performance, flexibly dynamic and open source. A case study to evaluate its operation was carried out with the properties: data simulation of biomass (X), substrate (S) and product (P) in batch, with cell growing by Monod equation. The parameters are  $\mu_{max} = 0.3 \text{ h}^{-1}$ ,  $K_S = 0.2 \text{ g/l}$ ,  $Y_{XS} = 0.06 \text{ g/g}$ ,  $Y_{PS} = 0.46 \text{ g/g}$  and initial values of  $X_0 = 0.06 \text{ g}$ ,  $S_0 = 9.54 \text{ g}$  and  $P_0 = 0.00 \text{ g}$  (Doran, 2013).

## 3. Results and discussion

### 3.1. Software architecture: UML diagrams

A prototype representing the system and some definitions are shown in a section of the class diagram in Figure 3.1a, and the use case diagram, which describes an overview of actions, is shown in Figure 3.1b. In Figure 3.1.a, the bioreactor mass balance is represented by the *OperationMode* class, which is divided according to the data they will receive from Equation 2.1. Each operating mode is a subclass, which inherits mass balance characteristics and are called by the *Continuous*, *BatchMode* and *FedBatchMode* subclasses.

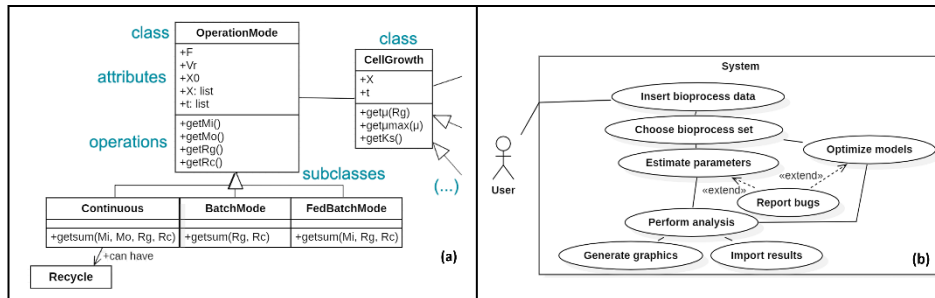


Figure 3.1. UML diagrams of bioprocess system: (a) class and (b) use case diagrams.

The equations for each of these classes depend on mass that enters or leaves the system. Within the *Continuous* subclass, recycle can be also considered. Other classes that have been generated are *CellGrowth* and *CellDeath* and may also be present in operating modes. *CellGrowth* subclasses were determined in a similar way as seen in *OperationMode* and include the different limitations that can occur for cell growth. These subclasses are represented by three dots to also indicate that the class diagram is dynamic, as the software architecture is, and can grow depending on what is added to the model.

Figure 3.1.b shows the cascade of actions, from the insertion of bioprocess data in the system to the generation and import of results. The idea is that the users indicate what they would like to do in the system according to the data they have: simulate or estimate parameters, or even optimizing them. For simulation they should indicate the operating mode and the type of growth, plus the kinetic parameters necessary; for parameter estimation, bioprocess data should be provided; and for optimization, a complete model and objective function should be defined, choosing which variables will be free and what will be optimized, according to the purpose of the model.

During the development of this structure, it was noticed that the abstraction of concepts involving bioprocess must be done with care, in order to generate several options, and it should allow for increasing complexity and flexibility of the analysis.

### 3.2. Simulation modelling

In the case study, a batch bioreactor is simulated and shown in Figure 3.2.  $X$ ,  $S$  and  $P$  data were generated considering total time of 100 h.

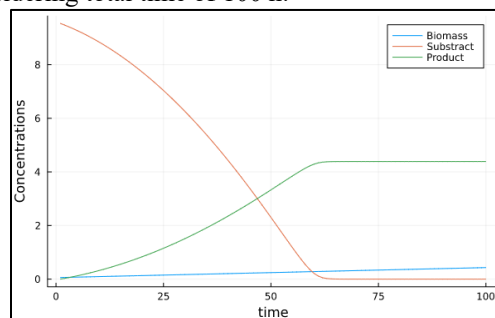


Figure 3.2. Data simulation for batch bioprocess.

The equations used to generate this simulation data can be recycled to perform, in other contexts, parameter estimation and optimization. In addition, the system can grow in number of equations to be considered in the model, which could make the differentiating of equations more laborious. To cover these issues without compromising the software's



functionality, the code implemented in Julia was adapted to the JuMP language in advance. JuMP is a domain-specific algebraic modeling language embedded in Julia and has a syntax that mimics natural mathematical expressions (Dunning et al, 2017). Because it only comprises algebraic equations, the differential equations (ODE) were discretized using Implicit Euler, which is a first-order orthogonal collocation method. Higher orders of this method can be implemented to deal with stiff problems (Biegler, 2010), which is interesting to be considered in this work to cover more complex bioprocesses that may involve stiff ODE systems. Although JuMP is mainly used to perform optimization, it can be adapted for simulation, as it happened here. The premeditated use of it will positively impact the software's development to deal with more complex problems and also involving optimization cases.

#### 4. Conclusions

This prototype can bring new perspectives for bioprocesses modelling, as currently the software community has given great attention to architectures as a key abstraction in design process. The use of UML for software generalization is demonstrated with a case study, showing that the software could serve as a model or example of good practice of software development for guiding simulations and parameter estimations of bioprocess in a structured way. In summary, this prototype is an alternative tool for model-based integration and analysis of data, which can be the seed of a new, robust and sophisticated tool for systems biology.

#### References

- V. Arora et al., 2020, Orientation-based ant colony algorithm for synthesizing the test scenarios in uml activity diagram, *Information and Software Technology*, v. 123, p. 1–21.
- M. Bassalo and R. Gill, 2016, Directed evolution and synthetic biology applications to microbial Systems, *Current Opinion in Biotechnology*, v. 39, p. 126–133.
- L. T. Biegler 2010, *Nonlinear programming: concepts, algorithms, and applications to chemical processes*. MOS-SIAM Series on Optimization.
- P. Doran, 2013, *Bioprocess engineering principles: Second edition*, Elsevier.
- I. Dunning et al., 2017, JuMP: A modeling language for mathematical optimization. *SIAM Review*, v. 59, p. 295-320.
- G. Guedes, 2018, *UML 2 - Uma abordagem prática*, Novatec.
- J. Hemmerich et al., 2021, pyfoomb: Python framework for object oriented modeling of bioprocesses, *Engineering in Life Sciences*, v. 21, p. 242-257.
- W. Marquardt et al., 2010, *OntoCAPE*, Springer Berlin Heidelberg, ISBN 978-3-642-04654-4.
- P. Meyer and J. Saez-Rodriguez, 2021, Advances in systems biology modeling: 10 years of crowdsourcing dream challenges, *Cell Systems*, v. 12, p. 636–653.
- M. Shuler and F. Kargi, 2002, *Bioprocess Engineering: Basic Concepts*, Pearson, second edition.
- A. Yurin and N. Dorodnykh, 2020, Personal knowledge base designer: Software for expert systems prototyping, *SoftwareX*, v. 11, p. 1–6.
- J. Zhang et al., 2013, A logic-reasoning based system to harness bioprocess experimental data and knowledge for design, *Biochemical Engineering Journal*, v. 74, p. 127–135, ISSN 1369703X.
- Z. Zuo and K. Zhao, 2018, The more multidisciplinary the better? - the prevalence and interdisciplinarity of research collaborations in multidisciplinary institutions, *Journal of Informetrics*, v. 12, p. 736–756.

#### Acknowledgements

This research was financially supported by the Coordination of Superior Level Staff Improvement (CAPES) grant no. 88887.464619/2019-00, PROEX program and National Council for Scientific and Technological Development (CNPq) for Priscila M. da Paz scholarship (142147/2019-2).

# Knowledge modelling framework for per-and poly-fluoroalkyl substances (PFAS) treatment solutions

Madeleine Bussemaker<sup>a\*</sup>, Nikos Trokanas<sup>b</sup>, Ioannis Kavakiotis<sup>b</sup> Franjo Cecelja<sup>a</sup>

<sup>a</sup>*University of Surrey, Guildford, UK*

<sup>b</sup>*Toolboks, Athens, Greece*

[m.bussemaker@surrey.ac.uk](mailto:m.bussemaker@surrey.ac.uk)

## Abstract

Remediation of per-and poly-fluoroalkyl substances (PFAS) is challenged with complexities of solutions, recalcitrance of end products and stringent, evolving regulations. Grouping, characterization and classification of PFAS compounds, environmental contaminations and treatment technologies through knowledge modelling has potential to overcome these challenges. Treatment technologies are often required to work in sequences, called treatment trains to achieve complete removal of PFAS from the environment i.e. a removal/separation stage followed by a degradation stage. Here, an ontology framework is presented to classify PFAS compounds and treatment technologies. Potential applications for the knowledge model to support decision making in environmental remediation and technology research and development is discussed.

**Keywords:** Per-and poly-fluoroalkyl substances (PFAS); remediation; ontology; knowledge modelling; treatment trains

## 1. Introduction

Per- and poly-fluorinated alkyl substances (PFAS) are a class of ~10,000 persistent chemicals, used for >50 years to produce, surfactants, surface treatments, aqueous fire-fighting foams (AFFFs), and plastics such as polytetrafluoroethylene (PTFE) (Buck 2021). Complete remediation of PFAS is challenged by their recalcitrance and bioaccumulation in flora and fauna (Ahrens 2011). PFAS are released into the environment from landfill, manufacturing plants, ground contaminations and sewage treatment plants. Common PFAS are now restricted under the Stockholm Convention on Persistent Organic Pollutants. However, as PFAS are banned, replacement PFAS compounds (next-gen PFAS, e.g. GenX) are introduced with unknown, toxicity and persistence. PFAS are estimated to cost the EU €52-€84 bn per year in health-related issues (Goldenman 2019). PFAS degradation technologies are often limited by i) incomplete mineralisation leading to smaller PFAS fragments; ii) select efficacy, unable to degrade all PFAS; iii) toxic by-products; iv) high energy and/or long treatment times; v) impact of the matrix composition; and vi) integration into end-to-end treatment trains.

Ontology engineering is proposed to support implementation of treatment regimens for PFAS remediation. Approaches to treatment design (e.g. superstructure optimization in wastewater treatment plants) are challenged by the extensive, up-to-date domain-specific knowledge required. For example, variable contamination characteristics, emerging technologies, developing policies and new PFAS types. Ontologies have been used to address a wide range of knowledge-intensive tasks, from linked data (Samwald et al., 2011) to molecule modelling for drug discovery (Lin et al. 2017) and to industrial symbiosis, matching of waste streams to waste processing technologies (Trokanas 2014) with the aim to create processing paths. Ontology engineering can enable formalization of data and knowledge representation of complex data sets from chemical elements to

biological and genomic data, to enable the identification of missing relationships (or steps), support classification and clustering models, as well as graph learning approaches. However, within the domain of environmental remediation knowledge modelling is limited with some examples in management of resources (Goel et al 2013), planning (Safavi 2015), and water flow modelling (Chau 2007). The objective of this work is to present a classification framework that combines the use of knowledge modelling for representation of complex chemistries and PFAS treatment options. The framework is a first step towards a decision support tool for PFAS remediation.

## 2. Framework Development

### 2.1. PFAS Classification

PFAS are defined as any compound with a fully fluorinated carbon, i.e.  $-CF_2-$  ethyl group or  $-CF_3$  methyl group and can be classed as non-polymer and polymer substances. In the environment, non-polymer PFAS are more common with detailed information on nomenclature available in literature (e.g. Buck 2011). In 2018 The OECD identified 4,730 PFAS substances (OECD, 2018), however fewer than 6% are reported to be globally relevant (Buck 2021). For the development of the knowledge model, established chemical classifications are used to manually group classes and subclasses, (e.g. Table 1). PFAS compounds may belong to two or more classes, e.g., a long chain perfluoroalkyl acid (PFAA) may be a carboxylic or sulfonic acid. Chemical classifications allow integration with existing Chemical ontologies, e.g. Chemical Entities of Biological Interest (ChEBI).

Table 1. Example classes of PFAS compounds

Perfluoroalkyl Substances	Subclass	Chemical structure
<b>Perfluoroalkyl acids (PFAAs)</b> $C_nF_{2n+1}-R$ (where R is a hydrophilic functional group)	Perfluoroalkyl carboxylic acids (PFCAs)	$C_nF_{2n+1}COOH$
	Perfluoroalkyl sulfonic acids (PFSAs)	$C_nF_{2n+1}SO_3H$
	Short chain (SC) PFAA	$C_nF_{2n+1}-R \quad n \leq 5$
	Long chain (LC) PFAA	$C_nF_{2n+1}-R \quad n \geq 6$
Polyfluoroalkyl Substances		
<b>Fluorotelomer (FT) substances</b> Used or made from the fluorotelomer process	(n:2) Fluorotelomer alcohols ((n:2) FTOHs)	$C_nF_{2n+1}CH_2CH_2OH$
	(n:2) Fluorotelomer sulfonic acids ((n:2) FTSAs)	$C_nF_{2n+1}CH_2CH_2SO_3H$
<b>Perfluoroalkyl ether acids (PFEAs)</b> Perfluoroalkyl moieties connected by ether bonds	Perfluoroether carboxylic acids (PFECAs)	e.g. GenX or HFPO-DA with the same protonated form: $C_3F_7OCF(CF_3)-COOH$
	Perfluoroether sulfonic acids (PFESAs)	e.g. n:2 Cl-PFESA: $Cl(CF_2)_nO(CF_2)_2SO_3H$

Chemical classifications also create classes that underpin treatment efficacy, manufacturing methods, persistence, and degradation products. Initial manufacturing of PFAS was based on the PFAAs perfluorooctane sulfonic acid (PFOS) and perfluorooctanoic acid (PFOA) (where  $n=7$  and  $n=8$ , respectively) known collectively as PFOX. Now almost all bodies of water contain at least 1 ng/L of PFOX (Ahrens 2014), and most treatment and toxicology research is based on these two compounds. However, fluorotelomers (FTs) and perfluoroalkylether acids (PFEAs) have replaced PFOS and PFOA, leading to new profiles of persistent PFAS in groundwater (Backe 2013). FTs are

based on shorter perfluorinated carbon chains (usually C4 or C6), linked to a hydrophilic end group via what is usually a two-carbon alkane chain. The FT compound class leads to formation of shorter chain PFAAs, or fluorotelomer sulfonates, depending on the initial structure. PFEAs are manufactured from short perfluoroalkyl moieties connected by ether bonds and have worldwide persistence (Pan 2018). PFEAs are resistance to hydroxyl radical mediated oxidation, photolysis and biodegradation, and are highly mobile, hence have similar remediation issues to their predecessor compounds.

## 2.2. Classification of treatment technologies

Treatment technologies initially focus on treatments for water contaminations. In general treatments can be classed into two types; i) separation/concentration treatment where low concentration (<ng/L) waste streams can be treated, producing a clean water that meets regulatory standards, and ii) degradation treatments where the main aim is to degrade the PFAS compounds into smaller PFAS fragments and/or to complete mineralization. The difficulty in achieving complete mineralization of PFAS comes from their extremely strong C–F bond and the shielding of carbons by fluorine components of the molecule (Kirsch 2004). Hence developing appropriate treatment trains that also mineralize PFAS are one of the key challenges for PFAS remediation. Numerous reviews exist into PFAS treatment technologies (e.g. Ross et al 2018; Horst et al, 2018; Kucharzyk et al 2017) here we aim to discuss the nuances of different regimes to underpin the ontology framework.

### 2.2.1. Separation / concentration treatments

Separation / concentration methods are grouped into granular/powder activated carbons (GAC/PAC), resins and silicas, membrane separation and foam fractionation (Table 2). Treatment properties can be attributed to each treatment class, as defined in Table 2 (Rahman 2014). Within each class, data properties will be further identified to capture the nuances of the treatment regimen and subclasses identified.

Table 2. Separation / concentration treatments, with relevant properties. NR = not reported.

TREATMENT	Efficacy for different PFAS			Matrix effects	Waste/side streams
	PFEA	SC	LC		
<b>GAC / PAC</b>	None to moderate	Moderate /ineffective	Moderate (not all precursors)	Organics compete	Contaminated solid (PAC), solid for regeneration (GAC)
<b>Resins and silicas</b>	NR	IX is less efficient	Yes	Depends on absorbent	Regenerate solutions
<b>Membrane separation</b>	NR	Yes	Yes	Impacted by organics	Membrane rejectate, spent membranes
<b>Foam-fractionation</b>	NR	Yes	Yes	Unlikely to interfere	Concentrated PFAS solution (ppm range), sedimentation

### 2.2.2. Degradation treatment technologies

Degradation mechanisms can be thermal, oxidative, biological, physical and reductive. In each case the mode of operation / technology used varies, along with the efficacy and matrix effects. For example, advanced oxidative processes (AOPs) and biological treatments are not effective at complete PFAS mineralization (Ross et al, 2018) and leave shorter chain PFAAs. Incineration, a thermal treatment can convert ~20-45% of the fluorine to toxic hydrogen fluoride (HF) gas (Aleksandrov et al 2019; Taylor et al 2014) and can emit gaseous fluorocarbons and unconverted PFAS with insufficient incinerator temperatures or residence times (Meng et al, 2017). Electrochemical reduction can be non-selective and reduction of matrix elements must be considered (Horst et al 2018).

Classifications must allow for technologies that support multiple mechanisms. For example, reductive methods include metal-mediated reduction (e.g. catalysis in

supercritical water, zero valent metals) (Kucharzyk et al 2017) and methods that promote the generation of the hydrated electron (photolytic, electrolytic and plasma) (Ross et al 2018). However electrolytic and photolytic technologies can also use oxidative degradation. Similarly ultrasonic technologies (i.e. sonolysis) are considered to be oxidative at low frequency, and argued to be pyrolytic (thermal) or reductive via the aqueous electron at high frequencies (Sidnell et al 2022). The ontology structure is therefore developed to allow such classifications (Figure 1).

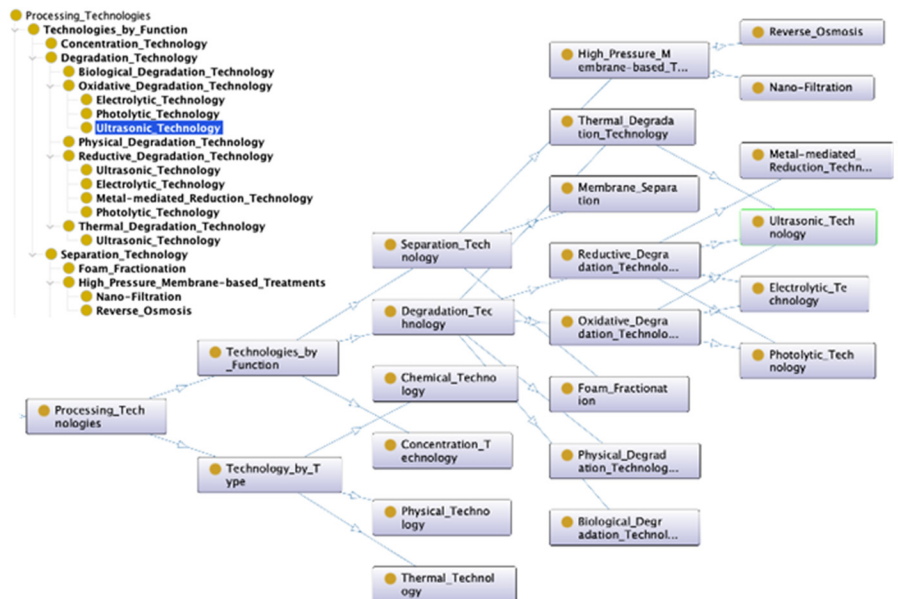


Figure 1. The PFAS Treatment technology ontology structure for a selection of treatment types.

### 2.2.3. Proposed ontology structure

Identified classes and object properties for technology classification (Table 3) are designed to capture the necessary considerations for overall remediation efficacy. The characteristics of function and type are chosen to allow for later development of knowledge graphs and inferencing around novel treatment technologies and/or new PFAS compounds, i.e. to handle missing information. Each technology may have more than one output stream, then each output stream can be classed according to environmental release laws, drinking water standards and potential further treatment options.

Within each treatment technology further object properties must be assigned to avoid sweeping generalizations around a class of technologies, often seen in literature. For example, ultrasonic degradation of PFAS can be highly dependent on frequency, additives, and reactor geometry. At low frequencies (<100 kHz) very little degradation is seen without additives, whereas at high frequency (100-1000 kHz) complete mineralization has been reported at ambient temperatures (Sidnell et al, 2022). Similarly electrolytic methods will depend on the electrolyte, electrode material and reactor configuration (Ryan et al 2021). Approaching the comparison and selection of technologies through knowledge modelling therefore highlights such nuances and enables better-informed decision making.

Table 3 Classes and object properties underpinning the PFAS treatment technology modelling

Classes	Examples
Function	Separation, concentration, degradation
Type	Thermal, physical, biological, chemical
By mechanism	Oxidation, reduction, membrane separation, ion exchange etc
Object properties	Examples
Output stream	F, CO <sub>2</sub> , SC PFAS, LC PFAS, PFAAs, co-contaminants...
Input composition	PFAS concentration, co-contaminants
Efficiency	Fluoride release rate, PFAS degradation / removal rate
Common parameters	Temperature, volume, pressure
Developmental stage	Laboratory, pilot scale, industrial scale
Cost	Capital, operational, energy consumption

### 3. Future applications and development of PFAS ontology

The classification framework for the PFAS ontology has potential application in supporting selection of treatment trains, provide predictions about technology efficacy, and focusing research and development of treatment technologies. Through modelling the inputs and outputs separation / concentration technologies can be matched with appropriate degradation technologies using matching algorithms (Trokanas 2014). This is especially useful since degradation technologies tend to be expensive at treating dilute solutions and more efficient at higher concentrations. For example, at concentrations in the mg/L range high frequency sonolysis tends to be zero order, but at lower concentrations is first order. Hence the zero-order regime is more energy efficient per gram of PFAS destroyed. Technology matching will show potential treatment trains given a composition profile of a PFAS contamination and desired PFAS output profile. Thus, has potential to inform economically and environmentally optimal treatment regimens.

One of the key challenges of PFAS remediation is the complete mineralization of all PFAS types. Due to the difficulty in destroying PFAS, many of the technologies applied are in various stages of development, with the majority of research on PFOS and PFOA degradation. Effective technologies often come at high energy and capital cost and more research and development is required to understand efficacies for new and emerging PFAS in a range of matrices. The nuances of each emerging technology are often only fully appreciated by subject-matter experts and there is danger in over-generalization of different treatment regimes. The ontology will inform knowledge graph development to facilitate machine learning to highlight necessary experiments and/or likely outcomes of various treatment regimes. If a known treatment has overlapping characteristics with an unknown treatment technology then the knowledge graph can be used to infer and predict the outcomes of the unknown treatment technology.

### 4. Conclusions

Here we presented a framework for developing an ontology around PFAS remediation. The ontology will be able to be integrated with existing knowledge models and is useful for creating a common language to model PFAS treatment regimes, support decision making on PFAS treatment trains and underpin research, development and learning around emerging PFAS degradation technologies.

### 5. References

- L. Ahrens. 2011, Polyfluoroalkyl compounds in the aquatic environment: A review of their occurrence and fate. *J Environ Monit.* 13(1):20-31.
- L. Ahrens, Bundschuh M., 2014 Fate and effects of poly- and perfluoroalkyl substances in the aquatic environment: A review. *Environ Toxicol Chem.*,33(9):1921-1929

- K. Aleksandrov, Gehrmann HJ, Hauser M, Mätzing, H., Pigeon, D., Stapf, D., & Wexler, M Waste incineration of polytetrafluoroethylene (PTFE) to evaluate potential formation of per- and poly-fluorinated alkyl substances (PFAS) in flue gas. *Chemosphere*. 2019;226:898-906.
- W. J. Backe, Day TC, Field JA., 2013, Zwitterionic, cationic, and anionic fluorinated chemicals in aqueous film forming foam formulations and groundwater from U.S. military bases by nonaqueous large-volume injection HPLC-MS/MS. *Environ Sci Technol*. 47(10):5226-34.
- R. Buck, Korzeniowski, S. H., Laganis, E., & Adamsky, F., 2021. Identification and classification of commercially relevant per-and poly-fluoroalkyl substances (PFAS). *Integrated environmental assessment and management*, 17(5), 1045-1055.
- K. W. Chau, 2007. An ontology-based knowledge management system for flow and water quality modeling. *Advances in Engineering Software*, 38(3), 172-181.
- D. Goel, Chaudhury, S., & Ghosh, H. 2017. Smart water management: An ontology-driven context-aware IoT application. In *International Conference on Pattern Recognition and Machine Intelligence* (pp. 639-646). Springer, Cham.
- J. Horst, McDonough J, Ross I, Dickson, M., Miles, J., Hurst, J., & Storch, P., 2018 Water treatment technologies for PFAS: The next generation. *Groundw Monit Remediat*. 38(2):13-23.
- P. Kirsch, 2004, *Modern Fluoroorganic Chemistry, Synthesis, Reactivity, Applications*. Weinheim: Wiley-VCH;
- K. H. Kucharzyk, Darlington R, Benotti M, Deeb R, Hawley E., 2017, Novel treatment technologies for PFAS compounds: A critical review. *J Environ Manage*. 204:757-764.
- Y. Lin, S. Mehta, H. Küçük-McGinty, H., Turner, J.P., Vidovic, D., Forlin, M., Koleti, A., Nguyen, D.T., Jensen, L.J., Guha, R. and Mathias, S.L., 2017 Drug target ontology to classify and integrate drug discovery data. *J Biomed Semant* 8, 50
- J. Meng, Lu Y, Wang T, Wang, P., Giesy, J. P., Sweetman, A. J., & Li, Q., 2017, Life cycle analysis of perfluorooctanoic acid (PFOA) and its salts in China. *Environ Sci Pollut Res*. 24(12):11254-64.
- OECD. *Toward a New Comprehensive Global Database of PAFs*. Vol 2018.; 2018
- Y. Pan Y, Zhang H, Cui Q, Sheng, N., Yeung, L.W., Sun, Y., Guo, Y. and Dai, J. 2018, Worldwide distribution of novel perfluoroether carboxylic and sulfonic acids in surface water. *Environ Sci Technol*. 52(14):7621-7629.
- M. Rahman, Peldszus S, Anderson WB., 2014, Behaviour and fate of perfluoroalkyl and polyfluoroalkyl substances (PFASs) in drinking water treatment: A review. *Water Res*. 50:318-340.
- I. Ross, McDonough J, Miles J, Storch, P., Thelakkat Kochunarayanan, P., Kalve, E., Hurst, J., S. Dasgupta, S., Burdick, J., 2018, A review of emerging technologies for remediation of PFASs. *Remediation*. 28(2):101-126.
- D. Ryan, Mayer, B., Baldus C, McBeath S, Wang Y, & McNamara P, 2021, Electrochemical technologies for per-and polyfluoroalkyl substances mitigation in drinking water and water treatment residuals. *AWWA Water Science*, 3(5), e1249.
- H. Safavi, Golmohammadi, M, & Sandoval-Solis S, 2015, Expert knowledge based modeling for integrated water resources planning and management in the Zayandehrud River Basin. *J Hydrology*, 528, 773-789.
- C. Samwald, C. Jentzsch, C.S. Bouton, E. Kallesøe, J. Willighagen, M.S. Hajagos, E. Marshall, O. Prud'hommeaux, E. Hassanzadeh. Pichler and S. Stephens, 2011. Linked open drug data for pharmaceutical research and development. *Journal of cheminformatics*, 3(1), p.19.
- T. Sidnell, Wood R, Hurst J, Lee J, & Bussemaker M, 2022. Sonolysis of Per-And Poly Fluoroalkyl Substances (PFAS): A meta-analysis. *Ultrasonics sonochemistry*, 105944.
- P. Taylor, Yamada T, Striebich RC, Graham JL, Giraud RJ., 2014 Investigation of waste incineration of fluorotelomer-based polymers as a potential source of PFOA in the environment. *Chemosphere*. 110:17-22.
- N. Trokanas, Cecelja, F., & Raafat, T., 2014. Semantic input/output matching for waste processing in industrial symbiosis. *Computers & Chemical Engineering*, 66, 259-268.

# Combined metabolic modeling and experimental data for enhanced biotechnological production

Anita L. Ziegler,<sup>a\*</sup> Melanie Filbig,<sup>b\*</sup> Johannes Parschau,<sup>a,b</sup> Till Tiso,<sup>b</sup> Lars M. Blank,<sup>b</sup> Alexander Mitsos,<sup>a,1</sup>

<sup>a</sup>Process Systems Engineering (AVT.SVT), RWTH Aachen University, Aachen, Germany

<sup>b</sup>Institute of Applied Microbiology (iAMB), Aachen Biology and Biotechnology (ABBt), RWTH Aachen University, Aachen, Germany

\*Equally contributing authors <sup>1</sup>Corresponding author, email: [amitsos@alum.mit.edu](mailto:amitsos@alum.mit.edu)

## Abstract

Biotechnological products are promising sustainable alternatives to fossil-derived fuels and plastics. A modified strain of the bacterium *Pseudomonas putida* KT2440 produces 3-(3-hydroxyalkanoyloxy)alkanoic acids (HAA) from glucose, which can serve as a precursor for bioplastic and biofuel production. We investigate the use of acetate as alternative, sustainable, and low-cost substrate. We use an expanded version of the genome-scale metabolic model *i*JN1463 to improve the understanding of the metabolism. Rates and yields calculated from experimentally obtained data were used as input for flux balance analysis (FBA) and compared to computational data. Despite substantial model-experiment mismatch, FBA provides insights into the acetate metabolization. Finally, we utilize the optimization-based gene knockout method OptKnock to propose potential targets for strain engineering with the aim of improved product yields.

**Keywords:** Metabolic modeling, Biotechnology, Acetate, HAA, *Pseudomonas putida*.

## 1. Introduction

Genome-scale metabolic models (GEMs) contain the entire set of genes of an organism and the metabolic reactions encoded by those genes. GEMs are built based on experimentally obtained data and genome annotation data and are used to predict cellular responses and metabolic fluxes under different environmental conditions, *e.g.*, different substrates, by optimization techniques such as flux balance analysis (FBA) (Varma & Palsson 1994). Thus, GEMs help to understand metabolic processes in a cell and increasingly find applications in biotechnology to support and reduce experimental approaches. FBA formulates a cellular objective and metabolic constraints as a linear optimization program to quantify the metabolic state of a cell under the assumption of a steady state (Orth 2010). Thereby, optimal metabolic flux distributions are calculated to predict growth rates, as well as maximal theoretical biomass or product yields. The experimental results can in turn be used to validate and optimize the model predictions. Improving the performance of a biotechnological strain often requires time-consuming genetic engineering. To reduce experimental effort, the elimination of reactions that result in a beneficial metabolic configuration for improved product yields can be predicted with GEMs. The bilevel programming framework OptKnock was developed by Burgard *et al.* (2003) to propose optimal gene deletions in *Escherichia coli*.

We investigate the biosynthesis of 3-(3-hydroxyalkanoyloxy)alkanoic acids (HAA) with an engineered *Pseudomonas putida* KT2440 strain on acetate in an approach combining



experimental and computational methods. HAA can be produced recombinantly using *P. putida* KT2440 KS3 (Blesken 2020). The current production, however, relies on glucose. Herein, we focus on acetate as an alternative substrate, as it is a low-cost, biotechnological substrate due to its abundance in several side streams (Kiefer 2021) and can be produced from C1-gases, such as CO<sub>2</sub> (Kiefer 2021). Thus, acetate contributes to an improved CO<sub>2</sub> balance of the overall process. Although *P. putida* KT2440 KS3 can use acetate as the sole carbon source, the growth behavior of the strain is impaired by the presence of the C2-compound and product titers lag behind titers achieved from glucose. FBA and OptKnock are used to help understand the altered metabolism on acetate when compared to metabolizing glucose and find potential targets for strain engineering toward optimized product formation. As a basis for FBA and OptKnock, several GEMs are available for *P. putida*. The most current and complete model available is *iJN1463* (Nogales 2020). In this study, *iJN1463* was expanded to contain the synthesis and export reactions for HAA. Experimental data including growth rate, biomass yield, product yield, substrate uptake rate, as well as CO<sub>2</sub> production rate were obtained from shake flask cultivations of *P. putida* KT2440 KS3 on glucose and acetate. While the substrate uptake rate was used as input for FBA, the other data were used to compare and evaluate computationally obtained data. Internal fluxes were calculated and displayed with special interest on CO<sub>2</sub>-forming reactions. Moreover, beneficial gene deletions resulting in higher HAA yields were predicted using OptKnock in several rounds comprising one to five target reactions. This work highlights the possibilities and advantages of a combined approach of experimental and computational work.

## 2. Methods

### 2.1. Experimental methods

The bacterium *P. putida* KT2440 KS3 (Blesken 2020) was cultivated in mineral salts medium as described by (Hartmans 1989) with threefold buffer concentration when 0.17 Cmol of glucose or fourfold buffer concentration when 0.17 Cmol acetate, were used as the sole carbon source. Cultivations were performed in 500 ml shake flasks with 10 % filling volume at 250 rpm and a shaking diameter of 50 mm. Main cultures were inoculated with a defined volume of the pre-cultures to start with an optical density at 600 nm (OD<sub>600</sub>) of 0.1. Cultivations with online monitoring of CO<sub>2</sub> resulting from bacterial metabolic activity were performed with 0.13 Cmol glucose or acetate in 1 L shaking flasks with a 50 ml filling volume. Biomass formation was monitored offline by measurement of the OD<sub>600</sub>, which was converted into cell dry weight (CDW) using a correlation of 0.31 g L<sup>-1</sup>.

Concentrations of glucose and acetate were determined using HPLC as described by Blesken *et al.* (2020). HAA titers in cultivation broth were determined using an Ultimate3000 HPLC system with a dual gradient pump, connected to a Corona Veo Charged Aerosol Detector (all Thermo Scientific). An analytical and an inverse gradient set up of 0.2 % formic acid in acetonitrile and 0.2 % formic acid in ultrapure water were applied for chromatographic separation with a NUCLEODUR C18 Gravity column (Macherey-Nagel) with 150 mm length and a diameter of 3 mm.

CO<sub>2</sub> production during cultivation was monitored using BSP-CO<sub>2</sub> sensors and the BlueVIS software (BlueSens gas sensor GmbH). Yields of biomass per substrate were calculated by division of the maximal reached CDW by the applied substrate concentration. Product yields were determined by dividing maximal product concentration by applied substrate concentration. Substrate uptake rates were calculated by dividing the growth rate  $\mu$  by the biomass yield per substrate.

## 2.2. Flux balance analysis and visualization

To obtain insights into the metabolism, FBA was performed according to

$$\begin{aligned} \max_{\mathbf{v} \in \mathbb{R}^n} \quad & \mathbf{c}^T \mathbf{v} & \mathbf{v}: & \text{flux vector; } n: \text{ number of reactions} \\ \text{s. t.} \quad & \mathbf{S} \mathbf{v} = 0 & \mathbf{c}: & \text{parameter vector to choose objective flux} \\ & v_i^{lb} \leq v_i \leq v_i^{ub} \quad \forall i \in \{1, \dots, n\}, & \mathbf{S}: & \text{stoichiometric matrix from GEM} \\ & & lb: & \text{lower bound; } ub: \text{ upper bound} \end{aligned}$$

where the maximization of the biomass flux or the maximization of the HAA flux was chosen as objective function. The FBA is based on the stoichiometric matrix  $\mathbf{S}$  of a GEM. The GEM applied in this work is *iJN1463* (Nogales 2020), which contains 2,927 reactions, 2,153 metabolites and 1,462 genes. We extended the network by the gene *rhlA* and the reactions specific to the strain KS3, namely four synthesis reactions in the cytosol, one for each HAA congener, and four transport reactions from the cytosol to the extracellular space. With the extracellular HAA congeners as reactants, one exchange reaction was added that simulates the production of HAA in the congener composition that was experimentally determined. Exchange reactions represent the exchange of the cell with the environment and enable closing the mass balance. The stoichiometric factors in the exchange reactions were adjusted on each substrate to display the altered HAA composition on glucose and acetate. The software COBRAPy ([opencobra.github.io/cobrapy](https://opencobra.github.io/cobrapy)) serves to conduct the FBA and Gurobi 9.5.2 serves as solver. Substrate uptake rates determined experimentally were set as input. To visualize the fluxes through the network, the software Escher ([escher.github.io](https://escher.github.io)) was utilized.

## 2.3. Optimal knockouts

Genes encode enzymes that catalyze reactions. When knocking out a certain gene by means of genetic engineering, the associated reaction is not catalyzed anymore, *i.e.*, cannot carry flux anymore. To suggest gene knockouts, the formulation OptKnock was applied, which suggests reactions that are advantageous to block, *e.g.*, reactions towards a byproduct. OptKnock is a bilevel optimization program, where the upper level maximizes the flux of the target chemical, and the lower level maximizes the flux of biomass as follows:

$$\begin{aligned} \max_{\mathbf{y} \in \{0,1\}^r} \quad & v_{HAA} & \mathbf{y}: & \text{knockout vector} \\ \text{s. t.} \quad & \max_{\mathbf{v} \in \mathbb{R}^n} v_{biomass} & r: & \text{number of (possibly reversible) reactions} \\ & \text{s. t. } \mathbf{S} \mathbf{v} = 0 & \mathbf{B}: & \text{mapping matrix} \\ & & & (\mathbf{B}\mathbf{y})_i v_i^{lb} \leq v_i \leq (\mathbf{B}\mathbf{y})_i v_i^{ub} \quad \forall i \in \{1, \dots, n\}. \quad (r \text{ to } n) \end{aligned}$$

In this study, the HAA exchange reaction flux  $v_{HAA}$  was set as objective function of the upper level and the biomass flux  $v_{biomass}$  as objective in the lower level. For numerical reasons, the stoichiometric network with  $r$ , possibly reversible, reactions was transformed to an irreversible network with  $n$  irreversible reactions. OptKnock was implemented in our in-house software libALE (Djelassi 2020) and solved using the solver Gurobi 9.5.2.

## 3. Results and discussion

### 3.1. Cultivation on acetate is possible, but achieves lower yields than on glucose

*P. putida* KT2440 KS3 was cultivated on 0.17 Cmole glucose and acetate to determine growth rate, HAA production rate, biomass, and product yield, as well as substrate uptake rate (Figure 1). While on glucose, the strain reaches a maximal cell dry weight (CDW) of 1.6 g L<sup>-1</sup> after 9 h with a growth rate of 0.47 h<sup>-1</sup>, the strain reached a final biomass of 1 g L<sup>-1</sup> CDW after 14 h cultivation time on acetate, exhibiting a growth rate of 0.37 h<sup>-1</sup>. Thus, the biomass yield per substrate is 35 % lower on acetate than on glucose. HAA formation is threefold higher on glucose than on acetate: while on glucose, a final HAA

titer of  $270 \text{ mg L}^{-1}$  was reached after 9 h, on acetate a maximal titer of  $90 \text{ mg L}^{-1}$  was reached after 14 h. The pH increases during the course of the cultivation to a pH above 8 due to the consumption of acetic acid from the medium.

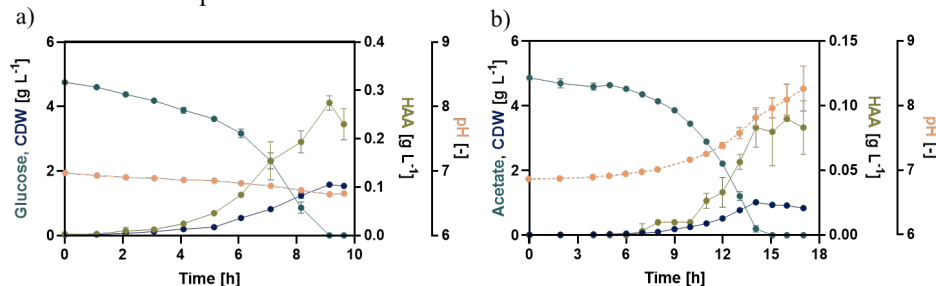


Figure 1: Biomass, HAA and substrate concentration, and pH during cultivation of *P. putida* KT2440 KS3 on a) 0.17 Cmol glucose and b) 0.17 Cmol acetate.

Measurements of the  $\text{CO}_2$  produced by the microbes during cultivation revealed a slightly higher  $\text{CO}_2$  production on acetate than on glucose. While on glucose *P. putida* KT2440 KS3 produced  $0.35 \text{ Cmol CO}_2$  per Cmol substrate, on acetate  $0.37 \text{ Cmol CO}_2$  were produced per Cmol substrate. All in all, *P. putida* KT2440 KS3 exhibits lower rates and yields on acetate than on glucose, which might result from the lower enthalpy of combustion of acetate ( $437 \text{ kJ Cmol}^{-1}$ ) compared to glucose ( $467 \text{ kJ Cmol}^{-1}$ ). However, growth as well as product formation is possible from acetate as sole carbon source. In order to elucidate acetate metabolism and calculate maximal theoretical yields, FBA was performed in a next step. For FBA, substrate uptake rates are required. These substrate uptake rates were determined on  $0.17 \text{ Cmol glucose}$  and  $0.17 \text{ Cmol acetate}$  to be  $7.8 \text{ mmol g}_{\text{CDW}}^{-1} \text{ h}^{-1}$  ( $0.04 \text{ Cmol g}_{\text{CDW}}^{-1} \text{ h}^{-1}$ ) and  $28.5 \text{ mmol g}_{\text{CDW}}^{-1} \text{ h}^{-1}$  ( $0.057 \text{ Cmol g}_{\text{CDW}}^{-1} \text{ h}^{-1}$ ), respectively.

### 3.2. Flux balance analysis gives insights into the altered metabolic configuration

FBA was used to calculate the theoretical yield of biomass and HAA per substrate. Those theoretical yields were then compared to wet lab results that were calculated from experimental data (Table 1).

Table 1: Yield of biomass and HAA on glucose and on acetate: wet lab results, i.e., calculated from experimental data, vs. dry lab results, i.e., from flux balance analysis (FBA). The errors represent the standard deviations of triplicate cultivations ( $n=3$ ). For FBA, the uptake rates were set to equal experimentally determined uptake rates. In case (a) of the FBA, the objective function is the flux of biomass; in case (b) the flux of HAA.

	Substrate	Wet lab	Dry lab
$Y_{\text{Biomass/Substrate}} [\text{g}_{\text{CDW}} \text{g}^{-1}]$	Glucose	$0.34 \pm 0.01$	0.54 (a)
	Acetate	$0.22 \pm 0.01$	0.46 (a)
$Y_{\text{HAA/Substrate}} [\text{Cmol Cmol}^{-1}]$	Glucose	$0.097 \pm 0.005$	0.68 (b)
	Acetate	$0.030 \pm 0.001$	0.59 (b)

When cultivating on glucose, the biomass yield reaches 63 % of the theoretical biomass yield, which is a small difference, when considering that the theoretical yield does not account for any product or byproduct formation. On acetate, 48 % of the theoretical biomass yield is reached experimentally, which is a larger difference. Overall, we conclude that FBA is a valid means to predict growth of *P. putida* KT2440 KS3, especially on glucose. One possible explanation for the greater deviation with acetate as a substrate is the toxic effect of the substrate (Kiefer 2021), which is not considered in the metabolic network. When looking at the HAA yields, FBA predicts values seven times higher or more, than determined experimentally for both substrates. Hence, the

experimental yield has much potential for improvement. However, the overall trend of a lower HAA yield on acetate than on glucose is predicted in the FBA.

To find a reason for this trend, in a next step, the altered tricarboxylic acid (TCA) cycle of the cell on glucose vs. on acetate was visualized in Escher maps (Figure 2). The magnitude of the fluxes in the TCA cycle gives insights into the energy demand of a cell.

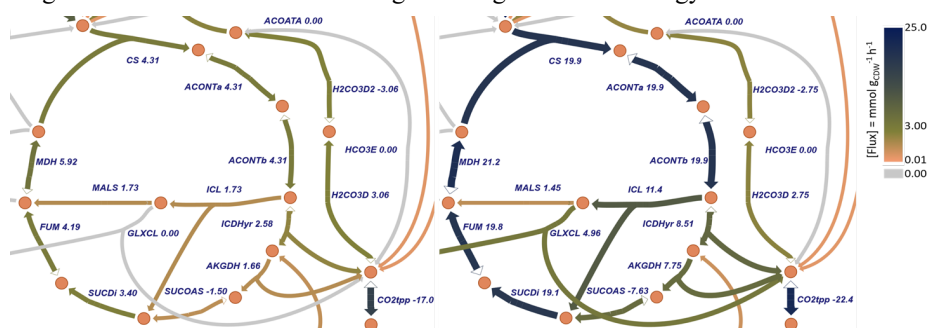


Figure 2: Visualization of the tricarboxylic acid cycle with glucose as carbon substrate (left) and acetate as carbon source (right). Biomass flux is the objective function; a substrate uptake of  $0.05 \text{ Cmol g}_{\text{CDW}}^{-1} \text{ h}^{-1}$  and a threshold of  $0.01 \text{ mmol g}_{\text{CDW}}^{-1} \text{ h}^{-1}$  on the HAA exchange reaction were set as constraints to the FBA. Reactions are labeled with BIGG ID (bigg.ucsd.edu) and the flux value. The figure was adapted from Parschau (2022).

The Escher maps reveal that the cell exhibits a different metabolic configuration on glucose than on acetate, meaning altered reaction fluxes inside the cell. The TCA cycle exhibits fluxes of  $1.5\text{-}5.9 \text{ mmol g}_{\text{CDW}}^{-1} \text{ h}^{-1}$  on glucose, whereas it displays fluxes of  $7.6\text{-}21.2 \text{ mmol g}_{\text{CDW}}^{-1} \text{ h}^{-1}$  on acetate. The higher fluxes through the TCA cycle probably result from the higher energy demand on acetate and the lower energy content of this carbon source. The  $\text{CO}_2$  data derived from FBA follow the same trend as  $\text{CO}_2$  data obtained experimentally: with acetate as substrate, more  $\text{CO}_2$  is produced than on glucose, namely  $22.4 \text{ mmol g}_{\text{CDW}}^{-1} \text{ h}^{-1}$  vs.  $17.0 \text{ mmol g}_{\text{CDW}}^{-1} \text{ h}^{-1}$  on glucose. Hence, more byproduct is formed on acetate. This may cause less HAA production on acetate than on glucose.

### 3.3. Two gene deletions may lead to improved yield

The optimization formulation OptKnock was used to suggest beneficial blocking of reactions, resulting in higher HAA yields. When only one reaction was set to be blocked, no solution was found that improved HAA yield. For two reactions, the HAA yield per substrate increases to  $0.53 \text{ Cmol Cmol}^{-1}$  while reducing the biomass production to the set threshold of  $0.08 \text{ h}^{-1}$ . The HAA yield thereby reaches 90 % of the theoretical yield (see Table 1). The suggested enzymatic reactions to be blocked are catalyzed by dihydromonapterin reductase (NADPH) and dihydromonapterin reductase (NADH). These reactions are located in the biomass synthesis pathway. Hence, by reducing the carbon flux to biomass, the carbon flux is redirected towards HAA synthesis. However, experimental experience indicates that HAA production is highly connected to biomass production. Presumably, knocking out these reactions would impede the organism. The next gene knockout suggestion addresses five reactions due to numerical problems with three and four reactions. The result for five reactions shows no improvement compared to the two knockouts. We conclude that two knockouts are sufficient to improve the HAA yield. In conclusion, the results show that there is potential for improving the yield by applying metabolic engineering. However, more advanced optimization formulations should be considered that guarantee for a minimal product yield (Tepper & Shlomi 2009) and that better represent the microorganism (Apaydin 2017) to improve the predictions.

#### 4. Conclusion

In our study, we have shown that HAA can be produced on acetate as sole carbon source by an engineered *P. putida* strain, however with a lower yield than on glucose. We have explained this by the toxicity of acetate and the altered metabolic configuration of the cell. The integration of experiments and simulations has brought confidence because both methods displayed the same trend. We found that more CO<sub>2</sub> is produced during growth on acetate, which reduces the carbon atoms available for HAA yield. To reduce the model-experiment mismatch, regulatory effects should be considered in the model formulation. We have suggested gene knockouts to improve the yield of HAA on acetate, which must be elaborated deeper by more advanced formulations and proven in the laboratory as a next step. Moreover, fermentation conditions with lower acetate concentrations can mitigate the toxicity of acetate to the cell and further improve HAA yield.

**Author Contributions:** ALZ designed the computations and MF designed the laboratory experiments. JP performed the computational and experimental laboratory work under the supervision of ALZ and MF. ALZ, MF and JP visualized, analyzed, and discussed the data. ALZ and MF wrote the manuscript draft. TT, LMB and AM conceptualized the project, discussed the data, and reviewed the draft. All authors read and approved the final manuscript.

**Acknowledgements:** This project was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – Cluster of Excellence 2186 „The Fuel Science Center” – ID: 390919832

#### References

- M. Apaydin, L. Xu, B. Zeng, X. Qian, 2017, Robust mutant strain design by pessimistic optimization, *BMC Genomics*, 18, 6, 677.
- C. C. Blesken, I. Bator, C. Eberlein, H. J. Heipieper, T. Tiso, L. M. Blank, 2020, Genetic Cell-Surface Modification for Optimized Foam Fractionation, *Front. Bioeng. Biotech.*, 8, 1116.
- A. P. Burgard, P. Pharkya, C. D. Maranas, 2003, Optknock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization, *Biotechnol. Bioeng.*, 84, 6, 647-57.
- H. Djelassi, O. Stein, A. Mitsos, 2020, Discretization-based algorithms for the global solution of hierarchical programs, *Lehrstuhl für Systemverfahrenstechnik, RWTH Aachen University*.
- S. Hartmans, J. P. Smits, M. J. van der Werf, F. Volkering, J. A. de Bont, 1989, Metabolism of Styrene Oxide and 2-Phenylethanol in the Styrene-Degrading *Xanthobacter* Strain 124X, *Appl. Environ. Microbiol.*, 55, 11, 2850-5.
- D. Kiefer, M. Merkel, L. Lilge, M. Henkel, R. Hausmann, 2021, From acetate to bio-based products: underexploited potential for industrial biotechnology, *Trends Biotechnol.*, 39, 4, 397-411.
- J. Nogales, J. Mueller, S. Gudmundsson, F. J. Canalejo, E. Duque, J. Monk, A. M. Feist, J. L. Ramos, W. Niu, B. Ø. Palsson, 2020, High-quality genome-scale metabolic modelling of *Pseudomonas putida* highlights its broad metabolic capabilities, *Environ. Microbiol.*, 22, 1, 255-69.
- J. D. Orth, I. Thiele, B. Ø. Palsson, 2010, What is flux balance analysis?, *Nature Biotechnol.*, 28, 3, 245-48.
- J. Parschau, 2022, Numerical optimization supported by experimental data towards improved production of 3-(3-hydroxyalkanoyloxy)alkanoic acids by *Pseudomonas putida*, *Bachelor Thesis, RWTH Aachen University*.
- N. Tepper, T. Shlomi, 2009, Predicting metabolic engineering knockout strategies for chemical production: accounting for competing pathways, *Bioinformatics*, 26, 4, 536-43.
- A. Varma, B. Ø. Palsson, 1994, Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110, *Appl. Environ. Microbiol.*, 60, 10, 3724-3.

# Machine learning models for predicting membranolytic anticancer peptides

Fatemeh Alimirzaei,<sup>a</sup> Chris A. Kieslich,<sup>a</sup>

<sup>a</sup>*Chemical Engineering Department, Auburn University, Auburn, AL 36849, USA  
cak0071@auburn.edu*

## Abstract

After heart disease, cancer is the second leading cause of death worldwide. Recently, membranolytic anticancer peptides (ACPs) have received considerable attention for their ability to target and kill cancer cells. Identification of ACPs is costly and usually time-consuming. Therefore, the development of efficient computational methods is of a great importance to aid in the identification of potential ACP candidates. In the current study, we developed multiple models using support vector machines (SVMs), gradient boosting classifiers (GB), and random forest classifiers (RF) to predict membranolytic anticancer activity given a peptide sequence. Oscillations in physiochemical properties in protein sequences have been shown to be predictive of protein structure and function, and in this work, we are taking advantage of these known periodicities to predict ACP sequences. To this end, Fourier transforms were applied to the property factor vectors to measure the amplitude of the physiochemical oscillations, which served as the features for our models. Peptides targeting breast and lung cancer cells were collected from the CancerPPD database and converted into physiochemical vectors using 10 property factors for the 20 natural amino acids. Using these datasets, cross-validation has been applied to train and tune the models based on multiple training and testing sets. Additionally, feature selection has been performed to further optimize our SVM models. To evaluate the models, performance has been quantified based on cross-validation classification accuracy. Furthermore, to try our prediction accuracy, we have also considered other sets of physiochemical features and properties of amino acids from the literature into our models.

**Keywords:** Membranolytic anticancer peptides, Computer-aided drug design, Machine learning

## 1. Introduction

Cancer is the second leading cause of death worldwide, and it is anticipated the number of deaths caused by cancer will be more than 13 million by 2030 (Liscano et al. 2020). Not only are conventional cancer therapies expensive, but poor specificity often leads to devastating side effects, such as anemia, gastrointestinal mucositis, alopecia, cardiotoxicity, or osteoporosis (Manavalan et al. 2017; Gabernet et al. 2016). Moreover, cancer cells may develop resistance to chemotherapy and receptor-targeted anticancer agents (Gabernet et al. 2019). Therefore, developing innovative anticancer drugs is essential to attenuate cancer cell proliferation (Manavalan et al. 2017).

Recently, membranolytic anticancer peptides (ACPs) have been proposed as a potential strategy in the fight against cancer by disrupting cancer cell membranes and destroying them. ACPs are typically short peptides composed of 10–60 amino acids that can inhibit tumor cell proliferation or migration or suppress the formation of tumor blood vessels and are less likely to cause drug resistance (Xie et al. 2020). They tend to be amphipathic,

positively charged, and many of them change their structure from random to helical conformation upon associating with a preferred lipid bilayer (Queme-Pena et al. 2021). It has been hypothesized that the cationic nature leads to their selectivity for cancer with negatively charged lipid membranes (Grisoni et al. 2019). However, only a limited number of potent anticancer peptides are known to date. For instance, just 7% of the 2981 peptides annotated in the antimicrobial peptide database (APD3) (Wang et al. 2016), a group sharing physiochemical properties with ACPs, have been tested for anticancer activity (Grisoni et al. 2019), or only 53 peptides in the cancerPPD database possess low micromolar activity ( $EC_{50} < 5\mu M$ ) to be useful for cancer treatment (Grisoni et al. 2019). Furthermore, their experimental identification and development are laborious, expensive, and time-consuming (Yu et al. 2020). Last but not least, the underlying structure-activity relationship that explains the membranolytic properties of these peptides is not completely understood (Queme-Pena et al. 2021). Therefore, innovative tools like sequence-based computational methods are needed for the fast identification and design of potent ACP candidates (Manavalan et al. 2017). To identify ACPs, existing methods try to discriminate ACPs from non-ACPs and then predict their anticancer activity based on the features extracted from known ACPs which their identities are held by publicly available datasets like CancerPPD (Shoobuatong et al. 2018).

In the current study, we are developing models that predict whether a peptide will exhibit ACP activity given only an amino acid sequence of the peptide. For this purpose, we are trying to identify patterns in the physiochemical properties of ACPs by using Fourier Transform-based features that measure the oscillations of physiochemical properties along the length of peptides. Previously we have shown that oscillations in physiochemical properties in peptide sequences are predictive of antiviral peptide activity (Kieslich et al. 2021). Here we aim to take advantage of these known periodicities of amino acid properties to develop models that predict anticancer peptide activity. Furthermore, by performing feature selection, we can determine which features are the most important predicting ACP activity that have the potential of forming a design template.

## 2. Methods

In this work, Python has been used to perform all steps of our analysis, which includes the generation of physiochemical vectors for each peptide, Fourier-based feature extraction, training/validation of classification models for predicting membranolytic anticancer activity, as well as SVM-based feature selection. Below are more elaborate descriptions of how these elements were implemented in this study.

### 2.1. ACP Dataset

One of the crucial requirements for developing data-driven classification models is having access to adequately large datasets containing both amino acid sequence and function labels. Currently, there are multiple publicly available databases which hold the identities of some known anticancer peptides, including APD (Wang et al. 2004), CancerPPD (Tyagi et al. 2015), DADP (Novkovic et al. 2012), DBAASP (Pirtskhalava et al. 2021), LAMP (Zhao et al. 2013). In our study, two peptide datasets targeting breast and lung cancer were chosen owed to their relevance for human health (2.9 million cases for each cancer type in 2018) (Grisoni et al. 2019). These datasets were previously assembled and curated manually from CancerPPD (Grisoni et al. 2019), so that the final training sets contained 232 peptides targeting breast cancer cells, including 116 active and 116 nonactive ACPs, and 198 peptides targeting lung cancer cells containing 99 active and 99 nonactive ACPs.

## *2.2. Feature Extraction*

One of the prerequisites for training a machine learning-based classifier is transforming the original raw training data into a set of “features”. For this purpose, we first converted the amino acid sequences of each peptide in our raw dataset into physicochemical vectors using 10 physical properties factors of the 20 naturally occurring amino acids. These 10 physicochemical property factors were generated based on principle component analysis of a large set of amino acid physicochemical properties (Kidera et al. 1985). Kidera et al. describe the property factors as  $\alpha$ -helix or bend-structure preference-related, bulk-related, structure preference-related, hydrophobicity-related, and 5 additional property factors that represent mixtures of several physicochemical properties. Each peptide is converted into 10 physicochemical vectors (numerical vectors) based on each of the physicochemical property factors. Afterwards, we used the `fft` function in Python to apply Fourier transforms to each of the property factor vectors. It is worth mentioning, in order to make sure that the same number of frequency (Fourier) components have been generated for each property vector, we have applied zero-padding with the assumption of a maximum sequence length of 64 amino acids. The maximum ACP sequence length across both peptide datasets is 38 amino acids. In this way, the amino acid properties corresponding to positions that are longer than a given peptide were set to zero. After using zero padding and the Fourier transform, 64 Fourier coefficients are generated for each property of each peptide. However, Fourier transform spectrums are in mirror shape, and only half of the coefficients (i.e., 32) were considered in addition to the center (0). Ultimately there were 33 features for each of the ten property factors in each peptide, which means 330 features are extracted for each peptide (feature set1). In other words, each sequence was represented by 33 Fourier coefficients measuring the oscillation of the 10 property factors. We also generated two more feature sets for further evaluation of our prediction accuracies. For feature set 2, we added 10 additional features to feature set 1 based on amino acid composition. The additional features were based on the percentage of amino acid positions that fit categories of positively charged, negatively charged, aliphatic, aromatic, polar, hydrophobic, size (tiny, small, large), as well as the total charge of the peptide. The amino acid categories were based on those that have been previously proposed for a related task (Manavalan et al. 2017). For generating the feature set3, instead of 10 amino acids property factors used to create feature set1, we applied 8 newly reported properties of amino acids by Medina-Ortiz et al. that are related to  $\alpha$  structure,  $\beta$  structure, hydrophobicity, volume, energy, hydropathy, secondary structure, and other indexes (Medina-Ortiz et al. 2022).

## *2.3. Support Vector Machines and Grid-search*

In this study, all support vector machines have been trained by using the `svc` function in Python based on the radial kernel (Gaussian radial basis function (RBF)). In order to tune the cost and gamma hyperparameters of the SVM models, a grid search has been used with cost and gamma values based on powers of ten,  $10^n \forall n \in [0, \dots, 6]$  for cost values and  $10^n \forall n \in [-6, \dots, 0]$  for gamma values, where  $n$  is an integer.

## *2.4. SVM Feature Selection*

Feature selection plays a critical role in data science since it leads to the identification of the fundamental set of predictive descriptors (features). Furthermore, it can increase the robustness of the model to prevent overfitting (Kieslich et al. 2021). Previously, a feature selection algorithm has been developed based on non-linear SVM, which is common in nature and has been applied to predict faults in chemical plants (Onel et al. 2018) and HIV-1 viral entry (Kieslich et al. 2016). It is a model-based algorithm requiring first training an SVM model prior to computing a criterion quantifying the contribution



between each single feature to the SVM objective function to determine which features to eliminate. The criterion (Eq. 1) is derived in terms of sensitivity analysis of the dual formulation of SVM models.

$$crit_k = -\frac{1}{2} \sum_i \sum_j \alpha_i^* \alpha_j^* y_i y_j \left. \frac{\partial K(x_i \circ z, x_j \circ z)}{\partial z_k} \right|_{z=1} \quad (1)$$

This algorithm applies a greedy approach in order to rank the features, so that at the beginning it starts with training a model based on all of the features, computes the criteria for all features and then removes a fraction of the features with the largest criteria values. (In other words, the largest criteria values have the least importance for the model). In the current work, %33 of the remaining features has been removed after each iteration of the algorithm, and the hyperparameters have been retuned after each iteration of the algorithm. The feature ranking procedure was applied to each of the five training sets and a consensus ranking was generated based on the average rank of each feature across five training sets.

### 2.5. Cross Validation and SVM Models Accuracy

Five-fold cross-validation, based on two balanced training datasets, was applied to tune and validate the models based on the first sorting of the peptide sequences according to length, and then choose five training and testing sets with an equal number of samples for each peptide class in each peptide datasets, separately. For all datasets, model performance was evaluated based on classification accuracy which is reported as the fraction of classes (active or nonactive ACPs) that have been predicted accurately in the test sets.

### 2.6. Gradient Boosting and Random Forest Models

In this study, all gradient boosting and random forest models have been trained and evaluated using “sklearn.ensemble” and “sklearn.metrics” packages in Python. Also, like the SVM model, 5-fold cross-validation was performed to generate train/test sets for them. It is worth mentioning for both of these models, we used the default values of the parameters controlling the size of the tree publicly available on the “scikit\_learn” website ([Scikit-learn: Machine Learning in Python](https://scikit-learn.org/stable/tutorial/tutorial.html)) (Pedregosa et al. 2011).

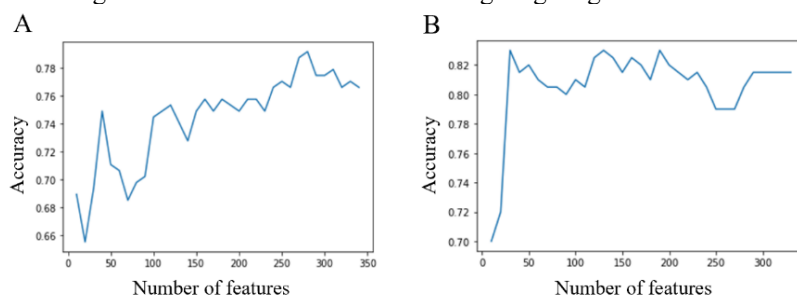
**Table 1.** Statistics of the performance of machine learning-based models predicting ACP activities.

Cancer type	Feature set/Number of features	SVM model accuracy/Minimum number of features to maintain the accuracy	Gradient Boosting model accuracy	Random Forest model accuracy
Breast	Set1/330	76%/~330	68%	68%
	Set2/340	79%/~280	59%	60%
	Set3/264	68%/~240	62%	62%
Lung	Set1/330	83%/~25	75%	72%
	Set2/340	82%/~270	74%	73%
	Set3/264	75%/~130	66%	65%

## 3. Results

Based on the proposed feature sets (3 sets in total), we developed models using SVM, Gradient Boosting, and Random Forest methods to distinguish the active and nonactive ACP peptides based on the Breast and Lung datasets. For SVM classification tasks, we performed feature selection to rank the physicochemical features. To measure the contribution of each feature to model accuracy we performed five-fold cross-validation

after adding each feature one at a time, starting with the highest-ranked feature. Also, for gradient boosting and random forest models, we have done five-fold cross-validation for all datasets and calculated the average accuracy across 5 train/test sets for each predictive model, discriminating two classes of ACPs targeting breast and lung cancer cells. A summary of obtained results of all the predictive model performances is shown in Table 1. Among them, the SVM model using feature sets2 has the highest accuracy (79%) for discriminating active and nonactive anticancer peptides targeting breast cancer cells. Also, for distinguishing between active and nonactive ACPs targeting lung cancer cells, the SVM model using feature set 1 has the highest accuracy (83%) compared to the rest of the predictive models. These reported cross-validation accuracies are the average of the classification accuracies for the five training and testing sets. As it is seen in Figure 1, the minimum number of features required to maintain those accuracies is around 280 features for the model predicting anticancer activities of the peptides targeting breast cancer cells, and approximately 25 features are needed to maintain the highest accuracy of the model predicting ACP targeting lung cancer cells. It is worth mentioning features in the x-axis on the plots are ordered based on the feature selection algorithm meaning adding more and more features. For example, those 25 features are the most important ones to distinguish active and nonactive ACPs targeting lung cancer cells.



**Figure 1.** Feature contribution to model accuracy based on feature selection ranking. A) Feature selection results for classifying active vs. nonactive anticancer peptides targeting breast cancer cells; B) Feature selection results for classifying active vs. nonactive anticancer peptides targeting lung cancer cells.

#### 4. Conclusion

In this study, we have developed a support vector machine, gradient boosting, and random forest models that distinguish between active and nonactive ACP sequences for two membranolytic ACP datasets targeting breast and lung cancer cells. To develop these models, we used different sets of amino acid property factors of naturally occurring amino acids from the literature to convert ACP sequences into property vectors that served as the input for the generation of Fourier-based features. Our results show that SVM models perform better in discriminating ACP activity. In addition, by ranking the importance of the developed Fourier-based features, we were able to train SVM models with improved accuracy and generalizability, while also starting to gain some insights into the importance of oscillations of physiochemical properties for ACP function. It can be concluded that our approach for feature extraction and model development, specifically including the SVM feature selection algorithm, has the potential for use in predicting peptide properties and functions as well as designing innovative anticancer peptides in the future. However, given the remaining challenges in discriminating between active and nonactive ACPs, more efforts are required that might need both further computational and even experimental investigations.

## References

- G. Gabernet, A. T. Muller, J. A. Hiss, G. Schneider, 2016, Membranolytic anticancer peptides. *Medchemcomm*, 7(12): 2232-2245.
- G. Gabernet, G. Gabernet, D. Gautschi, A.T. Müller, C.S. Neuhaus, L. Armbrrecht, P.S. Dittrich, J.A. Hiss, G. Schneider, 2019, In silico design and optimization of selective membranolytic anticancer peptides. *Sci Rep*, 9(1), 11282.
- F. Grisoni, C. S. Neuhaus, M. Hishinuma, G. Gabernet, J. A. Hiss, M. Kotera, 2019, De novo design of anticancer peptides by ensemble artificial neural networks. *J Mol Model*, 25(5), 112.
- A. Kidera, Y. Konishi, M. Oka, T. Ooi, H.A. Scheraga, 1985, Statistical-Analysis of the Physical-Properties of the 20 Naturally-Occurring Amino-Acids. *Journal of Protein Chemistry*, 4(1), 23-55.
- C.A. Kieslich, F. Alimirzaei, H. Song, M. DO, P. Hall, 2021, Data-driven prediction of antiviral peptides based on periodicities of amino acid properties, in *Computer Aided Chemical Engineering*, M. Türkay and R. Gani, Editors, Elsevier, 2019-2024.
- C.A. Kieslich, P. Tamamis, Y.A. Guzman, M. Onel, C.A. Floudas, 2016, Highly accurate structure-based prediction of HIV-1 coreceptor usage suggests intermolecular interactions driving tropism. *PLOS ONE*, 11(2), e0148974.
- Y. Liscano, Y., J. Onate-Garzon, and J.P. Delgado, 2020, Peptides with Dual Antimicrobial-Anticancer Activity: Strategies to Overcome Peptide Limitations and Rational Design of Anticancer Peptides. *Molecules*, 25(18), 4245.
- B. Manavalan, S. Basith, T.H. Shin, M. O. Kim, G. Lee, 2017, MLACP: machine-learning-based prediction of anticancer peptides. *Oncotarget*, 8(44), 77121-77136.
- D. Medina-Ortiz, S. Contreras, J. Amado-Hinojosa, J. Torres-Almonacid, J.A. Asenjo, M. Navarro, A. Olivera-Nappa, 2022, Generalized Property-Based Encoders and Digital Signal Processing Facilitate Predictive Tasks in Protein Engineering. *Frontiers in Molecular Biosciences*, 9, 898627.
- M. Novkovic, M. Novković, J. Simunić, V. Bojović, A. Tossi, D. Juretić, 2012, DADP: the database of anuran defense peptides. *Bioinformatics*, 28(10), 1406-7.
- M. Onel, C.A. Kieslich, Y.A. Guzman, C.A. Floudas, E.N. Pistikopoulos, 2018, Big Data Approach to Batch Process Monitoring: Simultaneous Fault Detection and Identification Using Nonlinear Support Vector Machine-based Feature Selection. *Comput. Chem. Eng.*, 115, 46-63.
- M. Pirtskhalava, A. A Armstrong, M. Grigolava, M. Chubinidze, E. Alimbarashvili, B. Vishnepolsky, A. Gabrielian, A. Rosenthal, D. E Hurt, M. Tartakovsky, 2021, DBAASP v3: database of antimicrobial/cytotoxic activity and structure of peptides as a resource for development of new therapeutics. *Nucleic Acids Research*, 49(D1), D288-D297.
- M. Queme-Pena, T. Juhasz, G. Kohut, M. Ricci, P. Singh, I. C. Szigyarto, Z. I. Papp, L. Fülöp, T. Beke-Somfai, 2021, Membrane Association Modes of Natural Anticancer Peptides: Mechanistic Details on Helicity, Orientation, and Surface Coverage. *Int J Mol Sci*, 22(16).
- W. Shoombuatong, N. Schaduangrat, C. Nantasenamat, 2018, Unraveling the bioactivity of anticancer peptides as deduced from machine learning, 17, 734-752.
- A. Tyagi, A. Tuknait, P. Anand, S. Gupta, M. Sharma, D. Mathur, A. Joshi, S. Singh, A. Gautam, G. P.S. Raghava, 2015, CancerPPD: a database of anticancer peptides and proteins. *Nucleic Acids Research*, 43(D1), D837-D843.
- G.S. Wang, X. Li, and Z. Wang, 2016, APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Research*, 44(D1), D1087-D1093.
- Z. Wang, G. Wang, 2004, APD: the Antimicrobial Peptide Database. *Nucleic Acids Res*, 32(Database issue), D590-2.
- M. Xie, D. Liu, Y. Yang, 2020, Anti-cancer peptides: classification, mechanism of action, reconstruction and modification. *Open Biol*, 10(7), 200004.
- L. Yu, R. Jing, F. Liu, J. Luo, 2020, DeepACP: A Novel Computational Approach for Accurate Identification of Anticancer Peptides by Deep Learning Algorithm. *Mol Ther Nucleic Acids*, 22, 862-870.
- X. Zhao, H. Wu, H. Lu, G. Li, Q. Huang, 2013, LAMP: A Database Linking Antimicrobial Peptides. *PLoS One*, 8(6), e66557.

# Data-driven prediction of peptide-MHC binding using oscillations of physicochemical properties

Hyeju Song,<sup>a</sup> Chris A. Kieslich,<sup>a</sup>

<sup>a</sup>*Auburn University, Auburn, AL 36849, USA*  
*kieslich@auburn.edu*

## Abstract

The Major Histocompatibility Complex (MHC) molecules play a major role in T-cell immunogenicity through the recognition of 'non-self' peptides derived from foreign antigens. Therefore, predicting peptides that trigger immune responses is of great interest for the general understanding of T-cell-mediated immunity and the design of peptide-based vaccines and cancer immunotherapy treatments. The presented work tests the performance of allele-specific SVM classification models in the prediction of pMHC binding. The models aim to classify MHC class II binding and non-binding peptides based on their amino acid sequences and derived features. In developing the models, we take advantage of underlying periodicities in physicochemical properties along the sequence of a peptide that has been shown to be predictive of protein structure and function. Once the physicochemical descriptors are generated, Fourier transforms are then applied to be able to encode peptide sequences of varying lengths. In training and testing the model, a comprehensive dataset of MHC class II binding peptides that includes 44 unique MHC class II alleles (molecules) with 60630 binding affinities is taken from the IEDB database. Cross-validation and hyperparameter tuning are applied across multiple train and test datasets. A feature selection algorithm is also incorporated into the model development to identify an essential set of predictive features. The blind test set prediction accuracy of the developed allele-specific models ranges from 0.58- 0.93, with an average classification accuracy of 0.73.

**Keywords:** Computational biology, data-driven, machine learning, support vector machines, Major Histocompatibility Complex, feature selection

## 1. Introduction

Understanding peptide-protein binding in T-cell immunity has potential applications in many biomedical fields, including the development of peptide-based vaccines and personalized cancer immunotherapy. Identifying T-cell epitopes first involves predicting if peptides derived from an antigen (e.g. viral protein) bind to specific Major Histocompatibility Complex (MHC) molecules since the formation of peptide-MHC (pMHC) complexes is a key determinant in the T-cell immunogenicity. Over the past years, a number of computational prediction methods have been developed to identify MHC-binding peptides as an alternative to costly and time-consuming experimental methods (Sanchez-Trincado et al. 2017). In general, the prediction of interactions between MHCs and peptides is challenging due to the highly polymorphic nature of

MHCs and also the varying length of binding peptides. Moreover, it is further compounded by non-binding residues, also known as peptide flanking regions (PFRs), which have been found to affect its binding affinity.

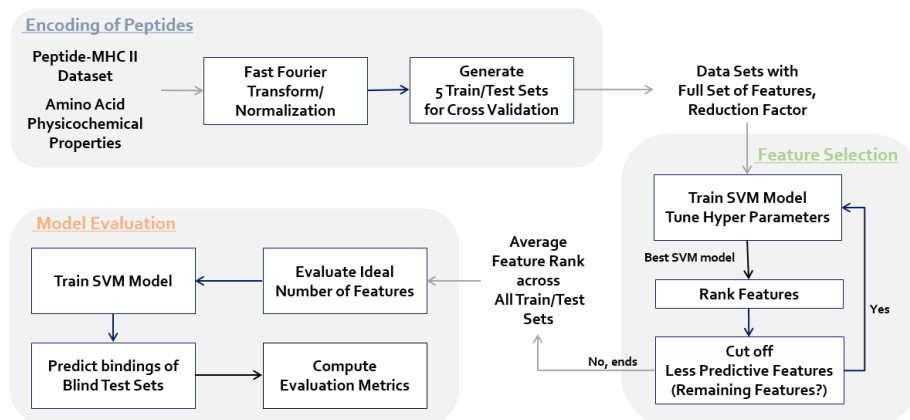
Current state-of-the-art binding prediction tools typically utilize artificial neural networks trained on sequence data with experimental binding affinity measurements. Although those sequence-based neural network models perform highly accurate predictions, they are hardly interpretable, limiting both biological insights and analysis of feature contributions. An alternative approach is machine learning methods such as support vector machine (SVM), which can be more interpretable and generalizable across peptide-MHC interactions.

One other limitation of previously developed models is that most do not incorporate information regarding the physicochemical properties of amino acids and, even when they do, they ignore the order of residues, which is known to be closely related to protein structure and function. Previously, it has been shown that families of protein structure and function can be categorized by underlying periodicities in protein or peptide physicochemical properties along the amino acid sequence (Eisenberg et al. 1984; Rackovsky 1998). Based on this observation, as well as numerical tricks such as fast Fourier transforms and zero-padding, these oscillations of physicochemical properties can be incorporated into prediction models, while also allowing the encoding of peptide sequences with varying lengths into the same number of features. A recent study on predicting antiviral peptides has shown the potential of using such Fourier-based coefficients as features for SVMs (Kieslich et al. 2021).

Here, we have developed allele-specific SVM models for the prediction of pMHC binding using the oscillations of physicochemical properties along the sequence of peptides. Our models were trained to classify binders and non-binders using supervised learning based on quantitative binding affinity experimental data. Furthermore, a previously developed feature selection algorithm was applied to rank the importance of the derived features to further optimize the developed models.

## **2. Methods**

An overview of the model development procedure used in the presented work is shown in Figure 1. All steps of model development and analysis which include data preprocessing, feature selection, training of SVMs, and model evaluation are performed using Python programming language. Each step is described in detail in the sections that follow.



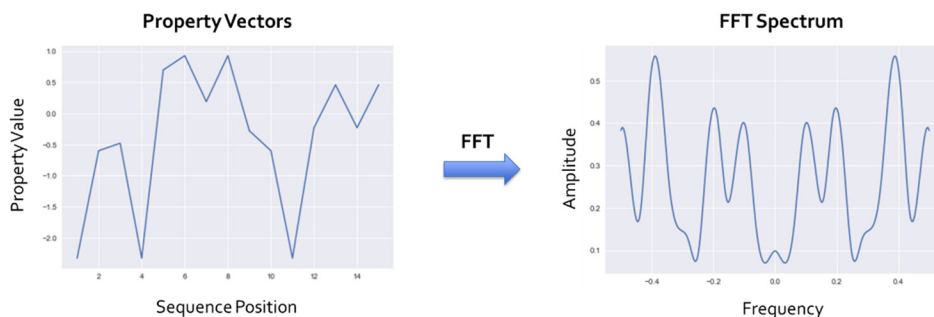
**Figure 1.** Overview of model development and evaluation procedures.

### 2.1. Peptide-MHC Class II Binding Data

The human MHC dataset was sourced from the Immune Epitope Database (IEDB) on Sep 30, 2022 (Fleri et al. 2017). Each peptide-MHC entry in the dataset consists of the peptide sequence, the MHC allele, and the measured binding affinity. For our dataset, only the quantitative type of measurements (e.g., IC<sub>50</sub>, EC<sub>50</sub>, and K<sub>D</sub>) were chosen, and entries with measurement values larger than 50000 nM were excluded from the dataset. Note that a lower value indicates higher binding affinities and vice versa. We also limited the length of peptides to range from 13 to 21 amino acids. The resulting binding affinity datasets consisted of 60630 entries with 44 unique MHC alleles and 10806 unique peptides. A binding affinity cutoff was set to 1000 nM to distinguish between binders and non-binders, which has been widely used for MHC class II binding classification. The final dataset was split into a training set and a blind test set for the evaluation of model performance with a ratio of 80:20 in a stratified sampling manner. The training set was then split again to generate 5 cross-validation sets of training/testing sets using the same approach.

### 2.2. Encoding Strategy

The first step in encoding peptide sequences is to convert them into numerical values. In this work, we adopt 10 physicochemical property factors that have been shown to describe 86% of the variance of the original 188 properties from which they were derived (Kidera et al, 1985). For each physicochemical property factor, a given peptide is converted into a corresponding physicochemical property vector (Figure 2, left), resulting in 10 physicochemical property vectors per peptide. To ensure that every peptide is represented by the same number of features, the 10 physicochemical property vectors are zero-padded assuming a maximum sequence length of 32 amino acids. Then, a Fourier transform is applied to each property vector using the Fast Fourier Transform (FFT) function of the SciPy package in Python, and we took the moduli (Figure 2, right) of the complex Fourier coefficients for frequency values between 0 and 0.5 as final features for training the SVMs which consists of 170 elements.



**Figure 2.** Fourier transform of property vectors into FFT Spectrum.

### 2.3. Support vector machines

We trained a total of 44 allele-specific binary classification SVM models based on the nonlinear radial basis kernel using the SVC function in the Python Scikit-learn package. For each SVM model, the hyper-parameters cost (C) and gamma (g) are tuned by grid-search, with values of  $10^n \forall n \in \{-3, \dots, 3\}$ , where n is an integer. In validating the SVM models, performance was assessed with 5-fold cross-validation with each training/testing set containing equal ratios of binders to non-binders. The accuracy of the SVM models was evaluated by computing classification accuracy, which is computed as to the percentage of entries in a testing set where the predictions were correct.

### 2.4. Feature selection

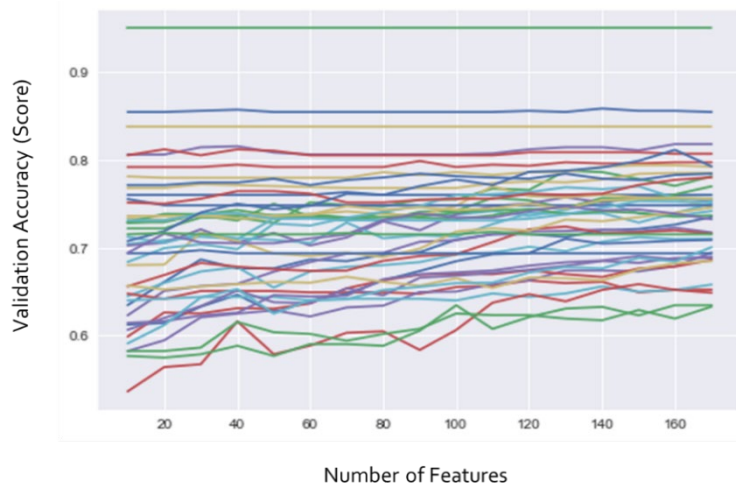
Feature selection, also known as variable selection, is a process by which the number of features is reduced to a subset of predictive features by eliminating redundant or irrelevant ones when developing prediction models. The process benefits machine learning tasks by decreasing the chances of over-fitting while potentially improving its accuracy and reducing training time. In this work, we have adopted a feature selection algorithm based on non-linear SVMs, which was proposed by Kieslich et al. (Kieslich et al. 2016). The algorithm has been applied to the broad area of prediction study including HIV-1 viral entry and fault detection in chemical plants (Kieslich et al. 2016, Onel et al. 2018, Onel et al. 2019).

The algorithm uses a greedy approach and ranks the features based on the criterion (Eq. 1), which characterizes the importance of feature k. The criterion was derived from the objective function of the dual formulation of the SVM model since solving the formulation, in general, is required to employ non-linear kernels for feature selection.

$$crit_k = -\frac{1}{2} \sum_i \sum_j \alpha_i^* \alpha_j^* y_i y_j \left. \frac{\partial K(x_i \circ z, x_j \circ z)}{\partial z_k} \right|_{z=1} \quad (1)$$

The algorithm initially starts with an SVM model with all of the features and computes the criteria for each feature. Then the features are ranked based on their criteria values, and given the feature reduction factor, the algorithm iteratively removes a set of features

that has the largest criteria values. The presented work used the reduction factor of 1/3, meaning 33% of the remaining features are removed after each iteration. Hyper-parameters are tuned again based on the remaining features after each iteration, and a consensus feature rank is also generated by averaging ranks across the five training sets.



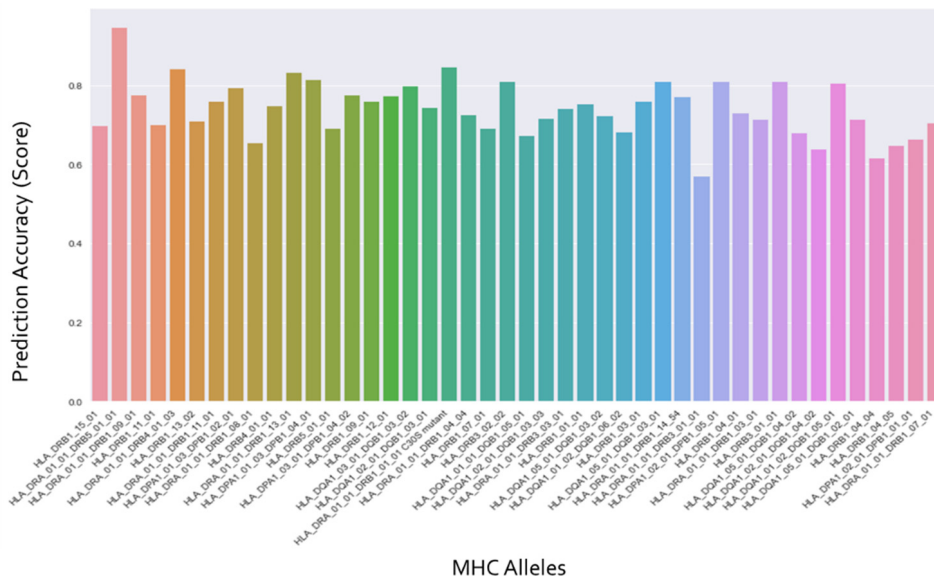
**Figure 3.** Feature contribution of 44 MHC alleles obtained from feature selection.

### 3. Results

The feature selection analysis was implemented for 44 allele-specific SVMs based on the generated physicochemical features. The feature selection algorithm provided the feature rank of each MHC allele model, and the contribution of each feature was measured by training and testing individual SVMs with five-fold cross-validation after adding 10 features at a time. The validation accuracies were computed as the mean accuracy of test set prediction over cross-validation with best hyper-parameters. As can be seen in Figure 3, overall, there was a trend that the addition of features leads to a steady increase in validation accuracy across most of the MHC alleles. Those models achieve the accuracies of 0.53-0.94 only with the first 10 features, and they improve up to 0.62-0.94 as we incorporate more features. In general, maximum accuracies were achieved when almost all features are used to train each model. Meanwhile, horizontal lines are also observed for some alleles, where we can infer that the models are biased due to the limited amount of data and, for the most part, are the result of largely imbalanced classification datasets.

For each MHC allele, the optimal number of features was selected from the above feature selection analysis (Figure 3). The final allele-specific SVM classification models were trained and optimized with those features identified as minimum features required for maximum validation accuracy and then evaluated by performing prediction on the blind dataset. As shown in Figure 4, the prediction accuracy is relatively consistent across the feature-optimized allele-specific models, ranging from 0.58- 0.93 with an average of 0.73. Overall, those models exhibit sound prediction performance, indicating the use of physicochemical property-derived features enables SVMs to capture the interactions between peptides and MHC molecules.





**Figure 4.** Blind set prediction accuracy of the feature-optimized allele-specific SVM models.

#### 4. Conclusions

In this work, we have developed 44 allele-specific binary classification SVM models that were trained based on the binding affinity datasets to classify MHC binding and non-binding peptides. Features for the SVMs were derived from the physicochemical properties of amino acids and extracted from Fourier analysis, assuming the oscillation of properties along the sequence can be predictive for peptide-MHC binding. Nonlinear SVM-based feature selection analysis was also performed during the workflow to measure the contributions of features and identify an optimal list of features. As a result, the proposed approach has shown the potential of physicochemical properties derived features in the prediction of peptide-protein interactions or functions. Future work will need to investigate the generalizability and interpretability of the prediction models for a better understanding of pMHC interactions across MHC alleles. The developed models with further analysis could be used to provide a guideline for developing and designing peptide vaccines and immunotherapeutics.

#### References

- Eisenberg, D., Weiss, R. M., & Terwilliger, T. C. (1984). The hydrophobic moment detects periodicity in protein hydrophobicity. *Proceedings of the National Academy of Sciences*, 81(1), 140-144.
- Fleri, W., Paul, S., Dhanda, S. K., Mahajan, S., Xu, X., Peters, B., & Sette, A. (2017). The immune epitope database and analysis resource in epitope discovery and synthetic vaccine design. *Frontiers in immunology*, 8, 278.  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5348633/pdf/fimmu-08-00278.pdf>

- Kidera, A., Konishi, Y., Oka, M., Ooi, T., & Scheraga, H. A. (1985). Statistical analysis of the physical properties of the 20 naturally occurring amino acids. *Journal of Protein Chemistry*, 4(1), 23-55.
- Kieslich, C. A., Alimirzaei, F., Song, H., Do, M., & Hall, P. (2021). Data-driven prediction of antiviral peptides based on periodicities of amino acid properties. In *Computer Aided Chemical Engineering* (Vol. 50, pp. 2019-2024). Elsevier.
- Kieslich, C. A., Tamamis, P., Guzman, Y. A., Onel, M., & Floudas, C. A. (2016). Highly accurate structure-based prediction of HIV-1 coreceptor usage suggests intermolecular interactions driving tropism. *PloS one*, 11(2), e0148974.  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4747591/pdf/pone.0148974.pdf>
- Onel, M., Kieslich, C. A., Guzman, Y. A., Floudas, C. A., & Pistikopoulos, E. N. (2018). Big data approach to batch process monitoring: Simultaneous fault detection and diagnosis using nonlinear support vector machine-based feature selection. *Computers & chemical engineering*, 115, 46-63.
- Onel, M., Kieslich, C. A., & Pistikopoulos, E. N. (2019). A nonlinear support vector machine-based feature selection approach for fault detection and diagnosis: Application to the Tennessee Eastman process. *AIChE Journal*, 65(3), 992-1005.  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7202572/pdf/nihms-1010061.pdf>
- Rackovsky, S. (1998). "Hidden" sequence periodicities and protein architecture. *Proceedings of the National Academy of Sciences*, 95(15), 8580-8584.
- Sanchez-Trincado, J. L., Gomez-Perosanz, M., & Reche, P. A. (2017). Fundamentals and methods for T-and B-cell epitope prediction. *Journal of Immunology Research*, 2017.



# A Computational Pipeline to Optimize 3D Scaffolds for Cancer Immunotherapy

Lucy Todd,<sup>a</sup> Matthew Chin,<sup>a</sup> Marc-Olivier Coppens,<sup>a\*</sup>

<sup>a</sup> *Centre for Nature Inspired Engineering & Department of Chemical Engineering,  
University College London, Torrington Place, London WC1E 7JE, United Kingdom*

*\*Email of the corresponding author: [m.coppens@ucl.ac.uk](mailto:m.coppens@ucl.ac.uk)*

## Abstract

Cancer diagnostic and treatment research continue to expand at unprecedented rates with new and promising discoveries being published regularly. However, within the realm of exciting cancer treatment improvements often lays the overlooked need for streamlined engineering systems and efficient data analytical methods. This research looks to achieve both in the framework of improving the *ex vivo* T-cell culturing environment required for Adoptive Cell Transfer (ACT) cancer immunotherapy. A unique set of Nature Inspired Solutions (Coppens, 2021) was applied to design and analyze 9600, unique, 3D Voronoi scaffolds as potential cell culturing environments. The streamlined design process involved a complex combination of computer languages and platforms and yet, was able to design and calculate various physical parameters of the 9600 scaffolds in just a few hours. To efficiently analyze and process the scaffolds, a novel systems engineering web-based networking software (the 4D Kauffman Fitness Network) was designed. This software integrates both key geometrical characteristics of the scaffolds (such as porosity, surface area and average mean curvature) and biological aspects related to the cells (such as proliferation rate). During testing, this software efficiently processed the 9600 scaffolds in under 20 s, designing a wide variety of networks which filtered the scaffolds selected for future wet lab testing. The highly adaptable and accessible nature of this software provides multiple potential applications in a wide range of scientific, data management and business sectors.

**Keywords:** Cancer Immunotherapy, Process Management, Network Design

## 1. Introduction

### 1.1. Background

There are very few people today who have not been affected by a cancer diagnosis either personally or through a loved one. Cancer can often appear as a growing, undefeatable threat. However, the portfolio of life-saving treatments continue to expand with ever increasing success rates. One of the most promising cancer immunotherapy treatments today is the Chimeric Antigen Receptor (CAR) T-cell Adoptive Cell Transfer (ACT). During this treatment, a patient's own immune-cells (T-cells) are extracted, proliferated, and activated *ex vivo*, genetically engineered to attack the specific tumour cells and then re-infused back into the patient. One of the main challenges with this treatment is the large number of activated T-cells required to proliferate (Jin et al., 2021). Currently, *ex vivo* T-cell culturing environments are often 2D, which neglect to facilitate the 3D mechanical and physical interactions required for efficient T-cell proliferation and activation (Jensen & Teng, 2020). We posit that a holistic, systems approach is essential.

By applying the Nature Inspired Solution (NIS) methodology developed at the UCL Centre for Nature-Inspired Engineering (Coppens, 2021), over 9000 unique, lymph node inspired, small-world, 3D T-cell culturing scaffolds have been designed.

A small-world network is a graph which requires a small number of steps to progress from one node to any other node within the network. This small-world feature is mathematically demonstrated through sigma ( $\sigma$ ), a combination of the path length between nodes and a clustering coefficient, where  $\sigma > 1$  shows a small-world network (Humphries et al, 2008). Examining the network topology of the fibroblastic reticular cell network (FRCN) found in the lymph nodes, Novkovic et al. demonstrated how the FRCN had  $\sigma > 1$  and a number of vertices approximated by the following equation:

$$v(6.52 * 10^{-5}) = x \quad [1]$$

where  $v$  was the volume of the design space and  $x$  was the approximate number of vertices (Novkovic et al., 2016). Therefore, a small-world network, 3D scaffold was designed. The breadth of scaffolds examined as potential cell culturing environments was then expanded upon to efficiently streamline the design and analysis process. A Voronoi pattern was chosen as the next structural type to be examined due to the similarities in network structure to the FRCN found in the lymph node.

### 1.2. Voronoi Structure and Lloyd's Algorithm

Voronoi diagrams are designed by populating a defined space (such as a square or cube) with a certain number of points. Lines are then drawn at average distances between adjacent points resulting in a collection of Voronoi cells. For this research, twenty orientations (seeds) of each of the 3D Voronoi scaffolds generated from 5-17 points were designed. Lloyd's "relaxation" algorithm was then applied to each of those scaffolds to produce a larger set of structures with differing geometric and topological parameters.

Lloyd's algorithm is a popular classification algorithm that was first designed by Stuart P. Lloyd in 1982 (Lloyd, 1982). This algorithm has since been applied in a wide variety of fields, including astronomy (Ordovás-Pascual & Sánchez Almeida, 2014) and biology (Ralf-Herwig et al., 1999). Applied to the designed Voronoi scaffolds, Lloyd's algorithm iteratively averages the volume of each Voronoi cell within the 3D structure and repositions the Voronoi generating point to the centre of the average volume. Over the course of a number of iterations, a "relaxed" state is achieved where no further iterations will shift the Voronoi's generating points. For this research, 40 iterations were selected for each Voronoi scaffold; very little change was noticed after about 15 iterations. Therefore, the combination of 12 different selections of Voronoi generating points, each with 20 different seeds, each with 40 Lloyd iterations, resulted in the 9600 Voronoi scaffolds designed.

### 1.3. The Streamlined Design Process

The process involved in designing and calculating the various geometric and topological parameters for each of the 9600 scaffolds required a combination of code written in the Grasshopper plugin of the Rhino software (McNeel, 2010), Python and R. This highly complex, multi-step procedure was efficiently streamlined to be capable of designing and calculating around 10 different geometric and topological parameters for 9600 Voronoi scaffolds in just a few hours. This process along with the specific parameters calculated will be discussed further in Section 2.

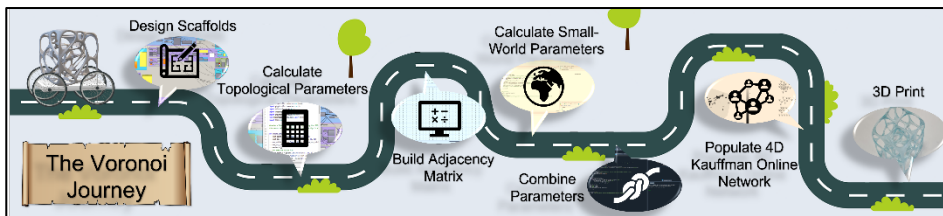
### 1.4. The 4D Kauffman Fitness Network

The pure magnitude of scaffolds being designed required an innovative computational method to analyze and process the parameters calculated for each of the scaffolds. Therefore, a novel, web-based data analytical method, inspired by Stuart Kauffman's NK fitness landscape, was designed (Kauffman, 1993). Kauffman visually demonstrated the likely evolutionary path of a series of simple proteins by stating each protein as a node in a network where edges were drawn between proteins of similar amino acid composition (Kauffman, 1993). Fitness parameters (such as the life span of the protein) were listed for each protein, such that, for any protein selected, it was possible to track the various evolutionary paths likely to occur for a protein to achieve a higher fitness and evolutionary state. Applying this methodology, where scaffolds were treated as the nodes and the various parameters calculated as the criteria by which the edges were drawn, a 4D Kauffman Fitness Network was designed. During testing, this software efficiently and successfully processed the 9600 scaffolds, designing a wide selection of networks which filtered the scaffolds selected for wet lab testing. This software, along with how it was applied, will be discussed in more detail in Section 3.

## 2. The Scaffold Design Process Journey

### 2.1. The Map

Every successful journey requires an informative map, and such was the case for these Voronoi scaffolds. Figure 1 demonstrates the various stages of the Voronoi scaffold's design and calculation process. The following sections will provide further information into the process systems required for each of the steps visualized in Figure 1.



**Figure 1:** A map of the various steps taken during the design of the Voronoi scaffolds.

### 2.2 The Design

Voronoi structures, based off varying numbers of generating points, were iterated 40 times using Lloyd's algorithm to create a collection of similar and yet unique Voronoi structures. This design code was originally written in a visual programming language (VPL). However, the VPL code included over 50 separate components and, as Table 1 demonstrates, the computational process time was considerable in view of the large number of Voronoi structures to be designed. Therefore, this code was transferred into a Python plugin component within the VPL which was subsequently parallelized. Traditionally, Grasshopper executes code on a single thread, therefore, the parallelization of the code had to be completed through a Python plugin package which was capable of multi-threading the code. Parallelizing within the Python plugin of Grasshopper initially provided a considerable challenge as the Python version within Rhino is locked in at version 2.7. It was also important to distinguish at what point the code could be parallelized as the Lloyd algorithm iterations required the Voronoi pattern from the previous iteration. However, once the code was adequately parallelized over the 64 CPU cores, this resulted in a 2- to 3-fold increase in efficiency (as shown in Table 1) which

ensured 9600 Voronoi scaffolds were designed and their parameters calculated, in just over an hour.

Coding Type	Number of Structures Generated	Time (s)	Avg. Time per Structure (s)
<b>VPL</b>	12	17.5	1.49
<b>Python (code parallelized)</b>	12	6.2	0.52
<b>VPL</b>	40	55.5	1.39
<b>Python (code parallelized)</b>	40	19.3	0.48
<b>Python (code parallelized)</b>	800	477.3	0.60

**Table 1:** The average time the computer required to complete the various parallelized and non-parallelized codes in seconds. Three trials of each calculation were completed with the average times appearing in the table above. Visual programming language has been abbreviated to VPL.

### 2.3 Mathematical Properties

In order to comprehensively understand and examine the scaffolds, a wide selection of the geometric and topological features of the Voronoi scaffolds were calculated. These included the porosity, surface area, average mean curvature, the number of generating points, the number of vertices, the number of edges, a topological parameter (the average number of edges connected to the vertices), Euler's characteristic and the minimum, maximum and average edge length. All of these were included in the Python code shown in Table 1 allowing any initial calculation errors to be easily changed and re-calculated for the thousands of scaffolds. This would also allow any future geometric, topological or physical parameters to be added to the code and efficiently calculated for all the scaffolds. In addition to this list of calculations, an adjacency matrix was also calculated for a selection of the scaffolds. This adjacency matrix was required to calculate the small-world parameter of sigma ( $\sigma$ ) for each of the scaffolds. Table 2 shows an example of the parameters calculated for one of the Voronoi scaffolds.

Points	Seed	Iteration	Min EL	Max EL	Avg. EL	SA
13	5	0	0.0263 $\mu\text{m}$	86.7 $\mu\text{m}$	31.5 $\mu\text{m}$	60600 $\mu\text{m}^2$
Vertices	Edges	Euler	Topology	Porosity	Sigma	AMC
65	126	14	3.88	83.5%	2.13	-0.0751

**Table 2:** The various topological and geometric parameters calculated for a Voronoi structure generated from 13 points before any Lloyd algorithm iterations. The abbreviated titles are edge length (EL), surface area (SA) and average mean curvature (AMC).

## 3. 4D Kauffman Fitness Network

### 3.1. The Design

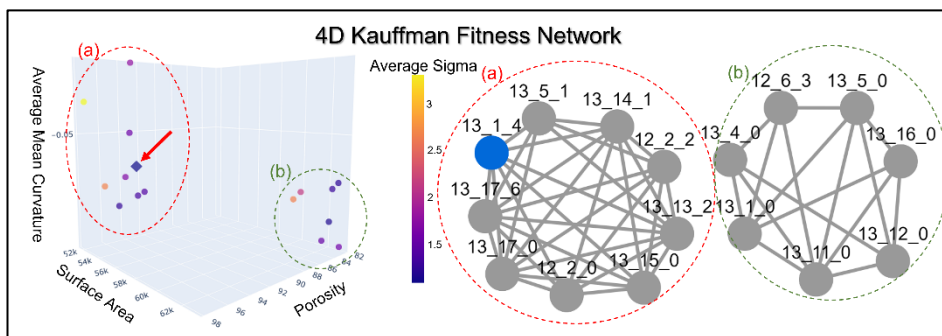
As discussed in Section 1.4, a networking software was designed to assist in analyzing and processing the 9600 Voronoi scaffolds. Designing this networking software required satisfying a variety of user and systems parameters. Firstly, the software needed to be capable of handling 9600 unique environments with around 10 separate topological characteristics. Secondly, the platform needed to allow for user integration and adaptability so networks and graphs could be re-designed based off prioritizing different geometric parameters. Thirdly, these new iterations of the networks and graphs needed to be effectively processed and displayed for user ease and efficiency. Finally, this data

needed to be easy to export in a workable format. To satisfy these various parameters, an online, web-based platform written in Python was selected. The web-based environment was chosen as this would be highly user friendly and allow the software to be easily applied in other sectors outside of computational research.

Initially the Python code, cycling through the dozen Excel files containing the 9600 scaffolds, was unable to produce an online network capable of redesigning, based off of user inputs, without crashing. It was only when each scaffold was redefined as a unique node, with its geometric and topological characteristics listed within the node's dictionary, that the software began to operate smoothly. This was likely due to each scaffold being simplified to one unit (a single node) rather than a series of 10 separate data units (1 for each of the characteristics). This node data was also highly efficient in being converted back into a data frame to produce a downloadable Excel file (something it completed in a matter of seconds).

### 3.2. The Application

Although the 4D Kauffman Fitness Network was capable of processing thousands of scaffolds, data gleaned from the graphs and networks produced, resulted in a significant filtering of the scaffolds. Of the 9600 Voronoi scaffolds designed, 952 contained  $65 \pm 2$  vertices (Equation 1) and only 73 of those had  $\sigma$  values greater than 1 (small-world networks). These 73 were further narrowed down to 36 “unique” scaffolds, as scaffolds designed from later Lloyd iterations were often indistinguishable and, therefore, the software combined them into a single node. Figure 2 demonstrates one of the generated online web-pages produced from the 4D Kauffman Fitness Network. Each node in the network on the right is a scaffold that corresponds to a point in the graph on the left. Two separate clusters can be clearly visualized in both the network and the graph demonstrating which scaffolds have similar surface area, average mean curvature and porosity. Network cluster (a) and (b) correspond to the circles (a) and (b) in the graph.



**Figure 2:** A generated 4D Kauffman Fitness Network showing 2 clusters of scaffolds in both the graph and network (indicated by the circles (a) & (b)) demonstrating the scaffolds similar in surface area, average mean curvature and porosity. The blue scaffold has been selected (“clicked-on” in the software) with its respective point highlighted as diamond, and shown with a red arrow, in the graph on the left. The labels for the nodes in the network on the right are the scaffold names (given as X\_Y\_Z) which denote the number of Voronoi generating points (X), the seed number with which the orientation of those points was saved (Y) and the number of Lloyd’s algorithm iterations (Z).



#### 4. Future Work

Future applications for both the design process and analysis software are now being explored. Firstly, the design process will be applied to other structural patterns including Lattice and Delaunay structures. Secondly, the analysis software will be utilized to select the scaffolds to be 3D printed and tested with live T-cells for cancer immunotherapy applications. Finally, the 4D Kauffman Fitness Network can be tested and applied in other scientific and business sectors where collections of similar environments or products need to be processed and graphically analyzed.

#### 5. Conclusion

The nature-inspired solution methodology applied to both the 3D T-cell culturing environments (through a lymph-node inspired scaffold design) and the software designed for the analysis of the scaffolds (inspired through Kauffman's evolutionary protein mapping) provided creative insight into the unique engineering solutions described. The computational pipeline designed and calculated 10 separate topological and geometric parameters for 9600 Voronoi scaffolds in under a few hours. These scaffolds were then filtered to 36 unique, small-world scaffolds using the novel 4D Kauffman Fitness Network with new networks and graphs being generated in a few seconds. The efficient design pipeline and analysis software provided the foundational framework to test these 3D scaffolds as *ex vivo* T-cell culturing environments for CAR ACT cancer immunotherapy. The software presented here has wider applicability for systems engineering research, beyond cancer immunotherapy.

#### References

- M.-O. Coppens, 2021, Nature-Inspired Chemical Engineering for Process Intensification, *Annu. Rev. Chem. Biomol. Eng.*, 12, 1, 187–215.
- M. Humphries, K. Gurney, 2008, Network 'Small-World-Ness': A Quantitative Method for Determining Canonical Network Equivalence, *PLoS ONE*, 3, 4.
- C. Jensen & Y. Teng, 2020, Is It Time to Start Transitioning From 2D to 3D Cell Culture?, *Frontiers in Molecular Biosciences*, 7, 33.
- Z. Jin, X. Li, X. Zhang, P. DeSousa, T. Xu, & A. Wu, 2021, Engineering the fate and function of human T-Cells via 3D bioprinting, *Biofabrication*, 13, 3.
- S. A. Kauffman, 1993, *The origins of order: self-organization and selection in evolution*, Oxford University Press.
- S. P. Lloyd, 1982. Least Squares Quantization in PCM, *IEEE Transactions on information theory*, 28, 2.
- M. Novkovic, L. Onder, J. Cupovic, J. Abe, D. Bomze, V. Cremasco, E. Scandella, J. v. Stein, G. Bocharov, S. J. Turley, & B. Ludewig, 2016, Topological Small-World Organization of the Fibroblastic Reticular Cell Network Determines Lymph Node Functionality, *PLoS Biol.*, 14, 7, e1002515.
- I. Ordovás-Pascual, & J. Sánchez Almeida, 2014, A fast version of the k-means classification algorithm for astronomical applications, *Astronomy and Astrophysics*, 565.
- A. Ralf-Herwig, J. Poustka, C. Müller, C. Bull, H. Lehrach, & J. O'Brien, 1999, Large-Scale Clustering of cDNA-Fingerprinting Data, *Genome Research*, 9, 11, 1093.
- R. McNeel, 2010, *Rhinoceros 3D, Version 6.0*, Robert McNeel & amp, Associates, Seattle, WA.

# Developing Cole-Cole model for *Bacillus subtilis* fermentation

Kumar Rajan Gopa<sup>a</sup>, Tom Wenzel<sup>a</sup>, Sankalp Jena<sup>a</sup>, Supasuda Assawajaruwan<sup>b</sup>,  
Valentin Khaydarov<sup>a</sup>, Leon Urbas<sup>a</sup>

<sup>a</sup>Technische Universität Dresden, Dresden 01062, Germany

<sup>b</sup>Evonik Operations GmbH, Rodenbacher Chaussee 4, 63457 Hanau, Germany

[kumar.rajan\\_gopa@tu-dresden.de](mailto:kumar.rajan_gopa@tu-dresden.de)

## Abstract

During the fermentation of *Bacillus Subtilis* it is unperceivable to observe the formation of spores since, following the evolution of spores is time consuming due to various sampling and analytical procedures. Dielectric spectroscopy has shown promising results, where the measured permittivity is correlated with the cell concentration. Most of the existing works focus on determining cell concentration and not on the determination of spores. In this work, we develop a permittivity model for cells in the fermentation solution based on the Cole-Cole equations that can be used to detect the sporulation phase in future work. Utilizing the data from several fermentation runs of *Bacillus Subtilis*, a superposition hypothesis is proposed. In this hypothesis, the model equations from Cole and Cole (1941) are used and the theoretical permittivity response of the cells is calculated. As a first step, the Cole-Cole model is developed, and the modeled permittivity values are fitted to the measured data and the results are analyzed.

**Keywords:** Cole-Cole, Permittivity, Sporulation, Superposition

## 1. Introduction

*Bacillus Subtilis* is an industrially significant gram-positive spore-forming bacteria which is used in the large-scale production of various bioproducts (Abi *et al.*, 2010). The biological system along with the spore formation and evolution mechanism is complex and difficult to model. In addition, the drawbacks of offline measurement like longer and infrequent sampling make the online monitoring and detection of spores highly desirable. One of the most promising online monitoring techniques is the dielectric spectroscopy, especially in its scanning frequency mode. There have been several works where the permittivity measurements have been found to reveal information related to changes in morphological phases. Opel, Li and Amanullah (2010) and Randek and Mandenius (2020) found the influence of cell sizes on the permittivity data.

In general, there are several modeling approaches applied on the permittivity data obtained from dielectric spectroscopy. Among them Konakovsky *et al.* (2015) used principal component analysis to predict the viable cell concentration (VCC). Dabros *et al.* (2009) applied Cole-Cole modeling and compared the results with linear and multivariate regression models. These works establish Cole-Cole modeling as a viable method for modeling cell concentrations from permittivity data.

Since the bacterial spores are much smaller and their presence is difficult to detect, works on the application of dielectric spectroscopy and identifying the presence of spores are rarely present. Abi *et al.* (2010) focused on the *Bacillus* species and identified the

physiological phases using the scanning frequency data. Studied parameters are capacitance, conductivity, and Cole-Cole parameters to monitor and predict the effect of operating conditions. In our study to determine the effect of spores on the permittivity data, we have applied a superposition hypothesis in conjunction with Cole-Cole modeling. This methodology is expected to result in a residual permittivity that can be indicative of the spore dynamics. This will enable the online monitoring of sporulation and to gain insight about the current fermentation phase in progress thereby reducing the time and effort required for sampling and other analytical procedures.

## 2. Theory

During the fermentation process of *Bacillus Subtilis*, the bacteria present in the culture consume the nutrients and undergo cell division. This is known as the vegetative cycle. In the initial period of fermentation, the culture broth has enough nutrients, resulting in exponential growth of bacteria. Once the nutrient depletion occurs, the cells cannot maintain the exponential growth rate. Normally this would result in a slowdown and the cells which are deprived of the nutrients will start dying. This phase is known as the stationary phase. In case of bacterial species like *Bacillus Subtilis*, a depletion of nutrients or any other adverse condition leads to a secondary cycle. This is called the sporulation cycle where a part of the population preserves its genetic material by forming dormant spores. In the last phase, a part of the vegetative cell population may convert itself into spores and the remaining cell population undergoes cell lysis. This phase is known as the death phase.

### 2.1. Dielectric spectroscopy

Over the years dielectric spectroscopy has been identified as a viable method for the online detection of cells (Teixeira *et al.*, 2009). It relies on the principle that under the influence of an electric field an exposed medium can act as an electric capacitor – in this case the cells and spores. This permittivity is measured by the probe. Since the amount of electric charge stored is related to the cell volume and the cell size this is a sufficient indicator on detecting the cells (Markx and Davey, 1999).

Dielectric spectroscopy could be conducted in two modes: 1. Dual frequency mode where a single frequency and a baseline frequency are used in parallel. 2. The scanning frequency mode where the permittivity is measured over a range of frequencies called the beta frequency (Dabros *et al.*, 2009). In the present case we use scanning frequency mode that consists of 25 frequencies ranging from 50 kHz to 20000 kHz that are scanned in one time step.

### 2.2. Experimental

*Bacillus subtilis* strain 168 (DSM 23778) was applied in this study. The inoculum was prepared in CASO-Bouillon as a pre-culture medium. Then it was inoculated into the minimal medium containing glucose as a carbon source. Batch cultivation was carried out in a 2-L stirred tank reactor with 1-L working volume. The 100 mL pre-culture was incubated at 30°C for 8 h and then it was inoculated into the bioreactor. The fermentation was run for 48 h at different temperature profiles. The bioreactor was equipped with dielectric sensor (Futura, Aber) to gain an insight of fermentation in real time. Temperature and glucose levels were changed for each experiment, with glucose levels in the range of 12.5-20 g/l and temperature levels in the range of 25 °C to 45°C. For some

experiments multiple temperature changes were carried out during the course of the fermentation process. The following data were used:

1. Spectral data – Permittivity measurement with sampling rate of around 2 per minute and spans over 25 frequencies from 50 kHz to 20000 kHz.
2. Offline data – Cell/spore count and the substrate added. About 14 samples.

Maskow *et al.* (2008) have pointed out that other intracellular products could be detected from the dielectric spectroscopy signals. We expect spores to be such products that could be detected. We hypothesize that the residual permittivity between the measured and (cell-) modeled data can provide information about the spore concentration.

$$\varepsilon_{\text{sensor}}(\omega) = \varepsilon_{\text{cells}}(\omega, N_{\text{cells}}) + \varepsilon_{\text{spores}}(\omega, N_{\text{spores}}) + u \quad (1)$$

Where  $\varepsilon$  is the total permittivity of the fermentation solution [pF/cm],  $\omega$  is the angular frequency [rad/s],  $N$  is the number density of either cells or spores [CFU/ml], and  $u$  is an unknown factor. An Illustration of the proposed superposition hypothesis is shown in Figure 1 on the left side. We propose that the sensor data can be modeled and fitted during the growth phase ideally since there is no contribution of spores to the permittivity signal. To model the permittivity of the cells we use a spherical cell model without cell size distribution. These assumptions could lead to a model error, since the cells are not ideally spherical, and their size is distributed. In order to compensate this model error, a correction factor  $\theta_0$  is introduced. We assume this linear correction factor as a first approach. The permittivity of the cells is then described as the product of the spherical cell model and the correction factor:

$$\varepsilon_{\text{cells}}(\omega) = \theta_0 \varepsilon_{\text{model}}(\omega, N_{\text{cells}}) \quad (2)$$

Using the approach from Maskow *et al.* (2008) the growth phase can be defined as the time span until the maximum of the permittivity is reached (see Figure 1). Following the equation system for Cole-Cole model present in Dabros *et al.* (2009), the permittivity equation (3) is used.

$$\varepsilon_{\text{model}}(\omega, N_{\text{cells}}) = \Delta\varepsilon \frac{1 + \left(\frac{\omega}{\omega_c}\right)^{1-\alpha} \sin\left(\frac{\pi}{2}\alpha\right)}{1 + \left(\frac{\omega}{\omega_c}\right)^{2-2\alpha} + 2\left(\frac{\omega}{\omega_c}\right)^{1-\alpha} \sin\left(\frac{\pi}{2}\alpha\right)} + \varepsilon_{\infty} \quad (3)$$

Where  $\Delta\varepsilon$  is the span in which the permittivity signal varies,  $\varepsilon_{\infty}$  is the permittivity at infinite frequency,  $\omega_c$  is the critical angular frequency where half of  $\Delta\varepsilon$  is reached and  $\alpha$  is the empirical Cole-Cole factor. The Cole-Cole  $\alpha$  for biological cells was reported to be in the range 0.1 – 0.2 (Dabros *et al.*, 2009). In the current work, a constant value throughout the batch process of 0.1 is used. The calculation of the constant  $\Delta\varepsilon$  is based on the Pauly-Schwan spherical cell model.

$$\Delta\varepsilon = \frac{3 N_{\text{cells}} \pi r_{\text{cells}}^4 C_m}{\varepsilon_0} \quad (4)$$

Where  $r_{\text{cells}}$  is the volume equivalent diameter of the cells,  $C_m$  is the membrane permittivity, and  $\varepsilon_0$  is the vacuum permittivity.  $C_m$  is assumed to 0.01 F/cm<sup>2</sup> Yardley *et al.* (2000). For the rod shaped bacteria the radius of the volume equivalent idealized sphere of 0.78  $\mu\text{m}$  was used (Errington and Aart, 2020). As suggested by Downey *et al.* (2014) an exponential growth is expected and validated with offline measurements.

$$N_{\text{cells}}(t) = N_{\text{cells},t=0} e^{kt} \quad (5)$$

### 3. Results and discussion

#### 3.1. Implementation of the Cole-Cole-model

In our approach we use the assumptions mentioned to determine the value of  $\Delta\varepsilon$ . For  $\alpha$  a value of 0.1 is assumed. As shown in Figure 1 (right), the values for  $\omega_c$  and  $\varepsilon_\infty$  are taken from the graph. For  $\varepsilon_\infty$  the permittivity value for the highest frequency is used. This procedure assumes that there is no significant change in the permittivity when the frequency is increased. To calculate  $\Delta\varepsilon$  information cell size and number are required. About 13 offline measurements were obtained during a 50-hour fermentation batch. To increase the number of the data points exponential curve fitting during the growth phase is applied.

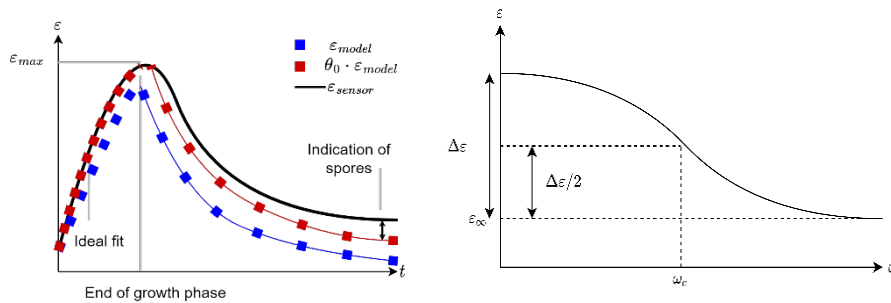


Figure 1: Illustration of the superposition hypothesis (left) and graphical representation of the Cole-Cole parameters (right).

#### 3.2. Fitting part – growth phase

Three batches KW1, KW2 and KW3 were considered for analysis. For KW1, the starting conditions were 12.5 g/l glucose and 30 °C temperature. No temperature change was made during the experiment. KW2 and KW3 included multiple temperature changes, with KW2 having a starting glucose concentration of 12.5 g/l and KW3 with 20 g/l. This ensured that each of the considered batch had differing batch conditions. The data set for the cell numbers can be increased when exponential curve fitting for equation (5) is used. Figure 2 compares the measured and the modelled permittivity during the growth phase at two frequencies for the batch number KW1. In a first step the Cole-Cole model  $\varepsilon_{model}$  without the fitting factor  $\theta_0$  is shown.

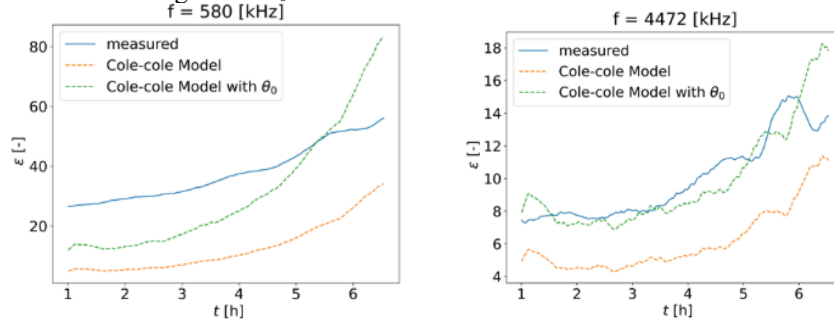


Figure 2: Permittivity data measured, model for  $g_1$  and extended model for different frequencies at KW1.

As outlined in the superposition hypothesis the model can be fitted better to the data when the factor  $\theta_0$  is introduced. It can be shown that there is a significant difference between the Cole-Cole model and the sensor data. We assume that the deviation is mostly down to either the inaccurate spherical model or the cell size assumptions. In Table 1 the corresponding values for  $\theta_0$  and the root mean square error (RMSE) before (0) and after (1) the introduction of  $\theta_0$  is shown for the fermentation batches KW1, KW2 and KW3. The proposed procedure shows that the modeled data after the introduction of  $\theta_0$  have a better fit at higher frequencies. Although all three batches had different conditions, all three achieved better fit with this method. The lowest  $RMSE_1$  and  $\theta_0$  closest to one were obtained for KW1 where no temperature variation was applied. No perfect fit for  $\epsilon_{cells}$  is achieved during the growth phase, but the mean error has been significantly reduced. The temperature change applied in the initial hours might have caused a higher  $RMSE_0$  values for KW2 and KW3. The reasons for deviations ( $RMSE_1$ ) could be due to imprecise assumptions, especially for the cell radius.

KW	1			2			3		
<b>f [kHz]</b>	580	1120	4472	580	1120	4472	580	1120	4472
<b><math>\theta_0</math> [-]</b>	2.5	2.3	1.6	4.2	3.8	1.6	7.7	-0.8	0.6
<b>RMSE<sub>0</sub> [-]</b>	25.3	18.5	4.0	34.4	23.7	3.9	38.5	21.4	6.4
<b>RMSE<sub>1</sub> [-]</b>	13.2	9.3	1.5	7.0	3.7	1.8	17.6	34.5	5.4

Table 1: RMSE values for the scalar factor  $\theta_0$  before and after optimization

In the following work, this approach will be extended to the remaining time periods of the fermentation batch till the end of the batch. The estimated  $\theta_0$  will be applied in equation 1 to obtain the residual permittivity that could be analyzed to establish the presence of spores.

#### 4. Conclusion

Cole-Cole modeling was identified as a suitable method that can be used to model the permittivity during the fermentation of bacteria. Using the superposition hypothesis and Cole-Cole modeling, an alternate model to describe the fermentation process of *Bacillus subtilis* was developed and the various components that contribute towards the total permittivity of the system were outlined. The linear correction parameter  $\theta_0$  has been used in this first approach to compensate for the inaccuracies in the cell model. It is observed that the modelled permittivity is in better agreement with the sensor data for higher frequencies. This work will be extended in the next stage to calculate the residual permittivity which might provide information on the development of spores, especially after the growth phase. Further work with increased number of offline measurements would provide a better fit with the current model. This model could be extended to determine the cell radius of the species more precisely. Since *Bacillus subtilis* is ellipsoidal in nature, models other than the spherical shell model could be used to improve the model fit.

#### Acknowledgements:

We want to thank the BMWK (German Federal Ministry of Economic Affairs and Climate Action, Support codes 01MK20014T) for their financial support.

## References

- Abi, A., Sarrafzadeh, M. H., Mehrnia, M. R., & Ghommidh, C. (2010). Application of dielectric permittivity measurements in physiological state monitoring of bacillus subtilis culture. 2010 2nd International Conference on Chemical, Biological and Environmental Engineering, 26–29. <https://doi.org/10.1109/ICBEE.2010.5650641>
- Cole, K. S., & Cole, R. H. (1941). Dispersion and Absorption in Dielectrics I. Alternating Current Characteristics. *The Journal of Chemical Physics*, 9(4), 341–351. <https://doi.org/10.1063/1.1750906>
- Dabros, M., Dennewald, D., Currie, D. J., Lee, M. H., Todd, R. W., Marison, I. W., & Stockar, U. (2009). Cole–Cole, linear and multivariate modeling of capacitance data for on-line monitoring of biomass. *Bioprocess and Biosystems Engineering*, 32(2), Art. 2. <https://doi.org/10.1007/s00449-008-0234-4>
- Downey, B. J., Graham, L. J., Breit, J. F., & Glutting, N. K. (2014). A novel approach for using dielectric spectroscopy to predict viable cell volume (VCV) in early process development. *Biotechnology Progress*, 30(2), 479–487. <https://doi.org/10.1002/btpr.1845>
- Errington, J. and Aart, L.T. van der (2020) ‘Microbe Profile: Bacillus subtilis: model organism for cellular development, and industrial workhorse: This article is part of the Microbe Profiles collection.’, *Microbiology*, 166(5), pp. 425–427. Available at: <https://doi.org/10.1099/mic.0.000922>.
- Konakovsky, V., Yagtu, A., Clemens, C., Müller, M., Berger, M., Schlatter, S., & Herwig, C. (2015). Universal Capacitance Model for Real-Time Biomass in Cell Culture. *Sensors*, 15(9), Art. 9. <https://doi.org/10.3390/s150922128>
- Markx, G. H., & Davey, C. L. (1999). The dielectric properties of biological cells at radiofrequencies: Applications in biotechnology. *Enzyme and Microbial Technology*, 25(3), Art. 3. [https://doi.org/10.1016/S0141-0229\(99\)00008-3](https://doi.org/10.1016/S0141-0229(99)00008-3)
- Maskow, T., Rölllich, A., Fetzer, I., Ackermann, J.-U., & Harms, H. (2008). On-line monitoring of lipid storage in yeasts using impedance spectroscopy. *Journal of Biotechnology*, 135(1), Art. 1. <https://doi.org/10.1016/j.jbiotec.2008.02.014>
- Opel, C. F., Li, J., & Amanullah, A. (2010). Quantitative modeling of viable cell density, cell size, intracellular conductivity, and membrane capacitance in batch and fed-batch CHO processes using dielectric spectroscopy. *Biotechnology Progress*, NA-NA. <https://doi.org/10.1002/btpr.425>
- Randek, J., & Mandenius, C.-F. (2020). In situ scanning capacitance sensor with spectral analysis reveals morphological states in cultures for production of biopharmaceuticals. *Sensors and Actuators B: Chemical*, 313, 128052. <https://doi.org/10.1016/j.snb.2020.128052>
- Teixeira, A. P., Oliveira, R., Alves, P. M., & Carrondo, M. J. T. (2009). Advances in on-line monitoring and control of mammalian cell cultures: Supporting the PAT initiative. *Biotechnology Advances*, 27(6), Art. 6. <https://doi.org/10.1016/j.biotechadv.2009.05.003>
- Yardley, J. E., Kell, D. B., Barrett, J., & Davey, C. L. (2000). On-Line, Real-Time Measurements of Cellular Biomass using Dielectric Spectroscopy. *Biotechnology and Genetic Engineering Reviews*, 17(1), 3–36. <https://doi.org/10.1080/02648725.2000.10647986>

# Improving accuracy scores of neural network driven QSAR models of mutagenicity

Alexander D. Kalian,<sup>a</sup> Emilio Benfenati,<sup>b</sup> Olivia J. Osborne,<sup>c</sup> Jean-Lou C.M. Dorne,<sup>d</sup> David Gott,<sup>c</sup> Claire Potter,<sup>c</sup> Miao Guo,<sup>a\*</sup> Christer Hogstrand,<sup>a\*</sup>

<sup>a</sup>*King's College London, Franklin-Wilkins Building, 150 Stamford St., London SE1 9NH, United Kingdom*

<sup>b</sup>*IRCCS - Istituto di Ricerche Farmacologiche Mario Negri, Via Mario Negri 2, 20156 Milano, Italy*

<sup>c</sup>*Food Standards Agency, 70 Petty France, London SW1H 9EX, United Kingdom*

<sup>d</sup>*European Food Safety Authority (EFSA), Via Carlo Magno 1A, 43126 Parma, Italy*  
<sup>\*</sup>*christer.hogstrand@kcl.ac.uk (CH); miao.guo@kcl.ac.uk (MG)*

## Abstract

Multiple QSAR models of mutagenicity were created and compared, using knowledge graph approaches to train and test multi-layer perceptron classifiers, following dimensionality reduction from several thousand dimensions to hundreds of dimensions via principal component analysis. Such knowledge graphs were built in one case using molecular fingerprint based structural similarities, while in another case using molecular fragments found via application of the Girvan-Newman algorithm. A simple hybrid model was also explored. However, both competing QSAR models performed with comparable accuracies, with both sensitivity and specificity scores for each occurring within range of 70%. The predictions of both models were in agreement in an average of 71% of cases, meaning that each could offer a related yet notably different perspective of toxicological space; hence a simple hybrid model was trialed, which only output predictions agreed between both constituent models, which averaged at 78% accuracy.

**Keywords:** QSAR, deep-learning, neural-networks, classification, toxicology

## 1. Introduction

Food and drink products contain a variety of chemicals, of which some are intentionally present (e.g. preservatives) while others are unintentionally present (e.g. environmental contaminants) [1]. The possible risks posed by consumption of food and drink chemicals to human and animal health, are ideally evaluated via chemical risk assessments (CRAs) carried out by regulatory bodies [1]. Current CRA frameworks are overly reliant on *in-vivo* studies on live animals [1], which pose significant ethical, scientific relevancy, and scalability-based concerns for the future of this space. New Approach Methodologies (NAMs), such as Quantitative Structure-Activity Relationship (QSAR) modelling, may be used to address some of these limitations and contribute to future frameworks that regulatory bodies may utilise [2].

QSAR modelling is an *in-silico* approach which relies on the premise that molecular structures are correlated with corresponding biological activities [2]; hence, in a toxicological context, QSAR models quantify relationships between toxicological properties of molecules and quantitative metrics concerning molecular structures. There is no universally agreed upon metric for reliably quantifying molecular structures [2],



however a variety of different approaches exist, such as structural similarity coefficients (SCs – which quantify structural similarities between pairs of molecules via techniques such as molecular fingerprinting [2] or alternatively applications of graph theory), as well as molecular fragmentation (of which numerous algorithms exist in order to split a molecular structure into smaller substructures, which may then be correlated with certain toxicological properties). QSAR models frequently use machine learning (ML) to construct models; in a toxicological context, this typically entails regression algorithms that can predict continuous metrics [3] (e.g. benchmark doses, for hazard assessments), however classification algorithms are more appropriate for certain endpoints that are thresholded and hence associated with discrete predictions [3] – e.g. mutagenicity is a thresholded endpoint where molecules may be assessed into discrete classes: A (strongly mutagenic), B (weakly mutagenic) and C (non-mutagenic) [4].

QSAR models of mutagenicity were compared as part of the 2014 Ames/QSAR International Challenge Project, involving a variety of rule-based QSAR frameworks such as SARpy and Toxtree, as well as statistical QSAR frameworks such as CEASAR and AMBIT. Sensitivity and specificity scores varied, however none were able to consistently surpass accuracy scores of 80% (considering both sensitivities and specificities), furthermore with many displaying imbalance across classes [4]. Neural networks (NNs) are a popular ML technique, using layers of artificial neurons with weights that may be optimised to form complex models. A 2021 study used feed-forward NNs to build a QSAR model of mutagenicity which was 84% accurate [5], while a separate 2021 study used graph convolutional NNs (GCNs) in order to obtain sensitivity scores that were consistently matching or below 70% and specificity scores that were consistently above 90% [6].

The aim of this study was to use NNs to construct and optimize novel QSAR models of mutagenicity, using data from *in-vitro* Ames mutagenicity studies. While *in-vitro* methods are NAMs that are frequently used in studies concerning mutagenicity, with certain advantages over *in-vivo* studies [2], they nonetheless continue to be limited in terms of efficiency and scalability, which *in-silico* techniques such as QSAR modelling may help address [2], as part of a potential future ecosystem of NAMs used in CRA. Furthermore, mutagenicity is a comparatively more thoroughly covered endpoint, in terms of data available in open source, as well as with literature concerning competing QSAR models [4], when compared to comparatively lesser covered endpoints such as neurotoxicity and cardiotoxicity; hence mutagenicity may serve as a more intuitive starting endpoint, for the construction and optimisation of novel QSAR models that may eventually be adapted and applied to more challenging endpoints in future.

## 2. Methods

### 2.1. SC-Based QSAR Model

The first QSAR model constructed used SCs (via molecular fingerprinting) to build a knowledge graph encoding similarities between constituent molecules of a dataset, which was then reduced in dimensionality via principal component analysis (PCA) and subsequently used to train and test a multi-layer perceptron (MLP) model.

Six SC algorithms that used molecular fingerprinting were separately used to create competing SC-based QSAR models; these were the Cosine, Tanimoto, Dice, Kulczynski, McConnaughey and Rogot-Goldberg SCs, each with their own unique methodology.

Each SC algorithm first required molecular fingerprints to be generated for each molecule, for which Morgan fingerprints were used. The Morgan fingerprinting algorithm was used as a function of the open-source cheminformatics Python library RDKit, with a specified radius of 3 and maximum number of 2048 bits. Following obtention of Morgan fingerprints for all possible pairs of molecules, the six algorithms were each used to calculate SCs between all Morgan fingerprint pairs.

The dataset of molecules used was that associated with the 2014 Ames/QSAR International Challenge Project, with Ames mutagenicity data collected from a variety of different studies, for a variety of organic molecules provided in canonical SMILES (simplified molecular-input line-entry) notation. Following data curation measures (such as excluding molecules with inconsistencies in descriptors, or with incomplete mutagenicity data), the final dataset used in this study contained 11,268 molecules. The occurrences of A, B and C classes in the data were imbalanced, which was alleviated via combining classes A and B into a single “mutagenic” class, while assigning class C as a “non-mutagenic” class (an approach similar to those used in past studies [4]). The two classes nonetheless remained imbalanced, hence further balancing was carried out via assignment of training and testing data.

The 11,268 molecules were split into approximately 80% training data and 20% testing data, using random sampling via pseudorandom numbers, with each molecule treated as a sample to be used in an MLP classifier model. Knowledge graphs for training data were constructed, using individual molecules as nodes, connected via edges to every present node (including to themselves) and weighted according to SCs between corresponding molecule pairs. Feature vectors for each training sample molecule were derived from the weightings of edges in the knowledge graph connected to the corresponding node. Feature vectors for training data similarly contained calculated edge data to the existing nodes in the knowledge graph, effectively describing how each test sample molecule would be positioned in the knowledge graph, if included.

Each feature vector was naturally of several thousand dimensions, which would have proven problematic for training effective models with simple architectures. PCA was implemented via the Python-based ML package scikit-learn, to reduce dimensionality, while conserving as much information contained within the multidimensional data as possible. Final dimensionalities were varied between 2-220 (see Figure 2), for the sake of finding an optimal dimensionality.

The NNs used were MLP classifiers, also implemented via scikit-learn in Python, with two hidden layers containing 500 artificial neurons, while using the rectified linear unit (ReLU) activation function, along with the Adam optimisation algorithm (a stochastic gradient-based algorithm, which is typically suitable for larger datasets) for weight optimisation. The strength of the L2 regularization term was set to  $1 \times 10^{-5}$ , while the number of epochs was set to 1,000. All other parameters remained as default.

## 2.2. Fragmentation-Based QSAR Model

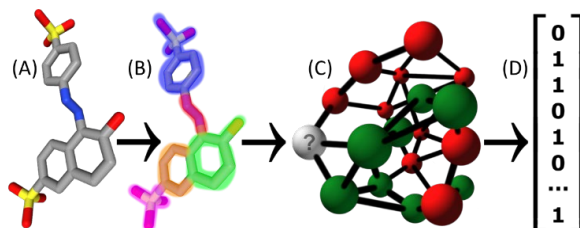


Figure 1: schematic diagram demonstrating how (A) an input molecule [7] would be (B) fragmented before (C) being placed in a graph of mutagenic/non-mutagenic molecules that shared substructures. (D) A feature vector would then be derived, based on connections.

A second QSAR model was constructed and applied to the same dataset, using knowledge graphs, PCA and MLP classifiers in an identical nature to that of the SC-based QSAR model, but instead using molecular fragmentation to construct knowledge graphs.

Fragments of molecules were obtained via consideration of canonical SMILES descriptors as molecular graphs, using graph theory (the Python library `pysmiles` was used to convert between canonical SMILES and Python Graph objects). Fragmentation was treated as a community detection problem, with communities of each molecular graph identified and regarded as molecular fragments via the Girvan-Newman algorithm; a widely used hierarchical community detection method available as a function of the `NetworkX` library of Python. The algorithm required a specified number of communities to be identified, hence a separate algorithm output ascending values for each new molecule, until all fragments found by the Girvan-Newman algorithm were within a specified size range in terms of constituent atoms (excluding hydrogen atoms). Three different size ranges were considered: 2-7, 3-12 and 3-15.

Once lists of fragments for every molecule were obtained, knowledge graphs were constructed (see Figure 1). Training and testing data assignment, as well as subsequent PCA and use of MLP classifiers, were identical to as described in Section 2.1.

### 2.3. Hybrid QSAR Model

Following training and testing of the QSAR models described in Sections 2.1 and 2.2, a simple hybrid model was explored via testing both models in parallel (with the same training data for each) and only outputting predictions that were agreed by both models. Within the hybrid model, the SC-based QSAR model used the cosine similarity, whereas the fragmentation-based QSAR model used a size range of 3-12 atoms for finding fragments. Both constituent models also used reduced dimensionalities of 100.

## 3. Results and Analysis

### 3.1. Individual QSAR Model Results

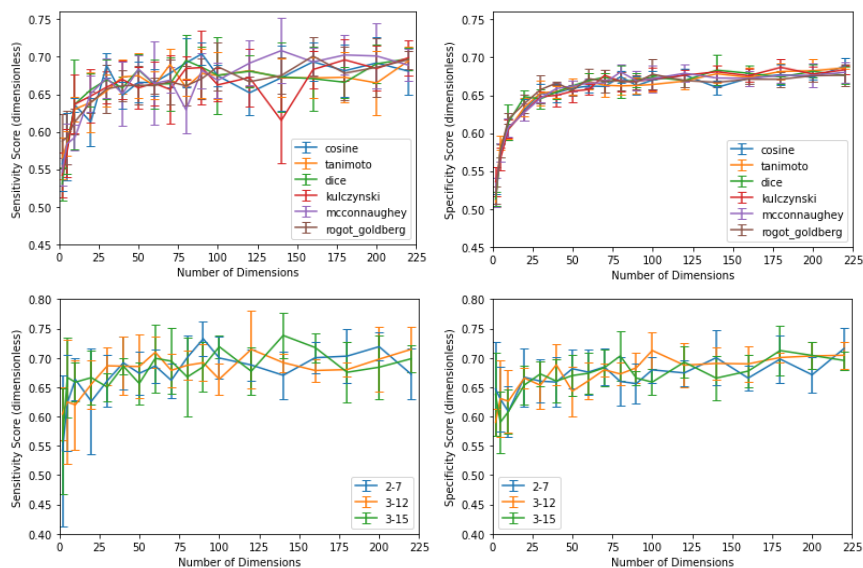


Figure 2: sensitivity (for correctly classified mutagenic molecules) and specificity (for correctly classified non-mutagenic molecules) scores, for the SC-based QSAR model (top row) and the fragmentation-based QSAR model (bottom row).

As observed from Figure 2, increased dimensionalities initially improved accuracy scores, as higher dimensionalities of data from knowledge graphs were presumably necessary for sufficiently conserving information regarding a complex toxicological space. For the SC-based model, convergence within range of 70% accuracy was apparent from approximately 100 dimensions onwards, as sufficient dimensionality was reached and inclusion of further dimensions contributed comparatively less to describing the complex toxicological space. Convergence for the fragmentation-based model appeared to have occurred less extremely and at lower dimensionalities. No consistent differences in performance were visible, concerning different SC algorithms or fragment size ranges (the latter may indicate that smaller substructures could have been responsible for causing mutagenic/non-mutagenic effects), however it should be noted that higher variances were generally present for sensitivity scores than for specificity scores; this is deemed to be due to the data equalisation measures that were performed, which successfully equalised the average sensitivity and specificity scores, but however caused greater numbers of non-mutagenic test samples to give rise to greater certainty in specificity scores than for sensitivity scores. The fragmentation-based model displayed marginally higher converged accuracy scores, than for the SC-based QSAR model, however the two models are overall deemed as closely comparable in performance.

### 3.2. Hybrid QSAR Model Results

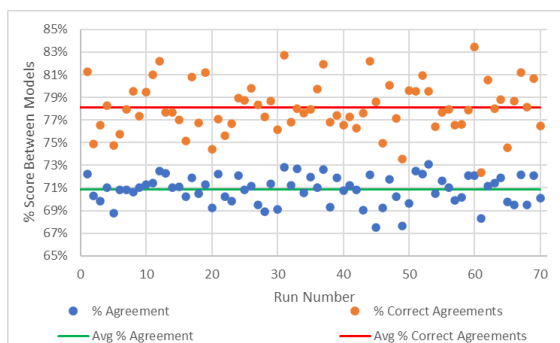


Figure 3: percentage agreements between constituent models, along with overall hybrid model performance, over 70 repeat runs.

Figure 3 demonstrates that an average of  $71\% \pm 1\%$  of predictions were agreed by constituent models, which gave rise to a hybrid model with an average accuracy score of  $78\% \pm 2\%$ . It was furthermore found that percentage agreement between the models was strongly positively correlated with percentage accuracy of the resulting hybrid model, with a Pearson correlation coefficient calculated at 0.8.

### 4. Conclusions

Our results suggested that the novel QSAR approaches explored were effective in modelling mutagenicity. Dimensionality reduction via PCA was a critical step for improving accuracy scores, via better enabling MLP classifiers to navigate a multidimensional toxicological space. Explored parameters such as individual SC algorithms, along with searched fragment sizes, were found to be less relevant for improving accuracy scores. Compared to existing QSAR models in this space, the two

individual QSAR models underperformed, however the hybrid model formed by both outperformed certain studies and matched others in average accuracy score [4-6]. Accuracy scores may be further improved by developing additional constituent QSAR models for the final hybrid model, as well as exploring parameter optimization for the NNs involved, alongside novel NN architectures and potentially more advanced and specialised NNs than MLPs. Exploration of differing dimensionality reduction algorithms may also prove beneficial, especially non-linear techniques such as autoencoders and locally linear embedding, which may perform better in preserving any complex multidimensional decision boundaries between classes that may potentially exist in higher dimensional toxicological space.

## References

- [1] J.C. Larsen, 2006. Risk assessment of chemicals in European traditional foods. *Trends in Food Science & Technology*, 17(9), pp.471-481.
- [2] S.E. Escher, et al., 2019. Towards grouping concepts based on new approach methodologies in chemical hazard assessment: the read-across approach of the EU-ToxRisk project. *Archives of toxicology*, 93(12), pp.3643-3667.
- [3] P. Gramatica, 2013. On the development and validation of QSAR models. In *Computational toxicology*, pp. 499-526.
- [4] M. Honma et al., 2019. Improvement of quantitative structure–activity relationship (QSAR) tools for predicting Ames mutagenicity: outcomes of the Ames/QSAR International Challenge Project. *Mutagenesis*, 34(1), pp.3-16.
- [5] R. Kumar et al., 2021. A deep neural network–based approach for prediction of mutagenicity of compounds. *Environmental Science and Pollution Research*, 28(34), pp.47641-47650.
- [6] C. Hung et al., 2021. QSAR modeling without descriptors using graph convolutional neural networks: the case of mutagenicity prediction. *Molecular diversity*, 25(3), pp.1283-1299.
- [7] National Center for Biotechnology Information (2022). PubChem Compound Summary for CID 17730, Sunset Yellow FCF. Retrieved December 8, 2022 from <https://pubchem.ncbi.nlm.nih.gov/compound/Sunset-Yellow-FCF>.

## Acknowledgements

This work was supported by grants from the Biotechnology and Biological Sciences Research Council [grant number BB/T008709/1] and the Food Standards Agency [Agency Project FS900120]

The views expressed in this article do not reflect the views of the European Food Safety Authority (EFSA) and/or are a reflection of the views of the authors only.

This paper aims to contribute to the international network on Advancing the Pace of Chemical Risk Assessment (APCRA), to contribute to the use of New Approach Methodologies (NAMs) in chemical risk assessment and ultimately reduce animal testing.

# Investigating tunable experiment variable effects on hiPSC-CMs maturation via unsupervised learning

Shenbageshwaran Rajendiran, Mohammadjafar Hashemi, Ferdous Frinklea,  
Nathan Young, Elizabeth Lipke, Selen Cremaschi\*

*Department of Chemical Engineering, Auburn University, Auburn, AL 36849, USA  
selen-cremaschi@auburn.edu*

## Abstract

Cardiomyocytes (CMs) are heart cells responsible for heart contraction and relaxation. CMs can be derived from human induced pluripotent stem cells (hiPSCs) with high yield and purity. Mature CMs can potentially replace dead and dysfunctional cardiac tissue and be used for screening cardiac drugs and toxins. However, hiPSCs-derived CMs (hiPSC-CMs) are immature, which limits their utilization. Therefore, it is crucial to understand how experimental variables, especially tunable ones, of hiPSC expansion and differentiation phases affect the hiPSC-CM maturity stage. This study applied clustering algorithms to day 30 cardiac differentiation data to investigate if any maturity-related cell features could be related to the experimental variables. The best models were obtained using k-means and Gaussian mixture model clustering algorithms based on the evaluation metrics. They grouped the cells based on eccentricity and elongation. The cosine similarity between the clustering results and the experimental parameters revealed that the Gaussian mixture model results have strong similarities of 0.88, 0.94, and 0.93 with axial ratio, diameter, and cell concentration.

**Keywords:** hiPSC encapsulation, CM differentiation, hiPSC-CM maturation, clustering techniques, K Means clustering

## 1. Introduction

Cardiovascular disease (CVD) is the leading cause of death worldwide (Ahmed et al., 2020). A heart attack can cause the loss of more than one billion heart cells, initiating blood flow overload and overstretching on viable cardiac cells, potentially leading to death. Because the human heart has limited regenerative capacity, it cannot replace damaged cells. Engineered heart tissue is a potential alternative for heart failure due to the difficulties associated with Cardiac Implantable Electronic devices (CIEDs) and heart transplants, two available therapies (Kempf et al., 2016). Human-induced pluripotent stem cells (hiPSCs) can be differentiated into CMs with the potential to produce therapeutic CMs (Tani et al., 2022). However, CMs produced by current differentiation protocols are immature. Immature CMs differ substantially from mature ones. For example, immature CMs have underdeveloped mitochondria, limited fatty acid oxidation capacity, less elongated cells, and disorganized sarcomeres and myofibrils. Research is ongoing to develop protocols that lead to mature hiPSC-CMs (Hamledari et al., 2022). This paper applied clustering algorithms to group hiPSC-CMs based on maturity-relevant features and investigated the relationship between clustering results and tunable experimental variables of the differentiation protocol with the aim of identifying variables that produce more matured CMs. Clustering, instead of classification, was employed

because the hiPSC-CMs cannot be labeled to train a classifier. The CMs were produced through hiPSC hydrogel encapsulation and direct differentiation within a 3D-engineered tissue microenvironment (Chang et al., 2020). The unbiased data were collected from 17 different batches, with each batch having different experimental variables, such as axial ratio (AR), cell concentration, PEG-fibrinogen (PF) concentration, and microspheroid size (i.e., diameter). Eight clustering algorithms from Scikit-learn (Fabian Pedregosa) were applied to day 30 cardiac differentiation data. The maturity-relevant features considered for clustering were cell area, cell circularity, eccentricity, elongation, sarcomere length, sarcomere organization score, and orientation index. The k-means and Gaussian mixture clustering algorithms yielded the best clusters based on evaluation metrics.

## 2. Experimental Procedure for hiPSC-CM Production and Data Collection

### 2.1. HiPSC Culture, Encapsulation, and Cardiomyocyte Differentiation

Un-Arc 16 Facs II (Shinnawi et al., 2015) was cultured on Geltrex (Gibco) with E8 media. The hiPSCs were resuspended in PF precursor solution at 30-60 million cells/mL and encapsulated within PF by using a novel microfluidic system as described previously (Finklea et al., 2021). Microspheroids with different sizes and ARs were produced (Seeto et al., 2019). The D-optimal experimental design was used, assuming Gaussian process model. Microspheroids were cultured for an additional 2 days in E8 media with daily media changes (days -2 and -1). To analyze the initial size and AR of the microspheroids, the autofluorescence of the photoinitiator Eosin Y in PF was captured using the FITC filter on the Nikon Eclipse Ti fluorescence microscope at low magnification. Standard plugins in ImageJ were used for quantification. To initiate cardiac differentiation on day 0, microspheroids were transferred to chemically defined cardiac differentiation media (CDM3) supplemented with CHIR99021 (5 – 7.5  $\mu$ M, STEMCELL Technologies). Exactly 24 h later, the media was exchanged for CDM3 supplemented with 5  $\mu$ M IWP2 (STEMCELL Technologies). Fresh CDM3 was added on days 3, 5, 7, and 10; following day 10, microspheroids were cultured in RPMI/B27 (Gibco). The differentiation outcomes were assessed using flow cytometry on day 10 using the primary antibodies, such as cTnT (Invitrogen) and MF20 (DSHB), and secondary antibody (1:300, AlexaFluor 647 goat anti-mouse IgG (ThermoFisher)).

### 2.2. Microspheroid dissociation and replating

Microspheroids on day 30 of differentiation were dissociated by incubating in a Collagenase-B (1 mg/mL, Roche) supplemented by DNase (0.05 mg/mL, Worthington) in PBS dissociation solution at 37 °C for about 8 minutes. The cells were plated on a Matrigel-coated coverslip in RPMI 20 medium (RPMI 1640 medium with 20% FBS, Atlanta Biologicals) supplemented with 5 $\mu$ M RI for 2 days before starting staining. After fixation, the cells were washed with PBS, permeabilized with PBS-T containing 0.2% Triton X-100 in PBS for 30 minutes, and blocked with 10% FBS in PBS blocking buffer for 45 minutes at room temperature. Cells were incubated in primary antibody (aSA) overnight at 4 °C. After washing, cells were incubated in the secondary antibody (Alexa Fluor 568, 1:200), and nuclei stained with Bisbenzimidazole Hoechst 33342 (MilliporeSigma) for 1 hour at room temperature. All samples were visualized with Nikon Eclipse TE2000 Inverted Microscopes equipped with a Nikon A1 plus Confocal Microscope System. To assess hiPSC-CMs cell morphology and sarcomere structure, SarcOmere Texture Analysis (Sutcliffe et al., 2018) was used for immunofluorescent confocal images.

### 3. Clustering Algorithms

#### 3.1. K-means Clustering

Given the number of clusters ( $|C|$ ), the algorithm (Likas et al., 2003) randomly initializes  $|C|$  cluster centroids ( $\mu_j, j = 1, 2, \dots, |C|$ ) and assigns a class membership to each point ( $x_i, i = 1, 2, \dots, n$ ) based on its closest centroid. Then, the centroids are reestimated using Eq. (1). Each point is reassigned a new class membership based on the new centroids. The steps are repeated until class memberships stop changing.

$$\sum_{i=0}^n \min_{\mu_j \in C} (||x_i - \mu_j||^2) \quad (1)$$

#### 3.2. Gaussian Mixture Model Clustering

The Gaussian mixture model (Yang et al., 2012) assumes that all data points come from a finite number of Gaussian distributions with unknown parameters. The distribution parameters are estimated by the expectation maximization (EM) algorithm. The EM algorithm first assigns the mean and variance of the distributions either randomly or based on the centroids of k-means clustering results. The probability of a point belonging to a cluster is calculated using the distribution parameters. Cluster means and variances are improved by maximizing the likelihood of the data given those parameters, and the procedure is repeated until the mean and variance of each distribution stop changing.

#### 3.3. Agglomerative Clustering

Agglomerative clustering (Murtagh & Legendre, 2014) is a hierarchical algorithm that uses a bottom-up approach. Each data point is initially considered a “cluster.” The algorithm proceeds by successively merging clusters using a selected linkage criterion. Criterion using ward, complete, average, and single linkage minimizes the sum of squared distances within all clusters, the maximum distance between observations of pairs of clusters, the average of the distances between all observations of pairs of clusters, and the distance between the closest observations of pairs of clusters, respectively.

### 4. Metrics Used for Evaluating Clustering Results

The three evaluation metrics used to assess the cluster results were Silhouette score, Calinski-Harabasz index (CH), and the Davies-Bouldin index (DB). Based on the results of these metrics, the best three cluster models were selected. Silhouette score (Shahapure & Nicholas) for a data point,  $s$ , is calculated using the distance between that point and all other points in the same and nearest clusters. The average silhouette score of all the data points gives the Silhouette score for the clusters. Silhouette score is between -1 and 1, with 1 indicating that the clusters are well separated. The CH (Maulik & Bandyopadhyay, 2002) is the ratio of the sum of dispersion between and within clusters for all clusters. The CH is higher for dense and separated clusters. The DB (Maulik & Bandyopadhyay, 2002) is defined as the average similarity between each cluster  $C_a$  and its most similar one  $C_b$ . The similarity score,  $R_{ab}$ , given in Eq. (2), is calculated between cluster  $a$  and  $b$ ,  $d_a$  is the average distance between each point in cluster  $a$  and its centroid, and  $D_{ab}$  is the distance between cluster centroids  $a$  and  $b$ . The DB is calculated by Eq. (3). Values closer to zero indicate better clustering.

$$R_{ab} = \frac{d_a + d_b}{D_{ab}} \quad (2)$$

$$DB = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} R_{ab} \quad (3)$$



## 5. Results and Discussion

### 5.1. Clustering Results

Five hundred thirteen (513) hiPSC-CMs from 17 batches were analyzed to obtain cell morphological features, such as cell area, eccentricity, circularity, elongation, and sarcomere properties, such as sarcomere length, orientation index, and organization score. Only k-means, Gaussian mixture model, and agglomerative clustering algorithms were found to be applicable to the dataset. The agglomerative clustering algorithm grouped almost all hiPSC-CMs in one cluster. These clustering results were not further analyzed. Because this study aims to investigate the relationship between mature and immature hiPSC-CMs and experimental variables, the number of clusters was set to two.

Based on the evaluation metrics, k-means clustering yielded the best model with a Silhouette score of 0.17, CH of 107, and DB of 1.99. Gaussian mixture model clustering with spherical, full, and diagonal covariance yielded the following models in that order. The evaluation metrics for these models were similar to those for the k-means algorithm. Analyzing the two clusters yielded by the k-means and Gaussian mixture model algorithms revealed that the hiPSC-CMs were clustered based on eccentricity and elongation. Mature CMs typically have an elliptical shape with larger eccentricity and elongation values (Karbassi et al., 2020).

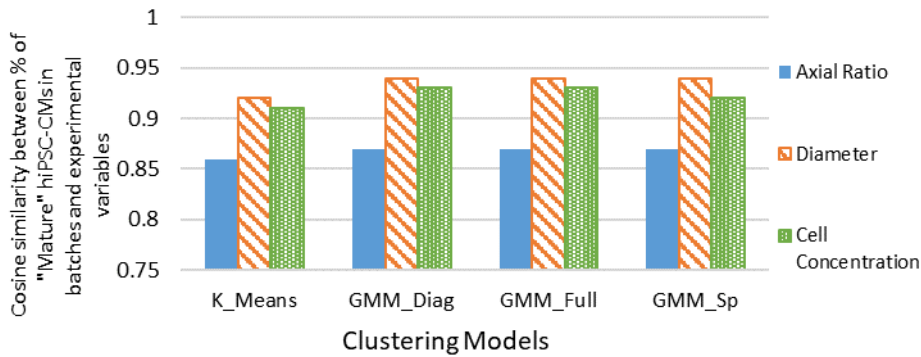
The k-means algorithm grouped hiPSC-CMs with eccentricity values between 0.7 and 1 and elongation between 2 and 10 in one cluster, which we named the “Mature” cluster, and the remaining cells in another, we named the “Immature” cluster. The Gaussian mixture model algorithms yielded clusters with the same cell property ranges. The overall clustering results suggest that eccentricity and elongation were significant features for separating mature hiPSC-CMs from immature cells by unsupervised learning.

### 5.2. Relationship between hiPSC-CM Maturity and Tunable Experimental Variables

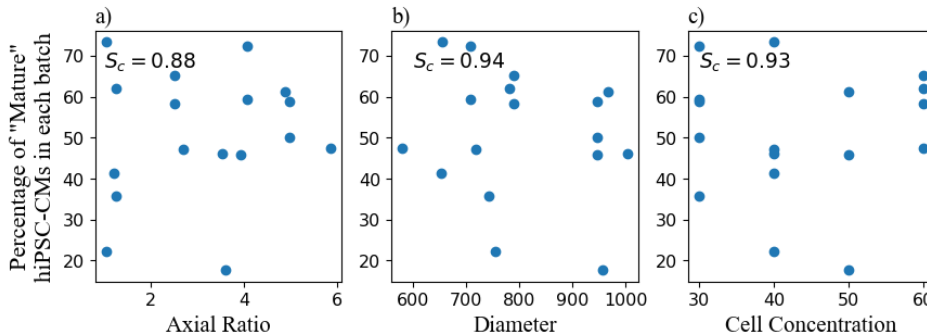
To investigate the potential relationship between hiPSC-CM maturity and tunable experimental variables, we calculated the percentage of hiPSC-CMs clustered as “Mature” in each of the 17 batches. A set of tunable experimental variables, namely AR, diameter, cell concentration PEG concentration, and PF concentration, were used in each batch. Then, the Cosine similarity (Xia et al., 2015) between this percentage and each tunable experimental variable for each clustering result was calculated. Cosine similarity is a measure of similarity between two vectors. Using cosine similarity, we quantify how similar the percentage of mature CMs to the experimental variables. The resulting similarity values are plotted in Figure 1. The PEG and PF concentration results were not considered because there were only three values for these variables. Figure 1 illustrates that the strongest similarity metric values were observed for the results obtained using the Gaussian mixture model clustering, with a cosine similarity of 0.88 with AR, 0.94 with diameter, and 0.93 with the cell concentration. The cosine similarity values for the results obtained with the k-means algorithm are similar, with a cosine similarity of 0.86 with AR, 0.92 with diameter, and 0.91 with cell concentration.

Figure 2 plots the percentage of “Mature” hiPSC-CMs in each batch calculated using Gaussian mixture model clustering results versus AR (Figure 2(a)), diameter (Figure 2(b)), and cell concentration (Figure 2(c)). Figures 2(a), 2(b), and 2(c) suggest that there is a relationship between the percentage of “Mature” hiPSC-CMs and AR, diameter, and cell concentration. Three of the five batches with ARs less than two have less than 45 % of hiPSC-CMs clustered as “Mature”. In contrast, batches with an AR greater than two have at least 45% hiPSC-CMs clustered as “Mature”, except for one. Figure 2(c) shows that when the concentration is equal to 60 million cells/mL, the percentage of mature cells is greater than 50%. These observations suggest the existence of non-linear and

potentially complex relationships between the mature hiPSC-CMs and the ARs, diameters, and cell concentration, which should be further investigated.



**Figure 1** – Cosine similarity between the percentage of “Mature” hiPSC-CMs in batches based on the clustering results and the experimental parameters. K\_Means – k-means clustering, GMM\_Diag, GMM\_Full, and GMM\_Sp – Gaussian mixture model with diagonal, full, and spherical covariance function.



**Figure 2** - Percentage of mature hiPSC-CMs clustered by k means clustering vs. (a) AR (b) Diameter (c) Cell concentration.  $S_c$  – Cosine similarity between the percentage of mature CMs and each experimental parameter.

## 6. Conclusions and Future Directions

Using day 30 cardiac differentiation data, we identified eccentricity and elongation as the significant features for clustering hiPSC-CMs into two groups via k-means and Gaussian mixture model clustering algorithms. The results of the Gaussian mixture and k-means models suggested that there is a concurrent relationship between ARs, diameter, and cell concentration, all experimentally tunable variables, and hiPSC-CMs with higher eccentricity and elongation, which are associated with maturity. Future work will investigate using different clustering approaches that relax the compact set assumption for determining better “Mature” cell cluster boundaries. Then, the relationships between the experimental variables and mature hiPSC-CMs will be modeled to aid in directing hiPSC differentiation towards mature CMs.

## 7. Acknowledgments

This work is funded by NSF grant number 2135059.

## References

- Ahmed, R. E., Anzai, T., Chanthra, N., & Uosaki, H. (2020). A Brief Review of Current Maturation Methods for Human Induced Pluripotent Stem Cells-Derived Cardiomyocytes. *Front Cell Dev Biol*, 8, 178.
- Chang, S., Finklea, F., Williams, B., Hammons, H., Hodge, A., Scott, S., & Lipke, E. (2020, Jul). Emulsion-based encapsulation of pluripotent stem cells in hydrogel microspheres for cardiac differentiation. *Biotechnol Prog*, 36(4), e2986.
- Fabian Pedregosa, G. V., Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Édouard Duchesnay;. *Scikit-learn: Machine Learning in Python*. <https://scikit-learn.org/stable/about.html#citing-scikit-learn>
- Finklea, F. B., Tian, Y., Kerscher, P., Seeto, W. J., Ellis, M. E., & Lipke, E. A. (2021, Jul). Engineered cardiac tissue microsphere production through direct differentiation of hydrogel-encapsulated human pluripotent stem cells. *Biomaterials*, 274, 120818.
- Hamledari, H., Asghari, P., Jayousi, F., Aguirre, A., Maaref, Y., Barszczewski, T., Ser, T., Moore, E., Wasserman, W., Klein Geltink, R., Teves, S., & Tibbits, G. F. (2022). Using human induced pluripotent stem cell-derived cardiomyocytes to understand the mechanisms driving cardiomyocyte maturation. *Front Cardiovasc Med*, 9, 967659
- Karbassi, E., Fenix, A., Marchiano, S., Muraoka, N., Nakamura, K., Yang, X., & Murry, C. E. (2020, Jun). Cardiomyocyte maturation: advances in knowledge and implications for regenerative medicine. *Nat Rev Cardiol*, 17(6), 341-359.
- Kempf, H., Andree, B., & Zweigerdt, R. (2016, Jan 15). Large-scale production of human pluripotent stem cell derived cardiomyocytes. *Adv Drug Deliv Rev*, 96, 18-30.
- Kerscher, P., Kaczmarek, J. A., Head, S. E., Ellis, M. E., Seeto, W. J., Kim, J., Bhattacharya, S., Suppiramaniam, V., & Lipke, E. A. (2017, Aug 14). Direct Production of Human Cardiac Tissues by Pluripotent Stem Cell Encapsulation in Gelatin Methacryloyl. *ACS Biomater Sci Eng*, 3(8), 1499-1509.
- Likas, A., Vlassis, N., & J. Verbeek, J. (2003). The global k-means clustering algorithm. *Pattern Recognition*, 36(2), 451-461.
- Maulik, U., & Bandyopadhyay, S. (2002). Performance evaluation of some clustering algorithms and validity indices. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(12), 1650-1654.
- Murtagh, F., & Legendre, P. (2014). Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *Journal of Classification*, 31(3), 274-295.
- Seeto, W. J., Tian, Y., Pradhan, S., Kerscher, P., & Lipke, E. A. (2019). Rapid Production of Cell-Laden Microspheres Using a Flexible Microfluidic Encapsulation Platform. *Small*, 15(47), 1902058.
- Shahapure, K. R., & Nicholas, C. (2020). Cluster Quality Analysis Using Silhouette Score.
- Shinnawi, R., Huber, I., Maizels, L., Shaheen, N., Gepstein, A., Arbel, G., Tijssen, A. J., & Gepstein, L. (2015, Oct 13). Monitoring Human-Induced Pluripotent Stem Cell-Derived Cardiomyocytes with Genetically Encoded Calcium and Voltage Fluorescent Reporters. *Stem Cell Reports*, 5(4), 582-596.
- Sutcliffe, M. D., Tan, P. M., Fernandez-Perez, A., Nam, Y. J., Munshi, N. V., & Saucerman, J. J. (2018, Jan 19). High content analysis identifies unique morphological features of reprogrammed cardiomyocytes. *Sci Rep*, 8(1), 158.
- Tani, H., Tohyama, S., Kishino, Y., Kanazawa, H., & Fukuda, K. (2022, Mar). Production of functional cardiomyocytes and cardiac tissue from human induced pluripotent stem cells for regenerative therapy. *J Mol Cell Cardiol*, 164, 83-91.
- Xia, P., Zhang, L., & Li, F. (2015). Learning similarity with cosine similarity ensemble. *Information Sciences*, 307, 39-52.
- Yang, M.-S., Lai, C.-Y., & Lin, C.-Y. (2012). A robust EM clustering algorithm for Gaussian mixture models. *Pattern Recognition*, 45(11), 3950-3961.

# On the systematic development of large-scale kinetics using stability criteria and high-throughput analysis of curated dynamics from genome-scale models

Konstantinos Mexis, Stefanos Xenios, and Antonis Kokosis\*

*Department of Process Engineering, NTUA, Iroon Politechniou 6 Zografou, Athens, Greece*

*\*Corresponding author: [akokossis@mail.ntua.gr](mailto:akokossis@mail.ntua.gr)*

## Abstract

This paper is part of a general effort to integrate disjoint stages of the Design-Build-Test-Learn cycle that addresses the simultaneous design of biocatalysts with process engineering. Engineering kinetics are currently based on regression studies based on experimental data with limited reference to the underlying reaction pathways. Instead, the ORACLE framework offers an attractive environment to generate populations of large-scale (curated) dynamics and a platform for in-silico kinetic models to check for physiological relevance and stability. Stability checks have reported low rates of success and the paper explains a systematic approach that combines deterministic methods and data analytics to accelerate realizable kinetics that could be set a basis to connect the dynamics of the cell with the dynamics of the process. The work is demonstrated with the production of muconic from *S. cerevisiae* which is achieved by shunting the shikimic pathway. The ORACLE framework consists of two reduction stages where the first one ensures that the generated kinetic models are physiologically relevant and the second one checks the model's stability. The conventional method produced 370 physiologically relevant models out of which 70 were stable (19.2%) whereas our approach increased uptake of acceptable solutions to 97.7%.

**Keywords:** metabolic engineering; machine learning; rule extraction; *S.Cerevisiae*; muconic acid;

## 1. Introduction

The typical design process towards a sustainable bio-production of chemicals follows the iterative Design-Build-Test-Learn (DBTL) cycle. The design involves the selection of the platform organism and the heterologous reactions that need to be expressed to produce the desired chemicals as well as strategies (enzyme upregulations or downregulations) to enhance product yield. The build module involves the genetic transformation and gene editing of the platform organism in accordance to the strategies of the design step. The test module we gather information on the cloning results, omics data and help comprehend cellular behavior. Small batch experiments showcase the developed strain's desired product yield as well as some key metabolites secretion rates. The learn module takes into account the generated information from the test step and incorporates it into new metabolic strategies to further optimize and increase the desired flux.

Optimization and Risk Analysis of Complex Living Entities (ORACLE) is an attractive methodology to formalize development stages and set a basis for the development of large-scale kinetics. The approach combines mathematical optimization and data analytics. The ORACLE framework is a venue to connect genome-scale models with process engineering. Thermodynamic Flux Analysis is used to predict potential fluxes at steady states; reduction methods and parameter sampling can be applied to produce candidates for observable dynamics. Pruning tests and stability checks can be applied to determine candidates for large-scale kinetic models. Such models are valuable in scale-up studies, in the design and optimization of bioreactors, also to manipulate pathways as required to increase yields and selectivity. The screening of candidates and the analysis of uncertainty are critical issues in the current methodology that currently embraces hypotheses and heuristics to tackle the large size and complexity of the problem. The challenges offer room for systems methods to add rigor, automate the analysis, and improve the quality and relevance of the high-throughput experiments. The paper explains the application of a systems approach to tackle stability of curated kinetics demonstrating significant improvements in the quality of the results.

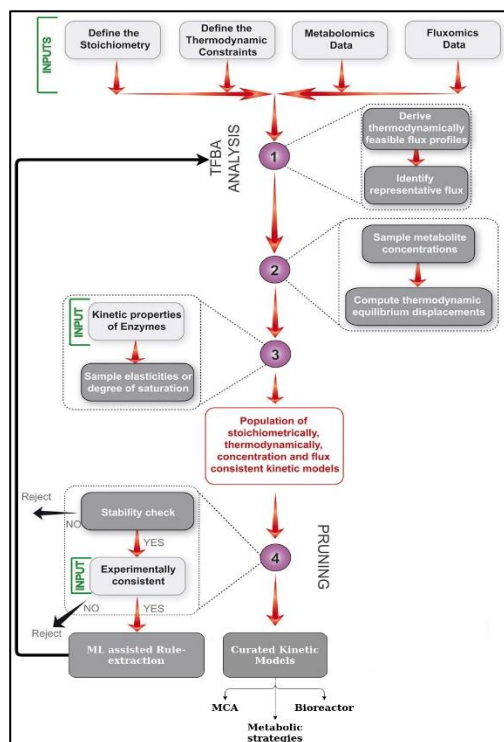


Figure 1. ORACLE methodology pipeline

## 2. Pilot study: Muonic acid production from *S.cerevisiae*

The pilot study involves a Genome-Scale model produced for the mutant strain of *S. Cerevisiae* that produces muonic acid via the shunting of shikimic pathways. The model is thermodynamically curated and reduced in order to generate kinetic models through the ORACLE framework. A total of 23500 kinetic models was generated; 370 of them remained within physiological bounds. Only 70 out of 370 models showed great stability to random initial metabolite concentration. Due to the small percentage of stable models, we deployed a decision tree classifier to extract rules on key kinetic parameters that affect the system stability. The decision tree's input kinetic parameters were the enzyme saturation ( $\sigma$ ) and thermodynamic displacement

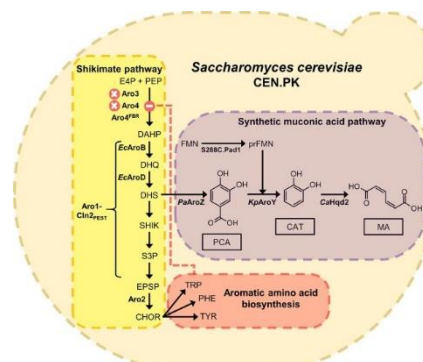


Figure 2. Muonic acid pathway

The decision tree's input kinetic parameters were the enzyme saturation ( $\sigma$ ) and thermodynamic displacement

(gamma) parameters. The differences in the development of decision trees motivated the remaining analysis where we explored different rule extraction approaches.

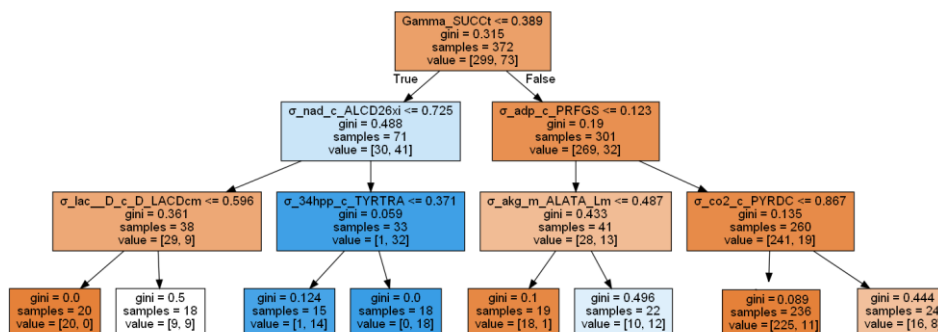


Figure 3. Decision tree of kinetic parameters

### 3. Methods and Materials

#### 3.1. Parameter classification problem

The parameter classification problem is the identification of a subspace in the parameter space wherein the parameters satisfy a given property (GP). Let us consider an  $n$ -dimensional space of parameters  $p_1, p_2, \dots, p_n$ , and assume that the GP is satisfied provided a function of these parameters satisfies  $f(p_1, p_2, \dots, p_n) > 0$ . The task is to find ranges of  $p_1, p_2, \dots, p_n$ , for which the GP is satisfied without any knowledge of the exact functional form of  $f(p_1, p_2, \dots, p_n)$ . As applied in the population of the curated kinetic models, the parameters  $p_1, p_2, \dots, p_n$  account for the saturation and thermodynamic displacement parameters of the kinetic model; GP would be satisfied if the referred kinetic model passes the stability tests.

#### 3.2. Learning-based Rule-extraction Algorithms

Rule-extraction is important to understand how the system functions. Rule extraction methods are typically used in conjunction with learning algorithms that feature an inherent explanation capability (surrogate model). In our case several ML methods (black-box models) are trained using saturation and thermodynamic displacement parameters of curated kinetic models to predict whether the curated kinetic model is stable or not. Then, a CART decision tree (surrogate model) was trained to approximate predictions of the black box models.

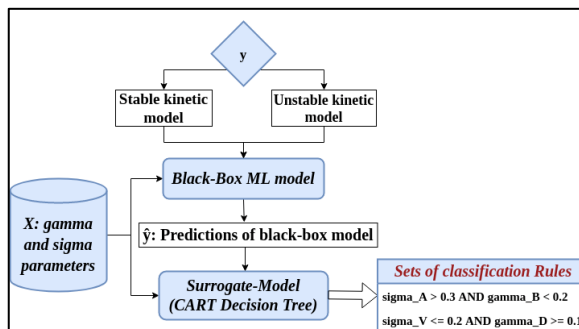
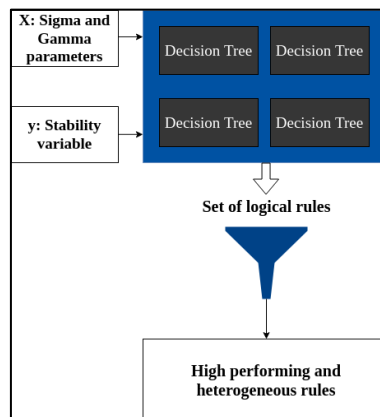


Figure 4. Pipeline to obtain a surrogate model

### 3.3. Skope-Rules

*Skope-Rule* extraction is a rule-based interpretable model that allocates logical rules with high precision to fuse them. The difference from other rule-extraction methods stands in that decision rules are actual inputs. Rules are extracted from tree ensembles; similar or duplicated rules are removed based on similarity thresholds. A fitted tree defines a set of rules; rules are then tested selecting and merging rules with higher precision. In our case, the Skope-Rules algorithm is implemented on training tree ensemble models on saturation and thermodynamic displacement parameters. They algorithm is used to predict if a given kinetic model is stable or not.



**Figure 5.** Skope-Rules pipeline

## 4. Methodology

The adopted rules-extraction procedure is outlined in the following steps:

**Step 1a.** The physiologically relevant generated kinetic models are used to define a *dataset* that consists of an *input* ( $X$ ) and an *answer set* ( $y$ ). The *input set* contained all the *gamma* (thermodynamic displacement) and *saturation* (*sigma*) parameters for each kinetic model; the *answer set* answers if the model is stable or not.

**Step 1b.** The dataset was split into a *training set* and a *validation set* in an 80:20 ratio. The *training set* was then used to train and extract the rules from the machine learning models, while the *validation set* was used to validate the accuracy of the extracted rules. We computed the *Stability Index* (SI) as the ratio between the number of kinetic parameters sets that passed the *stability check*.

**Step 2.** Several machine learning classifiers (“black-box” models) were built using this *training set*. Then, these models were used to predict the labels, i.e. the *answer set*, for the *input set* mentioned in *Step 1*. The employed machine learning algorithms used as “black-box” models were; Logistic Regression and Support Vector Machines.

**Step 3.** A decision tree (*surrogate model*) was built on the *input set* of the training data and the *predictions* of the black-box model. The rules generated from the decision tree represented the generalization behavior of the black-box model. We ranked the classification rules based on the number of training samples (kinetic parameter sets) that they enclose and we selected the rule enclosing the most samples.

**Step 4.** From the *validation set*, we selected only the kinetic models that satisfied the extracted rules and computed the SI over these sampled models. If the obtained SI is improved, the extracted rules and the general methodology as well are valid.

**Step 5.** Finally, we generated new populations of kinetic models based on the inferred rules from *Step 3* independently of the training set. We calculated the SI obtained and compared it with the one from the *training set*. If the SI and the number of feasible kinetic models are improved, the rules extracted from the black-box model are valid and the uncertainty in kinetic parameters is reduced.

## 5. Results and Discussion

Different sets of rules were extracted from each black-box model and the rule enclosing the most training examples was finally selected for each model. These inferred rules

constrained the saturations and the corresponding kinetic parameters of some critical enzymes, reducing the sampling space and, as a result, the uncertainty in the model analysis. The critical parameters that affect the stability of the generated kinetic models and the inferred rules according to every black-box model are presented in *Table 1*.

Black Box Model	Rules
Logistic Regression	Gamma_GLUDC ≤ 6.27e-09 Sigma_[chor_c]_CHORS ≤ 0.126
Decision Tree	Gamma_HCO3E > 0.856 Sigma_[nad_c] ≤ 0.524
SVM	Gamma_GLUDC ≤ 6.27e-09 Sigma_[2oxoadp_m] ≤ 0.646 Sigma_[g6p_c] ≤ 0.522 Sigma_[gmp_c] ≤ 0.999
Skope rules	Gamma_FBA ≤ 0.527 Sigma_[nad_m] > 0.484 Sigma_[glu_L_m] ≤ 0.862

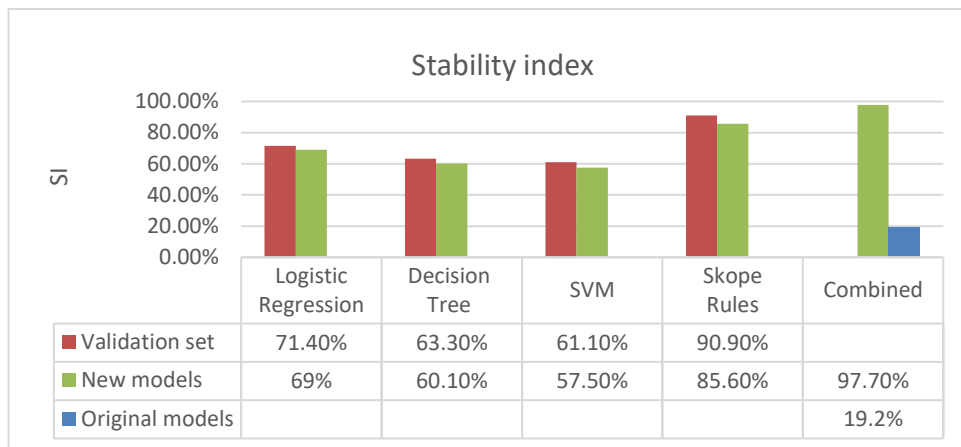
**Table 1.** Rules extracted from each machine learning algorithm

Reaction Name	Reaction Stoichiometry
GLUDC	glu_L_c + h_c ⇌ 4abut_c + co2_c
CHORS	3psme_c → chor_c + pi_c
H3COE	co2_c + h2o_c ⇌ h_c + hco3_c
GAPD	g3p_c + nad_c + pi_c ⇌ 13dpg_c + h_c + nadh_c
2OXOADPTm	2oxoadp_m + akg_c ⇌ 2oxoadp_c + akg_m
LMPD biomass	Biomass reaction
GK1	atp_c + gmp_c ⇌ adp_c + gdp_c
FBA	fdp_c ⇌ dhap_c + g3p_c
ALCD26xi	h_c + nadh_c + id3acald_c ⇌ nad_c + ind3eth_c
ASPTA	akg_c + asp_L_c ⇌ glu_L_c + oaa_c

**Table 2.** List of reactions that came up during the Machine Learning analysis

In order to quickly validate the accuracy of the extracted rules, we sampled from the *validation set* (as described in paragraph 2.4) only the kinetic models that satisfied these rules and computed the Stability Index over them. The obtained SI for every implemented machine learning algorithm was improved over the SI computed on the training test, meaning that our method is able to determine and constrain the critical kinetic parameters that lead to more feasible kinetic models. We proceeded with the generation of new kinetic models implementing the extracted rules for the parameters that seem to affect stability. A total of 200 physiologically relevant kinetic models were produced for every different set of rules. In addition, we implemented all the rules combined and generated again 200 physiologically relevant models. The generated models underwent the ORACLE stability check where 100 random concentration perturbations were performed using a wide range ( $0.5[X_{ref}] \leq [X] \leq 2[X_{ref}]$ ) of initial values. We deemed a kinetic model stable if 95/100 times the metabolite concentrations returned to steady state. The results obtained are presented on *Figure 6*.





**Figure 6.** Stability Index obtained from extracted rules for every machine learning model. *Validation Set:* SI is computed on samples following extracted rules on validation set. *New Models:* SI is computed on new kinetic models, generated based on extracted rules. *Original Models:* SI is computed on kinetic models generated originally, with no inferred rules.

## 6. Conclusion

The production of large scale metabolic kinetic models is hindered by the uncertainty of predicting the kinetic parameters and producing physiologically relevant and robust kinetic models. Using the ORACLE framework, we are able to generate large populations of kinetic models but the biggest percentage of them are not stable. In this work the uncertainty in the model analysis is reduced through the use of machine learning principles. Using machine learning classification and explainability techniques we were able to raise the stability index of the generated models. We were able to reach up to 97.7% stability on our generated models implementing simultaneously all the extracted rules which leads to postulate that it is better to constrain such systems as much as possible to get more feasible results.

## Acknowledgments

The authors acknowledge support from the EPFL team of Prof. Hatzimanikatis, especially Miskovic Ljubisa, who assisted in applying the ORACLE methodology. The research project was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the “2nd Call for H.F.R.I. Research Projects to support Post-Doctoral Researchers” (Project Number: 3817).

## References

- Xenios, S., Weilandt, D., Vasilis, H., Miskovic, L., & Kokosis, A. (2022). On the integration of process engineering with metabolomics for the production of muconic acid: the case for *Saccharomyces Cerevisiae*. *Computer Aided Chemical Engineering*, 541–546.
- Andreozzi, S., Miskovic, L., & Hatzimanikatis, V., 2016, iSCHRUNK – In Silico Approach to Characterization and Reduction of Uncertainty in the Kinetic Models of Genome-scale Metabolic Networks. *Metabolic Engineering*, 33, 158–168.
- Barakat, N., & Bradley, A. P., 2010, Rule extraction from support vector machines: A review. *Neurocomputing*, 74(1–3), 178–190

# Development and assessment of intensification alternatives on the lignocellulosic bioethanol production process

Yulissa M. Espinoza-Vázquez<sup>a</sup>, Fernando Israel Gómez-Castro<sup>a\*</sup>, Eduardo Sánchez-Ramírez<sup>a</sup>, Araceli G. Romero-Izquierdo<sup>b</sup>.

<sup>a</sup>*Departamento de Ingeniería Química, División de Ciencias Naturales y Exactas, Campus Guanajuato. Universidad de Guanajuato, Noria Alta S/N, Guanajuato, Gto. 36050, México. [fgomez@ugto.mx](mailto:fgomez@ugto.mx)*

<sup>b</sup>*Facultad de Ingeniería, Universidad Autónoma de Querétaro, Cerro de las Campanas S/N, Querétaro, Qro. 76010, México.*

## Abstract

In this work, the design of the conventional process to produce bioethanol from lignocellulosic biomass is developed. The conventional process consists of pretreatment, hydrolysis, and fermentation of the biomass to obtain bioethanol and other bioproducts. The bioethanol is purified up to 99% through extractive distillation. Different options for intensification are proposed to improve such conventional process. The analysis of the developed production schemes is performed by simulation in the software Aspen Plus V.11. The first intensified scheme (I-1) consisted in the elimination of the glycerol purification column. The second intensified scheme (I-2) consisted in the elimination of the column that purifies the ethanol-ammonia mixture, by means of the implementation of a side stream. Scheme I-2 reduces the thermal load by 2.36%, which is associated with a lower capital cost.

**Keywords:** bioethanol, lignocellulosic biomass, process intensification, process simulation.

## 1. Introduction

Fossil fuels contributed with 79.7 % of the total energy used in the world in 2015 (Banco Mundial, 2022). In addition, the U.S. Energy Information Administration (EIA) estimated that in 2021, the consumption of gasoline in the U.S. transportation sector resulted in the emission of about 1,018 million metric tons (MMmt) of carbon dioxide (CO<sub>2</sub>) (Energy Information Administration, 2022). To reduce the dependence on fossil fuels and their environmental impact, the use of biofuels has been proposed as a viable solution. Those biofuels are obtained from renewable biomass and are considered one of the most feasible options to reduce CO<sub>2</sub> emissions in the transportation sector (Alam and Tanveer, 2020). Bioethanol stands out among the fuels that can be obtained from lignocellulosic biomass. This alcohol is used in blends of 10 vol% with gasoline (Costagliola et al., 2016) and can be used to partially replace the use of gasoline.

Biomass is renewable organic feedstock that can be obtained from plants and animals. The composition of the plant biomass generally comprises cellulose, hemicellulose, and lignin; referred to as lignocellulosic biomass. By 2020, the production of agricultural residues in the world has been reported as  $2.14 \times 10^9$  tons (FAO, 2022), with a high contribution of residues from corn, sorghum, wheat, and barley crops. Sorghum is

important at national level since Mexico is its main producer in Central America. To be economically feasible, the conversion of biomass into biofuels must take place in a biorefinery scheme. A biorefinery is a facility that integrates biomass conversion processes and equipment to produce fuels, energy, and chemicals. In the biorefinery scheme based on lignocellulosic materials, it is necessary to break down the lignin structure on the feedstock, then obtaining fermentable sugars from cellulose and hemicellulose. The initial lignin removal requires a pretreatment stage. In the literature there are several studies related to the production of lignocellulosic bioethanol, such as the work developed by Conde-Mejía et al. (2012), in which different types of pretreatments as well as different conversion technologies were analyzed to find the best configuration of this process in economic and environmental terms. Similarly, Duque et al. (2020) analyzed the production of ethanol from barley straw combining a soda and enzyme-catalyzed extrusion pretreatment, reporting a production of 15.8 g ethanol/100 g raw barley straw. In these studies, several types of biomass and different conversion pathways were analyzed, with the major objective of obtaining feasible pathways with high yields. Nevertheless, biofuels must also be economically competitive with their fossil counterparts; this implies that the biofuels production costs must be reduced. In this context, process intensification (PI) plays a key role, since it can help to generate compact, energy-efficient, safer, and more environmentally friendly processes. Therefore, in this work, the design of the conventional process to produce bioethanol from lignocellulosic biomass is developed. Areas of opportunity are detected for the conventional process, proposing and evaluating intensification alternatives, such as thermally coupled columns and columns with side streams, aiming to the improvement of the energetic performance of the conventional process.

## 2. Case Study

The case study is based on the work of Conde-Mejía et al. (2013), considering 87,539 kg/h of sorghum residues as feedstock. Data in Table 1 show the composition of the lignocellulosic material. The carbohydrates profile of the feedstock is shown in Table 2.

**Table 1.** Composition of lignocellulosic biomass (Conde-Mejía et al., 2013).

Component	% wt.	Formula
Cellulose	27.56	$C_6H_{10}O_5$
Hemicellulose	15.37	$C_5H_8O_4$
Lignin	14.11	$C_{7.3}H_{13.9}O_{1.3}$
Carbohydrates	42.93	

**Table 2.** Carbohydrates profile (Conde-Mejía, et al., 2013).

Component	% wt.	Formula
Galactan	9.41	$C_6H_{10}O_5$
Manan	3.62	$C_6H_{10}O_5$
Arabinan	29.90	$C_5H_8O_4$

The pretreatment is carried out with dilute acid, due to the low sulfuric acid prices and the high yields of the treatment. Table 3 shows the reactions occurring in the pretreatment reactor, as well as their yields. In this reactor, 0.49 wt% of  $H_2SO_4$  is used at 160 °C and 1 atm. On the other hand, acid hydrolysis is carried out in a reactor at 220 °C and 1 atm. The hydrolysis reactions are presented in Table 4. As a next step, fermentation is carried

out using urea (CH<sub>4</sub>N<sub>2</sub>O) in a CH<sub>4</sub>N<sub>2</sub>O/sorghum residues mass ratio of 1:32.18, at 34°C and 1 atm. Fermentation reactions are presented in Table 5.

**Table 3.** Reactions for the pretreatment process (Conde-Mejía et. al., 2013).

Reaction	Yield	Reference reagent	Equation
$C_6H_{10}O_5 + H_2O \rightarrow C_6H_{12}O_6$	0.065	Cellulose	(1)
$C_6H_{10}O_5 + 0.5H_2O \rightarrow 0.5C_{12}H_{22}O_{11}$	0.007	Cellulose	(2)
$C_5H_8O_4 + H_2O \rightarrow C_5H_{10}O_5$	0.75	Hemicellulose	(3)
$C_5H_8O_4 \rightarrow C_5H_4O_2 + 2H_2O$	0.15	Hemicellulose	(4)
$C_6H_{10}O_5 + H_2O \rightarrow C_6H_{12}O_6$	0.75	Galactan	(5)
$C_6H_{10}O_5 \rightarrow C_6H_6O_3 + 2H_2O$	0.15	Galactan	(6)
$C_6H_{10}O_5 + H_2O \rightarrow C_6H_{12}O_6$	0.75	Manan	(7)
$C_6H_{10}O_5 \rightarrow C_6H_6O_3 + 2H_2O$	0.15	Manan	(8)
$C_5H_8O_4 + H_2O \rightarrow C_5H_{10}O_5$	0.75	Arabinan	(9)
$C_5H_8O_4 \rightarrow C_5H_4O_2 + H_2O$	0.10	Arabinan	(10)

**Table 4.** Reactions for the hydrolysis process (Conde-Mejía et. al., 2013).

Reaction	Yield	Reference reagent	Equation
$C_6H_{10}O_5 + 0.5H_2O \rightarrow 0.5C_{12}H_{22}O_{11}$	0.012	Cellulose	(11)
$C_6H_{10}O_5 + H_2O \rightarrow C_6H_{12}O_6$	0.8	Cellulose	(12)
$C_{12}H_{22}O_{11} + H_2O \rightarrow 2C_6H_{12}O_6$	1	Cellobiose	(13)
$C_5H_8O_4 + H_2O \rightarrow C_5H_{10}O_5$	0.99	Hemicellulose	(14)
$C_5H_{10}O_5 \rightarrow C_5H_4O_2 + H_2O$	0.8	Xylose	(15)

**Table 5.** Reaction for the fermentation process (Conde-Mejía et. al., 2013).

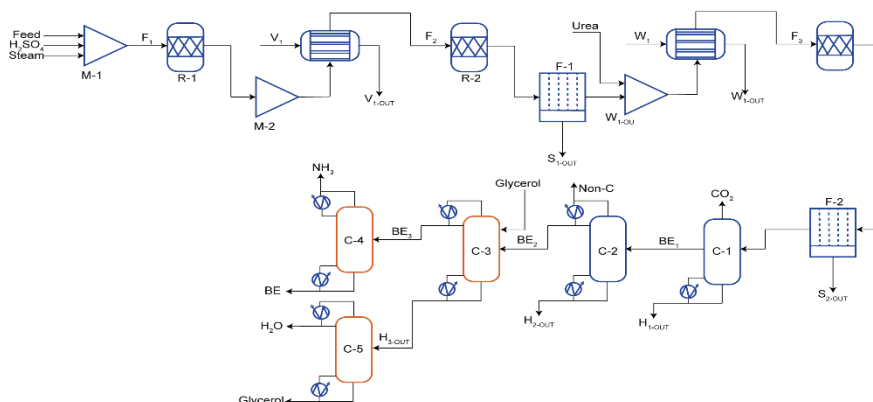
Reaction	Yield	Reference reagent	Equation
$CH_4N_2O + H_2O \rightarrow 2NH_3 + CO_2$	0.99	Urea	(16)
$C_6H_{12}O_6 \rightarrow 2C_2H_6O + 2CO_2$	0.92	Glucose	(17)
$C_6H_{12}O_6 + 2H_2O \rightarrow 2C_3H_8O_3 + O_2$	0.002	Glucose	(18)
$C_6H_{12}O_6 + CO_2 \rightarrow 2C_4H_6O_4 + O_2$	0.008	Glucose	(19)
$C_6H_{12}O_6 \rightarrow 3C_2H_4O_2$	0.022	Glucose	(20)
$C_6H_{12}O_6 \rightarrow 2C_3H_6O_3$	0.013	Glucose	(21)
$3C_5H_{10}O_5 \rightarrow 5C_2H_6O + 5CO_2$	0.85	Xylose	(22)
$3C_5H_{10}O_5 + 5H_2O \rightarrow 5C_3H_8O_3 + 2.5O_2$	0.029	Xylose	(23)
$3C_5H_{10}O_5 + 5CO_2 \rightarrow 5C_4H_6O_4 + 2.5O_2$	0.009	Xylose	(24)
$2C_5H_{10}O_5 \rightarrow 5C_2H_4O_2$	0.024	Xylose	(25)
$3C_5H_{10}O_5 \rightarrow 5C_3H_6O_3$	0.014	Xylose	(26)
$C_5H_{10}O_5 \rightarrow 5C_3H_6O_3$	1	Arabinose	(27)
$C_5H_{12}O_6 \rightarrow 2C_3H_6O_3$	1	Galactose	(28)
$C_6H_{12}O_6 \rightarrow 2C_3H_6O_3$	1	Mannose	(29)
$C_6H_{12}O_6 \rightarrow 2C_3H_6O_3$	1	Glucose	(30)
$3C_5H_{10}O_5 \rightarrow 5C_3H_6O_3$	1	Xylose	(31)
$C_2H_4O_2 + NH_3 \rightarrow C_2H_7NO_2$	1	Acetic Acid	(32)
$H_2SO_4 + 2NH_3 \rightarrow N_2H_8SO_4$	1	Sulfuric Acid	(33)

Through these conversion steps, it is possible to obtain 15,482.61 kg/h of ethanol; however, the fermentation effluent still requires a purification process. The description of

the strategy used to simulate and design the whole process will be presented in the following section.

### 3. Methodology

Bibliographical research was carried out to determine the parameters required for the simulation of the solid components present in the biomass, *i.e.* cellulose, hemicellulose and lignin; these data has been obtained from the database of the National Renewable Energy Lab (1996). NRTL was used as the thermodynamic model, due to the complexity of the involved mixtures and the operating conditions (Conde-Mejía et al., 2013). Case study was simulated in the software Aspen Plus V.11 using RYield reactors. Process flow diagram is shown in Figure 1. In the purification section, C-1 is the distillation column for the separation of CO<sub>2</sub> as top product with a purity of 95% mass, with heavy components at the bottom. The bottoms stream is fed to a rectifier (C-2) with partial condenser, for the removal of by-products. This is followed by an extractive distillation column (C-3), using glycerol as extractant agent with a mass ratio of 1:1 with relation to the amount of ethanol in the stream. A column is used for the extraction of the ammonia present (C-4) in the stream, obtaining ethanol with a purity of 0.99. Finally, a column for the purification of glycerol (C-5) is required, aiming to a 99% recovery. Operating conditions are shown in Table 6.



**Figure 1:** Conventional process.

**Table 6.** Feed conditions for the columns of the conventional process.

Unit	Temperature (°C)	Pressure (bar)
C-1	34	1
C-2	100.46	1.33
C-3	57.51	1
C-4	41.73	1.01
C-5	170.76	1.70

Following the simulation of the conventional process, two schemes are proposed for the intensification of the ethanol distillation sequence, aiming to reduce the number of equipment and the thermal load. Figure 1 highlights the area to be intensified in red. Characteristics of the schemes are obtained by recursive simulation and sensitivity

analysis to determine the configuration with the lowest heat duty, while maintaining the specified ethanol purity.

#### 4. Results

The first scheme (I-1) consists of replacing C-5 by a side rectifier, intending to reduce the heat duty while reducing the number of units. Figure 2 shows the scheme. The ethanol-water mixture leaving the column C-2 is fed into the thermally coupled extractive column C-3. The design consists of 30 stages, where glycerol is feed at stage 2 and the BE<sub>2</sub> stream is fed at stage 10. Interconnecting flows are located at stage 29.

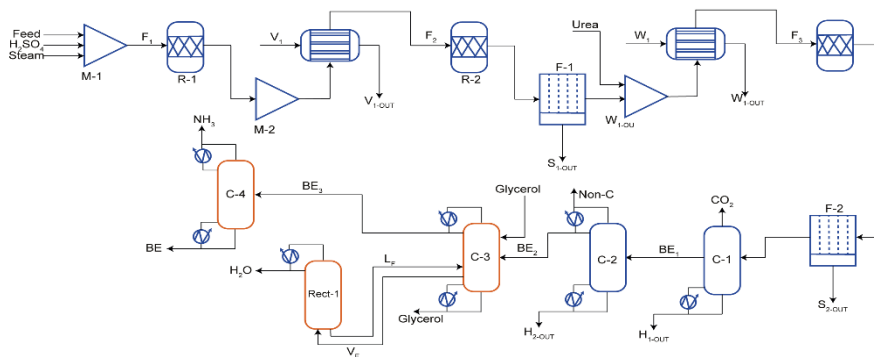


Figure 2: Intensified scheme 1 (I-1).

The second intensified scheme (I-2) involves replacing column C-4 by a side stream where ethanol is concentrated, as shown in Figure 3. In this scheme, the ammonia present in the mixture is obtained as top product in C-3. The column C-3 has 30 stages, where the glycerol enters in stage 10 and the stream BE<sub>2</sub> is fed at stage 25. The side stream is obtained from stage 4. Operating conditions of both schemes are shown in Table 7.

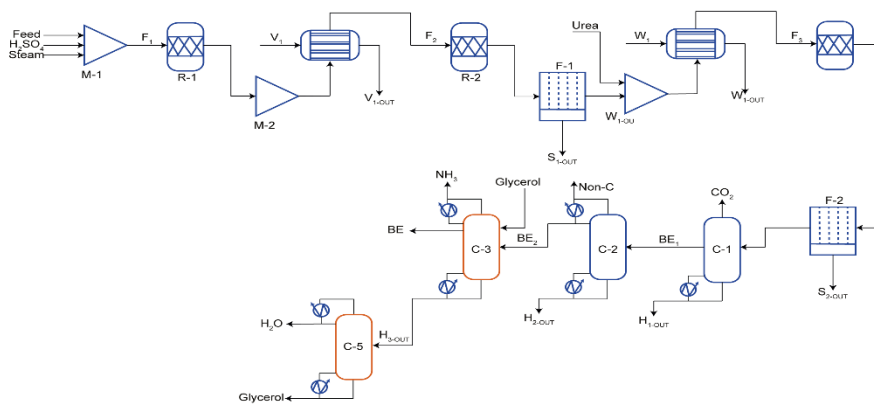


Figure 3: Intensified scheme 2 (I-2).

Table 7. Feed conditions to the columns of the intensified schemes.

Scheme, unit	Temperature (°C)	Pressure (bar)
I-1, C-4	42.01	1
I-2, C-5	192.02	2.18

Table 8 shows that I-1 consumes 2.36 % less energy than the conventional scheme. Meanwhile, I-2 has an energy reduction of 1.82 %. In both intensified schemes, 99% of the available bioethanol and 99% of the glycerol are recovered, being the heat duty in the reboilers the representative variable of these implementations.

**Table 8.** Comparison of schemes

Variable	Scheme	Conventional	I-1	I-2
Q (kW)		150,200.63	146,654.3	147,463.47
Bioethanol Recovery (kg/h)		14,472.67	14,435.15	14,431.04
Glycerol Recovery (kg/h)		14,540.74	14,542.23	14,536.96

## 5. Conclusions

The design of the conventional bioethanol production process from sorghum residues has been carried out. Two intensification alternatives have been proposed for the purification stage, with the goal of reducing the number of units and the thermal duty without compromising the recovery of the biofuel and the solvent. Scheme I-1 achieves a reduction of 3,546.33 kW as well as the incorporation of the glycerol recovery in a thermally coupled system. The recovery of bioethanol and glycerol is maintained at the target values of 99%. Finally, it is important to mention that the implementation of a rigorous optimization algorithm can achieve lower thermal duties with higher recoveries for the proposed schemes.

## References

- M.S., Alam, M. S. Tanveer, 2020, Conversion of biomass into biofuel: a cutting-edge technology, In L. Singh, A. Yousuf, D. Madhab (Eds.), *Bioreactors*, Elsevier, p. 55–74.
- Banco Mundial, 2022, Consumo de energía procedente de combustibles fósiles (% del total), <https://datos.bancomundial.org/indicador/EG.USE.COMM.FO.ZS>, last consulted on November 23, 2022. (Spanish).
- C. Conde-Mejía, A. Jiménez-Gutiérrez, M. M. El-Halwagi, 2013, Assessment of combinations between pretreatment and conversion configurations for bioethanol production, *ACS Sustainable Chemistry & Engineering*, 1, 8, 956–965.
- C. Conde-Mejía, A. Jiménez-Gutiérrez, M. M. El-Halwagi, 2012, A comparison of pretreatment methods for bioethanol production from lignocellulosic materials, *Process Safety and Environmental Protection*, 90, 189–202.
- M. A. Costagliola, M. V. Prati, S. Florio, P. Scorletti, D. Terna, P. Iodice, D. Buono, A. Senatore, 2016, Performances and emissions of a 4-stroke motorcycle fuelled with ethanol/gasoline blends, *Fuel*, 183, 470–477.
- A. Duque, P. Doménech, C. Alvarez, M. Ballesteros, P. Manzanares, 2020. Study of the bioprocess conditions to produce bioethanol from barley straw pretreated by combined soda and enzyme-catalyzed extrusion, *Renewable Energy*, 158, 263–270.
- Energy Information Administration, 2022, How much carbon dioxide is produced from U.S. gasoline and diesel fuel consumption?, <https://www.eia.gov/tools/faqs/faq.php?id=307&t=10>, last consulted on November 23, 2022.
- Food and Agriculture Organization of the United Nations (FAO), 2022, Crops and livestock products, <https://www.fao.org/faostat/en/#data/QCL>, last consulted on November 20, 2022.
- National Renewable Energy Lab, 1996, Development of an aspen plus physical property database for biofuels components (no. nrel/tp-425-20685), Tech. rep., NREL, United States. <https://www.nrel.gov/docs/legosti/old/20685.pdf>, last consulted on November 23, 2022.
- J. G. Segovia-Hernández, M. Vázquez-Ojeda, F. Gómez-Castro, C. Ramírez-Márquez, M. Errico, S. Tronci, B. G. Rong, 2014, Process control analysis for intensified bioethanol separation systems, *Chemical Engineering and Processing: Process Intensification*, 75, 119–125.

# Optimal ship-based CO<sub>2</sub> transport chains from Mediterranean emission points to the North Sea

Federico d'Amore\*, Matteo C. Romano

*Politecnico di Milano, Department of Energy, via Lambruschini 4, IT-20156 Milano (Italy).*

*federico.damore@polimi.it*

## Abstract

Carbon capture and storage is one option for reducing industrial CO<sub>2</sub> emissions. Ship CO<sub>2</sub> transport is gaining interest due to its potential competitive cost, lower risk, and higher flexibility with respect to pipelines. This work focusses on the cost of ship CO<sub>2</sub> transport chains in the European context. The objective is to develop an economic model of CO<sub>2</sub> transport via ship, including the chain stages: liquefaction, buffer storage, loading, ship, conditioning, and unloading. An optimisation model is proposed to determine the minimum transport cost from Southern European ports to permanent geological storage in the North Sea. The minimum cost is found around 26 €/t of CO<sub>2</sub> for a total transported CO<sub>2</sub> of 103 Mt/y, which may be cost-competitive with pipelines for long distance routes via large vessels (greater than 50 kt).

**Keywords:** Carbon capture and storage; CO<sub>2</sub> ship transport; chain logistics; optimisation; mixed integer linear programming.

## 1. Introduction

Energy and industry carbon dioxide (CO<sub>2</sub>) sources generate more than half of the annual worldwide CO<sub>2</sub> emissions (IEA, 2022). Carbon capture and storage (CCS) consists in a range of technologies that can separate the CO<sub>2</sub> from these concentrated sources, to avoid its release to the atmosphere (Bui et al., 2018). Then, the CO<sub>2</sub> is compressed, purified, transported, and permanently stored in geological sequestration sites (IPCC, 2005).

Europe has significant potential for CO<sub>2</sub> geological sequestration, especially in the offshore area of the North Sea (Holler and Biebahn, 2011) where different CCS research projects are in an advanced phase of implementation (e.g., Northern Lights, 2022), while storage basins identified in the Mediterranean area (e.g., in the Adriatic Sea and off the coasts of Greece) currently exhibit an early-stage degree of investigation and limited capacity (Donda et al., 2011; Koukouzas et al., 2021). Ship CO<sub>2</sub> transport could foster the development of CCS projects in European regions that are far from the North Sea (e.g., the Mediterranean), and such marine alternative to pipelines may be key due to its potentially competitive cost (Weihs et al., 2014), greater flexibility (Neele et al., 2014), and lower risk (Kjarstad et al., 2016).

A ship-based CO<sub>2</sub> transport chain is constituted by multiple stages: liquefaction, intermediate storage, naval shipping, conditioning, and unloading (IEAGHG, 2020); each stage comprises several options in terms of process parameters (e.g., the transport pressure, Roussanaly et al., 2021) and chain design alternatives (e.g., the unloading strategy, Roussanaly et al., 2013; 2014). This leads to the necessity of addressing such combinatorial complexity from a systems perspective. The objective of this work is to provide a detailed economic model of CO<sub>2</sub> transport by ship; the economic model is exploited for the development of a mixed integer linear programming (MILP)



optimisation framework, to minimise the total cost of a European ship-based CO<sub>2</sub> transport chain. This optimisation model is tested on a transport chain including the most significant industrial CO<sub>2</sub> emission points in southern Europe located in the vicinity of strategic CO<sub>2</sub> collection ports, and geological sequestration in the North Sea. Different unloading scenarios and carbon reduction targets are investigated. The resulting minimum unitary transport cost is 26 €/t, for transporting more than 103 Mt/y of CO<sub>2</sub>, which may be competitive with pipeline transport for very long-distance routes.

## 2. Modelling framework

A ship-based CO<sub>2</sub> transport chain is constituted by several sequential stages  $i = \{liq, stor, load, ship, cond, unload, pipe\}$  (Figure 1). This study assumes that the CO<sub>2</sub> arrives from the capture plant as a pressurised gas at pressure  $P_{in}$ . In the liquefaction plant (*liq*) the CO<sub>2</sub> is liquefied through compression at its transport pressure  $P_t$ . This work is based on a low  $P_t$  design (7 barg, -50°C), as low  $P_t$  benefits from lower costs with respect to medium  $P_t$ , if considering the whole transport chain (IEAGHG, 2020; Roussanaly et al., 2021). Then, the CO<sub>2</sub> is stored (*stor*) in buffer tanks to ensure continuous operation, to be subsequently loaded (*load*) onto ships (*ship*). This study considers vessel capacities  $C_{ship}$  [t] from 10 kt to 100 kt per ship. The transport stages downstream ship transport depend on the unloading strategy (*unload*): (S1: port-to-port) the CO<sub>2</sub> is unloaded onshore, stored in an unloading buffer storage facility, then sent to offshore sequestration via an offshore pipeline (*pipe*); (S2: port-to-FSI) the CO<sub>2</sub> is unloaded, through equipment onboard vessels, to a floating storage and injection (FSI) platform with intermediate storage tanks; and (S3: port-to-direct injection) the CO<sub>2</sub> is directly injected to the offshore basin without any unloading buffer storage hence, lower capital investment than S1 and S2, but higher risk in operation due to batch-wise unloading. The final stage is conditioning (*cond*), which increases the CO<sub>2</sub> pressure and temperature for injection.

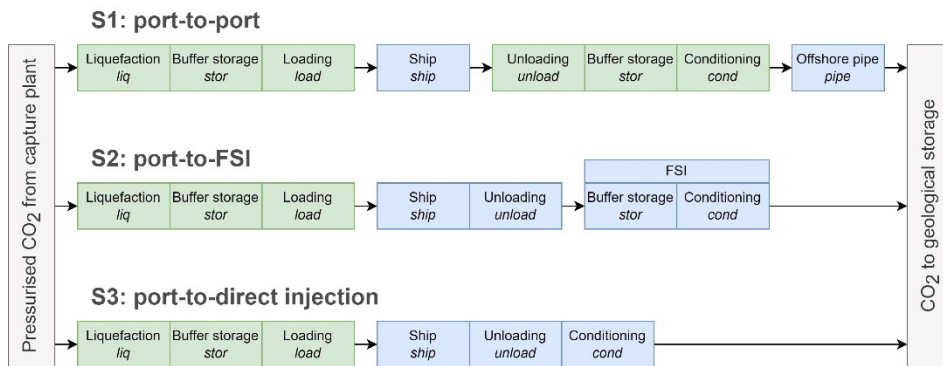


Figure 1. Schematic representation of a ship-based CO<sub>2</sub> transport chain.

## 3. Mathematical formulation

This study optimises a ship-based CO<sub>2</sub> transport chain at a European-scale. It is assumed that the CO<sub>2</sub> is captured from large-scale industrial emitters located in Southern Europe: 5 steel plants for a total emission of 26.5 Mt/y (17 % of sectorial European CO<sub>2</sub>), 25 refineries for a total of 43.8 Mt/y (33 % of sectorial European CO<sub>2</sub>), and 42 cement plants for a total of 33.1 Mt/y (23 % of sectorial European CO<sub>2</sub>) (d'Amore et al., 2021). The spatially-explicit features are described by means of nodes  $n$ , comprising ports and geological sequestration in the North Sea. The CO<sub>2</sub> collection ports are chosen among

large-scale harbours located in the vicinity of selected CO<sub>2</sub> emission points, and the port of Stavanger (Norway) is chosen as a CO<sub>2</sub>-to-storage hub for unloading S1 (Figure 2).

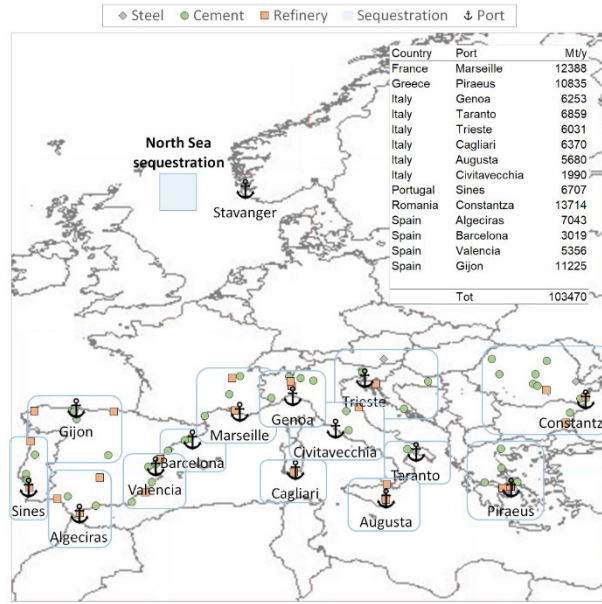


Figure 2. Geographic location of CO<sub>2</sub> emission clusters and collection ports, and yearly CO<sub>2</sub> emissions available at each port for ship transport.

The objective is to minimise the total annual cost  $TC$  [€/y] of the ship transport chain:

$$objective = \min(TC) \quad (1)$$

$$TC = \sum_i TC_i = \sum_i \left( \frac{CAPEX_i}{lifetime} + OPEX_i \right) \quad \forall i \quad (2)$$

being  $TC_i$  [€/y] given by the investment cost ( $CAPEX_i$  [€/y], scaled over a 20 years' *lifetime*) and operational expenditure ( $OPEX_i$  [€/y]) of the stages  $i$  of the transport chain. Costs of stages  $i$  are based on data from IEAGHG (2020) and the methodology is reported in d'Amore et al. (2023). The operational design input parameters are the inlet pressure  $P_{in}$ , the transport pressure  $P_t$ , the vessel capacity  $C_{ship,n,n'}$ , the number of ships  $N_{ship,n,n'}$ , the sailing distance  $d_{n,n'}$  [km], and the operational transit speed  $v_{ship}$  [km/h], to determine the amount of CO<sub>2</sub> transported  $\dot{Q}_{ship,n,n'}$  [t/y] through vessel  $ship$  from  $n$  to  $n'$ , which depends on  $C_{ship,n,n'}$ ,  $N_{ship,n,n'}$ , on the number of roundtrips per year  $N_{ship,n,n'}^{rtrip}$  [ $y^{-1}$ ], and on the maximum vessel filling  $\eta_{tank}$  (set at 95 %):

$$\dot{Q}_{ship,n,n'} = C_{ship,n,n'} \cdot N_{ship,n,n'} \cdot N_{ship,n,n'}^{rtrip} \cdot \eta_{tank} \quad \forall ship, n, n' \quad (3)$$

The total yearly amount of CO<sub>2</sub> transported via ships from  $n$  to  $n'$  is equal to:

$$\dot{Q}_{n,n'} = \sum_{ship} \dot{Q}_{ship,n,n'} \quad \forall n, n' \quad (4)$$

Considering the yearly amount of CO<sub>2</sub> transported  $\dot{Q}_n^{tr}$  [t/y] from port  $n$  and that sequestered  $\dot{Q}_n^{seq}$  [t/y] in node  $n$ , it is possible to impose the mass balance at each port  $n$ :

$$\dot{Q}_n^{tr} + \sum_{n'} \dot{Q}_{n',n} = \sum_{n'} \dot{Q}_{n,n'} + \dot{Q}_n^{seq} \quad \forall n \quad (5)$$

where  $\dot{Q}_n^{tr}$  is bounded by the amount of CO<sub>2</sub> available  $\dot{Q}_n^{capt}$  [t/y] at port  $n$  (Figure 2). The final aim is to sequester a fraction  $\alpha$  [%] of the total CO<sub>2</sub> from capture plants:

$$\sum_n \dot{Q}_n^{tr} = \alpha \cdot \sum_n \dot{Q}_n^{capt} \quad (6)$$

#### 4. Results

The MILP problem was optimised on a 3.0 GHz computer (32 GB RAM), by using GAMS software and CPLEX solver (optimality gap lower than 2 %). It emerges that liquefaction and ship represent about 10 % and 40 %, respectively, of the total investment, independently from the analysed unloading scenario. Other significant contributions to capital expenditures were found to be case-specific: namely offshore pipelines in S1 (up to 39 % of the total investment), FSI unit in the S2 option (comparable with the cost of ships), and conditioning in S3. Variable costs are dominated by ships operation and by liquefaction. Direct unloading (S3) emerges as the most cost-effective scenario. In fact, the results show that the unitary cost for transporting via ship 103 Mt/y of CO<sub>2</sub> is equal to: 26.1 €/t (S1), 26.4 €/t (S2), and 25.6 €/t (S3, though it could be an unfavourable design due to discontinuous injection) (Figure 3). As  $\alpha$  increases the selected ports and distances increase as well, so the unitary cost rises because: (i) more ports determine the installation of more onshore infrastructure; and (ii) more ships are needed as each journey takes longer. In fact, the optimal chain design avoids as much as possible (i.e., for low values of  $\alpha$ ) ports far from Stavanger, to minimise ship transport costs. For instance (Figure 4), the optimal chain for S3 involves the use of few but large ships from Spain and Portugal for  $\alpha=30$  %, while the network expands to more ships and ports located in Mediterranean France and Italy for  $\alpha=60$  %, and towards Greece and the Black Sea for  $\alpha=100$  %.

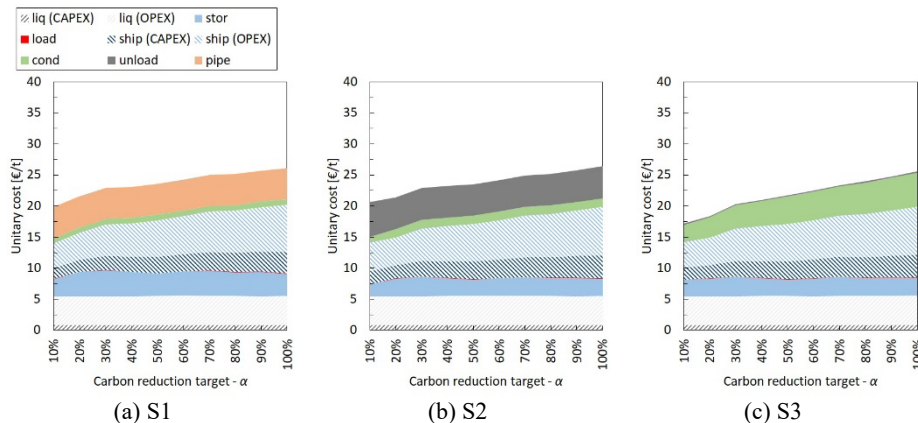


Figure 3. Unitary cost [€/t] depending on unloading scenario and carbon reduction target  $\alpha$ .

Ships investment cost depends on the choice in the vessel size: a decrease in the maximum size from 100 kt to 50 kt produces a +38 % increase in the investment for ships, though with limited effect on the overall cost of the transport chain (+10 %). Differently, deploying small vessels (10 kt) would double the chain cost. At the same time, optimised ship CO<sub>2</sub> transport through large-scale vessels (over 50 kt) is cost competitive with pipelines for long-distance routes, e.g., to connect Greece with the North Sea (Figure 5).

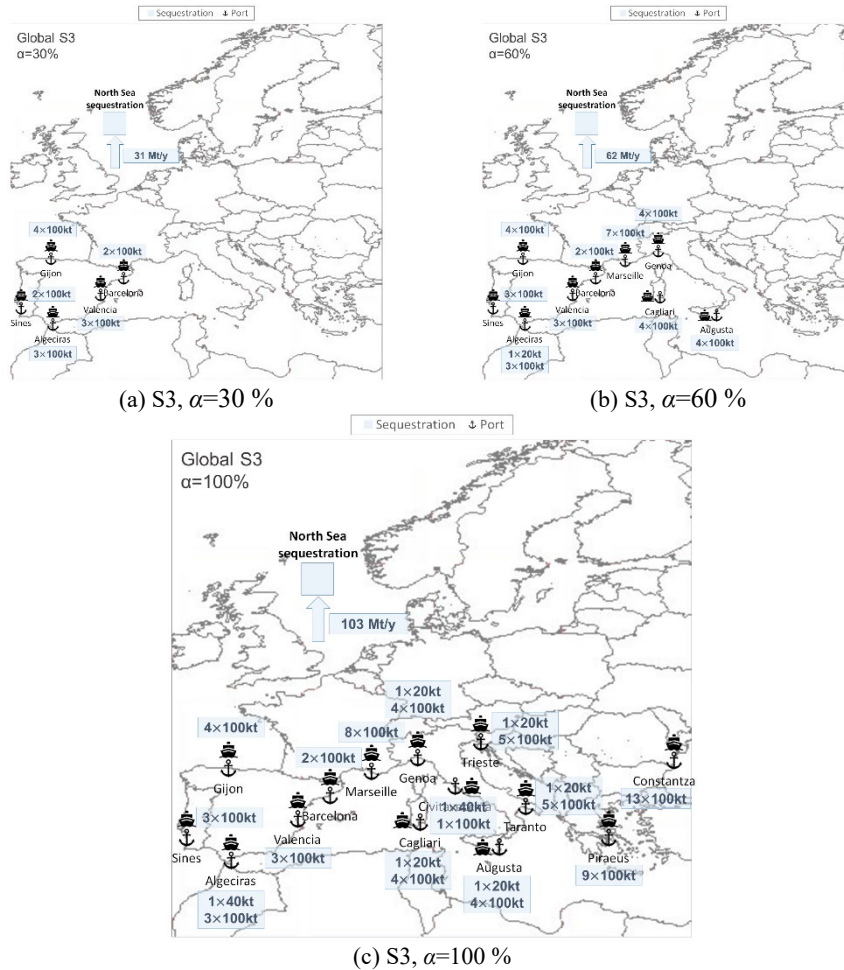


Figure 4. Optimal ship transport chains for S3 under increasing carbon reduction target  $\alpha$ .

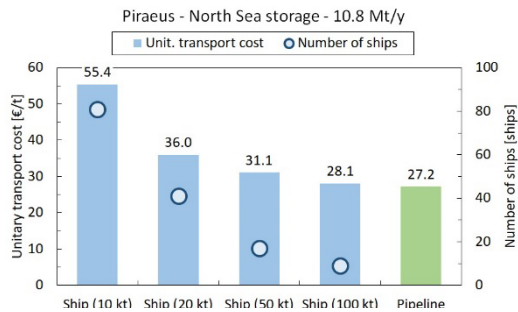


Figure 5. Ship vs. pipeline long-distance CO<sub>2</sub> transport from Greece (Piraeus) to North Sea storage. Unitary transport costs for pipelines are computed from d'Amore et al. (2021).

## 5. Conclusions

This study proposed a modelling tool for the economic and logistic optimisation of ship CO<sub>2</sub> transport chains from large-scale Southern Europe emitters to the North Sea offshore

storage. Depending on the unloading strategy, we found that the minimum unitary cost for transporting via ship 103 Mt/y of CO<sub>2</sub> is equal to 26.1 €/t (port-to-port), 26.4 €/t (port-to-floating storage and injection), and 25.6 €/t (port-to-direct injection), though this latter does not comprise a buffer storage prior to injection and is designed as a batch-wise process. Importantly, it was demonstrated that an optimal design of ship transport via large-scale vessels (50-100 kt) can be cost competitive with pipeline transport for very long distance routes (e.g., from Greece to the North Sea), while the use of small ships (10 kt) results in twice the cost of pipelines.

## References

- M. Bui, C.S. Adjiman, A. Bardow, E.J. Anthony, A. Boston, S. Brown, P.S. Fennell, S. Fuss, A. Galindo, L.A. Hackett, J.P. Hallett, H.J. Herzog, G. Jackson, J. Kemper, S. Krevor, G.C. Maitland, M. Matuszewski, I.S. Metcalfe, C. Petit, G. Puxty, J. Reimer, D.M. Reiner, E.S. Rubin, S.A. Scott, N. Shah, B. Smit, J.P.M. Trusler, P. Webley, J. Wilcox, N. Mac Dowell, 2018, Carbon capture and storage (CCS): The way forward, *Energy Environ. Sci.*, 11, 1062-1176
- F. d'Amore, L. Natalucci, M.C. Romano, 2023, Optimisation of ship-based CO<sub>2</sub> transport chains from Southern Europe to the North Sea, In peer review
- F. d'Amore, M.C. Romano, F. Bezzo, 2021, Optimal design of European supply chains for carbon capture and storage from industrial emission sources including pipe and ship transport, *Int. J. Greenh. Gas Control*, 109, 103372
- F. Donda, V. Volpi, S. Persoglia, D. Parushev, 2011, CO<sub>2</sub> storage potential of deep saline aquifers: The case of Italy, *Int. J. Greenh. Gas Control*, 5, 327-335
- S. Holler, P. Viebahn, 2011, Assessment of CO<sub>2</sub> storage capacity in geological formations of Germany and Northern Europe, *Energy Proc*, 4, 4897-4904
- IEA, 2022, Data and statistics, <https://www.iea.org/data-and-statistics/data-browser?country=WORLD&fuel=Energyconsumption&indicator=CO2Industry>
- IEAGHG, 2020, The status and challenges of CO<sub>2</sub> shipping infrastructures, <https://ieaghg.org/ccs-resources/blog/new-ieaghg-report-the-status-and-challenges-of-co2-shipping-infrastructures>
- IPCC, 2005, IPCC Special Report on Carbon Dioxide Capture and Storage. Prepared by Working Group III of the Intergovernmental Panel on Climate Change, <https://www.ipcc.ch/report/carbon-dioxide-capture-and-storage/>
- J. Kjarstad, R. Skagestad, N.H. Eldrup, F. Johnsson, 2016, Ship transport - A low cost and low risk CO<sub>2</sub> transport option in the Nordic countries, *Int. J. Greenh. Gas Control*, 54, 168-184
- N. Koukoulas, P. Tyrologou, D. Karapanos, J. Carneiro, P. Pereira, F. de Mesquita Lobo Veloso, P. Koutsovitis, C. Karkalis, E. Manoukian, R. Karametou, 2021, Carbon capture, utilisation and storage as a defense tool against climate change: Current developments in west Macedonia (Greece), *Energies*, 14, 3321
- F. Neele, H.A. Haugen, R. Skagestad, 2014, Ship transport of CO<sub>2</sub> - Breaking the CO<sub>2</sub>-EOR deadlock, *Energy Procedia*, 63, 2638-2644
- Northern Lights, 2022, Website, <https://norlights.com/>
- S. Roussanaly, J.P. Jakobsen, E.H. Hognes, A.L. Brunsvold, 2013, Benchmarking of CO<sub>2</sub> transport technologies: Part I - Onshore pipeline and shipping between two onshore areas, *Int. J. Greenh. Gas Control*, 19, 584-594
- S. Roussanaly, A.L. Brunsvold, E.S. Hognes, 2014, Benchmarking of CO<sub>2</sub> transport technologies: Part II - Offshore pipeline and shipping to an offshore site, *Int. J. Greenh. Gas Control*, 28, 283-299
- S. Roussanaly, H. Deng, G. Skaugen, T. Gundersen, 2021, At what pressure shall CO<sub>2</sub> be transported by ship? An in-depth cost comparison of 7 and 15 barg shipping, *Energies*, 14, 5635
- G.A.F. Weihs, K. Kumar, D.E. Wiley, 2014, Understanding the economic feasibility of ship transport of CO<sub>2</sub> within the CCS chain, *Energy Procedia*, 63, 2630-2637

# DynHeat: Heat Exchanger Network Design for Batch Processes via Dynamic Optimization

Dörthe Franzisca Hagedorn,<sup>a</sup> Sören Demandt,<sup>a</sup> Florian Joseph Baader,<sup>b,c</sup>  
Christiane Reinert,<sup>a</sup> Niklas von der Aßen<sup>a</sup>

<sup>a</sup> *Institute of Technical Thermodynamics, RWTH Aachen University, Schinkelstr. 8, 52062 Aachen, Germany*

<sup>b</sup> *Energy & Process Systems Engineering - Department of Mechanical and Process Engineering, ETH Zurich, 8092 Zürich, Switzerland*

<sup>c</sup> *Institute of Energy and Climate Research - Energy Systems Engineering (IEK-10), Forschungszentrum Jülich GmbH, Jülich, Germany*

## Abstract

Heat integration can reduce the consumption of external utilities in batch processes. Batch processes often include process steps with dynamic temperature progression, such as heating in vessels. During heat integration between those process steps, the temperature difference decreases over time. Exploiting the thermodynamic potential for heat integration can, therefore, lead to long process durations. Hence, there is a trade-off between the total cost of heat integration and the process duration. A method is missing to map this trade-off while considering the dynamic temperature progression. We introduce the DynHeat optimization method for heat exchanger network design in dynamic batch processes based on dynamic optimization. We apply the DynHeat method to a case study and perform a multi-objective optimization regarding the process duration and the total cost. We find that the DynHeat method can propose a suitable heat exchanger network design and operation mode for the regarded case study.

**Keywords:** Energy Efficiency, Heat Integration, Pinch Method, Logarithmic Mean Temperature Difference, Dynamic Processes

## 1. Introduction: Heat Integration in Batch Processes

Meeting thermal energy demands in industrial processes by external utilities is economically and environmentally expensive. Heat integration can reduce the consumption of external utilities by reusing waste heat. For heat integration, a heat exchanger network needs to be built. To obtain cost-optimal heat exchanger networks, the total cost should be considered, comprising operational costs for utility consumption and investment costs for heat-exchangers.

Common existing approaches for heat exchanger network design, based on the Pinch Method (Linnhoff and Flower, 1978) or the superstructure model by Yee and Grossmann (1990) and most of their extensions, reviewed by Klemeš and Kravanja (2013), assume stationary process streams. In practice, many industries work with instationary batch processes. In batch processes, the products are often cooled down or heated up in vessels, which leads to temperatures changing over time. When using heat integration in such processes, the transferred heat between two products decreases over time as the temperature difference between those products decreases. Thus, exploiting the full thermodynamic potential for heat integration between two products leads to long process durations. Hence, heat integration in dynamic batch processes causes a trade-off between process

duration and total cost for heat integration. This trade-off needs to be considered when designing heat exchanger networks for dynamic batch processes. A suitable method for heat exchanger network design for dynamic batch processes should, thus, represent the dynamic temperature profiles accurately to determine both the amount of integrated heat and the process duration.

In practice, the Pinch Method is commonly used for estimating the thermodynamic potential for heat integration of a process. The Pinch Method can be adapted to varying thermal demands by averaging the demands over time (time-average model) or by dividing the time horizon into several time slices (time-slice model) (Kemp and Lim, 2020). However, both approaches still assume temporary stationary process streams. Vaselenak et al. (1986) propose a heuristic method to handle different types of temperature progression of batch products in tanks in heat integration problems. Dowidat et al. (2014) extend the time-slice model to dynamic temperature progression by introducing additional time slices. However, both approaches do not explicitly design a heat exchanger network. In Dowidat et al. (2016), the authors introduce a match ranking matrix for the design of economically efficient heat exchanger networks. However, the approach relies on heuristic rules instead of considering the actual costs. Furthermore, in all the approaches mentioned above, the timing of the processes needs to be known in advance, and the trade-off between process duration and total costs, is not taken into account. Castro et al. (2015) introduce an optimization model for combined heat integration and scheduling in order to optimize the makespan of the process. However, in their approach, matches between process streams can be derived but no design optimization is included.

In conclusion, no optimization method exists that can represent the trade-off between process duration and total cost for thermal energy supply in heat integrated dynamic batch processes. In this contribution, we bridge this gap by developing the DynHeat optimization method for heat exchanger network design of dynamic batch processes while taking the trade-off between total cost and the process duration into account.

## 2. The DynHeat Method

The aim of the DynHeat method is to design a heat exchanger network for dynamic batch processes while taking the trade-off between total cost, comprising operational costs and investment costs, and the process duration into account. In the corresponding DynHeat optimization model, we model products that either need to be heated or cooled in vessels. The thermal demands can be provided by external utilities or via heat integration. For heat integration, there is the option of building an external heat exchanger between every two vessels. A scheme of the DynHeat model is shown in *Figure 1*.

To address the trade-off between cost and duration, we perform a multi-objective optimization with the total cost  $C^{total}$  of one process run as one objective function and the process duration  $\tau^{final}$  as the other objective function. The total cost  $C^{total}$  for one process run are composed of the investment cost for the heat exchangers and operational cost for external utility:

$$C^{total} = \left( \sum_{h \in H} (A_h^{cu} + A_h^{HI}) + \sum_{c \in C} (A_c^{hu} + A_c^{HI}) + \sum_{h \in H} \sum_{c \in C} A_{h,c}^{HE} \right) c^A + \sum_{h \in H} Q_h^{cu} c_h^{cu} + \sum_{c \in C} Q_c^{hu} c_c^{hu} \quad (1)$$

With the surface areas of the heat exchangers for hot utility ( $A_h^{cu}$ ), cold utility ( $A_c^{hu}$ ), hot product to heat transfer media (HTM) ( $A_h^{HI}$ ) and cold product to HTM ( $A_c^{HI}$ ) and the

external heat exchanger ( $A_{h,c}^{HE}$ ). The cost factor  $c^A$  refers to the specific investment costs for the heat exchanger surface areas scaled down to one process run. The operational costs comprise the amount of heat supplied by the cold utility  $Q_h^{cu}$  and hot utility  $Q_c^{hu}$  and the specific price for cold utility  $c^{cu}$  and hot utility  $c^{hu}$ .

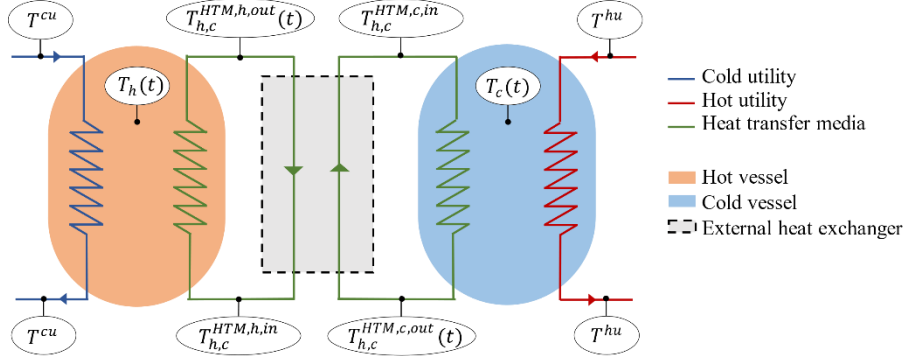


Figure 1: Exemplary scheme of the DynHeat model for one hot and one cold product. The left vessel (orange) contains the hot product of the temperature  $T_h(t)$  which can be cooled down by cold utility of the temperature  $T^{cu}$  (dark blue). Analogously the cold product in the right vessel (light blue) of the temperature  $T_c(t)$  can be heated up by hot utility (red) of the temperature  $T^{hu}$ . Both vessels can be connected by an external heat exchanger (grey) with heat transfer media (HTM) (green line). The temperature of the HTM at the inlet of the vessels  $T_{h,c}^{HTM,h,in}$  and  $T_{h,c}^{HTM,c,in}$ , is constant.

In the DynHeat model, we make the following assumptions: The vessels are perfectly mixed, and no heat losses occur. The external heat exchangers are counter-flow heat exchangers. All heat capacities, heat transfer coefficients, and the temperature of the utilities are constant. The HTM is not allowed to function as a storage, i.e., the total integrated heat flow must be transferred from the hot product to the cold product immediately, and the temperature of the HTM at the inlets of the vessels  $T_{h,c}^{HTM,c,in}$  and  $T_{h,c}^{HTM,h,in}$  is constant. The HTM between a hot and a cold product has the same heat capacity rate on both sides of the external heat exchanger, i.e., the temperature difference along the external heat exchanger  $(T_{h,c}^{HTM,h,in} - T_{h,c}^{HTM,h,in}(t)) = (T_{h,c}^{HTM,h,out}(t) - T_{h,c}^{HTM,c,in})$  is constant along the heat exchanger surface area.

In the following, we describe the DynHeat model in detail. The most important equations are given for hot products that need cooling. The equivalent equations for cold products that need heating can be derived analogously.

The temperature profile  $T_h(t)$  of a hot product  $h$  with the mass  $m_h$  and heat capacity  $c_{p,h}$  changes due to heat transferred to cold utility  $\dot{Q}_h^{cu}(t)$  or by heat integration with a cold product  $\dot{Q}_{h,c}^{HI}(t)$ :

$$\frac{dT_h(t)}{dt} = - \frac{(\sum_{c \in C} \dot{Q}_{h,c}^{HI}(t)) + \dot{Q}_h^{cu}(t)}{m_h \cdot c_{p,h}} \quad \forall h \in H \quad (2)$$

The heat flow from a hot product to the HTM  $\dot{Q}_{h,c}^{HI}(t)$  is defined by

$$\dot{Q}_{h,c}^{HI}(t) = \dot{m}_{h,c}^{HTM} c_p^{HTM} (T_{h,c}^{HTM,h,out}(t) - T_{h,c}^{HTM,h,in}) \quad \forall h \in H, \forall c \in C \quad (3)$$

, where  $\dot{m}_{h,c}^{HTM}$  is the given mass flow of the HTM between hot vessel  $h$  and cold vessel  $c$  and  $c_p^{HTM}$  is the given constant heat capacity of the HTM.

The temperature of the HTM at the outlet of the hot vessel  $T_{h,c}^{HTM,h,out}(t)$  is defined by



the temperature progression along the heat exchanger between the hot product and the HTM (Glück, 2017):

$$T_{h,c}^{HTM,h,out}(t) = T_{h,c}^{HTM,h,in}(t) \cdot e^{-\frac{k A_{h,c}^{HTM,h}}{\dot{m}_{h,c}^{HTM} c_p^{HTM}}} + T_h(t) \cdot \left(1 - e^{-\frac{k A_{h,c}^{HTM,h}}{\dot{m}_{h,c}^{HTM} c_p^{HTM}}}\right) \quad \forall h \in H, \forall c \in C \quad (4)$$

, where  $A_{h,c}^{HTM,h}$  is the surface area of the heat exchanger between the hot product and the HTM, and  $k$  is the constant heat transfer coefficient of the heat exchanger.

Finally, the required surface area of the external heat exchanger  $A_{h,c}^{HE}$  can be derived by the transferred heat and the temperatures of the heat transfer medium by

$$\dot{Q}_{h,c}^{HI}(t) \geq k A_{h,c}^{HE} \Delta T_{h,c}^{HTM}(t) \quad \forall h \in H, \forall c \in C \quad (5)$$

, where the temperature difference  $\Delta T_{h,c}^{HTM}(t)$  describes the temperature difference in the external heat exchanger. As in Verheyen and Zhang (2006), we assume that the biggest surface area needed at any time is installed and that all other operation modes can be realized by bypassing a part of the streams. The heat exchanger surface area for utility supply can be derived analogously. We assume a sequential process, meaning that external utilities are only used after heat integration is complete. Logical big-M constraints denote the beginning and end of utility supply and heat integration.

To solve the resulting dynamic-algebraic optimization problem, we discretize the time by orthogonal collocation on finite elements using `pyomo.dae` (Nicholson et al., 2018). Through the collocation, the discretized problem is a fully algebraic mixed-integer nonlinear optimization problem (MINLP) (Biegler, 2010).

### 3. Application to a Case Study

We apply the DynHeat method to a case study with one hot and one cold product. We adapt our process parameters to a milk pasteurization process, where fresh milk is first heated to a certain temperature and, afterwards, is cooled down to its final storing temperature. We assume that the heating of one batch and the cooling of another batch take place simultaneously. We adapt the process parameters from a case study in Fellows (2017).

We compute the case study on an Intel(R) Xeon(R) CPU E5-1660 v4 with 3.30GHz running on Microsoft Windows Server 2016 Standard with kernel version 10.0.14393. We use python 3.7 as a modeling language. As a solver, we use baron version 20.4.14 (Kılınç and Sahinidis, 2018). We apply eight threads. The relative gap is set to 1E-8. For the collocation method, we choose seven finite elements and 3 collocation points per element. We set a limit of 1000 *sec* to the time horizon to facilitate the solution of the problem.

With the DynHeat model, we perform a multi-objective optimization of the given case study, minimizing the process duration and the total cost of one process run. Figure 2 (left) shows the resulting Pareto front for the selected case study. The Pareto front visualizes the trade-off between the total cost for one process run and the process duration.

In the left anchor point, the total cost is minimal. Heat integration takes place as long as possible, such that the smallest possible amount of utility is used. However, thereby the maximal allowed time horizon of 1000 *sec* is required. In contrast, the right anchor point shows the results for minimal process duration. For minimal process duration, the thermal demands are covered by utility only. By the increased use of external utility, the total cost increase by almost 60% compared to the minimal total cost. Accordingly, heat integration can significantly reduce the total cost in the regarded case study.

Furthermore, there is one solution point between the two anchor points, representing a compromise between process duration and total cost. Figure 2 (right) shows the temperature profiles and the heat flows for the compromise solution. The heat integration is stopped before the temperature of the hot and the cold product reach the same level. Instead, a temperature difference of 19.02 K remains and utility is used to cover the remaining demands. This compromise increases the total cost only slightly, while reducing the process duration by almost 600 sec compared to the left anchor point.

All in all, the Pareto front shows that heat integration can reduce the total cost significantly in the regarded case study. However, heat integration extends the process duration notably. Given this trade-off, the DynHeat method also offers solutions with a finite process duration that still significantly benefit from heat integration.

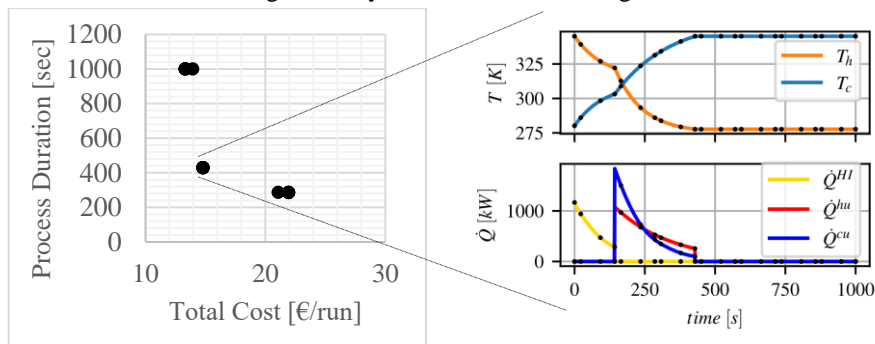


Figure 2: Left: Resulting Pareto front of the DynHeat method for the selected case study. Right: Process operation for one middle point on the Pareto front. The upper diagram shows the temperature progression of the hot product  $T_h(t)$  (orange) and the cold product  $T_c(t)$  (blue) over time. The lower diagram shows the corresponding heat flows (heat integration  $\dot{Q}_{h,c}^{HI}(t)$  (yellow), cold utility  $\dot{Q}_h^{cu}(t)$  (blue) and hot utility  $\dot{Q}_c^{hu}(t)$  (red)).

While the orthogonal collocation in the DynHeat method allows for a more accurate representation of the dynamic temperature profiles than existing approaches, it still leads to some discretization errors. The dynamic profiles are not represented exactly but approximated by polynomials. In our case, the DynHeat method leads to a discretization error concerning the integrated heat of 1.1 %. Thus, we conclude that the chosen settings are sufficient for accurate representation of the temperature profiles. Furthermore, due to the discretization, the process can only switch from heat integration to utility supply, and also finish the utility supply, at the boundaries of a finite element of the discretized time horizon. Consequently, the obtained solutions highly depend on the number of finite elements. The discretization, thus, restricts the solution space, but all solutions of the discretized problem are feasible to the original problem.

With the chosen settings, we find that the DynHeat model does not terminate to global optimality within the allowed computing time of 3000 sec for all points on the pareto front. Thus, further research on the model complexity and solvability needs to be done. However, the results show that the DynHeat model can propose reasonable solutions for heat integration in batch processes while considering dynamic temperature changes.

#### 4. Conclusion

In this contribution, we introduce the DynHeat method for heat exchanger network design in dynamic batch processes. We model the temperature dynamics of the batch process and discretize the problem by orthogonal collocation to a fully algebraic MINLP. To the

best of our knowledge, the DynHeat method is the first optimization method that meets the trade-off between process duration and total cost for heat integration in dynamic batch processes while taking design decisions for heat exchangers into account. We find that the DynHeat method enables us to find solutions that offer a compromise between total cost and process duration.

### Acknowledgements

This study is funded by the German Federal Ministry of Economic Affairs and Energy (ref. no.: 03EN2031D). FB received financial support from the Swiss Federal Office of Energy through the project "SWEET PATHFINDER". CR received financial support by the Ministry of Economics, Innovation, Digitalization and Energy of North-Rhine Westphalia (ref. no.: EFO 0001G). We gratefully acknowledge all the support.

### References

- L.T. Biegler, 2010, *Nonlinear programming*. Society for Industrial and Applied Mathematics, Philadelphia, Pa.
- P.M. Castro, B. Custódio, H.A. Matos, 2015, Optimal scheduling of single stage batch plants with direct heat integration, *Comput Chem Eng*, vol. 82, 172
- C. Dowidat, M. Kalliski, G. Schembecker, C. Bramsiepe, 2016, Synthesis of batch heat exchanger networks utilizing a match ranking matrix, *Appl Therm Eng*, vol. 100, 78
- C. Dowidat, K. Ulonska, C. Bramsiepe, G. Schembecker, 2014, Heat integration in batch processes including heat streams with time dependent temperature progression, *Appl Therm Eng*, vol. 70, 1, 321
- P. Fellows, 2017, *Food processing technology*. Woodhead Publishing, Duxford, United Kingdom
- B. Glück, 2017, *Wärmeübertrager, Raumheizflächen, Behälteraufheizung*, Verl. für Bauwesen, Berlin
- I.C. Kemp, J.S. Lim, 2020, *Pinch Analysis for Energy and Carbon Footprint Reduction: User Guide to Process Integration for the Efficient Use of Energy*. Butterworth-Heinemann
- M.R. Kılınç, N.V. Sahinidis, 2018, Exploiting integrality in the global optimization of mixed-integer nonlinear programming problems with BARON, *Optimization Methods and Software*, vol. 33, 3, 540
- J.J. Klemeš, Z. Kravanja, 2013, Forty years of Heat Integration: Pinch Analysis (PA) and Mathematical Programming (MP), *Curr Opin Chem Eng*, vol. 2, 4, 461
- B. Linnhoff, J.R. Flower, 1978, Synthesis of heat exchanger networks: I. Systematic generation of energy optimal networks, *AIChE J*, vol. 24, 4, 633
- B. Nicholson, J.D. Sirola, J.-P. Watson, V.M. Zavala, L.T. Biegler, 2018, pyomo.dae: a modeling and automatic discretization framework for optimization with differential and algebraic equations, *Math Prog Comp*, vol. 10, 2, 187
- J.A. Vaselenak, I.E. Grossmann, A.W. Westerberg, 1986, Heat integration in batch processing, *Ind Eng Chem Process Des Dev*, vol. 25, 2, 357
- W. Verheyen, N. Zhang, 2006, Design of flexible heat exchanger network for multi-period operation, *Chem Eng Sci*, vol. 61, 23, 7730
- T.F. Yee, I.E. Grossmann, 1990, Simultaneous optimization models for heat integration—II. Heat exchanger network synthesis, *Comput Chem Eng*, vol. 14, 10, 1165

# Process Simulation Approaches to Reduce Commercial Risk of CCS Projects

Ryan Muir<sup>a</sup>, Ralph Cos<sup>b</sup>, Chloe Smith<sup>b</sup>, Stelios Papastratos<sup>c</sup>

<sup>a</sup>*AVEVA, Lake Forest, California, United States of America*

<sup>b</sup>*AVEVA, Munich, Bavaria, Germany*

<sup>c</sup>*AVEVA, Thessaloniki, Greece*

## Abstract

Sustainability is increasingly emerging as a license to operate for Owner-Operator and EPC companies. As industry looks towards sustainable processes that reduce carbon emissions, Carbon Capture and Storage (CCS) has emerged as a viable way of increasing sustainability while reducing operational costs from emissions. For CCS technologies, process simulation is an essential tool for evaluating the economic and environmental feasibility of greenfield and retrofit projects with accelerated schedules and limited capital. When integrated with plant data, process simulation also enables real-time accounting of carbon emissions from operations. Accurate and timely accounting of process emissions is critical for justifying CCS project costs, evaluating different CCS technologies, and quantifying emission and cost savings after project execution. In this paper, we will show how this new approach enables engineers to quickly evaluate CCS alternatives by considering cost vs. relative capture rates to further reduce the commercial risk of new projects.

**Keywords:** Process simulation, sustainability, reducing commercial risk, emission monitoring, economics of CCUS.

## 1. Carbon Capture and Storage

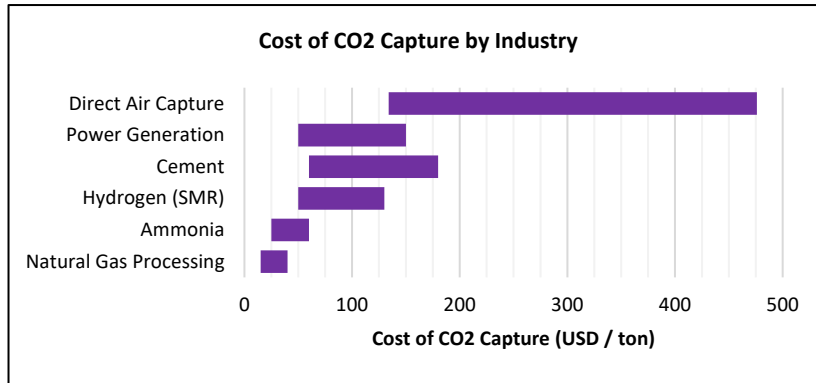
The global carbon capture market is expected to grow from \$1.9 billion in 2020 to \$7.0 billion by 2030. This growth represents a Compound Annual Growth Rate (CAGR) of nearly 13.8% between 2021 and 2030 (Allied Market Research, 2021). This growth is primarily attributed to emission reduction goals set by governments around the world and the companies responsible for reducing emissions to achieve these goals. CCS technologies allow for emission reduction with limited impact on existing processes, making this technology a good candidate for reducing emissions in the short- to medium-term. For some sectors like steel and cement, where renewable technologies are not readily available or scalable, CCS may be the only cost-effective strategy for reducing emissions.

In addition to reducing emissions from existing processes, carbon capture technologies are critical for supporting the ongoing hydrogen energy transition. Green hydrogen technologies like electrolysis offer a long-term solution to decarbonization. However, many of these technologies are not mature enough today to meet the rising demand for hydrogen. In 2021, Steam Methane Reforming (SMR) of natural gas accounted for nearly 60% of hydrogen production. However, SMR can have an emission factor (kg of CO<sub>2</sub>-equivalent emissions per kg of hydrogen produced) of over 12 kg/kg (Mehmeti et al., 2018). Despite the high emission factor, SMR will be critical to meeting the short- to medium-term demands of the hydrogen market. Carbon capture offers a ‘best of both

worlds' approach that allows for continued use of SMR for hydrogen production while offering a path to reduce emissions from the process by up to 90% (IEAGHG, 2017).

## 2. Commercial Challenges

One common challenge of CCS projects is the cost of operating these units. The cost of CCS, measured as the cost per ton of CO<sub>2</sub> captured, is inversely related to the CO<sub>2</sub> concentration in the process streams. **Figure 1** shows the cost of carbon capture for a variety of industries (Balylin-Stern and Berghout, 2021).



**Figure 1-** Relative cost of carbon capture by industry (Balylin-Stern and Berghout, 2021)

In addition to the operating cost of carbon capture units, companies must also weigh the capital cost and commercial risk of these projects. Recent analysis found that, on average, large capital projects overrun planned capital expenditure by 80% and are delivered 20 months beyond the original schedule (Agarwal et al., 2016). Due to the immediate need for carbon emission reduction, accelerated schedules for CCS projects will remain important for the success of this technology. Digital tools provide opportunities to reduce cost and schedule risk by integrating modern work process with next-generation modeling technologies.

## 3. Reducing Commercial Risk with Digital Tools

### 3.1. Introduction

Government incentives, the growing cost of carbon emissions from carbon tax programs, and emerging technologies have made CCS projects more feasible today than ever before. However, significant capital risk for these projects still exists. Government incentives are tied closely to current political ideology and do not provide a long-term guarantee of success; carbon tax programs are difficult to forecast throughout the length of a project; and emerging technologies are unproven and may present technical challenges that increase cost and project timelines.

Process simulation stands to reduce many of these risks by allowing owner-operator and EPC companies to gain greater insight into the technical and economic impact of decisions made while designing a new CCS unit. Solvent system, process flow (for example, pre-combustion vs. post-combustion carbon capture), and operating condition selections all impact the operating and capital cost of CCS projects. As shown in this study, the impact of these selections can be modeled and studied in detail using a next-generation process simulation platform.

### *3.2. Emission Calculations and Cost Accounting*

As a first step in any CCS project, companies must understand the magnitude of their existing carbon emissions and the associated operational expenditure from these emissions. As Carbon Tax and Cap-and-Trade programs become more widespread, accurate real-time accounting of process emissions will be critical for justifying the additional capital and operational costs of CCS technologies. Many companies today rely on complex spreadsheets for this accounting. However, collecting accurate data for these spreadsheets is cumbersome and does not allow for real-time emission monitoring and accounting. Modern process simulation platforms allow engineers to easily calculate carbon emissions, operational cost, and capital expenditure for any process.

Adding both emission and cost calculations to a single process simulation model allow for process-wide optimization of both values. For example, AVEVA Process Simulation allows users to simultaneously calculate overall emissions for any process and include these emissions in the total operating cost. Alongside built-in operating and capital cost estimates, users can leverage the application's optimization features to find the ideal tradeoff between carbon capture rates and operating/capital spend for a CCS project.

These values can also be monitored in real-time without the need for additional monitoring equipment (i.e., soft sensing) by connecting the simulation to a process historian and live process data. Accurate and timely accounting of process emissions is critical for justifying CCS project costs, evaluating different CCS technologies, and quantifying emission and cost savings after project execution.

### *3.3. Digital Twins: Integrated Process Simulation Models*

Process simulation has historically been used during the Conceptual Design and FEED stages of projects. At later stages of the project, simulation models remain largely disconnected from other project data and are often not used to diagnose operational issues. Next-generation simulation tools, however, break down the silos between process engineers and other project disciplines by connecting models to engineering and operational databases. These connections enable first-principle driven real-time monitoring and 'evergreen' models that receive project and asset data from common engineering databases. Interfaces between process simulation models, engineering databases, and real-time operational data are the basis for Engineering and Operating Digital Twin models. Such a unified approach to project design, with integrated simulation, engineering, and design data increases design efficiency by up to 30% and may reduce the total installed cost of a CCS project by up to 5% (AVEVA, 2020).

## **4. Process Simulation Case Study**

### *4.1. Introduction*

While traditional process simulation tools have been sufficient to model well-understood technologies, they have limited data and equipment models available to support emerging technologies like carbon capture and new energy markets.

Overcoming limitations of traditional process simulators developed in the 1980s, new simulation packages enable companies to explore a range of CCS technologies within a single simulation platform. New simulation platforms provide an open and extendible model writing infrastructure and customizable thermodynamic methods that can incorporate complex heat and mass transfer equations without custom programming. This infrastructure allows users to optimize the design and operation of CCS technologies using new and/or proprietary solvents and solvent systems. An open model writing

framework, paired with rigorous thermodynamics and dozens of standard customizable equipment models, allows engineers to quickly study the scale-up of new process developments with limited commercial risk.

To demonstrate the utility of simulation platforms when designing carbon capture processes, a case study was performed comparing a variety of process metrics for pre-combustion and post-combustion carbon capture for a typical SMR process. The processes were modeled in AVEVA Process Simulation utilizing custom reaction submodels and thermodynamic properties for the first-generation monoethanolamine (MEA) solvent. While newer amine solvents are now available, reaction kinetics and thermodynamic property data is widely available for MEA which makes it a good candidate for this illustrative study.

For purposes of this study, three simulation models were built: a SMR process with no carbon capture, SMR with pre-combustion carbon capture, and SMR with post-combustion carbon capture. Operating parameters and Key Performance Indicators (KPIs) were compared across the three simulations. SMR process specifications, MEA loading, and CO<sub>2</sub> capture rates were kept consistent across the simulations for easier comparison of results.

#### 4.2. Modeling Methodology

One of the greatest challenges of building process simulation models for emerging technologies is the lack of thermodynamic data and unit operation models for these processes. Historically, users have leveraged CAPE-OPEN and similar interfaces to incorporate new unit operations into existing process simulation packages. However, these models often rely on expert users with programming knowledge to build and distribute the packages. This is a significant hurdle to process simulation users who would like to incorporate custom models but do not have programming expertise or the time to build such detailed models.

Next-generation process simulation tools eliminate this obstacle by providing an open modeling framework that all simulation users can access. Modern frameworks require no programming and give users access to the same model writing tools used by internal developers of the simulation platform. For this study, the kinetic rate expressions proposed by Xu and Froment (1989) are used to model the SMR reactor. The equilibrium expressions proposed by Hla et al. (2009) and Moe (1962) were used to model the high- and low-temperature water gas shift (WGS) reactions, respectively. **Figure 2** illustrates how these rate expressions were added to the open model writing framework of AVEVA Process Simulation.

Condition	Name	Status	Formula	Start	End	Description
	Eqn2	●	Rate["CH4"] = -RateRxn[1] - RateRxn[3]			
	Eqn3	●	Rate["H2O"] = -RateRxn[1] - RateRxn[2] - 2 * RateRxn[3]			
	Eqn4	●	Rate["H2"] = 3 * RateRxn[1] + RateRxn[2] - 4 * RateRxn[3]			
	Eqn5	●	Rate["CO"] = RateRxn[1] - RateRxn[2]			
	Eqn18	●	den = 1 + Ka["CO"] * Py["CO"] + Ka["H2"] * Py["H2"] + Ka["CH4"] * Py["CH4"] + Ka["H2O"] * Py["H2O"] / Py["H2"]			
	Eqn19	●	RateRxn[1] * den^2 * 3600 = k[1] * Py["H2"]^(-2.5) * (Py["CH4"] * Py["H2O"] - (Py["H2"]^3 * Py["CO"])) / Keq[1]			
	Eqn20	●	RateRxn[2] * den^2 * 3600 = k[2] * Py["H2"]^(-1) * (Py["CO"] * Py["H2O"] - (Py["H2"] * Py["CO2"])) / Keq[2]			
	Eqn21	●	RateRxn[3] * den^2 * 3600 = k[3] * Py["H2"]^(-3.5) * (Py["CH4"] * Py["H2O"]^2 - (Py["H2"]^4 * Py["CO2"])) / Keq[3]			
	Eqn7	●	k[1] = 4.225E15 * exp(-240100 / (R * T))			
	Eqn8	●	k[2] = 1.055E6 * exp(-67120 / (R * T))			

**Figure 2** – Model editor window for AVEVA Process Simulation, a next-generation process simulation application. Users define a set of variables, parameters, and conditions then write custom equations in an open modeling framework.

The simulations used the electrolyte NRTL (eNRTL) method available in AVEVA Process Simulation with custom thermodynamic parameters taken from Zhang et al. (2011). Equilibrium reactions were modeled based on the work of Bishnoi and Rochelle (2000). To account for the impact of reaction kinetics on mass-transfer in the film layer, additional enhancement factors were modeled in the Rate-based distillation column using the custom modeling framework.

#### 4.3. Results and Discussion

The results of the three models built in AVEVA Process Simulation are summarized in **Table 1**. As previously noted, the design specifications for the SMR and subsequent high- and low-temperature WGS reactions were kept the same across all three simulations to facilitate easier comparison between processes. Given the wide variation in raw material and energy prices across regions, this study uses the energy intensity (GJ / tCO<sub>2</sub> captured) and lean MEA makeup flow in place of operating cost. Sustainability metrics like Global Warming Potential (GWP) and Total CO<sub>2</sub>-equivalent (CO<sub>2</sub>e) emissions are also reported for each simulation.

**Table 1** – Results of a pre- and post-combustion carbon capture simulation built in the next-generation simulation platform AVEVA Process Simulation

Property	Baseline SMR with no Carbon Capture	Pre-Combustion Capture using MEA	Post-Combustion Capture using MEA
<i>Steam Methane Reforming</i>			
Steam-to-carbon ratio	2.75	2.75	2.75
Natural Gas Feed [Nm <sup>3</sup> /h]	65,000	65,000	65,000
Feed Water [kg/h]	107,730	107,730	107,520
Export Steam [kg/h]	17,150	17,150	17,180
Methane Conversion	0.870	0.870	0.867
<i>Carbon Capture</i>			
Syngas Feed [Nm <sup>3</sup> /h]	-	284,880	298,140
CO <sub>2</sub> Capture Rate [%]	-	80	80
Lean MEA Flow [m <sup>3</sup> /h]	-	982.6	2,564
Energy Intensity [GJ/tCO <sub>2</sub> ]	-	3.22	5.60
Lean MEA CO <sub>2</sub> Load	-	0.21	0.21
<i>Overall Process Sustainability Metrics</i>			
Global Warming Potential	7.79	4.38	4.73
Total CO <sub>2</sub> e Emissions [t/h]	134.6	80.86	81.60
Captured CO <sub>2</sub> [Nm <sup>3</sup> /h]	-	41,750	53,411

The results of this study illustrate the importance of process simulation when evaluating carbon capture processes. For example, the post-combustion unit in this study captures more total CO<sub>2</sub> compared to pre-combustion capture, but results in a higher overall GWP and CO<sub>2</sub>e emission footprint due to its elevated energy intensity. The makeup MEA requirement for post-combustion capture is also 2.6x that of the pre-combustion configuration. AVEVA Process Simulation accounts for emissions from both direct and indirect sources which enables a more complete comparison between these processes. While post-combustion capture allows for capture of both flue gas and process CO<sub>2</sub> emissions (produced as a byproduct of SMR and WGS reactions), the lower pressure flue gas stream makes the capture less efficient. A holistic simulation model that calculates process metrics alongside sustainability and economic indicators enables accurate comparison between process alternatives. Given the reduced energy intensity, GWP, and MEA requirements, the pre-combustion capture may be preferred. If the higher overall CO<sub>2</sub> capture rate from post-combustion capture is required, sourcing steam from nearby process units could offset the additional indirect emissions from this configuration.



When designing new carbon capture processes, companies should leverage next-generation simulation tools like AVEVA Process Simulation to evaluate a variety of design alternatives. Capital cost, operating cost, and overall capture rates should be compared for a variety of solvent systems, process configurations, and operating conditions. An open modeling framework like the one shown in **Figure 2**, paired with custom thermodynamic data, allows companies to incorporate proprietary data into simulation models of CCS processes. While this study uses a first-generation MEA solvent as an illustrative example, custom thermodynamics and the open modeling framework allow for extension to modern second- and third-generation solvents.

## 5. Conclusions

Carbon Capture and Storage (CCS) projects offer a path to reduce carbon emissions with minimal changes to existing processes. Despite this prospect, CCS projects are often considered high-risk due to the uncertainty of emerging technologies and relatively high operating cost of these units. Carbon tax programs and policy decisions will continue to increase the viability of CCS projects but must be accompanied by smart project execution and operation. When used effectively, next-generation process simulation platforms reduce the capital risk of CCS projects. Applications like AVEVA Process Simulation include open modeling frameworks that allow for integration of custom and proprietary process and equipment data with no programming required. This allows for rapid comparison of a variety of solvent systems and process configurations for CCS. Next-generation simulation platforms also include built-in carbon emission calculations alongside estimates of operating and capital cost for process-wide optimization. This gives owner-operators and EPCs better insight into the tradeoff between capital expenditure, operating cost, and emission capture rates. CCS is critical to achieving short- to medium-term emission reduction goals and next-generation simulation platforms will enable the adoption of these technologies with limited commercial risk.

## References

- R. Agarwal et al., 2016, *Imagining construction's digital future*, McKinsey & Company.
- Allied Market Research, 2021, *Carbon Capture, Utilization, and Storage (CCUS) Market by Service, Technology, and End-Use Industry: Global Opportunity Analysis and Industry Forecast, 2021-2030*.
- AVEVA, 2020, *Unified Engineering: A new proposition to break down the silos between FEED and Detailed Design to minimize risk and maximize return on Capital Investment*.
- A. Balylin-Stern and N. Berghout, 2021, *Is carbon capture too expensive?*, International Energy Agency
- S. Bishnoi and G. Rochelle, 2000, *Absorption of carbon dioxide into aqueous piperazine: reaction kinetics, mass transfer and solubility*, *Chemical Engineering Science*, 55, 22, 5531-5543.
- S. Hla et al., *Kinetics of high-temperature water-gas shift reaction over two iron-based commercial catalysts using simulated coal-derived syngases*, *Chemical Engineering Journal*, 146, 148-154.
- IEAGHG, 2017, *Techno-Economic Evaluation of SMR Based Standalone (Merchant) Plant with CCS*, IEA Greenhouse Gas and R&D Program, 2017-02.
- A. Mehmeti et al., 2018, *Life Cycle Assessment and Water Footprint of Hydrogen Production Methods: From Conventional to Emerging Technologies*, *Environments*, 5, 2, 24.
- J. Moe, 1962, *Design of water-gas shift reactors*, *Chemical Engineering Progress*, 58, 33-36
- Y. Zhang et al., 2011, *Thermodynamic modeling for CO<sub>2</sub> absorption in aqueous MEA solution with electrolyte NRTL model*, *Fluid Phase Equilibria*, 311, 67-75.
- J. Xu and G. Froment, 1989, *Methane Steam Reforming, Methanation and Water-Gas Shift: I. Intrinsic Kinetic*, *AIChE Journal*, 35, 88-90.

# Incorporation of non-linear efficiency constraints in green ammonia production design for cost minimisation

Nicholas Salmon,<sup>a</sup> René Bañares-Alcántara,<sup>a</sup>

<sup>a</sup>*Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, United Kingdom*

## Abstract

Affordable production of green ammonia will play a critical role in the decarbonization of fertilisers, the maritime industry, and the energy sector. Since renewable electricity can represent more than 50% of the levelised cost of ammonia, maximizing efficiency of electrolysis is an important stepping stone to achieving low costs. This paper analyses the increased efficiency which can be obtained by operating electrolyzers with low throughputs, to determine the extent of the impacts on the ammonia costs. It considers both a QP and LP approach, demonstrating the computational efficiency of LP calculation in this setting. We show the cost improvements from considering low-rate operation are between 1 and 5%, and that a significant change in plant design is not likely; because electrolyzers themselves are expensive, intentionally operating them at low rates costs more in wasted equipment capacity than it saves in electricity costs.

**Keywords:** Non-linear programming, Renewable Fuels, Green Hydrogen, Green Ammonia, Energy Efficiency

## 1. Introduction

The renewable energy transition will require large-scale installation of energy storage technologies in the form of chemical vectors such as green hydrogen and ammonia in order to convert variable renewables into dispatchable, portable fuels (Schmidt, 2019). The requirement that hydrogen and ammonia production be flexible to match the variable renewable power source introduces complexity into the design process, as sophisticated optimisation techniques must be used to design plants which minimise production costs.

However, variable operation also represents an opportunity for cost savings, since electrolyzers operate more efficiently, and with lower degradation, at reduced load factors (Taibi, 2020). This has led some projects to pursue low-load factor operation, which enables access to free or very cheap electricity from otherwise curtailed renewables, and which can then be converted into hydrogen with high efficiency.

The opportunity of reduced load factor operation has historically been neglected by production cost prediction models due to the non-linearity of variable electrolyser efficiencies, which conflicts with the Mixed Integer Linear Programming approach that is typically adopted for optimisation of system design (Fasihi, 2021). Using Non-Linear Programming (NLP) approaches may resolve this problem technically, but typically reduces the speed at which the model solves, and the maximum amount of input weather data that the problem can handle, itself introducing error.

This study compares this non-linear approach with an alternative method for linearising this constraint without introducing integer variables, or increasing the solution time, through the introduction of a ‘pseudo-power’ variable. This enables the power reduction associated with increased efficiency to be approximated using a series of linear constraints.

## 2. Methodology

### 2.1. Ammonia cost estimation method

Conventional ammonia production, using natural gas, typically operates all equipment at or close to its design rate continuously, in order to extract the maximum value from the capital expenditure into equipment. This operating approach is not suitable for green ammonia production, because it depends on variable renewable energy which is not always available. Selection of equipment sizes needs to account for the possibility of using both wind and solar as an energy source, and must include sufficient back-up storage of power and hydrogen to enable continuous operation of the Haber-Bosch synthesis loop, which is only partially flexible.

The process can, mostly, be represented using linear mass and energy balances, meaning the optimal size of equipment can be determined in a straightforward fashion. The inputs to the model are the local renewable weather profiles (which here include wind, fixed solar PV and single-axis tracking PV), as well as equipment efficiencies and costs. The model can then output optimal equipment size as well as the levelised cost of ammonia. For this research, the process was modelled using pyomo, and was solved using Gurobi, which can be accessed for free under an academic license.

### 2.2. Quadratic Constraint

The electrolyser is an exception to the rule of linear constraints. Electrolyser efficiency is better at low current densities, typically because there are fewer energy losses to electrical resistance or heat loss. Here, the relationship between efficiency and current density is treated as linear, although non-linear curves are discussed in subsequent sections.

The simplest approach to modelling the efficiency is to do so directly:

$$\kappa_{H_2}(t) = \kappa_{H_2}(t-1) + \pi\mu_{H_2} - \frac{3}{17}\dot{m}_{NH_3} \quad (1)$$

$$\mu_{H_2}C_{Elec} = \mu_{Best}C_{Elec} - (\mu_{Best} - \mu_{Worst})\pi \quad (2)$$

where  $\kappa_{H_2}$  is the amount of hydrogen stored in tons at time  $t$ ,  $\pi$  is the energy in MWh supplied to the electrolyser,  $\mu_{H_2}$  is the efficiency of the electrolyser in t/MWh (which in this model is a variable, not a parameter),  $\dot{m}_{NH_3}$  is the production rate of ammonia (in tons),  $C_{Elec}$  is the capacity of the electrolyser in MW, and  $\mu_{Best}$  and  $\mu_{Worst}$  are, respectively, the efficiencies of the electrolyser under the best operating condition (i.e. very low current densities) and the worst operating condition (i.e. the electrolyser rated capacity).

Both equations are modelled as strict equality constraints in the LP. Equation 1 is simply a mass balance over hydrogen, with the factor  $\frac{3}{17}$  representing the stoichiometric mass

relationship between hydrogen and ammonia. Equation 2 is an expression relating the electrolyser efficiency to its operating rate; this is simply a linear interpolation between  $\mu_{Best}$  and  $\mu_{Worst}$ , which depends on the operating ratio, given by  $\frac{\pi}{C_{elec}}$ . The operating ratio does not appear explicitly as the entire equation is multiplied by  $C_{elec}$  in order to render the non-linear constraint as a quadratic.

In theory, the above reaction can be modelled by Gurobi, which can handle both convex and non-convex QPs (Quadratic problems) using a bilinear solver. However, in practice, the non-convexity of this problem makes it very challenging to solve, and significantly slows down the model solution. For that reason, using one year of data at hourly resolution, the model has over 50,000 constraints and 30,000 variables, meaning the solver was unable to identify an optimal solution within 48 hours when solved on an i7 desktop with 16 GB of RAM. Even applying a significantly reduced dataset, containing just one hundred time steps (~500 constraints, 300 variables), the model solution takes over 16 hours.

Realistically, the plant design needs to be rapid, so it can solve under a large number of conditions, such as varying locations, and equipment costs. For that reason, a linear solution needs to be found which can converge more rapidly, even though the two constraints described above do not readily linearise.

### *2.2.1. Linearisation approach*

In order to resolve this problem, we introduce a new variable,  $\zeta$ , which we term the ‘pseudopower’. The plant is then designed as normal under a linear arrangement (i.e. using equation (1) as a constraint, but not equation (2)), with  $\mu_{H2}$  treated as a parameter whose value is equal to  $\mu_{Worst}$ . The true power consumed by the electrolyser is therefore slightly less than the power modelled by the LP. The difference between the true power and the modelled power is the pseudopower, and it is returned to the input side of the overall plant energy balance, along with the power which originates from the variable renewable supply. The optimisation solver can then reallocate the unused energy as it sees fit.

Formally, the pseudo-power is calculated by:

$$\zeta = \pi \left( 1 - \frac{\eta_{H2}}{\eta_{Worst}} \right) \quad (3)$$

Note that for the definition of the pseudopower, rather than direct efficiencies, ( $\mu_{H2}$  and  $\mu_{Worst}$ , measured in t/MWh), we use their reciprocals,  $\eta_{H2}$  and  $\eta_{Worst}$ , measured in MWh/t. It is strictly necessary to use  $\mu_{H2}$  for the non-linear modelling, and the problem can only be solved for electrolysers where the efficiency is linearly related to the operating ratio, such that the problem will be quadratic. This linearisation approach is more flexible, and we therefore use  $\eta_{H2}$  in place of  $\mu_{H2}$  only so that the variable  $\eta_{H2}$  appears in the numerator rather than the denominator in the definition of  $\zeta$ . For the approximation of  $\eta_{H2}$  (which we subsequently refer to as the reciprocal efficiency), a slightly wider range of input functions can be used – for demonstration, we use a linear one:

$$\eta_{H2} = \eta_{Best} + \Delta\eta \left( \frac{\pi}{C_{Elec}} \right) \quad (4)$$

where  $\Delta\eta$  is the difference in the reciprocal efficiencies. Note that equation (4) is linear in the reciprocal efficiencies, rather than the efficiencies themselves, as per (2); the results will be similar, and the alternative is adopted merely for simplicity of demonstration. This results in the following expression for the pseudopower:

$$\zeta = \frac{\Delta\eta}{\eta_{Worst}} \left( \frac{-\pi^2}{C_{Elec}} + \pi \right) \quad (5)$$

This does not apparently resolve the problem, since the expression is non-linear. However, the pseudo-power, unlike the electrolyser efficiency, can be treated using inequalities, rather than strict equalities. Therefore, the acceptable region of values of the pseudopower can be defined using a series of linear constraints which are tangent to the curve, as shown in Figure 1.



Figure 1 - Suitable region for the definition of the pseudopower. Acceptable values lie within the grey curve, and are constrained there by the linear constraints represented using dashed lines.

In general, the model will maximise the power input into the energy balance, so it will typically select the maximum value of the pseudopower as allowed within the constraint region, i.e. it will select a value on the curve, rather than underneath it. The case in which it selects a pseudopower which is less than any of the linear constraints is equivalent to curtailing electricity. The curtailed electricity is simply the slack variable in the energy balance constraint; whether the slack is allocated to the energy balance or the pseudopower constraint is arbitrary to the actual outcome. Here, the shape of the pseudopower curve is quadratic; other shapes could be used, although integer variables may need to be introduced to ‘turn off’ the constraints in certain parts of the operating ratio range if the pseudopower curve has a large number of critical points.

Even now, a non-linearity remains in the calculation of the operating ratio. We resolve this by pre-estimating the size of the electrolyser,  $C_{elec}$  by allowing the model to run without any pseudopower. For only the constraints which consider the pseudopower,  $C_{elec}$

is treated as a parameter, rather than a variable; the amount of error introduced by this adjustment is discussed in the results.

### 3. Results

The linearized model converges in around one minute using an i7 desktop and a full year of weather data. This enabled the code to be run in over 400 locations with a wide range of weather profiles, considering a case for both 2022 and 2050 (which accounts for falling costs of some equipment, as well as improving efficiencies of the electrolyser).

On average for 2020, the pseudopower represented 3% of the energy ‘supplied’ to the plant, with a maximum value of 4.8%. This improvement in efficiency enabled a reduction in the electrolyser size by on average 4%, and at most 11% (compared to a case without pseudopower). The results for 2050 were broadly similar to those for 2022, although the improvement is slightly smaller. This is because the general improvement in electrolyser efficiency over time, as well as the rapidly falling price of solar power, means that (i) the efficiency benefit from low power operation is smaller as a fraction of total contribution, and (ii) efficiency is a less important criteria than low capital costs for determining the optimum arrangement of the ammonia plant. For that reason, pseudopower only represents 2% of the energy supplied, and the electrolyzers only decreased in size by 2.5%.

The reduced electrolyser size impacts the accuracy of the pseudopower calculation, since it depends upon the original estimate of the electrolyser size. This will cause the pseudopower to be overestimated when the operating ratio is less than  $\frac{1}{2}$ , and will be underestimated when the operating ratio is greater than  $\frac{1}{2}$ . To some extent, these two impacts should cancel out (since the average electrolyser load fraction is 47%, they spend roughly the same amount of time above and below this point). In any case, the error in the pseudopower will not be larger than 10%, which is fairly small in the context of large uncertainties in the equipment costs (and given that the pseudopower is a small fraction of the total input power). If the error needs to be reduced further, the high speed of solution would enable the model to be converged several times until the initial guess of the electrolyser size was within a tolerance limit of the output electrolyser size.

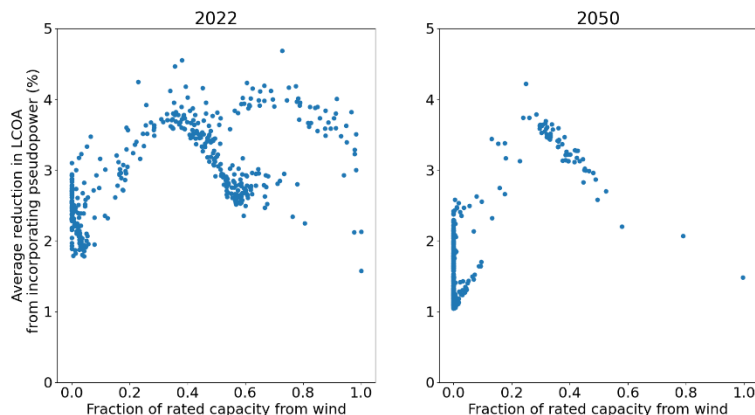


Figure 2 - Improvement in LCOA plotted against fraction of rated capacity from wind

Reducing both the electricity consumption and the electrolyser size also reduces the levelised cost of ammonia (LCOA), which is the objective of the optimisation. Unsurprisingly, the reduction in LCOA is well correlated to the amount of pseudo-power consumed.

Figure 2 shows the relationship between the size of this LCOA reduction and the fraction of the renewable energy farm which originates with wind. Very clearly for the 2050 locations, and slightly less clearly for the 2022 locations, the largest pseudo-power benefit originates at sites which use hybrid operation (i.e. rely on both wind and solar). At sites which predominantly rely on only one power source, the rated capacity of the electrolyser is similar in magnitude to the installed capacity of the renewables; this prevents excessive curtailment when the electrolyser is running without unduly oversizing it. To achieve the same outcome for the hybrid sites, the electrolyser may be larger in capacity than either of the renewables; only when they operate in tandem will the electrolyser be fully utilized. This means that the electrolyser more frequently operates at a lower rated capacity (i.e. when only one of the two installed renewables is available), and therefore benefits to a greater extent from the improved efficiency under those conditions.

Considering the efficiency as a function of electrolyser load has not radically changed system design, which is unsurprising given that it represents, proportionally, a small change in electrolyser performance. This suggests that low-load factor operation at high efficiencies is not likely to be a viable operating mode, predominantly because the electrolyser capital cost is quite high; therefore, it is better to use it at a higher rate where the availability of renewables allows it, even if doing so sacrifices some efficiency. From the perspective of design, therefore, suitable equipment sizes may be selected even without accounting for this non-linearity. However, from the perspective of comparing sites to each other, it may be necessary to consider the role of improved efficiency, since they have the greatest impact on sites with weather profiles that lend themselves to both wind and solar (as opposed to sites that rely on one particularly strong resource).

#### 4. Conclusions

This report has presented a novel method for dealing with a non-linearity present within optimisation of renewable ammonia production. It compared the possibility of using a QP with a linearized LP approach, demonstrating that the errors introduced from an LP are comparatively small, and enable a far more rapid solution than a non-convex QP. The improved efficiency of electrolyzers at low load factors may reduce costs between 1 and 5%; this is not sufficiently large to justify changes in the plant design, but may prove beneficial for operating plants, particularly if they use both wind and solar as power sources. Further investigation should consider the extent to which these results are affected by a connection to the electricity grid.

#### References

- Schmidt, J., et al., 2019, A new perspective on global renewable energy systems: why trade in energy carriers matters. *Energy & Environmental Science*. **12**(7): p. 2022-2029.
- Fasihi, M., et al., 2021, Global potential of green ammonia based on hybrid PV-wind power plants. *Applied Energy*. **294**: p. 116170.
- Taibi, E. et al., 2020, Green Hydrogen Cost Reduction: Scaling up electrolyzers to meet the 1.5°C climate goal, IRENA. p. 45

# Strategic participation of a gas-fired power plant in interdependent electricity and natural gas markets under carbon emission trading schemes

Christos N. Dimitriadis,<sup>a</sup> Evangelos G. Tsimopoulos,<sup>a</sup> Michael C. Georgiadis<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, Aristotle University of Thessaloniki, University Campus, Thessaloniki, 54124, Greece*

*Corresponding author, e-mail: [mgeorg@auth.gr](mailto:mgeorg@auth.gr)*

## Abstract

This work considers a bi-level optimization framework to model the strategic participation and derive optimal offers for a gas-fired power plant (GFPP) in interdependent pool-based electricity and natural gas markets, under high penetration of wind power generation. The upper-level problem ensures profit maximization for the strategic agent, while the lower-level problem represents the sequential electricity and natural gas market clearing process. Considering the worldwide establishment of low-carbon policies, a carbon emission trading scheme (CETS) is also incorporated into the proposed optimization framework. The bi-level algorithm is initially transformed into a mathematical programming with equilibrium constraints model and further recast into a mixed integer linear program, using Karush-Kuhn-Tucker optimality conditions and strong duality theory. Numerical simulations illustrate that the proposed methodology is capable of effectively mimicking the actual electricity and gas market-clearing process and derive the optimal amount of carbon emission allowances and strategic offers for the gas-fired power plant.

**Keywords:** bi-level optimization, strategic offering, electricity market, natural gas market, carbon emission trading

## 1. Introduction

In recent years, the development of mathematical approaches that ensure the strategic participation of a power producer in energy markets has been of great interest. Tsimopoulos and Georgiadis (2020) studied the optimal strategies for the participation of a market agent owning a mixed portfolio of conventional and wind power generation. In addition, Dimitriadis et al. (2022) proposed a bi-level algorithm to investigate the strategic bidding decisions of an energy storage agent in a joint energy and reserve market. Natural gas is considered an efficient and clean fuel that in the long-term can sufficiently replace more polluting fuels for power generation, such as coal and oil. However, while the GFPPs hold a significant share in the market, their optimal operation as power producers and natural gas consumers has not been thoroughly studied.

In this work, a sequential clearing mechanism for interdependent electricity and gas markets is adopted (Ordoudis et al., 2019), to derive the optimal capacity and offering strategies for a GFPP. The CETS is also incorporated into the algorithm to account for



the carbon emissions of each conventional producers and to investigate how carbon emission trading affects the operation of the GFPP.

## 2. Mathematical model

In this section, the bi-level formulation is presented to determine the optimal offering strategies for a strategic GFPP in interdependent electricity and gas markets, considering a cap-and-trade program for carbon emissions. As shown in (1), the hourly overall carbon emission allowances number  $Q_t$  varies with the net load i.e., demand load minus wind power production, while  $\eta$  represents the emission allowance factor (tCO<sub>2</sub>/MW). Constraint (2) defines the number of emission allowances  $Q_{h,t}^H$  that correspond to each conventional power plant  $h$ , depending on its allocated factor of allowances  $\alpha_h$ , which is calculated by (3).  $\zeta_h$  comprises the emission factor of each conventional power plant.

$$Q_t = \eta \cdot \left( \sum_{d \in L \cup DaN} L_{d,t}^E - \sum_{j \in JaN} w_{j,t} \right) \quad (1)$$

$$Q_{h,t}^H = \alpha_h \cdot Q_t \quad (2)$$

$$\alpha_h = \frac{\zeta_h}{\sum_h \zeta_h} \quad (3)$$

### 2.1. Upper-level problem: Profit maximization for the GFPP

$$\begin{aligned} \mathbf{maximize} \quad & \sum_t \left\{ \sum_{g \in GaN} \lambda_{n,t}^E \cdot v_{g,t} - \sum_{g \in GaR} \lambda_{r,t}^{NG} \cdot \varphi_g \cdot v_{g,t} - cp \cdot (\zeta_g \cdot v_{g,t} \right. \\ & \left. - Q_{g,t}^H) \right\} \end{aligned} \quad (4)$$

Objective function (4) aims at maximizing the profit of the GFPP  $g$ . The first term represents its income from selling its hourly power production  $v_{g,t}$ , while the second term corresponds to the cost of purchasing gas, with  $\varphi_g$  being the electricity-gas conversion factor for the GFPP.  $\lambda_{n,t}^E$ ,  $\lambda_{r,t}^{NG}$  and  $cp$  correspond to the electricity, natural gas and carbon prices, respectively. The third term represents GFPP's carbon trading cost, depending on its allocated ( $Q_{g,t}^H$ ) and actual ( $\zeta_g \cdot v_{g,t}$ ) carbon emission allowances.

### 2.2. Lower-level problem: Electricity market clearing

$$\begin{aligned} \mathbf{minimize} \quad & \sum_t \left\{ \sum_{g \in GaN} o_{g,t} \cdot v_{g,t} + \sum_{i \in IaN} C_i \cdot p_{i,t} \right. \\ & \left. + \sum_{g \in GaN} cp \cdot (\zeta_g \cdot v_{g,t} - Q_{g,t}^H) + \sum_{i \in IaN} cp \cdot (\zeta_i \cdot p_{i,t} - Q_{i,t}^H) \right\} \end{aligned} \quad (5)$$

s.t.

$$-\sum_{i \in IaN} p_{i,t} - \sum_{g \in GaN} v_{g,t} - \sum_{j \in JaN} w_{j,t} + \sum_{d^{EL} \in DaN} L_{d,t}^E + \sum_{m \in NaM} B_{n,m} \cdot (\delta_{n,t} - \delta_{m,t}) = 0 \quad : [\lambda_{n,t}^E] \quad \forall n, \forall t \quad (6)$$

$$0 \leq p_{i,t} \leq \overline{P}_i \quad : [\underline{a}_{i,t}, \overline{a}_{i,t}] \quad \forall i, \forall t \quad (7)$$

$$0 \leq v_{g,t} \leq \overline{V}_g \quad : [\underline{\beta}_{g,t}, \overline{\beta}_{g,t}] \quad \forall g, \forall t \quad (8)$$

$$0 \leq w_{j,t} \leq \overline{W}_j \quad : [\underline{\gamma}_{j,t}, \overline{\gamma}_{j,t}] \quad \forall j, \forall t \quad (9)$$

$$-Q_{h,t}^H + \alpha_h \cdot \eta \cdot \left( \sum_{d^{EL} \in DaN} L_{d,t}^E - \sum_{j \in JaN} w_{j,t} \right) = 0 \quad : [\rho_{h,t}] \quad \forall h, \forall t \quad (10)$$

$$-\overline{T}_{n,m} \leq B_{n,m} \cdot (\delta_{n,t} - \delta_{m,t}) \leq \overline{T}_{n,m} \quad : [\underline{\psi}_{n,m,t}, \overline{\psi}_{n,m,t}] \quad \forall (n, m) \in NaM, \forall t \quad (11)$$

$$-3.14 \leq \delta_{n,t} \leq 3.14 \quad \forall n, \forall t \quad : [\underline{\pi}_{n,t}, \overline{\pi}_{n,t}] \quad (12)$$

$$\delta_{n_1,t} = 0 \quad : [\eta_{n,t}^o] \quad \forall n = n_1, \forall t \quad (13)$$

Objective function (5) minimizes total operating cost and represents the electricity market clearing procedure, while  $o_{g,t}$  and  $C_i$  correspond to the strategic offers by the GFPP and the cost offers of non-GFPPs  $i$ , respectively. Constraint (6) applies the power balance at each electric bus  $n$ , where  $L_{d,t}^E$  is the power demand load and  $B_{n,m}$  the susceptance of power transmission line connecting buses  $n, m$ . Constraints (7) – (9) enforce the upper and lower capacity for non-gas-fired, gas-fired and wind power plants  $j$ , respectively. Constraint (10) derives the allocated emission allowances for each conventional power plant. Constraint (11) imposes transmission lines' capacity limits, while constraints (12), (13) limit the voltage angle  $\delta_{n,t}$  range of each electric bus and establish bus A as the electric grid's slack bus.

### 2.3. Lower-level problem: Natural gas market clearing

$$\text{minimize} \quad \sum_t \left\{ \sum_{k \in KaR} C_k \cdot f_{k,t} \right\} \quad (14)$$

s.t.

$$-\sum_{k \in KaR} f_{k,t} + \sum_{g \in GaR} \varphi_g \cdot v_{g,t} + \sum_{d^{NG} \in DaR} L_{d,t}^{NG} = 0 \quad : [\lambda_{r,t}^{NG}] \quad \forall r, \forall t \quad (15)$$

$$\underline{F}_k \leq f_{k,t} \leq \overline{F}_k \quad : [\underline{\varepsilon}_{k,t}, \overline{\varepsilon}_{k,t}] \quad \forall k, \forall t \quad (16)$$

Objective function (14) represents the natural gas clearing mechanism by minimizing total operating cost, with  $C_k$  corresponding to the cost offers of each natural gas supplier  $k$ . Constraint (15) constitutes the gas balance at each gas node  $r$  and  $L_{d,t}^{NG}$  represents natural gas demand of loads  $d^{NG}$ . Constraint (16) imposes upper and lower capacity for the natural gas production  $f_{k,t}$  of each supplier.

#### 2.4. Solution strategy

Considering the convexity and continuity of the lower-level problems, the bi-level formulation is reduced to a mathematical program with equilibrium constraints (MPEC), by employing Karush-Kuhn-Tucker optimality conditions and is further recast into a mixed-integer linear program (MILP), applying disjunctive constraints and the strong duality theorem.

### 3. Application study

The proposed algorithm is applied in a modified Pennsylvania – New Jersey – Maryland (PJM) 5-bus electric network, as illustrated by Fig.1, and a single-node natural gas network. Six non-gas-fired power plants, the strategic GFPP and a wind farm are located at the power network, while the total electricity demand is equally distributed on three load buses. Similarly, at the natural gas network, three gas suppliers and three gas loads are established. It is important to emphasize that the GFPP is considered as the fourth gas load, the capacity of which is directly linked to the power that the GFPP provides to the electricity market. Carbon emission price is set at 23 \$/t, while the maximum capacity and the susceptance of each power transmission line are equal to 1000 MW and 9.412, respectively.

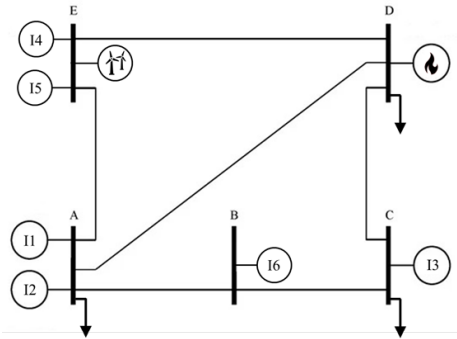


Figure 1. PJM 5-bus power network.

### 4. Results and Discussion

#### 4.1. Uncongested power network

The proposed MILP model is applied to the integrated electricity and natural gas system and solved using GAMS/CPLEX. The strategic gas-fired unit exerts its dominant position in the market and manipulated day-ahead prices to increase its profitability. Fig.2 depicts the market clearing prices (MCPs) during the 24-hour horizon, both for the case where carbon emissions of conventional producer are considered and for the opposite case. It is important to point out that the wind power plants and considered both cost- and emission-free. The value of CETS-embedded MCPs is higher compared to the normal MCPs for all time periods since each conventional producer due to its carbon emissions, increases its bidding cost by  $cp \cdot \zeta_h$ .

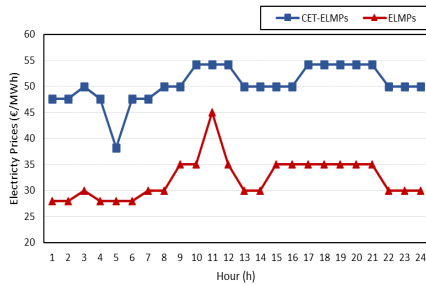


Figure 2. CETS-embedded and normal electricity MCPs.

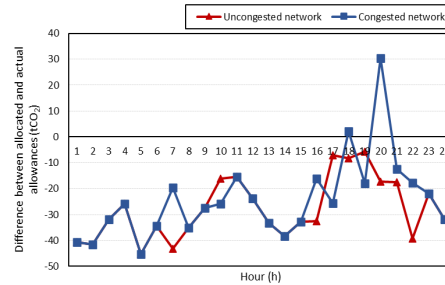


Figure 3. Net carbon emission allowances.

In the calculation of the expected profits for each conventional power plant, trading of carbon emission allowances plays a key role. The difference between the cost-free allowances provided by the government to the strategic gas-fired unit and its actual emissions based on its power production, is illustrated in Fig. 3. Positive values represent the number of allowances (in CO<sub>2</sub> tons) granted to the GFPP but not employed and thus constitute an income source. Conversely, negative values indicate the additional carbon allowances that the GFPP is worth paying for, in order to satisfy its desired electricity production. For the uncongested power network case, the GFPP strategically chooses to pay this extra allowance purchase cost, since the CETS-embedded MCPs are noticeably higher compared to the carbon trading price.

#### 4.2. Congested power network

In this case, the capacity of lines A-B and D-E is reduced to 400 and 300 MW equivalently. As a result, the power network becomes congested and different CETS-embedded electricity MCPs arise, as shown in Fig. 4. The GFPP acting strategically, manages to capitalize on congestion and further manipulate MCPs. Especially in bus D, where it is installed, a price cap needs to be established at 158.05 \$ by the MO, to prevent its excessive speculation. Bus E remains the only bus not affected by the bidding strategy of the GFPP. The reason seems to be the fact that the zero-cost wind farm is located at this bus. In fact, the wind farm generates electricity at full capacity during the 24-hour horizon and covers the largest percentage of the demand, thus holding electricity MCPs down. Furthermore, as depicted by Table 1, despite the fact that the GFPP generates less electricity compared to the uncongested network case, due to the significantly higher prices, manages to increase its daily profits by 21.5 %.

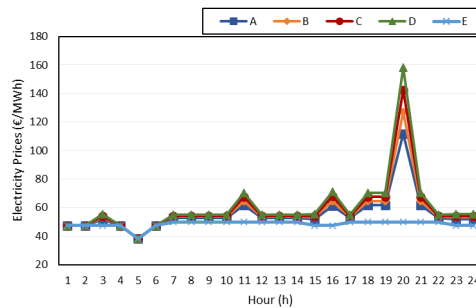


Figure 4. CETS-embedded electricity MCPs for the congested network.

Table 1: GFPP's electricity production and expected profits

	Electricity generation (MWh)	Profit (\$)
Uncongested power network	4,394	122,005.53
Congested power network	4,153	148,343.96

### 4.3. Natural gas price increase

This case study investigates the influence of natural gas price increase on the market outcomes and the optimal strategies of the GFPP. In particular, gas suppliers rise their bidding costs by 50%, 100%, 150% and 200% and consequently the gas MCPs. In order to balance this increase in its marginal cost, the strategic GFPP bids higher. As a result, for many time periods the electricity MCPs are also affected by following an upward trend. However, the GFPP cannot offset the rise of the gas MCPs and significantly withholds its power production experiencing a critical profit decrease, as shown in **Fig.5**.

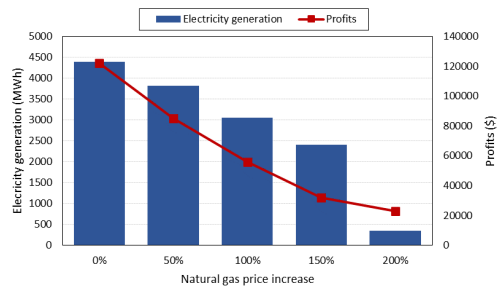


Figure 5. Power generation and profits for strategic GFPP.

## 5. Conclusions

In this work a bi-level mathematical algorithm is developed to derive optimal offering strategies and maximize the profit of a GFPP participating in interdependent low-carbon electricity and natural gas markets. The model derives the CETS-embedded electricity and gas MCPs and the optimal offering decisions and carbon emission management for the GFPP. Results demonstrate that power network congestion creates more arbitrage opportunities for the GFPP, compared to the uncongested network case, while an increase in gas prices leads the GFPP to withhold its production and thus experience profit losses. Future work should involve the investigation of the strategic behavior for an agent with mixed generation portfolio in an integrated electricity and gas market scheme.

## Acknowledgments

This research has been co-financed by (i) the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the specific call “Aquacultures” – “Industrial Materials” – “Open Innovation in Civilization” (project code: T6YBP-00251). Project title: Development of Computer-Aided Tools for Optimal Energy Consumption in Industrial Ceramics (CATOPEC-IC) and (ii) HELLENiQ ENERGY, under the framework of research program “HELLENiQ ENERGY Scholarships” (project number: 98665).

## References

- Dimitriadis, C. N., Tsimopoulos, E. G., & Georgiadis, M. C. (2022). Strategic bidding of an energy storage agent in a joint energy and reserve market under stochastic generation. *Energy*, 242, 123026. <https://doi.org/https://doi.org/10.1016/j.energy.2021.123026>
- Ordoudis, C., Pinson, P., & Morales, J. M. (2019). An Integrated Market for Electricity and Natural Gas Systems with Stochastic Power Producers. *European Journal of Operational Research*, 272(2), 642–654. <https://doi.org/10.1016/j.ejor.2018.06.036>
- Tsimopoulos, E. G., & Georgiadis, M. C. (2020). Withholding strategies for a conventional and wind generation portfolio in a joint energy and reserve pool market: A gaming-based approach. *Computers and Chemical Engineering*. <https://doi.org/10.1016/j.compchemeng.2019.106692>

# Modeling and optimization of hybrid P/VSA-membrane separation processes for CO<sub>2</sub> capture from post combustion flue gas

Christos C. Chatziasteriou,<sup>a,b</sup> Michael C. Georgiadis,<sup>a,b</sup> Eustathios S. Kikkinides,<sup>a,b</sup>

<sup>a</sup>*Department of Chemical Engineering, Aristotle University of Thessaloniki (AUTH), Thessaloniki 54124, Greece*

<sup>b</sup>*Chemical Process and Energy Resources Insitute (CPERI) Center for Research and Technology Hellas (CERTH), 6<sup>th</sup> km Charilaou-Thermi, Thessaloniki 57001, Greece*

## Abstract

This work presents a hybrid Pressure/Vacuum Swing Adsorption (P/VSA)-membrane process for CO<sub>2</sub> capture from flue gas. The proposed process combines the advantages of each individual process. Both P/VSA-membrane and membrane-P/VSA configurations have been considered and evaluated based on a modeling and optimization framework. In addition, a two-stage P/VSA case has been considered and optimized as a reference case for comparison purposes. In all cases, the goal of optimization is to minimize total energy requirements keeping the constraints of 95% total CO<sub>2</sub> purity and 90% total CO<sub>2</sub> recovery, separating a 30/70 mol % CO<sub>2</sub>/N<sub>2</sub> feed mixture. Optimization results demonstrate that the optimal membrane-P/VSA configuration leads to an energy requirement of 134.4 kWh/tCO<sub>2</sub>, representing ~8% reduction comparing to the respective value of 145.9 kWh/tCO<sub>2</sub> for the two-stage P/VSA case.

**Keywords:** Post-combustion CO<sub>2</sub> capture. P/VSA, Membrane, Hybrid process

## 1. Introduction

The capture of the post-combustion CO<sub>2</sub> emissions that come from various industrial applications (Abanades et al., 2015) is a decisive step to the direction of reducing global warming due to the increased concentration of CO<sub>2</sub> in the atmosphere. Chemical absorption is the most mature technology for this purpose. However, it constitutes a huge thermal heat duty process. As a result, research on more efficient processes is in progress. Pressure/Vacuum Swing Adsorption (P/VSA) and membrane separation are promising alternatives to absorption (Chatziasteriou et al., 2022). Unfortunately, single-stage membrane processes are unable to achieve the target of CO<sub>2</sub> Purity  $\geq 95\%$  and CO<sub>2</sub> Recovery  $\geq 90\%$  set by US Department of Energy (DOE/NETL) (Khurana and Farooq, 2017) utilizing the currently existing materials. On the other hand, industrially unapplicable low vacuum pressures are required for efficient one-stage P/VSA process configurations. Accordingly, two-stage P/VSA processes and three-stage or even four-stage membrane configurations have been studied to overcome the abovementioned limitations, resulting in significant increase in the capital expenditure (CAPEX) and the operational expenditure (OPEX) of the process (Chatziasteriou et al., 2022; Zanco et al., 2017).

An integrated or hybrid membrane-P/VSA separation process (Figure 1) may be capable of combining the advantages of each individual process and, at the same time, be a more efficient alternative to the abovementioned for the effective post-combustion CO<sub>2</sub> separation. P/VSA and membrane process are connected in such way that they constitute the two stages of the hybrid process. The flue gas is partially separated in the first stage and the CO<sub>2</sub> enriched product enters the second stage where the separation is completed. In this work, both P/VSA-membrane and membrane-P/VSA configurations are examined.

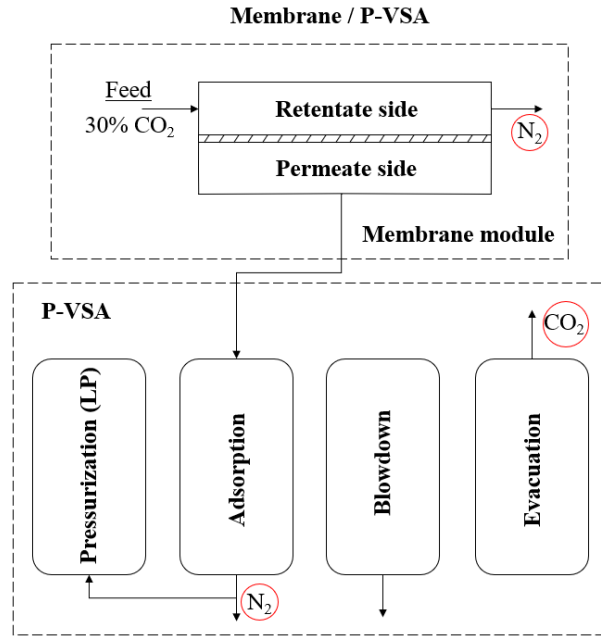


Figure 1. Hybrid membrane-P/VSA representation

## 2. Process modeling

### 2.1. Pressure/Vacuum Swing Adsorption

As partial separation of CO<sub>2</sub> is required in each one of the two stages, a simple 4-step P/VSA process has been considered in the present study, due to its high efficiency and the reduced computational cost compared to more complex cycles. The present 4-step cycle consists of an adsorption step, a co-current blowdown step, a counter-current evacuation step and a counter-current pressurization with light product (LP) step. A set of partial differential equations (Haghpanah et al., 2013) describes the whole process, while each step differs from the others due to its unique boundary conditions. The conditions at the end of each step are considered as the initial conditions of the next step. Zeolite 13X is considered as the adsorbent, as this is the most mature and efficient material used for CO<sub>2</sub>/N<sub>2</sub> mixtures separation at an industrial scale. Note that adsorbate/adsorbent equilibrium is described by the dual-site Langmuir isotherm.

#### 2.1.1. P/VSA Model Equations

Mass balance of component *i*:

$$\frac{\partial y_i}{\partial t} + \frac{y_i}{P} \frac{\partial P}{\partial t} - \frac{y_i}{T} \frac{\partial T}{\partial t} = \frac{T}{P} D_z \frac{\partial}{\partial z} \left( \frac{P}{T} \frac{\partial y_i}{\partial z} \right) - \frac{T}{P} \frac{\partial}{\partial z} \left( \frac{y_i P}{T} u \right) - \frac{R_g T (1 - \varepsilon_b)}{P \varepsilon_b} \frac{\partial q_i}{\partial t} \quad (1)$$

Mass transfer rate of component *i* (LDF model):

$$\rho_p \frac{\partial q_i}{\partial t} = \frac{c_{pi}}{q_i^*} k_{ci} (q_i^* - q_i) \quad (2)$$

Bed energy balance:

$$\begin{aligned} & \left[ (1 - \varepsilon_b) \rho_p c_{ps} + (1 - \varepsilon_b) \rho_g c_{pa} \sum_{i=1}^2 \frac{\rho_p q_i}{c_f} \right] \frac{\partial T}{\partial t} \\ &= \frac{\lambda_e}{\varepsilon_b} \frac{\partial^2 T}{\partial z^2} - \frac{c_{pg}}{R_g} \frac{\partial}{\partial z} (uP) - \frac{c_{pg}}{R_g} \frac{\partial P}{\partial t} - \frac{(1 - \varepsilon_b)}{\varepsilon_b} c_{pa} T \sum_{i=1}^2 \frac{\partial q_i}{\partial t} \\ &+ \frac{(1 - \varepsilon_b)}{\varepsilon_b} \sum_{i=1}^2 (-\Delta H_i) \frac{\partial q_i}{\partial t} - \frac{2h_{w,in}}{\varepsilon_b r_{in}} (T - T_w) \end{aligned} \quad (3)$$

Momentum balance (Darcy's law):

$$-\frac{\partial P}{\partial z} = \frac{150\mu(1 - \varepsilon_b)^2}{4R_p^2 \varepsilon_b^2} u \quad (4)$$

A wall energy balance is also used in order to describe the heat transfer between the column and the environment. In the above set of equations,  $y_i$  is the molar fraction of the component  $i$ ,  $P$  and  $T$  are the operating pressure and temperature, respectively,  $t$  and  $z$  are the time and the column's length,  $D_z$  is the axial dispersion coefficient,  $R_g$  is the universal gas constant,  $\varepsilon_b$  is the bed porosity,  $q_i$  is the concentration in the solid phase of the component  $i$ ,  $\rho_p$  is the adsorbent density,  $c_{pi}$  is the concentration of the component  $i$  in the adsorbent pores,  $q_i^*$  is the solid phase equilibrium concentration,  $k_{ci}$  is a mass transfer coefficient,  $c_{ps}$ ,  $c_{pa}$ ,  $c_{pg}$  are the specific heat capacities of the adsorbent, the adsorbed phase and the gas phase, respectively,  $\rho_g$  is the gas phase density,  $c_f$  is the feed concentration,  $\lambda_e$  is the gas axial thermal conductivity,  $\Delta H_i$  is the component  $i$  heat of adsorption,  $h_{w,in}$  is the heat transfer coefficient form the column to wall,  $r_{in}$  is the inner bed radius,  $T_w$  is the bed wall temperature,  $\mu$  is the gas viscosity and  $R_p$  is the solid particles radius.

### 2.1.2. P/VSA boundary conditions

The mass and energy balances boundary conditions constitute the most significant boundary conditions (BCs) that correspond to the inlet process streams and they are expressed as

$$u(y_i|_z - y_{i,f}) = D_z \frac{\partial y_i}{\partial z} \Big|_z \quad (5) \quad \text{and} \quad \rho_g c_{pg} u \varepsilon_b (T|_z - T_f) = \lambda_e \frac{\partial T}{\partial z} \Big|_z \quad (6)$$

At the outlet streams, the operating pressure is set to the desired value,  $P = P_t$ . Finally, at the column's closed ends, the gas velocity, is set  $u = 0$ .

## 2.2. Membrane process

In general, several types of flow models of various complexity have been proposed to describe gas separation by permeation through membranes (Geankoplis, 2003). In the present study we employ the cross-flow model which gives a high efficiency separation performance, that is very close to the more rigorous, but much more expensive computationally, counter-current flow model (Geankoplis, 2003).

### 2.2.1. Membrane model equations

Mass balance:

$$y_p = \frac{x_f - x_{ret}(1 - \theta)}{\theta} \quad (7)$$

Connection between retentate and permeate equation:



$$\frac{y}{1-y} = \frac{a^*[x - (P_L/P_h)y]}{(1-x) - (P_L/P_h)(1-y)} \quad (8)$$

where  $y_p$  is the permeate exit molar fraction,  $x_f$  is the feed molar fraction,  $x_{ret}$  is the retentate exit molar fraction,  $\theta$  is the membrane stage cut,  $y$  and  $x$  are the permeate and retentate molar fractions, respectively,  $a^*$  is the membrane permselectivity and  $P_L$  and  $P_h$  are the permeate and the retentate operating pressures, respectively.

### 2.3. Hybrid P/VSA – Membrane process

The hybrid process consists of a P/VSA and a membrane stage. In the present work, two potential configurations are examined: P/VSA-membrane and membrane-P/VSA configuration. The separation of 30 mol % CO<sub>2</sub> post combustion flue gas mixture from the cement industry is considered. The low pressure, that is, the pressure at the evacuation step in the P/VSA process and the permeate pressure in the membrane process, is fixed at 0.1 bar as this is the lowest applicable pressure at an industrial scale. The efficiency indicators in a two-stage configuration are described below (Haghpanah et al., 2014). The indicator indices represent the stage number, while the output of the first stage is the input of the second one.

$$Purity_{total} = Purity_2 \quad (9)$$

$$Recovery_{total} = Recovery_1 \cdot Recovery_2 \quad (10)$$

$$Energy_{total} = \frac{Energy_1}{Re_2} + Energy_2 \quad (11)$$

#### 2.3.1. Optimization problem

Optimization aims at investigating the optimal variable values which lead to the most efficient process performance. The separation processes energy requirement is one of the most significant indicators and as such, energy minimization is the target of the present optimization studies. The constraints and the decision variable bounds are summarized in Table 1 and Table 2.

Table 1. Optimization constraints and decision variables of the P/VSA process

	Stage 1	Stage 2
<b>Objective function</b>	Min Energy consumption	Min Energy consumption
<b>Constraints</b>	Purity <sub>CO2</sub> = 48.4% Recovery <sub>CO2</sub> = 100%	Purity <sub>CO2</sub> ≥ 95% Recovery <sub>CO2</sub> ≥ 95%
<b>Decision variables</b>	1.0 bar ≤ P <sub>h</sub> ≤ 1.5 bar 0.15 bar ≤ P <sub>1</sub> ≤ 0.95 bar 0.5 m/s ≤ u <sub>f</sub> ≤ 0.9 m/s	1.0 bar ≤ P <sub>h</sub> ≤ 1.5 bar 0.2 bar ≤ P <sub>1</sub> ≤ 0.95 bar 0.3 m/s ≤ u <sub>f</sub> ≤ 1.2 m/s

Table 2. Optimization constraints and decision variables of the membrane separation process

	Stage 1	Stage 2
<b>Objective function</b>	Min Energy consumption	Min Energy consumption
<b>Constraints</b>	Purity <sub>CO2</sub> = 48.4% Recovery <sub>CO2</sub> = 95.5%	Purity <sub>CO2</sub> ≥ 95% Recovery <sub>CO2</sub> ≥ 90%
<b>Decision variables</b>	1.0 bar ≤ P <sub>h</sub> ≤ 1.5 bar 0.1 bar ≤ P <sub>1</sub> ≤ 1.0 bar 0.0 ≤ x <sub>ret</sub> ≤ 0.3	1.0 bar ≤ P <sub>h</sub> ≤ 1.5 bar 0.1 bar ≤ P <sub>1</sub> ≤ 1.0 bar 0.0 ≤ x <sub>ret</sub> ≤ 0.3

#### 2.3.2. Numerical solution

The mathematical model of the P/VSA process consists of a set of partial differential and algebraic equations (PDAE). The system of PDAEs has been solved using a 2<sup>nd</sup> order

Orthogonal Collocation on Finite Elements Method (OCFEM) and/or Centered Finite Difference Method (CFDM) of 2<sup>nd</sup> order over a discretization grid of 150 intervals for the simulation and the optimization of the P/VSA process. For the case of the membrane cross-flow model, a semi-analytical solution has been employed (Geankoplis, 2003). The model is implemented in the gPROMS<sup>TM</sup> modeling environment, which has been previously proven to be a very efficient tool for simulation and optimization of P/VSA or membrane processes (Chatziasteriou et al., 2022).

#### 2.4. Optimization of the two-stage P/VSA process

Two-stage P/VSA processes have extensively been developed as promising alternatives to deep vacuum single-stage P/VSA separations. As a result, such a process has been used as reference in the present work. The results of this optimization study are presented in Table 3.

Table 3. Optimization results of the two-stage P/VSA reference case

Stage	P <sub>L</sub> (bar)	P <sub>I</sub> (bar)	P <sub>h</sub> (bar)	t <sub>ads</sub> (s)	u <sub>r</sub> (m/s)	Purity (%)	Recovery (%)	Energy (kWh/t)
1 <sup>st</sup>	0.1	0.92	1.2	30	0.69	48.4	100	85.4
2 <sup>nd</sup>	0.1	0.82	1.0	60	0.66	95.0	91.0	52.1
<b>Total</b>	-	-	-	-	-	<b>95.0</b>	<b>91.0</b>	<b>145.9</b>

It is seen that a two-stage P/VSA process can produce CO<sub>2</sub> meeting the DOE targets employing an industrially acceptable vacuum pressure of 0.1 bar. However, the use of two P/VSA units significantly increases the CAPEX and the OPEX of the process.

#### 2.5. Hybrid process optimization

Since a membrane unit has a lower capital and operating cost compared to a P/VSA unit, a two-stage process where a membrane unit replaces either the first or the second P/VSA stage of the reference case, may provide a more cost-effective total process, while the total energy requirement may be improved, as well. In the present work, we have examined the above design alternatives by simulation and optimization studies.

##### 2.5.1. Optimization of a P/VSA-membrane hybrid process

Optimization results of the membrane process and the total performance indicators of the P/VSA-membrane hybrid process are presented in Table 4. The total energy requirement is slightly larger than the reference case. As a result, the two configurations are comparable and they can provide the bases for further design and operating improvements.

Table 4. Optimization results of the P/VSA-membrane process

1 <sup>st</sup> stage	P <sub>L</sub> (bar)	P <sub>I</sub> (bar)	P <sub>h</sub> (bar)	t <sub>ads</sub> (s)	u <sub>r</sub> (m/s)	Purity (%)	Recovery (%)	Energy (kWh/t)
	0.1	0.92	1.2	30	0.69	48.4	100	85.4
2 <sup>nd</sup> stage	P <sub>L</sub> (bar)	P <sub>h</sub> (bar)	Permelectivity (-)	θ (-)	Purity (%)	Recovery (%)	Energy (kWh/t)	
	0.11	1.2	90	0.46	95	90	71.5	
<b>Total</b>	-	-	-	-	-	<b>95</b>	<b>90</b>	<b>166.4</b>

##### 2.5.2. Optimization of a membrane-P/VSA hybrid process

In the case of the membrane-P/VSA configuration (Table 5), the energy requirements are significantly lower comparing with both the reference process and the previously studied hybrid process. The energy reduction can be attributed to the lower energy needed by the first stage membrane compared to the P/VSA of the reference case, as well as the greater

recovery index of the second stage which affects the total energy demand. Note that the membrane's selectivity is, also, much lower in this case.

Table 5. Optimization results of the membrane-P/VSA process

1 <sup>st</sup> stage	P <sub>L</sub> (bar)	P <sub>h</sub> (bar)	Permelectivity (-)	θ (-)	Purity (%)	Recovery (%)	Energy (kWh/t)	
	0.23	1.0	25	0.59	48.4	95.5	73.9	
2 <sup>nd</sup> stage	P <sub>L</sub> (bar)	P <sub>i</sub> (bar)	P <sub>h</sub> (bar)	t <sub>ads</sub> (s)	u <sub>r</sub> (m/s)	Purity (%)	Recovery (%)	Energy (kWh/t)
	0.1	0.43	1.0	60	0.4	95.4	94.7	56.4
Total	-	-	-	-	-	95.4	90.4	134.4

### 3. Conclusions

A hybrid P/VSA-membrane process is a promising alternative for post-combustion CO<sub>2</sub> capture compared to a two-stage P/VSA process. Two different hybrid process configurations examined in this work provide high separation efficiency with low energy requirements. Specifically, the P/VSA-membrane configuration leads to a process with energy consumption of 166.4 kWh/tCO<sub>2</sub> (equivalent to 0.599 GJ/tCO<sub>2</sub>), greater than the reference two-stage P/VSA 145.9 kWh/tCO<sub>2</sub> (0.525 GJ/tCO<sub>2</sub>). However, the energy consumption in this case is not considerably higher. On the other hand, the membrane-P/VSA configuration results to an energy demand of 134.4 kWh/t CO<sub>2</sub> (0.484 GJ/tCO<sub>2</sub>), which is lower than the reference case. The results illustrate that a P/VSA process is more efficient for the separation of high CO<sub>2</sub> concentration mixtures, while membranes are preferable at CO<sub>2</sub> dilute mixtures. Combining the advantages of the two individual processes it can be revealed that a hybrid membrane-P/VSA process can be more efficient in terms of CAPEX and OPEX compared to the standard two-stage P/VSA process.

### Acknowledgements

The research project was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "1st Call for H.F.R.I. Research Projects to support Faculty Members and Researchers and the procurement of high-cost research equipment grant" (project number 2090).

### References

- Abanades J.C., Arias B., Lyngfelt A., Mattisson T., Wiley D.E., Li H., Ho M.T., Mangano E., Brandani S., 2015. Emerging CO<sub>2</sub> capture systems. *Int. J. Greenh. Gas Control.* 40, 126-166.
- Chatziasteriou C.C., Kikkinides E.S., Georgiadis M.C., 2022, Recent advances on the modeling and optimization of CO<sub>2</sub> capture processes, *Comput. Chem. Eng.*, 165, 107938.
- Geankoplis C. J., 2003, *Transport Processes and Separation Process Principles*, Prentice Hall PTR, Upper Saddle River, New Jersey.
- Haghpanah R., Majumder A., Nilam R., Rajendran A., Farooq S., Karimi I. A., Amanullah M., 2013, Multiobjective Optimization of a Four-Step Adsorption Process for Postcombustion CO<sub>2</sub> Capture Via Finite Volume Simulation, *Ind. Eng. Chem. Res.* 52, 11, 4249-4265.
- Haghpanah, R., Rajendran, A., Farooq, S., Karimi, I. A., 2014, Optimization of One- and Two-Stage Kinetically Controlled CO<sub>2</sub> Capture Processes from Postcombustion Flue Gas on a Carbon Molecular Sieve, *Ind. Eng. Chem. Res.* 53 (22): 9186-9198
- Khurana M., Farooq S., 2017, Integrated adsorbent-process optimization for carbon capture and concentration using vacuum swing adsorption cycles. *AIChE J.*, 63, 7, 2987–2995
- Zanco S.E., Joss L., Hefti M., Gazzani M., Mazzotti M., 2017, Addressing the criticalities for the deployment of adsorption-based CO<sub>2</sub> capture processes, *Energy Procedia*, 114, 2497-2505

# Integration of renewable energy and CO<sub>2</sub> capture and utilization technologies for decarbonization of energy intensive process industries

Calin-Cristian Cormos<sup>a</sup>, Letitia Petrescu<sup>a</sup>, Ana-Maria Cormos<sup>a</sup>, Simion Dragan<sup>a</sup>, Cristian Dinca<sup>b</sup>, Marius Sandru<sup>c</sup>

<sup>a</sup> Babes-Bolyai University, Faculty of Chemistry and Chemical Engineering, 11 Arany Janos, Cluj-Napoca, Postal code: RO-400028, Romania

<sup>b</sup> University Politehnica of Bucharest, Faculty of Power Engineering, 313 Splaiul Independentei, Postal code: RO-060042, Bucharest, Romania

<sup>c</sup> Sintef Industry, Sem Sælands vei 2 A, Trondheim, Norway

## Abstract

Renewable energy sources and CO<sub>2</sub> utilization technologies are predicted to play an important role in achieving the climate neutrality by the half of this century. This work evaluates from technical and environmental point of view the integration of renewable energy and CO<sub>2</sub> capture and utilization technologies for production of sustainable energy carriers / chemicals. As illustrative cases, Synthetic Natural Gas (SNG) and methanol production using renewable-based hydrogen and captured CO<sub>2</sub> were assessed at 100 MW thermal output. The overall processes have significant technical advantages (e.g., high overall energy efficiency up to 60%, more than 97% H<sub>2</sub> and CO<sub>2</sub> utilization yields etc.) as well as environmental benefits in term of reducing specific CO<sub>2</sub> emissions (nearly zero carbon emissions at the plant level and an overall negative emissions to contribute to the reduction of atmospheric CO<sub>2</sub> concentration), but significant scale-up efforts are required. **Keywords:** Renewable energy; Green hydrogen; CO<sub>2</sub> capture and utilization; Decarbonized energy carriers; Technical and environmental analysis.

## 1. Introduction

Currently the energy-intensive industrial processes are facing significant economic and environmental challenges. The large-scale utilization of renewable energy sources (e.g., solar, wind, biomass) and their integration with CO<sub>2</sub> Capture and Utilization (CCU) technologies are predicted to play an important role for development of low carbon economy to achieve the global climate neutrality (Alok et al., 2022). Figure 1 presents the overall integration of renewable energy sources and CO<sub>2</sub> capture and utilization technologies for production of sustainable carbon-neutral energy carriers / chemicals.

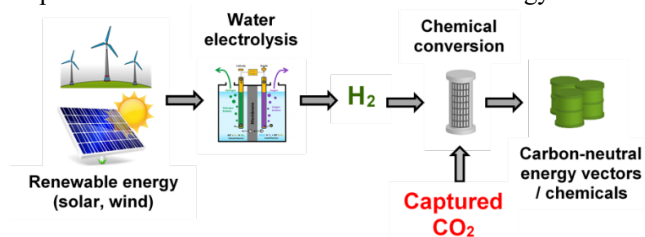
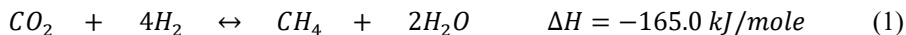


Figure 1. Integration of renewable power and CO<sub>2</sub> capture and utilization technologies

This work evaluates the potential decarbonization of several industrial processes (e.g., power generation, cement, petro-chemistry) by integration of renewable energy sources and the CO<sub>2</sub> capture and utilization technologies. As investigated systems, methanol and synthetic methane production were evaluated from captured CO<sub>2</sub> and green hydrogen (generated from water electrolysis using renewable power) were assessed at relevant industrial sizes (100 MW thermal output for both evaluated options). As key novelty aspects, the overall technical and environmental implications of integrated renewable and CO<sub>2</sub> utilization technologies at relevant industrial scales were evaluated using process modeling, model validation and thermal integration for global energy optimization (Szima and Cormos, 2018). For comparison reason, the current non-capture fossil-based processes were also considered to assess the overall benefits of proposed CCU systems.

## 2. Plant configurations and main design assumptions

The first investigated system integrating green hydrogen (produced from renewable power) and captured CO<sub>2</sub> from various energy-intensive industrial applications refers to the production of Synthetic Natural Gas (SNG) according to Sabatier chemical reaction:



The conceptual layout of SNG plant from renewable power (wind and solar) and captured CO<sub>2</sub> is presented in Figure 2. The methanation reaction is strongly exothermic and the reaction heat is recovered as steam which is then expanded to produce electricity. The produced electricity is totally covering the ancillary plant consumption and the excess is exported to the grid. This co-generation capability of SNG and power is a particular important aspect of this process contributing to the optimization of plant performances.

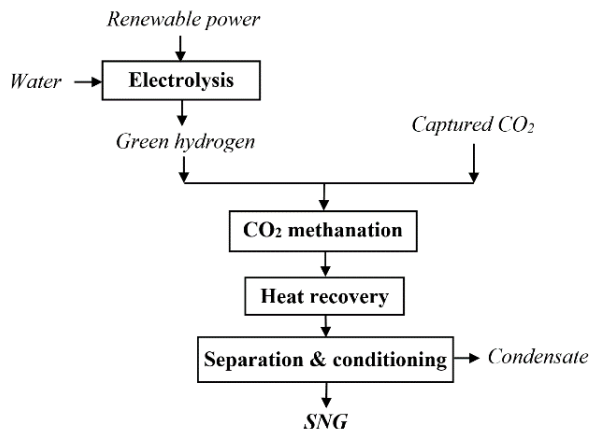
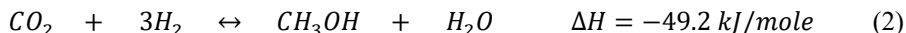


Figure 2. Conceptual layout of Substitute Natural Gas (SNG) plant

The second investigated system involving renewable-based green hydrogen and captured CO<sub>2</sub> is referring to the production of methanol that can be used both as energy carrier and chemical. The chemical reaction for producing methanol from is presented below:



The conceptual layout of methanol plant from renewable power (wind and solar) and captured CO<sub>2</sub> (from various energy-intensive processes) is presented in Figure 3.

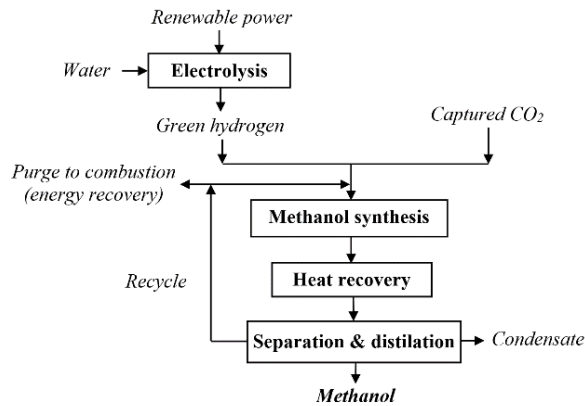


Figure 3. Conceptual layout of methanol plant

As in the case of SNG, the methanol plant has also available heat that is converted into steam and then in electricity to cover the ancillary consumption and the excess was used for export to the grid. The main design assumptions of SNG and methanol plants are presented in Table 1. Both CO<sub>2</sub> utilization systems were simulated using process flow modeling - ChemCAD, then systems were subject to detailed thermal integration analysis.

Table 1. Key design assumptions of SNG / methanol plants

<b>Unit</b>	<b>Design parameters</b>
Captured CO <sub>2</sub> composition (vol.)	99.00% carbon dioxide, 0.99% non-condensable gases (nitrogen, argon), 0.01% water
Water electrolysis unit	Power consumption: 54 kWh/kg Hydrogen purity: 99.90% vol.
CO <sub>2</sub> methanation unit	Ni-based catalyst Reactor temperature & pressure: 300°C & 50 bar Kinetic reactor Thermal mode: heat exchanger type Pressure drop: 1 bar
Methanol synthesis unit	Cu/ZnO/Al <sub>2</sub> O <sub>3</sub> catalyst Reactor temperature & pressure: 220°C & 80 bar Kinetic reactor Thermal mode: heat exchanger type Pressure drop: 1 bar

SNG compression unit	Delivery pressure: 60 bar Compressor efficiency: 85%
Methanol distillation unit	Methanol purity: 99.90% Distillation column: 12 stages Column feed temperature: 80°C
Heat recovery and steam cycle	Steam temperature & pressure: 200 - 280°C & 3 bar Condensation pressure: 0.046 bar Steam turbine efficiency: 86% Turbine steam wetness: max. 10%
Heat exchangers	$\Delta T_{\min.} = 10^{\circ}\text{C}$ Pressure drop: 2 - 5% of inlet pressure
Thermodynamic package	Soave-Redlich-Kwong (SRK)

---

### 3. Results and discussions

After modeling and simulation, the developed models were subject to model validation in comparison to experimental / industrial data. The model validation was done only for the methanation and methanol reactors since the rest of the processes (e.g., heat recovery, gas separation, distillation) is already common in chemical applications. The main simulated performance indicators (e.g.,  $\text{H}_2$  and  $\text{CO}_2$  conversion rates) were fully in line with experimental data (Koytsoumpa and Karellas, 2018; Lombardelli et al., 2022). For optimization of overall energy efficiency, the thermal integration analysis using pinch methodology was used (Smith, 2016). For SNG plant, the composite curves are presented in Figure 4 showing a significant heat recovery potential in form of steam generation.

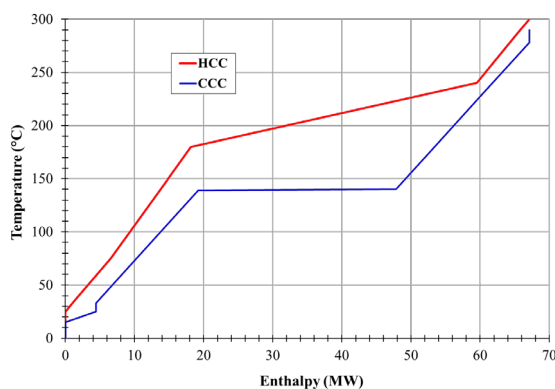


Figure 4. Thermal integration analysis of SNG plant

The main performance indicators of SNG plant are presented in Table 2. One can noticed the high conversion rates (>98%) of both used reactants (CO<sub>2</sub> and H<sub>2</sub>), a promising overall energy efficiency of the global system as well as an almost zero specific CO<sub>2</sub> process emissions at the plant level. In fact, considering the renewable nature of the power used for water electrolysis to produce the required green hydrogen stream, the overall CO<sub>2</sub> emissions are negative (contributing to reduction of CO<sub>2</sub> concentration from atmosphere).

Table 2. Main performance indicators of SNG plant

<b>Parameter</b>	<b>Unit</b>	<b>Value</b>
CO <sub>2</sub> flowrate	t/h	18.45
CO <sub>2</sub> conversion yield	%	98.90
H <sub>2</sub> flowrate	t/h	3.38
H <sub>2</sub> conversion yield	%	98.10
Specific power consumption for water electrolysis	kWh/kg	54.00
Power consumption for H <sub>2</sub> production	MW <sub>e</sub>	182.52
SNG production capacity	t/h	6.88
SNG lower heating value	MJ/kg	52.30
SNG thermal output	MW <sub>th</sub>	100.00
Gas decompression	MW <sub>e</sub>	0.60
Gross power output	MW <sub>e</sub>	5.95
Net power output	MW <sub>e</sub>	6.55
Overall energy efficiency	%	58.37
Specific CO <sub>2</sub> emissions	kg/kg SNG	0.05

Similar for methanol synthesis, the thermal integration was used for overall energy optimization (see Figure 5 for hot and cold composite curves). One can noticed comparing Figures 4 and 5 that the SNG case has the available heat sources at higher temperature than for the methanol case which will have an influence on the overall energy efficiency.



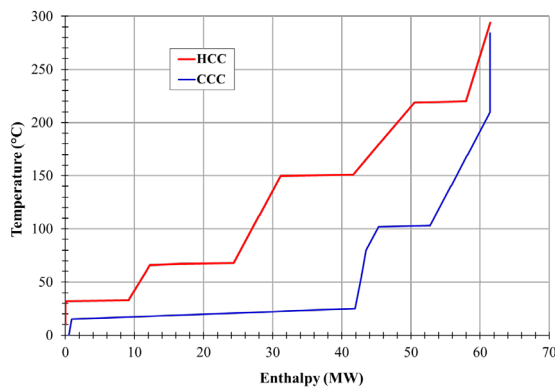


Figure 5. Thermal integration analysis of methanol plant

The main performance indicators of methanol plant are presented in Table 3. As in the case of SNG plant, high conversion rates for both reactants as well as overall energy efficiency coupled with almost zero CO<sub>2</sub> emissions were reported. For comparison reason, the conventional methanol process starting from natural gas (fossil source) without carbon capture has a specific CO<sub>2</sub> emission of about 95 - 110 kg/GJ which correspond to about 0.5 - 0.6 kg/kg (IRENA, 2021; Khojasteh-Salkuyeh et al., 2021). Comparing both investigated CO<sub>2</sub> utilization processes, one can noticed that the methanol plant has a slightly less overall energy efficiency which can be explained by lower temperature of available heat sources (as discussed above) as well as a more complex design of the overall system (e.g., involving a distillation of raw methanol which requires additional thermal energy for separation) with inherent higher energy consumptions.

Table 3. Key technical performance indicators of methanol plant

Parameter	Unit	Value
CO <sub>2</sub> flowrate	t/h	25.50
CO <sub>2</sub> conversion yield	%	97.35
H <sub>2</sub> flowrate	t/h	3.50
H <sub>2</sub> conversion yield	%	99.85
Specific power consumption for water electrolysis	kWh/kg	54.00
Power consumption for H <sub>2</sub> production	MW <sub>e</sub>	187.20
Methanol production capacity	t/h	18.07
Methanol lower heating value	MJ/kg	19.92
Methanol thermal output	MW <sub>th</sub>	100.00
Plant ancillary consumption	MW <sub>e</sub>	-4.95

Gross power output	MW <sub>e</sub>	6.00
Net power output	MW <sub>e</sub>	1.05
Overall energy efficiency	%	53.98
Specific CO <sub>2</sub> emissions	kg/kg MeOH	0.04

Although very promising in delivering better technical and environmental performance indicators, the evaluated CO<sub>2</sub> capture and utilization (CCU) systems still requires significant scale-up efforts from the current development level (up to several MW) to full industrial sizes (Bailera et al., 2017). In addition, the utilization of green hydrogen as raw material for CO<sub>2</sub> utilization technologies requires an efficient energy storage facility considering the time intermittency of solar and wind resources (Miehling et al., 2022).

#### 4. Conclusions

The present work evaluates the potential benefits of integration the renewable energy sources with CO<sub>2</sub> capture and utilization (CCU) technologies for production of sustainable carbon-neutral fuels at industrial level capacities. As illustrative cases, substitute natural gas and methanol production from green hydrogen and captured CO<sub>2</sub> were assessed using process modeling and thermal integration. As the results show, the integration of renewable energy and CO<sub>2</sub> capture and utilization technologies brings significant technical advantages (e.g., high overall energy efficiency 54 - 58%, more than 97% hydrogen and CO<sub>2</sub> utilization yields etc.) as well as positive environmental benefits in term of reducing specific CO<sub>2</sub> emissions (nearly zero carbon emissions at the plant level and overall negative emissions to contribute to the reduction of atmospheric CO<sub>2</sub> concentration) compared to the current non-capture fossil-based concepts. However, significant scale-up is still needed to deploy these technology to relevant industrial level.

#### Acknowledgements

This work was supported by the NO Grants 2014 - 2021, under project contract no. 13/2020 and a grant of the Romanian Ministry of Education and Research, CCCDI - UEFISCDI, project number PN-III-P4-ID-PCE-2020-0032, within PNCDI III.

#### References

- A. Alok, R. Shrestha, S. Ban, S. Devkota, B. Uprety, R. Joshi, 2022, Technological advances in the transformative utilization of CO<sub>2</sub> to value-added products, *Journal of Environmental Chemical Engineering*, 10, 106922
- M. Bailera, P. Lisbona, L.M. Romeo, S. Espatolero, 2017, Power to Gas projects review: Lab, pilot and demo plants for storing renewable energy and CO<sub>2</sub>, *Renewable and Sustainable Energy Reviews*, 69, 292-312
- International Renewable Energy Agency (IRENA), 2021, Innovation outlook - Renewable methanol
- Y. Khojasteh-Salkuyeh, O. Ashrafi, E. Mostafavi, P. Navarri, 2021, CO<sub>2</sub> utilization for methanol production; Part I: Process design and life cycle GHG assessment of different pathways, *Journal of CO<sub>2</sub> Utilization*, 50, 101608
- E.I. Koytsoumpa, S. Karellas, 2018, Equilibrium and kinetic aspects for catalytic methanation focusing on CO<sub>2</sub> derived Substitute Natural Gas (SNG), *Renewable and Sustainable Energy Reviews*, 94, 36-50

- G. Lombardelli, M. Mureddu, S. Lai, F. Ferrara, A. Pettinau, L. Atzoric, A. Conversano, M. Gatti, 2022, CO<sub>2</sub> hydrogenation to methanol with an innovative Cu/Zn/Al/Zr catalyst: Experimental tests and process modeling, *Journal of CO<sub>2</sub> Utilization*, 65, 102240
- S. Miehling, S. Fendt, H. Spliethoff, 2022, Optimal integration of Power-to-X plants in a future European energy system and the resulting dynamic requirements, *Energy Conversion and Management*, 251, 115020
- R. Smith, 2016, *Chemical process design and integration*, second edition, Wiley, Hoboken, USA
- S. Szima, C.C. Cormos, 2018, Improving methanol synthesis from carbon-free H<sub>2</sub> and captured CO<sub>2</sub>: A techno-economic and environmental evaluation, *Journal of CO<sub>2</sub> Utilization*, 24, 555-563

# Simultaneous optimisation of energy recovery and water recirculation in a ceramic plant

Miguel Castro Oliveira,<sup>a,b</sup> Muriel Iten,<sup>a</sup> Henrique A. Matos,<sup>b</sup>

<sup>a</sup> *Low Carbon and Resource Efficiency, R&Di, Instituto de Soldadura e Qualidade, 4415 491 Grijó, Portugal*

<sup>b</sup> *Centro Recursos Naturais e Ambiente (CERENA), Department of Chemical Engineering, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal*

## Abstract

The ongoing energy crisis and water shortages occurring in many EU Member countries have reinforced the importance of the achievement of the aims associated to sustainability policies. In this work, a non-linear programming (NLP) model developed with the Python language for the virtual implementation in a ceramic plant of a Water and Energy Integration System (WEIS). This is an innovative process integration and system retrofitting-based concept aiming for simultaneous energy efficiency improvement and water minimisation in industry. A set of indicators were assessed in post-processing for the results obtained to the ceramic plant WEIS project, namely a payback period of about 1 year and 4 months and an emission reduction level of 2.97 kton CO<sub>2,eq</sub>/year. These are highly favourable compared to industrial benchmarks. The developed optimisation model is part of a research & development initiative designated ThermWatt, set to create a customised computational tool for water, energy and waste management in end-use sectors and an Engineering consultancy service around it.

**Keywords:** Water and Energy Integration Systems, energy efficiency, water minimisation, ceramic industry, Python NLP-model.

## 1. Introduction

The improvement of water and energy efficiency is an important concern in the context of the most recent sustainable development requirements in the world. In the scope of the reduction of the significant use of both energy and water in end-use sectors, the authors have developed the innovative concept of Water and Energy Integration Systems (WEIS) (Castro Oliveira et al., 2022). These are types of systems that include a certain number of water-using processes and combustion-based thermal processes and all potential stream recirculation between these processes.

Several theoretical approaches have been proposed in academia for the problems related to heat recovery systems (fuel and electricity use reduction) (Castro Oliveira et al., 2022a) and water allocation and heat exchanger networks (freshwater and hot/ cold utilities use reduction) (Ibrić et al., 2016). While the first one has been proved theoretically and practically implemented (although not in a holistic perspective, considering all potential technologies and scenarios), the latter has been essentially studied in theory only, with the existing studies only focusing on a small number of end-use sectors.

The concept of WEIS has been introduced to encompass both the phenomenon of waste heat stream recirculation to produce fuel savings in combustion-based processes (as well as electricity generation in thermodynamic cycles) and the concept of water allocation and heat exchanger network mentioned afore. Furthermore, it also includes the integration of energy recovery from wastewater technologies for the production of additional fuel quantities to be fed to combustion-based processes. The wastewater treatment units to be

included within water systems are heat-driven ones (such as Multi-effect distillation) (Castro Oliveira et al., 2022b), so to make possible the allocation of enthalpy from waste heat streams originating from combustion-based processes. The conceptualization of WEIS consists on the analysis of all potential interdependencies between water and energy streams to generate savings in both, thus subsisting on the promotion of the water-energy nexus character associated to this type of process (Castro Oliveira et al., 2019). This work presents the development of an optimisation model considering simultaneous energy recovery and water recirculation in a ceramic plant. A post-processing assessment is performed based on the determination of indicators associated to economic and environmental impact viabilities.

## 2. Contextualization of the Case-study

The case-study is set within a ceramic industry plant installed in Portugal. The plant is characterized by the existence of 2 combustion-based processes (tunnel kilns) and 4 water-using processes (each one being set to remove a determinate quantity of salt contaminant). In Table 1, the operational data for the case-study is presented.

**Table 1.** Operational data for each combustion-based process and water-using line

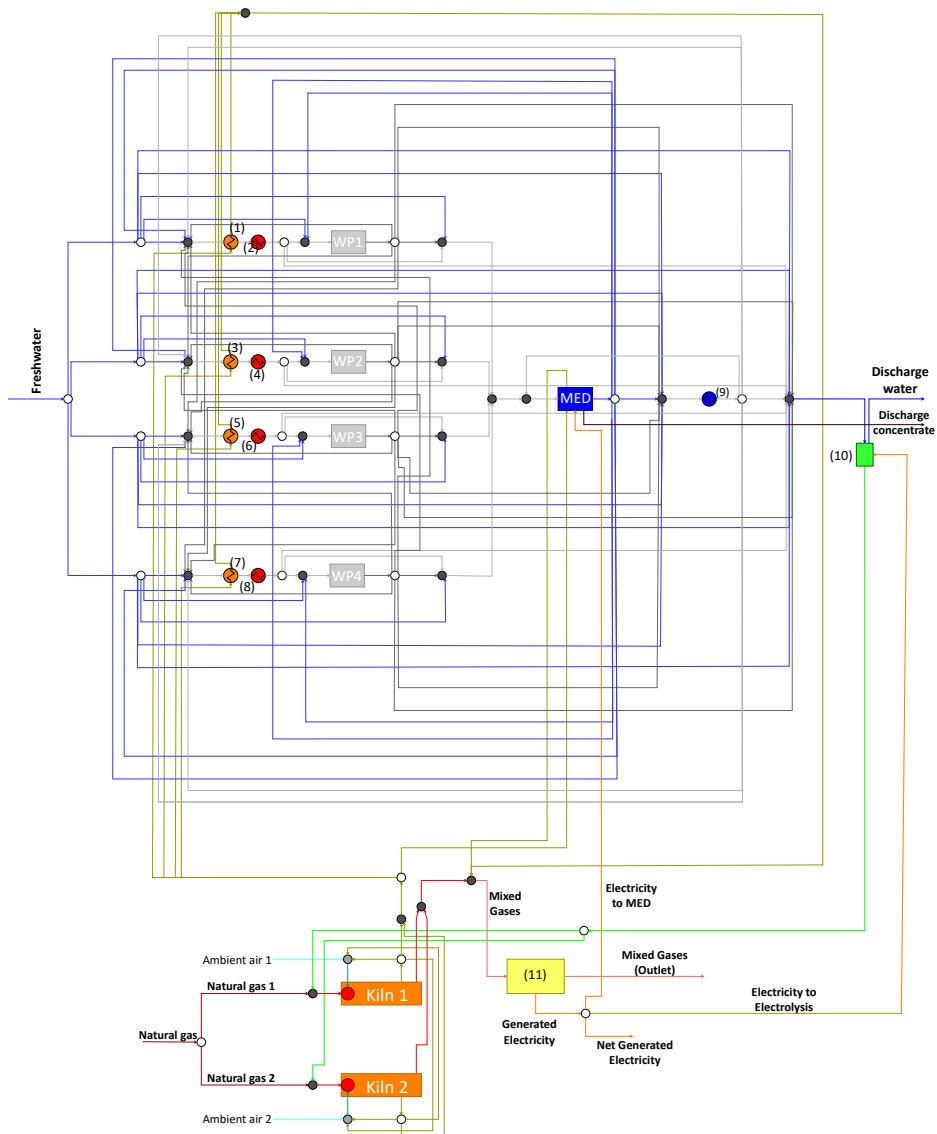
Process	Natural gas flow rate (kg/h)	Combustion-based processes			
		Hot Air		Exhaust gases	
		Flow rate (kg/h)	Temp. (°C)	Flow rate (kg/h)	Temp. (°C)
Kiln 1	196.5	20711.7	193.2	16480.5	230.0
Kiln 2	103.5	23583.1	95.0	14454.5	159.3

Water-using Lines	Water system						Hot Utility (MJ/year)	Cold Utility (MJ/year)
	Water			Salt				
	Flow rate (kg/h)	Final Temp. (°C)	Flow rate (kg/h)	Concentration (ppm)				
				WP Inlet	WP Outlet			
			Max.	Min.	Max.			
Line 1	474.83	95	0.47483	200	800	1000	495.00	
Line 2	477.12	95	0.47712	200	800	1000	495.00	
Line 3	365.43	95	0.36540	200	800	1000	495.00	
Line 4	58.79	100	0.05873	200	800	1000	130.00	
Discharge					0.00	1000		396.00
Total	1376.17						1615.00	396.00

The presented data is the foundation for the project of a WEIS in the ceramic plant. In terms of hot air/ exhaust gas streams allocation, the configuration may be described by the following sequence steps: a part of the hot air stream at the outlet of a kiln may be recirculated to the same kiln or the other as part of the total combustion air; the remaining quantities of the two hot air streams are mixed and then directed to the water system, with the stream being divided for each water-using line, and in its turn for each heater and saline water treatment Multi-effect distillation (MED) unit (5 effects and 1 condenser); the hot air streams at the heaters and MED's outlets are mixed with themselves and the two exhaust gas stream from each kiln and then directed to the HRSG unit of an Organic Rankine cycle (ORC) electricity generation; an Electrolysis unit is installed to produced hydrogen fuel from discharge water stream from the water system, which is then feed as additional fuel to each kiln in the form of hydrogen-enriched natural gas (HENG).

The allocation of water streams (recycling, reuse and by-pass) occurs as follows: from the water-using process (WP) to the MED unit inlet, heater inlet, cooler inlet, the water system outlet, other 3 heaters and other 3 coolers; from the MED unit to water-using line inlet, heater inlet, cooler inlet, the water system discharge, other 3 heaters and other 3 coolers; from the heater to water-using line inlet, MED unit inlet and water system discharge; from the cooler to water-using line inlet, MED unit inlet and water system discharge. The described strategy (in the form of stream allocation-based superstructure) is pictorially represented in Figure 1.



**Figure 1.** Superstructure for the conceptualized Water and Energy Integration System. Legend: ● -stream mixing; ○ -stream splitting ; — Exhaust Gases ; — Hydrogen; — hot air; (1) – Economiser 1; (2) – Heater 1; (3) – Economiser 2; (4) – Heater 2 ; (5) – Economiser 3; (6) – Heater 3; (7) – Economiser 4; (8) – Heater 4; (9) – Cooler; (10) – Electrolysis Unit; (11) – Organic Rankine Cycle

### 3. Optimisation Model

The optimisation model approached in this work was developed using the non-linear programming (NLP) methodology, with the GEKKO package built in the Python language. In Table 2, the developed NLP model is characterized, with the decision variables and constraints being presented in an aggregated and generic manner. The

optimal solution in the form of the optimal WEIS scenario flowsheet is presented in Figure 2, considering all the results for stream allocation and sizing parameters.

**Table 2.** Characterization of the optimisation problem

<b>Combustion-based Processes System</b>	
Decision Variables	
<ul style="list-style-type: none"> <li>Natural gas flow rates (<math>M_{\text{Fuel}}</math>)</li> <li>Ambient air flow rates (<math>M_{\text{Amb,Air}}</math>)</li> <li>Exhaust gases flow rates (<math>M_{\text{Ex}}</math>)</li> <li>Exhaust gases specific enthalpies (<math>h_{\text{Ex}}</math>)</li> <li>Exhaust gases temperatures (<math>T_{\text{Ex}}</math>)</li> <li>Hot air flow rates (<math>M_{\text{Hot Air}}</math>)</li> <li>Hot air specific enthalpies (<math>h_{\text{Hot Air}}</math>)</li> <li>Hot air specific temperatures (<math>T_{\text{Hot Air}}</math>)</li> </ul>	<ul style="list-style-type: none"> <li>Recirculated air flow rates (<math>M_{\text{Rec,Air}}</math>)</li> <li>Recirculated air specific enthalpies (<math>h_{\text{Rec Air}}</math>)</li> <li>Recirculated air temperatures (<math>T_{\text{Rec Air}}</math>)</li> <li>Air-to-fuel ratios (AF)</li> <li>Fuel lower heating value (<math>\text{LHV}_{\text{Fuel}}</math>)</li> <li>Generated Electricity (Elec)</li> <li>Appreciably Generated Electricity (<math>\text{Elec}_{\text{eff}}</math>)</li> <li>Electricity used in Electrolysis (<math>\text{Elec}_{\text{Electrolysis}}</math>)</li> </ul>
Relevant Inequality Constraints	
$T_{\text{Ex}} \leq 75^\circ\text{C}$	
Relevant Equality Constraints	
$M_{\text{Fuel,Baseline}} \cdot \text{LHV}_{\text{Natural Gas}} = M_{\text{Fuel}} \cdot \text{LHV}_{\text{Fuel}} + M_{\text{Comb,Air}} \cdot (h_{\text{Comb,Air}} - h_{\text{Comb,Air,Baseline}})$ (1)	
$M_{\text{Comb,Air}} = M_{\text{Recyc,Air}} + M_{\text{Amb,Air}}$ (2)	
$M_{\text{Comb,Air}} \cdot h_{\text{Comb,Air}} = M_{\text{Recyc,Air}} \cdot h_{\text{Recyc,Air}} + M_{\text{Amb,Air}} \cdot h_{\text{Amb,Air}}$ (3)	
$M_{\text{Fuel}} + M_{\text{Comb,Air}} = M_{\text{Ex}}$ (4)	
$M_{\text{Fuel}} \cdot \text{AF} = M_{\text{Comb,Air}}$ (5)	
$\text{AF} \cdot \text{LHV}_{\text{Natural Gas}} = \text{AF}_{\text{Baseline}} \cdot \text{LHV}_{\text{Fuel}}$ (6)	
$\text{LHV}_{\text{Fuel}} = y_{\text{Natural Gas}} \cdot \text{LHV}_{\text{Natural Gas}} + y_{\text{Hydrogen}} \cdot \text{LHV}_{\text{Hydrogen}}$ (7)	
$M_{\text{gas,in,ORC}} \cdot (h_{\text{gas,in,ORC}} - h_{\text{gas,out,ORC}}) \cdot 0.06 = \text{Elec} \cdot 3600$ (8)	
$\text{Elec} = \text{Elec}_{\text{eff}} + \text{Elec}_{\text{Electrolysis}}$ (9)	
<b>Water System</b>	
Decision Variables	
<ul style="list-style-type: none"> <li>Freshwater mass flow rate (<math>M_{\text{FW}}</math>)</li> <li>Each water stream mass flow rate (<math>M_{\text{w}}</math>)</li> <li>Each water stream specific enthalpy (<math>h_{\text{w}}</math>)</li> <li>Each water stream temperatures (<math>T_{\text{w}}</math>)</li> <li>Each water stream contaminant concentration (<math>C_{\text{w}}</math>)</li> <li>Heat transfer areas (A)</li> <li>Withdrawn enthalpy from hot air (<math>q_{\text{with}}</math>)</li> <li>Treated water mass flow rate produced in the second-to-last effects and condenser of the MED unit (<math>M_{\text{TW}}</math>)</li> </ul>	<ul style="list-style-type: none"> <li>Vapour mass flow rate from a MED unit effect (<math>M_{\text{v}}</math>)</li> <li>Concentrate water mass flow rate from a MED unit effect (<math>M_{\text{Conc}}</math>)</li> <li>Discharged water mass flow rate from the water system (<math>M_{\text{Disch}}</math>)</li> <li>Produced vapour (MED unit outlet) specific enthalpy (<math>h_{\text{v}}</math>)</li> <li>Produced hydrogen mass flow rate (<math>M_{\text{Hydrogen}}</math>)</li> <li>Water-to-hydrogen conversion ratio (<math>X_{\text{Hydrogen}}</math>)</li> </ul>
Relevant Inequality Constraints	
$C_{\text{w,WP,inlet}} \leq 200 \text{ ppm}$	
$C_{\text{w,WP,outlet}} \geq 800 \text{ ppm}$	
$C_{\text{w,WP,outlet}} \leq 1000 \text{ ppm}$	
Relevant Equality Constraints	
$M_{\text{W,to-be-split}} = \sum_{i=1} M_{\text{W,split},i}$ (10)	
$h_{\text{W,to-be-split}} = h_{\text{W,split},i}$ (11)	
$C_{\text{W,to-be-split}} = C_{\text{W,split},i}$ (12)	
$\sum_{i=1} M_{\text{W,to-be-mixed},i} = M_{\text{W,mixed}}$ (13)	
$\sum_{i=1} M_{\text{W,to-be-mixed},i} \cdot h_{\text{W,to-be-mixed},i} = M_{\text{W,mixed}} \cdot h_{\text{W,mixed}}$ (14)	
$\sum_{i=1} M_{\text{W,to-be-mixed},i} \cdot C_{\text{W,to-be-mixed},i} = M_{\text{W,mixed}} \cdot C_{\text{W,mixed}}$ (15)	
$M_{\text{W,in,eff}} \cdot 1/5 = M_{\text{TW,eff}} + M_{\text{Conc,eff}}$ (16)	
$q_{\text{with,MED}} = M_{\text{V,eff}} \cdot (h_{\text{V,eff}} - 418.896) + M_{\text{W,in,eff}} \cdot 1/5 \cdot (418.896 - h_{\text{w,in,eff}})$ (17)	
$M_{\text{TW,eff},k-1} \cdot (h_{\text{V,eff},k-1} - 418.896) = M_{\text{V,eff},k} \cdot (h_{\text{V,eff},k} - 418.896) + M_{\text{W,in,eff},k} \cdot 1/5 \cdot (418.896 - h_{\text{w,in,eff}})$ (18)	
$M_{\text{TW,eff}} \cdot (2675.43 - 418.896) = M_{\text{W,in,eff}} \cdot 1/5 \cdot (h_{\text{V,eff}} - 418.896)$ (19)	
$M_{\text{TW,eff},5} \cdot (h_{\text{V,eff},5} - 418.896) = M_{\text{W,in,MED}} \cdot 1/5 \cdot (h_{\text{w,in,eff}} - h_{\text{w,in,MED}})$ (20)	
$q_{\text{with}} = 400 \cdot A \cdot \left( (T_{\text{Air,in}} - T_{\text{w,out}}) \cdot (T_{\text{Air,out}} - T_{\text{w,in}}) \cdot \left( (T_{\text{Air,in}} - T_{\text{w,out}}) + (T_{\text{Air,out}} - T_{\text{w,in}}) \right) \cdot 0.5 \right)^{1/3}$ (21)	
$M_{\text{Disch,eff}} / 0.01801528 \cdot X_{\text{Hydrogen}} = M_{\text{Hydrogen}} / 0.002016$ (22)	
$M_{\text{Disch,eff}} / 0.01801528 \cdot 285.85 \cdot X_{\text{Hydrogen}} = \text{Elec}_{\text{Electrolysis}}$ (23)	
<b>Objective-function (€/h) (Natural gas, Electricity and Freshwater unitary prices for Portugal)</b>	

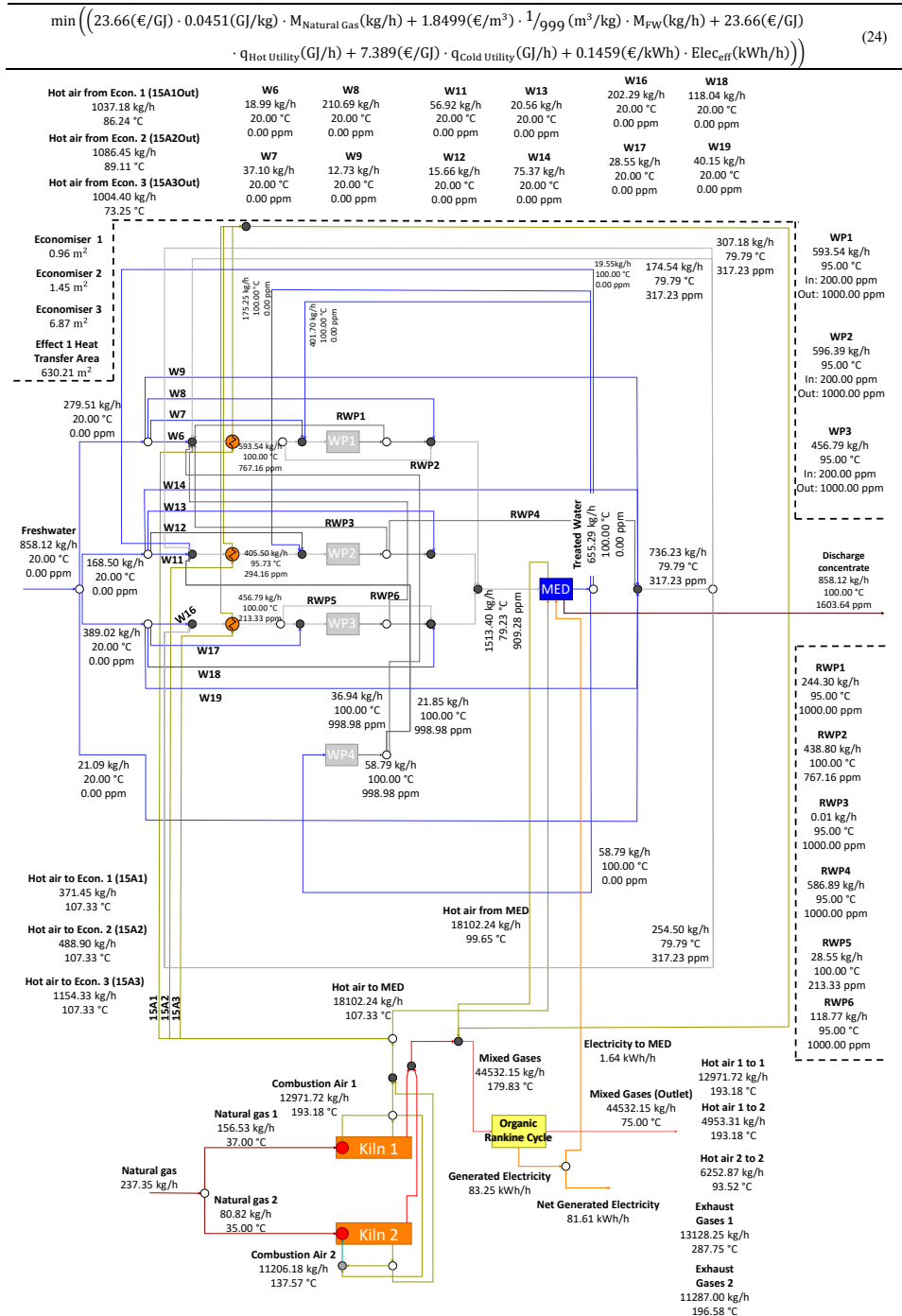


Figure 2. Flowsheet for the optimal WEIS scenario

#### 4. Post-processing: Economic and Environmental Impact Assessments

In Table 3, the results for economic and environmental impact assessment are presented.



**Table 3.** Economic and Environmental Impact Reduction Assessments

<b>Natural gas consumption (kg/h)</b>					
<b>Process</b>	<b>Initial</b>	<b>Improved</b>	<b>Relative Savings Share</b>	<b>Savings (€/h)</b>	
<b>Kiln 1</b>	196.50	156.53	20.34%	44.54	
<b>Kiln 2</b>	103.50	80.82	21.91%	25.27	
<b>Hot and Cold utilities consumption (MJ/h)</b>					
<b>Heater 1</b>	495.00	0.00	100.00%	11.71	
<b>Heater 2</b>	495.00	0.00	100.00%	11.71	
<b>Heater 3</b>	495.00	0.00	100.00%	9.55	
<b>Heater 4</b>	130.00	0.00	100.00%	3.08	
<b>Cooler</b>	396.00	0.00	100.00%	16.04	
<b>Water consumption (m<sup>3</sup>/h)</b>					
<b>Initial</b>	<b>Improved</b>	<b>Relative Savings Share</b>	<b>Savings (€/h)</b>		
1.38	0.86	23.50%	0.96		
<b>Electricity Balances (kWh/h)</b>					
<b>Net Electricity Generation (kWh/h)</b>				<b>Savings (€/h)</b>	
81.61				16.04	
<b>Final assessment</b>					
<b>CAPEX (k€)</b>	<b>OPEX (k€/year)</b>	<b>Savings (k€/year)</b>	<b>Payback Time (Years)</b>	<b>CO<sub>2,eq</sub> emissions reduction (kton/year)</b>	
1241.86	84.31	1054.47	1.28	2.97	

## 5. Conclusions

This work approaches the assessment of potential improvements of the overall water and energy efficiencies in a ceramic industry plant through the development and further use of an optimisation model using the Python language. The model was developed with the aim to apply the newly created methodology of Water and Energy Integration Systems (WEIS), which has only yet been proved in concept. The model proved to be valuable for the achievement of the proposed objective of reducing overall water and energy-related costs and associated pollutant emissions generated by water discharge and combustion gases discharge. A set of indicators were assessed in post-processing, having been obtained a payback period of about 1 year and 4 months and an emission reduction level of 2.97 kton CO<sub>2,eq</sub>/year, which are highly favourable compared to industrial benchmarks.

## Acknowledgements

The publishing procedure and conference participation have received funding through ISQ – Instituto de Soldadura e Qualidade through Programa Interface/ CIT – Centro de Interface Tecnológicos and CERENA under grant UIDB/04028/2020\_UIDP/04028/2020.

## References

- M. Castro Oliveira, M. Iten, H.A. Matos, J. Michels, 2019, Water-energy nexus in typical industrial water circuits, *Water* 11 (4), 699.
- M. Castro Oliveira, M. Iten, H.A. Matos, 2022, Review on Water and Energy Integration in Process Industry: Water-Heat Nexus, *Sustain.*, 14, 7954.
- M. Castro Oliveira, M. Iten, H.A. Matos, 2022a, Simulation and assessment of an integrated thermal processes and Organic Rankine Cycle (ORC) system with Modelica, 8, 764 – 770.
- M. Castro Oliveira, P. Coelho, M. Iten, H.A. Matos, 2022b, Modelling of Heat-Driven Water Treatment Systems: Multi-Effect Distillation (MED) model in Modelica. *Comput. Aided Chem. Eng.*, 51, 397–402.
- N. Ibrić, E. Ahmetović, Z. Kravanja, 2016, Mathematical programming synthesis of non-isothermal water networks by using a compact/reduced superstructure and an MINLP model. *Clean Technol. Environ. Policy*, 18, 1779–1813.

# Embedding Flexibility to the Design of Pressure-Vacuum Swing Adsorption Processes for CO<sub>2</sub> Capture

Steven Sachio<sup>a,b</sup>, Adam Ward<sup>a,b</sup>, Ronny Pini<sup>a,b</sup>, Maria M. Papathanasiou<sup>a,b</sup>

*<sup>a</sup>Sargent Centre for Process Systems Engineering, Imperial College London, London, United Kingdom SW72AZ*

*<sup>b</sup>Department of Chemical Engineering, Imperial College London, London, United Kingdom SW72AZ*

## Abstract

Capture of carbon dioxide (CO<sub>2</sub>) from post-combustion flue gases is an engineered solution to enable global transition to a sustainable energy system. For this, typically, pressure-vacuum swing adsorption (PVSA) is used. The design of PVSA processes traditionally relies on solving complex, constrained multi-objective optimization problems for the identification of good candidate operating points. Such processes are not agile as they are constrained to operate around a single optimal point. Examining process flexibility underpins the identification of suitable control strategies that handle disturbances. In this work, we deploy a machine learning-aided design space identification framework for simultaneous quantification of process flexibility and preliminary Pareto front identification to design flexible and controllable PVSA processes. The framework enables reducing the number of simulations required to identify good candidate operating points by 77%, with a difference of 1.1% in the cost optimum solution compared to performing rigorous optimization by NSGA-II. Additionally, the framework identifies that the cost optimal operation is highly non-flexible, suggesting that new design approaches are required to allow for the design of efficient and controllable PVSA processes.

**Keywords:** computing and systems engineering, process design, carbon capture

## 1. Introduction

One of the most significant contributions to anthropogenic carbon dioxide (CO<sub>2</sub>) emissions is the elution of flue gases from fossil fuel-fired power plants. It is likely that the combustion of fossil fuels will remain in widespread use over the next several decades while the world undergoes a transition to a sustainable energy system. Therefore, there is a need to design efficient processes for removing CO<sub>2</sub> from such flue gases to reduce emissions to the environment. Gas-phase separation by pressure-vacuum swing adsorption (PVSA) is a promising technology for performing post-combustion CO<sub>2</sub> capture (Raganati et al, 2021). In a PVSA process, flue gas is exposed to a fixed bed of solid adsorbent material to selectively remove CO<sub>2</sub>. The fixed bed then elutes a CO<sub>2</sub> deficient effluent stream which can be vented to the atmosphere. The CO<sub>2</sub> which has been adsorbed in the bed can be subsequently recovered at high purity to be sent for downstream utilization or geological sequestration. The computational design of PVSA processes for CO<sub>2</sub> capture is typically quite challenging (Balashankar et al, 2019). First, description of the highly non-linear process dynamics with a suitable mathematical model

requires a significant amount of computational time to evaluate the process performance for a given set of operating conditions. Second, the design problem requires the solution of constrained multi-objective optimization to find a Pareto front of optimum productivity/energy usage, while satisfying regulatory requirements on the purity and recovery of the extracted CO<sub>2</sub> (Haghpanah et al, 2013). Both factors together mean that several days of CPU time are required to design a single adsorbent material/process cycle PVSA configuration. Additionally, the status-quo approach for design of such processes does not account for the flexibility of the optimal process design. Assessment of the flexibility of a process is an underpinning element to understanding its controllability, which could prove to be an important factor in discrimination between competing PVSA configurations. Particularly, assessment of new configurations based on flexibility could improve robustness of process operations to disturbances in the flue gas feed conditions.

In this work, we present a framework for simultaneous design and flexibility assessment of PVSA processes applied to post-combustion CO<sub>2</sub> capture. The approach utilizes a high-fidelity dynamic process model as a virtual experimentation platform. In conjunction with quasi-random sampling, an artificial neural network (ANN) model, and an alpha shapes approach, we use the virtual experimentation environment to efficiently identify an initial set of Pareto fronts for the process, the design space of operating parameters where all the process constraints are satisfied and quantify the flexibility around the nominal operating point (NOP). The proposed methodology is both significantly more efficient than the status-quo design approach for PVSA presented in the literature and additionally generates a much richer set of outputs to inform the design and control of such systems.

## 2. The system

We consider a four-step PVSA process with feed pressurization for post-combustion CO<sub>2</sub> capture from a dry coal fired power plant flue gas using a fixed bed of zeolite 13X adsorbent. The feed gas is a binary mixture of CO<sub>2</sub>/N<sub>2</sub> with a molar composition of 15/85% which is available at ambient conditions. Regulatory requirements constrain the operation of the process to recover greater than 90% of the CO<sub>2</sub> fed to the process at a product purity of greater than 95%. The process has been described using a mathematical process model developed in a previous study, which is taken as a virtual experimentation platform for the purposes of this work (Ward & Pini, 2022a). Briefly, the model comprises a coupled system of partial differential and algebraic equations (PDAEs) which represent the governing balances of mass, momentum, and energy in the adsorption column. The PDAE system is solved in MATLAB using a finite volume scheme. The model equations are solved subject to a set of time-varying boundary conditions which describe the four-step adsorption cycle. The mathematical model has been validated against both experimental measurements (Ward & Pini, 2022b) and independent simulations conducted in the literature (Ward & Pini, 2022a). Design of such systems is cumbersome because simulation of the process performance for a given set of operating parameters takes up to several minutes of computational time.

As inputs, the process model requires specification of the operating parameters of the PVSA cycle. These are the intermediate operating pressure achieved in the blowdown step ( $p_I$ ), as well as the high operating pressure ( $p_H$ ) and the flow rate ( $v_F$ ) at which flue gas is fed to the bed during the adsorption step. For a given set of operating conditions, the process performance is simulated and a set of five key performance indicators (KPIs) are calculated. We evaluate the purity and recovery of the CO<sub>2</sub> extracted from the flue

gas to allow us to assess compliance with the regulatory constraints on the operation of the process. We also evaluate the productivity of the process and its specific energy usage to assess the efficiency of the process. The productivity specifies the capture rate of CO<sub>2</sub> per unit volume of the bed and the specific energy usage specifies the amount of energy consumed per tonne of CO<sub>2</sub> captured. Additionally, the dynamic process simulator is also coupled to a detailed techno-economic assessment framework. This allows for calculation of the capture cost associated with each set of operating parameters, given as the cost per tonne of CO<sub>2</sub> captured.

The design of PVSA processes for post-combustion CO<sub>2</sub> capture is classically conducted in three stages. First, the purity and recovery of CO<sub>2</sub> extracted from the flue gas is maximized. This allows for confirmation that the chosen process configuration has a design space for which the regulatory requirements on the purity/recovery of the extracted CO<sub>2</sub> are satisfied. Second, the productivity of the process is maximized while minimizing the specific energy usage, subject to the above regulatory constraints. In this step, we assess the efficiency of the process for capturing CO<sub>2</sub> from the flue gas. Third, a capture cost minimization is conducted subject to the purity/recovery constraints. The cost minimum point is usually taken to be the nominal operating point (NOP) of the process. The design of these systems is challenging because it requires constrained multi-objective optimization coupled with a detailed process simulation and economic assessment framework. The new approach presented in this study aims to significantly streamline the design by conducting all three design stages simultaneously, whilst also explicitly obtaining the design space and quantifying the flexibility of the NOP.

### 3. The framework

The framework has three main steps: problem formulation, design space identification, and design space analysis.

#### 3.1. Problem Formulation

The objective of this case study is to identify good candidate process operating points and assess their flexibility. As detailed in Section 2, the manipulated variables of the considered PVSA system are  $p_H$ ,  $p_I$  and  $v_F$ . The bounds of the manipulated variables are provided in Table 1 and were chosen based on good practice in the literature on PVSA process design. The KPIs of interest are the purity and recovery of CO<sub>2</sub> extracted from the flue gas, as well as specific energy consumption, productivity, and capture cost. The process must satisfy regulatory constraints on the CO<sub>2</sub> product stream with a purity > 95% and a recovery > 90%. Additionally, the operating conditions are constrained such that the intermediate pressure is always less than high pressure ( $p_I \leq p_H$ ).

Table 1. Bounds of the manipulated variables.

	$p_I$ (bar)	$p_H$ (bar)	$v_F$ (m/s)
Lower bound	0.05	1	0.1
Upper bound	3.00	10	2.0

#### 3.2. Design Space Identification

The quasi-random Sobol sequence is used to generate 4,096 parameter combinations within the parameter bounds. Owing to the constraint imposed on the operating pressures, only 3,458 of the sampled parameter combinations were feasible for use in virtual experimentation. The process model was evaluated at the operating point corresponding to each parameter combination to generate a dataset. Using this data, an artificial neural

network is trained, and an additional 260,000 samples are generated. This allows for the characterization of the design space with a fine alpha shape.

### 3.3. Design Space Analysis

Based on the model outputs for each of the Sobol sampled parameter combinations, we select the nominal operating point (NOP) for the process as that which minimizes the capture cost. Starting from the NOP, a uniform cube with respect to the parameter axes is formed and expanded until one of its vertices intersects the design space boundary. The resulting region is the uniform acceptable operating region (AOR). This allows for the extraction of acceptable operating ranges of the different parameters with respect to the NOP used, which is an underpinning element for analyzing process controllability.

### 3.4. Optimality-Related Analysis

In addition to the systematic design space and flexibility analysis, the framework also offers information regarding good candidate operating conditions and Pareto fronts. For benchmarking, we also perform rigorous optimization using the non-dominated sorting genetic algorithm II (NSGA-II). We solve to obtain purity/recovery and productivity/energy Pareto fronts, as well as a point of minimum capture cost. The productivity/energy and capture cost optimization problems are constrained by the purity/recovery requirements using a penalty function approach. For each problem, we use a population size of 72 with a maximum of 70 generations. Therefore, the solution of the three design problems requires  $72 \times 70 \times 3 = 15,120$  forward simulations by rigorous optimization, as compared to only 3,458 simulations by the proposed framework.

## 4. Results and Discussion

First, before performing design space and flexibility analysis, the dataset generated by virtual experimentation with the Sobol sequence can be used to obtain a first approximation of the Pareto fronts for the design problem. To this end, purity/recovery and productivity/energy Pareto fronts, as well as a cost-optimal point are identified and compared with the outputs of rigorous optimization using NSGA-II. Second, we present the design space analysis and associated flexibility analysis of the cost-optimal point.

### 4.1. Optimality-Related Analysis

The comparison between the Sobol sampled data points and the Pareto front obtained from the rigorous optimization with NSGA-II is shown in Figure 1. The boundary outlined by the data points obtained from the Sobol sampling shows strong agreement with the Pareto fronts obtained from the rigorous optimization using NSGA-II. Further, the cost-optimal point obtained from Sobol sampling is in very good agreement with the

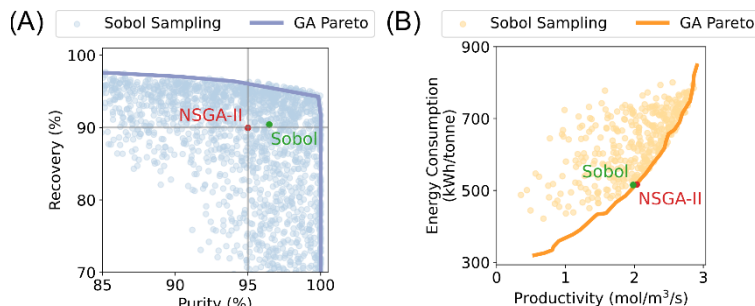


Figure 1. Pareto front comparison (A: purity-recovery and B: energy-productivity), solid lines: genetic algorithm, points: Sobol sampling. Red point is the cost-optimal point of NSGA-II and green is that of the Sobol sampling.

solution obtained using NSGA-II, as detailed in Table 2. The cost optimal point obtained from NSGA-II has a capture cost which is 1.1% lower than that obtained by Sobol sampling. However, solution of the set of design problems using NSGA-II requires greater than 4x the computational time to obtain the optimal process design as compared to Sobol sampling. This highlights a key advantage of the proposed framework, which is that it is effective at identifying good candidate process performance while requiring significantly less computational effort. In practical terms, the good candidate points obtained by Sobol sampling could be used to significantly tighten the parameter bounds before applying rigorous optimization to the system, which can reduce the burden of formal optimization quite significantly.

Table 2. Cost-optimal solution from NSGA-II and Sobol sampling.

	NSGA-II	Sobol sampling
$p_H$ (bar)	3.56	3.55
$p_L$ (bar)	1.72	1.25
$v_F$ (m/s)	1.44	1.40
Capture cost (\$/tonne)	61.67	62.35

#### 4.2. Design Space Identification & Analysis

The design space is identified and shown in Figure 2. The samples which violate the constraints are hidden to improve the clarity of the figure. We can see that at feed velocities greater than 1.5 m/s there are no samples which satisfied the purity and recovery constraints. Starting at low velocities, the design space is quite narrow. As the velocity increases, the size of the design space increases. However, after reaching a certain value, the design space reduces in size until the process becomes infeasible at higher velocities. The same can be seen when considering the other two manipulated variables. This is because there are trade-offs between achieving higher recovery or higher purity. Due to the non-linearity of the phenomena and highly coupled parameters affecting the process, it remains challenging to identify general heuristics for design. The proposed framework highlights these trade-offs through a systematic, data-supported methodology.

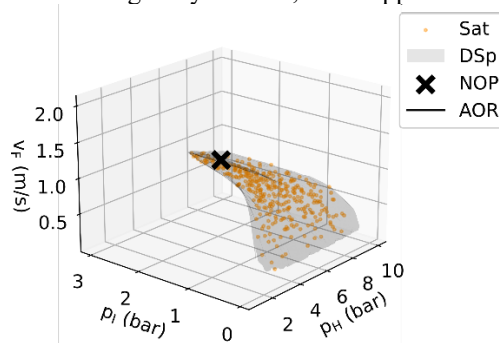


Figure 2. Identification of the optimal point acceptable operating region. Flexibility is very low therefore the AOR is not visible in the plot.

In Figure 2, the NOP used is the cost-optimal point obtained from Sobol sampling. The flexibility of the point is examined and an AOR is defined. Table 3 provides quantitative metrics from analysis of the design space. We can see that the process is non-flexible as only very small variabilities in the operating conditions are permissible without violation of the constraints. When the point of minimum capture cost is chosen to be the NOP the purity/recovery constraints are active at the NOP, meaning that it lies near the edge of the

design space and the process has very poor flexibility. Therefore, in practice, the process would be challenging to control while maintaining operation in compliance with regulatory constraints on the CO<sub>2</sub> product stream extracted from the flue gas. This insight remains unaccounted for in the adsorption process design literature, and we contend based on this case study that this is a central point of importance for the practical application of adsorption-based separations for CO<sub>2</sub> capture.

Table 3. Design space analysis metrics.

	Value
DSp size (bar <sup>2</sup> m/s)	4.25
Minimum capture cost (\$/tonne)	62.35
Average capture cost (\$/tonne)	76.41
Maximum capture cost (\$/tonne)	302.14
AOR size (bar <sup>2</sup> m/s)	$8.6 \times 10^{-5}$
Acceptable range $p_H$ (bar)	$3.55 \pm 0.054$
Acceptable range $p_I$ (bar)	$1.25 \pm 0.018$
Acceptable range $v_F$ (m/s)	$1.40 \pm 0.011$

## 5. Conclusion

In this work, a framework for flexible process design using design space identification has been presented and applied to an adsorption-based post-combustion CO<sub>2</sub> capture process. The proposed framework facilitates integrated analysis to quantify flexibility metrics with respect to any of the process KPIs and provide good candidate operating conditions. Acceptable operating regions and ranges of operating parameters based on any nominal point can be quantified. We used the framework to identify a good candidate operating point with a very similar capture cost to that obtained by rigorous optimization while requiring 77% fewer forward model evaluations. We show that the status-quo design approach of minimizing capture cost leads to a non-flexible design which would be challenging to operate within regulatory requirements on the CO<sub>2</sub> product stream. Herein, we provide a framework for efficiently evaluating process flexibility of any NOP which can therefore be used to assess new design approaches which account for flexibility from an early stage. Thereby, we will ultimately be able to design more highly controllable PVSA CO<sub>2</sub> capture processes and accelerate the propagation of this technology from research into industry. Future work on this framework includes the exploitation of the Sobol sequence for the identification of better bounds for rigorous optimization.

## References

- Riganati et al. (2021): "Adsorption of carbon dioxide for post-combustion capture: A review". *Energy & Fuels* (35).
- Balashankar et al. (2019): "Analysis of a batch adsorber analogue for rapid screening of adsorbents for postcombustion CO<sub>2</sub> capture". *Industrial & Engineering Chemistry Research* (58).
- Haghpanah et al. (2013): "Multiobjective optimization of a four-step adsorption process for post combustion CO<sub>2</sub> capture via finite volume simulation". *Industrial & Engineering Chemistry Research* (52)
- Ward & Pini (2022a): "Efficient bayesian optimization of industrial-scale pressure-vacuum swing adsorption processes for CO<sub>2</sub> capture". *Industrial & Engineering Chemistry Research* (61).
- Ward & Pini (2022b): "Integrated uncertainty quantification and sensitivity analysis of single-component dynamic column breakthrough experiments". *Adsorption* (28).

# Multiple-Model Fault Detection of PEMFC

A.Allam<sup>a</sup>, M.Mangold<sup>b</sup>, P.Zhang<sup>a</sup>

<sup>a</sup> *Institute of Automatic Control, Technische Universität Kaiserslautern,  
aallam@rhrk.uni-kl.de, pzhang@eit.uni-kl.de, Kaiserslautern, 67653, Germany*

<sup>b</sup> *Bingen University of Applied Sciences, m.mangold@th-bingen.de, Bingen 55411,  
Germany*

## Abstract

Catalytic degradation and flooding are classified as critical cell failures that significantly impact the durability of proton exchange membrane fuel cells (PEMFC). Delayed detection of both failures can have detrimental effects on the fuel cell. The purpose of this work is to achieve early fault detection by integrating a fault diagnosis system onto a PEMFC degradation model. The framework of the diagnosis system is based on the concept of multiple-model fault detection with a main task of selecting the underlying model which most likely represents the current operating condition of the system. Primarily, optimal state estimates are computed using the Square Root Unscented Kalman filter. Then fault indication signals are generated by the Bayes' Rule updated with conditional probabilities propagated by each filter model. Finally, a threshold function is derived based on the fault indication signals. Simulation results demonstrate that the proposed strategy successfully achieves fault detection and isolation.

**Keywords:** PEMFC, Fault Diagnosis, Multiple Model, Catalytic Degradation, Flooding

## 1. Introduction

Favorable characteristics of Proton Exchange membrane fuel cells (PEMFC), such as low operating temperature and compact design, enable them as an alternative solution in the electrification of the transportation sector. However, PEMFCs exhibits performance limitations when implemented as electric propulsion systems. A main contributing factor is load cycling which are rapid transient operating conditions that are unsustainable by the slow internal dynamical behavior of the fuel cell. Experiments have demonstrated that load cycling is a root cause of accelerated cell deterioration (Steiner, 2009).

Flooding and induced catalytic degradation are both symptoms of improper management of load cycling. Extreme conditions of both failures disrupt the chemical stability of the fuel cell and can lead to the loss of electrochemically active surface area (ECSA). Therefore, integrating an effective diagnosis system is key to enhance cell durability and to maintain optimal stack performance.

Several studies have considered the task of fault diagnosis in PEMFCs. In (Aitouche, 2011; Escobet, 2008; Lira, 2012) model-based approaches are adopted to detect flooding and catalytic degradation. Meanwhile, Hernandez (2010) constructs an equivalent circuit model and achieves fault detection through parameter identification of critical system parameters. In (Ibrahim, 2015; Zheng, 2014) fault diagnosis of flooding is accomplished by analyzing experimental data using the Discrete Wavelet Transform and a Clustering double-fuzzy approach respectively.



Motivated by recent developments in fault diagnosis (Dinge, 2013), in this paper the diagnosis of flooding and catalytic degradation in PEMFCs is investigated. Primarily, a PEMFC multiscale catalytic degradation model is developed based on the radii of platinum particles and the surficial coverage of platinum oxide. Principally, in multiple-model fault detection the main task of the diagnosis system is to accurately distinguish the correct model (nominal or faulty behavior) which most likely represents the current operating condition of the system. At first, optimal state estimates of a bank of models under different assumptions are obtained using the Square Root Unscented Kalman filter (SRUKF) given in (Merwe, 2001). Then the Bayes' Rule is implemented to generate fault indication signals based on the conditional posteriori probabilities of each model. A decision logic is then derived to determine the current operating condition of the system.

The paper is organized as follows. In section 2 a 5-state PEMFC multiscale catalytic degradation model is presented. The framework of the diagnosis system is developed in Section 3. Finally, simulation results are demonstrated in Section 4 to evaluate the performance of the diagnosis system under the events of flooding and catalytic degradation.

## 2. PEMFC Multiscale Degradation Model

Testing the developed diagnosis system requires reference data from specific failures. As it is quite hard to induce desired types of failures to a real system, simulated measurement data are used in this work, which are generated by a reference PEMFC model. The model consists of a macroscopic part derived from macroscopic mass and charge balances and of a microscopic kinetic model accounting for changes in the morphology of the catalytic structure.

Without loss of generality, the macroscopic part of model is based on the following assumptions: 1) Gases behave like ideal gases. 2) PEMFC is an isothermal system 3) The anodic reaction is fast enough and is not considered as a rate determining step. The macroscopic model has three states: the molar concentration of oxygen in gas flow channels ( $x_1$ ) and in the catalyst layer ( $x_2$ ) and the cathodic activation losses ( $x_3$ ). The model is a multiple-input multiple-output (MIMO) system with three inputs, the current ( $u_1$ ), the inlet air flow ( $u_2$ ) and the partial molar concentration of oxygen in the supplied air ( $u_3$ ). The two measurements are the molar concentration of oxygen in the catalyst layer ( $y_1$ ) and the cell voltage ( $y_2$ ).

From mass conservation principles, the molar concentrations of oxygen in the gas flow channel is computed by,

$$\dot{x}_1 = \frac{1}{V_C} (u_2(u_3 - x_2) - A_E \beta (x_1 - x_2)) \quad (1)$$

where  $V_C$  is the cathode bulk volume,  $A_E$  is the area of mass exchange and  $\beta$  is the mass transfer coefficient of the gas diffusion layer. The molar concentration of oxygen in the catalyst layer is derived from

$$\dot{x}_2 = \frac{1}{V_{CL}} \left( A_E \beta (x_1 - x_2) - k_{CD} \frac{i_{c,o} A_E}{4F} \left( 1 - e^{-\frac{-2Fx_3}{RT}} \frac{x_2}{u_3} \right) \right), \quad (2)$$

where  $V_{CL}$  is the volume of the catalyst layer,  $i_{c,o}$  is the exchange current density,  $F$  is the Faraday constant,  $R$  is the universal gas constant and  $T$  is stack temperature. The final state is the cathodic activation loss  $x_3$ , defined as the energy required to overcome the ionic energy barrier to initiate the electrochemical reaction

$$\dot{x}_3 = \frac{1}{\phi_C} \left( -u_1 - k_{CD} \frac{i_{c,o} A_E}{F} \left( 1 - e^{-\frac{-2Fx_3}{RT}} \frac{x_2}{u_3} \right) \right) \quad (3)$$

The microscopic model is motivated by the work in (Darling, 2003). An assumption imposed on the microscopic model characterizes the catalyst layer as finely distributed uniformly sized spherically shaped carbon-supported platinum (Pt) particles. Suboptimal operating conditions in the catalyst layer induce morphological changes to the Pt particles and imminently lead to loss of ECSA. The proposed model consists of two states: PtO coverage on the spherical Pt particle ( $x_4$ ) and the Pt particle radius ( $x_5$ ), which are computed from

$$\dot{x}_4 = \left( \frac{r_2 - r_3}{\Gamma_{max}} \right) - \left( \frac{2x_4}{x_5} \right) \dot{x}_5 \tag{4}$$

$$\dot{x}_5 = -\frac{M}{\rho}(r_1 - r_2), \tag{5}$$

where  $r_1$  and  $r_2$  expresses the rate of the electrochemical reactions of Pt ionization and PtO formation respectively and  $r_3$  is the rate of the overall electrochemical reaction. Due to limitation of space the electrochemical reactions of Pt dissolution and PtO formation ( $r_1, r_2$  and  $r_3$ ) described by the Butler Volmer kinetics equation are not covered. Both sub-models are connected by the degradation rate  $k_{cd}$  which is defined as the ratio between current particle radius  $x_5$  and initial radius  $x_{5,0}$ .

Finally, the two system measurements are the molar concentration of oxygen in the catalyst layer  $y_1 = x_2$  and the cell voltage. The latter is calculated by subtracting the activation losses  $x_3$  and the ohmic resistance losses  $R_{ohm}$  from the open circuit voltage  $U_0$ ,

$$y_2 = U_0 + x_3 - R_{ohm}u_1 \tag{6}$$

### 3. Fault Diagnosis Approach

#### 3.1. Fault Analysis and Modelling

Determining the correct system parameters to model the multiplicative faults of flooding and catalytic degradation is essential for the proposed diagnosis system. In literature various studies investigate the impact of both failures on the transient behavior

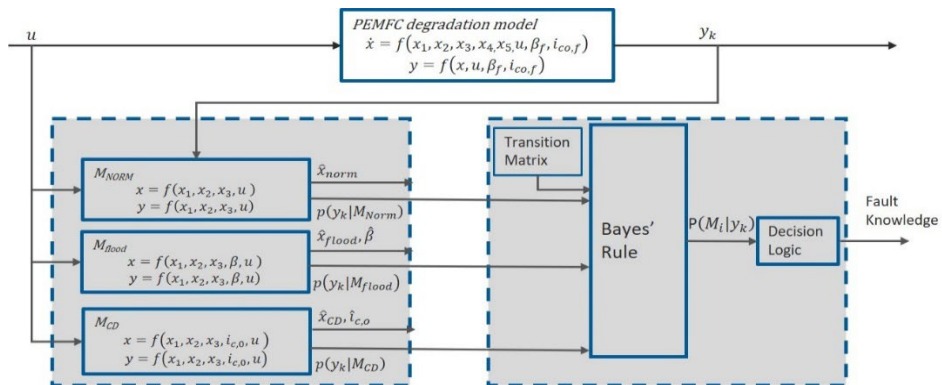


Fig 1. Schematic description of FD system

of PEMFCs. It is determined that early stages of flooding result in the accumulation of water within the porous structure of the gas diffusion layer compromising its diffusive

properties (Grotsch, 2008). Accordingly, the parameter chosen to model flooding is  $\beta$  which influences the mass flux of oxygen within the stack. The asserted assumption is that the cell loses 50% of its diffusive properties defined as  $\beta_f$  as illustrated in Fig.1.

Meanwhile, a qualitative analysis was conducted by Steiner et al (2008) to identify the various mechanisms of catalytic degradation. Severe conditions of the failure induce a non-uniform current distribution which is attributed to the exchange current density  $i_{c,o}$ . The fault is provoked by reducing  $i_{c,o}$  from its original and is assigned as  $i_{c,o,f}$ .

### 3.2. Fault Diagnosis Scheme

In this subsection the proposed fault diagnosis system will be developed. Unlike in (Blackmore, 2006) where a set of finite models are differentiated by their operating state (normal or faulty conditions), the presented diagnosis system is based on a bank of models that are distinguished by a desired system parameter  $\gamma$  to be estimated, i.e.  $\gamma = \beta$  or  $\gamma = i_{c,o}$ . The system parameter  $\gamma_i$  is derived from the analysis conducted on the multiplicative fault and is described as an additional state in the system model with zero dynamics i.e.  $\dot{\gamma}_i = 0$ . Consider that the system dynamics of a bank of models  $M_i$  are described by,

$$\begin{aligned} M_i: \dot{x}^i(t) &= f_i(x^i(t), u(t), \gamma_i) + w^i(t) \\ y_k^i &= h(x^i(t_k), u(t_k), \gamma_i) + v_k^i \end{aligned} \quad (7)$$

where  $i=0$  represents the normal macroscopic three-state model with  $\gamma_i=0$  (no additional state or system parameter to be estimated),  $i=1,2,\dots,\alpha$  represents  $i$ -th multiplicative fault described by the  $\gamma_i$  system parameter to be estimated,  $x^i(t)$ ,  $u(t)$ ,  $y_k^i$ ,  $w^i(t)$ ,  $v_k^i$  are respectively, the state vector, the control input vector, the output vector, the process noise and the measurement noise when the system incorporates the  $i$ -th multiplicative fault. As shown in Fig.1,  $M_0=M_{norm}$  represents a model with nominal system parameters with no additional state.  $M_1=M_{flood}$  is a model that incorporates flooding through the additional system state  $\beta$ .  $M_2=M_{CD}$  describes catalytic degradation which does not define a specific mechanism but is manifested as a deprecation of the exchange current density  $i_{c,o}$ .

The proposed diagnosis strategy is composed of two parts as mapped in Fig 1. Primarily, the SRUKF is adopted to compute optimal state estimates of each model  $M_i$  and to predict the probability density function of the output conditioned by each model

$$p(y_k | M_i) = \frac{1}{(2\pi)^{\frac{\alpha}{2}} |S_{yy_i}|} \exp\left(-\frac{(y_k - \bar{y}_k^i)^T S_{yy_i}^{-1} (y_k - \bar{y}_k^i)}{2}\right), \quad (8)$$

where  $\alpha$  is the number of models,  $\bar{y}_k^i$  is the estimated output mean and  $S_{yy_i}$  is output estimation error covariance matrix of the  $i$ -th SRUKF of  $M_i$ . Secondly, the Bayes' Rule is implemented to update the propagated prior information and to compute the conditional posteriori probability that  $M_i$  is true given the current measurement  $y_k$ ,

$$P(M_i | y_k) = \frac{p(y_k | M_i)P(M_i | y_{k-1})}{\sum_{j=1}^m p(y_k | M_j)P(M_j | y_{k-1})}, i = 1, \dots, m \quad (9)$$

A probability transition matrix is introduced to define the likelihood of the different models to transition among each other. Finally, to classify each failure the decision logic is derived as,

$$i = \arg \max_i P(M_i | y_k) \text{ and } P(M_i | y_k) > J_{th} \quad (10)$$

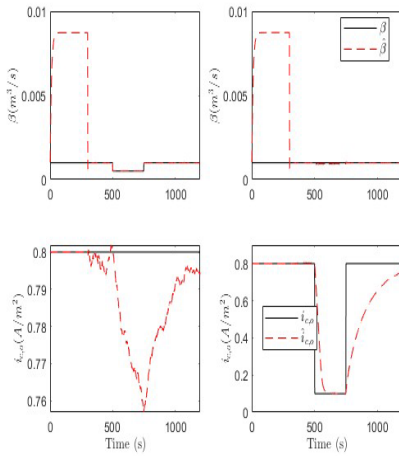


Fig 2. Parameter estimation provided by  $M_{flood}$  (top figures) and  $M_{CD}$  (bottom figures) under flooding (left) and catalytic degradation (right)

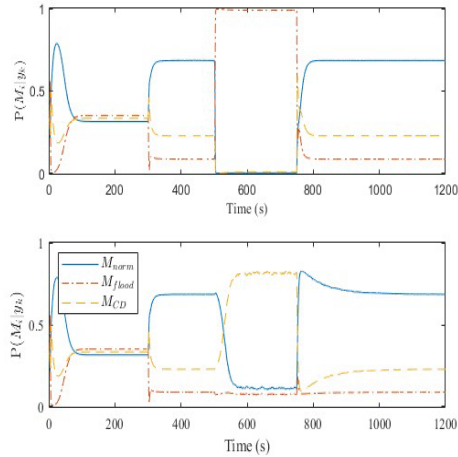


Fig 3. Posteriori Probabilities for the case of flooding (top) and catalytic degradation (bottom)

where  $J_{th}$  is a threshold introduced to increase the reliability of the diagnosis system. Specifically, satisfying both conditions indicates that the PEMFC system is operating under the  $i$ -th condition. In the presented case  $J_{th} = 0.7$ .

### 4. Simulation Results

In this section the validity of the developed fault diagnosis system is evaluated in the MATLAB environment. Accurate fault detection of the proposed diagnosis system significantly relies on correct parameter estimation ( $\beta, i_{c,o}$ ) provided by the SRUKF. The initial conditions of both system parameters are assumed to be equal to the nominal values.

Figure 2 plots the system parameters  $\beta$  and  $i_{c,o}$  and their respective estimates based on  $M_{flood}$  and  $M_{CD}$  in response to a step change input current from 0 to 1 A at 300s, a failure event introduced at 500s and cell recovery at 750s. Specifically, the left plots illustrate the response of both models to a flooding event between 500s and 750s. During the first 300s, the fuel cell is in idle conditions and due to the electrochemical inactivity the SRUKF diverges and fails to estimate  $\beta$ . However, after the introduction of the unity step input current the SRUKF identifies the system parameter and converges to its true value. At 500s, it can be observed that  $M_{flood}$  reacts to the failure event and correctly estimates  $\beta$ . Meanwhile, the SRUKF based on  $M_{CD}$  weakly recognizes the failure and a slight decrease in  $\hat{i}_{c,o}$  can be observed. A similar behavior is also exhibited by both models under the event of catalytic degradation (right plots of Fig.2). The SRUKF of  $M_{CD}$  reacts to the failure event of catalytic degradation and correctly estimating  $i_{c,o}$ . However, the filter model based on  $M_{flood}$  weakly responds to the induced failure event.

The diagnosis of the failure events is demonstrated in Fig 3. using the posteriori probabilities given by (9) and the decision logic derived in (10). For the first 300s the stack is in idle conditions and the three models share a relatively similar probability. As current is introduced at 300s the fuel cell is electrochemically activated and the diagnosis system is able to distinguish between the models with  $P(M_{norm} | y_k) > 0.7$  correctly indicating that the system is operating under fault-free conditions. The top plot demonstrates that flooding was successfully detected and isolated at 505s with

$P(M_{flood} | y_k) > 0.7$  and a detection delay of 5s. Meanwhile, the bottom plot indicates that catalytic degradation was detected at 553s with a detection delay of 53s.

A challenging task when implementing the proposed diagnosis scheme is maintaining the stability of the multiple Kalman filter models. Under certain failure events often the Kalman filter exhibits unstable dynamics and diverge. For instance, under catalytic degradation the SRUKF based on  $M_{flood}$  would diverge while estimating  $\beta$ . Two solutions were implemented to overcome this problem. Firstly, by tuning the parameters of the Kalman filter (system and measurement noise) according to the transient impact of the failure on the fuel cell system. Secondly, a constrained SRUKF is constructed based on bounding the additional state within a feasible range. Imposed by these constraints a new Kalman gain is computed by solving an optimization problem based on the unconstrained estimated states and covariance matrices using *fmincon* on MATLAB. Then an updated constrained state space is constructed using the new Kalman gain.

## 5. Conclusion

In this paper, a fault diagnosis scheme based on multiple-model fault detection is developed to classify two major PEMFC failures, namely flooding and catalytic degradation. Models are distinguished by an additional system state which represents a predefined system parameter corresponding to the multiplicative fault under consideration. Therefore, the effectivity of the diagnosis system significantly relies on the correct state estimation provided by the SRUKF. The Bayes' Rule is then implemented to generate fault indication signals to achieve fault detection and isolation.

## References

- A. Aitouche, Q. Yang, and B. O. Bouamama, 2011, Fault detection and solution of PEM fuel cell system based on nonlinear analytical redundancy application via parity space approach, *The European Physical Journal Applied Physics*, 54, 2, 23408.
- T. Escobet et al., 2009, Model-based fault diagnosis in PEM fuel cell systems, *Journal of Power Sources*, 192, 1, 216–223.
- A. Hernandez, D. Hissel, and R. Outbib, 2006, Fuel cell fault diagnosis: A stochastic approach, in *Proc. IEEE Int. Symposium on Industrial Electronics*, 1984–1989.
- M. Ibrahim et al. 2015, Signal-based diagnostics by wavelet transform for proton exchange membrane fuel cell, *Energy Procedia*, 74, 1508-1516.
- Z. Zheng, et al., 2014, A double-fuzzy diagnostic methodology dedicated to online fault diagnosis of proton exchange membrane fuel cell stacks. *Journal of Power Sources*, 271, 570-581.
- S. X. Ding, 2013, *Model-based fault diagnosis techniques: design schemes, algorithms, and tools*. Springer.
- R. Van Der Merwe and E. A. Wan, 2001, The square-root unscented kalman filter for state and parameter-estimation, in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, 3461–3464.
- R. Darling and J. Meyers, 2003, Kinetic model of platinum dissolution in PEMFCs. *Journal of the Electrochemical Society*, 150, 11, A1523
- M. Groetsch, R. Hanke-Rauschenbach, and M. Mangold, 2008, Bifurcation analysis of a two-phase PEMFC model, *Journal of Fuel Cell Science and Technology*, 5, 2.
- N. Yousfi-Steiner et al., 2009, A review on polymer electrolyte membrane fuel cell catalyst degradation and starvation issues: Causes, consequences and diagnostic mitigation, *Journal of Power Sources*, 194, 1, 130–145.
- L. Blackmore and B. Williams, 2006, Finite horizon control design for optimal discrimination between several models, *Proceedings of the IEEE Conference on Decision and Control*, 45, 1147-1152.

# Implications of Heat and Work Integration in Biomass to Fischer-Tropsch Fuels: A simulation Approach

Samukeliso Dube\*, Baraka Celestin Sempuga, Xinying Liu

*Institute for Development of Energy for African Sustainability (IDEAS), University of South Africa, Private Bag X6, Florida, Johannesburg, South Africa*

## Abstract

The sustainability of biomass and biogas use relies on the efficiency of conversion processes and produced products. This work determines specific process targets that have the highest efficiency in terms of material, energy, and work balances. Many authors have suggested biomass and biogas to be sustainable sources of hydrogen. However, the analysis in this work shows that conversion of these to hydrogen has fundamental drawbacks because of the inevitable CO<sub>2</sub> emissions and zero carbon efficiency. Co-production of hydrogen with methanol and/or Fischer-Tropsch fuels can significantly increase the efficiency but maximizing the carbon, energy and work efficiencies requires zero hydrogen production, with the potential to achieve net CO<sub>2</sub> consumption. There is need to rigorously integrate heat and work energies between the gasification/reformer and synthesis section. We propose the use of heat engine-heat pump configuration with a theoretical Carnot efficiency of 0.43. We also explore using solar energy to maximize CO<sub>2</sub> consumption.

**Keywords:** Biomass, Biogas, Fischer Tropsch synthesis, Targeting techniques, Process synthesis.

## 1. Introduction

Solid and gas bio-energy resources, are relatively evenly distributed across the globe and have been identified as renewable resources from which hydrogen and hydrocarbons can be produced sustainably. However, hydrogen production from biomass can be considered inefficient because of the CO<sub>2</sub> emissions which, from a material balance perspective, are a result of feeding material containing carbon, hydrogen and oxygen to only make hydrogen. Consequently, CO<sub>2</sub> is the preferred product to get rid of carbon and oxygen. Lepage et al. (2021) reviewed biomass-to-hydrogen production routes and concluded that biomass conversions need improvements and adaptations to be implemented at a larger scale (Lepage et al., 2021). Therefore, one needs to understand how to maximize the potential of biomass by utilizing all its components namely, carbon, hydrogen, and oxygen, and minimize CO<sub>2</sub> emissions. A system's approach, adopted in this work, can give a better understanding of the process options and assist in setting specific performance targets which will guide the development and design of the process. Performance targets may be understood as theoretical process performance in terms of material, energy, and work balances, indicating the best feasible options against which a process can be measured or as an indication of opportunities for improvement. Patel et al.

(Patel et al., 2005) pointed out that a high level mass and energy flow analysis across a process is a pre-cursor to designing processes that are cost- and energy- efficient. When dealing with carbonaceous feedstocks, it is imperative to consider the H/C and O/C ratios, both in the feed and the product streams; this has an impact on the hydrogen and carbon efficiencies (Liu & Patel, 2017) and will also impact the energy and work flows across the process. Therefore, in this work, we co-feed biomass with biogas as an alternative renewable source of hydrogen and carbon elements which allows varying their ratios in the feed and also provides more degrees of freedom and a much bigger region where we can set the optimisation targets.

## 2. Methodology

The high-level analysis of a process starts by developing its material balance attainable region. To develop the material balance attainable region of the biomass to hydrogen and hydrocarbons process we start by the overall material balance that includes all the components we have identified.

$$aCH_{1.4}O_{0.6} + bCO_2 + cCH_4 + dH_2O + eO_2 + fH_2 + gC + hCO + iCH_3OH + kCH_2 = 0 \quad (1)$$

From this, we deduce the atomic species balances with three independent equations and 10 variables; thus, the system has 7 degrees of freedom. However, too many degrees of freedom make it difficult identifying optimal targets as it leads to searching within a multidimensional region. To systematically narrow down the attainable region into sub-regions of interest we only specify a limited number of variables representing the main targets and then use an optimization method to determine the rest of the variables to meet a specific objective function. Thus, in this method, the objective function and the constraints will represent the main targets for the process and then the optimization variables will be solved to meet these targets. We can express an optimization problem where we vary the ratio of biomass to biogas in the feed as follows:

for $R_{B/BG} = [0: 1]$ ; $max N_{CH_2}$ subject to	$\left\{ \begin{array}{l} A \cdot N_d^T \leq 0 \\ A_{eq} \cdot N^T = b_{eq} \\ \sum N_i \Delta H_{T_o} \leq 0 \\ \sum N_i \Delta G_{T_o} \leq 0 \\ CO = 0 \end{array} \right.$	$(2)$
---	--	-------

Where:

$R_{B/BG}$ , is the ratio of biomass to biogas in the feed,  $A \cdot N_d^T \leq 0$  is the inequality constraint;  $A_{eq} \cdot N^T = b_{eq}$  is the material balance constraints in equations;  $\sum N_i \Delta H_{T_o} \leq 0$  and  $\sum N_i \Delta G_{T_o} \leq 0$  are the energy and work constraints, respectively.

We determined the extent to which a non-idealised process deviates from the optimal targets by simulating the process in Aspen Plus® and the flowsheet is shown in Figure 1 below.

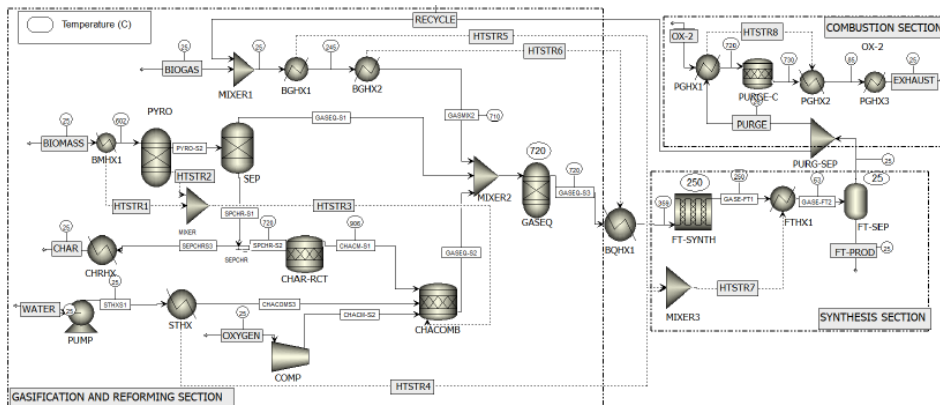


Figure 1: FT-synthesis from a mixture of biomass and biogas feedstock: Aspen Plus Simulation

### 3. Results and Discussions

#### 3.1 MATLAB Results

We targeted maximising the hydrogen and carbon efficiencies of the system by matching the H/C and the O/C ratios in the feed to the H/C and O/C ratios in the valuable product streams, respectively. These produce a broad product spectrum. Matching the H/C ratios produced  $\text{CH}_3\text{OH}$ , char,  $\text{H}_2\text{O}$  (when the  $\text{CH}_4$  fraction  $< 0.4$ ),  $\text{H}_2$  and  $\text{CH}_2$  products. Significant amounts of  $\text{O}_2$  and  $\text{H}_2\text{O}$  are consumed and  $\text{CO}_2$  is consumed from a  $\text{CH}_4$  fraction of 0.2. This produces less amounts of  $\text{CH}_2$ , but favours  $\text{CH}_3\text{OH}$  production. The amount of  $\text{CH}_3\text{OH}$  produced increases with an increase in the  $\text{CH}_4$  fraction, this is due to the increase in the H/C ratio in the feed. At a  $\text{CH}_4$  fraction of 0.4, only  $\text{CH}_3\text{OH}$  and char are produced and this is the best achievable target in terms of mass, energy and work efficiencies; results are shown in Figure 2a. Furthermore, both the hydrogen and carbon efficiencies are 100% at this point. Most of the chemical potential potential of the system is conserved ( $-10 \text{ kJ/mol}$ ) and the enthalpy change is zero, hence no energy is lost to or required from the surroundings as shown in Figure 2b.

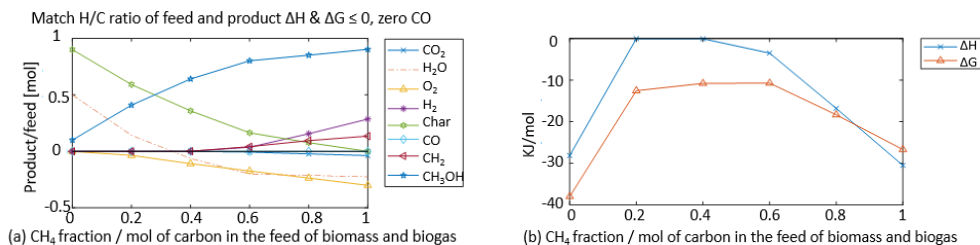


Figure 2: (a) Material targets and (b) energy and work targets that match the H/C ratio in the feed to the H/C ratio in the products

Figure 3a shows that matching the O/C ratio in the feed to the O/C ratio in the valuable product streams favours the production of, mainly Fischer-Tropsch hydrocarbons ( $-\text{CH}_2$ ),  $\text{H}_2\text{O}$  and char.  $\text{CH}_3\text{OH}$  and  $\text{H}_2$  start being produced at a  $\text{CH}_4$  fraction of 0.2.  $\text{CH}_2$  production increases with an increase in the  $\text{CH}_4$  fraction.  $\text{CO}_2$  and  $\text{O}_2$  are consumed within such a process. The system achieves 100% carbon efficiency and 67% hydrogen



efficiency because some of the hydrogen is oxidised to  $H_2O$  to get rid of the oxygen in the system. Both the enthalpy and Gibbs free energy are negative throughout, as illustrated in Figure 3b, therefore, neither energy nor work input is required to achieve such a target. This makes the production of FT-fuels an attractive target that needs to be maximised.

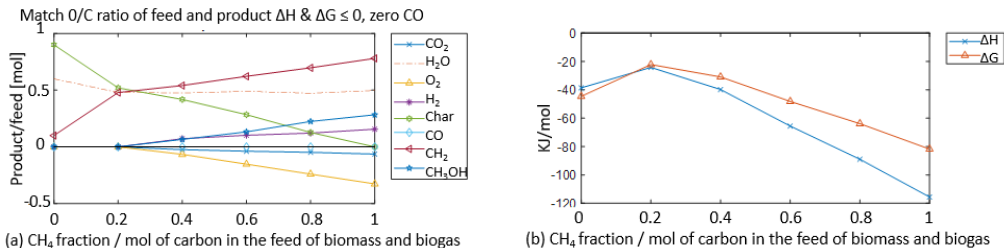


Figure 3: (a) Material targets and (b) energy and work targets that match the O/C ratio in the feed to the O/C ratio in the products

Setting the target to maximise the production of  $CH_2$  led to consumption of a significant amount of  $CO_2$  from a methane fraction of 0.4 and less  $O_2$  is required as illustrated in Figure 4a.  $CH_2$ ,  $H_2O$  and char are the only products obtained and  $CH_2$  production increases with an increase in the fraction of  $CH_4$  in the feed. From Figure 4b, we concluded that the process is thermodynamically feasible. Both the enthalpy and the Gibbs free energy are less or equal to zero throughout the system. A zero Gibbs free energy means that the process fully utilises the chemical potential of the feed material (Sempuga et al., 2010).

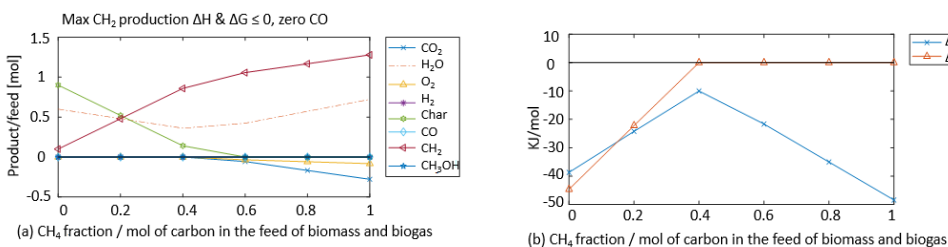


Figure 4: (a) Material targets and (b) energy and work targets that maximize the production of  $CH_2$

### 3.2 Aspen Plus Simulation Results

The results from the simulation on Aspen Plus, shown in Table 1, show that it is not possible to achieve a net  $CO_2$  consumption unless we rigorously integrate the heat and work energies within the system, particularly, between the reformer and the synthesis reactors. Case0 shows a system with no energy integration, the energy released from the FT-synthesis (-53.26kW) reactor is simple lost, and it can be seen that the process deviates by approximately 200% from the theoretical  $CO_2$  consumption target of 0.17kmol and the carbon and hydrogen efficiencies are significantly low, 84% and 48%, respectively.

Table 1: Results obtained when the thermal efficiency is varied from 0 to 43% at a synthesis operating pressure of 1 bar.

	Case0	Case1	Case2	Case3	Case4	Case5	Target
<i>Carnot-eff</i>	0,00	0,10	0,25	0,35	0,43	0,43	
<i>Q-Synthesis (kW)</i>	-53,26	-55,26	-58,57	-61,07	-63,33	-75,10	

<b><i>Q-Gas<sub>Ref</sub></i> (kW)</b>	51,93	53,93	57,26	59,76	62,01	73,63	
<b><i>Q-Combustion</i></b> <b>(kW)</b>	-51,93	-46,03	-36,35	-29,23	-23,10	-5,79	
<b><i>Q<sub>deff</sub></i> @ 720 (kW)</b>	0,00	0,00	0,00	0,00	0,00	21,69	
<b><i>Recycle fr</i></b>	0,61	0,65	0,71	0,75	0,78	0,92	
<b><i>CH<sub>4</sub></i></b>	-0,80	-0,80	-0,80	-0,80	-0,80	-0,80	-0,80
<b><i>H<sub>2</sub></i></b>	0,00	0,00	0,00	0,00	0,00	0,00	0,00
<b><i>CO</i></b>	0,00	0,00	0,00	0,00	0,00	0,00	0,00
<b><i>CO<sub>2</sub></i></b>	0,16	0,13	0,07	0,03	0,00	-0,10	-0,17
<b><i>H<sub>2</sub>O</i></b>	0,80	0,77	0,70	0,66	0,62	0,51	0,57
<b><i>O<sub>2</sub></i></b>	-0,50	-0,45	-0,36	-0,30	-0,24	-0,09	-0,06
<b><i>CH<sub>2</sub></i></b>	0,84	0,87	0,93	0,97	1,00	1,10	1,17
<b><i>Biomass</i></b>	-0,20	-0,20	-0,20	-0,20	-0,20	-0,20	-0,20
<b><i>C<sub>eff</sub></i></b>	0,84	0,87	0,93	0,97	1	1	1
<b><i>H<sub>eff</sub></i></b>	0,48192	0,50108	0,53259	0,555	0,575	0,631	0,671
				7	5	1	
<b><i>Delta H</i></b>	-223,66	-200,51	-162,46	-135	-111	-43	-35,01
<b><i>Delta G</i></b>	-187,45	-165,14	-128,45	-102	-78	-13	0

A heat engine and heat pump configuration, as shown in Figure 5, is adopted to integrate the heat from a low temperature exothermic process to a high temperature endothermic process.

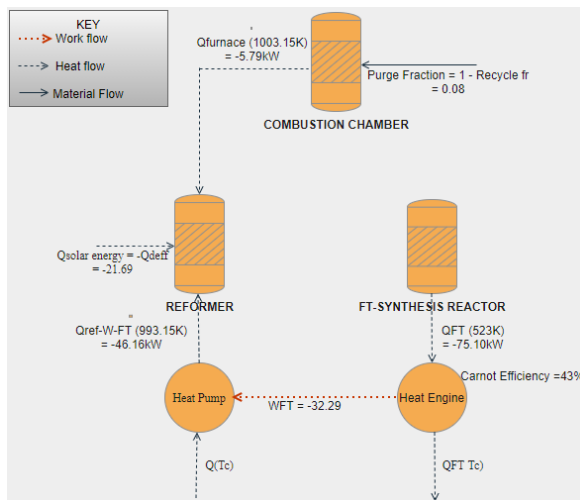


Figure 5: Illustration of the heat engine-heat pump configuration

A 19% decrease in CO<sub>2</sub> production was achieved when the heat engine was operated at a carnot efficiency of 10%, shown in Case1. As the carnot efficiency is increased from 10% to 43%, CO<sub>2</sub> production decreases while the carbon and hydrogen efficiencies increase significantly. A net-zero CO<sub>2</sub> production was obtained at Carnot efficiency of 0.43 as shown in Table 1, Case4. However, the system still deviates by 100% from the 0.17 kmol CO<sub>2</sub> consumption target. This is because, at 43% Carnot efficiency, the synthesis section cannot fully supply the heat energy needed in the reformer, therefore, some of the hydrocarbon product is burnt to provide the additional heat, hence, produces CO<sub>2</sub>. The only way the target can be achieved is through integration of solar energy as a clean, renewable energy source to compensate for the energy shortfall, shown as 21.69kW in

Case5, Table 1. Case5 is illustrated in Figure 5 and shows heat integration from the FT-synthesis reactor, combustion chamber and solar energy.

## Conclusions

Biomass and biogas are considered to be green and sustainable sources of hydrogen. In this work, we show that the process suffers a significantly low carbon efficiency and we propose the production of methanol and/or Fischer-Tropsch fuel instead of hydrogen to achieve higher conversion efficiencies. The production of methanol and FT-fuels is an attractive target and is thermodynamically feasible throughout the biogas to biomass ratios considered (CH<sub>4</sub> fraction). There is potential to achieve a net CO<sub>2</sub> consumption in the biomass to FT-fuels process, however Aspen Plus® simulation clearly shows that without energy integration throughout the system, the process will always emit CO<sub>2</sub> and have low carbon and hydrogen efficiencies. A zero net CO<sub>2</sub> emission can be attained when the synthesis heat, converted to work at the maximum Carnot efficiency, is integrated into the reformer. Furthermore, the simulation clearly shows that introduction of external energy is necessary to attain a net CO<sub>2</sub> consumption even if the process is fully energy integrated. This is a direct consequence of reaction kinetics and equilibrium limitations. In this instance, solar is considered as a renewable energy source. During the targeting analysis, we noted that matching the H/C ratio in the feed to H/C ratio in the products favours the production of methanol. This process should be investigated further to gain insight on the effect of heat and work integration in the biomass to methanol process. The methanol can be used in blended fuels to improve the octane number and hence, the quality of the fuel.

## References

- Lepage, T., Kammoun, M., Schmetz, Q., & Richel, A. (2021). Biomass-to-hydrogen: A review of main routes production, processes evaluation and techno-economical assessment. *Biomass and Bioenergy*, *144*(December 2020), 105920. <https://doi.org/10.1016/j.biombioe.2020.105920>
- Liu, X., & Patel, B. (2017). Carbon efficiency targets for methanol production from a hybrid solar-carbonaceous feedstocks process. *AIP Conference Proceedings*, *1850*(June 2017). <https://doi.org/10.1063/1.4984471>
- Patel, B., Hildebrandt, D., Glasser, D., & Hausberger, B. (2005). Thermodynamics analysis of processes. 1. Implications of work integration. *Industrial and Engineering Chemistry Research*, *44*(10), 3529–3537. <https://doi.org/10.1021/ie048787f>
- Sempuga, B. C., Hausberger, B., Patel, B., Hildebrandt, D., & Glasser, D. (2010). Classification of chemical processes: A graphical approach to process synthesis to improve reactive process work efficiency. *Industrial and Engineering Chemistry Research*, *49*(17), 8227–8237. <https://doi.org/10.1021/ie100288h>

# Modeling and Assessment of Electrochemical Reduction of CO<sub>2</sub> in Amine-based Capture Solvents

Alexios-Spyridon Kyriakides,<sup>a</sup> Panagiotis Kazepidis,<sup>a</sup> Athanasios I. Papadopoulou,<sup>a</sup> Panos Seferlis<sup>b</sup>

<sup>a</sup>*Chemical Process & Energy Resources Institute (C.P.E.R.I.), Center for Research and Technology Hellas (CE.R.T.H.), P.O. Box 60361, 57001, Thessaloniki, Greece*

<sup>b</sup>*Department of Mechanical Engineering, Aristotle University of Thessaloniki, P.O. Box 484, 54124 Thessaloniki, Greece*

*alexkyr@certh.gr*

## Abstract

Electrochemical CO<sub>2</sub> reduction is a promising utilization process. The use of amine solvents to deliver the captured CO<sub>2</sub> into the cell is receiving experimental attention, but models are also needed to investigate the performance of such systems. This work aims to model and assess the performance of the electrochemical reduction of CO<sub>2</sub> into CO in an amine capture solvent. The developed process model is validated using a conventional electrolyte (KHCO<sub>3</sub>), exhibiting good match and low deviation from experimental data. The process is evaluated for monoethanolamine (MEA), and compared to KHCO<sub>3</sub> within a range of 2-20 cm<sup>3</sup>/min of total inlet cell flow. MEA enables lower undesirable H<sub>2</sub> formation (less than 5% of the total current density) and faster reaction, as implied by the higher CO current density (from 45%, up to 98.7% within the considered flow range). KHCO<sub>3</sub> exhibits higher conversion rates than MEA, but MEA's higher CO production rate within the entire inlet flow range reaches a value of 30.8% at the lower flow end.

**Keywords:** CO<sub>2</sub> utilization, electrochemical reduction of CO<sub>2</sub>, CO<sub>2</sub>RR, amines

## 1. Introduction

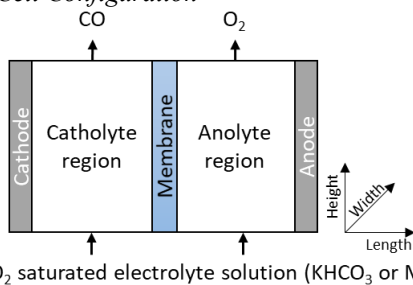
Solvent-based, post-combustion CO<sub>2</sub> capture is a mature technology for short-term industrial implementation (Nessi et al., 2021). The high capture costs that mainly result from the intense, thermal solvent regeneration prohibit its wide deployment. Electrochemical CO<sub>2</sub> reduction processes for the production of valuable products (e.g., CO, CH<sub>3</sub>OH, etc.) are gaining attention as a means of off-setting the capture costs (De Luna et al., 2019). These systems enable the *in situ* production of H<sub>2</sub>, which reacts with CO<sub>2</sub> over catalysts to attain the desired products. Conventionally, the CO<sub>2</sub> is transferred into the electrochemical cell in a purely aqueous solution. The low solubility of CO<sub>2</sub> in water and the mass transfer and kinetic limitations result in low production rates and conversion of feedstock. Experimental reports propose the use of an amine solvent (e.g., monoethanolamine (MEA)-Lee et al., 2020; and 1-cyclohexylpiperidine (CHP)-Diaz et al., 2018) directly from the absorber of the capture process to transfer the captured CO<sub>2</sub> into the electrochemical cell. There the CO<sub>2</sub> may be released and react with hydrogen, while the solvent is regenerated. This approach enables the delivery of a richer CO<sub>2</sub> stream, than using water as the carrier, and allows the complete replacement of the thermal solvent regeneration process.

The mathematical models developed for the electrochemical CO<sub>2</sub> reduction process largely focus on aqueous CO<sub>2</sub> solutions as feedstock, whereas additives like KOH (Weng *et al.*, 2019) or KHCO<sub>3</sub> (Kotb *et al.*, 2017) are also used as electrolytes to increase performance. Although there are several studies on the experimental evaluation of electrochemical reduction of CO<sub>2</sub> in amine-based media, there are no studies on the mathematical modeling of the process, that could aid in process design and optimization. The aim of this study is the development and validation of a mathematical model for the electrochemical reduction of CO<sub>2</sub> into CO (selected due to the availability of experimental data), using an aqueous amine-CO<sub>2</sub> mixture as feedstock, with the model also used to compare the amine performance with KHCO<sub>3</sub> solutions. The evaluation considers process indicators such as partial current densities, Faradaic efficiency and CO<sub>2</sub> conversion.

## 2. Model Development

This section aims to discuss the modeling approach, governing equations, occurring reactions and boundary conditions. First, the one-dimensional, isothermal model of the planar electrochemical cell configuration is presented, followed by the governing equations of all underlying physicochemical phenomena (mass and current balances, Nernst-Planck and Ohm). Then all reactions taking place both in the electrolyte and electrodes (charge transfer reactions) are described, followed by the boundary conditions.

### 2.1. Cell Configuration



**Figure 1.** Schematic representation of the electrochemical CO<sub>2</sub> reduction cell configuration.

The electrochemical cell (Figure 1) is composed of two electrodes and a proton exchange membrane. A channel filled with the CO<sub>2</sub>-saturated electrolyte solution separates the cathode and anode electrodes and the membrane, namely the catholyte and anolyte regions. Both regions are considered to be well-mixed and are constantly fed with a CO<sub>2</sub>-saturated electrolyte solution (either KHCO<sub>3</sub> or MEA in this work, but the model

is generic), whereas the proton exchange membrane separates the two regions.

### 2.2. Governing Equations

In both electrolyte channels (anolyte and catholyte), mass continuity is considered:

$$\frac{\partial C_i}{\partial t} + \nabla \cdot N_i = R_i \quad (1)$$

where  $N_i$  is the species molar flux [ $mol/(m^2s)$ ], and  $R_i$  is the source term [ $mol/(m^3s)$ ], for  $i = CO_2, H^+, OH^-, HCO_3^-, CO_3^{2-}, K^+, RNH_2, RNHCOO^-, CO, O_2, H_2$ . At steady state the first term is equal to zero, whereas only the distance between the two electrodes is considered, in terms of spatial dimensions. The species flux is computed with the Nernst-Planck equations, which consider diffusion, electric migration and convection transportation, as described by Eq. (2):

$$N_i = -D_i \nabla C_i - z_i u_{mi} F C_i \nabla \phi_i + \mathbf{u} C_i, \quad u_{mi} = \frac{D_i}{RT} \quad (2)$$

where  $D_i$  is the diffusion coefficient [ $m^2/s$ ],  $C_i$  is the species concentration [ $mol/m^3$ ],  $z_i$  is the species charge number,  $u_{mi}$  is the mobility of ions [ $(s\ mol)/kg$ ],  $F$  is the Faraday constant [ $C/mol$ ],  $\varphi_l$  is the electrolyte potential [ $V$ ], and  $\mathbf{u}$  is the velocity vector [ $m/s$ ]. The ionic mobility is based on the Nernst-Einstein relation (Eq. (2)), where  $R$  is the universal gas constant [ $J/(K\ mol)$ ],  $T$  is the temperature [ $K$ ]. The charge conservation at the electrolyte is considered, while the current density is described by Ohm's law:

$$\nabla \cdot \mathbf{i}_l = 0, \mathbf{i}_l = -\sigma_l \nabla \varphi_l, \sigma_l = \frac{F^2}{RT} \sum_i z_i^2 D_i C_i \quad (3)$$

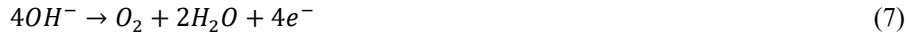
where  $\mathbf{i}_l$  is the electrolyte current density [ $A/m^2$ ], and  $\sigma_l$  is the electronic conductivity at the electrolyte [ $S/m$ ]. The source term in Eq. (1) is given by the following equation:

$$R_i = R_{CT,i} + R_{EL,i} \quad (4)$$

where  $R_{CT,i}$  is related to the charge transfer reactions, and  $R_{EL,i}$  is related to the electrolyte reactions (described in detail in section 2.3).

### 2.3. Charge Transfer (Anode and Cathode) and Electrolyte Reactions

Hydrogen (HER, Eq. (5)) and carbon evolution reactions (COER, Eq. (6)), and the anode-related reaction is the oxygen evolution reaction (OER, Eq. (7)).



The reaction kinetics are assumed to follow the Tafel kinetics. The reaction current densities ( $i_k$ ) are calculated by Eq. (8),(9) and (10), for the three reactions, respectively.

$$i_{HER} = i_{o,HER} \exp\left(\frac{-\alpha_{HER} F \eta_{HR}}{RT}\right) \quad (8)$$

$$i_{COER} = i_{o,COER} \frac{C_{CO_2}}{C_{CO_2,ref}} \exp\left(\frac{-\alpha_{COER} F \eta_{COER}}{RT}\right) \quad (9)$$

$$i_{OER} = \exp\left(\frac{\alpha_{OER} F \eta_{OER}}{RT}\right) \quad (10)$$

where  $i_{o,k}$  is the exchange current density,  $\alpha_k$  is the charge transfer coefficient,  $C_{CO_2,ref}$  is the saturation concentration of dissolved CO<sub>2</sub> at ambient conditions, and  $\eta_k$  is the overpotential for reaction  $k$  [ $V$ ], given by Eq. (11):

$$\eta_k = \varphi_s - \varphi_l - E_k^0 \quad (11)$$

where  $\varphi_s$  is the solid phase (cathode or anode) applied potential, and  $E_k^0$  is the reaction equilibrium potential [ $V$ ]. The charge transfer reaction source term is defined as:

$$R_{CT,i} = \sum_k \frac{\nu_i i_k}{n_k F} \quad (12)$$

where  $\nu_i$  is the stoichiometric coefficient of component  $i$  in reaction  $k$ , and  $n_k$  is the number of electrons transferred in reaction  $k$ .

### 2.3.1. Utilization of $\text{KHCO}_3$ as an electrolyte

For  $\text{KHCO}_3$  the following reactions are considered at the electrolyte phase:

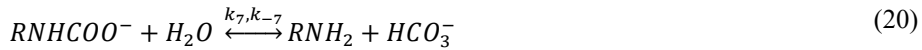
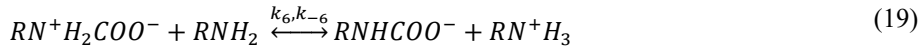


Additionally, water dissociation reaction takes place:



### 2.3.2. Utilization of MEA as an electrolyte

For aqueous MEA solution additional reactions are considered for the equilibrium at the electrolyte solution. Primary amines capture  $\text{CO}_2$  by reacting with it to form carbamic acid, carbamate ion and bicarbonate through the following reaction mechanism:



Then, bicarbonate ( $\text{HCO}_3^-$ ) reacts with  $\text{H}^+$  (Eq. (13)) in both electrolyte channels to produce  $\text{CO}_2$  to be used for the reduction reactions taking place on the electrodes. The bulk reaction source term for component  $i$  is defined as:

$$R_{EL,i} = \sum_n v_i \left( k_n \prod_{v_j < 0} C_j - k_{-n} \prod_{v_j > 0} C_j \right) \quad (21)$$

where  $k_n$  and  $k_{-n}$  is the forward and backward reaction rate of the  $n^{\text{th}}$  reaction.

### 2.4. Boundary Conditions

At the boundary between the electrode and the electrolyte, current conservation applies:

$$\nabla(-\sigma_l \nabla \phi_l) = S_l, S_{l,c} = \alpha_{\text{COER}} i_{o,\text{COER}} + \alpha_{\text{HER}} i_{o,\text{HER}}, S_{l,a} = -\alpha_{\text{OER}} i_{o,\text{OER}} \quad (22)$$

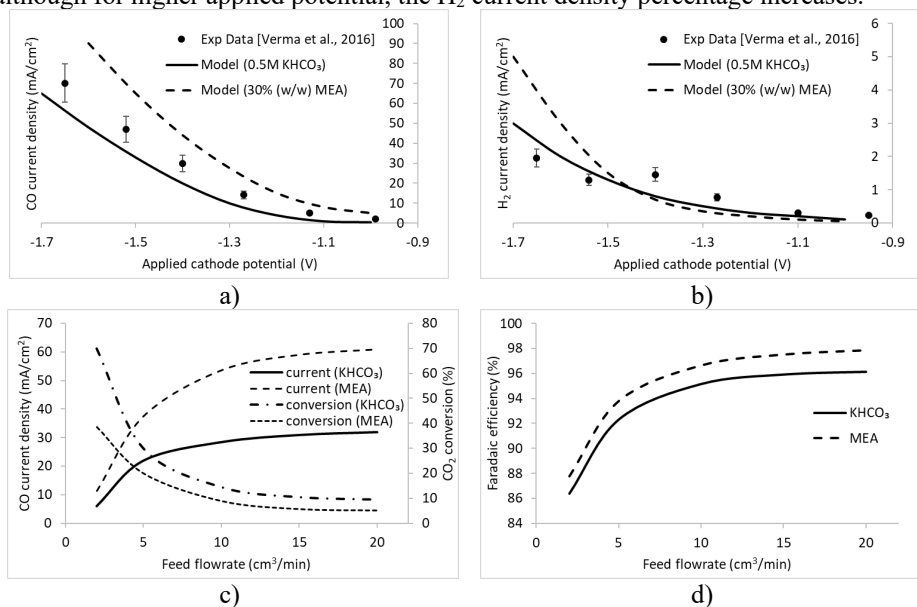
where  $S_l$  is the current source term resulting from the electrochemical reactions occurring at each electrode.

## 3. Results and discussion

The electrochemical cell dimensions are  $3.4 \times 3.4 \times 8.1$  cm in height, width and length, with a membrane of 1.5 mm thickness (Verma *et al.*, 2016). The electrolyte flowrate is equal to  $20 \text{ cm}^3/\text{min}$ , at 298.15 K and 1 atm. 0.5 M  $\text{KHCO}_3$  and 30% w/w MEA aqueous solutions are considered. Reaction kinetics and constants are from Borhani *et al.* (2018), Gupta *et al.* (2013) and Hebelmann *et al.* (2022).

### 3.1.1. Partial Current Density vs Applied Potential

The model is validated for KHCO<sub>3</sub> with experimental data from Verma et al. (2016) (Figure 2a and Figure 2b). The partial CO (Figure 2a) and H<sub>2</sub> (Figure 2b) current densities are presented as a function of applied cathode potential. The simulated results are in good agreement with the experimental data. Higher discrepancies are observed at higher applied potential for the CO partial current density. In both cases, both CO and H<sub>2</sub> partial current densities increase when the applied potential in the cathode is increased (becomes more negative). Higher partial current densities are observed in the case of MEA, resulting in higher reaction rates. The undesirable H<sub>2</sub> formation is less than 10% of the total current density for all potentials tested for KHCO<sub>3</sub> and less than 5% for MEA, although for higher applied potential, the H<sub>2</sub> current density percentage increases.



**Figure 2.** Comparison between experimental and simulation results for a) CO and b) H<sub>2</sub> partial current densities as a function of applied cathode potential. Effect of feed flowrate on c) CO current density and CO<sub>2</sub> conversion rates and d) on Faradaic efficiency.

### 3.1.2. Effect of MEA solution feed flowrate

The effect of the feed flowrate on electrochemical cell performance is shown in Figure 2c and Figure 2d. As the flowrate increases (2 to 20 cm<sup>3</sup>/min), the current density and Faradaic efficiency increase too, whereas the CO<sub>2</sub> conversion decreases. At flowrates higher than 10 cm<sup>3</sup>/min, the CO<sub>2</sub> concentration near the electrode is higher, resulting in higher reaction rates, current density and Faradaic efficiency. However, the effect of mass transport limitation increases, resulting in marginal increase in current density and Faradaic efficiency for further flowrate increase. Higher flowrate results in lower residence time, which affects the overall CO<sub>2</sub> conversion rate. At flowrates higher than 10 cm<sup>3</sup>/min, the current density achieved in the MEA solution is significantly higher than that achieved when KHCO<sub>3</sub> is used as an electrolyte. An increase between 45% and 98.7% is observed for 2 and 20 cm<sup>3</sup>/min, respectively. The CO<sub>2</sub> conversion is lower by 40% on average. However, the lower conversion rate when combined with the higher inlet molar flowrate of CO<sub>2</sub> (0.77 and 0.40×10<sup>-3</sup> mol/s for MEA and KHCO<sub>3</sub>) results in higher CO production rates. For example, the outlet CO flowrates for MEA are 0.34 and



$0.06 \times 10^{-3}$  mol/s, whereas for  $\text{KHCO}_3$  they are 0.28 and  $0.04 \times 10^{-3}$  mol/s (for 2 and 20  $\text{cm}^3/\text{min}$ ). At the higher flowrate range, the conversion is similar for both electrolytes, however the higher  $\text{CO}_2$  concentration and reaction rate that MEA exhibits result in 30.8% higher CO production rate than  $\text{KHCO}_3$ . The Faradaic efficiency presents similar behavior for both solutions.

#### 4. Conclusions

The comparative study and the parametric analysis indicate that higher CO and lower  $\text{H}_2$  current densities are observed for MEA compared to  $\text{KHCO}_3$ , implying higher reaction rates and less undesirable product. Lower  $\text{CO}_2$  conversion ratios are observed for MEA, attributed to higher initial  $\text{CO}_2$  quantities and mass transport limitations. However, when these are combined with the initial higher  $\text{CO}_2$  concentration at the inlet streams, they result in higher CO production for MEA. In future work, optimization studies should investigate the existence of an intermediate optimal potential, flowrate pair and electrochemical cell dimensions, such that a trade-off between operating cost (Faradaic efficiency) and capacity cost (flowrate, conversion, dimensions) can be achieved.

#### Acknowledgements

This research has been co-financed by the European Union and Greek national funds under the call ERANET 2021a-European E&T Cooperation in European Network ACT, Project T12EPA5-00040.

#### References

- Borhani T.N., Oko E., Wang, M., 2018. Process modelling and analysis of intensified  $\text{CO}_2$  capture using monoethanolamine (MEA) in rotating packed bed absorber. *Journal of Cleaner Production*, 204, pp.1124-1142.
- De Luna P., Hahn C., Higgins D., Jaffer S.A., Jaramillo, T. F., Sargent, E. H., 2019. What would it take for renewably powered electrosynthesis to displace petrochemical processes? *Science*, 364 (6438), eaav3506.
- Diaz L.A., Gao N., Adhikari B., Lister T.E., Dufek E.J. and Wilson A.D., 2018. Electrochemical production of syngas from  $\text{CO}_2$  captured in switchable polarity solvents. *Green Chemistry*, 20(3), pp.620-626.
- Gupta M., da Silva E.F., Hartono A., Svendsen, H.F., 2013. Theoretical Study of Differential Enthalpy of Absorption of  $\text{CO}_2$  with MEA and MDEA as a Function of Temperature. *The Journal of Physical Chemistry B*, 117(32), pp.9457-9468.
- Hebelmann M., B.C. Bräsel B.C., R.G. Keller R.G., M. Wessling M., 2022, Simulation-based guidance for improving  $\text{CO}_2$  reduction on silver diffusion electrodes, *Electrochemical Science Advances*, 2100160.
- Kotb Y., Fateen S.E.K., Albo J. and Ismail I., 2017. Modeling of a Microfluidic Electrochemical Cell for the Electro-Reduction of  $\text{CO}_2$  to  $\text{CH}_3\text{OH}$ . *Journal of The Electrochemical Society*, 164(13), p.E391.
- Lee G., Li, Y.C., Kim J.-Y., Peng T., Nam D.-H., Rasouli., A.S., Li F., Luo M., Ip A.H., Joo Y.-C., Sargent E.H., 2021, Electrochemical upgrade of  $\text{CO}_2$  from amine capture solution. *Nature Energy*, 6, p46-53.
- Nessi E., Papadopoulos A.I., Seferlis P., 2021, A Review of Pilot and Commercial Plants for Post-Combustion  $\text{CO}_2$  Capture: Packed Bed, Phase-Change and Rotating Processes, *International Journal of Greenhouse Gas Control*, 111, 103474.
- Verma S., Lu X., Ma S., Masel R.I., Kenis P.J.A., 2016, The effect of electrolyte composition on electroreduction of  $\text{CO}_2$  to CO an Ag based gas diffusion electrodes. *Physical Chemistry Chemical Physics*, 18 p.7075.
- Weng L.-C., Bell A.T., Weber A.Z., 2019, Towards membrane-electrode assembly systems for  $\text{CO}_2$  reduction: a modeling study. *Energy & Environmental Science*, 12, p.1950.

# Design of CO<sub>2</sub> Capture and Mineralization Systems: Integrated Process Optimization and Controllability Assessment in Parallel Infrastructures

Thomas Prousalis<sup>a</sup>, George Gkizas<sup>b</sup>, Athanasios I. Papadopoulos<sup>b</sup>, Panos Seferlis<sup>c</sup>

<sup>a</sup>*Y Squared P.C., 57019, Perea, Greece*

<sup>b</sup>*Chemical Process and Energy Resources Institute, Centre for Research and Technology Hellas, 57001, Themi, Thessaloniki, Greece*

<sup>c</sup>*Department of Mechanical Engineering, Aristotle University of Thessaloniki, 54124, Thessaloniki, Greece*  
*spapadopoulos@certh.gr*

## Abstract

This work presents the use of an optimization algorithm with approximate computing capabilities in the simultaneous CO<sub>2</sub> capture (CC) and utilization (CU) process design and controllability assessment. The CC design problem includes 6 solvent and process flowsheet combinations, followed by the use of a Rotating Packed Bed (RPB) process for the production of precipitated calcium carbonate (PCC) minerals as the CU option. The algorithm employs memoization and task dropping to accelerate execution in a high-performance cluster. The results indicate up to 141.7 h of saved CPU time due to the use of the approximate computing techniques. When the CC and CU processes are considered simultaneously, the generated Pareto front changes compared to the one of the CC process. The 2-Aminoethanol 30 wt. % solvent with intercooling and side feeds in the desorber results in almost equally profitable PCC production to the 3-Amino-1-Propanol 35 wt. % solvent in simpler flowsheets.

**Keywords:** CO<sub>2</sub> capture and utilization, process optimization, controllability assessment.

## 1. Introduction

CO<sub>2</sub> capture (CC) technologies promise to help mitigate the effects of global warming, but their wide industrial deployment is elusive due to high costs (Bui et al., 2018). The latter can be offset by the profitability induced by CO<sub>2</sub> utilization (CU) processes. The transformation of CO<sub>2</sub> into carbonated salts (minerals) is a mature technology and there is a wide market for products such as precipitated calcium carbonate (PCC) (Jimoh et al., 2018). Such CC and CU systems need to be designed simultaneously, considering their performance under disturbances, but the calculations are challenging. The combination of algorithms for process design and control with parallelization techniques could accelerate the computations and widen the scenarios leading to superior designs.

In the area of CO<sub>2</sub> capture, there are published approaches for integrated process design and control (Sahraei and Ricardez-Sandoval, 2014), while advanced control approaches have also been presented (Jung et al., 2020). A framework for integrated process design and control has also been proposed where parallel capabilities were considered, without specific application details or benefits (Miller et al., 2017). Multi-scale design frameworks for CC, CU and sequestration have been proposed (Hasan et al., 2015), without considering the controllability of the CC or CU systems. The effects of different

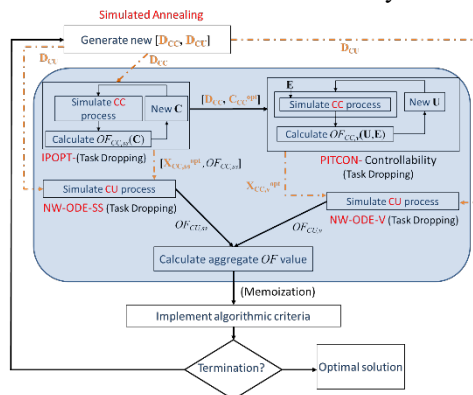
capture materials (e.g., solvents) and process flowsheets in the simultaneous design and controllability assessment of CC and CU systems have not been reported previously. Vasilas et al. (Vasilas et al., 2022) developed an algorithm for simultaneous process design and controllability in distributed and heterogeneous infrastructures, applying it to the design of CC systems. The work showed the benefits that can be attained when several decision levels are addressed simultaneously and also exhibited significant reductions in the computational effort due to the use of advanced parallelization tools.

Building on these developments, this paper incorporates for the first time in the above algorithm a mineralization process that uses the captured  $\text{CO}_2$  as raw material to produce PCC in a rotating packed bed (RPB). The latter is another innovation introduced here, as PCC production and CC process design have never been integrated for different solvents and process layouts or operating variability. The innovation is enhanced by the use of an RPB, which is an intensified reactor in terms of size and efficiency (Borhani et al., 2019).

## 2. Models and methods

### 2.1. Algorithm and problem formulation

The proposed algorithm (Figure 1) comprises an outer loop that handles discrete design parameters through Simulated Annealing (SA) (Kirkpatrick et al., 1983) and two inner loops, whereby steady-state process optimization is performed, followed by process optimization under the influence of disturbances. Two vectors  $D_{CC}$  and  $D_{CU}$  with the discrete model input parameters are generated for both the CC and CU processes. The CC model undergoes optimization for the given  $D_{CC}$  within the IPOPT block (Wächter and Biegler, 2006), which considers only the continuous parameters as design variables (vector  $C$ ), targeting the minimization of the objective function  $OF_{CC,ss}$ . The optimized state variables  $X_{CC,ss}^{opt}$  and the  $OF_{CC,ss}$  from IPOPT, as well as the  $D_{CU}$  vector from SA are used as input to the CU model that generates the  $OF_{CU,ss}$ . This block is designated as “NW-ODE-SS”, representing the inclusion of a Newton-based (NW) step and an ordinary differential equations (ODE) solution step for the steady-state (SS) case. The  $OF_{CC,ss}$  value is the cost per ton of captured  $\text{CO}_2$  that is used in the CU model as raw material. These two blocks represent the SS instance of the CCU system in each SA iteration.



**Figure 1:** Main algorithmic steps. Vector  $D$  represents the discrete variables, vector  $C$  the continuous variables, vector  $X$  the state variables, vector  $E$  the disturbances, vector  $E$  the manipulated variables,  $OF$  the objective function,  $ss$  is steady state, and  $v$  variability.

Vector  $C_{CC,ss}^{opt}$  of the continuous variables and  $D_{CC}$  of the discrete variables are passed to the PITCON block (Rheinboldt and Burkardt, 1983) for the controllability assessment of

the CC process. The latter is achieved through a non-linear sensitivity analysis approach (Seferlis and Grievink, 2004) and implemented in an integrated design fashion (Vasilas et al., 2022). The model is simulated under the effect of disturbances to calculate the CC process performance under variability, represented through  $OF_{CC,v}$ . The latter expresses the deviation of the operating point of the CC process from the set-point and the resources required to maintain this set-point under the effect of exogenous disturbances. Then, the  $D_{CU}$  vector, the CC model state variables  $X_{CC,v}^{opt}$  and the calculated cost per ton of CO<sub>2</sub> under variability are passed to the CU process (NW-ODE-V block). The CU model is simulated for the optimum solution that was attained by the CC model under the influence of disturbances, providing the  $OF_{CU,v}$  value which is different than the corresponding  $OF_{CU,ss}$  one. The effects of the disturbances on the CC process are therefore propagated indirectly to the CU model calculations.

The optimum design and controllability assessment problem formulation is represented through the objective functions shown below. The addition of equality and inequality constraints are omitted due to space restrictions. Details for the CC part are reported in Vasilas et al. (2022), and for the CU part in Dimoliani et al. (2021). The overall  $OF$  of Eq. (1) represents maximization of the profit from selling the product, considering the maximization of the sum of the steady-state and under variability profits. For a vector of disturbances that induce deterioration of the  $OF_{CU,v}$  it is desired to target a value that is as close as possible to the  $OF_{CU,ss}$ . The CC optimization and controllability assessment objective functions are shown in Eq. (2) and (3).

Overall optimization problem (SA loop), considering the CU model:

$$\max OF = OF_{CU,ss} + OF_{CU,v} \quad (1)$$

Optimization of the CC process at steady-state (IPOPT block):

$$\min OF_{CC,ss} = \frac{TAC_{CC}}{\dot{m}_{CO_2} \cdot t_a} \quad (2)$$

Controllability assessment of CC process (PITCON block):

$$\min OF_{CC,v} = (Y - Y_{sp})^T W_Y (Y - Y_{sp}) + (U - U_{ss})^T W_U (U - U_{ss}) \quad (3)$$

where  $TAC_{CC}$  is the total annual cost (M€),  $\dot{m}_{CO_2}$  is the flow rate of the captured CO<sub>2</sub> (t/h),  $t_a$  is the annual equipment deployment (h/y),  $Y$  is the vector of controlled variables and  $Y_{sp}$  the vector of set-points.  $W_Y$  and  $W_X$  are the weight matrices (Vasilas et al., 2022).  $OF_{CC,ss}$  is used as input to simulate the CU model at steady-state, whereas the same term as that of Eq. (2) is calculated based on the result of the PITCON block (it is the cost per ton of CO<sub>2</sub> for the CC process under variability) and used as input to the CU model in the NW-ODE-V block. The objective functions of the CU process are the following:

$$OF_{CU,i} = (SP - CP_{Prod,i}) \cdot P \text{ (NW-ODE-}i \text{ block)}, i=\{ss, v\} \quad (4)$$

where  $SP$  is the product selling price (€/t),  $CP_{Prod}$  is the product cost (€/t) and  $P$  is the total yearly production (t/y).

## 2.2. Acceleration and parallelization

The algorithm implements the approximate computing techniques of memoization and task dropping (Mittal, 2016) to accelerate the execution, on top of the parallelization. Memoisation refers to the development of a stored record of already visited and converged designs. The record is used to directly attain the stored solutions when the

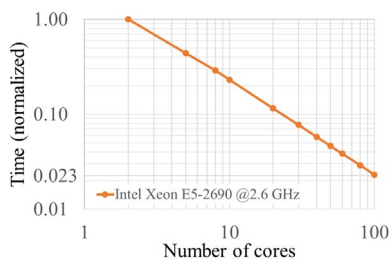
corresponding designs are re-visited by SA (a typical feature in this algorithm), instead of repeating the time-consuming simulations. Task dropping refers to avoiding tasks which incur intense computational effort. These include the IPOPT, PITCON and NW-ODE algorithms, that take longer to converge or do not converge as the values of the discrete variables shift the simulations away from the fixed starting point used here. A task is automatically dropped as unconverged when a time threshold, set from offline tests, is overcome. Vectors  $D_{CC}$  and  $D_{CU}$  are distributed to an equivalent number of parallel computing units which execute the IPOPT, PITCON and NW-ODE algorithms.

### 3. Implementation

We consider 6 cases that result from the combination of 2 solvents (Sol1 and Sol2) and 3 process structures (Proc1, Proc2, and Proc3) for the CC case. Sol1 and Sol2 are aqueous 2-Aminoethanol 30 wt. % and 3-Amino-1-Propanol 35 wt. %. Proc1 is the conventional absorption/ desorption flowsheet, in Proc2 the rich  $\text{CO}_2$  stream enters the desorber at different locations, and Proc3 is a combination of Proc2 with cooling along the absorber height (Damartzis et al., 2014). The absorber and the stripper are divided in 3 elements, using the orthogonal collocation on finite elements method (Damartzis et al., 2014). The length of each element (ranging between 1-20) is considered in vector  $D_{CC}$ , whereas  $D_{CU}$  includes the inner radius of the RPB (0.2-1.2 m, with step of 0.1 m). Vector  $C$  (IPOPT) includes the solvent make-up flowrate (50-140 mol/s), the reboiler temperature (363-393 K), the absorber and desorber pressure (1-3 bar), the stream split ratios (0-100%), and the duties of the cooler and of the two intercoolers (0-2 MW). The controlled variables  $Y$  are the percentage of captured  $\text{CO}_2$ , the lean stream loading, and the column temperatures. The manipulated variables  $U$  are all the continuous variables, except for pressures. The disturbance  $E$  is on the  $\text{CO}_2$  concentration in the feed stream, that is allowed to be up to 2% lower than the nominal value of a feed stream with 12 % vol.  $\text{CO}_2$  and a total flow of 320.22 mol/s. A 10 wt. % aqueous  $\text{Ca}(\text{OH})_2$  is the raw material for RPB-based PCC production (Dimoliani et al., 2021). The number of packing layers is 10, the packing specific area and porosity are 2,500  $\text{m}^2/\text{m}^3$  and 0.76, and the rotating speed is 3,000 rpm. The  $t_a$  is 8,000 h/y and the  $SP$  is 1,000 €/t. The implemented annual variability scenario (VS-A) considers that 20 % of the yearly operation is at nominal operating conditions and 80 % under variability. The latter includes 24 %, 19.2 %, 16%, 12.8% and 8% of the time, disturbance of -0.4%, -0.8%, -1.2%, -1.6% and -2%, respectively.

## 4. Results and discussion

### 4.1. Algorithmic scalability



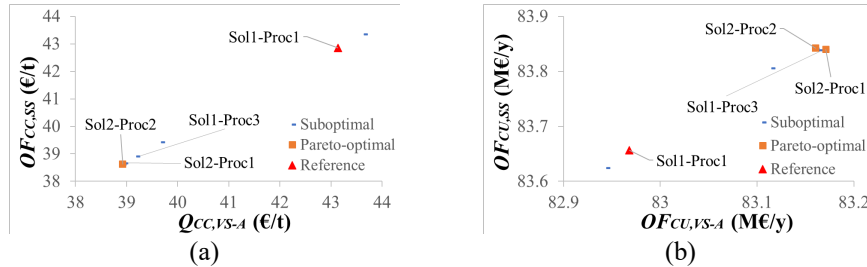
**Figure 2:** (a) Scalability analysis of the proposed algorithm as computation time against number of cores utilized.

The computational time from execution in 100 CPU cores was up to 72,000 s (20.0 h) for the performed runs. Memoization and task dropping, reduced the computational time of the algorithm by up to 510,000 s (141.7 h), by avoiding the repetitive simulations of already tested discrete variable combinations and computationally intensive tasks. Figure 2 exhibits a desirable scalability trend, as the computational time decreases linearly with the addition of parallel cores. Utilizing up to 100 CPU cores, we gradually achieve a 97.5 % reduction in execution time, compared to

running in 2 cores.

#### 4.2. CC results

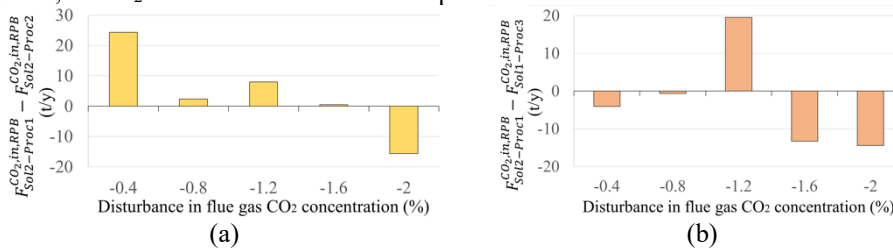
Figure 3a illustrates a Pareto analysis of the cost per ton of CO<sub>2</sub> at steady-state ( $OF_{CC,SS}$ ) vs. under VS-A ( $Q_{CC,VS-A}$ ) for all of the CC solvent-process combinations. The Pareto front indicates that Sol2-Proc2 achieves the best trade-off between operation at steady-state and under variability, overcoming the reference Sol1-Proc1. Sol2-Proc2 achieves a cost of 38.62 €/t at nominal conditions and 38.92 €/t CO<sub>2</sub> under VS-A. Sol2-Proc1 achieves very similar values, but is not Pareto optimal.



**Figure 3:** Pareto analysis of designed (a) CC process cost per ton CO<sub>2</sub> at steady-state vs. operation under VS-A (both OFs are minimized), (b) CU process annual profit at steady-state vs. operation under VS-A (both OFs are maximized).  $Q_{CC,VS-A}$  stands for cost per ton of CO<sub>2</sub> under VS-A.

#### 4.3. CU results

Figure 3b illustrates a Pareto analysis for the annual profit of the CU process (tailored after each tested CC process case) at steady-state vs. operation under VS-A. The reference case Sol1-Proc1 results again in suboptimal behavior. The Pareto front indicates that Sol2-Proc2 remains Pareto optimal, but Sol2-Proc1 is now part of the Pareto front. The latter may be considered the best choice for the CU process, due to its higher annual profit under variability (83.171 M€/y vs. 83.160 M€/y for Sol2-Proc2, although the values are very close). Even more pronounced is the case of Sol1-Proc3 that was clearly suboptimal in the CC process (Figure 3a) and has moved between the two Pareto optimum systems in Figure 3b. These changes can be explained based on the results of Figure 4. Although all of the tested solvent-process combinations have the same captured CO<sub>2</sub> flowrate at steady-state, disturbances may cause the captured CO<sub>2</sub> to deviate from case to case. For instance, Sol2-Proc1 captures about 25 tons of CO<sub>2</sub> per year more than Sol2-Proc2 in the disturbance magnitude of -0.4 %, and about 16 tons of CO<sub>2</sub> less in the magnitude of -2 %. Similar differences are observed for Sol1-Proc3. This behavior changes the annual profit, as CO<sub>2</sub> is the raw material in PCC production.



**Figure 4:** Deviation of CC process captured CO<sub>2</sub> flow vs. disturbance range in VS-A between a) Sol2-Proc1 and Sol2-Proc2, and b) Sol2-Proc1 and Sol2-Proc3.  $F_i^{CO_2,in,RPB}$  stands for the mass flow rate of captured CO<sub>2</sub> fed to the RPB reactor.

## 5. Conclusions

The presented work has illustrated the use of a parallel algorithm for solvent selection, CC and CU process design and controllability assessment, considering an RPB process as a CO<sub>2</sub> utilization option for production of PCC. The results showed that the proposed algorithm exhibits excellent scalability and significant acceleration is attained due to the use of approximate computing techniques. The Pareto optimal solvent-process structure combinations that were attained in the CC process exhibits differences with the Pareto front attained when the CU process is considered simultaneously, highlighting the need for algorithms for such complex design problems.

## Acknowledgements

The experiment “CCUPAR” has received funding from the European High-Performance Computing Joint Undertaking (JU) through the FF4EuroHPC project under grant agreement No 951745. The JU receives support from the European Union’s Horizon 2020 research and innovation programme and Germany, Italy, Slovenia, France, Spain.

## References

- T.N. Borhani, E. Oko, M. Wang, 2019. Process modelling, validation and analysis of rotating packed bed stripper in the context of intensified CO<sub>2</sub> capture with MEA, *J. Ind. Eng. Chem.*, 75, 285–295.
- M. Bui, C.S. Adjiman, A. Bardow et al., 2018, Carbon capture and storage (CCS): The way forward, *Energy Environ. Sci.*, 11, 1062–1176.
- T. Damartzis, A.I. Papadopoulos, P. Seferlis, 2014, Optimum synthesis of solvent-based post-combustion CO<sub>2</sub> capture flowsheets through a generalized modeling framework, *Clean Technol. Environ. Policy*, 16, 1363–1380.
- M. Dimoliani, A.I. Papadopoulos, P. Seferlis, 2021, Modeling and Parametric Investigation of Rotating Packed Bed Processes for CO<sub>2</sub> Capture and Mineralisation, *Chem. Eng. Trans.*, 88, 187–192.
- M.M.F. Hasan, E.L. First, F. Boukouvala, C.A. Floudas, 2015, A multi-scale framework for CO<sub>2</sub> capture, utilization, and sequestration: CCUS and CCU, *Comput. Chem. Eng.*, 81, 2–21.
- O.A. Jimoh, K.S. Ariffin, H.B. Hussin, A.E. Temitope, 2018, Synthesis of precipitated calcium carbonate: a review, *Carbonates and Evaporites*, 33, 2, 331–346.
- H. Jung, D. Im, S. Heo, B. Kim, J.H. Lee, 2020, Dynamic analysis and linear model predictive control for operational flexibility of post-combustion CO<sub>2</sub> capture processes, *Comput. Chem. Eng.*, 140, 106968.
- S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, 1983, Optimization by simulated annealing, *Science*, 220, 671–680.
- D.C. Miller, D. Agarwal, D. Bhattacharyya et al., 2017, Innovative Computational Tools and Models for the Design, Optimization and Control of Carbon Capture Processes, in: Papadopoulos, A.I., Seferlis, P. (Eds.), *Process Systems and Materials for CO<sub>2</sub> Capture*, John Wiley and Sons Ltd, Chichester, West Sussex, 311–339.
- W.C. Rheinboldt, J.V. Burkardt, 1983, A Locally Parameterized Continuation Process, *ACM Trans. Math. Softw.*, 9, 215–235.
- M.H. Sahraei, L.A. Ricardez-Sandoval, 2014, Simultaneous design and control of the MEA absorption process of a CO<sub>2</sub> capture plant, *Energy Procedia*, 63, 1601–1607.
- P. Seferlis, J. Grievink, 2004, Process design and control structure evaluation and screening using nonlinear sensitivity analysis, *Comput. Aided Chem. Eng.*, 17, 326–351.
- N. Vasilas, A.I. Papadopoulos, L. Papadopoulos et al., 2022, Approximate computing, skeleton programming and run-time scheduling in an algorithm for process design and controllability in distributed and heterogeneous infrastructures, *Comput. Chem. Eng.*, 164, 107874.
- A. Wächter, L.T. Biegler, 2006, On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, *Math. Program.*, 106, 25–57.

# Accurately Modeling Hydropower in the USA

Amanda Farnsworth,<sup>a,b</sup> Emre Gençer,<sup>a</sup>

<sup>a</sup> MIT Energy Initiative, Massachusetts Institute of Technology, 400 Main Street,  
Cambridge, MA 02142, USA

<sup>b</sup> Department of Chemical Engineering, Massachusetts Institute of Technology, 25 Ames  
Street, Cambridge, MA 02142, USA

## Abstract

This study marks the first comprehensive evaluation of hydropower across the US. Both installation limits and capacity factor profiles were extracted and processed from a variety of sources and incorporated into a capacity expansion model. Life-cycle assessment strategies were used to fairly assesses the three types of hydropower: conventional, run-of-river, and pumped storage. When only operational emissions are considered, 30-63% of system emissions go unaccounted for. As emissions caps become stricter, hydropower is relied upon more heavily. At a 50 gCO<sub>2</sub>-eq/kWh emissions limit, seven out of nine regions analyzed reach their installation limit in at least one hydropower technology type. Lastly, a carbon neutral power sector requires carbon negative technologies.

**Keywords:** hydropower, LCA, optimization, decarbonization, power sector

## 1. Introduction

Currently in the US, hydropower capacity is 80.25 GW, outputting 274 TWh of energy annually. This contribution accounts for 6.7% of generation capacity, and 6.6% overall power production and 38% of renewable power production (Uría-Martínez 2021). Around this technology, there is a heated debate regarding its value and impacts. It has been shown that hydropower can be leveraged to lower the cost of decarbonization, where hydropower integration provides increasing value to the system as decarbonization targets become more stringent (Dimanchev 2021, Bain 2018). Conversely, another faction of researchers emphasize that hydropower can damage the social and ecological ecosystems of a region, and that these impacts are not accurately captured in power system expansion and planning (Williams 2020, Moran 2018). This study aims to fairly assess the value of hydropower while penalizing it appropriately for its direct and indirect emissions.

The term “hydropower” primarily refers to a dammed river because the vast majority of hydropower resources are this type. However, hydropower technologies can fit into 3 main categories. Dammed hydropower will hereafter be referred to as “conventional.” Run-of-river hydropower (RoR) does not incorporate any water storage, so power generation is intermittent and non-dispatchable as it depends solely on river flowrate. Pumped hydropower storage (PHS) requires two nearby reservoirs at varying heights. Water is pumped up or drained down to the respective higher or lower reservoir to convert electricity to and from gravitational potential energy.

This study uses a capacity expansion model (CEM) to optimize power sector buildout by minimizing system cost while simultaneously reaching specific decarbonization targets. Life Cycle Assessment (LCA) and Technoeconomic Analysis (TEA) techniques are used



to track emissions and cost values throughout the life cycle. The last consideration of this study is that hydropower capacity buildout is limited in each region based on local geography and weather patterns.

## 2. Methodology

This section focuses on hydropower-related data and constraints, rather than overall model structure.

### 2.1. Ideal Grid

Ideal Grid (IG) is a capacity expansion model built within MIT's Sustainable Energy System Analysis Modelling Environment (SESAME). SESAME is a pathway analysis tool which utilizes LCA and TEA techniques to estimate the emissions and costs of producing a particular unit of product within the overall energy system, such as steel, hydrogen, gasoline, etc (Gençer 2020). IG references the electricity production pathways, so external LCA and TEA evaluation of technologies is not needed. Eleven different generator types are represented: wind, solar, RoR hydropower, conventional hydropower, 3 types of natural gas (combustion turbine, combined cycle, and combined cycle with carbon capture), 3 types of coal (boiler, IGCC, and boiler with carbon capture), and nuclear fission. Two different storage options are included: lithium-ion batteries (LIB), and PHS. Single-nodal, deterministic optimization is performed over a year, at an hourly timestep. Figure 1 shows the nine regions explored, roughly along the North American Electric Reliability Corporation boundary lines for convenience.

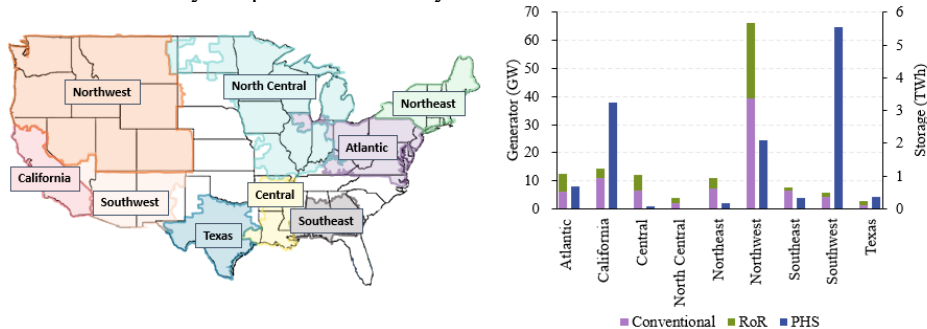


Figure 1. Regions and their respective hydropower capacity limits

### 2.2. Regional Capacity Limits

As mentioned above, hydropower adds an additional constraint to the optimization because its buildout is limited. Conventional hydropower is estimated by obtaining current capacity values from *Electric Power Annual* and adding upgrade and non-powered potentials from *Hydropower Vision* and *An Assessment of Energy Potential at Non-Powered Dams in the US*, respectively. Data is provided on a state-by-state basis, but is easily aggregated to represent regional potential. RoR regional potential is sourced from *New Stream-reach Development* assessment. This comprehensive review identified potential RoR hydropower sites across the US while eliminating sites which are close to national parks, wild and scenic rivers, and wilderness areas. Again, capacity estimates are provided in a state-by-state basis and aggregated to represent regional potential. PHS potential can be downloaded from NREL's *Closed-Loop Pumped Storage Hydropower Resource Assessment for the US*. Sites are aggregated into regional groups based on their latitudes and longitudes. The generation and storage potential of each region can be compared in Figure 1.

2.3. Hydropower Generation Profiles

Conventional hydropower generation is a dispatchable generation source, but must be balanced periodically to account for reservoir size. To account for this, monthly capacity factors (CFs) are calculated based on existing nameplate hydropower capacity and monthly power output, both sourced from EIA’s *Electric Power Monthly* reports. Equation (1) shows the simple relation used to track monthly CF, and Equation (2) shows how this CF is enforced within the model

$$CF_{m,y} = \frac{E_{m,y}}{24 * n_m * C_y} \quad \forall m \in jan, feb \dots dec \wedge y \in 2007, 2008 \dots 2013 \quad (1)$$

$$CF_{m,y} = \frac{\sum_{d=st}^{end+24} E_d}{C} \quad \forall m \in jan, feb \dots dec \wedge y \in 2007, 2008 \dots 2013 \quad (2)$$

where  $E$  is energy output in MWh,  $C$  is capacity in MW,  $n$  is days in the month, and  $st$  and  $end$  reflect the first and last day of each month.

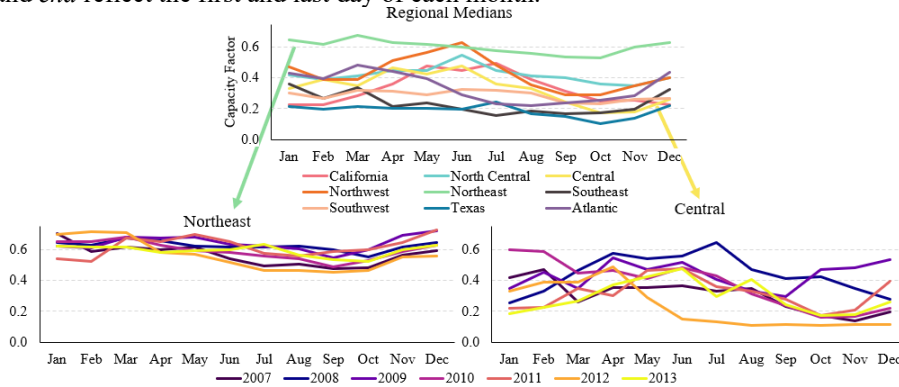


Figure 2. Conventional Hydropower CF Profiles

Figure 2 compares the conventional hydropower CF trends for all nine regions. Table A shows the standard deviation in CFs for each region, which reflects conventional hydropower resource predictability. Figure 2 also includes the yearly variation for the most and least predictable regions. This variability trend has a very low correlation with installed capacity with an  $R^2$  value of 0.1046, meaning that site aggregation is not the main reason for this high range of CF consistency.

Table A. Standard Deviation of CF in Each Region

	Atlantic	California	Central	North Central	Northeast	Northwest	Southeast	Southwest	Texas
Conventional	7.4%	8.8%	12.1%	9.7%	5.2%	6.7%	9.9%	5.4%	9.8%
RoR	17.6%	10.7%	13.8%	12.4%	9.9%	7.8%	17.6%	8.1%	24.6%

RoR CF curves are calculated based off daily river flowrate data from the United States Geological Survey (USGS). USGS provides data on over 1.9 million water-resources within the US. Each stream-site is aggregated into its respective region to produce a daily conglomerate river flowrate in each region. Equation (3) shows the calculation used to convert regional river flowrate to daily CF

$$CF_d = \begin{cases} \frac{Q_d}{Q_{max}} * \eta_{turb} & \text{for } Q_d \leq Q_{max} \\ \eta_{turb} & \text{for } Q_d > Q_{max} \end{cases} \quad (3)$$

where  $Q_a$  is hourly flowrate,  $Q_{max}$  is the flowrate which 30% of the hours exceed, and turbine efficiency,  $\eta_{turb}$ , is 85%. The values calculated with this method match the estimated values published in the *New Stream-reach Development* assessment with less than 15% discrepancy. Figure 3 shows regional RoR CF trends. Also, the yearly variation in the most and least predictable regions is shown, based on average standard deviation values provided in Table A. Just as in the conventional hydropower case, available river data is not the cause of variation, resulting in an  $R^2$  value of 0.0190. Lastly, PHS is assumed to have dispatchable operation, limited only by energy and power capacities.

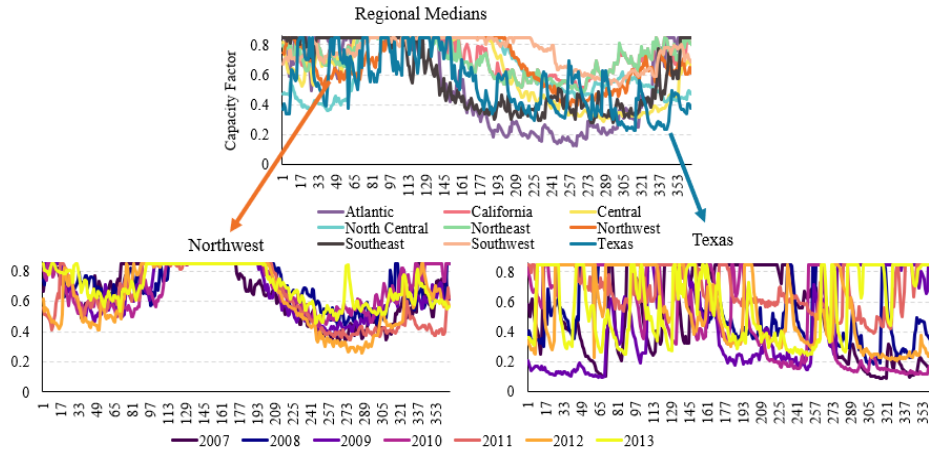


Figure 3. RoR Hydropower CF Profiles

#### 2.4. Solar and Wind Generation Profiles and Demand Profiles

Regional solar and wind CF curves are generated through the aggregation of 169 equidistant site-specific CF profiles in each region, spaced 30 miles apart. Hourly site-specific solar and wind outputs are sourced from ZEPHYR (Zero-emissions Electricity system Planning with Hourly operational Resolution) repository. ZEPHYR accesses weather data from NREL's *Wind Integration National Dataset* (WIND) Toolkit and NREL's *National Solar Radiation Database* (NSRDB), and calculates power output.

2021 regional demand profiles are sourced from the EIA *Hourly Electric Grid Monitor*. Demand curves were scaled down to 8760 kWh annually so that decarbonization strategies can be directly compared inter-regionally. This also allows for an easier evaluation of overbuilding where a generation capacity of 1kW means that all generators are operating at 100% capacity factor for all hours and no electricity is curtailed. This is only possible if demand is perfectly flat, or if there is ample energy storage without the consideration of charging/discharging or parasitic efficiency losses.

#### 2.5. LCA and TEA

Both construction and operational costs and emissions are considered in this model, so as to fairly assess hydropower. Most importantly, an estimation for emissions from inundated areas from conventional hydropower are tracked. Figure 4 maps the emissions-cost space of generators and storage technologies based on the construction and operational emissions and cost values which are combined based on median CF for renewables and 100% CF for fossil fuel generators.

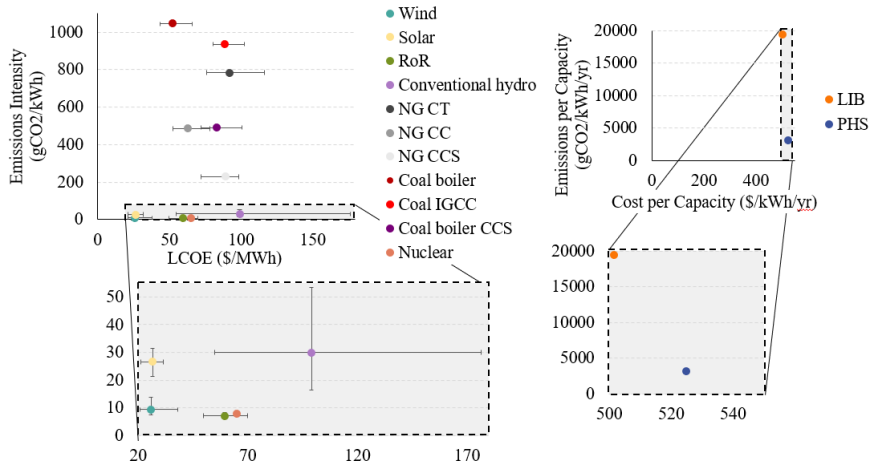


Figure 4. Mapping the Cost-Emissions Space for Generators and Energy Storage Technologies

Upon first glance, it seems that solar, wind, RoR, and nuclear are far better decarbonization options than conventional hydropower or any fossil fuel generators. This comparison is not fair because conventional hydropower and fossil fuel generators provide dispatchability to the grid which reduces energy storage requirements and limits overbuilding and curtailed energy. Additionally, fossil fuel generators costs and emissions are largely based on operation, so adjusting the power output level of these resources greatly impacts their location on the Figure 4 mapping.

### 3. Case Study

A simple case study illustrates the importance of LCA in assessing hydropower, and other decarbonization options. Figure 5 compares decarbonization strategies depending on differing emissions tracking schemes. First, it is clear that ignoring emissions from stages of the life-cycle outside operation allows for a significant increase in emissions, ranging from 30-63% increase depending on the region. The majority of this decarbonization from LCA enforcement comes from a reduction of natural gas operation, and carbon capture technologies introduced in the Atlantic, Southeast, Southwest, and Texas. Wind and solar technologies are increased by up to 25%, seeing a decrease in only the Californian region. This standalone decrease is due to the introduction of nuclear which supplies 18% of noncurtailed electricity. This shift to LCA increases system price by 1.7-3.9% depending on region. Lastly, it is shown that LIBs are not installed until PHS capacity is maximized.

LCA increases system reliance on hydropower in all regions, causing seven out of the nine regions to maximize their installation capacity in at least one of the hydropower technologies. When tracking only operational emissions, conventional hydropower is maxed out in three regions (Atlantic, North Central, and Northeast), but when LCA is applied, it is maximized in three additional regions (Central, Southeast and Southwest), while also increasing by significant amounts in two others (California and Northwest). LCA increases RoR in California and the Southwest. Also, LCA increases PHS by a significant amount in almost every region. When regional hydropower capacity limits are ignored, installed hydropower generator capacity increases by 102% and PHS decreases by 16%, resulting in a price decrease of about 1%. Lastly, this study shows that a truly carbon neutral power sector requires some carbon capture or carbon negative technologies.

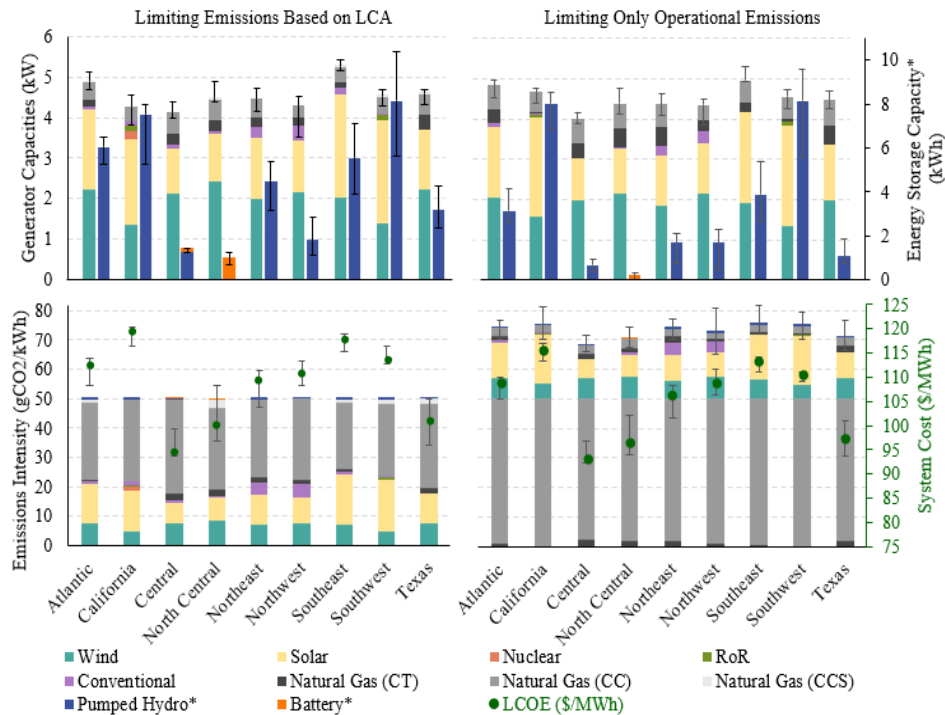


Figure 5. Decarbonization Strategies Depend on Emissions Tracking

#### 4. Conclusion

A wide range of sources are used to comprehensively evaluate hydropower resources across the USA. Regional capacity limits and performance profiles are identified for conventional, RoR, and PHS resources. LCA techniques are used to show that the optimal decarbonization strategy for every region involves the leveraging of at least one hydropower resource. Also, only monitoring operational emissions can lead to up to 63% error. Lastly, carbon negative technologies are needed to reach net neutrality.

#### References

- Bain, Dominique M, and Thomas L Acker. 2018. "Hydropower Impacts on Electrical System Production Costs in the Southwest United States." <https://doi.org/10.3390/en11020368>.
- Dimanchev, Emil G., Joshua L. Hodge, and John E. Parsons. 2021. "The Role of Hydropower Reservoirs in Deep Decarbonization Policy." *Energy Policy* 155 (August): 112369. <https://doi.org/10.1016/j.enpol.2021.112369>.
- Gençer, Emre, Sarah Torkamani, Ian Miller, Tony Wenzhao Wu, and Francis O'Sullivan. 2020. "Sustainable Energy System Analysis Modeling Environment: Analyzing Life Cycle Emissions of the Energy Transition." *Applied Energy* 277 (November). <https://doi.org/10.1016/j.apenergy.2020.115550>.
- Moran, Emilio F., Maria Claudia Lopez, Nathan Moore, Norbert Müller, and David W. Hyndman. 2018. "Sustainable Hydropower in the 21st Century." *Proceedings of the National Academy of Sciences of the United States of America* 115 (47): 11891–98. <https://doi.org/10.1073/pnas.1809426115>.
- Uriá-Martínez, Rocío, Megan M. Johnson, Rui Shan. 2021. "U.S. Hydropower Market Report."
- Williams, Jessica M. 2020. "The Hydropower Myth." *Environmental Science and Pollution Research* 27 (12): 12882–88. <https://doi.org/10.1007/s11356-019-04657-6>.

## Biogas valorisation to liquid fuels: modelling and setting up process targets.

Selusiwe Ncube,<sup>a\*</sup> Baraka C Sempuga,<sup>a</sup> Xinying Liu<sup>a</sup>

<sup>a</sup>*Institute for the Development of Energy for African Sustainability (IDEAS) Research Unit, University of South Africa (UNISA), Florida, Johannesburg, 1710, South Africa.*

### Abstract

The concept of attainable region and targeting techniques is used to explore the feasibility of co-producing Fischer-Tropsch (FT) fuels and methanol through biogas valorisation. The interest is particularly in sub-regions with a net consumption of CO<sub>2</sub> as we are aiming for a carbon negative process. One of the requirements to achieve a net CO<sub>2</sub> consumption is to be able to fully integrate the energy within the process. Heat integration between the reformer and the synthesis reactors is one of the factors that cause significant deviations from the target, especially the energy target. Adding large amounts of feed to compensate for the heat required would not be practical and the process would be a net CO<sub>2</sub> producer instead of a net CO<sub>2</sub> consumer. Comparatively, a heat pump uses as less power as possible. Thus, we propose the addition of a heat engine to the simulation process to maximise the efficiency of the process and improve CO<sub>2</sub> utilisation. A heat pump is needed to pump heat between the FT and methanol synthesis section to the reformer because of the temperature gradient. We compute the theoretical targets using MATLAB and then validate the theoretical solutions by modelling the processes using Aspen Plus®. A match of the mass balance simulation results with the theoretical stoichiometry showed a good correspondence. For heat integration, it was shown that 12.19 kW of heat can be absorbed from the synthesis reactor at 230 °C to put in 13.81 kW of work for the supply of 26 kW of heat required at the reformer which is at 801 °C for a carbon negative process.

**Keywords:** Biogas, Targeting techniques, Fischer-Tropsch fuels, Methanol, Simulation

### 1. Introduction

The abundance of carbon dioxide and methane either as flared gases or emissions from industrial sectors propels their use as raw materials. These greenhouse gases can be utilised in the form of biogas, a low-cost renewable resource [1]. There is a drive for converting CO<sub>2</sub> to fuel and chemicals such as methanol, FT (Fischer-Tropsch) fuels and ethanol [2], [3]; hence biogas is a cost and energy effective source of CO<sub>2</sub> ready to be utilised. However, the efficiency of the processes involved in valorising biogas to different fuels has to be taken into account while ascertaining that the process is carbon negative (net CO<sub>2</sub> consumption). The simulation process is a facile way to determine the heat or work that is required or can be expended by a process. One way in which work can be transferred to processes to make them feasible without altering their net energy requirement is using the Carnot temperature concept [4]. This paper unfolds from a process synthesis perspective through the use of graphical targeting techniques where we set theoretical targets in MATLAB to valorise biogas to FT fuels and methanol. The study

is substantiated by a modelling process in Aspen Plus® to validate the theoretical targets and a heat pump is introduced for heat and work integration purposes to try and improve the efficiency of the process while maintaining an overall carbon negative process.

## 2. Methodology

The methodology section is divided into two sections that include the theoretical and the simulation.

### 2.1. Setting up process targets

We first look at the theoretical approach which allows us to optimise certain targets called the Attainable Regions (AR) using MATLAB. We regard the overall process as a ‘black box’ by looking at only the input and output of the system as shown in Figure 1.

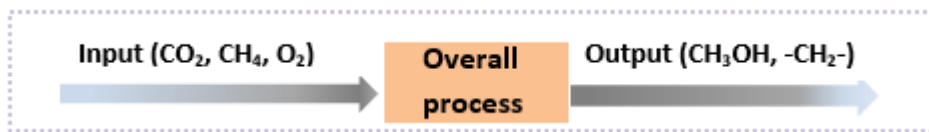


Figure 1: A schematic representation of the overall process

The basis for the production of the products (methanol and FT fuels) was set as 1 mol of methane. The targets identified in MATLAB were then simulated in Aspen Plus®.

### 2.2. Simulation process

The NRTL-RK thermodynamic property method was used for all the reactors used in the simulation process to allow for an accurate description of the phase equilibrium. The reformer was simulated as an equilibrium reactor, an RGibbs reactor model. The methanol synthesis reactor was modelled as an RPLUG model using a kinetic model derived from Skrzypek et al. [5]. The model has the form of the Langmuir-Hinshelwood–Hougen–Watson (LHHW) reaction Figure 2 illustrates the implemented flowsheet.

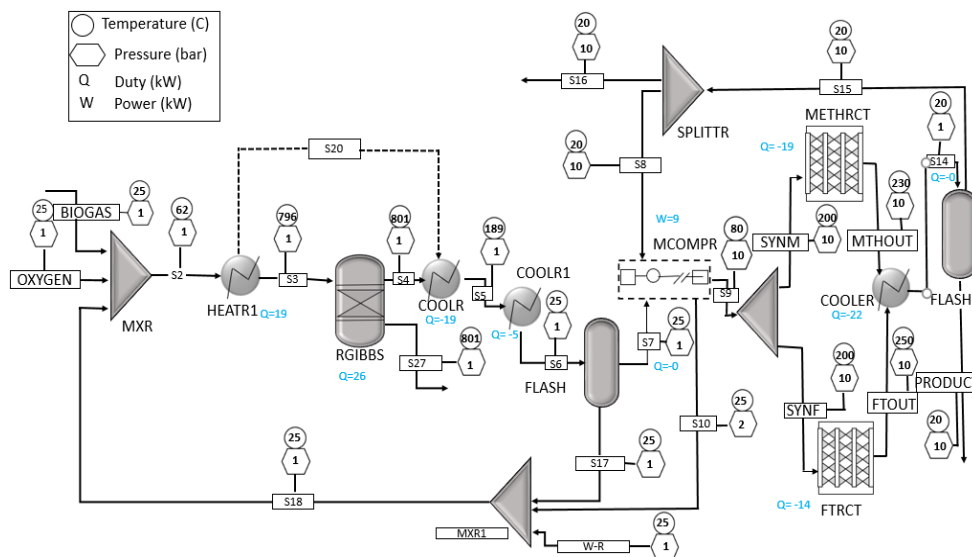


Figure 2: Process flow diagram for the synthesis of methanol and FT from biogas.

### 2.2.1. Heat integration computations

A heat pump will be included in the simulation process to integrate the heat between the reformer and the synthesis reactor. To use a heat pump, it is imperative to determine the amount of work needed to put into the pump to draw heat at a lower temperature ( $T_c$ ) and pump the resultant heat to a much higher temperature ( $T_h$ ) to supply the required heat. Given  $Q_h$  (heat load),  $T_h$  and  $T_c$  obtained from the simulation, the work ( $W$ ) required is calculated as shown in Equation 1.1.

$$W = Q_h * \left(1 - \frac{T_c}{T_h}\right) \quad (1.1)$$

The amount of heat ( $H$ ) to be absorbed from the synthesis reactor is computed as shown in Equation 1.2.

$$H = Q_h - W \quad (1.2)$$

## 3. Results and Discussion

Figure 3 shows the different production rates of  $CH_2$  where  $CH_3OH$  production is maximized using the MATLAB software. Point A illustrates the desired target.

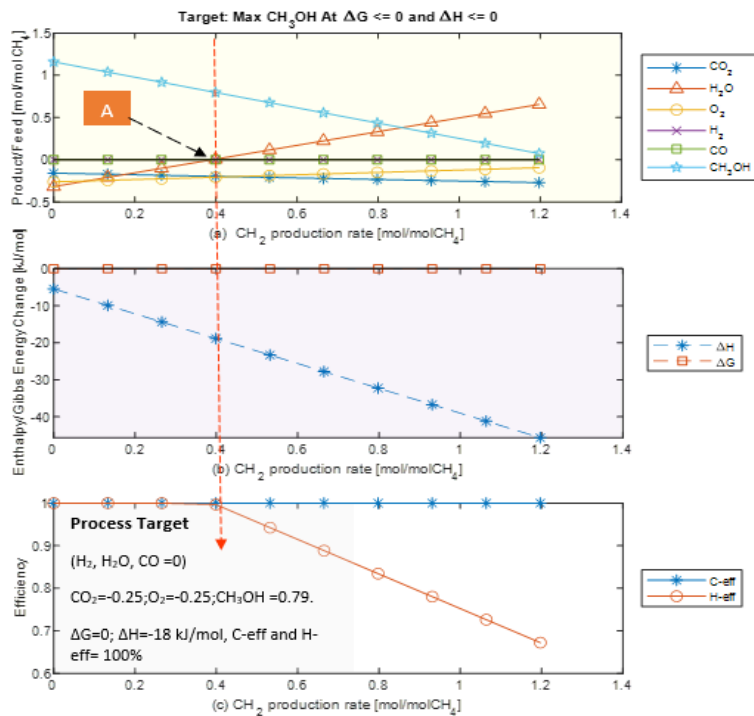


Figure 3: Process material balance targets (a) and their associated energy and work balances (b), and carbon and hydrogen efficiencies (c) for biogas conversion to FT-fuel at various production rates per mole of  $CH_4$  consumed and maximum  $CH_3OH$  consumption.



Tables 1 and 2 show the results obtained from MATLAB (Target) and Aspen Plus®.  
Table 1: Global Mass balance

Component	Simulation		Simulation (Net)(kmol/h)	Target (kmol/h)
	Inlet (kmol/h)	Outlet (kmol/h)		
CO	0.00	0.00	0.00	0.00
CO <sub>2</sub>	0.25	0.11	-0.14	-0.25
H <sub>2</sub> O	0.27	0.27	0.00	0.00
O <sub>2</sub>	0.25	0.00	-0.25	-0.25
CH <sub>4</sub>	1.00	0.05	-0.95	-1.00
CH <sub>3</sub> OH	0.00	0.73	0.73	0.79
-CH <sub>2</sub> -	0.00	0.31	0.31	0.40

Table 2: Global Energy balance

	Heat Duty/Energy (Q) (kW)		Power (kW)	
	Theoretical Target (MATLAB)	Simulation	Theoretical Target (MATLAB)	Simulation
Net Energy/Power	-18	-34	0	9

The simulation results achieved 78% of the theoretical target of -CH<sub>2</sub>- production while 92% was achieved for methanol production indicating a small deviation from the theoretical targets as shown in Table 1. As per the target, there was net CO<sub>2</sub> consumption. However, the slight discrepancy was that only 0.14 moles of CO<sub>2</sub> were consumed instead of 0.25 moles due to limitations in driving forces such as entropy generation and temperature approaches across heat exchangers. Considerable deviations were observed in the energy balance between the simulation (net) and the target results from MATLAB as shown in Table 2. The theoretical targeted  $\Delta\hat{H}$  value was -18 kJ/mol. Therefore, it is critical to ensure that  $\Delta H$  is recovered as work, which is high-quality energy. If it is recovered as heat, the quality of the recovered energy is poor, and the process is inefficient. However, the theoretical value for work ( $\Delta\hat{G}$ ) was 0 kJ/mol and a value of 9 kW was obtained from the simulation process as a result of the work required by the compressor (MCOMPR) to supply high pressure needed at the methanol synthesis reactor to increase chemical driving reaction forces which are one of the major sources of entropy generation and thus work losses. To supply some of the work needed by the process, a heat pump is needed to integrate the temperature gradient from low to high instead of burning an external energy source such as CH<sub>4</sub> in order to generate enough work. The consequence of burning CH<sub>4</sub> is that the process would be a net CO<sub>2</sub> producer instead of a net CO<sub>2</sub> consumer. As an optimisation strategy, the heat generated in situ was used for heat integration within the simulation process. Heat integration was carried out between

the two heat exchangers at the inlet and outlet of the reformer and the minimum temperature approach was set at 5°C (801°C-796°C) as shown in Figure 4.

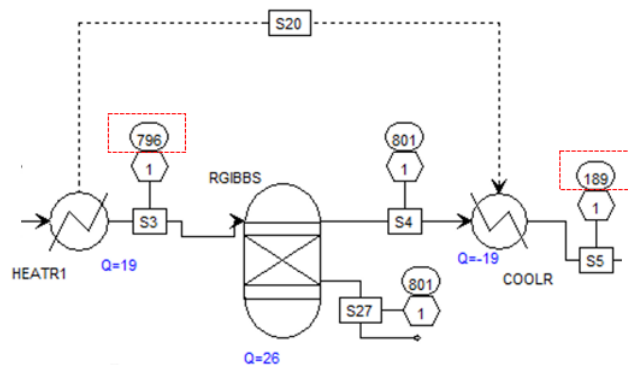


Figure 4: Heat integration between the heater and the cooler in the reformer unit (extracted from the process flowsheet in Aspen Plus®).

The overall process consisted of the reforming and the synthesis stage that could be optimised for energy integration.

- The reforming stage (RGIBBS) is endothermic at high temperatures hence the heat carries work into the process.
- The synthesis stage (METHRCT) is exothermic at lower temperatures and work is rejected from the process.

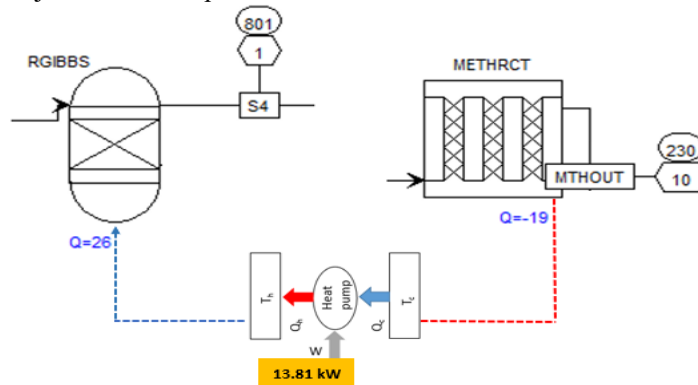


Figure 5: Heat integration between the reformer and the synthesis reactor using a heat pump (extracted from a process flowsheet in Aspen Plus®).

One would need a heat pump to integrate the heat between the reformer and the synthesis reactor because of the temperature gradient as a hot stream at 230 or 250 °C cannot be used to heat a cold stream at 801 °C as shown in Figure 5. Therefore, work is needed to pump heat in the opposite direction of its normal flow. Therefore, from the computations (Equations 1.1 and 1.2), 12.19 kW of heat can be absorbed from the synthesis reactor at 230 °C to put in 13.81 kW of work for the supply of 26 kW of heat required at the reformer which is at 801 °C. It should be noted that from a practical point of view, there are currently no heat pump technologies that are capable of reaching such high temperatures;

however, we seek to determine what is required in order to achieve the heat integration within the process or a carbon negative overall process.

#### 4. Conclusions

Simulated mass balance targets extrapolated to conventional methanol and FT fuels synthesis conditions were in good agreement with the results obtained from the theoretical targets. Of particular significance is that the process revealed a net CO<sub>2</sub> consumption, meaning that the overall process is carbon negative. However, to achieve a zero net CO<sub>2</sub> consumption negative process, it is necessary to fully integrate energy within the process. In this case, there is a need to integrate the heat and work between the synthesis reactors (CH<sub>2</sub> and Methanol) and the reformer. However, a heat pump is needed to overcome the heat flow temperature gradient from low to high. The simulation reveals that in order to get close to the target, external work supply to the process is required. This is one of the major deviations from the target which is 0 work requirement due to the work needed at the heat pump to integrate heat between the low temperature synthesis section to the high temperature reformer section; also due to the high pressure required at the methanol synthesis reactor which is essentially reaction driving forces to obtain high conversions for methanol synthesis. Thus, the additional work must be supplied from a renewable source to achieve a truly carbon negative process. In essence, heat integration between the reformer and the synthesis reactors is one of the factors that causes significant deviations from the target, especially the energy target, and therefore, should be a subject of optimisation. These limitations are a result of lack of better technologies or better catalysts to improve the efficiencies in processes. The discrepancies in this work can be taken as a reference point for future work.

#### References

- [1] Rahul Kadam, N. L. Panwar, "Recent advancement in biogas enrichment and its applications," *Renew. Sustain. Energy Rev.*, vol. 73, no. September 2016, pp. 892–903, 2017, doi: 10.1016/j.rser.2017.01.167.
- [2] Daniele Previtali, Antonio Vita, Andrea Bassani, Cristina Italiano, André Furtado Amaral, Carlo Pirola, Lidia Pino, Alessandra Palella, and others, "Methanol synthesis: A distributed production concept based on biogas plants," *Chem. Eng. Trans.*, vol. 65, pp. 409–414, 2018, doi: 10.3303/CET1865069.
- [3] Borja Hernandez, Mariano Martin, "Optimization for biogas to chemicals via tri-reforming. Analysis of Fischer-Tropsch fuels from biogas," *Energy Convers. Manag.*, vol. 174, no. August, pp. 998–1013, 2018, doi: 10.1016/j.enconman.2018.08.074.
- [4] Baraka Celestin Sempuga, Diane Hildebrandt, Bilal Patel, David Glasser, "Work to chemical processes: The relationship between heat, temperature, pressure, and process complexity," *Ind. Eng. Chem. Res.*, vol. 50, no. 14, pp. 8603–8619, 2011, doi: 10.1021/ie2004785.
- [5] J. Skrzypek, M. Lachowska, H. Moroz, "Kinetics of methanol synthesis over commercial copper/zinc oxide/alumina catalysts," *Chem. Eng. Sci.*, vol. 46, no. 11, pp. 2809–2813, 1991, doi: 10.1016/0009-2509(91)85150-V.

# Modeling Commercial Fleet Charging and Regulation Strategies via Data-driven Optimization

James Owens<sup>a,b</sup>, Emre Gençer<sup>a</sup>

<sup>a</sup>*MIT Energy Initiative, Massachusetts Institute of Technology, 400 Main Street, Cambridge, MA 02142, USA*

<sup>b</sup>*Department of Chemical Engineering, Massachusetts Institute of Technology, 25 Ames Street, Cambridge, MA 02142, USA*

## Abstract

While vehicle electrification offers an opportune path to transportation decarbonization, current tariffs can make the transition cost prohibitive for commercial fleets. To this end, providing frequency regulation has been proposed as a means for fleets to generate revenues while parked. However, regulation dispatch uncertainty can make participation difficult for fleets with strict schedules, and assuming a deterministic dispatch profile tends to overestimate a fleet's flexibility and underestimate the financial risk of its bids. Moreover, the most efficient retail rate and strategy for providing regulation, or simply charging alone, highly depends on fleet schedule and demand relative to other loads a fleet owner may have. Thus, we pose a two-stage, stochastic model, informed by real-world data, to realistically determine optimal charging and regulation behavior under different tariffs. A case study demonstrates model utility in analyzing a last-mile delivery truck fleet housed at a large shipping warehouse.

**Keywords:** vehicle, electrification, grid, tariffs, optimization

## 1. Introduction

As the world marches toward a net-zero future, transportation is poised to electrify. Apart from personal vehicles, light truck and bus fleets are anticipated to undergo a rapid transition. In addition to reducing fossil fuel consumption, electrified truck fleets can provide demand response and ancillary services to the grid, known as "Vehicle-to-Grid" (V2G). For example, commercial delivery fleets are prime V2G candidates given their large batteries, predictable schedules, and overnight co-location at shipping depots. Moreover, V2G revenues can be used to offset costs of fleet electrification. For the grid, V2G increases system flexibility, ultimately reducing firm generator emissions (Owens, 2022). However, traditional U.S. retail rates are not conducive to affordable EV charging and V2G for commercial entities. Monthly demand charges, which recoup fixed delivery costs and are determined by peak consumption over a 15-minute period, can represent as much as 75% of an EV fleet operator's bill despite relatively low consumption and demand peaks the remainder of the month (He, 2019). This can discourage charging or performing V2G at opportune times, or even make electrification prohibitively expensive.

While literature has probed the issue, limitations include (1) A lack of real regulation data or ignoring uncertainty, which overestimates performance, and (2) Optimizing charging subject to current retail rates, which limits cost-optimal outcomes. For example, DeForest et al. (2018) pose a charging and ancillary service optimization model for a 29-vehicle fleet housed within a U.S. Air Force base. The authors conduct scenario analysis to assess

optimal bidding strategy and its sensitivities to resource utilization, local load conditions, and retail electricity prices. However, the study’s deterministic approach and parametric nature limit insights into hedged bid behavior and market outcomes. Likewise, Al-Hanahi et al. (2022) pose an integrated charging and routing optimization problem for commercial vehicles traveling to both public and depot charging stations. The model considers both demand charges and time-of-use tariffs, but also takes a deterministic approach and does not consider ancillary service participation. Neither study considers the implications of alternative tariffs.

Here we pose a stochastic, two-stage charging and frequency regulation (the ancillary service with the most earning potential) bidding model to more accurately assess realistic smart charging and regulation opportunities for commercial EV fleets. Of particular note, the model considers the risks and rationale of a standalone fleet owner and utilizes random regulation signals derived from actual EVs, which are a significant source of uncertainty and key to informing hedged regulation bids (DeForest, 2018). The mixed-integer model, subject to different market prices and rules, makes depot level and single vehicle decisions over multiple time scales to minimize costs and maximize frequency regulation revenues. Informing the high-resolution model with real world delivery truck data and the nuances of corresponding warehouse load profiles, we demonstrate the model’s utility by evaluating different EV charging tariffs – from longstanding industrial rates to newly proposed, separately metered EV rates – and how they may enable V2G.

## 2. Methodology

The depot-centered model framework is tariff and fleet-agnostic, making it extendable to several combinations of fleets, building loads, time horizons, and charging schemes. As an illustrative case study, we formulate the optimization to minimize monthly demand charges and net energy costs for the overnight charging of a last-mile delivery fleet. Given the “return-to-base” nature and predictability of these fleets, the model uses data from an average summer day in 2022 and optimizes decisions for a single overnight (one day-ahead market settlement) period and assumes known vehicle arrival states. In the day-ahead (DA) stage, depot-wide charging demand is scheduled and regulation bids are submitted on behalf of the entire depot. The recourse stage sees to real time (RT) charge and capacity management, subject to stochastic regulation signals. For model tractability, the RT step resolution is taken to be 1-minute. Since regulation capacity is bid by the hour (market rules) and baseline charging is scheduled in 15-minute intervals, the DA capacity commitments and charging plan are mapped to the RT scale with additional constraints.

### 2.1. Nomenclature

#### Indices and Sets

$h/t$	Time step index
$T_H, T_b, T_m$	Optimization horizon in hours, quarter-hours (b), minutes (m)
$i/V$	Vehicle index / Set of vehicles
$s/S$	Scenarios index / Set of scenarios

#### Parameters

$\omega_s$	Probability of scenario $s$
$p_D(t), p_U(t)$	DA regulation prices, down, up (\$/kW)
$p_{E,ret}(t), p_{E,inj}(t)$	Price of retail electricity, Payment for injection (\$/kWh)
$p_{RD,vio}(t), p_{RU,vio}(t)$	Costs of violating regulation cap commitment down, up (\$/kW)
$p_{S,vio}(t)$	Costs of violating SOC departure targets (\$/kWh)
$p_{deg}$	Degradation cost associated with cycling from injection (\$/kWh)

$p_{demand}(d)$	Charge for peak demand during period $d$ (\$/kW)
$f_D^s(t), f_U^s(t)$	AGC utilization factor down, up
$\eta$	One-way charger efficiency (0.94, up and down)
$P_{max}$	Maximum charge/discharge power of individual vehicle (kW)
$S_{travel}$	Energy consumed in daily travel (kWh)
$SOC_{min}, SOC_{max}$	Vehicle state-of-charge minimum, maximum (kWh)
<b>Decision Variables</b>	
$R_D(h), R_U(h)$	DA bid/commitment regulation capacity down, up (kW)
$P_{Plan}(t)$	DA scheduled charging of full fleet (kW)
$dP^s(t)$	RT deviation from $P_{Plan}$ in scenario $s$ (kW)
$P_{RT}^s(t), R_{inj}^s(t)$	RT charging, injection of full fleet in scenario $s$ (kW)
$q_i^s(t), r_i^s(t)$	RT charging, injection of vehicle $i$ in scenario $s$ (kW)
$SOC_i^s(t)$	State-of-charge of vehicle $i$ in scenario $s$ (kWh)
$R_{D,RT}^s(t), R_{U,RT}^s(t)$	RT depot reserve availability in scenario $s$ down, up (kW)
$R_{D,vio}^s(t), R_{U,vio}^s(t)$	RT violation of cap commitment in scenario $s$ down, up (kW)
$D^s(d)$	Peak demand during period $d$ in scenario $s$ (kW)

## 2.2 Objective and Constraints

The model objective maximizes revenues from participation in the wholesale frequency regulation market while minimizing real-time charging costs and demand charges.

$$\max \sum_{h \in T_H} R_D(h) * p_D(h) + R_U(h) * p_U(h) - \sum_{t \in T_b} \frac{1}{4} P_{Plan}(t) * p_{E,ret}(t) \quad (1a)$$

$$+ \sum_{s \in S} \frac{\omega_s}{60} \left( \sum_{t \in T_m} R_{inj}^s(t) * (p_{E,inj}^s(t) - p_{Deg}) - (R_{ch}^s(t) + dP^s(t)) * p_{E,ret}(t) \right) \quad (1b)$$

$$- \sum_{s \in S} \omega_s \left( S_{vio}^s * p_{S,vio} + \sum_{t \in T_H} R_{D,vio}^s(t) * p_{RD,vio}(t) + R_{U,vio}^s(t) * p_{RU,vio}(t) \right) \quad (1c)$$

$$- \sum_{s \in S} \omega_s * \sum_{d \in D} D^s(d) * p_{demand}(d) \quad (1d)$$

In the DA stage, revenue is earned by offering up and down regulation capacity and a baseline charging plan is determined (1a). It is assumed that the depot operates as a price taker, submitting quantity-only capacity offers that optimally respond to deterministic regulation price forecasts and are accepted. In RT, the depot can deviate from baseline charging or discharge to the grid to satisfy regulation dispatch signals each minute (1b). Here we consider the fact that vehicles dispatching up regulation (discharging) can satisfy other depot loads/baseloads (instantaneously reducing metered load is identical to injection in this regard) and in turn further reduce demand charges. The EVs earn the retail energy rate when covering the baseload and the wholesale price for injection that exceeds the baseload ( $p_{E,inj}^s(t)$  is governed by unlisted constraints), but also incurs the cost of accelerated aging from increased battery cycling,  $p_{deg}$  (1b). For this study, per the California Public Utilities Commission, the EVs/depot participate in the wholesale regulation market, but all energy consumed by the EVs is settled at the local retail rate (DeForest 2018). If DA capacity commitments are unmet or SOC departure targets are violated, penalty charges are incurred (1c). Demand charges are set across one or more demand periods, depending on the specific tariff (1d). Since demand charges are applied once per month, price terms (1a-1c) are all scaled to a one-month basis in the objective.

The charge/discharge decisions of individual vehicles are bound by their battery capacities and charging infrastructure as well as their state-of-charge upon arrival and energy requirements for the next day travel (Eq. 2-5).  $P_{RT}^s(t)$  and  $R_{reg}^s(t)$  are the RT charging and net regulation of all vehicles, respectively (Eq. 6-7).

$$SOC_i^s(t) = SOC_i^s(t-1) + [q_i^s(t) + r_i^s(t)]/60, t \in T_m, i \in V, s \in S \quad (2)$$

$$SOC_i^s(t_F) = SOC_i^s(t_0) + S_{travel}, t \in T_m, i \in V, s \in S \quad (3)$$

$$-P_{max} \leq q_i^s(t) + r_i^s(t) \leq P_{max}, t \in T_m, i \in V, s \in S \quad (4)$$

$$SOC_{min} \leq SOC_i^s(t) \leq SOC_{max}, t \in T_m, i \in V, s \in S \quad (5)$$

$$P_{RT}^s(t) * \eta = \sum_{i \in V} q_i^s(t), t \in T_m, s \in S \quad (6)$$

$$R_{reg}^s(t) = \sum_{i \in V} r_i^s(t), t \in T_m, s \in S \quad (7)$$

Available regulation capacity is governed by battery SOC, maximum energy capacity, charger capacity, and planned baseline charging level (Eq. 8-11). Per market regulations, the fleet must bid a minimum capacity into the regulation market,  $R_{Min}$  (i.e. 100 kW in California, binary constraints not shown for brevity), and is required to maintain the bid capacity over the entire hour in real-time. If the day-ahead commitments are not satisfied by the grid-connected vehicles in RT, the shortfall is acquired on the spot market (for the entire hour) and is reflected as a penalty in the objective ( $R_{D,viol}^s$  and  $R_{U,viol}^s$ ).

$$R_{D,RT}^s(t) \leq \sum_{i \in V} SOC_{max} - SOC_i^s(t) - P_{Plan}^s(t) + R_{D,viol}^s(t_h), t \in T_m, s \in S \quad (8)$$

$$R_{D,RT}^s(t) \leq N_V * P_{max} - P_{Plan}^s(t) + R_{D,viol}^s(t_h), t \in T_m, s \in S \quad (9)$$

$$R_{U,RT}^s(t) \leq \sum_{i \in V} (SOC_i^s(t) - SOC_{min}) + P_{Plan}^s(t) + R_{U,viol}^s(t_h), t \in T_m, s \in S \quad (10)$$

$$R_{U,RT}^s(t) \leq N_V * P_{max} + P_{Plan}^s(t) + R_{U,viol}^s(t_h), t \in T_m, s \in S \quad (11)$$

RT frequency regulation signal response is satisfied through a combination of deviation from the scheduled baseline charging,  $R_{chg}^s(t)$ , and energy injection into the depot or grid,  $R_{inj}^s$  (Eq. 13). When not providing regulation, the fleet can increase its rate of charge in real time,  $dP^s(t)$ , to meet future capacity commitments and/or SOC targets. Additional constraints govern the signs and distribution of regulation between  $R_{chg}^s(t)$  and  $R_{inj}^s(t)$ .

$$R_{D,RT}^s(t) * f_D^s(t) - R_{U,RT}^s(t) * f_U^s(t) = R_{reg}^s(t), t \in T_m, s \in S \quad (12)$$

$$R_{reg}^s(t) = R_{chg}^s(t) * \eta + R_{inj}^s(t) * (1/\eta), t \in T_m, s \in S \quad (13)$$

$$P_{RT}^s(t) = P_{Plan}^s(t) + dP^s(t), t \in T_m, s \in S \quad (14)$$

$$-P_{Plan}^s(t) \leq R_{chg}^s(t), t \in T_m, s \in S \quad (15)$$

$$-P_{Plan}^s(t) \leq dP_{chg}^s(t), t \in T_m, s \in S \quad (16)$$

Finally, for each demand period, containing time steps  $d$ , peak demand is set as a function of the net load from the depot baseload, charging, and regulation.

$$\max (P_{base}(t) + P_{RT}^s(t) + R_{reg}^s(t)) \leq D^s(d), t \in d, s \in S \quad (17)$$

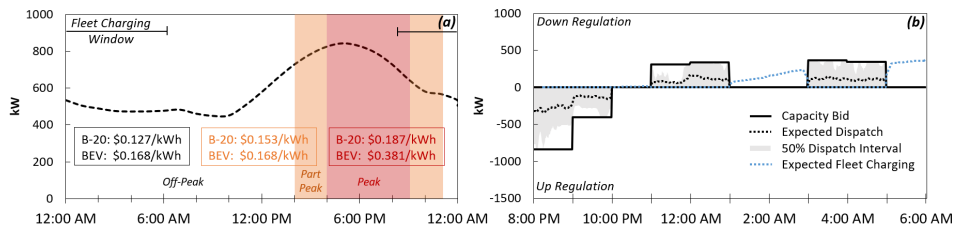
## 2.2. Stochastic Signal Data

Coupled frequency regulation signals,  $f_D^s(t)$  and  $f_U^s(t)$ , are uniformly sampled from a set of 500 synthetic 10-hour profiles generated using the SynAS Python module (Gehbauer 2020). The SynAS module was developed using 143 days of CAISO dispatch signals (4 second resolution) for a V2G field demonstration at the Los Angeles Air Force Base. The demo comprised a fleet of bi-directional EVs that bid 100 kW of symmetric regulation into the market. The synthetic profiles are statistically similar to the actual signals.

## 3. Case Study and Results

To demonstrate the model's utility, we compare both regulation opportunities and net charging costs under two commercial electricity tariffs offered by the Pacific Gas & Electric (PG&E) company in California, USA. The "B-20" is a traditional rate offered to customers with peak loads over 500 kW. Customers are billed on the basis of energy

consumption and maximum demands across “peak”, “part-peak” and “off-peak” periods, each with its own energy cost (Fig. 1a). B-20 demand charges are \$30.06/kW, \$6.13/kW, \$28.04/kW for maximum peak, part-peak, and overall demand. The BEV demand charge is \$1.91/kW for maximum overall demand only. Depot and EV loads are *not* separately metered. Designed to undue cost burdens from brief peaks in charging profiles, the “Business Electric Vehicle” (BEV) rate rolls delivery costs into volumetric energy prices with higher, more aggressive time-of-use rates, resulting in monthly charges more commensurate with overall consumption. Under this rate, EV loads are separately metered. Details of both rates can be found at [pge.com/tariffs](http://pge.com/tariffs). While the BEV rate will theoretically always outperform B-20 for standalone fleet charging, we uncover the nuances that emerge when considering corresponding warehouse/depot load profiles.



**Figure 1.** Depot baseload profile and time-of-use energy prices for each rate (a) Regulation and charging behavior for a 10-truck fleet under B-20 (Case A) (b)

We consider electrified delivery trucks charging overnight at an 800,000 square foot shipping depot/warehouse during summer months. The depot’s load profile shape and magnitude (Fig. 1a) are informed by publicly available data from the U.S. government’s National Renewable Energy Laboratory (NREL) and Energy Star program. The model’s energy prices are informed by the PG&E rates and the hourly wholesale regulation prices are the California Independent System Operator’s (CAISO) median day-ahead clearing prices for June-September 2022 (Fig. 1a). Capacity and SOC violation costs are tuned to reasonably minimize violations. All trucks are assumed to have 200 kWh batteries and 100 kW DC fast chargers, with 100-mile daily routes (~60 kWh overnight recharging for a 1.75 mi/kWh efficiency). For simplicity, all trucks are assumed to be available for charging/regulation strictly between 8PM and 6AM (based on a 75% vehicle availability threshold using NREL’s Fleet DNA data) and weekend activity is not considered. The number of EVs is varied across cases, from 10 to 50 trucks, and both “Smart Charge Only” and “Smart Charge + Regulation” scenarios are considered. The departing truck SOC is left to be optimized for the most profitable regulation capacity. When including regulation, sampling 50 regulation dispatch scenarios was found to sufficiently characterize the distribution. Key results are presented in the Table 1.

We begin our analysis examining the regulation and charging behavior of a 10 EV truck fleet under the B-20 rate (Case A, Fig. 1b). In this case, the EV charging load is small relative to depot load, meaning the fleet can leverage pre-existing peak depot demands. Fig. 1b shows the optimal DA bid profile and expected RT dispatch and charging behaviors. Up regulation is bid in the earliest hours, when prices are at their highest and the EVs still have several hours to regain discharged energy. Likewise, these bids are backed by an optimum departure SOC at full charge. Having to recover discharged energy, down regulation is bid in the hours immediately following, when overnight energy prices are low, and with the last hour left available for additional charging



corrections. Notice that existing depot peaks enable fleet charging on the order of 300 kW without incurring increased demand charges. Compared to the “Smart Charging Only” case in Table 1, regulation offers considerable net cost savings (43.4%). In addition to DA capacity revenues, savings are achieved by providing energy to the depot (reducing its overall demand) in early evening hours when retail rates are high. We also appreciate the more realistic (i.e. hedged) bidding behavior that the stochastic data offers. For instance, the expected value of perfect information (EVPI) for Case A is 452.8 – which would yield a 51.5% underestimate of net cost. In contrast to Case A, the fleet in Case B is separately metered (BEV) and thus cannot leverage the depot’s baseload peaks, which makes it the inferior option for smart charging alone. Still, the relatively low demand charges enable regulation participation that offers significant savings (63%) and net savings comparable to Case A.

**Table 1.** Scenarios Results for Monthly Fleet Charging Costs

Case	Fleet Size	Tariff	Smart Charging Only		Smart Charging + Regulation					Savings via Regulation
			Δ Demand Charge (\$)	Net Fleet Cost (\$)	Δ Demand Charge (\$)	RU Bid (MW)	RD Bid (MW)	DA Cap. Revs (\$)	Net Fleet Cost (\$)	
A	10	B-20	0	<b>1551</b>	0	1.2	1.4	508.9	<b>878</b>	<b>43.4%</b>
B	10	BEV	136	<b>2283</b>	217	0.9	0	306.3	<b>856</b>	<b>63%</b>
C	20	B-20	0	<b>3102</b>	107	2.1	0	629	<b>2265</b>	<b>27%</b>
D	20	BEV	271	<b>4566</b>	424	1.9	0	612.6	<b>1894</b>	<b>59%</b>
E	50	B-20	102	<b>8376</b>	102	0	0	0	<b>8375</b>	<b>0%</b>
F	50	BEV	677	<b>11416</b>	1039	4.7	0	1531.5	<b>4977</b>	<b>56%</b>
G (no depot)	10	B-20	2423	<b>4046</b>	2423	0	0	0	<b>4046</b>	<b>0%</b>

Cases C and D for a 20-truck fleet tell a similar story for only smart charging, but the doubled EV demand forces more conservative bids and therefore fewer net savings under the B-20 rate. We observe, however, that the V2G bids enabled by BEV rate scales linearly, with comparable overall savings due to the low marginal demand charge. We explore this trend up to 50 trucks, at which point no capacity is bid under B-20 due to the increased demand charges. Rather than fleet size, the outcome depends on vehicle loads relative to flexibility from the depot. For example, 10 standalone trucks see no regulation bids under B-20 (Case G) since all marginal demand peaks come at a prohibitive cost.

#### 4. Conclusions

A two-stage stochastic model is posed for assessing charging and regulation bid behavior under different tariffs. It is used to demonstrate that the most effective fleet charging strategies and tariffs are not one-size-fits-all, but context dependent. The model is easily adapted to other technologies and scenarios and can readily accommodate additional uncertainties (i.e. vehicle schedules, prices) in future work.

#### References

- J. Owens et al., 2022. Can vehicle-to-grid facilitate the transition to low carbon energy systems?. *Energy Advances*, 12.
- Y. He et al., 2019. Fast-charging station deployment for battery electric bus systems considering electricity demand charges. *Sustainable Cities and Society*, 48.
- C. Gehbauer et al., 2020. Synthetic CAISO Frequency Regulation Signal, *The American Modelica Conference 2020*.
- N. DeForest et al., 2018. Day ahead optimization of an electric vehicle fleet providing ancillary services in the Los Angeles Air Force Base vehicle-to-grid demonstration, *Appl. Energy*, 210.
- Al-Hanahi, Bassam, et al., 2022. An optimal charging solution for commercial electric vehicles. *IEEE Access*, 10.

# **Economic Benefits from Planned Renewable Installations in the US using Hydrogen and Modular Ammonia Production Units**

Apoorv Lal, Fengqi You

*Cornell University, Ithaca, New York 14853, USA*

## **Abstract**

Planned installations for renewable power generation are often associated with significant economic risks for the investors and thus require specific policies, incentives by the government, and risk reduction mechanisms in place. Each installation for renewable power generation has a period before its integration with the grid, which involves initiating and completing regulatory approvals and the construction phase. In the current scenario, a given installation generates no revenue before its commercial operation, even when it can produce power at less than its nameplate capacity. This study has analyzed the possibility of extracting added profits from planned installations which could negate the significant economic risks borne by the investors. Two chemical-based energy storage options, namely hydrogen and ammonia, have been evaluated in this work to harness economic benefits during the installation period of planned renewable energy systems.

**Keywords:** renewable installations, hydrogen, ammonia, modular manufacturing

## **1. Introduction**

The substantial dependence on fossil fuels to meet the ever-increasing energy demand is one of the burning questions the world is trying to answer (Kim et al., 2014). Accordingly, many countries incentivize renewable energy production to limit carbon emissions and facilitate the transition toward a more sustainable energy matrix (Tian et al., 2020). Despite the continued efforts by governments to facilitate the development of new facilities for renewable power generation (Sun et al., 2021), significant issues in operations and investment risk management, both pre and post-commercial operations, hinder an accelerated adoption of renewable power (Pistikopoulos et al., 2022). The anticipated expansion of renewable energy capacity is crippled by factors including high capital expenditure and the transient nature of renewable sources of energy (Ajagekar et al., 2022), which are considered two major roadblocks in its contribution to affordable electricity (Zhao et al. 2020). In the current scenario, before the commercial operation, a given installation generates no revenue even when it can produce power at lesser than nameplate capacity. This implies that a significant amount of renewable energy can be hedged into profits. The energy utilization can be localized at the installation sites or with minimal infrastructure to direct the power to the use site. To the best of our knowledge, previous works have

not evaluated the possibility of extracting added profits during the installation period of planned renewable energy systems. Therefore, this paper proposes a framework for added profits from planned installations using simultaneous investment in energy storage options, thus, creating incentives for early investment in planned installations in the US. Two scenarios for chemical-based energy storage, namely, hydrogen and ammonia, have been evaluated to harness economic benefits during the installation period of planned renewable energy systems.

## 2. Preliminaries

### 2.1. Electrolyzer and Hydrogen Tank

An electrolyzer (ELE) takes power as input and produces hydrogen. The electrolyzer considered in this study is an alkaline electrolyzer (AWE), a relatively mature and commercially available technology. AWE is capable of producing at large capacities with a life span of around 80,000 hrs (Bhandari and Shah, 2021). In addition to its economic viewpoint (Thomassen et al., 2019), AWE can withstand low current densities, making it flexible, so it can easily accommodate the dynamic nature of electrolyzers (Shams et al., 2021). Hydrogen produced can be stored in a hydrogen tank at high pressures, which comes in various sizes depending on the application.

### 2.2. Ammonia Production

Ammonia is an essential feedstock for fertilizer production, with the global annual demand expected to touch 230 Mt by 2025 (Apodaca, 2022). It is regarded as a potential carrier for "green hydrogen" owing to its higher energy density and easier storage and transport than gaseous hydrogen (Giddey et al., 2017; Hank et al., 2020). Nitrogen required to produce ammonia can be obtained using a cryogenic air separation unit (ASU). The Haber-Bosch (H-B) process can be used to combine the reactants to produce ammonia. This study evaluates ammonia production using an AWE electrolyzer to produce hydrogen and nitrogen from the ASU unit combined in the H-B loop. Modular manufacturing unit specifications for ammonia production are considered in this study.

## 3. Model Formulation

The proposed formulation aims to maximize the profitability for the different scenarios implemented for all the planned renewable installations considered in the study. The profit for different scenarios is defined based on the revenue generated for the entire life of the project and the associated costs during this period. The general equation to calculate the *PROFIT* for both scenarios can be defined as:

$$MAX (PROFIT) = REVENUE - CAPEX - O \& M - FC + SAL \quad (1)$$

where *REVENUE*, *CAPEX*, *O & M*, *FC*, and *SAL* represent the revenue, total capital expenditure, operation and maintenance, fixed cost, and salvage value for the investment at the end of project life, respectively, for the scenarios considered. The generic form for total revenue for the project can be calculated as the summation of the revenue for the hourly resolution considered in the study, represented as follows:

$$REVENUE = \sum_t PRODUCT(t) \quad (2)$$

where  $PRODUCT(t)$  represents the revenue generated in different time intervals for the respective scenarios. The following equation can describe the total load balance of the available power:

$$P_{available}(t) = P_{utilized}(t) + P_{surplus}(t) \quad (3)$$

where  $P_{available}(t)$ ,  $P_{utilized}(t)$ , and  $P_{surplus}(t)$  represent the available power from planned renewable installation, utilized power, and surplus power at different time intervals, respectively.

### 3.1. Hydrogen scenario

In the case of the hydrogen scenario, a combination of electrolyzer which splits water into hydrogen and oxygen (Zhao and You, 2021b, 2021c), and hydrogen tanks has been used to utilize the available power from the planned installations. The following equation can represent the total balance for the power which has been utilized:

$$P_{utilized}(t) = P_{ele}(t) + P_{storage}(t) \quad (4)$$

where  $P_{ele}(t)$  and  $P_{storage}(t)$  refers to the power consumed by the electrolyzer and storing the produced hydrogen. Corresponding to the power which has been utilized in the electrolyzer, the hydrogen produced ( $h2ele(t)$ ) can be calculated as follows (Niaz, Shams, et al., 2022):

$$h2ele(t) = P_{ele}(t) \cdot \eta^{ELE} / LHV \quad (5)$$

where  $\eta^{ELE}$  and  $LHV$  represent the efficiency of the electrolyzer and the lower heating value for hydrogen in kWh/kg of H<sub>2</sub> (Ahmed and Krumpelt, 2001). The power consumed in storing the produced hydrogen depends on the amount produced in the given time interval. Thus, Eq.(6) calculates the storage power based on the hourly data for the hydrogen produced by the electrolyzer (Niaz, Liu, et al., 2022).

$$P_{storage}(t) = P^{HTANK} \cdot h2ele(t) \quad (6)$$

where  $P^{HTANK}$  is the factor showing the amount of power consumed to store the hydrogen produced. The total capital expenditure is the summation of capital expenditure on electrolyzer and storage units. The total capital expenditure on hydrogen storage can be calculated using a certain percentage of the capital expenditure on the electrolyzer unit (Lal and You, 2022). Based on the total life for the electrolyzer and storage units, the salvage value for both units can be calculated using the double declining depreciation method. The total operating expenditure ( $OPEX^{ELE}$ ) is the summation of the operating cost of the electrolyzer ( $O\&M^{ELE}(t)$ ) for all the time intervals considered in the study, as depicted below (Zhao and You, 2021a):

$$OPEX^{ELE} = \sum_t O \& M^{ELE}(t) \quad (7)$$

The fixed cost for the facility (*FC*) is essentially the fixed cost for the total electrolyzer units in operation, which can be calculated using the number of electrolyzer units, the maximum capacity of each electrolyzer, and the project life.

### 3.2. Ammonia scenario

The ammonia scenario requires power utilization in three main processes, i.e., producing hydrogen using the electrolyzer, generation of nitrogen in ASU, and combining the produced hydrogen and nitrogen in the HB-loop (Cheema and Krewer, 2018; Rouwenhorst et al., 2021). Thus, the total load balance of utilized power can be described using Eq.(8).

$$Putilized(t) = Pasu(t) + Pele(t) + Pmodular(t) \quad (8)$$

where  $Pasu(t)$  and  $Pmodular(t)$  represents the power consumed in the ASU and modular ammonia manufacturing unit, respectively. The amount of ammonia generated ( $NH3(t)$ ) can be calculated using the power given to the modular manufacturing unit using Eq.(9).

$$NH3(t) = Pmodular(t) \cdot \alpha_{NH3} \quad (9)$$

where  $\alpha_{NH3}$  refers to the factor for conversion between the power consumed in the modular unit (Gao and You, 2017). Using similar factors to convert the power consumption to the production rate in the electrolyzer and the ASU (Yang et al., 2018), we can get the amount of hydrogen and nitrogen needed. Now, the distribution of the total consumed power among the electrolyzer, ASU, and manufacturing unit depends on their respective ratios ( $\mu_1^{NH3}$ ,  $\mu_2^{NH3}$ ,  $\mu_3^{NH3}$ ) for the ammonia generation process (Wang et al., 2021).

$$Pele(t) = Putilized(t) \cdot \mu_1^{NH3} \quad (10)$$

$$Pasu(t) = Putilized(t) \cdot \mu_2^{NH3} \quad (11)$$

$$Pmodular(t) = Putilized(t) \cdot \mu_3^{NH3} \quad (12)$$

The total capital expenditure is the summation of capital expenditure on the electrolyzer, ASU, and modular units for ammonia production, which can be obtained using the respective number of units employed. Based on the total life of the equipment used, the salvage value can be calculated using the double-declining depreciation method, similar to the hydrogen scenario. The total operating expenditure for the ammonia scenario ( $O \& M^{NH3}$ ) is the summation of the operating cost of the electrolyzer and the storage cost for the hydrogen and ammonia produced for all the time intervals considered in the study, as depicted below:

$$O \& M^{NH_3} = \sum_t (O \& M^{ELE}(t) + O \& M^{HTANK}(t) + O \& M^{NH3TANK}(t)) \quad (13)$$

The storage cost for the hydrogen produced ( $O \& M^{HTANK}(t)$ ) by the electrolyzer unit can be calculated based on the hydrogen production in the given interval and the per-unit storage cost. Similarly, the ammonia storage cost ( $O \& M^{NH3TANK}(t)$ ) can be calculated based on the ammonia production in the given interval and the per unit storage cost. The fixed cost for the facility ( $FC$ ) is essentially the fixed cost for the total electrolyzer units in operation, like the hydrogen scenario, which can be calculated using the number of electrolyzer units, the maximum capacity of each electrolyzer, and the project life.

#### **4. Results and Discussion**

The hydrogen scenario was found to be profitable for the planned renewable installations of most counties considered in the study, with the maximum profit observed for the Texas-Roseland Solar Project and the Texas Atkina Solar project. These cases have a renewable power utilization of around 0.86 with 165 MW of electrolyzers. A scenario can lead to more profit considering the demand for its generated product is higher, which drives up the revenue generated. In contrast to the ammonia scenario, hydrogen has a relatively higher selling price, thus, making its profit margins much higher. It was observed that planned installations with higher nameplate capacity resulted in higher utilization rates. The lowest utilization rate for the hydrogen scenario was 12%. The planned renewable installation in Oregon was the only case that did not result in any profit generation, reflecting it not being an appropriate state due to a lower wind potential. In the case of the ammonia scenario, the planned renewable installations were profitable; however, the profitability in comparison to hydrogen was relatively lower, as depicted in Figure 1. Texas-Roseland Solar Project was found to be a prominently profitable installation for both hydrogen and the ammonia scenario. The most profitable installations in the ammonia scenario had the majority of their utilization ranging from 50%-59%, with only one touching the 70% mark. In contrast to the hydrogen scenario, there is no clear trend of similarity between the profitability and the investment in electrolyzers and ammonia production units. The cost-intensive nature of ASU and H-B units can be attributed to the ammonia scenario being less profitable for the planned installations considered in the study. The project life of the pre-operational planned installation was two years, which is small compared to the normal project life of 30 years for an ammonia production unit, thus aggravating the relatively fewer profit margins for many installations in this scenario.

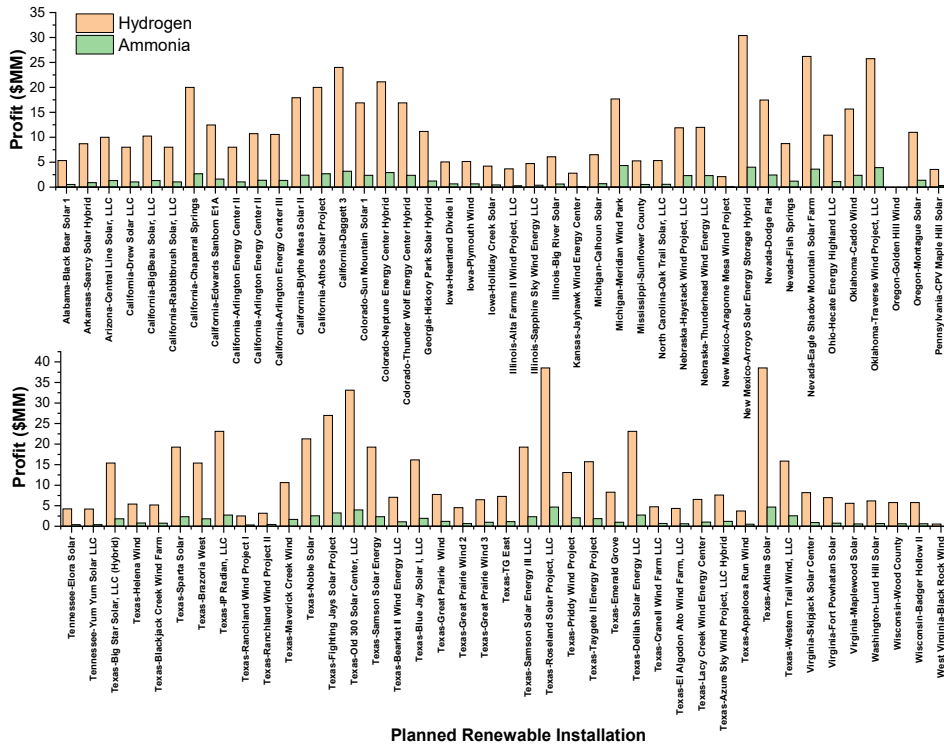


Figure 1. Profit generated from different planned renewable installations based on hydrogen and ammonia scenarios for utilizing the available renewable energy.

## 5. Conclusion

In the current practice, the planned installations for upcoming renewable projects like wind and solar power do not generate any revenue before they are integrated with the grid for commercial operation. This study has examined the potential economic benefits of 83 planned renewable installations in the US before supplying electricity to the grid. It has been analyzed if it can be profitable to operate the planned installations at less than nameplate capacity utilizing two chemical-based energy storage options, namely, hydrogen and ammonia. Results suggest that producing hydrogen using the available renewable power from the planned installations was more profitable due to the higher selling price of the green hydrogen in the market, coupled with the cost-intensive nature of the equipment required to produce ammonia.

## References

- S. Ahmed and M. Krumpelt, 2001. Hydrogen from hydrocarbon fuels for fuel cells. *International Journal of Hydrogen Energy*, 26, 291-301.
- A. Ajagekar, F. You, 2022, Quantum computing and quantum artificial intelligence for renewable and sustainable energy: A emerging prospect towards climate neutrality. *Renewable & Sustainable Energy Reviews*, 165, 112493.
- L. Apodaca, 2022. Nitrogen Statistics and Information. <https://www.usgs.gov/centers/national-minerals-information-center/nitrogen-statistics-and-information#>

- R. Bhandari and R. R. Shah, 2021. Hydrogen as energy carrier: Techno-economic assessment of decentralized hydrogen production in Germany. *Renewable Energy*, 177, 915-931.
- I. I. Cheema and U. Krewer, 2018. Operating envelope of Haber-Bosch process design for power-to-ammonia. *RSC Adv*, 8, 34926-34936.
- J. Gao, F. You, 2017, Can Modular Manufacturing Be the Next Game-Changer in Shale Gas Supply Chain Design and Operations for Economic and Environmental Sustainability? *ACS Sustainable Chemistry & Engineering*, 5, 10046-10071.
- S. Giddey, S. P. S. Badwal, C. Munnings, and M. Dolan, 2017. Ammonia as a Renewable Energy Transportation Media. *ACS Sustainable Chemistry & Engineering*, 5, 10231-10239.
- J. Gong, F. You, 2015, Sustainable design and synthesis of energy systems. *Current Opinion in Chemical Engineering*, 10, 77-86.
- C. Hank, A. Sternberg, et al., 2020. Energy efficiency and economic assessment of imported energy carriers based on renewable electricity. *Sustainable Energy & Fuels*, 4, 2256-2273.
- K. T. Kim, D.J. Lee, and S.J. Park, 2014. Evaluation of R&D investments in wind power in Korea using real option. *Renewable and Sustainable Energy Reviews*, 40, 335-347.
- A. Lal and F. You, 2022. Targeting climate-neutral hydrogen production: Integrating brown and blue pathways with green hydrogen infrastructure via a novel superstructure and simulation-based life cycle optimization. *AIChE Journal*, 69, e17956.
- H. Niaz, J. J. Liu, et al., 2022. Can Texas mitigate wind and solar curtailments by leveraging bitcoin mining? *Journal of Cleaner Production*, 364, 132700.
- H. Niaz, M. Shams, et al., 2022. Mining Bitcoins with Carbon Capture and Renewable Energy for Carbon Neutrality Across States in the USA. *Energy & Environmental Science*, 15, 3551-3570.
- E.N. Pistikopoulos, A. Barbosa-Povoa, J.H. Lee, et al., 2021, Process systems engineering – The generation next? *Computers & Chemical Engineering*, 147, 107252.
- K. H. Rouwenhorst, A. G. Van der Ham, and L. Lefferts, 2021. Beyond Haber-Bosch: the renaissance of the Claude process. *International Journal of Hydrogen Energy*, 46, 21566-21579.
- M. H. Shams, H. Niaz, J. Na, A. Anvari-Moghaddam, and J. J. Liu, 2021. Machine learning-based utilization of renewable power curtailments under uncertainty by planning of hydrogen systems and battery storages. *Journal of Energy Storage*, 41, 103010.
- L. Sun, F. You, 2021, Machine Learning and Data-Driven Techniques for the Control of Smart Power Generation Systems: An Uncertainty Handling Perspective. *Engineering*, 7, 1239-1247.
- X. Tian, T. Meyer, H. Lee, et al., 2020. Sustainable design of geothermal energy systems for electric power generation using life cycle optimization. *AIChE Journal*, 66, e16898.
- G. Thomassen, M. Van Dael, S. Van Passel, et al., 2019, How to assess the potential of emerging green technologies? Towards a prospective environmental and techno-economic assessment framework. *Green Chemistry*, 21, 4868-4886.



- M. Wang, M. A. Khan, et al., 2021. Can sustainable ammonia synthesis pathways compete with fossil-fuel based Haber–Bosch processes? *Energy & Environmental Science*, 14, 2535-2548.
- M. Yang, F. You, 2018, Modular methanol manufacturing from shale gas: Techno-economic and environmental analyses of conventional large-scale production versus small-scale distributed, modular processing. *AIChE Journal*, 64, 495-510.
- N. Zhao, F. You, 2020, Can renewable generation, energy storage and energy efficient technologies enable carbon neutral energy transition? *Applied Energy*, 279, 115889.
- X. Zhao and F. You, 2021. Consequential Life Cycle Assessment and Optimization of High-Density Polyethylene Plastic Waste Chemical Recycling. *ACS Sustainable Chemistry & Engineering*, 9, 12167-12184.
- X. Zhao and F. You, 2021. Waste respirator processing system for public health protection and climate change mitigation under COVID-19 pandemic: Novel process design and energy, environmental, and techno-economic perspectives. *Applied Energy*, 283, 116129.
- X. Zhao and F. You, 2021. Waste high-density polyethylene recycling process systems for mitigating plastic pollution through a sustainable design and synthesis paradigm. *AIChE Journal*, 67, e17127.

# Data-Driven Robust Model Predictive Control on Building Climate Control with Renewable Energy

Wei-Han Chen, Fengqi You

*Cornell University, Ithaca, New York, 14853, USA*

## Abstract

In this work, we propose a data-driven robust model predictive control (DDRMPC) framework to address building climate control with renewable hybrid energy systems under weather forecast uncertainty. The control and energy system configurations include heating, ventilation, and air conditioning, geothermal heat pump, photovoltaic panel, and electricity storage battery. Historical weather forecast and measurement data are gathered from weather station to identify the forecast errors and for the use of uncertainty set construction. The data-driven uncertainty sets are constructed with multiple machine learning techniques, including principal component analysis with kernel density estimation, K-Means clustering, density based spatial clustering of applications with noise, and Dirichlet process mixture model. Lastly, a data-driven robust optimization problem is developed to obtain the optimal control inputs for a building with renewable energy systems. A case study on controlling a building with renewable energy located on the Cornell campus is used to demonstrate the advantages of the proposed framework.

**Keywords:** building climate control, data-driven robust MPC, renewable energy systems.

## 1. Introduction

While the implementation of renewable energy systems and model predictive control (MPC) could reduce the non-renewable energy consumption (Killian and Kozek, 2016), one challenge to building climate control using MPC is the weather forecast uncertainty. A deterministic predictive control framework to regulate the climate of a sustainable building using hybrid energy systems has been developed (Chen and You, 2022). However, thermal comfort would violate the constraints if uncertainties in weather forecasts are not effectively accounted for. Robust MPC (RMPC) could protect the system states from violating the constraints (Yang et al., 2019). Although RMPC could protect the indoor climate from becoming uncomfortable to occupants, it may lead to over-conservative results. To reduce the conservatism of RMPC, data-driven RMPC (DDRMPC) is a popular approach that adopts machine learning techniques or statistical hypothesis tests to construct uncertainty sets that capture high-density regions of the uncertain forecast errors from historical weather data (Shang and You, 2019). Therefore, DDRMPC is an appropriate approach to ensure building climate by taking account of uncertainties within weather forecast errors.

## 2. Building Climate Dynamic Model and RMPC Framework

The overview of the DDRMPC framework for building climate control with the adoption of renewable energy systems is presented in Fig. 1. Temperature, humidity, and predicted mean vote index are the controlled variables in this control framework. The predicted mean vote index, which is derived from temperature, humidity, and other indoor factors, is regulated to ensure the occupants' thermal comfort. The renewable energy systems

adopted in this work include ground source heat pumps, photovoltaic panels, and battery energy storage.

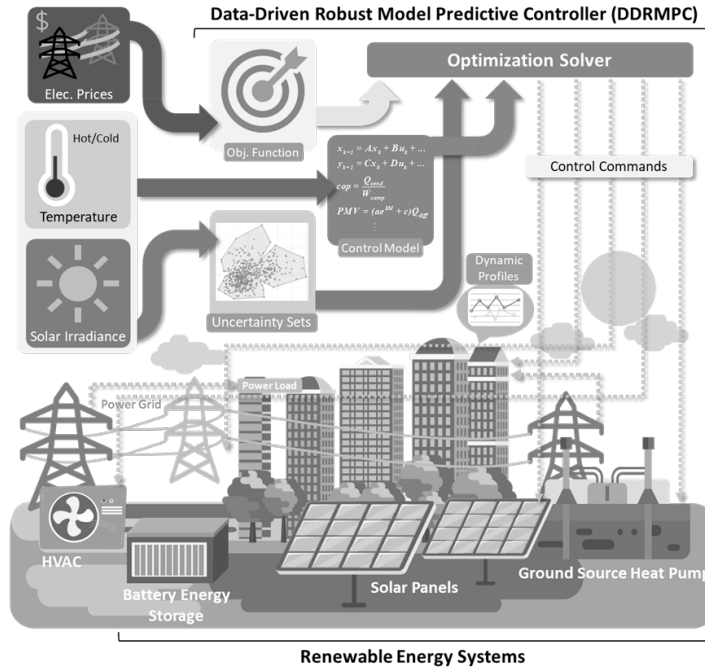


Figure 1. The overview of data-driven robust model predictive control framework for building climate control utilizing renewable energy sources considering weather forecast uncertainties and dynamic pricing.

The most important factor for thermal comfort is indoor temperature. A dynamic temperature model is needed for the DDRMPC to predict future temperatures (Hu and You, 2022). A popular approach to constructing a building dynamic temperature model is to take building components analogous to resistances and capacitances within an electric circuit. The dynamic temperature model for a multi-zone building is generated through Building Resistance-Capacitance Modeling (BRCM) Toolbox in this work (Sturzenegger et al., 2016). The dynamic temperature model for a multi-zone building generated by BRCM is based on the building materials, structures, and geometry. Another important factor in building climate relating to thermal comfort is humidity. The dynamic model for absolute humidity can first be developed using the mass balance equation on water. The equation considering the water amount through ventilation, respiration, and control actuators is given as,

$$\rho V \frac{dh_i}{dt} = m_{vent} + m_{res} + m_{hum} \quad (1)$$

where  $h_i$  denotes absolute humidity,  $m_{vent}$  represents the ventilation,  $m_{res}$  is the occupant respiration, and  $m_{hum}$  denotes the net flow of control actuators for humidity.

Many studies on building energy control only measure thermal discomforts for occupants by the difference between the set point temperature and the room temperature, where other factors that would affect thermal comfort (e.g., humidity, wind speed, occupant's activity level) are neglected. To address this issue, the predicted mean vote index is

adopted in this work to better measure the thermal comfort of the occupants (Enescu, 2017). The predicted mean vote index is estimated by temperature, humidity, clothing, action level of occupants, airspeed, and mean radiant temperature (Yang et al., 2018),

$$PMV = (ae^{bM} + c)Q_{diff} \quad (2)$$

where  $Q_{diff}$  is the net heat of the occupant's body;  $M$  is the occupant's metabolic rate (Yang et al., 2022). The dynamic models are first linearized, and the linearized state-space model for building energy control can be expressed in a compact form with the prediction horizon  $H$ ,

$$\mathbf{x} = \mathbf{A}\mathbf{x}_0 + \mathbf{B}_u \mathbf{u} + \mathbf{B}_v \mathbf{v} + \mathbf{B}_w \mathbf{w} \quad (3)$$

where  $\mathbf{A}$ ,  $\mathbf{B}_u$ ,  $\mathbf{B}_v$ , and  $\mathbf{B}_w$  are the system matrices with the size according to prediction horizon  $H$ .  $\mathbf{x}$ ,  $\mathbf{u}$ ,  $\mathbf{v}$ , and  $\mathbf{w}$  are the sequence vectors of the system state, control input, weather forecasts that are known a priori, and disturbances from forecast errors.

The uncertain disturbances in this work are weather forecast errors. Since weather forecast errors have a disjoint structure, clustering the data has been proven helpful and could be adopted (Zhao and You, 2022). One effective clustering technique is K-means clustering. The forecast errors of temperature, humidity, and solar radiation are first clustered by K-means clustering. Principal component analysis with density estimation (PCADE), which can capture the shape of weather forecast error distributions, is then adopted to construct the data-driven uncertainty sets (Chen and You, 2021). The PCADE-based data-driven uncertainty set for weather forecast errors is given by,

$$D^{PCADE} = \left\{ \mathbf{w} \left| \begin{array}{l} \mathbf{w} = \boldsymbol{\mu}_0 + \mathbf{P}\boldsymbol{\eta}, \boldsymbol{\eta} = \bar{\boldsymbol{\eta}}z^+ + \underline{\boldsymbol{\eta}}z^- \\ z^+ + z^- \leq \mathbf{1}, \mathbf{0} \leq z^+, z^- \leq \mathbf{1}, \mathbf{1}^T(z^+ + z^-) \leq \phi \\ \underline{\boldsymbol{\eta}} = \left[ \hat{F}_{KDE}^{(1)-1}(\alpha), \dots, \hat{F}_{KDE}^{(m)-1}(\alpha) \right]^T \\ \bar{\boldsymbol{\eta}} = \left[ \hat{F}_{KDE}^{(1)-1}(1-\alpha), \dots, \hat{F}_{KDE}^{(m)-1}(1-\alpha) \right]^T \end{array} \right. \right\} \quad (4)$$

Another popular machine learning approach, Dirichlet process (DP) mixture model, has gained interest in density estimation and clustering (Ning and You, 2019). The DP mixture model approach has the advantage of systematically determining the number of clusters without specifying by users in advance. The data-driven uncertainty set constructed by DP mixture model with a variational inference algorithm is shown as,

$$D^{DPMM} = \left\{ \mathbf{w} \left| \mathbf{w} = \boldsymbol{\mu} + s\boldsymbol{\Psi}^{1/2}\boldsymbol{\Lambda}\mathbf{z}, \|\mathbf{z}\|_\infty \leq 1, \|\mathbf{z}\|_1 \leq \phi \right. \right\} \quad (5)$$

The DDRMPC adopts the affine disturbance feedback policy to obtain the optimal control inputs. However, the problem might be infeasible. The feasibility issue can be handled by softening the system state constraints (Lu et al., 2020). The control input constraints should remain hard constraints as the control input cannot surpass the designed limit. The objective function of this work is a linear cost function formulated by the dynamic electricity pricing and the consumption of the actuators. The goal is to minimize the total electricity costs. The soft-constrained data-driven robust optimization problem is presented as,

$$\begin{aligned} & \min_{\mathbf{M}, \mathbf{h}, \boldsymbol{\varepsilon}} \mathbf{c}^T \mathbf{h} + \boldsymbol{\varepsilon}^T \mathbf{S} \boldsymbol{\varepsilon} \\ & \text{s.t. } \mathbf{G}_x [\mathbf{A}\mathbf{x}_0 + \mathbf{B}_u \mathbf{h} + \mathbf{B}_v \mathbf{v} + (\mathbf{B}_u \mathbf{M} + \mathbf{B}_w) \mathbf{w}] \leq \mathbf{g}_x + \boldsymbol{\varepsilon}, \forall \mathbf{w} \in D \\ & \quad \mathbf{G}_u [\mathbf{M}\mathbf{w} + \mathbf{h}] \leq \mathbf{g}_u, \forall \mathbf{w} \in D \\ & \quad \boldsymbol{\varepsilon} \geq \mathbf{0} \end{aligned} \quad (6)$$

where  $\mathbf{cc}$  denotes the electricity costs of the actuators.  $\mathbf{S}$  and  $\boldsymbol{\varepsilon}$  are the matrix of penalty weight and the vector of slack variable added to soften the problem.

### 3. Case Study on a Campus Building Adopting Renewable Energy Systems

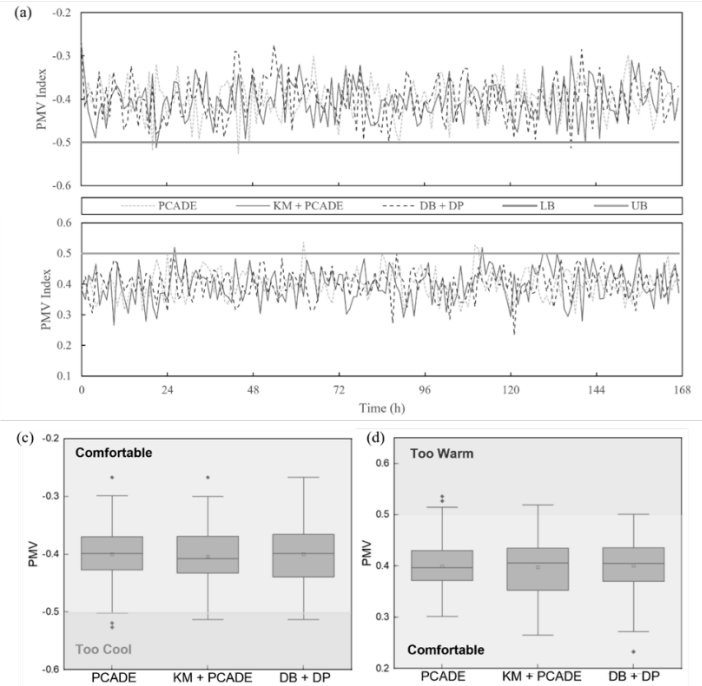


Figure 2. Predicted mean vote index profiles under clustering-based robust MPC comparing three uncertainty sets, including PCADE, K-means followed by PCADE, and DBSCAN followed by DP mixture model, in (a) winter and (b) summer. Box charts that show the maximum, 75<sup>th</sup> quantile, median, 25<sup>th</sup> quantile, and minimum of predicted mean vote index with three different data-driven sets in (c) winter and (d) summer.

**Table 1.** Control Performance of DDRMPC with data-driven uncertainty sets using PCA with KDE, K-means followed by PCA with KDE, DBSCAN, and by DPMM.

	PCA-KDE	KM + PCA-KDE	DB + DPMM
Electricity cost (\$)	20,016	18,914	18,629
Savings	-	-5.5%	-6.9%
Violation percentage	0.36%	0.32%	0.34%

We consider a multi-zone campus building using renewable energy sources at Cornell University in Ithaca, New York, following the proposed DDRMPC framework. Carpenter Hall, which has two stories and multiple zones, is selected as the target building with the goal of decarbonizing the energy systems of the building. The uncertainty sets of weather forecast errors are clustered by different machine learning techniques including PCADE, K-Means clustering, density based spatial clustering of applications with noise (DBSCAN) (Hu and You, 2023), and DP mixture model. Fig. 2(a) and 2(b) show the predicted mean vote index profile of three different approaches, including PCADE, K-

Means coupled with PCADE, and DBSCAN coupled with DP mixture model (Ning and You, 2016) in winter and summer. Fig. 2(c) and 2(d) are the box charts that show the maximum, 75<sup>th</sup> quantile, median, 25<sup>th</sup> quantile, and minimum of predicted mean vote index using data-driven robust MPC approach with three different data-driven sets, including PCADE, K-Means coupled with PCADE, and DBSCAN followed by DP mixture model in winter and summer. The constraints of the predicted mean vote index are set to be between 0.5 and -0.5. The indoor climate would be cool if the predicted mean vote index is below -0.5 and too warm if the predicted mean vote index is above 0.5. During the winter, heat pumps frequently operate to ensure the indoor climate is warm enough for the occupants. The adoption of geothermal energy could effectively decarbonize building energy. The predicted mean vote index is mostly above -0.5, the lower bound. However, when forecast errors are too large, the predicted mean vote index constraints are occasionally violated. The DBSCAN followed by DP mixture model approach has the least violation among the three approaches. The reason is that the outliers are handled best with the help of the DBSCAN approach.

Table 1 presents the control performances of three approaches, including PCADE, K-means coupled with PCADE, and DBSCAN coupled with DP mixture model in terms of electricity cost, savings compared with the PCDE approach, and violation percentage. Constraint violations occasionally occur due to the slight chances of forecast errors not within the constructed uncertainty sets. While all three approaches avoid the constraint violation most of the time with 0.36%, 0.32%, and 0.34% of constraint violations under PCADE, K-means coupled with PCADE, and DBSCAN coupled with DP mixture model, respectively, the DBSCAN coupled with DP mixture model approach results in the least electricity cost and PCADE approach ends up the most. The DBSCAN coupled with the DP mixture model approach achieves 6.9% less electricity cost than the PCADE approach. The least electricity cost under similar constraint violation percentages of DBSCAN coupled with the DPMM approach suggests that DBSCAN is able to better capture the shape of uncertainty data and handle the outliers more efficiently.

#### **4. Conclusion**

In this work, we developed a DDRMPC framework for a multi-zone building with renewable energy systems to address weather forecast uncertainties, reduce electricity costs, and ensure occupants' thermal comfort. Instead of regulating the temperature solely, humidity and predicted mean vote index were also considered to better ensure the occupants' thermal comfort. Machine learning techniques, including DBSCAN and K-means clustering, were first adopted to cluster the weather forecast data. Uncertainty sets of temperature, humidity, and solar radiation forecast errors with a disjunctive structure were then constructed using PCADE and DP mixture model. Finally, a data-driven robust optimization problem was formulated and could be solved to yield the optimal inputs for control actuators at each time step. A year-round simulation case of regulating the temperature, humidity, and predicted mean vote index in a two-story multi-zone campus building in Ithaca, New York was studied to demonstrate the effectiveness of the developed energy optimization framework. The control performances of the uncertainty sets constructed by different machine learning approaches were compared. The proposed DDRMPC framework could effectively minimize the total control cost and ensure the predicted mean vote index within the region. The advantage of DBSCAN coupled with DP mixture model was shown with a smaller constraint violation percentage and less electricity cost than the approach without clustering methods.

## References

- W.-H. Chen, F. You, 2021, Smart greenhouse control under harsh climate conditions based on data-driven robust model predictive control with principal component analysis and kernel density estimation, *Journal of Process Control*, 107, 103-113.
- W.-H. Chen, F. You, 2022, Semiclosed Greenhouse Climate Control Under Uncertainty via Machine Learning and Data-Driven Robust Model Predictive Control. *Ieee Transactions on Control Systems Technology*, 30, 1186-1197.
- W.-H. Chen, F. You, 2022, Sustainable building climate control with renewable energy sources using nonlinear model predictive control, *Renewable and Sustainable Energy Reviews*, 168, 112830.
- D. Enescu, 2017, A review of thermal comfort models and indicators for indoor environments, *Renewable and Sustainable Energy Reviews*, 79, 1353-1379.
- G. Hu, F. You, 2022, Renewable energy-powered semi-closed greenhouse for sustainable crop production using model predictive control and machine learning for energy management, *Renewable and Sustainable Energy Reviews*, 168, 112790.
- G. Hu, F. You, 2023, Multi-zone building control with thermal comfort constraints under disjunctive uncertainty using data-driven robust model predictive control. *Advances in Applied Energy*, 9, 100124.
- M. Killian, M. Kozek, 2016, Ten questions concerning model predictive control for energy efficient buildings, *Building and Environment*, 105, 403-412.
- S. Lu, J.H. Lee, et al., 2020, Soft-constrained model predictive control based on data-driven distributionally robust optimization. *AIChE Journal*, 66, e16546.
- C. Ning, F. You, 2017, Data-Driven Adaptive Nested Robust Optimization: General Modeling Framework and Efficient Computational Algorithm for Decision Making Under Uncertainty. *AIChE Journal*, 63, 3790-3817.
- C. Ning, F. You, 2018, Data-driven decision making under uncertainty integrating robust optimization with principal component analysis and kernel smoothing methods. *Computers & Chemical Engineering*, 112, 190-210.
- C. Ning and F. You, 2019, Data-driven adaptive robust unit commitment under wind power uncertainty: A bayesian nonparametric approach, *IEEE Transactions on Power Systems*, 34, 2409-2418.
- C. Ning, F. You, 2019, Optimization under uncertainty in the era of big data and deep learning: When machine learning meets mathematical programming. *Computers & Chemical Engineering*, 125, 434-448.
- C. Shang, F. You, 2019, Data analytics and machine learning for smart process manufacturing: Recent advances and perspectives in the big data era, *Engineering*, 5, 1010-1016.
- C. Shang, F. You, 2019, A data-driven robust optimization approach to scenario-based stochastic model predictive control. *Journal of Process Control*, 75, 24-39.
- D. Sturzenegger, D. Gyalistras, M. Morari and R. S. Smith, 2016, Model predictive climate control of a Swiss office building: Implementation, results, and cost-benefit analysis, *IEEE Transactions on Control Systems Technology*, 24, 1-12.
- S. Yang, M. P. Wan, B. Ng, et al., 2018, A state-space thermal model incorporating humidity and thermal comfort for model predictive control in buildings, *Energy and Buildings*, 170, 25-39.
- S. Y. Yang, M. P. Wan, W. Y. Chen, et al. 2019, An adaptive robust model predictive control for indoor climate optimization and uncertainties handling in buildings, *Building and Environment*, 163.
- S. Yang, H. Oliver Gao, et al., 2022, Model predictive control in phase-change-material-wallboard-enhanced building energy management considering electricity price dynamics. *Applied Energy*, 326, 120023.
- S. Yang, H. Oliver Gao, et al., 2022, Model predictive control for Demand- and Market-Responsive building energy management by leveraging active latent heat storage. *Applied Energy*, 327, 120054.
- N. Zhao, F. You, 2022, Sustainable power systems operations under renewable energy induced disjunctive uncertainties via machine learning-based robust optimization, *Renewable and Sustainable Energy Reviews*, 161, 112428.

# Thermodynamic Modelling and Performance Analysis of Power Plant for Switching from Coal to Biomass

Romuald Coupan,<sup>a,\*</sup> Ahmed Baccouche,<sup>a</sup> Pierre Szymanski,<sup>a</sup>

<sup>a</sup>*ALBIOMA, Tour Opus 12 – La Défense 9, 77 Esplanade du Général de Gaulle, La Défense F-92914, France*

## Abstract

Converting coal plants to biomass is an economical and environmental solution to continue power generation operations with former coal power plants. Indeed, such fuel change allows significant reduction in greenhouse gas emissions. Switching from coal to biomass, impacts however the process operating conditions and it is required to understand how to adjust these parameters. In this work, we present both modelling, calibration, and simulation of the 60 MW twin thermal power plant of Albioma Bois Rouge (ABR) at La Reunion Island (a French overseas territory). The ABR power plant is currently achieving the replacement of coal by biomass, and a detailed process model of ABR is useful to predict the future performances of the plant. The process model is achieved by using Epsilon® Professional, a computational thermodynamic modelling software. Process simulations allow obtaining the plant performances in terms of boiler efficiency, overall plant energy efficiency, and specific fuel consumption. A comparison of the main process parameters is also discussed for both coal and biomass process operations.

**Keywords:** Power Plant, Energy Efficiency, Biomass, Process Modelling.

## 1. Introduction

The main goal of the energy transition is to fight against global warming. All energy and industry sectors have thus to address plant conversion in an environmentally responsible manner that ensure minimizing greenhouse gas emissions. Accordingly, fossil fuels are being abandoned in favour of renewable energy sources such as solar, wind, hydraulic, and biomass. The energy transition is country specific. For instance, France has set an ambitious goal of increasing its installed renewable energy capacity by 50% by 2023. Accordingly, Albioma – a French independent renewable energy producer, actor in solar, geothermal and biomass energies with more than 1 GW installed capacity – is currently achieving fuel switch <sup>1</sup> of its main fired thermal power plant assets located in French overseas territories. Albioma has reached 75% of renewable energy in its energy production and is also phasing out coal to 100% by 2025 using sustainable biomass. The fuel substitution aims at maintaining power production with former coal power plants with significant reduction in carbon footprint <sup>2</sup>. About 130 to 420 g-CO<sub>2</sub> equivalent per kWh are emitted from biomass power plants whereas 740 to 910 g-CO<sub>2</sub> equivalent per kWh are measured from coal power plant.



Biomass energy is one of the ways to produce renewable and low-carbon electricity by mean of thermal power plant. In this latter, biomass is first burned in presence of air comburant to provide heat which allows the generation of high-pressure steam. This steam then drives a turbine to produce work. The rotation of the turbine shaft finally drives an alternator, which generates electricity. In this process, we can distinguish different successive stages in which chemical energy is first converted into thermal energy, then kinetic energy, mechanical energy, and at this end into electric energy.

To achieve such process, a thermal power plant comprises four main subsystems: (i) the boiler section corresponding to the hot source, (ii) the turboalternator section, (iii) the condenser section corresponding to the cold source, and (iv) the so-called balance-of-plant (BOP) where the feed water is set at suitable operating conditions.

In this work, we present both modelling, calibration, and simulation of the 60 MW twin thermal power plant (2 units of 30 MW each) of Albioma Bois Rouge (ABR) at La Reunion Island (a French overseas territory). A detailed process model of ABR is useful to predict the future performances of the plant with only biomass as fuel, especially for different load cases since the island situation imposed the variation of load every hour. Switching from coal to biomass, impacts indeed the process operating conditions. Process simulations allow obtaining the plant performances in terms of boiler efficiency, overall plant energy efficiency, and specific fuel consumption. A comparison of the main process parameters (e.g., steam turbine inlet flowrate, exhaust turbine condenser pressure, outlet flue gas temperature, etc.) is also discussed for both coal and biomass process operations.

## 2. Process Modelling, Calibration and Validation

The ABR process model is achieved by using Ebsilon® Professional, a computational thermodynamic modelling software <sup>3</sup>. The fluids considered in the model are water, steam, flue gas, air, and fuel (i.e., coal and solid matter). Water and steam are considered pure streams. Their thermodynamic properties are given by the revised formulation of 1997 of The International Association for the Properties of Water and Steam (IAPWS-IF97) <sup>4</sup>. The other flue gas, air and fuel are characterized by their composition (see **Table 1**). Low heating value (LHV) is also specified for the fuel. The model is based on a thermodynamic Hirn cycle with steam extraction. Each system of the power plant – boiler, turbine, condenser, pump – is described according to available input design and process data. The model calibration is realized thanks to real coal-based power generation operations.

*Table 1: Fuels parameters.*

Raw material properties	Coal	Biomass
Carbon (wt.%)	67.1	25.1
Hydrogen (wt.%)	3.8	1.8
Oxygen (wt.%)	6.4	9.6
Nitrogen (wt.%)	1.6	0.2
Sulfur (wt.%)	0.6	0.1
Chlorine (wt.%)	0.007	0.000
Ash (wt.%)	12.6	1.0
Water (wt.%)	7.9	62.2

LHV (kJ/kg)	26215	7935
-------------	-------	------

The power plant is modelled by using predefined blocks and objects for each process conditions changes (see **Figure 1**). The steam once produced is discharged in two successive steam turbine blocks (**1** and **2**) both connected to alternator (**3**) and generator blocks (**4**). The gross electric production stream is obtained from the generator block. A medium-pressure (MP) steam extraction from the first turbine block (**1**) feeds three different blocks corresponding to an air preheater (**5**), a heat consumer (**6**) and a thermal deaerator (**7**). The heat consumer block (**6**) is a way to consider the cogeneration mode in which both electricity and heat can be produced. The exhaust steam from the second turbine block (**2**) feeds a condenser block (**8**), where the steam is fully condensed to water and directed to the deaerator (**7**) via a pump block (**9**). The condensation is ensured by mean of circulating cooling water from a cooling tower block (**10**) fed with ambient air. At the deaerator, four streams are thus recovered: (i) the condensed water from the steam turbine condenser, (ii) the MP steam extraction from the turbine, (iii) the condensate water from the air preheater, and (iv) the return condensate water from the heat consumer. The water stream exiting the deaerator is compressed via a pump block (**11**) and is directed to a heat exchanger economizer block (**12**). The economizer allows preheating the water by recovering heat from flue gas. The hot water from economizer can be used in a second air preheater (**13**) before being vaporized and superheated in a steam generator block (**14**). Optionally, an amount of water from the feed water pump (**11**) is used as de-superheating water stream in the steam generator (**14**). This latter is linked to a combustion zone block (**15**) where fuel reacts with the preheated air to provide useful heat. A controller block (**16**) calculates the suitable fuel flowrate to ensure the equalization of heat flux at both combustion zone and steam generator blocks.

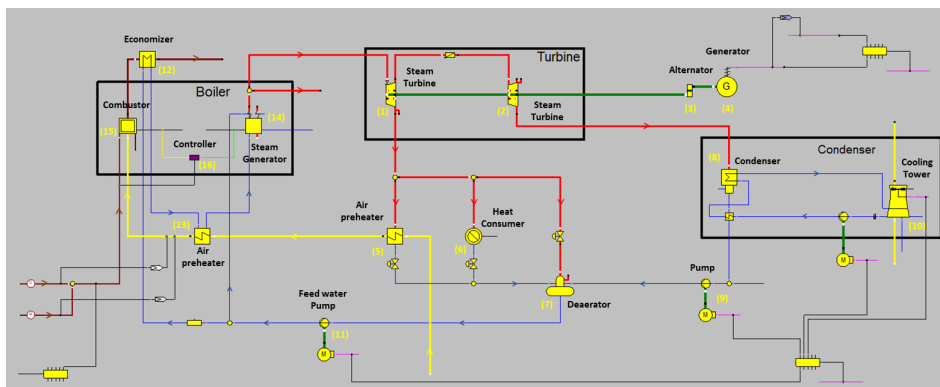


Figure 1: Ebsilon® process model of the ABR power plant. Flowsheet showing (yellow objects) equipment and process streams: (red line) steam, (blue line) liquid water, (yellow line) air, (brown line) coal or biomass fuel, (green line) work, (pink line) electricity, (black line) logic connection.

According to the type of blocks or objects, specific parameterizations are performed. Heat exchanger blocks (economizer, air preheater, condenser, heat consumer) are modelled as counter-current flow exchangers and are specified by terminal temperatures (i.e., either cold or hot outlet temperatures). Pump blocks are specified by their suction and discharge pressures. Turbine blocks are characterized by their discharge ratio and isentropic efficiency. For the steam generator, the pressure drops, and the steam specifications are

filled. For the combustion zone block, it is specified the exhaust gas temperature, the slag temperature, the air / fuel ratio, the thermal and the combustion efficiency. Note that these combustion parameters are dependent on the fuel input. The main process parameters of the ABR plant are given in **Table 2**.

To calibrate and represent part-load mode (operation at flowrate conditions below the nominal case), specific parameters and characteristic curves are applied to turbine, heat exchanger, and boiler blocks. In non-nominal operating cases, the turbine model follows the Stodola's cone law, and it is applied an isentropic efficiency curves function of either mass flowrate, volume flowrate or pressure discharge ratio <sup>5</sup>. Furthermore, pre-established performance curves of steam turbines can be used. These are the basis of a so-called SCC turbine model developed by Spencer, Cotton, and Cannon in 1962 and revised in 1974. Note that it however exists a limitation for the SCC model where it is not valid for powers lower than 16.5 MW and for turbine bodies with an inlet pressure lower than 1.34 bar <sup>6</sup>. For heat exchangers, a characteristic curve is given to represent heat transfer coefficients function of either cold or hot flowrate. For the boiler, the exhaust flue gas temperature is determined with a characteristic curve function of the air flowrate.

*Table 2: Main process parameters of the ABR power plant.*

Process specifications	Nominal parameters
Gross electric production	30 MW
Steam turbine admission pressure	81 bara
Steam turbine admission temperature	525°C
Turbine exhaust pressure	0,10 bara
Water condenser outlet temperature	50°C
Deaerator pressure	2.3 bara
Fluegas economizer outlet temperature	150°C
Fluegas boiler outlet temperature	400°C
Ambient temperature	27°C

From all the used input measurement data, an average uncertainty of 3% is estimated on the output values of the model. Moreover, a validation step is performed to assess the reliability of the simulated performances. To do so, the evolution of a performance index named "specific fuel consumption" (SFC) – corresponding to the equivalent mass of a standard coal required for 1 kWh net electric production – is tracked and compared with real data from the production site. In the range of 5 to 30 MW, it is thus found 2% of deviation between the simulated and the real SFC. Accordingly, this deviation being inferior to the average uncertainty, the model is validated for the corresponding electric production range. Note that the model validation is only done for coal operations according to the data availability. Fuel switch being in progress, the results for biomass are still subject to validation.

### 3. Results and discussions

The model once build, is used to simulate SFC of the power plant for two different fuels: coal and biomass. Switching from coal to biomass, implies some changes in the operating

conditions, but also impacts the performances. This modification of operating conditions is considered via the combustion parameters specifications such as the exhaust gas temperature, the air / fuel ratio, and the thermal / combustion efficiency. The simulated performances are given in term of SFC function of the net electric production (**Figure 2 – a**) and energy efficiency (EE) function of the thermal power input (**Figure 2 – b**). Note that EE is the ratio between the net electric power production and the thermal power input. The SFC of coal is thus found lower than the SFC of the biomass for all the covered range of net electric production. It is observed a difference of about 3% to 6% of SFC from low to high net electric production. The same trend is obtained for EE in the range of thermal power input. The biomass power plant operation is less energy-efficient compared to coal operation. Accordingly, to produce the same amount of electricity than coal operation, more energy is needed with biomass as fuel.

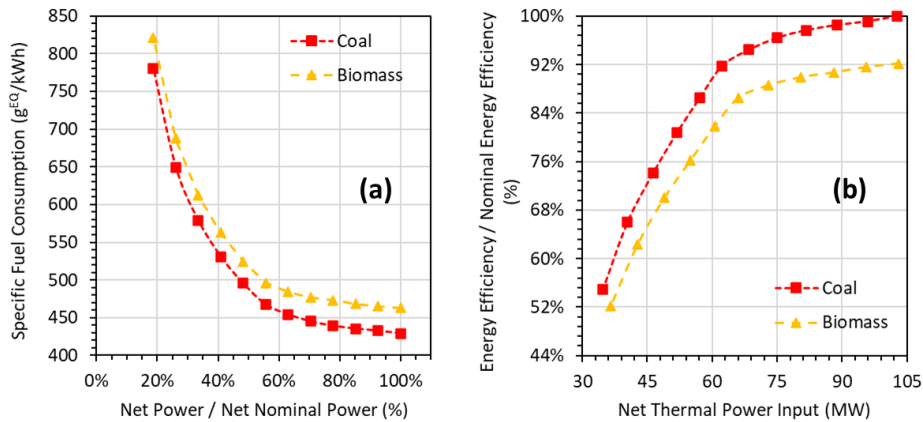


Figure 2: Simulation of (a) the specific fuel consumption function of the net electric production and (b) the energy efficiency function of the net thermal power input of the ABR power plant: (red) coal, (yellow) biomass.

Knowing that the turbine, condenser, and BOP sections are not affected by fuel changes, only the boiler efficiency can be pointed out to explain the SFC reduction by switching from coal to biomass. The boiler efficiency is defined as the ratio of the useful heat output to the total energy input. From the simulations in the range of 10 to 30 MW, the boiler efficiencies are calculated at about 92% and 88% for coal and biomass operations respectively. These values are in line with those observed in literature<sup>7</sup>. When the boiler efficiency decreases, more fuel must be burned to reach the steam specifications for the electricity production.

Working with biomass instead of coal, several observations can be done: (i) the air / fuel ratio is higher, leading to dilution of the heat produced; (ii) the thermal and combustion efficiency are lower, mainly due to less carbon content, high moisture content and heterogeneous size of the biomass; and (iii) the exhaust flue gas temperature is higher, as it is correlated to the former coal-fired boiler design. Therefore, the main reason for poorer performance of biomass boiler is the thermal losses and depends on the fuel composition.

The main loss of the boiler being the flue gas losses, a particular attention is generally addressed for heat exchanger economizer design<sup>8-9</sup>. The economizer allows the heat recovery from hot exhaust flue gas to preheat the feed water addressing operating

constraints such as: the variation of hot and cold flowrates and temperature profiles, the dust content, the limit temperature on both sides i.e., the flue gas acid dew point and the water vaporization temperature <sup>9</sup>.

The biomass composition and particularly its high moisture content, imply a very low LHV. Higher is the moisture content, lower the LHV. Significant effects can thus be expected on the boiler efficiency according to the typical variation of moisture content of biomass with the different sources and seasons. Note that biomass energy density can be improved by mean of a biomass dryer upstream the boiler.

#### 4. Conclusion

In this study, coal-to-biomass power plant conversion is assessed by mean of process modelling using the computational thermodynamic modelling software Ebsilon® Professional. Switching from coal to biomass fuels, requires changing the operating conditions, and leads to modifications of the power plant performances. A process model is thus built, calibrate, and validate to simulate the specific fuel consumption of the plant on different load cases. Using this model, the biomass operations is found less energy efficient than coal operations. The boiler efficiency is pointed out with explanations based on the increase of heat losses with biomass as fuel. With these considerations, process intensification strategy will be proposed to improve the efficiency.

From energy efficiency aspect, this analysis does not point to positive considerations for switching from coal to biomass. However, one must remind that the main objective of this fuel substitution is the environmental aspects to continue power generation operations with former coal power plants with significant reduction in greenhouse gas emissions. The carbon footprint of biomass power plants is about 130 to 420 g-CO<sub>2</sub> equivalent per kWh compared to coal power plant with about 740 to 910 g-CO<sub>2</sub> equivalent per kWh <sup>2</sup>.

#### References

- [1] D. W. Bunn, J. Redondo-Martin, J. I. Muñoz-Hernandez, P. Diaz-Cachinero, **2019**, Analysis of coal conversion to biomass as a transitional technology, *Renewable Energy*, *132*, 752 – 760.
- [2] Steffen Schlömer (ed.), **2014**, Technology-specific Cost and Performance Parameters, Annex III of Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.
- [3] P. Madejski, P. Zymelka, **2020**, Calculation methods of steam boiler operation factors under varying operating conditions with the use of computational thermodynamic modeling, *Energy*, *197*, 117221 – 117232.
- [4] The International Association for the Properties of Water and Steam (IAPWS), **2007**, Revised Release on the IAPWS Industrial Formulation 1997 for the Thermodynamic Properties of Water and Steam, Lucerne, Switzerland.
- [5] S. Naderi, M. Banifateme, O. Pourali, A. Behbahaninia, I. MacGill, G. Pignatta, **2020**, Accurate capacity factor calculation of waste-to-energy power plants based on availability analysis and design/off-design performance, *Journal of Cleaner Production*, *275*, 123167 – 123170.
- [6] R. C. Spencer, K. C. Cotton, C. N. Cannon, **1974**, A method for predicting the performance of steam turbine-generators...16500kW and larger.
- [7] Esa Kari Vakkilainen (Ed.), **2017**, 3 – Boiler processes, Steam Generation from Biomass, Butterworth-Heinemann, 57 – 86.
- [8] V. D. Stevanovic, M. M. Petrovic, T. Wala, S. Milivojevic, M. Ilic, S. Muszynski, **2019**, Efficiency and power upgrade at the aged lignite-fired power plant by flue gas waste heat utilization: High pressure versus low pressure economizer installation, *Energy*, *187*, 115980 – 115992.

- [9] G. Miliauskas, E. Puida, R. Poskas, V. Ragaisis, L. Paukstaitis, H. Juhara, L. Mingilaite, **2022**, Experimental investigations of water droplet transient phase changes in flue gas flow in the range of temperatures characteristic of condensing economizer technologies, *Energy*, *256*, 124643 – 124655.



# Bottleneck-identification methodology and debottlenecking strategy for heat exchanger network with disturbance

Liwen Zhao,<sup>a</sup> Guilian Liu,<sup>a</sup>

*<sup>a</sup>School of Chemical Engineering and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi Province, 710049, China*

## Abstract

In chemical processes, source/sink parameters of the HEN change due to catalyst deactivation, industrial restructuring, and other reasons. The bottleneck of energy utilization exists and changes with these parameters. A general bottleneck-identification method and the corresponding debottleneck strategy are proposed. Based on topology analysis and algebraic reasoning, feasible Disturbance Response Schemes (DRSs) of the system and load-shift rules of heat exchangers under single-stream heat capacity flowrate disturbance are clarified. Relations among the overall thermal conductance demands, total annualized cost (TAC), and parameter fluctuation are deduced. Corresponding diagrams are constructed to identify the bottleneck, and the TAC is taken as the evaluation index to target the optimal scheme for eliminating the bottlenecks. The proposed method is intuitive and efficient and can guide the design, optimization, or retrofit of chemical processes. A benzene alkylation process is analyzed to illustrate its application.

**Keywords:** heat exchanger network, bottleneck, debottleneck, economy

## 1. Introduction

Along with the process improvement and product/raw material adjustment, the streams related to the reactor and separator will change in composition and flowrate, leading to constant changes in their heat capacity flowrates ( $CP$ ). These streams are generally sources and sinks of the Heat Exchanger Network (HEN) and form multiple heat flow paths with utilities (Smith, 2010). System disturbances transfer along these paths, affecting the utility consumption, the heat load requirements, and the overall thermal conductance demands ( $UA$ )<sup>req</sup> of some heat exchangers. In this process, the  $UA$  of some heat exchangers may be insufficient, resulting in bottlenecks restricting energy utilization and even affecting the feasibility of the process. Identifying the locations of these bottlenecks and proposing economic debottleneck strategies is of great significance to the optimization of HENs and stable production of a chemical process.

Towards practical chemical processes with disturbances, approaches and algorithms were proposed for matching heat transfer areas in a multiperiod HEN retrofit (Kang and Liu, 2014), optimizing topology and handling continuous heat load variables (Pavao et al., 2017), and optimizing the placement of the heater/cooler and relocating utilities within HEN (Zamora et al., 2020). These optimization/retrofit methods rely heavily on topology modifications, might result in a long-retrofit duration, and could be impractical. To overcome this disadvantage, heat transfer enhancement technology was deployed in the HEN retrofit (Klemes et al., 2020). However, the heat exchanger's area demand is generally set as a constant to reduce the difficulty in solving, and this conflicts with the



actual production condition (Tian et al., 2018) and may lead to unreasonable retrofit. Besides, the best location in which to implement the retrofit and the treatment of downstream effects after the retrofit are left out of consideration.

This paper aims to study the locations of energy bottlenecks under disturbance and propose the corresponding debottlenecking strategy. The disturbance response scheme and the change rules of load-shift and  $(UA)^{req}$  will be studied in Section 2. In Section 3, methods are developed for identifying bottlenecks and debottlenecking strategy. A case is investigated in Section 4 to illustrate the proposed method.

## 2. Disturbance response scheme and load-shift rules

For a disturbed stream ( $S_D$ ), variations in flow rate and composition cause its  $CP$  changes from  $CP^{ini}$  to  $\alpha CP^{ini}$ , affecting the system's utility consumption. The variations in heating and cooling utilities are related to the position and properties of  $S_D$  in the temperature-enthalpy (T-H) diagram and can be determined according to the location rules of sources and sinks (Zhang and Liu, 2017), as shown in Table 1. In this table,  $\Delta H$  denotes the enthalpy change of the disturbed stream and can be calculated by Eq. (1);  $\Delta H_1$  and  $\Delta H_2$  represent the enthalpy change above and below the pinch and equal  $(\alpha - 1)CP^{ini}(T^S - T^P)$  and  $(\alpha - 1)CP^{ini}(T^P - T^T)$ , respectively.

Table 1 Variation of utilities when  $CP$  of  $S_D$  changes

Cases	Disturbed stream		Variation of utility	
	Position	Property	Heating utility ( $\Delta Q_H$ )	Cooling utility ( $\Delta Q_C$ )
Case 1	Above the pinch	Source/sink	$-\Delta H$	0
Case 2	Below the pinch	Source/sink	0	$\Delta H$
Case 3	Across the pinch	Source	$-\Delta H_1$	$\Delta H_2$
Case 4	Across the pinch	Sink	$-\Delta H_2$	$\Delta H_1$

$$\Delta H = \Delta H_1 + \Delta H_2 = (\alpha - 1)CP^{ini}(T^S - T^P) + (\alpha - 1)CP^{ini}(T^P - T^T) \quad (1)$$

Where  $T$  is the temperature of  $S_D$ ; superscripts  $S$  and  $T$  represent the supply and target values;  $T^P$  is the average pinch temperature; superscript  $ini$  denotes the initial value;  $\alpha$  represents the fluctuation coefficient of  $CP$ .

### 2.1. Disturbance response scheme

The heat balance of the disturbed process can be maintained by controlling the flow rate of the cooler or heater's medium, which is defined as a Response Variable (RV). The heater/cooler whose heat duty changes under disturbance is defined as a Response Object (RO) located at the response stream  $S_R$ . This scheme responding to the disturbance by adjusting the RV is called the Disturbance Response Scheme (DRS). The fewer the ROs, the lower the scheme's operational complexity.

The disturbance propagates through downstream paths, which are uninterrupted and have a mature determination procedure (Zhu et al., 1996). In a DRS, the shortest path connecting  $S_D$  and  $S_R$  is defined as the Disturbance Propagation Path (DPP). It is worth mentioning that, affected by the topology structure of the HEN, there may be no DPP between  $S_D$  and  $S_R$ . In this case, the RO on  $S_R$  and the DRS are invalid.

For the HEN with  $n$  heaters and  $m$  coolers, if  $S_D$  lies above the pinch, there are at most  $n$  DRSs. When heater  $HE_i$  ( $i=1, 2, \dots, n$ ) is taken as the RO, the DRS is recorded as  $DRS_i$ , and the corresponding disturbance propagation path is marked as  $DPP_i$ . Similarly, if  $S_D$  locates below the pinch and cooler  $CE_j$  ( $j=1, 2, \dots, m$ ) is set as the RO, the response scheme and propagation path are recorded as  $DRS_j$  and  $DPP_j$ , and there are at most  $m$  valid DRSs; if  $S_D$  crosses the pinch, the maximum number of possible DRSs is  $(n \cdot m)$

since the disturbed system has two ROs, one heater, and one cooler, and the variations of their duties are  $\Delta Q_H$  and  $\Delta Q_C$ , respectively. If  $HE_i$  and  $CE_j$  are taken as ROs, the control scheme is  $DRS'_{i,j}$ , corresponding to two disturbance propagation paths,  $DPP'_i$  and  $DPP'_j$ .

### 2.2. Load-shift rules of heat exchangers

With disturbance migration, the duty of heat exchangers on the DPP will change. By referring to the analysis methods proposed by Zhao and Liu (2022), the load variation of each heat exchanger under various disturbance-response cases is investigated, and general load-shift rules of heat exchangers along the DPP are summarized.

If  $S_D$  lies above or below the pinch, a DRS has a single RO and DPP. The  $K$  heat exchangers in  $DPP_i$  or  $DPP_j$  are numbered and divided into odd and even groups. Load-shift rules inside and outside the disturbance propagation path are as follows:

- (1) If  $K$  is even (odd), the heat load variation of the odd (even) group's heat exchanger is  $\Delta Q_H$  or  $\Delta Q_C$ , and that of the even (odd) group is  $-\Delta Q_H$  or  $-\Delta Q_C$ .
- (2) The heat load of RO changes by  $\Delta Q_H$  or  $\Delta Q_C$ , and that of the heat exchanger outside the path remains unchanged.

For  $S_D$  crosses the pinch, a DRS has two ROs and DPPs, and the device's load variation follows the rules introduced above. It is worth noting that if a heat exchanger locates on  $DPP'_i$  and  $DPP'_j$  simultaneously, its heat load variation equals the superposition of those in two paths.

## 3. Heat transfer bottleneck identification and debottleneck strategy

### 3.1. Changes in the overall thermal conductance of heat exchangers

According to the heat transfer equation shown in Eq. (2),  $UA$  needs to be changed to meet the heat transfer requirements of the heat exchanger. However, Eq. (2) cannot determine the dynamic change of the  $(UA)^{req}$  under disturbance due to its strong temperature dependence. In this case, the heat transfer efficiency,  $\varepsilon$ , is introduced. It is the ratio of the heat exchanger's actual heat transfer rate to the maximum value,  $Q_{E, \max}$ , and can be transformed into Eq. (3) for a counterflow heat exchanger. Besides,  $\varepsilon$  can be expressed by Eq. (4) (Incropera and Dewitt, 1996).

$$Q_E = UA\Delta T_m = UA[(T_{hi} - T_{co}) - (T_{ho} - T_{ci})] / \ln[(T_{hi} - T_{co}) / (T_{ho} - T_{ci})] \quad (2)$$

$$\varepsilon = Q_E / Q_{E, \max} = (Q_E / CP_{\min}) / (Q_E / CP_{\min} + \Delta TE_{\min}) \quad (3)$$

$$\varepsilon = \left| \operatorname{sgn} \left( 1 - \frac{CP_{\min}}{CP_{\max}} \right) \right| \frac{CP_{\max} - CP_{\min} \exp \left[ (UA)^{req} (1/CP_{\max} - 1/CP_{\min}) \right]}{CP_{\max} - CP_{\min} \exp \left[ (UA)^{req} (1/CP_{\max} - 1/CP_{\min}) \right]} + \frac{(UA)^{req}}{(UA)^{req} + CP_{\min}} \quad (4)$$

Where  $U$  is the heat transfer coefficient;  $A$  indicates the heat exchange area;  $\Delta T_m$  denotes the logarithmic mean temperature difference; subscripts  $h$  and  $c$  denote the hot and cold stream; subscripts  $i$  and  $o$  indicate the inlet and outlet stream;  $CP_{\max}$  and  $CP_{\min}$  indicate the maximum and minimum one between  $CP_h$  and  $CP_c$ ;  $\Delta TE_{\min}$  denotes the minimum temperature difference of a heat exchanger and remains unchanged.  $\operatorname{sgn}(x)$  is a sign function, and it equals -1, 0, and 1 when  $x$  is less, equal, and greater than 0.

With Eq. (3) and Eq. (4) combined, relations among  $(UA)^{req}$ ,  $CP$ , and  $Q_E$  are deduced, as shown in Eq. (5). According to section 2,  $Q_E$  can be described as a function of  $\alpha$ , i.e.,  $Q_E = Q(\alpha) = r + q\alpha$ , where  $r$  and  $q$  are constants and can be calculated based on Table 1. In addition, different initial relations between  $CP_h$  and  $CP_c$  result in different representations

of  $CP_{\max}$  and  $CP_{\min}$ , which will change as the variation of  $\alpha$ . This leads to four possible initial-disturbed combinations, as shown in Figure 1.

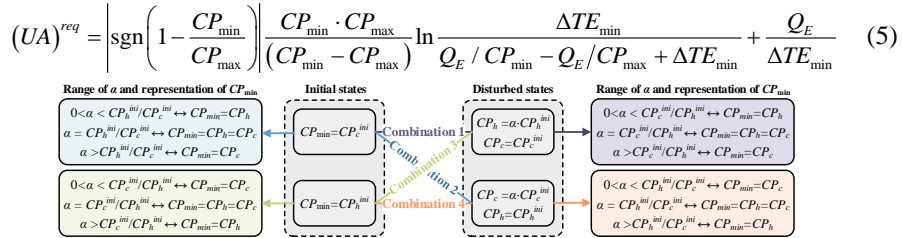


Figure 1 Four initial-disturbed cases and the representation of  $CP_{\min}$  under different ranges of  $\alpha$ . The relationship between  $(UA)^{req}$  and  $\alpha$  is analyzed for each combination based on Eq. (5), and the general correlation is summarized and shown in Eq. (6).

$$(UA)^{req} = \left| \operatorname{sgn}(\alpha - R) \right| \left\{ \frac{\alpha CP_u}{|\alpha - R|} \ln \left[ \frac{|\alpha - R| Q(\alpha)}{\alpha CP_u \Delta TE_{\min}} + 1 \right] - \frac{Q(R)}{\Delta TE_{\min}} \right\} + \frac{Q(R)}{\Delta TE_{\min}} \quad (6)$$

Where  $R$  denotes the ratio between the initial  $CP$ s of streams  $u$  and  $v$  and is a constant;  $u$  and  $v$  represent streams with constant and changing  $CP$ .

### 3.2. Identification of heat exchange bottleneck and debottleneck strategy

With the monotonicity of  $(UA)^{req}$  analyzed, the  $(UA)^{req} \sim \alpha$  curve can be constructed, as shown in Figure 2. For an actual process, the curve can be plotted to intuitively illustrate the variation of  $(UA)^{req}$  along with the disturbance, find out the bottleneck restricting energy utilization, and lay a foundation for proposing the debottleneck strategy.  $(UA)^{req} \leq (UA)^{ini}$  indicates that the designed heat exchanger can meet the heat exchange demand, and there is no need to retrofit this equipment. If  $(UA)^{req} > (UA)^{ini}$ , the heat exchanger's initial overall thermal conductance is insufficient and cannot meet the heat exchange demand, and it needs to be modified. A common modification strategy is to increase its area, and the area increment ( $\Delta A$ ) can be calculated by Eq. (7) in the case of constant  $U$ . This figure can also be applied in the design stage to determine the optimal  $UA$  of the installed heat exchanger, i.e.,  $(UA)_{\max}$ , to satisfy the heat transfer requirements throughout the whole production cycle.

$$\Delta A = 0.5 \left[ (UA)^{req} - (UA)^{ini} + \left| (UA)^{req} - (UA)^{ini} \right| \right] / U \quad (7)$$

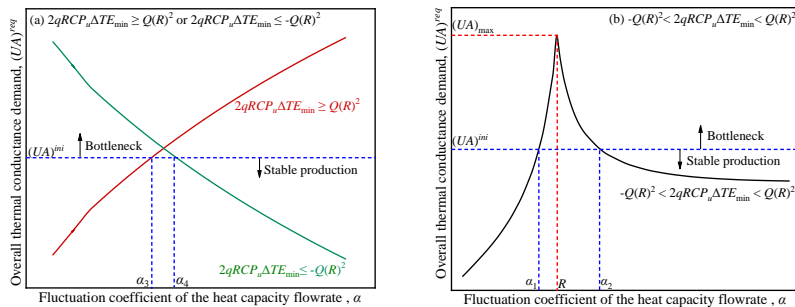


Figure 2  $(UA)^{req}$  versus  $\alpha$  curves

In summary, multiple DRSs exist when the system fluctuates and correspond to different ROs and DPPs. For different DRSs, the system's utility demands are different, as well as the annual operation cost ( $Co$ ). With single-stream  $CP$  disturbance, the change in  $Co$  can be expressed by Eq. (8). With the relation between  $\Delta Q_H / \Delta Q_C$  and  $\alpha$  substituted, the relation between  $Co$  and  $\alpha$  is determined. Besides, changes in the  $(UA)^{req}$  of each heat

exchanger lead to variations in the annual investment cost ( $C_i$ ). For HEN containing  $X$  heat exchangers, the serial number of each heat exchanger is denoted as  $Ex$  ( $x=1, 2, \dots, X$ ). On the basis of Eq. (7),  $C_i$  can be calculated by Eq. (9), consisting of the purchase and the installation costs (Al-Riyami et al., 2001).

$$Co = t_A \left[ p_{HU} (\Delta Q_H + Q_H^{ini}) + p_{CU} (\Delta Q_C + Q_C^{ini}) \right] \quad (8)$$

$$Ci = \left\{ a \cdot Y + \sum_{x=1}^X 0.5^c b \left[ (UA)_{Ex}^{req} - (UA)_{Ex}^{ini} + \left| (UA)_{Ex}^{req} - (UA)_{Ex}^{ini} \right|^c / U^c \right] \right\} / D_p \quad (9)$$

Where  $t_A$  is the annual running time, h·year<sup>-1</sup>;  $p$  denotes the price of the utility, USD·kJ<sup>-1</sup>;  $a$ ,  $b$ , and  $c$  are cost coefficients;  $Y$  indicates the number of heat exchangers with  $(UA)^{req}$  greater than  $(UA)^{ini}$ ;  $D_p$  represents the depreciable period of heat exchangers, year.

With Eq. (8) and Eq. (9) superimposed, the relationship between TAC and  $\alpha$  can be deduced for each DRS of a disturbed system (with  $z$  DRSs), and the economic variation diagram can be plotted on this basis. For a given  $\alpha$ , the DRS with the smallest TAC should be chosen as the optimal bottleneck-solving strategy.

#### 4. Case study

For a simplified alkylation process (Zhao and Liu, 2022), there are one source, three sinks, and five heat exchangers (including two heaters) in the HEN; their initial data are displayed in Figure 3. As the catalyst deactivates, the heat capacity flowrate of the reactor's outlet stream (H1), located above the pinch, changes. In this case, the disturbed system has two DRSs, DRS<sub>1</sub> (HE1 and C1 are RO and S<sub>R</sub>) and DRS<sub>2</sub> (HE2 and C2 are RO and S<sub>R</sub>).

According to Table 1, the cooling utility keeps unchanged while the heating utility changes as  $CP_{H1}$  fluctuates and  $\Delta Q_H=4706(1-\alpha)$ . Based on load-shift rules deduced in section 2.2, relations between the  $Q_E$  and  $\alpha$  of each heat exchanger are determined for DRS<sub>1</sub> and DRS<sub>2</sub>, also shown in Figure 3.

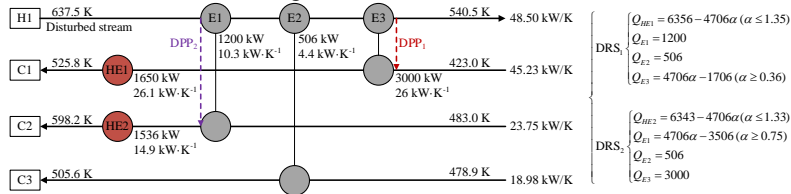


Figure 3 Simplified HEN of the alkylation process from benzene to ethylbenzene

With these initial data substituted into Eqs. (6)~(9), relations among heat exchanger's  $(UA)^{req}$ ,  $Co$ ,  $C_i$ , TAC, and  $\alpha$  can be clarified for DRS<sub>1</sub> and DRS<sub>2</sub>. Based on this, the  $(UA)^{req} \sim \alpha$  curves and the economic variation diagram are plotted, as shown in Figure 4, where  $\Delta TAC$  indicates the difference between  $TAC_{DRS1}$  and  $TAC_{DRS2}$ .

Figure 4 (a) shows that when  $\alpha$  changes to different values, the bottlenecks of the system are different. Besides, different DRS corresponds to different bottlenecks. For example, when  $\alpha$  equals 0.75, E1, E2, and HE1 are bottlenecks for DRS<sub>1</sub> due to the insufficient  $UA$ , but for DRS<sub>2</sub>, E2 and HE2 are bottlenecks; when  $\alpha$  varies to 1.33, the bottlenecks of DRS<sub>1</sub> and DRS<sub>2</sub> are E3 and E1. Affected by this characteristic, the variations of  $C_i$  and TAC, along with  $\alpha$ , are discontinuous, as shown in Figure 4 (b). When  $\alpha$  locates in the interval [0.91,1.09],  $C_i$  is zero, indicating that the initial  $UA$  of heat exchangers can satisfy the heat exchange demands of the system. According to the  $\Delta TAC \sim \alpha$  curve, if  $\alpha$  lies in the interval [0.75,0.88], DRS<sub>2</sub> should be chosen as the debottleneck strategy since  $\Delta TAC > 0$ . Compared to DRS<sub>1</sub>, this scheme can save up to 19,400 USD·year<sup>-1</sup> (point A<sub>1</sub>). When  $\alpha$

locates at the interval  $[0.88, 1.33]$ ,  $\Delta TAC \leq 0$ . From an economic point of view, DRS<sub>1</sub> should be selected as the bottleneck-solving strategy and can save up to 24,024 USD·year<sup>-1</sup> (point A<sub>2</sub>) compared with DRS<sub>2</sub>.

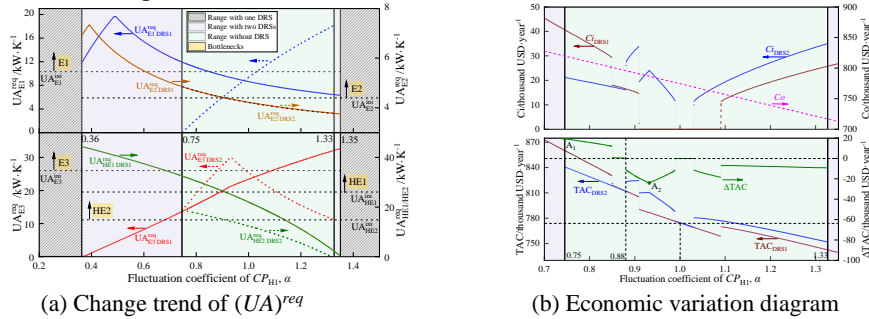


Figure 4 Dynamic change diagram of  $(UA)^{req}$  and system cost

## 5. Conclusions

Locations of energy bottlenecks under system disturbance are studied with the constraint of the overall thermal conductance considered, and the corresponding debottlenecking strategy is proposed. The method can be applied in the design, optimization, or modification stage to evaluate the system's operational flexibility, determine the energy bottlenecks, and propose the optimal bottleneck elimination scheme without complex simulation and calculation. For the studied case, optimal bottleneck elimination schemes are determined when  $\alpha$  is in different intervals. Although the study aims at a single stream's  $CP$  fluctuation, the method can be extended to analyze other critical parameters, which will be studied in future work.

## References

- Al-Riyami BA, Klemes J, Perry S, 2001, Heat integration retrofit analysis of a heat exchanger network of a fluid catalytic cracking plant, *Applied Thermal Engineering*, 21, 13/14, 1449.
- Incropera FP, Dewitt DP, 1996, Fundamentals of heat and mass transfer, Wiley, New York.
- Kang LX, Liu YZ, 2014, Target-oriented methodology on matching heat transfer areas for a multiperiod heat exchanger network retrofit, *Industrial & Engineering Chemistry Research*, 53, 45, 17753-17769.
- Klemes JJ, Wang QW, Varbanov PS, Zeng M, Chin HH, Lal NS, Li NQ, Wang BH, Wang XC, 2020, Heat transfer enhancement, intensification and optimisation in heat exchanger network retrofit and operation, *Renewable & Sustainable Energy Reviews*, 120, 109644.
- Pavao LV, Costa CBB, Ravagnani M, 2017, Heat exchanger network synthesis without stream splits using parallelized and simplified simulated annealing and particle swarm optimization, *Chemical Engineering Science*, 158, 96-107.
- Smith R, 2010, Chemical process design and integration, Wiley, West Sussex, England.
- Tian X, Yin CF, Lv DH, Wang P, Liu GL, 2018, Effect of catalyst deactivation on the energy consumption of gasoline-diesel hydrotreating process, *Energy & Fuels*, 32, 10, 10879-10890.
- Zamora JM, Hidalgo-Munoz MG, Pedroza-Robles LE, Nunez-Serna RI, 2020, Optimization and utilities relocation approach for the improvement of heat exchanger network designs, *Chemical Engineering Research & Design*, 156, 209-225.
- Zhang D, Liu GL, 2017, Integration of heat exchanger network considering the pressure variation of distillation column, *Applied Thermal Engineering*, 116, 777-783.
- Zhao LW, Liu GL, 2022, Dynamic coupling of reactor and heat exchanger network considering catalyst deactivation, *Energy*, 260, 125161.
- Zhu J, Han Z, Rao M, Chuang K, 1996, Identification of heat load loops and downstream paths in heat exchanger networks, *The Canadian Journal of Chemical Engineering*, 74, 6, 876-882.

# A Power-to-Gas energy system: modeling and operational optimization for seasonal energy supply and storage

Yifan Wang,<sup>a</sup> Luka Bornemann,<sup>a</sup> Christiane Reinert,<sup>a</sup> Niklas von der Aßen<sup>a,\*</sup>

<sup>a</sup>*Institute of Technical Thermodynamics, RWTH Aachen University, Schinkelstr. 8, 52062 Aachen, Germany*

## Abstract

Power-to-Gas (PtG) technologies show high potential as a long-term solution for future energy systems to compensate for seasonal fluctuations in energy demand and renewable energy supply. Despite growing attention to PtG technologies, their modeling is typically simplified during energy system optimization for computational tractability.

This work mathematically describes a power-to-gas energy system for seasonal energy supply and storage (ES<sup>4</sup>) with great technical and temporal detail based on a real-world pilot PtG ES<sup>4</sup>. The proposed optimization model is highly transparent and versatile in its adjustability for various applications and solution methodologies.

We use the model to evaluate the operational performance of the PtG ES<sup>4</sup>, considering energy prices of the German day-ahead markets since the energy crisis in late 2021, and find that 36% of the overall methane consumption comes from PtG technologies. Additionally, the PtG ES<sup>4</sup> can achieve net-zero direct greenhouse gas emissions with a cost increase of a factor of 2.3.

**Keywords:** low-carbon, long-term, large-scale MI(N)LP, linearization, decomposition.

## 1. Introduction

Future energy systems typically have a high need for energy storage, specifically long-term energy storage, to account for seasonal fluctuations in energy demand and renewable energy supply. Due to their capability to shift energy on seasonal time scales, Power-to-Gas (PtG) technologies show high potential (Blanco and Faaij, 2018). Through electrolysis and methanation, electricity can be chemically stored for later use, including reconversion to electricity.

In the literature, PtG technologies are receiving increased attention in the community of energy system modelers. Wang and Yuan (2020) build an operational optimization model for integrated electrical and natural-gas systems with PtG. In their work, however, the PtG technologies are modeled with low levels of technical detail and optimized with low temporal resolution. Gabrielli et al. (2018) propose a detailed mixed-integer linear programming (MILP) formulation of PtG technologies to design multi-energy systems with an hourly resolution of an entire year but do not count methanation options in. Zhang and Zhang (2020) present a detailed scheduling model for integrated energy systems, including an electrolyzer, a methanation reactor, and a carbon capture process with carbon dioxide (CO<sub>2</sub>) storage. Yet, they still use a typical day with an hourly resolution.

The description of seasonal cycles requires a time horizon of one year, while the operation decisions call for hourly resolution (Gabrielli et al., 2018). Still, PtG technologies are

often considered with limited technical or temporal detail to balance the model complexity and computational effort.

In this contribution, we mathematically describe a power-to-gas energy system for seasonal energy supply and storage (ES<sup>4</sup>) with great technical and temporal detail. Firstly, we propose a mixed-integer nonlinear programming (MINLP) formulation for operational optimization. Then, the MINLP problem is automatically linearized. In this way, the mathematical formulation can be well adapted to suit different solution approaches for computational tractability. Finally, we solve the resulting large-scale MILP problems via a decomposition-based method. Section 2 explicates the modeling and optimization of the PtG ES<sup>4</sup>. Accordingly, in Section 3, we analyze how the economic competitiveness of PtG technologies has evolved since the energy crisis started in late 2021. Section 4 concludes the work.

## 2. Modeling of a Power-to-Gas energy system

### 2.1. Structure of a PtG ES<sup>4</sup>

Figure 1 shows the structure of a PtG ES<sup>4</sup>, which aims to fulfill electricity (e-) and heating (h-) demands via energy conversion or purchase from electricity and gas (g-) grids.

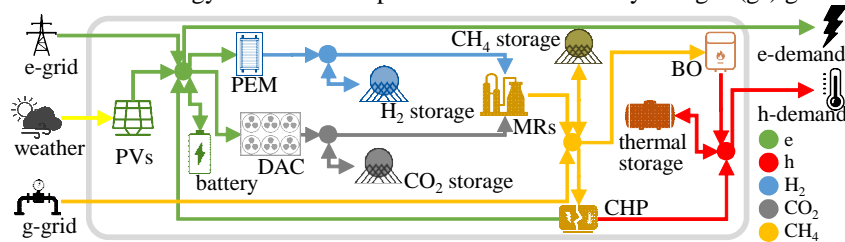


Figure 1: Structure of a PtG ES<sup>4</sup>.

In total, there are 11 components inside the PtG ES<sup>4</sup>: Photovoltaic panels (PVs) generate electricity under different weather conditions. The electricity can be used to cover the e-demand directly, be stored in a battery, or be utilized to operate a proton exchange membrane electrolyzer (PEM) and a direct-air-capture plant (DAC). The hydrogen (H<sub>2</sub>) and CO<sub>2</sub> produced by the PEM and the DAC can be fed into methanation reactors (MRs) for methane (CH<sub>4</sub>) production. Three gas tanks are available to store H<sub>2</sub>, CO<sub>2</sub>, and CH<sub>4</sub>, respectively. Via a boiler (BO) and a combined heat and power system (CHP), CH<sub>4</sub> can be burnt to provide the heat (h), which can be used to satisfy the h-demand or be stored in a thermal storage for later use. Another product of the CHP is electricity. Five buses indicate five balance constraints for the product e, h, H<sub>2</sub>, CO<sub>2</sub>, and CH<sub>4</sub>, respectively.

The described PtG ES<sup>4</sup> is based on a real-world pilot energy system located in Duisburg, Germany (Leitmarkt Agentur.NRW, 2019). We increase the nominal size of the components ( $\dot{E}_i^N$  and  $E_i^N$ ) to allow energy supply and storage on a seasonal scale, as summarized in Table 1. Please note that this work focuses on the components reflected in a real-world energy system. Further decarbonization potentials, such as heat integration and alternative CO<sub>2</sub> supply technologies, are outside the scope of this study but can be implemented with ease.

### 2.2. Mathematical description of the PtG ES<sup>4</sup>

#### 2.2.1. Exogenous inputs

Exogenous inputs encompass all external energy supply and demand: e- and g-grid, weather, e- and h-demand, as shown in Figure 1. They serve as time-dependent input data

for the energy system optimization. This work uses an hourly resolution with a time horizon of one year beginning on September 1<sup>st</sup>, 2021, as shown in Figure 2.

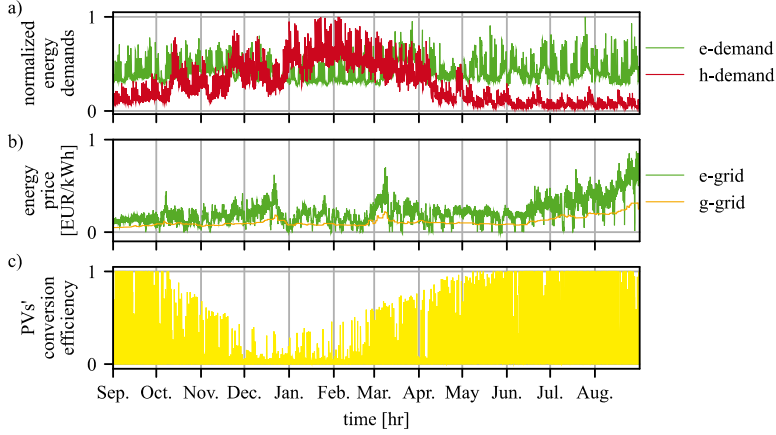


Figure 2: Input time series from September 1, 2021 to August 31, 2022: a) normalized energy demands, b) energy prices  $c_{b,t}^{\text{buy}}$  (EEX, 2022, SMARD.de, 2022), c) conversion efficiency of the photovoltaic panels (PVs)  $\eta_{PVs,t}$  (Perez et al., 1987, Eicker, 2003, Leuchner and Ketzler, 2022).

In our mathematical implementation,  $\dot{E}_{b,t}^D$  is used to indicate exogenous demands for each product  $b$  ( $\forall b \in B = \{e, h, H_2, CO_2, CH_4\}$ ) at every time step  $t$  ( $\forall t \in T$ ). Here,  $\dot{E}_{e,t}^D$  and  $\dot{E}_{h,t}^D$  are measured data of an office building;  $\dot{E}_{H_2,t}^D$ ,  $\dot{E}_{CO_2,t}^D$  and  $\dot{E}_{CH_4,t}^D$  equal zero. We take the energy prices from the German day-ahead markets (EEX, 2022, SMARD.de, 2022) as energy buying prices  $c_{b,t}^{\text{buy}}$ , where  $\forall b \in B' = \{e, CH_4\}$ . We apply the measured weather data from Leuchner and Ketzler (2022), and directly process the data to obtain the conversion efficiency of photovoltaic panels  $\eta_{PVs,t}$  (Perez et al., 1987, Eicker, 2003).

### 2.2.2. Conversion components

Conversion components include all components that convert one product to any other product(s) ( $\forall i \in I \setminus I^S$ ), where  $I = \{\text{PVs, PEM, DAC, MRs, BO, CHP, battery, thermal storage, } H_2 \text{ storage, } CO_2 \text{ storage, } CH_4 \text{ storage}\}$  and  $I^S = \{\text{battery, thermal storage, } H_2 \text{ storage, } CO_2 \text{ storage, } CH_4 \text{ storage}\}$ . We use input-output relationships to model the performance of conversion components in the PtG ES<sup>4</sup>. The main parameters, adapted models and their references are summarized in Table 1 (1a - 1h).  $\dot{E}_{i,b,t}^{\text{in}}$  and  $\dot{E}_{i,b,t}^{\text{out}}$  express the input and output of each component  $i$  ( $\forall i \in I \setminus I^S$ ) at each time step  $t$  ( $\forall t \in T$ ). The index  $b \in B$  indicates if there is more than one input-output relationship for the component  $i$ .  $\eta_i^N$  is the nominal conversion efficiency. Besides, the performance of each component is limited by its minimal and maximal part-load fraction  $p_i^{\text{min}}$  and  $p_i^{\text{max}}$ .

### 2.2.3. Storage components

Storage components transport products from one time step to other time steps ( $\forall i \in I^S$ ). The mathematical description of the five storage components is shown in Table 1 (1i - 1m). The state of charge (SOC) of a storage component  $i$  ( $\forall i \in I^S$ ) at the time step  $t + 1$  depends on its'  $SOC_{i,b,t}$ , product input  $\dot{E}_{i,b,t}^{\text{in}}$  and product output  $\dot{E}_{i,b,t}^{\text{out}}$  at the last time step  $t$  considering the efficiency of self-discharge  $\eta_i^{\text{self}}$ , charging  $\eta_i^{\text{in}}$  and discharging  $\eta_i^{\text{out}}$ , respectively (1i). The so-called cycling constraint ensures that the product is conserved (1j). At the same time, the product input and output flow is limited by the charging and discharging rate  $c_i^{\text{in}}$ ,  $c_i^{\text{out}}$  (1k, 1l). Also, the binary variables  $\delta_{i,t}^{\text{in}}$  and  $\delta_{i,t}^{\text{out}}$  prevent the charging and discharging behavior from occurring at the same time step (1m).



Table 1: Mathematical models of components inside the PtG ES<sup>4</sup>, photovoltaic panels (PVs), a proton exchange membrane electrolyzer (PEM), a direct-air-capture plant (DAC), methanation reactors (MRs), a combined heat and power system (CHP), a boiler (BO), and five storage technologies for electricity, heat, H<sub>2</sub>, CO<sub>2</sub> and CH<sub>4</sub>, respectively.<sup>a</sup>

$t \in \Gamma^S$	$E_t^N$	$E_{i,b,t}^{in}$	$E_{i,b,t}^{out}$	$\eta_i^N$	$p_i^{\min}, p_i^{\max}$	Input-output relationship(s)	Ref.
PVs	900 kW <sub>e</sub>	-	e	1	0	$E_{i,e,t}^{out} \leq \eta_{PVs,t}(\text{weather}) \cdot E_t^N$	Baumgärtner et al. (2020) (1a)
PEM	900 kW <sub>e</sub>	e	H <sub>2</sub>	0.63	0.2	$E_{i,H_2,t}^{out} = \eta_{i,H_2,t} \cdot E_{i,e,t}^{in}$ , with $\eta_{i,H_2,t} = \frac{607.68p^4 + 11007.9p^3 - 52.60p^2 + 1.49p}{1041.50p^4 + 686.68p^3 + 10.90p^2 - 2.33p + 0.12} \eta_i^N$ , $p = \frac{E_{i,H_2,t}^{out}}{E_t^N}$	Kopp et al. (2017) <sup>b</sup> (1b)
DAC	250 kg <sub>CO<sub>2</sub>}/hr</sub>	e	CO <sub>2</sub>	0.5	0	$E_{i,CO_2,t}^{out} = \eta_i^N \cdot E_{i,e,t}^{in}$	Fasih et al. (2019) (1c)
MRs	1500 kW <sub>CH<sub>4</sub>}}</sub>	H <sub>2</sub> CO <sub>2</sub>	CH <sub>4</sub>	0.78	0	$E_{i,CH_4,t}^{out} = \eta_i^N E_{i,H_2,t}^{in}$ $E_{i,CO_2,t}^{in} = 0.178 E_{i,CH_4,t}^{out}$	Gorre et al. (2020) (1d) (1e)
CHP	800 kW <sub>h</sub>	CH <sub>4</sub>	h e	0.87 (0.47 for h 0.4 fore)	0.5	$E_{i,h,t}^{out} = \eta_{i,h,t} E_{i,CH_4,t}^{in}$ $E_{i,e,t}^{out} = \eta_{i,e,t} E_{i,CH_4,t}^{in}$ , with $\eta_{i,h,t} = -0.135p^2 + 0.605p$ $\eta_{i,e,t} = -0.12p^2 + 0.52p$ , $p = \frac{E_{i,h,t}^{out}}{E_t^N}$	Baumgärtner et al. (2020) (1f) (1g)
BO	500 kW <sub>h</sub>	CH <sub>4</sub>	h	0.9	0.001	$E_{i,h,t}^{out} = \eta_{i,h,t} \cdot E_{i,CH_4,t}^{in}$ , with $\eta_{i,h,t} = \frac{84.35p^3 - 16.34p^2 + 2.68p}{79.14p^3 - 10.39p^2 - 0.79p + 0.73} \eta_i^N$ , $p = \frac{E_{i,CH_4,t}^{out}}{E_t^N}$	Baumgärtner et al. (2020) (1h)
$t \in \Gamma^S$	$E_t^N$	$E_{i,b,t}^{in/out}$	$\eta_i^{self}$	$\eta_i^{in} = \eta_i^{out}$	$c_i^{in} = c_i^{out}$	Input-output relationship(s)	Ref.
battery	5000 kW <sub>h</sub> <sub>e</sub>	e	0.000042 1/hr	0.92	0.36 1/hr	$SOC_{i,b,t+1} = SOC_{i,b,t}(1 - \eta_i^{self} \Delta t)$	Baumgärtner et al. (2020) (1i)
thermal storage	5000 kW <sub>h</sub>	h	0.005 1/hr	0.95	1 1/hr	$+ \Delta t \left( \eta_i^{in} E_{i,b,t}^{in} - \eta_i^{out} E_{i,b,t}^{out} \right)$ $SOC_{i,b,t+1} = SOC_{i,b,t} + \eta_i^{in} \Delta t$	Baumgärtner et al. (2020) (1j)
gas tanks <sup>c</sup>	5000 kW <sub>h</sub> <sub>H<sub>2</sub>}}</sub>	H <sub>2</sub>	0 1/hr	1	0.25 1/hr	$0 \leq E_{i,b,t}^{in} \leq \delta_{i,t}^{in} c_i^{in} E_t^N$ , $0 \leq E_{i,b,t}^{out} \leq \delta_{i,t}^{out} c_i^{out} E_t^N$ $\delta_{i,t}^{in} + \delta_{i,t}^{out} \leq 1$	Gabrielli et al. (2018) (1k,1l) (1m)

<sup>a</sup> Index H<sub>2</sub> and CH<sub>4</sub> indicate that the value is related to the high heating value of H<sub>2</sub> and CH<sub>4</sub>, respectively.

<sup>b</sup> The model is based on measured data in Kopp et al. (2017).

<sup>c</sup> We assume that the efficiencies, charging and discharging rates of the CO<sub>2</sub> and CH<sub>4</sub> storage tanks are equal to those of the H<sub>2</sub> storage tank.

### 2.3. MINLP formulation of the operational optimization problem

In Eqs. (1), we state the generic operational optimization problem using an MINLP formulation. Given the input time series (Section 2.2.1), technical constraints of each energy system component (Section 2.2.2 and Section 2.2.3) and other limitations (1p), the energy system aims to minimize the operational expenditures (1n) while fulfilling the energy demands at each time step (1o).  $\Delta t_t$  denotes the length of each time step  $t$ . All decision variables are summarized in (1q). The surrogate vectors  $\mathbf{x}$  and  $\mathbf{y}$  represent other decision variables that are not specified here, including on/off decisions of components.

$$\min \text{OpEx}^{\text{ES}} = \sum_{t \in T} \Delta t_t (\sum_{b \in B'} (c_{b,t}^{\text{buy}} \dot{E}_{b,t}^{\text{buy}})) \quad (1n)$$

$$\text{s.t.} \quad \sum_{i \in I} (\dot{E}_{i,b,t}^{\text{out}} - \dot{E}_{i,b,t}^{\text{in}}) + \dot{E}_{b,t}^{\text{buy}} = \dot{E}_{b,t}^{\text{D}} \quad \forall b \in B, \forall t \in T \quad (1o)$$

$$h(\dot{E}_{i,b,t}^{\text{in}}, \dot{E}_{i,b,t}^{\text{out}}, \text{SOC}_{i,b,t}, \delta_{i,t}^{\text{in}}, \delta_{i,t}^{\text{out}}, \mathbf{x}, \mathbf{y}) \leq 0 \quad \forall i \in I, \forall b \in B, \forall t \in T \quad (1p)$$

$$\begin{aligned} \dot{E}_{i,b,t}^{\text{in}}, \dot{E}_{i,b,t}^{\text{out}} &\in \mathbb{R}^+; 0 \leq \text{SOC}_{i,b,t} \leq E_i^{\text{N}}; \delta_{i,t}^{\text{in}}, \delta_{i,t}^{\text{out}} \in \{0, 1\}; & \forall i \in I, \forall b \in B, \forall t \in T & (1q) \\ \mathbf{x} \in \mathbb{R}^{\text{N}_x}; \mathbf{y} \in \{0, 1\}^{\text{N}_y} & & & \end{aligned}$$

### 2.4. Implementation and Solution

We model the PtG ES<sup>4</sup> on the COMANDO platform (Langiu et al., 2021). The MINLP formulation is transformed into an MILP formulation via the COMANDO built-in linearization function. We implement a decomposition-based method (Baumgärtner et al., 2020) to solve the resulting large-scale MILP problem with an optimality gap of 2 %.

We assess the PtG ES<sup>4</sup> from an economical and environmental perspective using two scenarios: with (1) and without (2) exogenous methane supply.

## 3. Results and discussion

Figure 3 illustrates the resulting operational performance of the PtG ES<sup>4</sup> of Scenario 1. The left side shows the energy consumption from the grids and the part-load fraction of the PtG components. At the beginning and end of the time horizon, where sufficient electricity of PVs is available, PtG components are operated frequently. In particular, after  $t = 7000$  hr (around April 8<sup>th</sup>, 2022), when energy prices increase dramatically, gas supply via PtG dominates. Overall, 36% of the total gas supply is from PtG components.

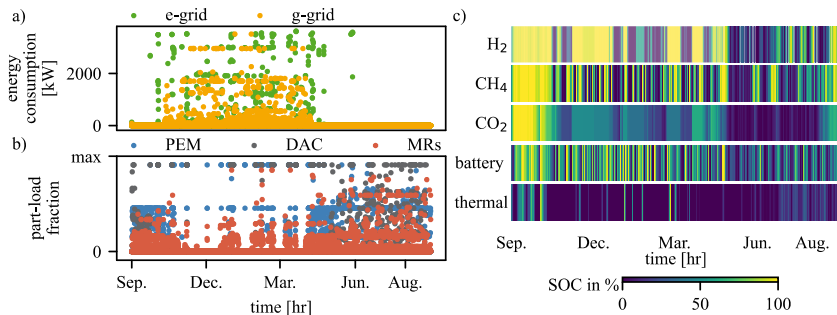


Figure 3: Operational performance: a) energy consumption from electricity and gas grid; b) part-load fraction of PtG components; c) state of charge (SOC) behavior of all storage components.

The right side shows the SOC behavior of storage components. Gas storage tanks perform as expected on large time scales. The battery is preferred over thermal storage systems. In our analysis, the current energy crisis makes PtG technologies economically viable. Please note that this work focuses on the operation of an existing energy system and thus does not consider capital costs.

The optimal operational performance in Scenario 2 shows a similar pattern to Scenario 1. As the exogenous energy supply is solely from PVs and the e-grid, the PtG ES<sup>4</sup> achieves net-zero direct greenhouse gas emissions; however, the cost increases by a factor of 2.3. Each computation involved in this work includes over 890 000 variables and 1 000 000 constraints after linearization. The computational time varies between 30 minutes to 2 hours depending on the configuration.

#### 4. Conclusions

PtG technologies are promising candidates for seasonal energy supply and storage for future energy systems. However, due to seasonal fluctuations, optimizing the operation of a PtG ES<sup>4</sup> is computationally challenging. We introduce a modeling and optimization approach based on a real-world PtG ES<sup>4</sup>. The proposed model involves large-scale operational decisions with great technical and temporal detail. Meanwhile, the resulting operational problems remain computationally tractable.

Considering the energy crisis since late 2021, we find that the natural gas supply from PtG has become economically competitive due to the dramatic increase in energy prices. We also show that a net-zero direct greenhouse gas emissions energy system is possible with a cost increase by a factor of 2.3. Overall, we expect our modeling and operational optimization approach to be transferable to other cases with different exogenous inputs.

#### Acknowledgments

This study is funded by the 'Europäischer Fonds für regionale Entwicklung (EFRE)' (EFRE-0801844). The support is gratefully acknowledged.

#### References

- N. Baumgärtner, D. Shu, B. Bahl, M. Hennen, D.E. Hollermann, A. Bardow, 2020, DeLoop: Decomposition-based Long-term operational optimization of energy systems with time-coupling constraints, *Energy*, vol. 198, 4, p. 117272
- H. Blanco, A. Faaij, 2018, A review at the role of storage in energy systems with a focus on Power to Gas and long-term storage, *Renewable and Sustainable Energy Reviews*, vol. 81, p. 1049
- EEX, 2022, Natural Gas Day Ahead Market data: GASPOOL and THE, <https://www.eex.com/en/market-data/natural-gas/spot>. Accessed 2022.10.13
- U. Eicker, 2003, *Solar technologies for buildings*. John Wiley & Sons
- M. Fasihi, O. Efimova, C. Breyer, 2019, Techno-economic assessment of CO<sub>2</sub> direct air capture plants, *Journal of Cleaner Production*, vol. 224, 3, p. 957
- P. Gabrielli, M. Gazzani, E. Martelli, M. Mazzotti, 2018, Optimal design of multi-energy systems with seasonal storage, *Applied Energy*, vol. 219, 7, p. 408
- J. Gorre, F. Ruoss, H. Karjunen, J. Schaffert, T. Tynjälä, 2020, Cost benefits of optimizing hydrogen storage and methanation capacities for Power-to-Gas plants in dynamic operation, *Applied Energy*, vol. 257, 1–2, p. 113967
- M. Kopp, D. Coleman, C. Stiller, K. Scheffer, J. Aichinger, B. Scheppat, 2017, Energiepark Mainz: Technical and economic analysis of the worldwide largest Power-to-Gas plant with PEM electrolysis, *International Journal of Hydrogen Energy*, vol. 42, 19, p. 13311
- M. Langiu, D.Y. Shu, F.J. Baader, D. Hering, U. Bau, A. Xhonneux, D. Müller, A. Bardow, A. Mitsos, M. Dahmen, 2021, COMANDO: A Next-Generation Open-Source Framework for Energy Systems Optimization, *Computers & Chemical Engineering*, vol. 152, 11, p. 107366
- Leitmarkt Agentur.NRW, 2019, Projekt PtG-MSE: Power to Gas-Modellvorhaben für sektorübergreifende Energiesysteme, [https://www.leitmarktagentur.nrw/lw\\_resource/datapool/\\_items/item\\_763/pdb\\_kesw-1-2-006.pdf](https://www.leitmarktagentur.nrw/lw_resource/datapool/_items/item_763/pdb_kesw-1-2-006.pdf). Accessed 2020 -10-15
- M. Leuchner, G. Ketzler, 2022, Klimamessstation Aachen-Hörn - Monatsbericht 2021-2022. RWTH Aachen University. Department of Geography. Physical Geography and Climatology Group. ISSN 1861-3993
- R. Perez, R. Seals, P. Ineichen, R. Stewart, D. Menicucci, 1987, A new simplified version of the perez diffuse irradiance model for tilted surfaces, *Solar Energy*, vol. 39, 3, p. 221
- SMARD.de, 2022, Strommarkt, <https://www.smard.de/home/downloadcenter/download-marktdaten>. Accessed 2022.09.01
- S. Wang, S. Yuan, 2020, Interval optimization for integrated electrical and natural-gas systems with power to gas considering uncertainties, *International Journal of Electrical Power & Energy Systems*, vol. 119, 1, p. 105906
- X. Zhang, Y. Zhang, 2020, Environment-friendly and economical scheduling optimization for integrated energy system considering power-to-gas technology and carbon capture power plant, *Journal of Cleaner Production*, vol. 276, 3, p. 123348

# Liquid organic hydrogen carriers (LOHC) as a solution for the energy transition: a reactor design assessment

Carlos Prieto<sup>a</sup>, Antonio Sánchez<sup>a</sup>, Mariano Martín<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, University of Salamanca, Salamanca, 37003, Spain*

## Abstract

Liquid organic hydrogen carriers (LOHCs) use for green hydrogen storage has become a promising option for an effective decarbonization. The core of the process relies on the hydrogenation/dehydrogenation reaction, therefore rigorous reactor models are required. In this work, slurry and trickle bed reactors are modelled for LOHCs hydrogenation/dehydrogenation. Kinetic models as well as mass, energy, and momentum transfer are considered. The influence of operational variables was evaluated jointly with design parameters. The benefits and limitations of each design were highlighted discovering in some cases the presence of two regions: mass controlling and a second where kinetic begins to contribute. For dehydrogenation reactors, the main barrier is the high endothermicity of the reaction limiting the kinetic step of the entire process.

**Keywords:** energy storage, energy carriers, green hydrogen, LOHC, reactor design.

## 1. Introduction

The introduction of renewable resources is essential to decarbonize the energy system which, including transportation, is responsible for around 2/3 of global greenhouse gas emissions. As a result, deployment of renewables will play an important role in the sustainable energy transition. The two most important renewable sources in the near future are expected to be wind and solar. Nevertheless, both are highly fluctuating, which unavoidably introduces a new challenge in a system with high penetration of renewables. Therefore, to achieve an effective decarbonization, the use of energy storage systems and energy carriers is mandatory.

At this point, green hydrogen emerges as one of the most important actors in the new energy system. This energy carrier can be used efficiently to decarbonize different energy demands that are hard to directly electrify, for example, heavy or maritime transport. In addition, hydrogen is also proposed as an energy storage system on a seasonal horizon. A major barrier is the difficulties regarding hydrogen storage and transportation as it is one of the lightest chemicals with a low volumetric energy density. To overcome these issues, different hydrogen storage technologies are available. Some of them are conventional such as high pressure or liquified hydrogen, although there also exist other arising options as metal hydrides or metal organic frameworks. Among the novel alternatives, liquid organic hydrogen carriers (LOHCs) are emerging as a promising option. The idea behind LOHCs is to hydrogenate a liquid chemical to improve the storage and transportation of hydrogen. Then, when hydrogen is required, the liquid is dehydrogenated. The main advantages of LOHCs are the reaction reversibility, moderate reaction temperatures, commercial availability of LOHCs, and compatibility with the existing oil infrastructure

(He et al., 2015). Different LOHCs have been proposed such as n-ethylcarbazole or dibenzyltoluene looking for the ideal one with good manageability at liquid conditions, reduced cost, safety, and non-toxicity (Niermann et al., 2021).

Different research works have investigated the use of LOHCs from two different perspectives. Firstly, different authors have explored the hydrogenation/dehydrogenation reactions from a kinetic point of view to determine the kinetic rate for the different LOHCs because reaction step is the core section since only additional stream conditioning is required in the process (Salman et al., 2022). In addition, the global performance of these LOHCs has also been evaluated from a process perspective (Niermann et al., 2019). But the details of the chemical hydrogenation/dehydrogenation reactions beyond kinetics have not been addressed. In this work, an intermediate step between these areas is explored focusing on the reactor design of the hydrogenation/dehydrogenation unit. Thus, an accurate assessment can be performed at process scale using realistic data of the reactor performance. Therefore, this evaluation is essential to support deployment of LOHCs for hydrogen storage systems. Particularly, in this work, the most extended three-phase reactor designs have been analyzed: trickle bed and slurry. For both, a rigorous model is developed including principles of mass, energy, and momentum transfer to describe the complex system of a hydrogenation/dehydrogenation reactor. Using this model, it is possible to determine the main variables that affect the design and performance of the LOHCs reactors determining the best operating conditions of each of the alternatives. Specifically, two of the most promising LOHCs systems are evaluated: dibenzyltoluene and a mixture of 1,2-dimethylindol and 7-ethylindole. However, this analysis can be easily extended to other LOHCs systems.

## 2. Reactor modelling

### 2.1 Trickle bed reactor model

Trickle bed reactors can be applied for both hydrogenation and dehydrogenation reactions. This is because liquid-to-particle mass transfer is intensified under trickle and high interaction conditions and a hydrogen gas acceptor will favor hydrogen transference to the gas phase, avoiding hydrogen accumulation in the liquid media.

The design complexity of trickle bed reactors is found in the operating regimes. They mainly depend on the gas and liquid flows. The most common regimes are high interaction and the transition between high interaction and trickle, which are used in industrial activities and, also, in the proposed model. Its flow can be co and counter-current, although in general, co-current with downward flow is the preferred option.

The performance of trickle bed reactors can be assumed to be a plug flow model with no axial dispersion. Also, the consideration that both phases remain at the same temperature is well accepted. With these assumptions, the mass transfer between phases and the chemical reaction and diffusion inside catalyst pores can be described given by Eqs (1-3) for hydrogen reaching the catalyst active center. Liquid species such as LOHCs are assumed not to be transferred to the gaseous phase because of their low volatility. They will only transfer into catalyst surface and then will perform parallel mechanisms of diffusion and reaction. Because of that, only Eqs (2-3) will be applied.

$$\frac{(dF_{H_2})_g}{dW} = (k_l a_g)_{H_2} \cdot ((C_{H_2})_g^* - (C_{H_2})_l) \quad (1)$$

$$\frac{(dF_i)_l}{dW} = (k_m a_m)_i \cdot ((C_i)_l - (C_i)_s) \quad \forall i \quad (2)$$

$$(k_m a_m)_i \cdot ((C_i)_l - (C_i)_s) = \eta \cdot r_{surface} \quad \forall i \quad (3)$$

F represents the molar flow rate of each component i, C is the concentration in the gas (g) or liquid (l) phase, denoting with an asterisk the equilibrium concentration,  $k_l a_g$  is the mass transfer coefficient for the gas-liquid transport and  $k_m a_m$  for the liquid-solid transport. Finally,  $r_{surface}$  is the surface reaction rate corrected using the effectiveness factor ( $\eta$ ). It must be noticed, that for the reaction to occur, both reactants must reach the catalytic center. Otherwise, the reaction cannot progress. Because of that, the model accounts for the transfer of both species and reaction velocity taking the lowest as the overall rate, assuring that all reactants are transferred. Additionally, it must be considered that ideally, catalyst particles will be surrounded by the liquid phase. However, to describe the actual situation, wetting factor is required to account for the catalyst real effectiveness.

Related to heat transfer, an adiabatic operation is selected, and the temperature radial gradient can be discarded. Also, the intraparticle gradient can be neglected, although others such as longitudinal (Eq.4) and interparticle (Eq.5) cannot be removed.

$$\frac{dT}{dW} = \frac{r \cdot (-\Delta Hr)}{\sum_{i=1}^n F_i \cdot Cp_i} \quad (4)$$

$$h \cdot a_m \cdot (T_s - T_b) = r \cdot (-\Delta Hr) \quad (5)$$

In these equations, h is the convection heat transfer coefficient and Ts and Tb the temperatures in the surface of the particle (s) and bulk liquid (b). To predict pressure drop along the bed, a modification of the Ergun's equations for a two-phase flow is selected. Another relevant parameter is the reactor volume fraction occupied by liquid, defined as liquid hold-up. It can be divided into internal and external, and this last into static and dynamic.

## 2.2 Slurry reactor model

In slurry reactors, a liquid phase holds finer particles than those in the trickle bed, thanks to the action of a bubbled gas introduced at the bottom and dispersed with a gas sparger. Slurry reactors hydrodynamics can be divided into slug flow, churn turbulent, and homogeneous. Industrial columns tend to prefer churn-turbulent, at large diameter and superficial gas velocities, as heat and mass transfer are intensified.

Slurry reactors are known to be well-mixed units. As a result, liquid and solid phases can be understood as a pseudo-phase and modelled considering perfect mixture. Nevertheless, the hydrogen is better described as a plug flow since it is not mixed with the liquid. Variables such as gas hold-up and bubble diameter have a great influence on hydrogen transference from the gas to the liquid. This mandates to characterize the features of the bubble. For the sake of simplicity, a bi-modal distribution (large and small bubbles) is considered. Thus, gas-to-liquid hydrogen transfer can be modelled with Eq.6. Assuming steady-state, reaction velocity can be computed following Eq.7. As same as with trickle bed reactors, the slurry model must assure that all species are able to perform all transfer steps. For that, both reactive velocities are calculated and the slowest is taken.

$$-\frac{(dF_{H_2})_g}{dV} = \left( k l a_{g_{small}} + k l a_{g_{large}} \right)_{H_2} \cdot \left( (C_{H_2})_g^* - (C_{H_2})_l \right) \quad (6)$$

$$(k_m a_m)_i \cdot ((C_i)_l - (C_i)_s) = \eta \cdot r_{surface} \quad \forall i \quad (7)$$

Where  $kla_{g_{small}}$  and  $kla_{g_{large}}$  are the gas-liquid transfer coefficient for the two bubble sizes. Slurry reactors are considered to have proper temperature control; therefore, an isothermal operation can be performed if heat transfer internals are used. Gas sparger design is also required. In general, it is preferred to use pipe spargers for higher diameter reactors. Its design is limited by weeping phenomena, fixing minimum hole velocity.

### 3. Results

With the proposed models, it is possible to simulate both reactor performance: trickle beds and slurry ones, for hydrogenation and dehydrogenation and evaluate how operational and design parameters impact on LOHCs conversion.

Hydrogenation reactors account for velocities of both reactants. In general, when LOHCs concentration is higher, the limiting step is hydrogen gas-liquid transfer. Figure 1 presents the results for the dibenzyltoluene hydrogenation slurry reactor. A first gas-to-liquid hydrogen transfer controlling region can be distinguished accounting for nearly 75-85% of the total resistance. The superficial gas velocity (directly related with the excess of hydrogen) seems to have beneficial effects on conversion as gas holdup and mass transfer coefficients are enhanced while kinetics has negligible influence. Mass transfer from bulk liquid to particle is responsible for the remaining 15-20%. This behavior ends up with a second tendency, in which, hydrogen transference reduces its importance and LOHCs concentration depletion becomes responsible for increasing kinetic contribution. This change of behavior ends up with all superficial gas velocity lines converging, which means that its influence becomes negligible.

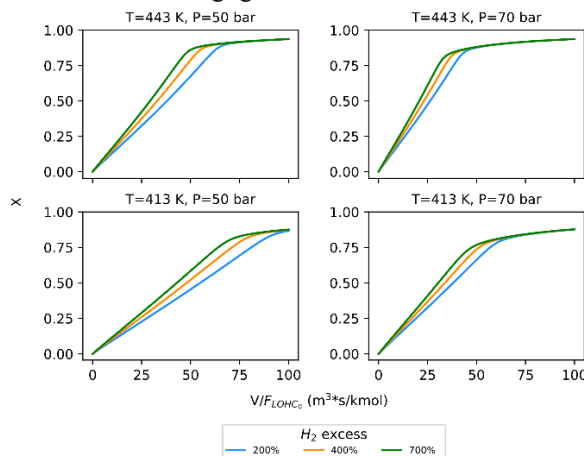


Figure 1 Hydrogenation conversion against normalized volume for dybenzyltoluene on slurry reactor with  $100 \text{ kg/m}^3$  catalyst load

Temperature and pressure influence can be also evaluated from Figure 1. Both will increase hydrogen solubility in the proposed liquid systems. Higher gas solubility favors the mass transfer, increasing conversion, and reaching the second operating window earlier. Higher conversions could be obtained if normalized volume increases. Nevertheless, it is preferred to avoid unit oversizing since the second region is not as favorable as the first one. In general, perfect mixture reactors are usually employed for intermediate-high conversions. In this specific case, conversion can range within 75-90% if the reactor works in the gas-liquid mass transfer controlling region. In fact, the first region slope is 12 times higher than the second one for dibenzyltoluene, and 17 and 3

times for 1,2-dimethylindole and 7-ethylindole respectively. It clearly indicates a better catalyst use and the convenience to operate and design the reactor in the first zone. Working on this region, expected capital costs range from 1.5 to 2.5 MM $\$/s/kmol$  H<sub>2</sub>. Related to trickle bed reactors, the conversion is plotted against normalized catalyst weight for 1,2-dimethylindole and 7-ethylindole system hydrogenation in Figure 2. In general, trickle bed reactors are considered for applications with low reaction velocities. Thus, feed is preferred not to be pure. In addition, temperature rise can damage the catalyst if there is an inadequate heat removal from catalyst particles. For that, a molar relation of reactant to product of 1:4 is fed to the reactor.

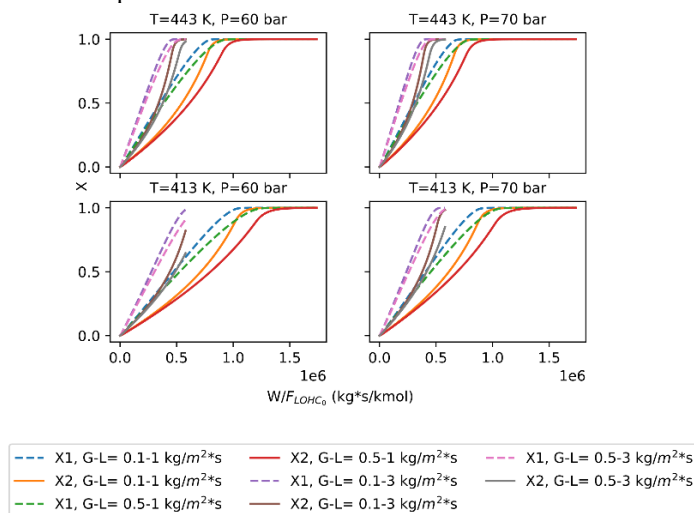


Figure 2 Hydrogenation conversion against normalized catalyst weight for 1,2-dimethylindole and 7-ethylindole 1:1 mixture on trickle bed reactor

Although LOHCs concentrations are lower, from this analysis, the gas-to-liquid step is attributed to be the main resistance contributor with approximately 95%. The remaining 5%, corresponds to liquid-to-particle transference. Only in the end, trickle bed reactors find a sharp increase in kinetic resistance due to extremely lower concentrations, becoming the main limiting step (60-80%). This fact is captured in Figure 2 with a change of tendency on conversion pattern. Notwithstanding, liquid-to-particle transference will be still relevant (20-40%).

Gas and liquid flowrates for a given diameter can be varied to determine their effect on conversion. While increasing liquid flows has evident benefits, the gas flow rate will have a contrary effect. Higher liquid flows increase mass transfer coefficients. It results in better catalyst use for the same amount of reactive. However, at upper gas flowrates, the catalyst bed pressure drop will be higher. It will cause a faster reduction of hydrogen solubility in the liquid phase. Undoubtedly, catalyst usage per hydrogenated LOHC is considerably higher on trickle beds than on slurry reactors. It has a clear impact on the global reactor cost, as from 20 to 40 MM $\$/s/kmol$  H<sub>2</sub> are estimated.

However, as it can be seen in Figure 2, this kind of reactor allows to reach higher conversions with lower concentrations of reactants as the kinetic controlling region is only found at the end of the reaction. Therefore, a proper combination of both designs can be explored to take advantage of the benefits of the reaction technologies. For instance, a first slurry reaching almost 80% conversion before the flattened curve region can be set, and then a trickle bed reactor to fully hydrogenate the LOHC.



During dehydrogenation, hydrogen transfer no longer limits reaction progress. Nonetheless, proper hydrogen mass transfer to gas must be assured to avoid accumulation. In general, it is easy to accomplish as lower pressures are used, which considerably reduces its solubility. Slurry dehydrogenation reactors are clearly limited in both systems by kinetics. Thus, the main variables affecting conversion are temperature and catalyst load. As dehydrogenation kinetics are considerably lower than hydrogenation, a great unit oversizing is required compared with previous designs. To achieve similar conversions of 80%, higher particle concentrations are used. Related to trickle bed dehydrogenation reactors, kinetics is also the main resistance. Reaction endothermicity highly limits the rate as the temperature is reduced through the catalyst bed. It makes necessary to introduce a high fraction of dehydrogenated LOHC to avoid higher temperature drops. In this case, the molar relationship between reactive and product is increased to 1:9. Both 1,2-dimethylindol and 7-ethylindole can be almost fully dehydrogenated with 10000 kg·s/mol. Nevertheless, reaching higher conversions is not possible for dibenzyltoluene as kinetics are slow, and is not possible to introduce higher catalyst weight without higher pressure drop in systems operating at lower pressures. It limits conversions to values up to approximately 25-40%.

#### 4. Conclusions

A detailed reactor modelling was performed for hydrogenation and dehydrogenation of LOHCs with slurry and trickle bed designs. Hydrogenation slurry was characterized by having two differentiated zones: a mass transfer controlled region where gas superficial velocity has a positive influence, and a second where kinetics increases its contribution to the total resistance and superficial gas velocity no longer has effect. It was found that operating in the first region will work from 3 to 17 times better than in the second one. Therefore, a combination of both designs could be explored in order to maximize the global rate in each design avoiding slurry oversizing. A similar configuration could be proposed for dehydrogenation reactors. Dehydrogenation-associated endothermicity is one of the bottlenecks for LOHCs deployment, from an energetic point of view but also due to a kinetic reduction as temperature falls. The study shows that trickle bed reactors will have worse catalyst use than slurry and the unit oversizing when dehydrogenation reactors are compared to hydrogenation ones. On the whole, this work shows the great importance of the transfer phenomena in the design of LOHCs reactor and the need to go beyond traditional kinetic studies for an accurate analysis of them.

#### Acknowledgments

The authors acknowledge H2MetAmo project by the financial support.

#### References

- T. He, Q. Pei, P. Chen, 2015, Liquid organic hydrogen carriers, *Journal of Energy Chemistry*, 24(5), 587-594.
- M. Niermann, S. Drünert, M. Kaltschmitt, K. Bonhoff, 2019, Liquid organic hydrogen carriers (LOHCs)-techno-economic analysis of LOHCs in a defined process chain, *Energy & Environmental Science*, 12(1), 290-307.
- M. Niermann, S. Timmerberg, S. Drünert, M. Kaltschmitt, 2021, Liquid Organic Hydrogen Carriers and alternatives for international transport of renewable hydrogen, *Renewable and Sustainable Energy Reviews*, 135, 110171.
- M.S. Salman, N. Rambhujun, C. Prathana, K. Srivastava, K.F. Aguey-Zinsou, 2022, Catalysis in Liquid Organic Hydrogen Storage: Recent Advances, Challenges, and Perspectives, *Industrial & Engineering Chemistry Research*, 61(18), 6067-6105.

## Powering Clean and Affordable Energy through Geothermal Integration: A Case Study on Dominica.

Keara Bhagaloo<sup>a</sup>, Rehannah Ali<sup>a</sup>, Anastasia Baboolal<sup>a</sup> and Keeran Ward<sup>\*b</sup>

<sup>a</sup>Department of Chemical Engineering, The University of the West Indies, St. Augustine, Trinidad and Tobago.

<sup>b</sup>School of Chemical and Process Engineering (SCAPE), University of Leeds, Leeds, UK, LS2 9JT.

### Abstract

Caribbean economies are particularly vulnerable to climate change. While developed nations are actively making strides in becoming more sustainable, small island developing states in the Caribbean face many hurdles. Hence, this study presents techno-economic and environmental assessments to inform on sustainable options for affordable and clean energy on the island of Dominica. Four cases were explored using both current and future indicators: leveled electricity (LCOE) costs, product annualized costs (TAC) and environmental impacts. Our results illustrate that with 100% deployment of geothermal energy, greenhouse gas (GHG) emissions and LCOE can be reduced by 99.5% and 70% respectively. Furthermore, diversification linked to production of methanol and ammonia gave environmental benefits, but were infeasible based on current fossil-based prices. Ultimately, the most sustainable option was energy exportation from Dominica to Martinique via subsea transmission, leading to a 20% decrease in LCOE and 2.5 million tonne CO<sub>2</sub>-eq/year avoided emissions.

**Keywords:** Geothermal Energy, Small Island Developing States, Sustainable Energy, Techno-economics, Life Cycle Assessments.

### 1. Introduction

The Organisation of the Eastern Caribbean States (OECS)- an 11-member group of islands within the easternmost region of the Caribbean Sea, is an inter-governmental association which aims to enhance regional integration by promoting trade along with economic, environmental, and social resilience<sup>1</sup>. The organisation has thus prioritised long-term use of renewable energy-namely geothermal energy, with which the region is endowed<sup>2,3</sup>. Transitioning to renewable energy for power generation is especially important given the region's vulnerability to the effects of climate change, such as rising sea levels and increased occurrence of intense weather systems such as hurricanes<sup>4</sup>-coupled with the longstanding issue of dependence on a volatile oil sector with high prices on the global market, making power generation expensive even at low production capacities<sup>5</sup>. Additionally, energy security is limited by aging infrastructure<sup>2</sup>- which in the face of natural disasters, puts a strain on reliant sectors.

Owing to its geographical location, many forms of renewable energy exist in the Eastern Caribbean region such as wind, solar and hydroelectric power<sup>6</sup>. These sources are dependent on weather conditions and may be intermittent in supply and thus, require other dispatchable power systems for island-wide power distribution. Geologically, the region is an active volcanic region, wherein geothermal energy, independent of weather and climatic situations, can be harnessed<sup>6</sup>. Also, the abundance of untapped geothermal allows for large-scale power generation while simultaneously eliminating the need for fossil fuel. Islands such as Guadeloupe (15 MW installed capacity) and Dominica (up to 1390 MW potential)<sup>6</sup> have begun exploring geothermal energy, with ambitions to reduce emissions from local power sectors. Dominica's geothermal potential surpasses its national demand, allowing neighbouring islands with higher energy needs to benefit by the creation of an interconnected grid via subsea transmission<sup>2</sup>. Furthermore, the island's high geothermal energy potential allows for considerable energy storage systems.

Six of the eleven OECS nations are involved in the Paris Agreement, which aims to increase renewable energy penetration and reduce greenhouse gas emissions. Thus, our study utilizes multiple decision criteria using techno-economic and environmental assessments to advise on sustainable operations of the OECS power sector through a case study approach. We propose case-specific scenarios for Dominica in achieving increased renewable energy penetration. By disseminating evidence-based results, solutions can be provided to aid OECS to promote sustainable development through the deployment of clean and affordable energy (SDG 7)<sup>8</sup>.

## 2. Methodology

The following cases are aligned to Dominica's national commitment targets on the integration of renewable energy sources by 2030, Each case is evaluated based on techno-economic and environmental assessments aligned to process design, economic viability and life cycle analysis.

### 2.1 Case 1: Dominica 2030 Outlook- 100% RE Deployment

This case explores the feasibility of utilising a fraction of Dominica's geothermal potential to replace the 63% grid share dominated by fossil fuels by 2030, while power from existing hydroelectric plants constitute the remaining 37%; thereby fulfilling 100% (27MW) renewable power.

### 2.2 Cases 2 and 3: Energy Storage Platforms: MeOH and NH<sub>3</sub> Production

These cases explore the sustainable production of Methanol (MeOH) and Ammonia (NH<sub>3</sub>), using electrochemical production of hydrogen via water electrolysis, using Polymer Electrolyte Membrane (PEM) technologies. The power required for the process was derived from geothermal energy integration on the island by 2030.

For Case 2, our renewable MeOH flowsheet was designed using Aspen Plus V10 in accordance with past studies done by some of us<sup>10,11</sup>, utilizing renewable H<sub>2</sub> and CO<sub>2</sub> from Direct air capture (DAC). For Case 3, the Haber process was considered whereby NH<sub>3</sub> was produced from renewable H<sub>2</sub> and N<sub>2</sub> from cryogenic air separation-modelled using Aspen Plus V10.

### 2.3 Case 4: Geothermal Energy Exportation: Martinique 2030 Outlook

Case 4 explores the feasibility of Dominica exporting power to Martinique through sub-sea cable transmission while sustaining their own geothermal energy grid by 2030.

### 2.4 Economic Feasibility

The economic feasibility was assessed by evaluating the levelized cost of electricity (LCOE) for all four cases as well as the Total Annualized Cost (TAC) for Cases 2 and 3. The LCOE for each case was determined using Eq (1) where P<sub>ij</sub> represents the energy (MWh) derived from technology j, and LCOE<sub>j</sub> is the fixed LCOE (USD/MWh) associated with technology j.

$$LCOE = \frac{\sum_j^n (P_{ij} \times LCOE_j)}{\sum_j^n P_{ij}} \quad (1)$$

Total Annualized Cost (TAC) was computed by estimating the capital (CAPEX) and operating costs (OPEX) associated with MeOH and NH<sub>3</sub> operations. CAPEX was calculated using bare module costing parameters, with purchased equipment costs extrapolated for the cost year 2020 using Chemical Engineering Plant Cost Index (CEPCI). OPEX was estimated from fixed costs (FC) using cost allocations and variable costs (VC) such as raw material prices based on current and future market analysis. TAC scores were subsequently compiled using Eq 2, whereby the annualized fixed costs (AFC) were calculated using the annual capital cost ratio (Eq 3) and a 330 day yearly on-stream factor.

$$AFC = FCI \times \frac{(1+i)^n \times i}{(1+i)^n - 1} \quad (2)$$

where, FCI -Fixed Capital Investment (CAPEX), interest rate i = 7% and the designated plant lifetime, n = 20 years

$$TAC = \frac{AFC + OPEX}{\text{Yearly production rate}} \quad (3)$$

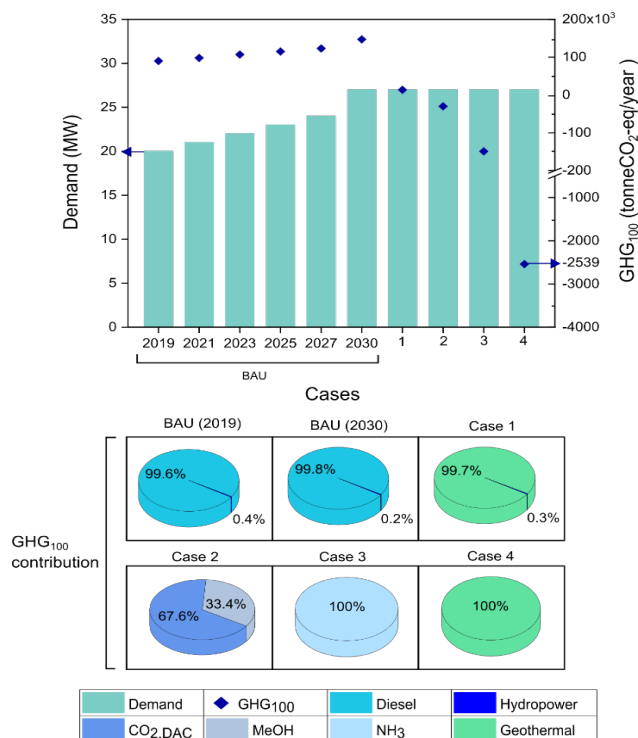
### 2.5 Life Cycle Assessment (LCA) Framework

A consequential cradle-to-power generation gate, as outlined in accordance with the ISO 14040:2006 methodology<sup>12</sup> was employed. The functional unit for all cases was set as 1MWh of energy production. Case-specific life cycle inventories were defined by mass and energy balances across system boundaries, and included electrical production from deep geothermal wells, hydroelectric run-of-river, petrol/oil, biomass, solar and wind supply chains, deionized water, natural gas (heating purposes) and flue gas emissions (MeOH/NH<sub>3</sub> process). The inventories associated with raw materials and utilities were obtained from Ecoinvent v3.4 databases. For Cases 2-4, a substitution allocation approach was employed to distribute environmental burdens among all products. Environmental benefits were evaluated at the midpoint level using ReCiPe 2016 hierarchist (H)<sup>10</sup> method within the SimaPro v9 LCA modelling platform -with global warming potential (GWP<sub>100a</sub>) as the main impact category. Finally, the LCA results were interpreted to provide sufficient evidence supporting the transition towards greater sustainable power generation.

### 3. Results and Discussion

#### 3.1 Environmental Performance

Dominica currently requires a peak demand output of 20MW (63% from diesel and 37% from hydroelectric sources)<sup>9</sup>. The business-as-usual (BAU) case assumes that existing hydroelectric power plants maintain operations at current capacity while diesel generation increases to satisfy a growing energy demand in the following years up to 2030. Increased fossil fuel use to satisfy this increasing demand leads to a 62% observed increase in GHG emissions from 90,366 tonnes CO<sub>2</sub> eq/year in 2019 to 146,831 tonnes CO<sub>2</sub> eq/year in 2030 as shown in **Figure 1**.



**Figure 1: Case specific Life Cycle GHG emission quotas (tonne CO<sub>2</sub>-eq/year) for Dominica as a function of peak power demand over the period 2019-2030.**

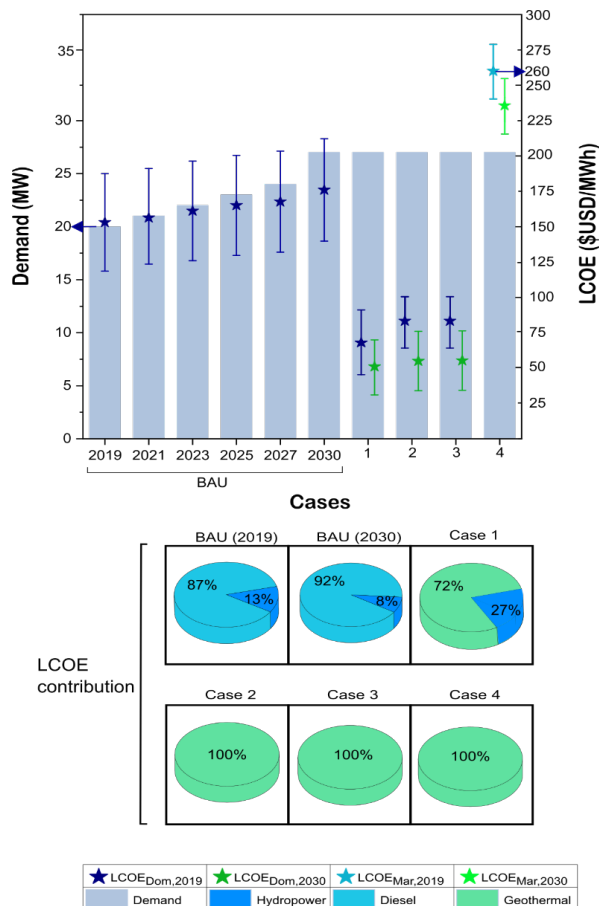
In Case 1, power from fossil fuels is completely substituted by geothermal sources, supporting a decline in total emission quotas to 734 tonnes CO<sub>2</sub> eq/year.

Cases 2 and 3 which explore opportunities for excess energy storage in the form of MeOH and NH<sub>3</sub> respectively achieve negative carbon emissions through avoided burdens- avoiding 36,871 and 149,013 tonnes CO<sub>2</sub>-eq/year respectively. Avoided burdens were mainly attributed to lower fossil fuel utilization- through the absence of conventional energy intensive steam methane reforming for hydrogen production as well as CO<sub>2</sub> utilization through direct air capture (DAC) operations. Case 4, which investigates the option of a RE export supply chain from Dominica to Martinique by harnessing geothermal energy to fulfil both countries' peak demand in 2030, shows an avoided life cycle GHG emission quota of 2.5 million tonnes CO<sub>2</sub>-eq- the highest amongst all scenarios investigated. Thus, Case 4 can significantly contribute to GHG emission reduction efforts for both countries under the Paris Agreement.

#### 3.2 Economic Performance

The economic feasibility of all cases was assessed by utilizing current and projected LCOE data for each technology as well as energy storage potential in the form of MeOH and NH<sub>3</sub> grassroots operations. For the BAU case, the LCOE for Dominica increased from 153 USD in 2019 to 175 USD/MWh in 2030 due to a 35% increase in power demand –resulting in an increase

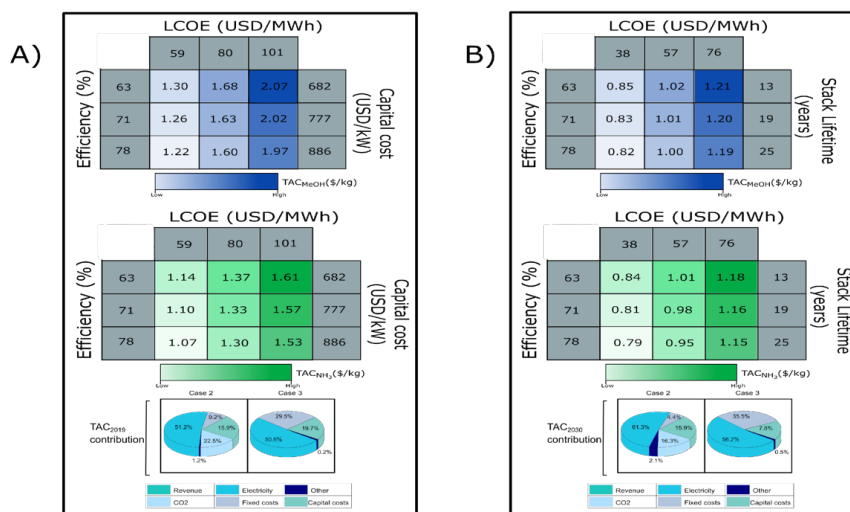
in diesel share contribution from 87% to 92%, shown in **Figure 2**. Although hydropower maintains 44% of the grid in 2019, it accounts for only 13% of the LCOE due to its lower energy costs over diesel.



**Figure 2: Case-specific LCOE as a function of net energy demand for 2019-2030, given with respect to technology grid share.**

The technology is, however, constrained by high reliance on weather events- impeding on its potential for dispatchable power generation and leading to a possible decrease in energy share to 33% in 2030.

Across Cases 1-3, our results show a drastic decrease in LCOE for Dominica. For Case 1, a 62% decrease in LCOE to 69 USD/MWh is attributed to the complete substitution of diesel power generation by geothermal energy- representing a 100% renewable grid; with 67% of the demand fulfilled by geothermal power. Cases 2 and 3 were analysed using TAC as a key performance indicator. **Figure 3** describes the respective current and future economic feasibility for each power-to-fuel scenario. Results illustrate TAC ranges for MeOH and NH<sub>3</sub> from 1.22 to 2.07 USD/kg<sub>MeOH</sub> and 1.07 to 1.61 USD/kg<sub>NH<sub>3</sub></sub> respectively. Thus, our data shows MeOH production to be the more expensive of the two scenarios- mainly attributed to energy and capital-intensive operations surrounding electrolysis and DAC. Overall, cost contributions to TAC for each process show major expenses attributed to geothermal electricity utilization (48%-51%) - with fixed and capital costs giving 9%-31% and 21%-22% respectively. Furthermore, DAC operations accrued up to 16% of the total TAC for MeOH production. Comparing current prices for both MeOH and NH<sub>3</sub> markets, green production routes utilizing geothermal energy can be regarded as uncompetitive - with 75-80% higher costs compared to fossil-based technologies<sup>13,14</sup>. However, with future reduction in electrolyser CAPEX and geothermal energy costs by 2030, TAC scores reveal strong MeOH and NH<sub>3</sub> competitiveness at near market value.



**Figure 3: Current-2020 (A) and future-2030 (B) TAC for MeOH (blue) and NH<sub>3</sub> (green) production as a function of LCOE (geothermal power), electrolyser efficiency, capital cost and stack lifetime. Costs associated with the category ‘other’ are natural gas fuel, process water, wastewater and catalyst charges.**

Lastly, the economic feasibility of Case 4- wherein Dominica exports power to the neighbouring island of Martinique- was assessed. Apart from achieving 100% RE penetration for both Dominica and Martinique by the exportation of 434MW of geothermal power via subsea transmission, the projected  $LCOE_{Mar}$  (Figure 2) was found to be 259 USD/MWh- 11% lower than the current consumer price in Martinique. Furthermore, with lower geothermal costs expected in the future, a further 20% decrease in  $LCOE_{Mar}$  to 236 USD/MWh in 2030 is anticipated. Among the case-specific cost contributions, subsea transmission accounts for 75% of the  $LCOE_{Mar}$ . Considering the vast energy losses during electrolysis when compared to transmission losses, energy exportation was found to be the most economically viable option. Notwithstanding, that the projected geothermal energy potential of Dominica surpasses the total deployment covered in this study (Proven potential = 1390 MW<sup>5</sup>), the nation can realize increased economic advantages from energy exportation and storage activities in the future. While our results do support a sustainable geothermal energy transition for Dominica, it accompanies several hurdles for the island linked to high capital investment such as upgrading of current power grid infrastructure, maintenance logistics linked to sub-sea transmission to Martinique and finally, exploration and geothermal production costs. Thus, to achieve its sustainable energy transition, Dominica will need to seek public and private sector investment while advocating for changes in current energy policies which align to strategic benefits for the country.

#### 4. Conclusion

A study on the island-nation of Dominica was conducted utilising techno-economic and environmental assessments to determine ways in which the nation can reduce anthropogenic emissions via the deployment of geothermal energy to fulfil its growing energy demand. Four cases were evaluated using LCA framework, LCOE and TAC with sensitivity to assess the most sustainable options for Dominica. Our results indicate that by 2030, without further renewable penetration, the island will see a 62% increase in GHG emissions together with an increase in LCOE<sub>2030</sub> up to 175 USD/MWh. However, with a 100% renewable grid supported by power generation from geothermal and hydropower sources, GHG emissions can be reduced by 99.5%, allowing the nation to meet its GHG commitments and achieve clean and affordable energy with a LCOE<sub>2030</sub> of 53 USD/MWh. While the production of MeOH and NH<sub>3</sub> allows for significant environmental benefits, TAC shows infeasibility compared to current market prices. Ultimately, the most sustainable case was the exportation of geothermal energy to fulfil 100% grid capability of neighbouring Martinique, resulting in a 20% decrease in LCOE and 2.5 million tonne CO<sub>2</sub>-eq/year avoided emissions.

#### References

1. Organisation of Eastern Caribbean States (OECS), OECS Strategic Objectives, <https://www.oecs.org/en/who-we-are/strategic-objectives>. (Accessed June, 2021).
2. OECS, Energy Issues and Options (2006), <https://openknowledge.worldbank.org/bitstream/handle/10986/17975/esm3170PAPER0OECS0energy01PUBLIC1.pdf?sequence=1&isAllowed=y>.
3. Joseph EP. Geothermal Energy Potential in the Caribbean Region. Retrieved July 4, 2021, from [https://sustainabledevelopment.un.org/content/documents/3339energy\\_joseph.pdf](https://sustainabledevelopment.un.org/content/documents/3339energy_joseph.pdf).
4. Glenn E, Comarazamy D, González JE, Smith T. Detection of recent regional sea surface temperature warming in the Caribbean and surrounding region. *Geophys Res Lett* 2015;42:6785–92. <https://doi.org/10.1002/2015gl065002>.
5. Manijean L, Saffache P. Geothermal energy slowly makes its entrance in the caribbean region. *Dynenviron* 2016:108–19. <https://doi.org/10.4000/dynenviron.695>.
6. Koon Koon R, Marshall S, Morna D, McCallum R, Ashtine M. A Review of Caribbean Geothermal Energy Resource Potential. *WIJE*. 2020;42(2):37-43.
7. Savaresi A. The Paris Agreement: a new beginning? *J. Energy Nat. Resour. Law*, 34;1:16-26. DOI: 10.1080/02646811.2016.1133983.
8. United Nations, Goal 7: Ensure access to affordable, reliable, sustainable and modern energy for all, <https://sdgs.un.org/goals/goal7>.
9. U.S Department of Energy- Energy Transition Initiative, Energy Snapshot: Dominica, <https://www.nrel.gov/docs/fy15osti/62704.pdf>. [accessed 7 July 2021].
10. Narine K, Mahabir J, Koylass N, Samaroo N, Singh-Gryzbon S, Baboolal A, et al. Climate smart process design for current and future methanol production. *J. CO2 Util* 2021;44:101399. <https://doi.org/10.1016/j.jcou.2020.1013>
11. Mahabir J, Bhagaloo K, Koylass N, Boodoo MN, Ali R, Guo M, et al. What is required for resource-circular CO<sub>2</sub> utilization within Mega-Methanol (MM) production? *J. CO2 Util* 2021;45:101451. <https://doi.org/10.1016/j.jcou.2021.101451>
12. International Organization for Standardization, Environmental management -Life cycle assessment - Requirements and guidelines; ISO 14044:2006 c2006 [updated 2016; cited 2020 December].
13. Zhang H, Wang L, Van herle J, Maréchal F, Desideri U. Techno-economic comparison of green ammonia production processes. *Appl Energy* 2020;259:114135. <https://doi.org/10.1016/j.apenergy.2019.114135>.
14. IRENA and Methanol Institute. Innovation Outlook: Renewable Methanol, International Renewable Energy Agency, 2021; ISBN 978-92-9260-32.

# Optimal design of a water energy food nexus system through a multi-objective approach

Daniel Peña-Torres<sup>a</sup>, Marianne Boix<sup>a</sup>, Ludovic Montastruc<sup>a</sup>

<sup>a</sup> LGC, UMR 5503, Toulouse INP – CNRS – Université Toulouse 3, 4 allée Emile Monso  
31432 Toulouse, France

## Abstract

Water, energy, and food resources have seen a growth in their global demand in the last decade and it is expected to continue for the year 2030 (National Intelligence Council, 2012). During the last 20 years, the scientific community have deeply studied the relations between these resources. Although it is at the center of governmental concerns, there is still a lack of studies regarding the WEFN systems with a nexus approach (FAO, 2021). Multi-objective optimization methods have the potential to conceive in a sustainable manner these complex systems. This study aims to design water, energy, and food optimal connections where different agents interact with each other. Two objective functions are taken into account for the model: total CO<sub>2</sub> emissions and net economic benefit. When minimizing total CO<sub>2</sub> emissions, the results indicate a total of 1 046 711 t CO<sub>2</sub>e and costs of 26 510 million USD. When minimizing net economic benefit, the emissions increase 2 000 times its ideal value, and the system perceives an economic benefit of 29 941 million USD. For future work it is expected to do a sensitivity analysis varying certain system parameters, as well as separating the economic objective function (one for each agent of the system), in order to study the competition that may exist among them.

**Keywords:** multi-objective optimization, decision support tool, water-energy-food nexus.

## 1. Introduction

Over the last decade, water, energy, and food resources grown in their global demand and it is expected that this tendency will continue for the year 2030 (FAO et al., 2021). Indeed, the reliance on water, energy, and food is projected to increase by 40, 50, and 35 percent, respectively, compared to their 2012 values (National Intelligence Council, 2012). The increment in the resources demand is attributable to the climate change that the earth is currently going through and rapid global population growth, which is expected to increase by 50% for the year 2050 (Ferroukhi et al., 2015). Given the above, there is a great pressure on these 3 resources. Therefore, in order to ensure safe access to water, energy, and food source, it is necessary to develop further studies for insuring a responsible resource management (Peña-Torres et al., 2022).

On the other hand, water, energy, and food are highly interrelated, presenting both synergisms and trade-offs along their supply chains. In the food production chain, the use of water and energy is necessary, from planting and feeding in agriculture and livestock, respectively, to cooking food in homes. Energy generation requires over half of the freshwater resources used (Bauer et al., 2014), which is transformed into steam to turn turbines, for example. It is also possible to use by-products of the food industry as biomass to generate energy. Both water treatment and water transport require the use of energy.



The interconnection of these three resources is referred in literature as the Water Energy Food Nexus (WEFN). This concept considers that, when a decision is taken for the management of one of the three, it will necessarily impact the decisions that will be taken for the two other resources. A recent report of the Food and Agriculture Organization of the United Nations (FAO et al., 2021) pointed out the negative effects of considering these three resources not-interconnected:

- Insecurity of access of water, energy, and food.
- Economic crises, geopolitical conflicts, public health concerns.
- Environmental impacts greater than expected.

The design of resources allocation systems through the nexus approach also addresses the use of land and soils by taking into account socio-economic factors. According to Carvalho et al. (2022), there is a real need for systemic thinking to achieve multipurpose design dedicated to WEFN. A systemic approach could also help to consider important indicators such as sustainability and resilience in terms of connections across different domains rather than just as individual components (Huntington et al., 2021). WEFN systems should then be conceived with a holistic approach so that they are resilient, durable, and organized to respond to current societal problems. Multi-objective optimization methods and tools have the potential to conceive in a sustainable manner these complex systems.

This study aims to design water, energy, and food optimal connections at a given territorial scale where different agents interact with each other.

## **2. Problem definition**

The overall goal of this study is to design water, energy, and food optimal connections at a given territorial scale, with two conflicting objectives: maximizing net economic benefit and minimizing total Greenhouse Gas (GHG) emissions. To optimize them, it is necessary to go through a step of modeling, formalization and formulation of optimization criteria. Based on a case study, we propose to optimize a territory in order to make it more respectful towards resource consumption, especially water, which is itself highly interconnected with energy generation and food production.

The optimization model describes every possible exchange with mass balance and conservation equations, and demand satisfaction restrictions. Linear models represent every material to energy conversion. The epsilon constraint method was used to obtain different points between the maximization of the net economic benefit (scenario A) and the minimization of the GHG emissions (scenario E). Three intermediate net economic benefit points were established between these two scenarios, minimizing the GHG emissions in each one of them to generate scenarios B, C, and D. The resulting mixed integer linear programming (MILP) model was solved using CPLEX v.20.1.0.1 solver (IBM, 2021) for GAMS 38.2.1 (General algebraic modeling system) (Bussieck and Meeraus, 2004).

## **3. Case Study**

This study describes the problem of water, energy, and food distribution at a macro level in the hydrological region of Sinaloa (58 000 km<sup>2</sup> approx.), located in the northwestern part of Mexico, and it is based on the previous work of Núñez-López et al. (2021). Within this region, four power plants satisfy the energy demand by burning natural gas or agricultural wastes. Aquifers, dams, and wells satisfy the water demand of the different

sectors (agriculture, livestock, domestic needs and industries). The total domestic food demand is fixed and obtained from agriculture, livestock or industrially processed food, or a combination of these three. In the system there is also the option of installing new actors:

- Three new desalination plants can be installed that can provide water and energy to different agents of the system.
- New rainwater collecting units for agriculture, livestock farmers, and the domestic sector that provide water to each one of them, respectively.
- New water treatment plants that offer fresh water to agriculture, industries and the domestic sector.
- A solar energy system for the domestic sector that offers energy outside of the power plants.
- A new gasification system that handles domestic waste and produces gas that feeds power plants and new desalination plants.

In a WEFN system, different types of resources are exchanged between different agents. In the water layer of the nexus, the power plants receive water from the sea and send it to livestock farmers, the agriculture sector, industries, the domestic sector, dams and aquifers. The same connections are proposed for the desalination plants. Dams receive from these two plants, in addition to water received from precipitations, and they send water to aquifers, livestock farmers, agriculture, industries, and the domestic sector. Aquifers receive water from dams, natural precipitations, and from water treatment units installed for the domestic, agricultural and industrial sector. Wells receive water only from aquifers, and they send water to livestock farmers, agriculture, industries and the domestic sector.

In the energy layer, power plants and desalination plants send energy to livestock farmers, the agriculture sector, industries and the domestic sector, while receiving fuel from natural gas suppliers, gasification system installed for domestic waste treatment, and agricultural wastes.

In the food layer of the system, livestock farmers, agriculture, and industries send food to the domestic sector. They have the option as well to sell their products to an external market. Agriculture sends its waste to the livestock sector in order to feed the cattle. Industries receive vegetables from agriculture to produce food. Finally, in order to grow the food in the agricultural sector industries sell fertilizer to them.

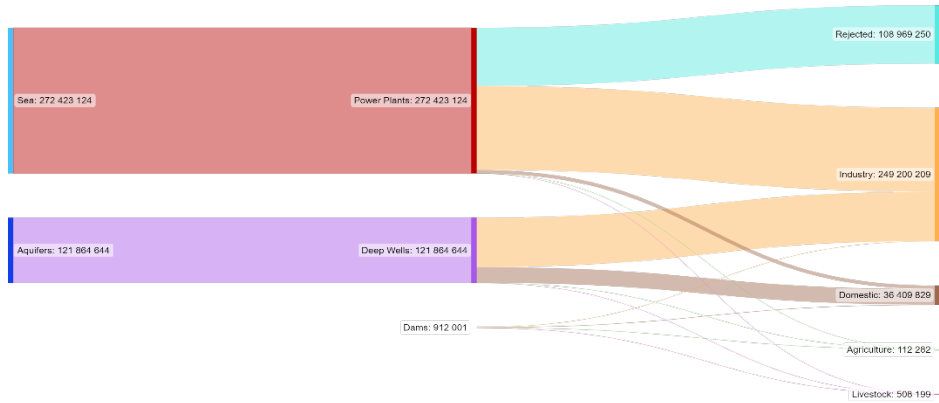
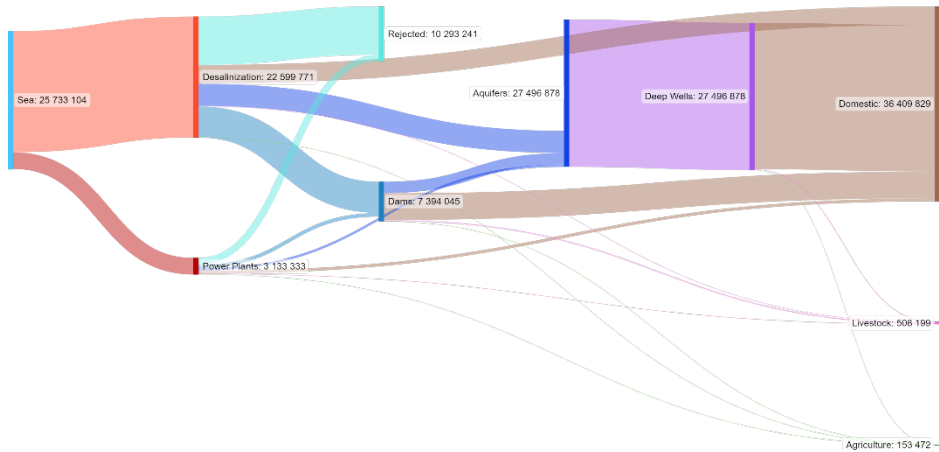
#### **4. Results**

The results are summarized in Table 1. In order to detail its analysis five different scenarios were selected. The extreme points A and E correspond to the results of the model when maximizing the net economic benefit and when minimizing GHG emissions respectively. The scenarios B, C, and D are obtained varying the trade-of index, which represents the importance of the economic goal in the multi-objective optimization problem, over the environmental goal (the importance of the environmental objective is obtained by subtracting this value from 1).

Comparing the scenarios, A and E, it can be seen that the economic function goes from 26 000 million USD of costs to 30 000 million of economic benefits. On the other hand, the environmental function increases 2 000 times its value from scenario E to scenario A.

Table 1: Objective function values when addressing economic and environmental goals.

Scenarios	Trade-off index	GHGC / t CO <sub>2</sub> e	Net economic benefit / 10 <sup>6</sup> USD
Maximizing Net Economic Benefit (A)	1	2 804 999 808	29 941
Scenario B	0.8	1 802 450 000	19 234
Scenario C	0.6	1 040 490 000	8 804
Scenario D	0.4	311 635 000	-1 627
Minimizing GHGC (E)	0	1 046 711	-26 510

Figure 1: Water flows in scenario A (in m<sup>3</sup>).Figure 2: Water flows in scenario E (in m<sup>3</sup>).

The main changes between the system configuration in scenarios A and E are the installation of 3 new desalination plants and 3 new gasification systems, both for scenario E (Figures 2 and 4). The installation of the new desalination plants entails the reduction of GHG emissions of 1 245 448 330 t CO<sub>2</sub>e, whereas the installation of gasification system accounts for a reduction of 311 564 647 t CO<sub>2</sub>e. On the other hand, the cost for the installation and operation of the new desalination plants is 21 564 million USD, and the profit of it is 9 million USD, meaning that the installation of new desalination plants implies a negative net economic benefit 21 555 million USD. This

new additions to the system also affect the number of connections within it. As it can be seen from Figures 1 and 2, when having desalination plants as a water source, the number of connections is higher, and therefore the cost for the system increases.

As the net economic benefit increases, the fertilizer production inside the system also does. When minimizing the GHG emissions, the fertilizer production within the system is 0, and the emissions related to its production are from the importation of fertilizer. When maximizing the net economic benefit, the fertilizer production increases up to 89 000 000 t, entailing an increase of GHG emissions of 1 247 848 375 t CO<sub>2</sub>e (Figures 3 and 4), but also an increase of the profits of 62 300 million USD.

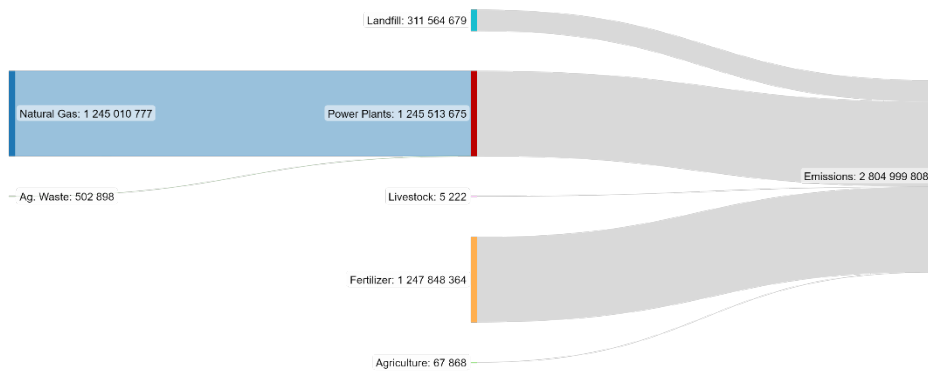


Figure 3: GHG emissions in scenario A (in t CO<sub>2</sub>e).

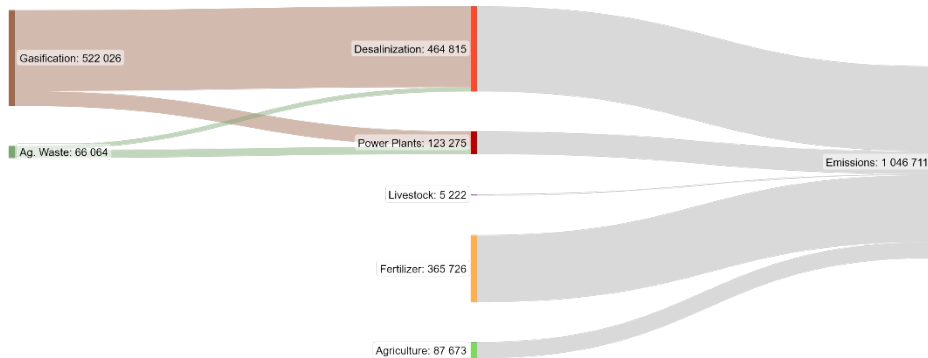


Figure 4: GHG emissions in scenario E (in t CO<sub>2</sub>e).

It is also possible to notice that between the points D and C the net economic benefit of the system changes sign, passing from costs to benefits. This means that when increasing the environmental function between 300 and 1 000 times its ideal value, the system already starts perceiving economic benefits. Between points D and C, a balanced solution could be obtained. For that it is necessary to include other indicators in the analysis, such as net present value and water footprint.

## 5. Conclusions

In this work, a MILP model is developed to design water, energy, and food optimal connections at a given territorial scale. The model considers energy flows coming from

power plants; water flows from deep wells, dams, and aquifers; and food flows from agriculture, livestock and industrial processed food. It also considers the installation of new desalinization plants for energy and water supply.

The model is applied to the design and plan of a WEFN system in Sinaloa, Mexico. Different scenario points were obtained using the epsilon constraint method between the maximization of the net economic benefit and the minimization of the GHG emissions. When minimizing the environmental objective function, the system configuration generates 1 046 711 t CO<sub>2</sub>e and has a cost of 26 510 million USD. When maximizing the economic objective function, the emissions result in 2 000 times its ideal value, and the economic benefit for the system is 29 941 million USD. The main difference between the system configuration of these two scenarios (A and E), is the installation of three new desalinization plants and three new gasification systems, both in scenario E.

For future work, it is intended to explore other scenarios in the model and do a sensitivity analysis, varying certain parameters of the system, such as domestic food demand and total rainfall in the region. It is also expected to consider an economic objective function for each one of the agents present in the system, thus analyzing the behavior and competition that may exist among these agents.

## References

- D. Bauer, M. Phillbrick, B. Vallario, H. Battey, Z. Clement, F. Fields, 2014. The water-energy nexus: Challenges and opportunities. US Department of Energy
- M. R. Bussieck, A. Meeraus, 2004. General algebraic modeling system (gams). In: Modeling languages in mathematical optimization. Springer, pp. 137-157.
- P. N. Carvalho, D. C. Finger, F. Masi, G. Cippolletta, H. V. Oral, A. Tóth, M. Regelsberger, A. Exposito, 2022. Nature-based solutions addressing the water-energy-food nexus: Review of theoretical concepts and urban case studies. *Journal of Cleaner Production* 338, 130652.
- FAO, IFAD, UNICEF, WFP, WHO, 2021. The State of Food Security and Nutrition in the World 2021. FAO, Rome, Italy.
- R. Ferroukhi, D. Nagpal, A. Lopez-Peña, T. Hodges, R. H. Mohtar, B. Daher, S. Mohtar, M. Keulertz, 2015. Renewable energy in the water, energy and food nexus. *International Renewable Energy Agency* (January), 1-125.
- H. P. Huntington, J. I. Schmidt, P. A. Loring, E. Whitney, S. Aggarwal, A. G. Byrd, S. Dev, A. D. Dotson, D. Huang, B. Johnson, J. Karenzi, H. J. F. Penn, A. Salmon, D. J. Sambor, W. E. Schnabel, R. W. Wies, M. Wilober, 2021. Applying the food-energy-water nexus concept at the local scale. *Nature Sustainability* 4(8), 672-679. URL <https://doi.org/10.1038/s41893-021-00719-1>
- IBM, 2021. IBM ILOG CPLEX 20.1 User's Manual.
- National Intelligence Council, 2012. Global Trend: Alternative Worlds. URL [www.dni.gov/nic/globaltrends](http://www.dni.gov/nic/globaltrends)
- J.M. Núñez-López, E. Rubio-Castro, J. M. Ponce-Ortega, 2021. Involving resilience in optimizing the water-energy-food nexus at a macroscopic level. *Process Safety and Environmental Protection* 147, 259-273.
- D. Peña-Torres, M. Boix, L. Montastruc, 2022. Optimization approaches to design water-energy-food nexus: A literature review. *Computers & Chemical Engineering*, 108025.

# Techno-economic comparison of peaking power plants: OCGT+CCS vs Lithium-ion batteries

Mathew Dennis Wilkes,<sup>a</sup> Diarmid Roberts,<sup>a</sup> Solomon Brown,<sup>a</sup>

<sup>a</sup> *Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom,*

## Abstract

To ensure electricity system security dispatchable power generation is required to provide balancing capacity in times of need, such as providing diurnal peak power supply. In order to achieve a cost-effective transition to Net Zero, these peaking power plants need to be low-cost and low-carbon. Herein, we compare two forms of dispatchable power: open-cycle gas turbine (OCGT) with carbon capture and storage (CCS), and Lithium-ion battery energy storage system (BESS). The results highlight the lower cost of BESS compared to OCGT+CCS, although this is only for systems that provide <8h duration. Hence, both forms of dispatchable power generation will be required in future resilient energy systems.

**Keywords:** CCS, Energy, Energy Storage, Energy Grid

## 1. Introduction

As we move towards Net-Zero emissions by 2050, all forms of power generation need to have low, zero, or negative carbon emissions. With an anticipated growth in variable renewable energy (VRE) (Heptonstall & Gross, 2021), Energy System Operators (ESOs) across the world are strengthening their balancing capacity. These services ensure a constant balance between supply and demand of energy, providing crucial system security and reliability of supply (NGESO, 2018). Balancing services include dispatchable power generation (peaking plants), demand side response, interconnectors, and frequency response (NGESO, 2022). Within the UK, the National Grid (NG) expects an increase in dispatchable electricity supply, which is currently dominated by gas turbines but by 2050 they will be replaced by energy storage, interconnectors, and hydrogen (NGESO, 2021). Similarly, the Australian Energy Market Operator (AEMO) has anticipated a 57 GW growth in dispatchable resources to enable an additional 125 GW of large-scale VRE's. Traditionally, Australia has relied on gas-fired generators for peaking plants but there is a growing interest in energy storage solutions (AEMO, 2022).

Australia's Clean Energy Council compared battery storage to gas turbines, and highlighted the batteries are superior peaking plant in terms of cost, flexibility, emissions, and the range of network services it can provide (CEC, 2021). The study did not include CO<sub>2</sub> abatement/tax, which will have a major impact on the cost of electricity from gas turbines. Also, the economics of energy storage is dependent on the mode of operation, capacity factor, electricity price, and location-based factors.

### 1.1. Aims and Objectives

Currently, dispatchable gas turbines are the main form of peaking technology; however, in order to meet Net-Zero targets their emissions need to be reduced. Hence, this study makes a techno-economic comparison of open-cycle gas turbine (OCGT) power generation with Carbon Capture and Storage (CCS), and a Lithium-ion battery energy storage system (BESS), specifically for operation in the UK. Herein, we utilize process models developed in our earlier studies to provide the necessary design and operating

parameters which are then used within the economic models. The main study objectives are:

- Utilize process modelling to identify key design/operating parameters.
- Define economic assumptions for both dispatchable UK power plants.
- Techno-economic comparison of **OCGT+CCS** against **BESS**.

## 2. Methodology

This section highlights the process design, costing methodology, and economic assumptions, for both dispatchable technologies (OCGT+CCS and BESS).

### 2.1. Process design

For the OCGT+CCS power plant, the gas turbine is a Siemens SGT-400 producing 10.4 MWe and 33.8 kg/s flue gas containing 6.78 wt.% CO<sub>2</sub>. In order to be low-carbon, the flue gas needs post-combustion CO<sub>2</sub> capture using the benchmark Monoethanolamine (MEA) process. More information on the power generation, capture plant design, and process model can be found in our previous study (Wilkes, et al., 2021a). The CO<sub>2</sub> stream is then conditioned ready for pipeline transportation, at 111 bar and 50 ppm moisture. More information on the conditioning train design and process model can be found in our previous study (Wilkes, et al., 2021b).

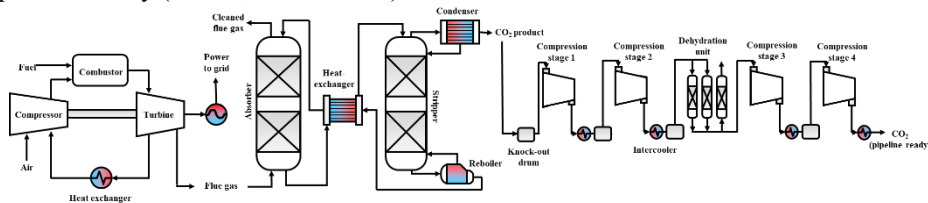


Figure 1: OCGT+CCS power plant flowsheet

The BESS is modelled as a turnkey stationary grid-connected system of 10 MWe rating, using the definitions in Mongird et al. (2020). The round-trip efficiency is set at 86%. The BESS generates revenue by charging the battery during periods of low price and supplying electricity to the grid when the price is high.

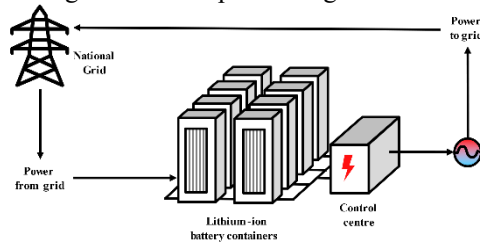


Figure 2: BESS plant flowsheet

### 2.2. Economic costing metric

In this study we use the levelized cost of electricity (LCOE) for the OCGT+CCS plant and the levelized cost of storage (LCOS) for the BESS. The LCOE is the breakeven point at which the electricity must be sold to counteract the cost of constructing and operating the plant. The LCOE is the sum of the net present value (NPV) of costs divided by the sum of the NPV of electricity generated then sold, shown in Equation 1 (BEIS, 2020).

$$LCOE = \frac{NPV_{total\ costs}}{NPV_{generation}} = \frac{\sum_n \left( \frac{TCC_n}{(1+r)^n} + \frac{FOM_n}{(1+r)^n} + \frac{VOM_n}{(1+r)^n} \right)}{\sum_n \frac{E_n^{Out}}{(1+r)^n}} \quad (1)$$

Where TCC is the total capital costs, FOM is the fixed operating and maintenance costs, VOM is the variable operating and maintenance costs,  $E_t^{Out}$  the total electricity generated,  $r$  is the discount rate, and  $n$  is the time period.

For the OCGT+CCS cost model, the TCC is calculated using correlations from Chauvel et al. (1981) and is based on the purchased equipment cost (PEC):

$$PEC = \sum_1^{total} E_i \quad (2)$$

This is for all of the major pieces of equipment ( $i$ ) identified in the process model, and are shown in Figure 1. The individual equipment cost ( $E_i$ ) is based on the correlation from Towler & Sinnott (2012):

$$E_i = a + bS^n \quad (3)$$

Where  $S$  is the sizing factor specific to each piece of processing equipment, and  $a$ ,  $b$ , and  $n$  are cost factors. The costs are scaled to the present day using Turton et al. (2018):

$$E_i^{2019} = E_i^{2010} \left( \frac{F^{2019}}{F^{2010}} \right) \quad (4)$$

The Chemical Engineering Plant Cost Index (CEPCI) for 2010 is 551 ( $F^{2010}$ ) and for 2019 is 607.5 ( $F^{2019}$ ) (Jenkins, 2020). The VOM and FOM are also calculated using correlations from Chauvel et al. (1981). Table 1 shows the economic assumptions for the OCGT+CCS plant. Table 3 shows the utilities prices used to calculate the VOM.

The LCOS is very similar to LCOE, in that the NPV of all costs is divided by the NPV of the electricity output. They can hence be used for a like-for-like comparison. LCOS is defined by:

$$LCOS = \frac{NPV_{total\ costs}}{NPV_{generation}} = \frac{TCC + \sum_n \left( \frac{VOM_n}{(1+r)^n} \right)}{\sum_n \frac{E_n^{Out}}{(1+r)^n}} \quad (5)$$

Which is adapted from Julch (2016), where it is assumed that  $TCC$  are accrued in year 0.  $VOM_n$ , the variable operating costs are further defined by:

$$VOM_n = OPEX_n + CAPEX_{re,n} + \bar{c}_{elec,n} \frac{E_n^{Out}}{\eta} - R_n \quad (6)$$

Where  $OPEX_n$  is the fixed operational cost (excluding degradation and charging)  $CAPEX_{re,n}$  covers any re-investments in equipment to maintain capacity following degradation,  $\bar{c}_{elec,n}$  is the average price of electricity used to charge the BESS,  $\eta$  the round-trip efficiency of the BESS, and  $R_n$  the residual value of the equipment. Table 2 shows the economic assumptions for the BESS facility.



This analysis neglects additional revenue that could be generated using the BESS for arbitrage, or ancillary services such as frequency response, and which would reduce the LCOS for the peaking application. This is done under the assumption that the pivotal role of the BESS is to provide incremental peaking capacity, and that performing other roles may push other capacity out of the market.

Table 1: OCGT power plant economic assumptions

Assumption	Description
Plant location	Yorkshire, England
Power rating	10 MWe
Plant lifetime	25 years
Construction time	2 years
Start of construction	2019
Discount rate	Proven technology = 7.8% (BEIS, 2018) Higher risk technology = 8.9% (BEIS, 2018)
Capacity factor	1500 annual operating hours (Millbrook Power Ltd., 2017)
Carbon price (CP)	£21.70/tCO <sub>2</sub> (BEIS, 2018)

Table 2: Li-ion BESS economic assumptions

Parameter	Description
Plant location	Yorkshire, England
Power rating	10 MWe
Project life	10 years
Start of operation	2021
$r$	7.8 % proven technology rate from Table 1.
$TCC$	£784-3825/kW, 2020 data in (Mongird, et al., 2020)
$OPEX_n$	£4.20/kW/year (Mongird, et al., 2020)
$\bar{c}_{elec,n}$	£100/MWh. Mean N2EX 2021 day-ahead price 20:00-07:00.
$CAPEX_{re,n}$	The system is assumed to last for the entire project life; therefore, this value is 0 (Roberts & Brown, 2022).
$R_n$	There is no value in the equipment once it has been used.
$\eta$	0.86 (Mongird, et al., 2020)

Table 3: Utility prices

Assumption	Description
Natural gas	33 £/MWh (Global Petrol Prices, 2021)
Electricity	140 £/MWh (BEIS, 2021)
Cooling water	1.39 £/m <sup>3</sup> (Yorkshire Water, 2021)
MEA solvent	5.0 £/L (Mistral Industrial Chemicals, 2013)

### 3. Results

The LCOE of the OCGT+CCS plant and the LCOS of the BESS are shown in Figure 5. For each technology, the levelized cost (£/MWh) drops as the operational hours increase, as the CAPEX is spread over a greater electricity output. It does not drop to zero, as there are variable operating costs in both cases. The operating hours range between 250-4,000 hours annually for both technologies.

The OCGT+CCS plant is considerably more expensive than unabated OCGT. It is also more expensive than BESS except when the required duration of daily response is greater than 8 h and less than 1000h per year. In this case, the TCC associated with the BESS is prohibitive. The results show that a BESS is more suitable than OCGT+CCS as a low

carbon solution for a diurnal peaking application where power must be provided for 2-4h every day. This situation will arise in grids predominantly powered by PV, such as those in Australia or California. In the UK in the future however, it is likely that wind power supply variation will mean that multi-day back-up power is more important than peaking power (Drax, 2022). It is clear from Figure 5 that BESS with sufficient duration will be more expensive than OCGT+CCS.

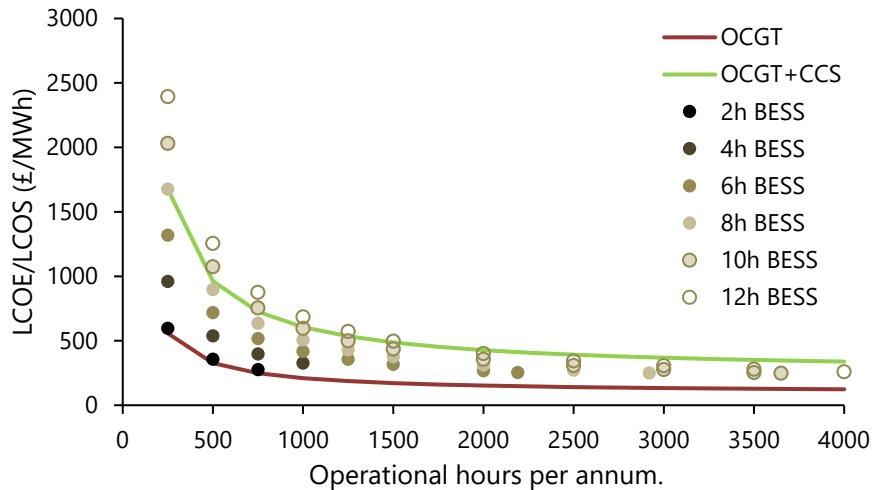


Figure 3: Modelled LCOE and LCOS for OCGT, OCGT+CCS, and BESS versus operational hours per year. For the BESS the operational hours are capped by assuming that the BESS performs 1 cycle per day.

Based on predicted pathways for future energy scenarios from NGENSO (NGESO, 2022), natural gas peakers are expected to decrease in line with Net Zero scenarios; however, the results in this study show they will still have a role to play. Depending on the NGENSO pathway, the Gas-CCS share of dispatchable electricity supply ranges between 3.2 to 20.8% or 3.3 GW to 23.1 GW. All NGENSO pathways show a large portion of flexibility will be provided by energy storage. The results shown in Figure 3 highlight the slightly lower cost of BESS (276-2,393 £/MWh) compared to Gas-CCS (290-1,675 £/MWh), but only for systems designed for <8h duration. Therefore, future work should focus on reducing the cost of BESS for long duration systems, and reducing the cost of dispatchable Gas-CCS by identifying low energy capture solutions, thermal integration potential, and CO<sub>2</sub> utilization options.

#### 4. Conclusion

Future energy systems that incorporate a high proportion of variable renewable energy technologies require balancing capacity to ensure security of electricity supply. This study is a techno-economic comparison of two forms of dispatchable power generation used during periods of peak power demand: OCGT+CCS and BESS. Including CCS for the OCGT plant significantly increase the LCOE, as the additional capital costs associated with capturing, transporting, and storing the CO<sub>2</sub> exacerbate the issues of small plant size and low capacity factor. Interestingly, the results show the lower cost of BESS, but only for low duration (<8h) operation. Therefore, future work is required to ensure the cost of longer duration BESS is lower than Gas-CCS, i.e., investigating alternate BESS technologies or medium-duration storage technologies (compressed air or pumped hydro). Thus, leading to a cost-effective Net Zero transition in the power sector.

## 5. References

- AEMO, 2022. *2022 Intergrated System Plan*, Melbourne, Australia: Australian Energy Market Operator.
- BEIS, 2018. *Assessing the Cost Reduction Potential and Competitiveness of Novel (Next Generation) UK Carbon Capture Technology*, s.l.: UK Department for Business, Energy & Industrial Strategy.
- BEIS, 2020. *Electricity Generation Costs 2020*, London: Department for Business, Energy & Industrial Strategy.
- BEIS, 2021. *Quarterly Energy Prices: UK January to March 2021*, s.l.: The Department for Business, Energy and Industrial Strategy.
- CEC, 2021. *Battery Storage: The new, clean peaker*, s.l.: Clean Energy Council.
- Chauvel, A. et al., 1981. *Manual of Economic Analysis of Chemical Processes: Feasibility Studies in Refinery and Petrochemical Processes.* s.l.:McGraw-Hill.
- Drax, 2022. *Electric Insights 24-11-22 to 01-12-22*. [Online]  
Available at: <https://www.electricinsights.co.uk/#/dashboard?period=7-days&start=2022-11-24&k=xw6no5>  
[Accessed 1 December 2022].
- Global Petrol Prices, 2021. *United Kingdom natural gas prices*. [Online]  
Available at: [https://www.globalpetrolprices.com/United-Kingdom/natural\\_gas\\_prices/](https://www.globalpetrolprices.com/United-Kingdom/natural_gas_prices/)  
[Accessed April 2021].
- Heptonstall, P. J. & Gross, R. J. K., 2021. A systematic review of the costs and impacts of integrating variable renewables into power grids. *Nature Energy*, Volume 6, pp. 72-83.
- Jenkins, S., 2020. *2019 CHEMICAL ENGINEERING PLANT COST INDEX ANNUAL AVERAGE*. [Online]  
Available at: <https://www.chemengonline.com/2019-chemical-engineering-plant-cost-index-annual-average/>  
[Accessed 25 October 2021].
- Julch, V., 2016. Comparison of electricity storage options using levelized cost of storage (LCOS) method. *Applied Energy*, Volume 183, pp. 1594-1606.
- Millbrook Power Ltd., 2017. *The Millbrook Power (Gas Fired Power Station) Order*, s.l.: Parsons Brinckerhoff.
- Mistral Industrial Chemicals, 2013. *MONOETHANOLAMINE MEA 2-AMINOETHANOL*. [Online]  
Available at: <https://mistralni.co.uk/products/mea-monoethanolamine>  
[Accessed October 2021].
- Mongird, K. et al., 2020. *2020 Grid Energy Storage Technology Cost and Performance Assessment*, s.l.: US DOE.
- NGESO, 2018. *ESO Balancing Services: A guide to contracting, tendering and providing response and reserve services*, Warwick, England: National Grid ESO.
- NGESO, 2021. *Operability Strategy Report*, London, England: National Grid ESO.
- NGESO, 2022. *List of all balancing services*. [Online]  
Available at: <https://www.nationalgrideso.com/industry-information/balancing-services/list-all-balancing-services>  
[Accessed 11 November 2022].
- Roberts, D. & Brown, S., 2022. The economics of firm solar power from Li-ion and vanadium flow batteries in California. *MRS Energy & Sustainability*.
- Towler, G. & Sinnott, R., 2012. *Chemical Engineering Design: Principles, Practice and Economics of Plant and Process Design*. 2nd ed. Oxford: Elsevier.
- Turton, R., Shaeiwitz, J. A., Bhattacharyya, D. & Whiting, W. B., 2018. *Analysis, Synthesis, and Design of Chemical Processes*. 5th ed. s.l.:Pearson Education Inc..
- Wilkes, M. D., Mukherjee, S. & Brown, S., 2021a. Transient CO<sub>2</sub> capture for open-cycle gas turbines in future energy systems. *Energy*, p. 119258.
- Wilkes, M. D., Mukherjee, S. & Brown, S., 2021b. Linking CO<sub>2</sub> capture and pipeline transportation: sensitivity analysis and dynamic study of the compression train. *International Journal of Greenhouse Gas Control*, Volume 111, p. 103449.
- Yorkshire Water, 2021. *Charges for customers with a meter*. [Online]  
Available at: <https://www.yorkshirewater.com/bill-account/how-we-work-out-your-bill/customers-with-a-meter/>  
[Accessed October 2021].

# Bioenergy with carbon capture and storage (BECCS) – power generation coupled with hydrogen production

Mathew Dennis Wilkes,<sup>a</sup> Oludayo Asuni,<sup>a</sup> Solomon Brown,<sup>a</sup>

<sup>a</sup> *Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom,*

## Abstract

Achieving Net Zero emissions by 2050 requires the utilization of negative emissions technologies such as Bioenergy with Carbon Capture and Storage (BECCS), as well as hydrogen to decarbonize challenging industrial CO<sub>2</sub> emitters. Within this paper we develop a BECCS-Power-Hydrogen process model in Aspen Plus. The analysis shows the key operating parameters and key performance indicators for the process. The base case electrical efficiency is very low (18.4%) but including the hydrogen production increases the overall system efficiency to 52%. Also included in this study is a sensitivity analysis investigating the effect gasification temperature has on the hydrogen production and net system efficiency.

**Keywords:** BECCS, Hydrogen, Energy, Gasification, Aspen.

## 1. Introduction

Bioenergy coupled with Carbon Capture and Storage (BECCS) is a negative emissions technology that can be applied to power (BECCS-Power) and hydrogen production (BECCS-Hydrogen). BECCS is featured heavily in energy scenarios designed to achieve Net-Zero, alongside the use of hydrogen to decarbonize power, industry, heating, and transportation (CCC, 2020; BEIS, 2021).

One option for generating hydrogen from biomass is via gasification, whereby the product syngas components are separated to produce high purity H<sub>2</sub>. Several review papers have highlighted the current state-of-the-art of biomass gasification technologies (Lepage, et al., 2021; Tezer, et al., 2022), including the integration of carbon capture, utilization and storage (CCUS) technologies (Shahbaz, et al., 2021) and the possibility of bio-renewable hydrogen production (Cao, et al., 2020). Mutlu and Zeng (Mutlu & Zeng, 2020) highlighted the challenges and opportunities of modelling biomass gasification. Process simulation tools such as ChemCAD, Matlab, gPROMS, Aspen, GAMBIT, and Fluent, can allow users to identify process limitations and optimize operation.

There are studies that have focused on modelling BECCS-Hydrogen in Aspen Plus (Marcantonio, et al., 2019; Safarian, et al., 2022) and the possibility of including heat integration to incorporate BECCS-Power (Cohce, et al., 2011; Ishaq & Dincer, 2020). Cohce et al. (2011) evaluated biomass gasification-based hydrogen production, with waste heat recovery and electricity generation to improve overall system efficiency (energy efficiency = 33% and exergy efficiency = 29%). Ishaq and Dincer (2020) showed an integrated energy system design for heat, electricity, and hydrogen production, based on biomass gasification and water electrolysis. The process model uses 0.4 kg/s of biomass feedstock and has the potential to produce 1.4 MW of energy and 10.74 mol/s H<sub>2</sub>. The overall energy efficiency is 53.7 2%; however, there is a very high capital

investment required for the cryogenic air separation unit, quantity of turbines, proton exchange membrane electrolyzer, and entrained flow gasifier.

Therefore, this study investigates the thermochemical conversion of biomass for the co-production of power and hydrogen. Utilizing process models developed in Aspen Plus to identify the processes key operating parameters (KOPs) and key performance indicators (KPIs). The relatively simplistic design, biomass gasification for hydrogen production coupled with heat recovery for power generation, provides a baseline for the conversion of biomass to power and hydrogen. A sensitivity analysis on the identified KOPs will enable effective process optimization and highlight potential process limitations. The key objectives are:

- Develop a process model, in Aspen Plus® v11.1, for BECCS-Power-Hydrogen.
- Highlight and analyze the processes KOPs and KPIs.
- Identify process limitations and options for optimization.

## 2. Process design and modelling

The process design is separated into five main sections: biomass gasification, heat recovery and steam cycle, syngas cleaning and upgrading, hydrogen separation, and CO<sub>2</sub> conditioning. The flowsheet is shown in the model topology (Figure 1) in Section 2.6.

### 2.1. Biomass Gasification

The biomass gasification and syngas production in a downdraft gasifier is based on the experiments conducted by Wei et al. (2009) using hardwood chips (Tauqir, et al., 2019), the key operating parameters are shown in Table 1. The proximate and ultimate analysis of the chips is shown in Table 2.

Table 1: Operating conditions for the gasification unit for run #14 from Wei et al. (2009)

Parameter	Value
Biomass feed flowrate (kg/hr)	20.58
Moisture content (wt.%)	10.05
Gasifying temperature (°C)	789
Air flowrate (m <sup>3</sup> /hr)	46.05
Equivalence ratio	0.289

Table 2: Hardwood chips proximate and ultimate analysis (Tauqir, et al., 2019)

Proximate Analysis	Value (wt.%)
Moisture	25
Fixed Carbon	79.85
Volatile Material	19.031
ASH	1.119
Ultimate Analysis	
Carbon	49.817
Hydrogen	5.556
Nitrogen	0.078
Chlorine	0.000
Sulfur	0.005
Oxygen	43.425
HHV (MJ/kg)	18.58

### 2.2. Power generation

Within the steam cycle, water is pumped to 125 bar and passes through a heat recovery steam generator (HRSG) unit, where the hot gases from the gasifier heat the steam to approximately 565°C (Chiesa, et al., 2005). The isentropic and mechanical efficiency of the high pressure (HP), intermediate pressure (IP), and low pressure (LP) steam turbines is 80% and 95%, respectively (Cohce, et al., 2011).

### 2.3. Syngas upgrading

The syngas upgrading using the water-gas shift (WGS) reaction unit is based on Marcantonio et al. (2019) and Moneti et al. (2016), both of which simulated the experiments conducted within the UNIfHY project (UNIfHY, 2016). The WGS reaction favors the production of CO and H<sub>2</sub>O at high temperatures, thus the UNIfHY project

looked at high temperature shift (HTS) at 400°C and low temperature shift (LTS) at 200°C. Both HTS and LTS reactors are operated at 27 bar (Cohce, et al., 2011).

#### 2.4. H<sub>2</sub> production

Sulfur within the biomass is converted into H<sub>2</sub>S and COS in the gasification unit, and all of the COS is converted in H<sub>2</sub>S in the WGS units. Therefore, prior to H<sub>2</sub>/CO<sub>2</sub> separation H<sub>2</sub>S is removed using Selexol, which also co-captures CO<sub>2</sub>. The separation efficiency for H<sub>2</sub>S is 100% and for CO<sub>2</sub> it is 95%. High purity H<sub>2</sub> (99.99%) is separated using a Pressure Swing Adsorption (PSA) unit (Chiesa, et al., 2005). Modelling the Selexol process and PSA units for H<sub>2</sub> production is possible in Aspen Plus (Zhu, et al., 2020), however, for this study we assume set separation efficiencies to reduce the complexity of the model.

#### 2.5. CO<sub>2</sub> conditioning

The CO<sub>2</sub> conditioning unit ensures the CO<sub>2</sub> stream is ready for pipeline transportation, it uses sub-critical liquefaction (at 66 bar) and liquid pumping (111 bar) to reduce the energy consumption (Wilkes, et al., 2021).

#### 2.6. Process model

The process model is developed in Aspen Plus® v11.1, the model topology is shown in Figure 1. Wet biomass enters and **DRYER** unit and the output moisture content is set to 10.05 wt.%. Within Aspen, biomass is a non-conventional solid and needs to be decomposed (**DECOMP**) into its constituent elements before gasification (**GASI**). An RGibbs reactor is chosen for the gasification unit, which calculates chemical equilibrium based on Gibbs free energy minimization (Tauqir, et al., 2019). The high temperature (**HT-WGS**) and low temperature (**LT-WGS**) WGS units are REquil reactors (Marcantonio, et al., 2019). The Selexol and H<sub>2</sub>-PSA processes are modelled as separation units (**SELEXOL1**, **SELEXOL2**, and **PSA**), with the separation efficiencies specified in Section 2.4. Heat is recovered in the **HRS** unit, and the high-pressure steam produces power in the **HP-ST**, **IP-ST**, and **LP-ST** turbines.

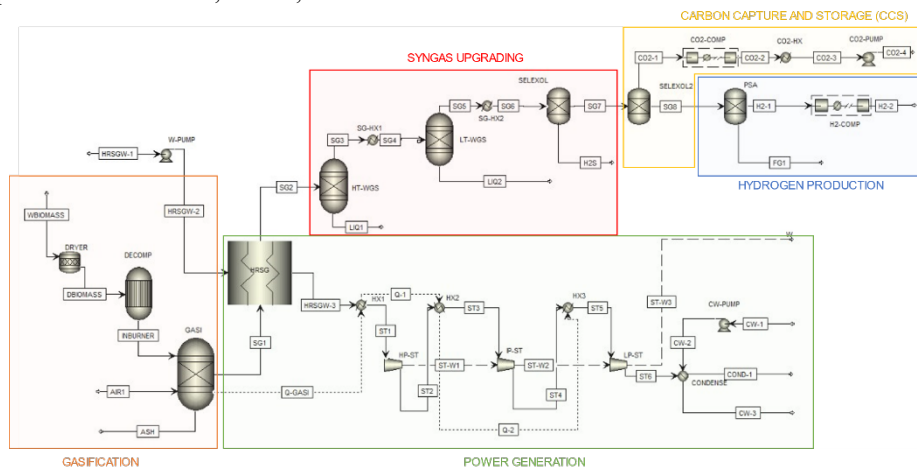


Figure 1: Biomass gasification for H<sub>2</sub> and power production, model topology using Aspen Plus

### 3. Results

#### 3.1. Validation

The validation is carried out on the gasification unit, based on the downdraft gasifier experiments (run #14) conducted by Wei et al. (Wei, et al., 2009), the key operating conditions are shown in Table 1. The comparison between the gas compositions in the

experiment and simulation is shown in Figure 2. The experimental study only investigated the composition of CH<sub>4</sub>, CO<sub>2</sub>, CO and H<sub>2</sub>, on a dry basis. The first simulation (Simulation #1) used the same ER and air flowrate as Wei et al. (Wei, et al., 2009), which resulted in underestimating the CO and H<sub>2</sub> composition and overestimating the CO<sub>2</sub> composition. As this study focusses on H<sub>2</sub> production, it is imperative the H<sub>2</sub> composition is accurate. Therefore, we modified the ER in Simulation #2 to 0.214 or 34.12 m<sup>3</sup>/hr of air. This is still within the range (0.2 to 0.28) specified in Wei et al. (Wei, et al., 2009). In Simulation #2, the composition of CO<sub>2</sub>, CO, and H<sub>2</sub> deviate 3.61%, 3.70%, and 0% compared to the experiment. The composition of CH<sub>4</sub> is much lower (0.0013 vol.%) than shown in Wei et al. (Wei, et al., 2009), this is due to the use of thermodynamic equilibrium calculations within Aspen's Rgibbs reactor (Tauqir, et al., 2019). As the model underpredicts the quantity of CH<sub>4</sub>, the simulations syngas LHV (4.58 MJ/Nm<sup>3</sup>) is 26% lower than expected (6.17 MJ/Nm<sup>3</sup>).

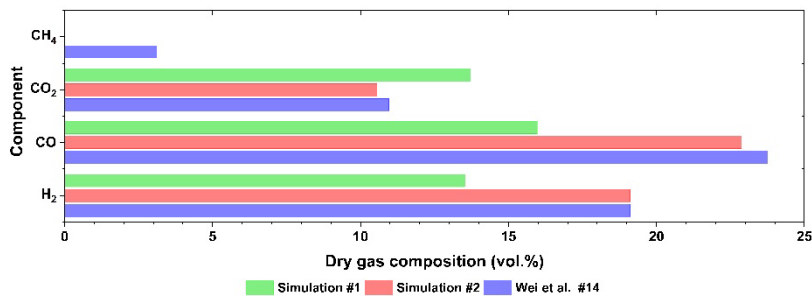


Figure 2: Downdraft gasification syngas composition comparison

It is also important to ensure the correct ratio of H<sub>2</sub>/CO is achieved after the WGS units. Typically, the syngas H<sub>2</sub>/CO ratio should range between 1.5:1 to 2:1, this is suitable for further processing in methanol or Fischer-Tropsch synthesis (Binder, et al., 2018). For this simulation the H<sub>2</sub>/CO ratio is 1.79, therefore, the model is in good agreement with experimental data and within the acceptable syngas composition range.

### 3.2. Base case

To ensure enough steam can be generated to power the turbines, the process is scaled to use 3600 kg/hr of biomass (hardwood chip analysis is shown in Table 2) and ER=0.214. The energy flows through the process are shown in Figure 3. The KOPs are shown in Table 3 and the KPI's are shown in Table 4.

The gasifier heat duty is 5,023 kW and the turbines combined power output is 3,846 kW. Chiesa et al. (2005) showed the electrical efficiency of the steam cycle (based on manufacturer quotes) is approximately 18.1-19.1%. In this study, the electrical efficiency of the gasification plant is 18.4%. Therefore, the process design is comparable to sources within the literature and industry. The power generation and electrical efficiency is low compared to conventional biomass combustion technologies, due to the much lower ER which reduces the amount of heat recoverable in the HRSG.

The thermal efficiency is based on the amount of hydrogen produced, which is 185 kg/hr or 6,176 kW. The net efficiency is the combination of the electrical and thermal conversion of the biomass feedstock, for this plant design it is 51.65%, similar to Ishaq and Dincer (2020).

Table 3: Key operating parameters

KOP	Value
Biomass input (kg/s)	3,600
Biomass energy (kW)	18,580
Air flowrate (kg/s)	6934
Equivalence Ratio	0.214
HRSG flowrate (kg/s)	10,497
Syngas/HRSG stream ratio	1

Table 4: Key performance indicators

KPI	Value
Power demand (kW)	425
Power generated (kW)	3,846
Net power output (kW)	3,421
Hydrogen output (kW)	6,176
Electrical efficiency (%)	18.41
Thermal efficiency (%)	33.24
Net efficiency (%)	51.65

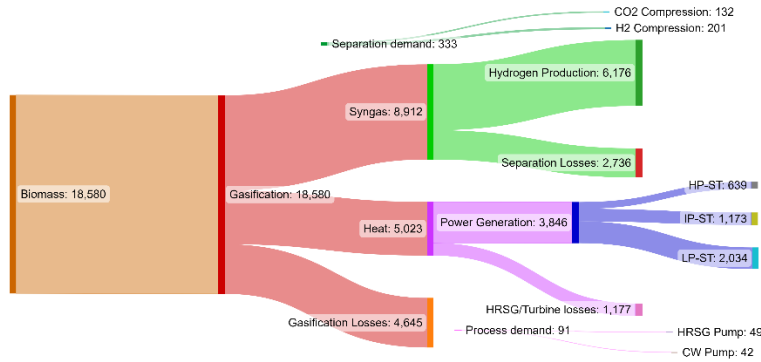


Figure 3: Sankey diagram showing energy flows

### 3.3. Sensitivity Analysis

The gasification process is one of the most important units, especially in terms of exergy destruction (Cohce, et al., 2011). Therefore, we investigated the effect of varying the gasification temperature from 700°C to 1000°C. Interestingly, the H<sub>2</sub> concentration in the syngas decreases due to an increase in combustion components, but the concentration post-WGS increases. Thus, more H<sub>2</sub> is produced which increase the net efficiency, as shown in Figure 4. More energy is available to recover in HRSG; however, this requires re-adjusting the syngas/HRSG steam ratio, and only the gasification temperature was modified in this study. Future work should focus on global optimization and maximizing the net efficiency without minimizing hydrogen production.

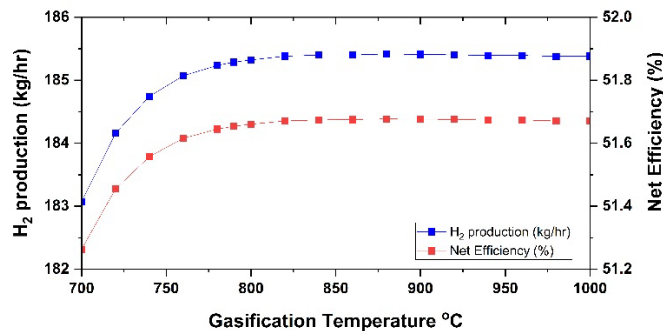


Figure 4: Effect of varying the gasification temperature on the H<sub>2</sub> production and net efficiency



#### 4. Conclusion

In order to achieved Net Zero targets by 2050, two important pathways are negative emissions with BECCS and industrial decarbonization via hydrogen. Within the study we evaluate a process for BECCS-Hydrogen-Power, using process modelling in Aspen Plus to identify the key operating parameters and key performance indicators. Using 3,600 kg/hr of wood chips can produce 185 kg/hr of H<sub>2</sub>, with an overall net efficiency of 52%. Within the study we also highlight the effects of increasing the gasification temperature. Future work should investigate additional heat integration options at high gasification temperatures, as well as analyze the effect of different separation technologies post-WGS.

#### 5. References

- BEIS, 2021. *Net Zero Strategy: Build Back Greener*, s.l.: Crown copyright.
- Binder, M., Kraussler, M., Kuba, M. & Luisser, M., 2018. *Hydrogen from biomass gasification*, s.l.: IEA Bioenergy.
- Cao, L. et al., 2020. Biorenewable hydrogen production through biomass gasification: A review and future prospects. *Environmental Research*, Volume 186, p. 109547.
- CCC, 2020. *The Sixth Carbon Budget: The UK's path to Net Zero*, s.l.: Committee on Climate Change.
- Chiesa, P., Consonni, S., Kreutz, T. & Williams, R., 2005. Co-production of hydrogen, electricity and CO<sub>2</sub> from coal with commercially ready technology. Part A: Performance and emissions. *International Journal of Hydrogen Energy*, Volume 30, pp. 747-767.
- Cohce, M. K., Rosen, M. A. & Dincer, I., 2011. Efficiency evaluation of a biomass gasification-based hydrogen production. *International Journal of Hydrogen Energy*, 36(17), pp. 11388-11398.
- Ishaq, H. & Dincer, I., 2020. A new energy system based on biomass gasification for hydrogen and power production. *Energy Reports*, Volume 6, pp. 771-781.
- Lepage, T., Kammoun, M., Schmetz, Q. & Richel, A., 2021. Biomass-to-hydrogen: A review of main routes production, processes evaluation and techno-economical assessment. *Biomass and Bioenergy*, Volume 144, p. 105920.
- Marcantonio, V. et al., 2019. Process analysis of hydrogen production from biomass gasification in fluidized bed reactor with different separation systems. *International Journal of Hydrogen Energy*, Volume 44, pp. 10350-10360.
- Moneti, M. et al., 2016. Influence of the main gasifier parameters on a real system for hydrogen production from biomass. *International Journal of Hydrogen Energy*, 41(28), pp. 11965-11973.
- Mutlu, Ö. Ç. & Zeng, T., 2020. Challenges and Opportunities of Modeling Biomass Gasification in Aspen Plus: A Review. *Chemical Engineering & Technology*, 43(9), pp. 1674-1689.
- Safarian, S., Unnthorsson, R. & Richter, C., 2022. Performance Investigation of Biomass Gasification for Syngas and Hydrogen Production Using Aspen Plus. *Open Journal of Modelling and Simulation*, Volume 10, pp. 71-87.
- Shahbaz, M. et al., 2021. A comprehensive review of biomass based thermochemical conversion technologies integrated with CO<sub>2</sub> capture and utilisation within BECCS networks. *Resources, Conservation and Recycling*, Volume 173, p. 105734.
- Tauqir, W., Zubair, M. & Nazir, H., 2019. Parametric analysis of a steady state equilibrium-based biomass gasification model for syngas and biochar production and heat generation. *Energy Conversion and Management*, Volume 199, p. 111954.
- Tezer, Ö. et al., 2022. Biomass gasification for sustainable energy production: A review. *International journal of Hydrogen Energy*, Volume 47, pp. 15419-15433.
- UNIFHY, 2016. *Final Report Summary - UNIFHY (UNIQUE gasifier for hydrogen Production)*, s.l.: s.n.
- Wei, L. et al., 2009. Syn-Gas Quality Evaluation for Biomass Gasification with a Downdraft Gasifier. *Transactions of the American Society of Agricultural and Biological Engineers*, 52(1), pp. 21-37.
- Wilkes, M. D., Mukherjee, S. & Brown, S., 2021. Linking CO<sub>2</sub> capture and pipeline transportation: sensitivity analysis and dynamic study of the compression train. *International Journal of Greenhouse Gas Control*, Volume 111, p. 103449.
- Zaman, S. A. & Ghosh, S., 2022. Thermo-economic and environmental performance analyses of a biomass-based carbon negative system integrating externally fired gas turbine and molten carbonate fuel cell. *Energy Conversion and Management: X*, Volume 14, p. 100187.
- Zhu, X. et al., 2020. Application of elevated temperature pressure swing adsorption in hydrogen production from syngas. *Adsorption*, Volume 26, pp. 1227-1237.

# Optimizing the Sustainable Energy Transition: A Case Study on Trinidad and Tobago

Sherard Sadeek,<sup>a</sup> Dhurjati Chakrabarti,<sup>a</sup> Maria M. Papathanasiou,<sup>bc</sup> Keeran Ward<sup>d</sup>

<sup>a</sup> *Department of Chemical Engineering, The University of the West Indies, St. Augustine, Trinidad and Tobago.*

<sup>b</sup> *Sargent Centre for Process Systems Engineering (SCPSE), Imperial College London, South Kensington Campus, London SW7 2AZ, UK*

<sup>c</sup> *Department of Chemical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, UK;*

<sup>d</sup> *School of Chemical and Process Engineering (SCAPE), University of Leeds, Leeds, LS29JT, UK;*

## Abstract

Trinidad and Tobago is one of the largest emitters of CO<sub>2</sub> per capita globally, with a significant reliance on oil and gas sectors. With the country's commitment, as a small island developing state (SIDS), to sustainable development goals and climate change agreements, rapid redesign of Trinidad and Tobago's power sector is critical to promoting a sustainable energy transition. Hence this study implements a mixed-integer linear programming model (MILP) to assess the levelized cost of electricity (LCOE) and life-cycle greenhouse gas emissions (GHG<sub>LC</sub>) across five scenarios. The results illustrate that with an improvement in power generation technology and resource efficiency, reduction of LCOE and GHG<sub>LC</sub> of up to 40% and 24% respectively are possible. For 2030, our results indicate an estimated increase of 29% and 5% for LCOE and GHG<sub>LC</sub>, respectively compared to the equivalent scenarios in 2019. Ultimately, through a multi-objective optimisation framework, our results highlight the value of systems-based planning and implementation in the sustainable energy transition across the Caribbean region, in accordance with sustainable development goals (SDGs).

**Keywords:** Sustainable development goals, Power generation, Multiple decision criteria, Life cycle assessment, Small Island Developing States.

## 1. Introduction

The Caribbean comprises of 13 island nations-most of which are classed as developing states<sup>1</sup>. Across the Caribbean archipelago, energy is an expensive necessity as most of the region is heavily reliant on fossil fuels and current market demand in meeting its energy needs. The deployment of clean and affordable energy (SDG 7) is integral in combatting the effects of climate change and promoting climate action (SDG 13). A holistic approach is required to aid in this transition, considering all aspects of a robust and sustainable energy system supporting the development of sustainable cities and communities (SDG 11)<sup>2</sup>. This transition, however, remains a challenging task as both economic and environmental constraints need to be optimized to guarantee SDGs are met across the wider global population.

Despite being a small island developing state (SIDS), Trinidad and Tobago is endowed with natural resources and is a net producer of energy, with an emphasis on - natural gas and oil. The country also relies heavily on its petrochemical sector (including methanol (MeOH) and ammonia (NH<sub>3</sub>)), which utilizes steam reforming to produce hydrogen (H<sub>2</sub>), along with significant amounts of grid electricity. This results in alarming rates of CO<sub>2</sub> at a national level and the country must emphasize sustainable forms of energy to meet its 15% GHG reduction pledge<sup>3</sup>.

By utilizing a multi-objective optimization framework, guidance on achieving clean and affordable energy through whole systems thinking can be fostered and utilized locally or domestically across the Caribbean Region, and globally among several SIDS. Hence, our work aims to utilize mixed integer linear programming (MILP) methodologies to minimize both levelized electricity costs (LCOE) and overall life-cycle greenhouse gas emissions (GHG<sub>LC</sub>). One of the key novelties of the presented work is that our MILP framework is tailored to resource availability, local infrastructure, and energy system capabilities of SIDS. This is of critical importance as previous works focused on developed nations.

## 2. Methodology

### 2.1. Overview

This work employs MILP to investigate LCOE and GHG<sub>LC</sub> in Trinidad and Tobago. The model involves the use of binary integer variables to represent each power plant, thus allowing the model the opportunity to select which ones can be used. Each plant was constrained to a minimum capacity factor of 0.3. The key variables used included the energy required by each plant for both peak and normal times (assumed to be 12 hours each), along with the natural gas utilized. These variables link directly with costs and emissions as seen in **Sections 2.6 and 2.7**. The system was constrained to meet the national electricity demand of 8.87 TWh (2019), with a daily peak load of 1350 MW. These factors are examined for various scenarios including the production of H<sub>2</sub> from polymer electrolyte membrane electrolysis (PEM) to provide an alternative to present reforming practices linked to the petrochemical sector for downstream products like MeOH and NH<sub>3</sub>.

### 2.2. Scenario 1A & 1B: Business-as-Usual (BAU)

This case examines the present power generation system in Trinidad and Tobago, with four natural gas-based power plants. There is a total power capacity of 1950MW, with two single cycle (SC) and two combined-cycle (CC) plants accounting for 990MW and 960MW, respectively. Scenario 1A highlights the use of the current take or pay power purchase agreement (PPA), ensuring that all power plants were online. Scenario 1B optimizes the present system without this PPA, which allows the model the choice of online power plants, if possible.

### 2.3. Scenario 2: Upgrade of Plant Efficiency

Currently, the largest power plant (Plant A) in the country employs SC gas turbine technology, with an installed capacity of 765 MW and an average efficiency of 23%<sup>4</sup>. The upgrade of this plant, results in a total of 1725 MW of CC grid power available, with Plant B (225 MW) being the sole SC plant serving the national power grid.

#### 2.4. Scenario 3: H<sub>2</sub> Production

Scenarios 3,4 and 5 examine the production of downstream products into the power sector supply and its ability to reduce LCOE and GHG<sub>LC</sub>. Scenario 3 employs H<sub>2</sub> generation via PEM. PEM is used as it is more efficient versatile, and commercial feasible over alkaline electrolysis and solid oxide electrolysis<sup>5</sup>. The power demand of this process is exclusively supplied by the upgraded CC plant (Plant A) mentioned in Scenario 2.

#### 2.5. Scenario 4 and 5: MeOH Production and NH<sub>3</sub> Production

The production of MeOH and NH<sub>3</sub> is explored via electrolytic H<sub>2</sub> as highlighted in Scenario 3. MeOH and NH<sub>3</sub> production, via CO<sub>2</sub> hydrogenation and the Haber-Bosch process respectively, was modelled in accordance to previous work<sup>6-7</sup>. Production of these downstream commodities with electrolytic H<sub>2</sub> provides a comparison between the country's use of resources (especially natural gas), costs and environmental impact, presently with this alternative method.

#### 2.6. Economic Feasibility

The economic impact of each scenario was defined by the levelized cost (\$<sub>USD</sub> per MWh), computed by estimating the total capital (CAPEX) and operating (OPEX) expenditure, for each process,  $r$ , in a given scenario. CAPEX costs were derived from literature<sup>6,8</sup> for each process technology while OPEX<sup>7-9</sup> was calculated from fixed and variable costs associated with maintenance, labour, capital overheads, purchasing feedstocks (including natural gas), insurance, royalties and other process-based costs. These costs were levelized based on the grid energy required for the national grid, on a yearly basis. This is shown in **Equation 1**.

$$LCOE = CAPEX_r + OPEX_r \quad \forall r \quad (1)$$

#### 2.7. Environmental Impact

A consequential life cycle assessment (LCA) was used to examine the environmental impacts related to GHG<sub>LC</sub> for a function unit of 1MWh of energy produced. A cradle-to-power generation gate approach was followed using the ISO 14040:2006 methodology<sup>10</sup>. Mathematically, each process  $r$  for a given scenario consumes feedstocks  $P_i$  ( $kg p_i$ /MWh) while emitting waste streams,  $W_r$  ( $kg w_r$ /MWh). Furthermore, each feedstock and waste stream is linked to a characterized environmental impact (GWP<sub>100</sub>)  $I_i^{IN}$  ( $kg CO_{2eq}/kg P_i$ ) and  $I_r^{OUT}$  ( $kg CO_{2eq}/kg W_r$ ), respectively. For multiple product systems (Scenarios 3-5), avoided burdens  $I_r^{COMM}$  ( $kg CO_{2eq}/kg pc_r$ ) relative to the BAU downstream processes,  $PC_r$  ( $kg pc_r$ /MWh) were recorded using Ecoinvent v3.4. The environmental impact, described as GHG<sub>LC</sub>, represents the total life cycle GHG emissions corresponding to the summation of embodied feedstock impacts and the impacts of each scenario as well as avoided burdens associated with the substitution of low carbon H<sub>2</sub>, MeOH and NH<sub>3</sub> and the utilization of CO<sub>2</sub>. This is seen in **Equation 2**.

$$GHG_{LC} = \sum_i P_i I_i^{IN} + \sum_r W_r I_r^{OUT} - PC_r I_r^{COMM} \quad \forall r, i \quad (2)$$

### 3. Results and Discussion

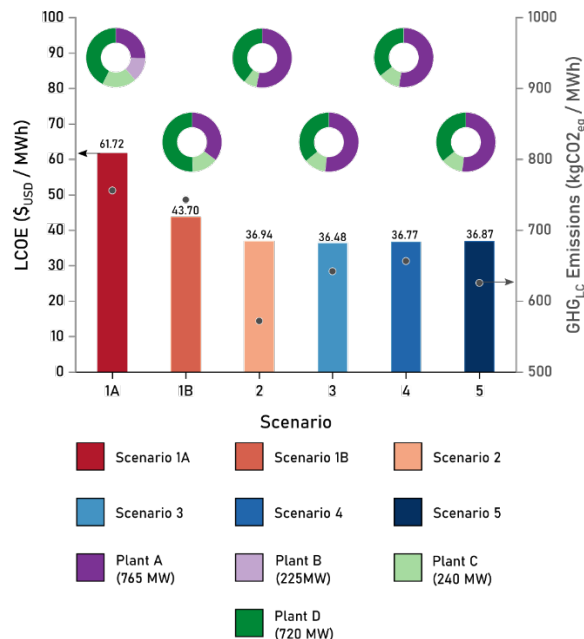
#### 3.1. 2019 SOO Outlook

Initially, our model was optimized using SOO with a focus on the LCOE. This is in line with the current national perspective on decision making.

Scenarios 1A and 1B employ the use of the present power generation grid technology. However, Scenario 1B explores the removal of the current take or pay PPA which results in an approximate LCOE reduction of \$<sub>USD</sub> 18 per MWh (29% decrease) compared to Scenario 1A (\$<sub>USD</sub> 61.72 per MWh). As seen in **Figure 1**, the elimination of the PPA allows the grid demand to be satisfied by only 3 power plants. The use of more efficient power plants also resulted in a reduction of GHG<sub>LC</sub> from 756 kgCO<sub>2eq</sub> to 743 kgCO<sub>2eq</sub>.

Scenario 2 resulted in a reduction of LCOE of approximately \$<sub>USD</sub> 25 per MWh (40% decrease) and GHG<sub>LC</sub> decreased to 572 kgCO<sub>2eq</sub> (24% reduction) compared to Scenario 1A. These improved results were as a result of the improved efficiency due to technological improvements on the power grid.

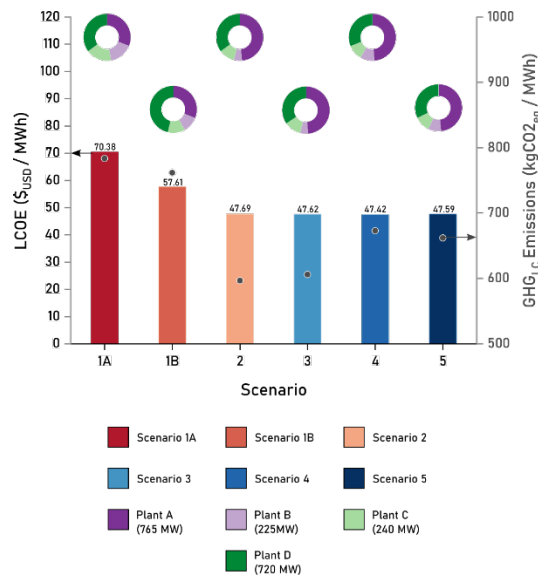
Scenarios 3-5 demonstrated a similar LCOE reduction as Scenario 2. However, these cases required improved commodity prices (H<sub>2</sub>, MeOH and NH<sub>3</sub>) for the production of downstream products to be considered feasible in this system. Also, GHG<sub>LC</sub> increased by 8 – 12% when compared to Scenario 2. Thus, it can be noted that Scenario 2 is the best performing case for both economic and environmental in assessment despite the use of single objective optimization (SOO).



**Figure 1: SOO 2019 highlighting LCOE, GHG<sub>LC</sub> and Grid share attributed to each Power Plant. Bars show LCOE prices while • illustrates GHG<sub>LC</sub>.**

### 3.2. 2030 MOO Outlook

The MOO analysis for 2030 (**Figure 2**) allows for insight into the current power system’s outlook and assessment with respect to the national goals. Due to the projected increase in grid demand and peak loads, four power plants were online across all scenarios. The results for 2030 followed a similar trend to 2019. However, there was an approximate increase in LCOE of \$<sub>USD</sub> (8.5 – 14) per MWh compared to the equivalent scenarios in 2019 SOO. The GHG<sub>LC</sub> increased by approximately (8 – 27) kgCO<sub>2eq</sub>, when compared to 2019 equivalent, with Scenario 3 being an exception. Scenario 3 demonstrated a reduction in the GHG<sub>LC</sub> of approximately 39 kgCO<sub>2eq</sub> when comparing 2030 to 2019. This was due to a reduction in electrolyser capacity, as a result of grid constraints. Overall, our results indicate an estimated increase of 29% and 5% for LCOE and GHG<sub>LC</sub>, respectively compared to the equivalent scenarios in 2019.



**Figure 2: MOO 2030 highlighting LCOE, GHG<sub>LC</sub> and Grid share attributed to each Power Plant. Bars show LCOE prices while • illustrates GHG<sub>LC</sub>.**

## 4. Conclusion

We present the use of multiple decision criteria, through multi-objective optimization, and its relevance in promoting clean and affordable energy among SIDS, using Trinidad and Tobago as a case study. This work demonstrates the ability to link energy supply chains for less developed countries in the Caribbean and globally, thus linking to sustainable development. Five scenarios were explored, utilizing MILP while incorporating a LCA framework, with LCOE and GHG<sub>LC</sub> in mind. Our results illustrate that the most optimal scenario would be the improvement of gas turbine technology from

SC to CC for the nation's largest power plant. This resulted in a LCOE reduction (for 2019) of 40% compared to BAU while  $\text{GHG}_{\text{LC}}$  reduces by 24% to 572  $\text{kgCO}_{2\text{eq}}$  per MWh. The introduction of downstream chemicals-  $\text{H}_2$ , MeOH and  $\text{NH}_3$  within the national energy system shows to be infeasible at current market prices. For 2030, our results indicate an estimated increase of 29% and 5% for LCOE and  $\text{GHG}_{\text{LC}}$ , respectively compared to the equivalent scenarios in 2019. These results validate the need for new policies linked to active decarbonisation such as CCS frameworks and carbon taxes, along with the reconsideration of present power purchase agreements. Our results underpin viable solutions-driving impact and distributing guidance on the sustainable energy transition across the Caribbean Region, through the deployment of clean and affordable energy.

## References

- United Nations, 2022, Office of the High Representative for the Least Developed Countries, Landlocked Developing Countries and Small Island Developing States. <https://www.un.org/ohrills/content/list-sids>
- United Nations, 2019, Take Action for the Sustainable Development Goals <https://www.un.org/sustainabledevelopment/sustainabledevelopment-goals/>.
- National Renewable Energy Laboratory, Energy Transition Initiative, 2015. Energy Snapshot Trinidad and Tobago. <https://www.nrel.gov/docs/fy15osti/64117.pdf>
- Inter-American Development Bank, 2015. A Unique Approach for Sustainable Energy in Trinidad and Tobago, <https://publications.iadb.org/en/publication/12365/unique-approach-sustainable-energy-trinidad-and-tobago>.
- International Renewable Energy Agency, 2020, Green Hydrogen Cost Reduction Scaling up Electrolysers to Meet the 1.5 °C Climate Goal [https://irena.org/-/media/Files/IRENA/Agency/Publication/2020/Dec/IRENA\\_Green\\_hydrogen\\_cost\\_2020.pdf](https://irena.org/-/media/Files/IRENA/Agency/Publication/2020/Dec/IRENA_Green_hydrogen_cost_2020.pdf).
- Narine, K., Mahabir, J., Koylass, N., Samaroo, N., Singh-Gryzbon, S., Baboolal, A., Guo, M., Ward, K., 2021, Climate smart process design for current and future methanol production. *J.CO2 Util.* 41, 101399, <http://dx.doi.org/10.1016/J.JCOU.2020.101399>.
- Bhagaloo, K., Ali, R., Baboolal, A., Ward, K., 2022b, "Powering sustainable transition with geothermal energy: A case study on Dominica," *Sus. Ener. Tech. Assess.* 51. <https://doi.org/10.1016/j.seta.2021.101910>.
- Bhagaloo, K., Baboolal, A., Ali, R., Razac, Z., Lutchmansingh, A., Mangra, A., Muhammad, T., Ward, K., 2022a. "Resource efficiency as a guide to clean and affordable energy: A case study on Trinidad and Tobago," *Chem. Eng. Res. Des.* 178: 405-420. <https://doi.org/10.1016/j.cherd.2021.12.026>.
- U.S Energy Information Administration, 2020, Capital Cost and Performance Characteristic Estimates for Utility Scale Electric Power Generating Technologies.
- International Organization for Standardization, 2016, Environmental Management – Life Cycle Assessment – Requirements and Guidelines. ISO 14044.

# Optimal operation of a sustainable water desalination unit integrated a green energy system

Ali Ahmadian,<sup>a</sup> Ali Almansoori,<sup>b</sup> Ali Elkamel,<sup>c</sup>

<sup>a</sup>*Department of Electrical Engineering, University of Bonab, Bonab, Iran.*

<sup>b</sup>*Department of Chemical Engineering, Khalifa University, Abu Dhabi 127788, United Arab Emirates*

<sup>c</sup>*Department of Chemical Engineering, University of Waterloo, 200 University Avenue West, Waterloo, ON, N2L 3G1, Canada*

## Abstract

In this paper, an optimization framework is proposed in order to optimal operation of a water desalination unit in a sustainable green energy network. The operation problem is modeled as a mixed integer linear programming (MILP) and optimized using a mathematical based optimization algorithm. In order to increase the flexibility of the studied sustainable energy system, a fuel cell unit and hydrogen-based vehicles have been included in the system. Moreover, the uncertainty of stochastic parameters has been considered and a machine learning based method namely long short term memory (LSTM) algorithm is utilized for forecasting of the parameters. The proposed methodology is applied on a typical case study and the simulation results are extracted and discussed.

**Keywords:** Water desalination unit, sustainable green energy, smart cities.

## 1. Introduction

Today's, the water demand addressing is a challenge in development countries and the policy makers and governments have been trying to increase the water resources. One of the main approach is utilization of water desalination units to produce water for domestic and industries. The water desalination units should be operated optimally in order to increase their benefits [1]. Several approaches have been proposed in the literature in order to increase the benefits of water desalination units. In [2], the renewable energy resources have been integrated with the water desalination units for producing a sustainable system. A geothermal energy based hybrid system is proposed in [3] and an experimental investigation is carried out in [4].

In this paper, the optimal utilization of desalination units and their interaction with the electricity network is implemented. In addition, to increase the flexibility of the proposed method fuel cell units and hydrogen-based vehicles are modeled. The optimization problem is a mixed-integer linear programming method that has the ability to solve large-scale problems efficiently and can be solved with well-known solvers such as CPLEX. Owing to the fact that uncertainty parameters especially those of RESs and load demand plays a decisive effect on the total cost of the system. In this paper, a deep learning-based method known as long short-term memory networks (LSTMs) a well-known method in time series forecasting is implemented for modeling different stochastic parameters.

## Nomenclature

$n, m$	Indices of buses
$e$	Index of emission types



$dg, r, es$	Index for DGs, RESs, and ESs
$t$	Index for time
$des$	Index for desalination unit
$\alpha_{dg}, \beta_{dg}, \gamma_{dg}$	DGs cost coefficients
$\Gamma_{e,dg}$	Emission factor of different emission types (\$/lb)
$G_{n,m}, B_{n,m}$	Conductance and susceptance of lines (p.u.)
$PD_{n,t}, QD_{n,t}$	Active and reactive load demands
$EP_t$	Electricity price (\$/kWh)
$PDG_{dg}^{MIN}, PDG_{dg}^{MAX}$	Minimum and maximum output power of DGs (kW)
$PCH_{es}^{MAX}, PDS_{es}^{MAX}$	Maximum charging and discharging power of ESs (kW)
$S^{BASE}$	Base apparent power (kVA)
$SL_{n,m}^{MAX}$	The maximum thermal limit of lines (kVA)
$SES_{es}^{MIN}, SES_{es}^{MAX}$	Minimum and maximum SOC of ESs (kWh)
$V_n^{MIN}, V_n^{MAX}$	Minimum and maximum voltage magnitude (p.u.)
$V_{des}^{T,MAX}$	Maximum volume of the tank (m <sup>3</sup> )
$V_{des}^{DES,MIN}, V_{des}^{DES,MAX}$	Minimum and maximum treated water by desalination unit (m <sup>3</sup> )
$W_t^D$	Water demand (m <sup>3</sup> )
$D^{EF}$	Efficiency of WDs
$\zeta_{dg}$	Externality emission cost of DGs (lb/kWh)
$\alpha_{es}^{ES}$	Degradation factor of ESs
$\alpha_r^{RES}$	Cost factor of RESs (\$/kW)
$\Delta t$	Length of time interval
$Q_{n,m,t}$	Linear load flow constant
$\eta_{es}^{CH}, \eta_{es}^{DCH}$	ES's charging and discharging efficiency
$\theta_n^{MIN}, \theta_n^{MAX}$	Minimum and maximum voltage angle (rad)
$PCH_{es,t}, PDS_{es,t}$	ES charge and discharge (kW)
$UP_t$	Exchanged power with the upstream network (kW)
$PDG_{dg,t}, PES_{es,t}, PRES_{r,t}$	Generated power by DG, ES, and RESs (kW)
$PDS_{des,t}$	Consumed power by WD unit (kW)
$PG_{n,t}, QG_{n,t}$	Generated active and reactive powers in each bus
$P_{n,m,t}, Q_{n,m,t}, S_{n,m,t}$	Active, reactive, and apparent power flow (kW)
$SES_{es,t}$	SOC of ESs (kWh)
$I_{es,t}^{CH}, I_{es,t}^{DIS}$	Binary variables for ESs

$U_{dg,t}$	Binary variables for on/off of DGs
$V_{n,t}$	Voltage magnitude (p.u.)
$V_{des,t}^T$	The volume of stored water in the tank (m <sup>3</sup> )
$V_{des,t}^{T,INP}$	Input water to the tank (m <sup>3</sup> )
$\theta_{n,t}$	Voltage angle (rad)

## 2. Problem Formulation

Due to the water crisis in the world, this paper aims to find optimal management for a smart city to satisfy the electric and water demand of residents. In the proposed method we consider water desalination units besides renewable energy resources (RESs). In this regard, we can get the most out of RESs by using their surplus power for making drinking water. The main objective is minimizing cost as demonstrated in equation (1).

$$\text{Minimizing Cost} = \sum_t UP_t EP_t + \sum_t \sum_{dg} ((\alpha_{dg} PDG_{dg,t}^2 + \beta_{dg} PDG_{dg,t} + \gamma_{dg}) U_{dg,t} + \sum_e \zeta_{dg} \Gamma_{e,dg} PDG_{dg,t}) + \sum_t \sum_{es} \alpha_{es}^{ES} PES_{es,t} + \sum_t \sum_r \alpha_r^{RES} PRES_{r,t} \quad (1)$$

The total cost consists of the cost of exchanging power with the upstream network, the cost of DGs considering their emission cost, the degradation cost of ESs, and the RESs cost. DGs constraints for the maximum and minimum generated active power is shown in (2). The other constraints such as ramp rate and on/off limits are modeled based on reference [5].

$$PDG_{dg}^{MIN} U_{dg,t} \leq PDG_{dg,t} \leq PDG_{dg}^{MAX} U_{dg,t} \quad (2)$$

Equations (3)-(8) show the operational constraints of ESs. (3) show the level of energy stored in the ES. The amount of power delivered or injected into the grid is calculated based on (4). Equations (5)-(7) demonstrate the charging and discharging state of these units. Finally, (8) restricts the state of charge of ESs.

$$SES_{es,t} = SES_{es,t-1} + (PCH_{es,t} \eta_{es}^{CH} - PDS_{es,t} / \eta_{es}^{DCH}) \Delta t \quad (3)$$

$$PES_{es,t} = PDS_{es,t} - PCH_{es,t} \quad (4)$$

$$0 \leq PCH_{es,t} \leq PCH_{es}^{MAX} I_{es,t}^{CH} \quad (5)$$

$$0 \leq PDS_{es,t} \leq PDS_{es}^{MAX} I_{es,t}^{DIS} \quad (6)$$

$$I_{es,t}^{CH} + I_{es,t}^{DIS} \leq 1 \quad (7)$$

$$SES_{es}^{MIN} \leq SES_{es,t} \leq SES_{es}^{MAX} \quad (8)$$

In order to model the water desalination units (WDs) equations (9)-(12) are taken into account. To increase the flexibility of WDs a tank for storing water is considered. The volume of water stored in this tank is calculated based on (9). The volume of input water and the minimum and maximum of stored water in this unit are controlled by (10) and (11). (12) indicate the relation between purified water and the power consumed by WD.

$$V_{des,t}^T = V_{des,t-1}^T + V_{des,t}^{T,INP} - WD_t \quad (9)$$

$$0 \leq V_{des,t}^T \leq V_{des}^{T,MAX} \quad (10)$$

$$V_{des}^{DES,MIN} \leq V_{des,t}^{T,INP} \leq V_{des}^{DES,MAX} \quad (11)$$

$$V_{des,t}^{T,INP} = PDS_{des,t} D^{EF} \quad (12)$$

For modeling the load flow constraints a linearized load flow method based on [6] is used. Active and reactive power balances are checked by (13) and (14). Equations (15)-(17) express active, reactive, and apparent power flows through the network. Equations (18)-(20) are used to control network limits.

$$(PG_{n,t} - PD_{n,t})/S^{BASE} = (2V_{n,t} - 1)G_{n,n} + \sum_{m \neq n} G_{n,m} (V_{n,t} + V_{m,t} - 1) + B_{n,m}(\theta_{n,t} - \theta_{m,t}) \quad (13)$$

$$(QG_{n,t} - QD_{n,t})/S^{BASE} = -(2V_{n,t} - 1)B_{n,n} + \sum_{m \neq n} -B_{n,m} (V_{n,t} + V_{m,t} - 1) + G_{n,m}(\theta_{n,t} - \theta_{m,t}) \quad (14)$$

$$P_{n,m,t} = G_{n,m}(V_{n,t} - V_{m,t}) + B_{n,m}(\theta_{n,t} - \theta_{m,t}) \quad (15)$$

$$Q_{n,m,t} = -B_{n,m}(V_{n,t} - V_{m,t}) + G_{n,m}(\theta_{n,t} - \theta_{m,t}) \quad (16)$$

$$S_{n,m,t} = P_{b,\hat{b},t} + Q_{n,m,t} \quad (17)$$

$$V_n^{MIN} \leq V_{n,t} \leq V_n^{MAX} \quad (18)$$

$$\theta_n^{MIN} \leq \theta_{n,t} \leq \theta_n^{MAX} \quad (19)$$

$$|S_{n,m,t}| \leq SL_{n,m}^{MAX} \quad (20)$$

### 2.1. Uncertainty Modeling

Nowadays, due to access to big data in power systems. Data-driven-based approaches especially deep-learning-based approaches known as promising tools for modeling different stochastic parameters. One of the well-known methods for forecasting time series stochastic parameters is long-short-term memory networks (LSTMs). The fundamentals and formulation of LSTMs are presented in [7].

## 3. Numerical results and discussion

### 3.1. Input data

To evaluate the proposed method a network as presented in Fig. 1 is considered. As shown in Fig. 1 the system contains two wind turbines (WTs), two photovoltaic panels (PVs), two WDs, and three DGs. Input data related to DG and ES and other data are extracted from reference [8] and the emission data are obtained based on [9].

### 3.2. Simulation Results

After running the optimization problem based on the above-mentioned data. The

results for the study case are as follows. The amount of consumption power by WD units is illustrated in Fig. 2. As stated in this figure and fig. 4 to utilize the maximum benefit of RESs, WDs are used during hours when the total power of RESs is high which proves the good performance of the proposed method.

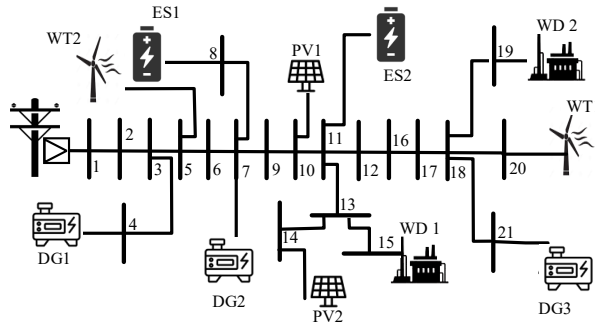


Fig. 1. Single-line diagram of the case study.

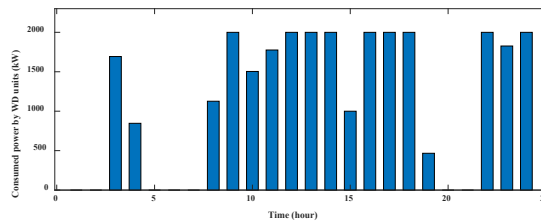


Fig. 2. Consumed power by WD units.

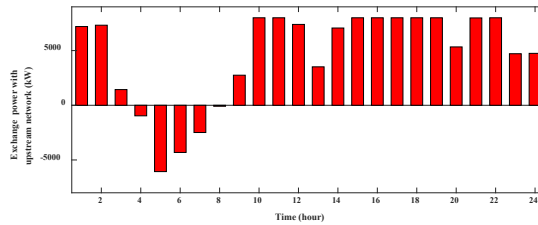


Fig. 3. exchange power with the upstream network.

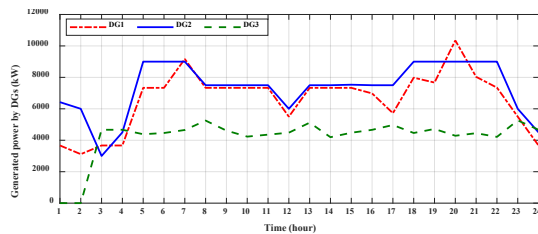


Fig. 4. Generated power by DGs.

Fig. 3 shows the amount of power exchange with the upstream network. By comparing this result with Fig. 2 it is obvious that the algorithm adjusts its power based on the

electricity price, RESs generation, and water demand. Fig. 4 indicates the output power of each DG during different hours. A detail of the resources' cost is depicted in Table 1.

**Table 1: Detail of distribution network cost.**

Parameters	Cost (\$)
DG cost	25509
Power exchange cost	9483
Total emission cost	720
RESs cost	1045
Total Cost	36757

#### 4. Conclusion

In this paper, a water desalination unit is operated optimally in a sustainable green energy system. The uncertainty of stochastic parameters has been considered using a machine learning based approach and the operation problem is solved as a mixed integer linear programming (MILP) using a mathematical based optimization algorithm. The simulation results verified the effectiveness and robustness of the proposed approach in reducing the operation cost of the sustainable energy system and the emission pollution.

#### References

- [1] Peng, Wanxi, Akbar Maleki, Marc A. Rosen, and Pouria Azarikhah. "Optimization of a hybrid system for solar-wind-based water desalination by reverse osmosis: Comparison of approaches." *Desalination* 442 (2018): 16-31.
- [2] Bundschuh, Jochen, Michał Kaczmarczyk, Noredine Ghaffour, and Barbara Tomaszewska. "State-of-the-art of renewable energy sources used in water desalination: Present and future prospects." *Desalination* 508 (2021): 115035.
- [3] Okati, V., Amir Ebrahimi-Moghadam, A. Behzadmehr, and Mahmood Farzaneh-Gord. "Proposal and assessment of a novel hybrid system for water desalination using solar and geothermal energy sources." *Desalination* 467 (2019): 229-244.
- [4] Mostafa, Marwa, Heba M. Abdullah, and Mohamed A. Mohamed. "Modeling and experimental investigation of solar stills for enhancing water desalination process." *IEEE Access* 8 (2020): 219457-219472.
- [5] Gougheri SS, Dehghani M, Nikoofard A, Jahangir H, Golkar MA. Economic assessment of multi-operator virtual power plants in electricity market: A game theory-based approach. *Sustain Energy Technol Assessments* 2022;53:102733.
- [6] Sadeghi S, Jahangir H, Vatandoust B, Golkar MA, Ahmadian A, Elkamel A. Optimal bidding strategy of a virtual power plant in day-ahead energy and frequency regulation markets: A deep learning-based approach. *Int J Electr Power Energy Syst* 2021;127:106646. <https://doi.org/https://doi.org/10.1016/j.ijepes.2020.106646>.
- [7] Jahangir H, Tayarani H, Gougheri SS, Golkar MA, Ahmadian A, Elkamel A. Deep Learning-based Forecasting Approach in Smart Grids with Micro-Clustering and Bi-directional LSTM Network. *IEEE Trans Ind Electron* 2020.
- [8] Gougheri SS, Jahangir H, Golkar MA, Ahmadian A, Golkar MA. Optimal participation of

a virtual power plant in electricity market considering renewable energy: A deep learning-based approach. *Sustain Energy, Grids Networks* 2021;26:100448.

- [9] Mohamed FA, Koivo HN. System modelling and online optimal management of microgrid using mesh adaptive direct search. *Int J Electr Power Energy Syst* 2010;32:398–407.



# Optimal production of green hydrogen with grid assistance for enhanced flexibility

Christopher Varela,<sup>a,b,c</sup> Mahmoud Mostafa,<sup>a</sup> Edwin Zondervan,<sup>d</sup>

<sup>a</sup>*Laboratory of Process Systems Engineering, University of Bremen, Leobener Str. 6, 28359 Bremen, Germany*

<sup>b</sup>*Facultad de Ciencias Naturales y Matemáticas, Escuela Superior Politécnica del Litoral, ESPOL, Campus Gustavo Galindo, Km. 30.5 Vía Perimetral, Guayaquil P.O. Box 09-01-5863, Ecuador*

<sup>c</sup>*Centro de Energías Renovables y Alternativas, CERA, Escuela Superior Politécnica del Litoral, ESPOL, Campus Gustavo Galindo, Km. 30.5 Vía Perimetral, Guayaquil P.O. Box 09-01-5863, Ecuador*

<sup>d</sup>*SPT-PSE, University of Twente, 7522NB Enschede, The Netherlands*

## Abstract

The use of grid power to compensate for shortages of renewable energy results in higher profit at electrolysis facilities. However, the carbon footprint of hydrogen increases in regions where fossil sources have a significant share in the energy mix. This contribution seeks to balance profit with carbon footprint through mathematical programming. The case study comprises a 100-MW<sub>p</sub> wind power plant, and data on electricity price and CO<sub>2</sub>-emissions content in the grid, covering a full-year of operation in northern Germany. The results show that the highest profit is attained at maximum load, with a minor influence of the electricity price (<75 €/MWh) and no influence of the CO<sub>2</sub>-emissions content of the grid. In this scenario, approximately 50% of the power stems from conventional sources. Therefore, a cost penalty is proposed to the use of fossil power to limit the grid assistance. This further reduces the grid contribution to around 30%, still allowing for a high profit.

**Keywords:** Energy transition, green hydrogen, Power-to-X, MILP.

## 1. Main Text

The main text can start here. Next paragraphs should start with heading as given in below example.

## 2. Introduction

The efficient conversion and long-term storage of renewable energy is a key aspect within the energy transition concept. This can be achieved by for example Power-to-X (PtX) technologies, using green hydrogen as renewable energy carrier. It is expected that green hydrogen will meet a global demand of 74 exajoules per year by 2050, with a production cost below 1.25 \$/kg [1]. While the economic figures seem promising, scaling-up the production of green hydrogen requires operation strategies to deal with volatile renewable energy and the dynamics of energy markets.

In earlier work a method to determine the optimal number of electrolyzers and their optimal operation schedule for an off-grid electrolysis facility was presented [2].



However, the connection to power grids provides additional flexibility to move the operation of electrolyzers towards more favorable conditions. This comprises demand-side management and provision of grid services [3]. On the other hand, loading power beyond the available renewable capacity increases the carbon footprint of hydrogen, especially in periods with low share of renewables in the energy mix. This work includes the grid power in the scheduling program of alkaline water electrolysis (AEL). The grid is expected to participate mainly in periods of low CO<sub>2</sub>-emissions content, low electricity prices, and to avoid unnecessary shutdowns. In addition, the program includes a penalty to avoid intensive use of the grid. This penalty is based on a time-variant profile of CO<sub>2</sub>-emissions content in the grid. In such way, the grid assists the PtX facility maximizing the profit with minimum carbon footprint.

### 3. Methodology

The water electrolysis optimization program (available in [2], [4]) determines the optimal schedule of electrolyzers with profiles of renewable energy and electricity price. It is based on a timed-automaton model that represents the states and power load of a commercial electrolysis stack. The constraints describe the physical and technical restrictions of the system (startup/shutdown time, load range, etc.). The current case study considers a reference 100-MW<sub>p</sub> wind power supply (Figure 1) and the power grid with a theoretical unlimited capacity. Agorameter [5] reports time-variant data on electricity price and CO<sub>2</sub>-emissions content in the grid for Germany. Data of the year 2019 was considered. The objective function is set to maximize the profit with a penalty on the fossil power loaded from the grid.

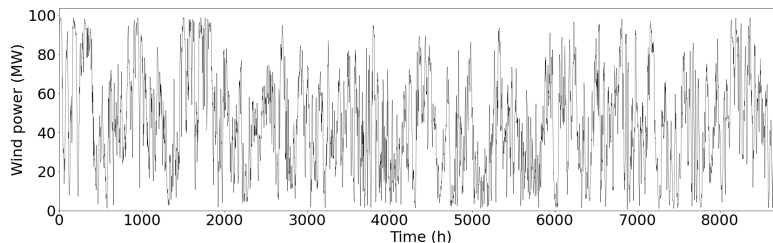


Figure 1. Wind power supply (54°10'29.2"N 9°03'51.9"E, Vestas V150, 80 m hub height) [6], [7]

The results are obtained for a base case scenario that reflects the operation of PtX within the current energy market and AEL performance. Furthermore, a sensitivity analysis allows estimating the effect on the profit and carbon footprint of the process for other scenarios. Such scenarios reflect different operation goals, technological development, and market trends. The selected parameters are shown in Table 1.

Table 1. AEL-model parameters

Parameter	Base case	Sensitivity analysis
Annual production (MNm <sup>3</sup> )	110	50 – 200
Investment cost per 3.5-MW electrolyzer (M€)	1.96	0.5 – 2.0
Penalty cost of CO <sub>2</sub> emissions (€/tCO <sub>2</sub> )	40	0 – 200

The annual demand is set to around 60% of the maximum production capacity for the base case, with an evenly distributed production per month. An increased demand of hydrogen requires either adding more electrolyzers or making intensive use of the grid. The investment cost of electrolyzers is expected to decrease in the next few years, for

which lower values are mainly considered. The penalty on CO<sub>2</sub>-emissions is a cost, which is set to around the order of magnitude of values reported by the European Trading System.

#### 4. Results and discussion

The program is solved using the optimizer GUROBI/PYTHON [8]. The number of variables is in the order of 1.6 million, as it considers a full-year horizon with time intervals of 1 hour and several electrolysis stacks. Despite the large extension of the program, finding optimal solutions in a few seconds is feasible due to its linear structure.

The program finds the optimal schedule for each electrolyzer considering the whole time horizon, while the number of installed electrolyzers has to be fixed at first. The base case scenario presents the solution with 27 installed electrolyzers of 3.1 MW with load ranges from 10 to 100% each. The sensitivity analysis, on the other hand, includes MILP-solutions generated by setting the number of electrolyzers from 20 to 34, while varying the parameters mentioned in Table 1.

##### 4.1. Base case scenario

The optimal solution shows that the facility should mainly operate at full capacity to increase the profit, despite the penalty added for the grid use. Figure 2 contains the load of the electrolysis facility (almost constant profile), while the available renewable power appears in the background. It shows many periods in which the facility uses the grid to compensate for the lack of renewable power. The grid contributed around 46.8% of the total power, while the overall production cost and carbon footprint were estimated to be 2.46 €/kg<sub>H2</sub> and 10.83 kg<sub>CO2</sub>/kg<sub>H2</sub>, respectively. Since fossil hydrogen (from SMR) has an approximate production cost of 1.5 €/kg<sub>H2</sub> and a carbon footprint of 10 kg<sub>CO2</sub>/kg<sub>H2</sub>, AEL with major grid assistance is economically and environmentally unviable under the case study conditions. Similar results were obtained with a different number of electrolyzers added to the facility.

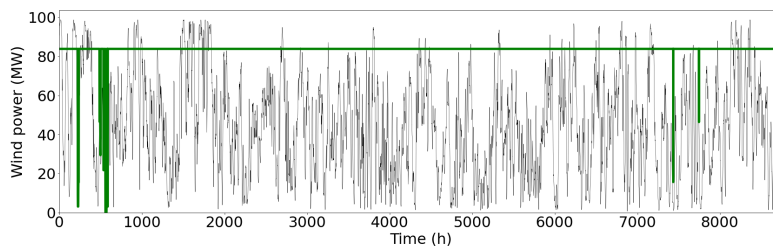


Figure 2: Optimal plant load with 27 electrolyzers

Moreover, there are few periods where the process undergoes a partial or total shutdown. These periods coincide with periods of electricity prices greater than 75 €/MWh. Thus, periods of high electricity price trigger total or partial shutdown of electrolyzers, regardless of the renewable power availability or CO<sub>2</sub>-emissions content in the grid. The process response to such scenario is shown in Figure 3, which contains the load of electrolyzers in the time interval of 500 to 600 hours. Few electrolyzers report partial loads, keeping rather their operation at maximum capacity or shutdown.

##### 4.2. Sensitivity analysis

The sensitivity analysis is performed to explore optimal solutions under more optimistic conditions to the energy transition, namely lower investment cost of electrolyzers or restricted use of fossil-based power. It is of special interest to identify under which

scenarios the process remains cost-competitive while the carbon footprint considerably decreases as compared to the base case scenario. The MILP considers 20 to 34 installed electrolyzers and parameter values within the range shown in Table 1. The sensitivity analysis is locally performed, meaning that it determines the individual influence of the parameters. The surface of the objective function in relation to the parameters are projected in the subplots of Figure 4, with the inline figure indicating the objective value in M€/a.

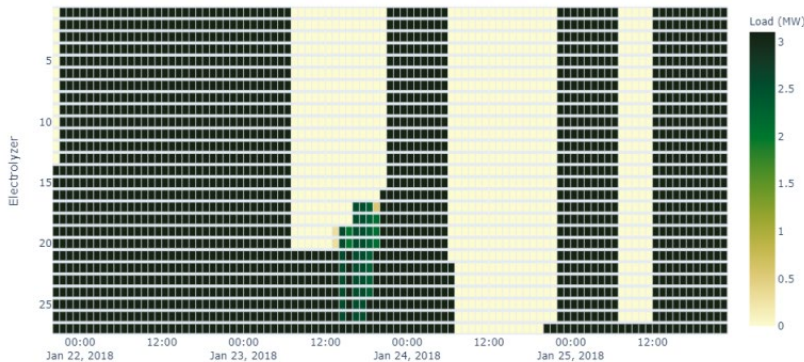


Figure 3: Optimal schedule of electrolyzers during the time 500 to 600 h

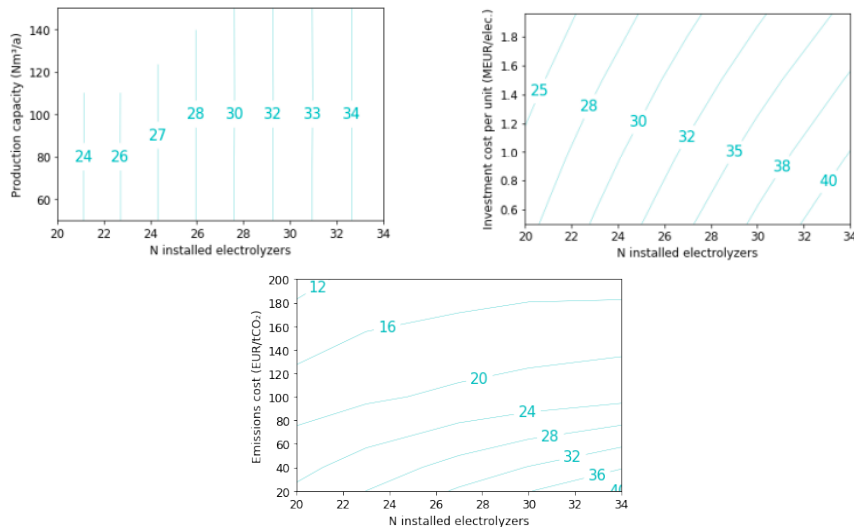


Figure 4: Parameters sensitivity to the objective function in M€/a: (top left) production capacity, (top right) purchasing cost of electrolyzers, and (bottom) cost of CO<sub>2</sub> emissions.

The effect of the production capacity is negligible, as the electrolysis facility always load at its full capacity to increase the profit. This means that optimal solutions correspond to steady-state operations with few executed startup/shutdown cycles, major grid assistance, and production rate that exceeds the monthly demand. The purchasing cost of electrolyzers has a relevant impact on the objective function. For the most optimistic scenarios (low capital cost) the objective value increases by around 15% as compared to the base case. This can be possible considering the expected technological developments in future years. However, both parameters have no effect on the production schedule, keeping the loading of electrolyzers always at their full capacity. On the other hand, the

objective function presents a high sensitivity to the CO<sub>2</sub>-emissions cost. For values lower than 40 €/tCO<sub>2</sub>, the facility opts operations at full capacity, similarly to the base case scenario. While for costs higher than 40 €/tCO<sub>2</sub>, lower carbon footprints are computed with considerably lower annual production of hydrogen. The results are shown in Table 2.

Table 2: Sensitivity analysis results

	CO <sub>2</sub> -emissions penalty cost (€/tCO <sub>2</sub> )				
	0	20	40	100	200
Annual production (MNm <sup>3</sup> )	156.10	155.83	155.46	135.29	110.00
Production cost (€/kgH <sub>2</sub> )	2.47	2.46	2.46	2.43	2.52
Carbon footprint (kgCO <sub>2</sub> /kgH <sub>2</sub> )	20.34	10.88	10.83	8.57	5.27
Annual profit without penalty (M€)	35.55	35.54	35.51	31.23	24.49
Grid assistance (%Power)	95.43	46.96	46.84	39.84	27.29
Operation strategy	Full load	Full load	Full load	Partial load	Load following

It is noted that the demand constraint (at 110 MNm<sup>3</sup>/a, evenly distributed each month) restricts the carbon footprint to a minimum of 5.27 kgCO<sub>2</sub>/kgH<sub>2</sub>. This also means that the grid had to provide at least 27.29% of the total power to cover the hydrogen demand. The operation however presents a different load profile as compared to the base case, as shown in Figure 5, for a snapshot of the operation for the period from 2000 to 3000 h.

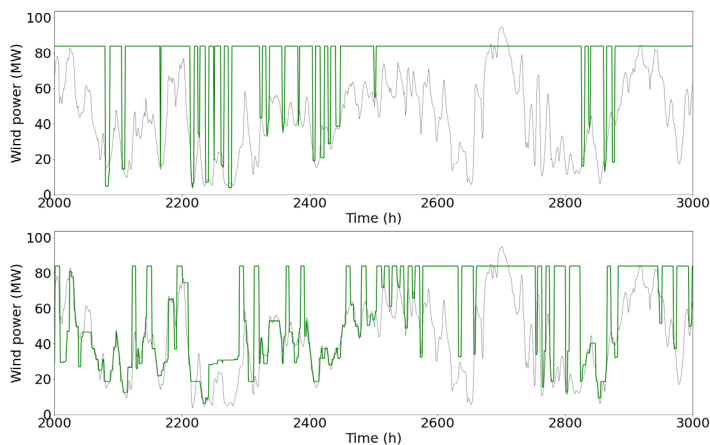


Figure 5: Load profile of the AEL facility for (top) emission cost of 100 €/tCO<sub>2</sub>, and (bottom) emission cost of 200 €/tCO<sub>2</sub>

This result shows how the penalty allowed moving the operation strategy from a steady-state load profile to a load-following profile (same behavior as the renewable power). An intermediate profile, denoted in this work as to partial load (Figure 5, top), is still a profit-driven operation in which the electricity price has a major influence as compared to the emissions content in the grid. The load-following profile (Figure 5, bottom) attempts to reduce the grid assistance but still exploits periods of low electricity price to increase the profit. The carbon footprint of the load-following operation decreases by almost a half of the estimated value at the base case, while the profit presents a reduction of 31%.

## 5. Conclusions and outlook

The operation of electrolysis facilities at their full capacity (renewable power plus grid compensation) is the most beneficial operation strategy in terms of profit, and therefore

the most obvious choice to industry nowadays. However, it substantially differs from a renewable-based operation, for instance, load-following operations with a reference renewable power plant.

This work presented an approach to restrict the grid assistance to specific scenarios, namely avoiding unnecessary shutdowns and benefit from periods of low electricity prices or low emissions content in the grid. It is to remark that the demand of hydrogen (model constraint) posed a minimum limit to the grid participation. In the case study, the grid contributed to at least 30% of the total power supply to the AEL facility, while the wind power plant of reference supplied the major part. Although green hydrogen plants are expected to fully operate on renewables, an efficient use of the grid could enhance the process economics while keeping a high renewable content in the product. In this study, the optimization was driven by including a cost penalty to the emissions content in the grid, which can be considered a carbon tax. This is an indirect form to promote the grid use when the renewable content in the grid is high, regardless the specific location, capacity and generation profile of those renewable power facilities. An alternative approach would be to add a constraint that restricts the CO<sub>2</sub> content to a certain level, for instance, to participate in green or low-carbon hydrogen markets.

There is still the opportunity to analyze the effect of hydrogen storage in such systems, as a buffer could reduce the extent of the grid assistance. Yet it is to remark that on-grid operations of PtX facilities should estimate the carbon footprint due to the emissions content in the grid, with its spatial and temporal distribution. In addition, this work suggests that a regulatory framework to support PtX projects is required to favor load-following operations. This can include establishing preferential electricity tariffs, markets for the provision of grid services, or taxes when using grid power at periods with a high emission factor.

## References

- [1] IRENA, *Global hydrogen trade to meet the 1.5°C climate goal: Part III – Green hydrogen cost and potential*. 2022.
- [2] C. Varela, M. Mostafa, and E. Zondervan, “Modeling alkaline water electrolysis for power-to-x applications: A scheduling approach,” *Int. J. Hydrogen Energy*, vol. 46, no. 14, pp. 9303–9313, 2021, doi: 10.1016/j.ijhydene.2020.12.111.
- [3] A. Gitis, M. Leuthold, and D. U. Sauer, “Chapter 4 - Applications and Markets for Grid-Connected Storage Systems,” in *Electrochemical Energy Storage for Renewable Sources and Grid Balancing*, P. T. Moseley and J. Garche, Eds. Amsterdam: Elsevier, 2015, pp. 33–52.
- [4] J. Bartels, C. Varela, T. Wassermann, W. Medjroubi, and E. Zondervan, “Integration of water electrolysis facilities in power grids: A case study in northern Germany,” *Energy Convers. Manag.*, vol. 14, no. March, p. 100209, 2022, doi: 10.1016/j.ecmx.2022.100209.
- [5] Agora Energiewende, “Agorameter,” 2019.
- [6] S. Pfenninger and I. Staffell, “Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data,” *Energy*, vol. 114, pp. 1251–1265, Nov. 2016, doi: 10.1016/j.energy.2016.08.060.
- [7] I. Staffell and S. Pfenninger, “Using bias-corrected reanalysis to simulate current and future wind power output,” *Energy*, vol. 114, pp. 1224–1239, Nov. 2016, doi: 10.1016/j.energy.2016.08.068.
- [8] Gurobi Optimization LLC, “Gurobi Optimizer Reference Manual.” 2022.

# Water sources evaluation for green hydrogen production: a case study in Brazil

Leonardo O. S. Santana,<sup>a</sup> Artur S. Bispo,<sup>a</sup> Gustavo S. Santos,<sup>a</sup> Chrislaine B. Marinho,<sup>a</sup> Ewerton E. S. Calixto,<sup>a</sup> José L. G. Almeida,<sup>a</sup> Fernando L. P. Pessoa<sup>a,b</sup>

<sup>a</sup> *SENAI CIMATEC university center, Avenida Orlando Gomes, Salvador 41650-010, Brazil*

<sup>b</sup> *Federal University of Rio de Janeiro (UFRJ), Av. Athos da Silveira Ramos, Rio de Janeiro 21941-909, Brazil.*  
*leosantana049@gmail.com*

## Abstract

A proper sustainability analysis must include technical, economical, environmental, and social criteria. In this regard, a look at the water consumption in the green hydrogen process is essential to guarantee that the process is indeed sustainable. This work presents an approach to evaluate the suitability of water sources (WS) for hydrogen production via water electrolysis. The approach was applied to a designed 1 MW green hydrogen cluster to be in the future in SENAI-CIMATEC Park in Brazil. Five different WS were evaluated; seawater, treated industrial wastewater, tap water, rainwater, and groundwater. 12 criteria were considered, and economical calculations were performed to evaluate each source. The sum of each measure defined the best and worst water source for the case study. The results have shown that the best WS were tap and rainwater and the worst was seawater, mainly due to the elevation of the hydrogen plant.

**Keywords:** Green hydrogen production, water electrolysis, water treatment, water scarcity, sustainable value.

## 1. Introduction

The production of green hydrogen from the electrolysis of water using renewable energy sources is essential. Beswick, et al. (2021) indicate that in a renewable future the need for hydrogen would be 2.3 Gt per year, where the carbon emissions from the energy sector could be reduced by up to 10,2Gt annually. However, according to Newborough & Cooley (2021), if the green hydrogen replaces all fossil fuels, the water requirement for electrolysis would be near 1x10<sup>14</sup> kg per year (1.8% of present global water consumption). Thus, as a result, water ecosystems can be under significant pressure if not well managed. With this, the scarcity of water, both in quality and quantity, generates the need to assess the adequacy of the sources. (Woods, et al. 2022)

The use of quantitative resources and qualitative factors becomes a priority. Competing with human uses, such as agriculture and urban consumption, requires alternative water sources, such as seawater or industrial wastewater, which have great potential for large-scale production of green hydrogen. The study of these potential WS can lead to a more sustainable one, creating the opportunity for support water security and reducing environmental contamination. (Winter et al., 2022)

Ultrapure water is required for electrolysis, typically water type I or II, as defined by the American society for testing and materials (ASTM) with ionic conductivity < 5µS/cm (ISO 2696) and this water quality must be guaranteed regardless the source of water otherwise water impurities can lead to irreversible damage to the electrolyzer. Thus,

designing a water treatment plant to meet the electrolyzer's requirements is crucial. (Woods, et al. 2022; Simões et al. 2021)

This work presents an approach to assessing the suitability of WS for hydrogen production via water electrolysis, applying a Sustainable Value methodology and an optimization method for aqueous liquid effluent treatment systems to support decision making, combining economic, environmental, and social criteria. The approach is applied to an 1MW green hydrogen cluster on SENAI-CIMATEC Park in Brazil. WS are evaluated for water availability, quality, transport options, abstraction costs, treatment, regulatory needs (including environmental restrictions), and social acceptance. The WS selected were grid water, seawater, rainwater, industrial wastewater (from a near centralized effluent treatment plant), and good quality groundwater (from the São Sebastião aquifer). The necessary treatments were optimally generated using a decision support tool based on heuristic rules for water and effluent treatment, and the water Quality Specification (pre- purification) was established according to the electrolyzer data (according to Directive (EU) 2020/2184- quality of water for human consumption).

## 2. Methodology

The methodology presented herein is based on the multi-criteria approach of sustainable value methodology proposed by Simões (2021). It was developed by integrating different concepts from distinct subjects such as value analysis, ecoefficiency, energy efficiency and cleaner production.

The first step was to identify and mapping all potential WS that can input the electrolyzer which were available around the green hydrogen cluster, and then information about distance and elevation (between the water source and the hydrogen plant) and treatment needs where summarized.

Five potential WS were identified (grid “Tap” water (TW), treated industrial wastewater (IW), sea water (SW), rainwater (RW), and ground water (GW)). The CAPEX and OPEX information were calculated taking in account captation, collection, transport. The table 1 presents investment costs for water captation, storage for each water source, and an overall water loss of 10% was taking in account (due to evaporation, leaks, etc.); water treatment (including installation costs, terrain preparation, etc.) were considered for supplying 700 m<sup>3</sup>/day. The individual treatment technologies where: Reverse Osmosis (RO); Reverse Osmosis (seawater) (RO\*); Ultrafiltration (MF); Fine screening (FS) and Filtration/Coagulation (Chempre), and heuristics from literature were used to create the best treatment train.

**Table 1** – Treatment and captation costs

WS	Treatment train	W. loses	Process CAPEX (€)	Energy consumption kWh/m <sup>3</sup>	Captation (€)	Captation OPEX kW/m <sup>3</sup>
TW	RO	10%	500,000.00	4.5000	0.00	0
IW	FS→MF→RO	25%	776,000.00	4.5635	25,000.00	180.0000
GW	Chempre→RO	20%	580,000.00	4.5500	150,000.00	600.0000
RW	FS→MF	10%	201,000.00	0.0635	-	-
SW	FS→RO*	37%	576,000.00	4.5010	80,000.00	-

For investment costs calculations for water transport, the specific distances and elevation from each potential WS to SENAI CIMATEC Park cluster (as in Table 4) were considered. For water transport via pipelines, purchase costs of welded and screwed pipe per unit length were assumed based in the equation on table 3. The construction of pipeline (considering the different transport distances in Table 2) includes stainless steel piping with an internal diameter of 50 cm.

**Table 2** – Considered distances and elevation for water transport

Potential water source	CIMATEC Park cluster (m)
Ground water	On site
Industrial wastewater	5.247/13
Seawater	4.686/40
Water grid	On site
Rainwater	On site

Water costs were calculated based on the work of Joksimovic (2007) and McGivney & Kawamura (2008) as shown by the equations on table 3.

**Table 3** – Water costs equations

	CAPEX	OPEX
<b>Water transport</b>	$CC = 21,715 \cdot H \cdot Q^{0,52}$ $CP = C_1 \cdot e^{C_2 \cdot D}$ <p>CC is the pumping station capital cost (R\$), H is the required pumping head (m) and Q the design flow rate. CP is the pipe unit cost (R\$/m), D is the diameter (m) and C<sub>i</sub> is the cost coefficients from literature.</p>	$CE = \theta_{hp} \cdot C_e \cdot (V_{ann} \cdot H / 2,7 \cdot \eta)$ <p>CE is the annual cost of energy required for pumping (R\$); <math>\theta_{hp}</math> is a conversion factor to kWh <math>\theta_{hp} = 0.746</math>; <math>C_e</math> is the electricity price [R\$/kWh]; <math>V_{ann}</math> the volume of water pumped annually (m<sup>3</sup>) and <math>\eta</math> the pump efficiency (65 %).</p>
<b>Storage</b>	$UCS = C_1 \cdot V^{C_2}$ <p>UCS is the CAPEX unit cost of storage facility (R\$/m<sup>3</sup>); C<sub>i</sub> is the cost coefficients from literature and V is the storage volume (m<sup>3</sup>).</p>	
<b>Treatment</b>	$CT = Cost_{(1)} \cdot f$ <p>CT is the capex cost for the treatment plant; <math>Cost_{(1)}</math> is the sum of CAPEX cost of individual treatment process and f is the factor that includes other capex cost of treatment plat (f=1,8226 )</p>	$OT_{El} = E \cdot C_e$ $OT_{Total} = OPEX_{El} / 0,275$ <p>OT is the OPEX of treatment plant in function of electricity and the total OPEX; E is the energy consumption [kWh/m<sup>3</sup>] and <math>C_e</math> the electricity price [R\$/kWh]</p>

For this study, values of 10.00 L per kg produced H<sub>2</sub> are considered as required to input for 1MW PEM electrolyser, and 25401,6 kWh per day of electricity consumption. The values are based on information publicly available from electrolyser suppliers.

After cost calculations each of the potential WS (Table 5) was qualitatively assessed for each site adopting a functional value approach where the function is to supply water for hydrogen production. For this, the following criteria were identified: a) Short-term



reliability of availability (effect of weather factors on WS as droughts); b) Long-term reliability of availability which can affect authorization on water use by environmental authorities (perceived future impact of climate change on water source); c) Reliability of supply (possibility for non-weather-related intermittenencies, such as maintenance pauses, that could damage ensuring continuity of water supply); d) Competition with other uses (at water abstraction/collection level); e) Complexity of abstraction/collection (number of involved entities and existence of previous experience with this type of water for a similar use); f) Transport distance from water source to H<sub>2</sub> production plant site; g) Degree of water treatment needed up to electrolyser input requirements; h) Social acceptance; i) Complexity of the permitting process required (number of involved entities and existence of previous experience with this type of water for a similar use, including transport). j) Proportional cost related to CAPEX; k) Proportional cost related to OPEX and l) Electricity consumption. The performance level of the qualitative assessment of water sources for Criteria A to I was taken from Simões, et al (2021), with exception of criteria J, K and L.

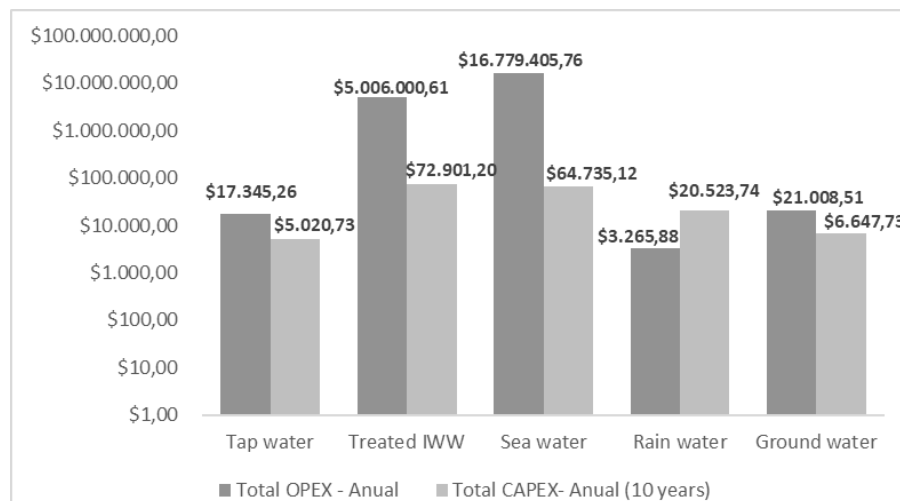
**Table 4** – Criteria for Sustainable Value analysis of water sources.

Criteria	Performance Level			
	1	2	3	4
<b>J</b>	75-100% of H2 plant CAPEX	50-74% of H2 plant CAPEX	25-49% of H2 plant CAPEX	0-24% of H2 plant CAPEX
<b>K</b>	75-100% of H2 plant OPEX	50-74% of H2 plant OPEX	25-49% of H2 plant OPEX	0-24% of H2 plant OPEX
<b>L</b>	75-100% of H2 plant electricity consumption	75-100% of H2 plant electricity consumption	75-100% of H2 plant electricity consumption	75-100% of H2 plant electricity consumption

### 3. Main results

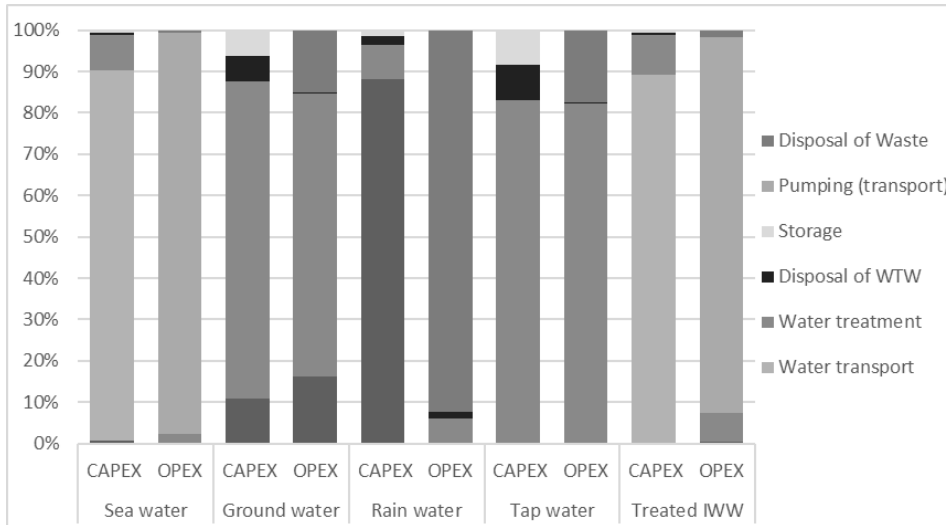
The performed cost of the different water sources shown that the more expensive source is sea water, specially in terms of OPEX, this is regarding the greater distance between the source and the treatment plant. Again because of transporting the treated industrial wastewater is also an expensive source.

**Figure 1** – Water sources CAPEX and OPEX



The analysis of the distributed costs of water presented in Figure 2. shows that costs related to transport (where necessary, for example sea water and treated IWW) are the most relevant in a general context. Another high cost is the cost of water treatment itself.

**Figure 2 – Water sources distributed costs.**

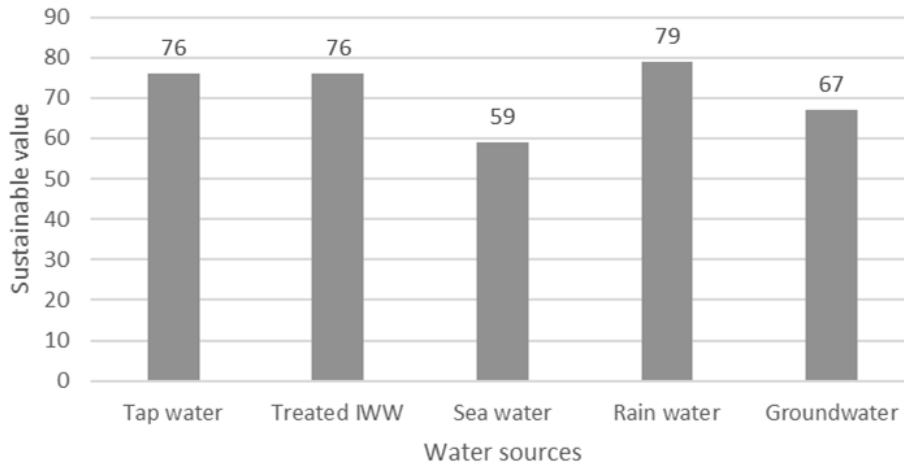


The results of the qualitative criteria are shown in Table 5, in which the grading (1–4 points) follows the criteria presented in Table 3. The best qualitative water source performance was obtained for rainwater followed by treated industrial wastewater and grid water. The water options with lower performance values for the site is sea water.

**Table 5 – Sustainable Value analysis of water sources.**

Weight	Criteria	TW	IW	SW	RW	GW
2	A	3	4	4	1	2
1	B	4	4	4	1	2
2	C	4	4	4	2	2
2	D	1	4	4	4	1
1	E	4	2	1	2	2
2	F	4	2	2	4	4
1	G	4	1	1	3	3
4	H	1	4	4	4	2
1	I	4	3	1	3	2
1	J	4	1	1	4	4
3	K	4	3	1	4	4
4	L	4	3	1	4	4
<b>Total points</b>		76	76	59	79	67

The Rainwater shown to be low costing easy to handle and abundant in the local of study. It's important to take in account that these results represents the best water source for this specific study case, once the performance level are location and capacity dependent, for example, the high proportional capital and operational cost, the high energy consumption and the complexity of collection make the seawater be a bad choice for the study case plant, however, as the hydrogen capacity increases and the hydrogen cluster is constructed near from the sea, the proportional cost decreases, making sea water being a better option.

**Figure 3** – Results of water sources evaluation

#### 4. Conclusion

A sustainable value indicator allows a relative quantitative comparison of the performance of different water sources for electrolysis and the costs involved. It appears that for this case study rainwater is the most suitable source of water for electrolysis due to the lower risk of supply, lower costs, and avoids complex licensing processes. Likewise, grid water and effluent from the effluent treatment plant proved to be possible sources of water, where the factors that most affect suitability are the costs of transporting water and disposing of waste from effluent treatment.

#### References

- M. Newborough & G. Cooley, 2021, Green hydrogen: water use implications and opportunities, Fuel Cells Bulletin, Volume 2021, Issue 12 , p 12-15.
- P. Woods, H. Bustamente, K. Zinsou, 2022, The hydrogen economy- where is the water? Energy nexus, volume 7.
- S. Simoes, J. Catarino, A. Picado, T. Lopes, S. di Bernardino, F. Amorim, F. Gírio, C. Rangel, T. Ponce de Leão, 2021, Water availability and water usage solutions for electrolysis in hydrogen production, Journal of Cleaner Production, 315, p 128124.
- R. Beswick, A. Oliveira, Y. Yan, 2021, Does the green hydrogen economy have a water problem?, ACS energy letters, 6, p 3167-3169.
- D. Joksimovic, 2007. Decision support system for planning of integrated water reuse projects.
- W. McGivney, S. Kawamura, 2008, Cost estimating manual: for water treatment facilities. New Jersey: Hoboken

# Sustainable Development Goals assessment of carbon capture on-board

Valentina Negri, Margarita A. Charalambous, Juan D. Medrano-García,  
Gonzalo Guillén-Gosalbez

*Institute for Chemical and Bioengineering, Department of Chemistry and Applied Biosciences, ETH Zürich, Vladimir-Prelog-Weg 1, Zürich 8093, Switzerland*  
[gonzalo.guillen.gosalbez@chem.ethz.ch](mailto:gonzalo.guillen.gosalbez@chem.ethz.ch)

## Abstract

The shipping industry of cargo containers is a very efficient freight mode, yet it is still entirely reliant on fossil fuels. While sustainable fuels penetrate the market, carbon capture and storage has been proposed as an interim solution for low-carbon shipping. In this work, we assess the technical feasibility, economic and environmental performance of a cargo ship with a retrofitted carbon capture plant on-board that captures 94 % of direct emissions at 85 \$<sub>2019</sub>/tCO<sub>2</sub>. Compared to the current scenario, our solution reduces climate change by 50 %, while the direct air capture technology stands at a 45 % reduction. Our environmental assessment is based on absolute thresholds to quantify the impacts relative to the Earth's carrying capacity and shows that burden-shifting occurs in all the categories but climate change. Finally, the results are analyzed in light of five Sustainable Development Goals using 16 life cycle impact assessment metrics and their associated absolute thresholds.

**Keywords:** container ship, maritime emissions, carbon capture on-board, life cycle assessment, Sustainable Development Goals.

## 1. Introduction

The shipping industry is one of the most efficient transportation modes: it only contributed to approximately 3 % of the global GHG emissions in 2018 (IMO, 2021) while moving 80 % of the total cargo by sea (Istrate et al., 2022). Nonetheless, this sector is still heavily reliant on fossil fuels and its decarbonization was never part of the goals set in the Paris Agreement (Rogelj et al., 2016). Despite the measures adopted to reduce the environmental impact of the shipping industry (IMO, 2022), its emissions are still projected to increase due to population and consequent freight demand growth. Long-term solutions are based on a complete switch of the current fleet to low- or zero-carbon fuels, with a non-negligible change in the infrastructure, i.e., they require compatible engines and bigger fuel storage tanks on-board to compensate for the lower volumetric energy densities. While these sustainable fuels, such as hydrogen, ammonia and methanol, are introduced into the market, interim solutions based on carbon (CO<sub>2</sub>) captured directly at the source (DNV, 2022) or from the air (DAC) could be deployed in parallel.

We conduct a comprehensive techno-economic and environmental assessment of a carbon capture plant retrofitted on-board a cargo ship. While the concept was first proposed by Det Norske Veritas and Process Systems Enterprise in 2013 and recently

evaluated by Feenstra and co-workers (Feenstra, 2019), its absolute environmental footprint remains unclear. Additionally, we compare this solution with the current business-as-usual scenario (BAU) and DAC as an alternative emissions abatement technology.

In this work, life cycle assessment (LCA)-based absolute sustainability methods quantify impacts relative to the planet's carrying capacities using environmental thresholds to interpret the results on a global scale. These methods are rooted in the planetary boundaries (PBs) concept introduced in the work by Rockström et al. (2009), where for the first time, a set of Earth system biophysical limits that should never be surpassed to operate our planet safely were defined. Later, Sala and co-authors (2020) built on the existing literature connecting 16 indicators of the Environmental Footprint (EF) method to five Sustainable Development Goals (SDGs), which is the approach followed here and the focus of this work.

## 2. Methodology

### 2.1. Reference ship

We consider an 8500 twenty-foot equivalent units (TEU) cargo ship powered by a standard internal combustion engine with heavy fuel oil. The reference vessel travels fixed routes and distances between commercial ports at an assumed speed of 26.5 knots. We retrofit a standard carbon capture process on-board the reference ship to reduce the direct emissions considering a week-long trip. We perform a detailed modeling of the system in Aspen HYSYS v11, which includes three main sections, namely the exhaust gas cleaning, CO<sub>2</sub> capture and its liquefaction, as displayed in Figure 1, that are heat-integrated.

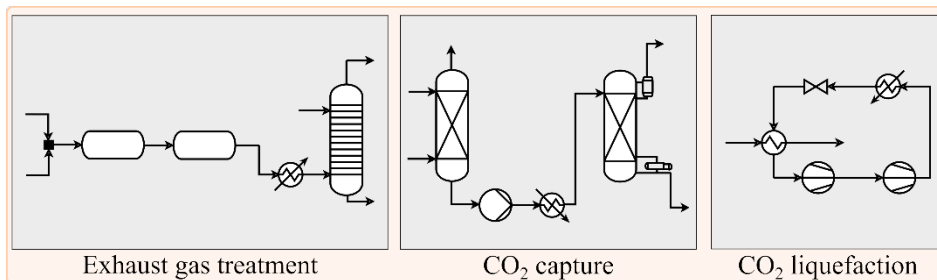


Figure 1. Sketch of the process modeling sections. The exhaust gas treatment, CO<sub>2</sub> capture and CO<sub>2</sub> liquefaction are heat-integrated on-board the vessel.

The pre-treatment section includes a scrubber for the removal of nitrogen and sulfur oxides from the combustion of the heavy fuel oil, as well as a particulate matter abatement technology to comply with the current regulations.

We retrofit a standard post-combustion carbon capture plant on the vessel. Among all the available options, absorption-based CO<sub>2</sub> capture is characterized by the highest technology readiness level, which leads to an easy installation on-board and high efficiency. However, this solution comes at the expenses of considerable space reduction on-board mainly allocated to the absorption and desorption column and the natural gas furnace needed to supply heat to the desorber reboiler. Alternative technologies for CO<sub>2</sub> capture might be considered to overcome these challenges; however, their assessment is not included in this work.

The exhaust gas coming out from the cleaning section is put in counter-current contact with a 30 % wt. monoethanolamine aqueous solution that absorbs the CO<sub>2</sub> and leaves

from the bottom of the first column. The clean gas containing less than 1 % mol. CO<sub>2</sub> is vented to the atmosphere. The CO<sub>2</sub> absorbed by the solvent is stripped in the second column employing heat supplied by the furnace and it is sent to the liquefaction section. Here, it is cooled with natural gas and an ammonia refrigeration cycle to storage conditions (22.0 bar and -16.6 °C) in commercial tanks until it can be unloaded at the port and transported to storage sites via pipelines. We note that the emissions of the furnace are combined with the exhaust gas, so the plant on-board captures the net CO<sub>2</sub> emissions of the system.

Further details about the techno-economic analysis summarized in the sections below can be found in Negri et al. (2022) where we report capital and operating costs of the capture plant, as well as a full LCA study based on Rockström et al. PBs.

### 2.2. Feasibility study

Firstly, we assess the technical feasibility of the configuration proposed to ensure that there is enough space on-board to retrofit the capture plant. In this step, we determine the weight and volume of each piece of equipment required on-board in addition to the ones already usually present on vessels, which are not accounted for. Then, we relate this information to the cargo displaced by the plant. Lastly, we calculate the number of extra ships that will be required to transport the cargo displaced in order to keep the total weight of the ship unchanged.

### 2.3. Economic assessment

The economic assessment is performed in accordance with the correlations in Towler and Sinnott for chemical plants (Towler, 2013), given the large scale of the equipment installed. We calculate the cost of each unit, considering installation and material factors, for the reference year 2019. The electricity required to drive pumps and compressors is given by the generator, while freshwater is supplied directly from the sea using a desalinator at no additional cost.

### 2.4. Absolute environmental assessment

We perform an absolute environmental sustainability assessment following the four phases of the LCA methodology where we quantify the total impact of the system on 16 LCA metrics based on the Environmental Footprint (European Commission, 2013) and updated LANCA (Bos et al., 2018 and Laurentiis et al., 2019) methods connected to the nine PBs and five SDGs. Then, we compare the environmental performance to the threshold values reported by Sala et al., where we updated the human health-related effects according to the population of 2019. The functional unit of the study is the global annual cargo demand (36 trillion tkm in 2019) with a well-to-propulsion scope. The life cycle inventory (LCI) phase is carried out in SimaPro v.9.2.0.2, combining data of the foreground from the process simulation and background system from the Ecoinvent v3.5 database. During the last LCA phase, where the results are interpreted, we analyze the relative impacts of the system in the full safe operating space (SOS %). Therefore, an impact above 100 % implies a transgression of the corresponding PB. By using the full SOS, we avoid allocating a share of the PBs to the container ships industry, which is often a controversial step.

The life cycle impacts are calculated for each flow  $i$  ( $imp_k$ ), based on the characterization factors ( $CF_{s_{ik}}$ ) of each elementary flow ( $E_i$ ) with respect to  $k$  PBs categories. The impact is then compared to the threshold values reported in Sala et al. ( $PB_k$ ) to determine the transgression level ( $tl_k$ ) according to Eq. (1). The LCA inventories include all the activities from the capture to the storage of the CO<sub>2</sub>.

$$tl_k = \frac{imp_k}{PB_k} = \frac{\sum_i CF_{ik} E_i}{PB_k} \quad \forall k \in K \quad (1)$$

The analyses performed in this study and the main findings that are discussed in the following paragraph are summarized in Figure 2.

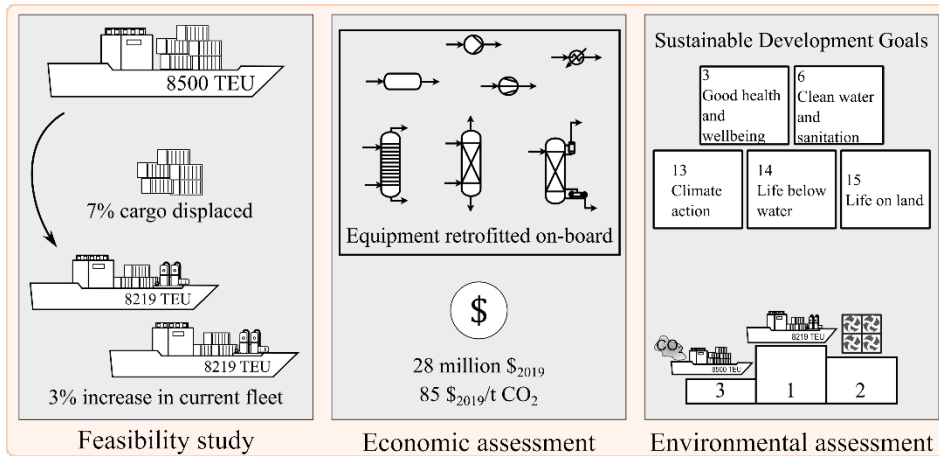


Figure 2. Analyses performed in this study: feasibility study based on the weights of the equipment retrofitted on-board, economic assessment of the units required for the CO<sub>2</sub> capture and absolute environmental assessment linked to the PBs and SDGs.

### 3. Results and discussion

The carbon capture plant on-board cargo vessels is a technically feasible and economically appealing solution to abate direct emissions compared to the current scenario and the alternative emissions reduction technology DAC.

The plant designed has an overall efficiency of 94 %, with a total cost of 85 \$<sub>2019</sub>/tCO<sub>2</sub> captured, without considering the capital expenses for the construction of the additional vessels. Given its lower technology maturity level and the low concentration of CO<sub>2</sub> in the air, DAC is still a more expensive option with estimates at 300-600 \$/tCO<sub>2</sub> captured (Fuss et al., 2018 and DAC, 2021). The total equipment required on-board displaces 7 % of the cargo on a mass-basis for a week-long trip, which implies an increase of 3 % in the current fleet to fulfill the annual cargo demand of 2019. The CO<sub>2</sub> captured during the trip is stored on-board until the vessel has reached the port, where it is unloaded and transported to saline aquifers storage sites.

We assess the environmental performance of the system with respect to the full SOS and we note that none of the categories is transgressed (Figure 3). The impact on climate change [kg CO<sub>2-eq</sub>] is reduced by 50 % in the capture on-board compared to BAU, while DAC reduces it by 45 % due to the capture of direct CO<sub>2</sub> emissions in both scenarios. Although the two alternatives to the BAU contribute to mitigating climate change, significant burden-shifting occurs in the other categories. Eutrophication freshwater, marine and terrestrial are the most affected by burden shifting, where capture on-board performs significantly worse than DAC in the last two. Resource use energy carriers, and minerals and metals also increase compared to BAU, in particular in the DAC scenario due to the additional energy inputs required by the capture process. If we considered

downscaling of the SOS to allocate a portion of it to the shipping industry only, e.g., using the sector GVA (3 % of the global GVA in 2010, OECD 2016), some of the categories such as respiratory inorganics and resource use energy carriers would show a transgression level above 95 %. On a broader perspective, trade-offs within SDG 3, SDG 13 and 14 appear when trying to reduce climate change. Respiratory inorganics, i.e., particulate matter, is the most impactful category in SDG 3 good health and wellbeing, where capture on-board performs worse than the BAU scenario. SDG 13 climate action benefits from the CO<sub>2</sub> emissions reduction, however, resource use energy carriers increases with both capture on-board and DAC. Lastly, SDG 14 life below water is affected by the eutrophication freshwater category, where both capture on-board and DAC perform worse than BAU, with DAC showing the worst performance.

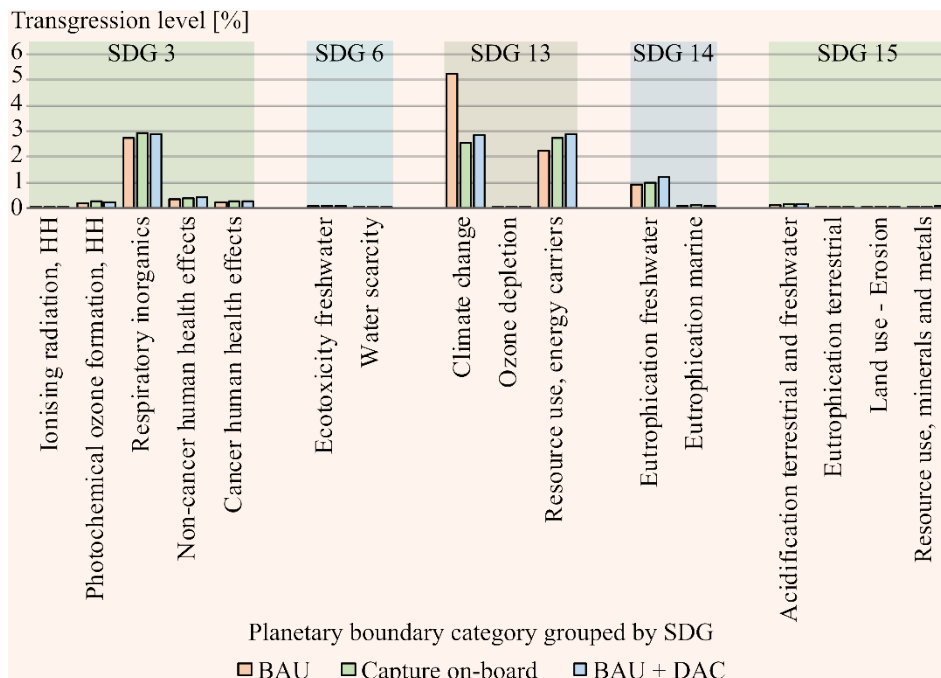


Figure 3. Environmental assessment results: planetary boundaries transgression [%] with respect to the full space and respective Sustainable Development Goals mapping.

#### 4. Conclusions

An interim solution to reduce direct emissions from the cargo shipping industry based on a retrofitted carbon capture plant on-board was proven to be technically feasible, economically competitive and environmentally appealing. The scenario presented is compared to the business as usual where no measures are implemented and direct air capture.

The plant achieves 94% net capture efficiency displacing 7% of the cargo on a mass-basis, which is transported by additional ships corresponding to a 3% increase in the current fleet in 2019. The cost of capture is 85 \$<sub>2019</sub>/tCO<sub>2</sub>, which is considerably cheaper than removing the same amount of CO<sub>2</sub> from the air using direct air capture technologies.

The absolute environmental assessment reveals that none of the categories is transgressed compared to the full safe operating space. While capture on-board reduces climate change



by 50% compared to the business-as-usual, direct air capture achieves a 45 % reduction. However, burden-shifting occurs in all the other categories in both scenarios, although the scenarios assessed occupy only a small percentage of the full safe operating space. In particular, capture on-board performs the worst in eutrophication marine and terrestrial. The resource use categories are also exacerbated compared to the current scenario, particularly using direct air capture.

In order to implement this solution, social and political barriers still exist regarding abatement technologies and should be overcome by putting legal frameworks in place. Additionally, the capital expenses faced by the shipping industry to retrofit the plant on-board and build new vessels that satisfy the demand should be supported by public incentives.

## References

- U. Bos, R. Horn, T. Beck, J. P. Lindner, M. Fischer, 2018, LANCA®-Characterization Factors for Life Cycle Impact Assessment, Version 2.5, Fraunhofer Verlag
- DAC, Direct Air Capture Summit 2021, 2021, <https://www.youtube.com/watch?v=ZoMz7dmfitY&t=1039s>
- DNV, 2022, Maritime Forecast to 2050
- European Commission, 2013, Recommendation 2013/179/EU on the use of common methods to measure and communicate the life cycle environmental performance of products and organisations
- M. Feenstra, J. Monteiro, J. T. van den Akker, M. R. M. Abu-Zahra, E. Gilling; E. Goetheer, 2019, Ship-Based Carbon Capture Onboard of Diesel or LNG-Fuelled Ships, *International Journal of Greenhouse Gas Control*, 85, 1–10
- S. Fuss, W. F. Lamb, M. W. Callaghan, J. Hilaire, F. Creutzig, T. Amann, T. Beringer, W. De Oliveira Garcia, J. Hartmann, T. Khanna, G. Luderer, G. F. Nemet, J. Rogelj, P. Smith, J. V. Vicente, J. Wilcox, M. Del Mar Zamora Dominguez, J. C. Minx, 2018, Negative Emissions - Part 2: Costs, Potentials and Side Effects, *Environmental Research Letter*, 13 (6)
- Á. Galán-Martín, V. Tulus, I. Díaz, C. Pozo, J. Pérez-Ramírez, G. Guillén-Gosálbez, 2021, Sustainability Footprints of a Renewable Carbon Transition for the Petrochemical Sector within Planetary Boundaries, *One Earth*, 4, 565–583
- IMO, Initial IMO Strategy on Reduction of GHG Emissions from Ships, <https://www.imo.org/en/MediaCentre/HotTopics/Pages/Reducing-greenhouse-gas-emissions-from-ships.aspx>
- IMO, 2021, Fourth IMO GHG Study 2020
- I. Istrate, D. Iribarren, J. Dufour, R. Ortiz Cebolla, A. Arrigoni, P. Moretto, F. Dolci, 2022, Quantifying Emissions in the European Maritime Sector
- V. De Laurentiis, M. Secchi, U. Bos, R. Horn, A. Laurent, S. Sala, 2019, *Journal of Cleaner Production*, 215, 63–74
- V. Negri, M. A. Charalambous, J. D. Medrano-García, G. Guillén-Gosálbez, 2022, Navigating within the safe operating space with carbon capture on-board, *ACS Sustainable Chemistry & Engineering*, 10, 51, 17134–17142
- OECD, 2016, *The Ocean Economy in 2030*; OECD Publishing Paris, Ed.
- J. Rockström, W. Steffen, K. Noone, Å. Persson, F. Stuart III Chapin, E. Lambin, T. M. Lenton, M. Scheffer, C. Folke, H. Joachim Schellnhuber, B. Nykvist, C. A. de Wit, T. Hughes, S. van der Leeuw, H. Rodhe, S. Sörlin, P. K. Snyder, R. Costanza, U. Svedin, M. Falkenmark, L. Karlberg, R. W. Corell, V. J. Fabry, J. Hansen, B. Walker, D. Liverman, K. Richardson, P. Crutzen, J. Foley, 2009, Planetary boundaries: Exploring the safe operating space for humanity, *Ecology and Society*, 14
- J. Rogelj, M. Den Elzen, N. Höhne, T. Fransen, H. Fekete, H. Winkler, R. Schaeffer, F. Sha, K. Riahi, M. Meinshausen, 2016, Paris Agreement Climate Proposals Need a Boost to Keep Warming Well below 2 °C, *Nature*, 534, 631–639
- M. W. Ryberg, M. Owsianiak, K. Richardson, M. Z. Hauschild, 2018, Development of a Life-Cycle Impact Assessment Methodology Linked to the Planetary Boundaries Framework, *Ecological Indicators*, 88, 250–262
- S. Sala, E. Crenna, M. Secchi, E. Sany-Mengual, 2020, Environmental sustainability of European production and consumption assessed against planetary boundaries, *Journal of Environmental Management*, 269, 110686
- G. Towler, R. K. Sinnott, 2013, *Chemical Engineering Design*; Elsevier

# Parametric study of the reactive absorption of CO<sub>2</sub> for soda ash production

Maria F. Gutierrez,<sup>a</sup> Peter Schulze,<sup>a</sup> Andreas Seidel-Morgenstern<sup>a</sup>, Heike Lorenz<sup>a</sup>

<sup>a</sup>*Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, 39106 Magdeburg, Germany*

## Abstract

The design of carbon-negative processes replacing traditional energy-intensive processes is desirable to achieve the UN sustainability goals. The CODA project aims to develop a sustainable process to produce soda ash by capturing airborne CO<sub>2</sub> using NaOH followed by a crystallization sequence. This work aimed to analyze the influence of three process variables, namely the Na<sub>2</sub>CO<sub>3</sub> concentration, the temperature and air humidity in the absorber. The CO<sub>2</sub> capture rate, the energy consumption and the water loss in the absorber were considered as process performance variables. High Na<sub>2</sub>CO<sub>3</sub> concentrations result in small capture rates in the absorber, which is directly related with larger absorbers and higher energy consumption for pumping. Absorption is favored at low temperatures because the Na<sub>2</sub>CO<sub>3</sub> saturation concentration is smaller and water loss is reduced. The assessment of two downstream schemes in terms of energy led to the identification of opportunities for further process optimization.

**Keywords:** absorption, carbon dioxide, capture, soda ash.

## 1. Introduction

The development of processes for carbon capture and utilization is one of the pathways that could help to limit global warming. To contribute to this goal, the new CODA (Carbon-negative sODA ash) project has the aim to develop a sustainable process to produce soda ash (FONA, 2022). In the classical Solvay process, the carbonate ion is obtained from lime stone in an energy-intensive reaction that finally generates ~500 kg CO<sub>2</sub>/ton soda. In contrast, the carbonate source in the CODA process is CO<sub>2</sub> captured directly from air using a NaOH solution. Soda ash (from this point mentioned as *soda* in this document) is crystallized from the obtained sodium carbonate solution.

The implementation of the CODA process requires technical studies allowing the design and optimization of a reactive absorber to capture CO<sub>2</sub> using a NaOH solution. Theoretical and experimental studies have reported the average CO<sub>2</sub> flux and potential costs of this operation (Keith et al., 2018; Mazzotti et al., 2013). However, some design variables important for the CODA process have not been studied in detail, e.g. the influence of the Na<sub>2</sub>CO<sub>3</sub> concentration on the absorption performance. The possibility of soda crystallization in the absorber sets new challenges for the equipment design, and the influence of other variables such as weather conditions become worth to study.

Our goal is to study the influence of the Na<sub>2</sub>CO<sub>3</sub> concentration, relative air humidity and temperature on the reactive absorption of CO<sub>2</sub> in an alkaline solution and on the downstream processing of the resulting liquid stream or slurry. To deal with the possible crystallization of soda inside the absorber, we studied the absorption in a column of droplets and without any packing material that could be blocked by crystals (similar to

(Cho et al., 2018)). The flow of droplets and the air are arranged in a co-current mode, in which no fan is required because the falling droplets induce the airflow inside the absorber, in a similar way that in an ejector venturi scrubber (Atay et al., 1987). A rate-based model capable to describe the reactive absorption of CO<sub>2</sub> in a NaOH solution was established using thermodynamic, kinetic and property parameters reported in the literature.

## 2. Mathematical model

### 2.1. Droplet absorber

The absorption of CO<sub>2</sub> in droplets of aqueous NaOH was modeled using the differential equation describing the concentration of CO<sub>2</sub> along the absorber (Eq. 1). The water loss was calculated using the differential equation describing the concentration of water in the gas phase along the absorber (Eq. 2) and the differential equation describing the temperature along the absorber (Eq.3). These equations were obtained from the component mass balance and the energy balance in the gas phase. The energy balance neglects the enthalpy of CO<sub>2</sub> dissolution and enthalpy of reaction because both enthalpies are less than 1% of the evaporation enthalpy. Table 1 presents a list of symbols used for these equations.

$$\frac{dy_{CO_2}}{dz} = -\frac{K_{G,CO_2}^c a}{\dot{G}'} (y_{CO_2} - y_{CO_2}^*) \frac{P}{RT} \quad (1)$$

$$\frac{dy_w}{dz} = \frac{k_{G,w}^c a}{\dot{G}'} (y_w^* - y_w) \frac{P}{RT} \quad (2)$$

$$\frac{dT}{dz} = -\frac{dy_w \Delta H_{vap} M_w}{dz C_{p,G} M_G} \quad (3)$$

Table 1: List of symbols used in equations 1, 2 and 3

Symbol	Meaning	Symbol	Meaning
$y_{CO_2}$	molar concentration of CO <sub>2</sub> in the gas phase	$y_w^*$	molar concentration of water in the gas phase in equilibrium with the liquid
$z$	height of the absorber from top to bottom	$y_w$	molar concentration of water in the gas phase
$K_{G,CO_2}^c$	overall mass transfer coefficient of CO <sub>2</sub> in the gas phase (in m/s)	$k_{G,w}^c$	local mass transfer coefficient of water in the gas phase (in m/s)
$a$	specific surface area (in m <sup>2</sup> /m <sup>3</sup> )	$\Delta H_{vap}$	Enthalpy of vaporization (in kJ/kg)
$\dot{G}'$	molar flow per cross sectional area (in kmol/h/m <sup>2</sup> )	$C_{p,G}$	Heat capacity of the gas phase (in kJ/(kg°C))
$y_{CO_2}^*$	CO <sub>2</sub> gas concentration in equilibrium with the CO <sub>2</sub> liquid bulk concentration	$M_w$	Molecular weight of water
$P$	pressure of the system	$M_G$	Molecular weight of the gas phase
$T$	temperature of the system	$R$	universal gas constant

Following assumptions were made:

- The change of the flow rate of gas due to the absorption and water loss is negligible. Due to the small concentration of CO<sub>2</sub> in the gas phase, if all the CO<sub>2</sub> on the gas stream would be absorbed, the molar flow rate would be reduced approx. by 0.0004%. In the worst-case scenario, the change in the molar flow rate of gas due to water evaporation is around 0.0043%.
- The equilibrium concentration is calculated from the isofugacity equation ( $y_{CO_2}^* P = H_{v,CO_2} c_{CO_2}$ ), which requires the bulk concentration of CO<sub>2</sub> in the liquid phase (calculated from the mass balance in the liquid phase considering the flux of CO<sub>2</sub> into the liquid phase and the amount of CO<sub>2</sub> reacting with NaOH).

- The equilibrium concentration  $y_w^*$  is calculated from the isofugacity equation ( $y_w^*P = a_w p_s$ ), which requires the activity of water ( $a_w$ ) and the vapor pressure of pure water ( $p_s$ ), both functions of the temperature.
- The reaction is irreversible because the hydroxide ions concentration is always much higher than the concentration of the dissolved CO<sub>2</sub> (Danckwerts, 1970).
- The droplet diameter, number of droplets and gas and liquid velocities are constant along the absorber. No coalescence is considered. Results of the model are somehow ideal in this sense.

The overall mass transfer coefficient was calculated using the two-resistance theory, which requires to compute the local mass transfer coefficients in the liquid and in the gas phase (see Eq. 4). The Enhancement factor ( $E$ ) was used to calculate the liquid mass transfer coefficient due to the reaction taking place in the liquid phase ( $CO_2 + 2NaOH \rightarrow Na_2CO_3 + H_2O$ ). This factor is a function of the Hatta number ( $Ha$  in Eq. 5) and depends on how fast is the reaction relative to the mass transfer. Our previous studies showed that the droplet absorber remains in the intermediate-fast reaction regime, where the penetration theory can be applied (Danckwerts, 1970). The equation to calculate  $E$  is reported in literature (van Swaaij & Versteeg, 1992).

$$\frac{1}{K_{G,CO_2}^c} = \frac{1}{k_{G,CO_2}^c} + \frac{H_{v,CO_2}}{E k_{L,CO_2}^c RT} \quad (4)$$

$$Ha = \frac{\sqrt{k' D_{CO_2,L}^m}}{k_{L,CO_2}^c} \quad (5)$$

Local mass transfer coefficients were calculated with Eq. 6 and Eq. 7 for the gas and liquid phase, respectively. The equation for the gas phase is used for convective flow surrounding spheres, while the equation for the liquid phase is based on the penetration theory. The Henry volatility constant ( $H_{v,CO_2}$ ) and the second order kinetic constant ( $k'$ ) were calculated as a function of the solution ionic strength (Gondal, 2014). For the diffusivity of CO<sub>2</sub> in the liquid mixture ( $D_{CO_2}^m$ ), a function of the viscosity. (Kucka et al., 2002) was used. The specific surface area was calculated with Eq. 8 (Dimiccoli et al., 2000). In this equation, the liquid volumetric flow rate ( $\dot{q}_L$ ), the absorber diameter ( $D$ ), the droplet diameter ( $d$ ) and the liquid velocity ( $v_L$ ) are required. Geometric, hydrodynamic and some operational variables of the absorber were constant in the parametric study (Table 2). Some of these variables were fixed based on our previous experimental studies. The differential-algebraic equation system was solved with Python.

$$k_{G,CO_2}^c = \frac{D_{CO_2}^m}{d} (2 + 0.6Re^{0.5}Sc^{1/3}) \quad (6)$$

$$k_{L,CO_2}^c = 2 \sqrt{\frac{D_{CO_2,L}^m v_L}{\pi d}} \quad (7)$$

$$a = \frac{24\dot{q}_L}{\pi d D^2 v_L} \quad (8)$$

Table 2: Variables kept constant in the parametric study

Absorber variable	Units	Value	Absorber variable	Units	Value
Absorber diameter	mm	28.4	Pressure	kPa	101.325
Absorber height	m	1.4	Concentration of NaOH	wt.%	5
Liquid velocity	m/s	2.38	Inlet concentration of CO <sub>2</sub>	ppm	400
Gas velocity	m/s	2.56	Gas molar flow ( $\dot{G}'$ )	kmol/s/m <sup>2</sup>	0.109
Droplet diameter	μm	333	Liquid volumetric flow rate	mL/min	922.4
Specific surface area	m <sup>2</sup> /m <sup>3</sup>	182.7	Pressure drop of liquid	mbar	129

## 2.2. Downstream processes

The solution out of the absorber should be processed to obtain anhydrous sodium carbonate ( $\text{Na}_2\text{CO}_3$ ). Different processing strategies can be used depending on the final concentrations obtained in the absorber. Here, two strategies were identified and evaluated. In the first strategy, the absorption is carried out until sodium hydroxide is completely consumed and  $\text{Na}_2\text{CO}_3$  concentration is below saturation (because the sodium hydroxide inlet concentration is set to 5 wt.% and the corresponding soda concentration is below the saturation condition). Then, anhydrous soda is obtained from an evaporative crystallization, using mechanical vapor recompression (MVR) for heat integration and reducing the primary energy demand.

In the second strategy, the absorption is carried out using a NaOH solution saturated with  $\text{Na}_2\text{CO}_3$ , in the metastable zone where no crystals are formed. Then, a crystallization of decahydrate ( $\text{Na}_2\text{CO}_3 \cdot 10\text{H}_2\text{O}$ ) at ambient conditions will follow. This scheme implies that the liquid out of the crystallizer (saturated in  $\text{Na}_2\text{CO}_3$ ) is recycled to the absorber. In this way, most of the water in the solution is eliminated from the product by crystallization (without evaporation), and only 10 moles of water remain with the product. Finally, decahydrate crystals undergo melting, followed by evaporative crystallization with MVR to obtain anhydrous soda ash.

## 3. Parametric study

Three parameters were changed in this study of reactive absorption of  $\text{CO}_2$  with NaOH:  $\text{Na}_2\text{CO}_3$  concentration in the inlet of the absorber, ambient temperature at the inlet and relative humidity of air.

Aiming to evaluate the performance of the absorber, the specific capture rate was calculated along the absorber using Eq. 9. This value expresses the flow of  $\text{CO}_2$  captured (kg/h) in one cubic meter of absorber. The smaller the capture rate, the bigger the absorber should be to achieve a certain capture demand. Based on our previous studies, the pump energy was related to the capture rate to obtain the energy consumption in the absorber (in kWh/kg  $\text{CO}_2$  captured). Using a carbon dioxide equivalency of 0.1237 kg of  $\text{CO}_2$  emitted per kWh electricity consumed (maximal value for onshore wind power (Amponsah et al., 2014)); the energy consumption was transformed into kg  $\text{CO}_2$  equivalent/kg of  $\text{CO}_2$  captured. Values of the capture rate reported and used here after correspond to the arithmetic average along the absorber (valid due to low  $y_{\text{CO}_2}$ ).

$$r = \frac{K_G^c a y_{\text{CO}_2} P}{RT} \quad (9)$$

Some water is also evaporated in the absorber depending on the conditions of the inlet air. The saturation concentration of water in the air was calculated with the isofugacity equation, which requires the activity of water in the liquid (with ions) and the vapor pressure of pure water at the system temperature. The water loss per mass of  $\text{CO}_2$  captured was calculated with Eq. 10.

$$W_{\text{loss}} = \frac{\dot{G}(y_w^{\text{out}} - y_w^{\text{in}})}{rV_{\text{abs}}} \quad (10)$$

The power consumed by the MVR compressors in each strategy was calculated using the amount of water needed to be evaporated and the enthalpy difference for the vapor in the MVR blower (compression ratio of 1.5). This energy was related with the amount of  $\text{CO}_2$  capture by stoichiometry (1 mol of  $\text{CO}_2$  should be captured to obtain 1 mol of  $\text{Na}_2\text{CO}_3$ ).

## 4. Results and Discussion

The capture rate and emission of  $\text{CO}_2$  in the absorber due to pumping are presented in Figure 1 as a function of the system outlet temperature,  $\text{Na}_2\text{CO}_3$  inlet concentration and

inlet air humidity. The highest Na<sub>2</sub>CO<sub>3</sub> concentration on each temperature and humidity case changes because it corresponds to the saturation concentration. The decrease in the capture rate when the Na<sub>2</sub>CO<sub>3</sub> concentration increases can be related with the decrease of CO<sub>2</sub> solubility in the liquid due to the presence of the carbonate ions in the solution (salting-out effect). Results show that the absorption using a liquid feed saturated in Na<sub>2</sub>CO<sub>3</sub> is favored at lower temperatures because the Na<sub>2</sub>CO<sub>3</sub> saturation concentration is smaller. The air humidity does not affect considerably the capture rate in the low temperature range. On the other hand, the CO<sub>2</sub> emission in the absorber is heavily affected (Figure 1b is in logarithmic scale) by the temperature and Na<sub>2</sub>CO<sub>3</sub> concentration in the absorber. According to our results, the maximal emission of this absorption process occurs at 30°C when the liquid is almost saturated with Na<sub>2</sub>CO<sub>3</sub> (smallest capture rate).

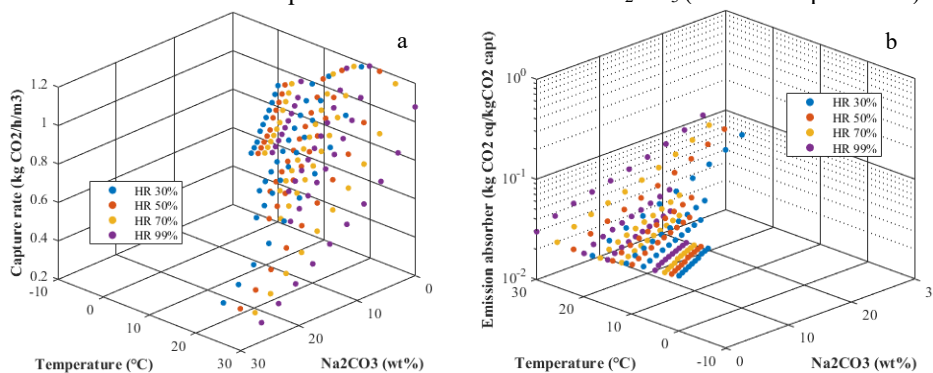


Figure 1: Results of the parametric study in the absorber (5 wt% NaOH) regarding a) CO<sub>2</sub> capture rate and b) emission of CO<sub>2</sub> due to energy consumption in the absorber.

The water loss in absorber is heavily affected by relative humidity and air (see Figure 2a). Processes at lower temperature and high humidity present the smallest water loss (~0.00024 kg water/kg CO<sub>2</sub> captured). Since the water used in the process should be highly pure (to achieve pure crystals), it is recommended to include an upstream humidifying tower, in which the air can be saturated with low grade water (cheaper) prior to the absorption (at dry and warm weather conditions).

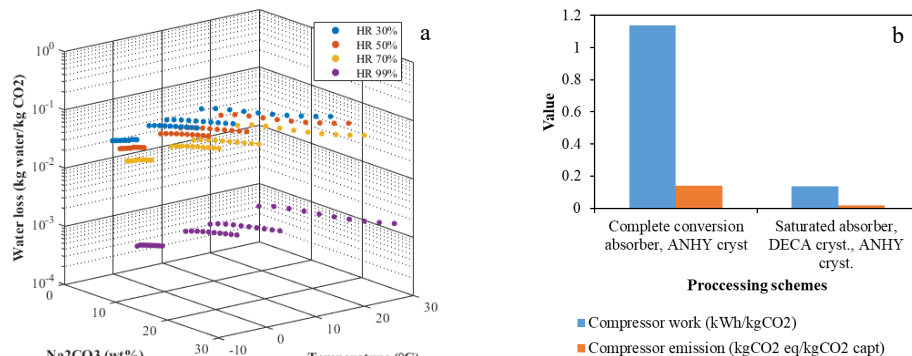


Figure 2: Results of the parametric study in the absorber (5 wt% NaOH) regarding a) water loss in the absorber and b) emission of CO<sub>2</sub> due to energy consumption in the downstream process.

Analysis of the downstream process shows in Figure 2b that the second strategy has a smaller energy consumption compared to the first strategy. Even though both downstream processing schemes are carbon-negative, the second strategy avoids the emission of around 0.123 kg CO<sub>2</sub> per kg of CO<sub>2</sub> captured.

However, the second strategy also implies to operate at smaller capture rates resulting in bigger absorbers and higher energy consumption in the absorber. An opportunity for process synthesis by optimization is evidenced from our results.

## 5. Conclusion and Outlook

The reactive absorption of CO<sub>2</sub> from air with carbonated NaOH has been analyzed using a parametric study in which the Na<sub>2</sub>CO<sub>3</sub> concentration and the temperature and humidity of air were varied. The effect of these parameter on the capture rate, water loss and energy consumption in the absorber were evaluated. In addition, two downstream processing schemes to obtain anhydrous soda were assessed in terms of energy consumption. The presence of high concentration of Na<sub>2</sub>CO<sub>3</sub> in the absorber reduces considerably the capture rate; but at the same time, this is required in the downstream process with the lowest energy consumption. The results also showed that absorption is favored at low temperatures, where the saturation concentration is smaller (the capture rate is higher) and the water loss is minimized. Under dry and warm weather scenarios, a pre-humidification process should be added to avoid the loss of high-quality (distilled) water. Analysis of the energy consumption in terms of CO<sub>2</sub> emitted per CO<sub>2</sub> captured allowed to identify scenarios in which the operation of absorber emits around 10% of the CO<sub>2</sub> captured. The results of this study will be used to optimize the conceptual design of the CODA process.

## References

- Amponsah, N.Y., Troldborg, M., Kington, B., Aalders, I., Hough, R.L., 2014. Greenhouse gas emissions from renewable energy sources: A review of lifecycle considerations. *Renewable and Sustainable Energy Reviews* 39, 461-475.
- Atay, L., Lewandowski, G., Trattner, R., 1987. Fluid Flow and Gas Absorption in an Ejector Venturi Scrubber. *Environmental Progress* 6, 198-203.
- Cho, M., Lee, S., Choi, M., & Lee, J. W. (2018). Novel Spray Tower for CO<sub>2</sub> Capture Using Uniform Spray of Monosized Absorbent Droplets. *Industrial and Engineering Chemistry Research*, 57(8), 3065-3075
- Danckwerts, P. v., 1970. *Gas-Liquid Reactions*. McGraw-Hill, New-York.
- Dimiccoli, A., di Serio, M., Santacesaria, E., 2000. Mass transfer and kinetics in spray-tower-loop absorbers and reactors. *Ind Eng Chem Res* 39, 4082-4093.
- FONA, 2022. CODA – Development of an environmentally friendly process for the production of soda. URL <https://www.fona.de/en/measures/funding-measures/KlimPro/coda.php>
- Gondal, S., 2014. Carbon dioxide absorption into hydroxide and carbonate systems (Ph. D. Thesis). NTNU-Trondheim, Trondheim.
- Keith, D.W., Holmes, G., Angelo, D. st., Heidel, K., 2018. A Process for Capturing CO<sub>2</sub> from the Atmosphere. *Joule* 2, 1573-1594.
- Kucka, L., Kenig, E.Y., Górak, A., 2002. Kinetics of the gas-liquid reaction between carbon dioxide and hydroxide ions. *Ind Eng Chem Res* 41, 5952-5957.
- Mazzotti, M., Baciocchi, R., Desmond, M.J., Socolow, R.H., 2013. Direct air capture of CO<sub>2</sub> with chemicals: Optimization of a two-loop hydroxide carbonate system using a countercurrent air-liquid contactor. *Clim Change* 118, 119-135.
- Stolaroff, J.K., Keith, D.W., Lowry, G. v., 2008. Carbon dioxide capture from atmospheric air using sodium hydroxide spray. *Environ Sci Technol* 42, 2728-2735.
- van Swaaij, W.P.M., Versteeg, G.F., 1992. Mass Transfer Accompanied With Complex Reversible Chemical Reactions In Gas-Liquid Systems: An Overview, *Chemical Engineering Science* 47, 3181-3195.

# Machine learning enabled modelling and sensitivity analysis for the power generation from a 660 MW supercritical coal power plant

Waqar Muhammad Ashraf, Vivek Dua\*

*Sargent centre for Process Systems Engineering, Department of Chemical Engineering,  
University College London, Torrington Place, London WC1E 7JE, UK*

\* Corresponding Author: [v.dua@ucl.ac.uk](mailto:v.dua@ucl.ac.uk)

## Abstract

The modelling of large-scale industrial power generation system is a challenging task owing to the hyperdimensional input space of the variables and the non-linear interactions among them. In this work, the power production operation from 50% to 100% generation capacity of a 660 MW supercritical coal power plant is modeled on almost three months' operational data by artificial neural network (ANN). The hyperparameters of the ANN model are optimized, an effective ANN model is developed and validated on the power generation conditions. The partial derivative-based sensitivity analysis is carried out and it reveals that main steam flow rate is the most significant input variable on the power production followed by coal flow rate, reheat steam temperature and main stem temperature. This research work presents a reliable utilization of ANN model for the modelling of large capacity power plants that can be extended to conduct the enterprise-level performance enhancement analytics.

**Keywords:** power generation modelling; machine learning; net-zero; coal power plants

## 1.0 Introduction

The power generation from a large-scale industrial complex is maintained under the synchronized operation of several devices and systems having associated non-linearities and interactions. The development of first principles to model the power generation under various operational constraints and state of equipment availability is a challenging task, and model-based optimization can be computationally prohibitive. Artificial neural network is a powerful modelling algorithm that is widely used in the area of machine learning, and it can effectively approximate nonlinear function constructed on hyperdimensional input space with reasonable accuracy. Moreover, ANN can build an effective functional map between the input and output variables of the given system, and thus, it can be deployed to model the complex power generation operation of the coal power plants.

Researchers have developed machine learning models for modelling the power generation operation from the coal-based power plants. A comprehensive review of the application of various machine learning models for fuel combustion to generate power is presented by (Mohammadi, Immonen et al. 2022). (Blackburn, Tuttle et al. 2022) presented the study on the machine learning assisted optimization of the coal power plant using nine input variables. An optimized model for the improved coal combustion in the boiler with respect to power generation and quality of coal is presented by (Li and Yao



2017). In another work, (Safdarnejad, Tuttle et al. 2019) presented the recurrent neural network assisted dynamic modelling and optimization of a coal-based furnace for the minimization of  $\text{NO}_x$  and CO emissions.

In the research studies reported in literature, many operating variables are deployed to model coal power plant's operation that can be reduced by considering the operational relevance and variables' independence. The application of ANN for modelling the power generation from a 660 MW power plant under reduced input space has not been fully explored. Moreover, the sensitivity of the input variables on the power generation is also required to be investigated that is potentially lacking in literature and can be of interest to the industrial community operating the coal-based power generation systems.

In this work, data from a 660 MW supercritical coal power plant is utilized to model the complex and multivariable dependent power generation operation under reduced but relevant input variables. Furthermore, the partial-derivative based sensitivity analysis of the input variables is performed to evaluate their significance on the power production. This research work demonstrates the reliable utilization of machine learning algorithm for the industrial relevance and competitiveness that can contribute to smart operation management of industrial complexes supporting the net-zero goal.

## 2.0 Variables identification, data-collection, visualization and pre-processing

A generic schematic diagram of a 660 MW coal power plant operation is presented in Figure 1, whereas the detailed description on the schematic is provided in (Muhammad Ashraf, Moeen Uddin et al. 2020). Four critically controlled and operational-relevant input variables, i.e., coal flow rate ( $\dot{m}_f$ ), main steam temperature (MST), main steam flow rate (MSF) and reheat steam temperature (RHT), which are backed by domain-knowledge and literature survey (Tunckaya and Koklukaya 2015, Kumar, Nikam et al. 2020), are identified to model the power generation from a 660 MW

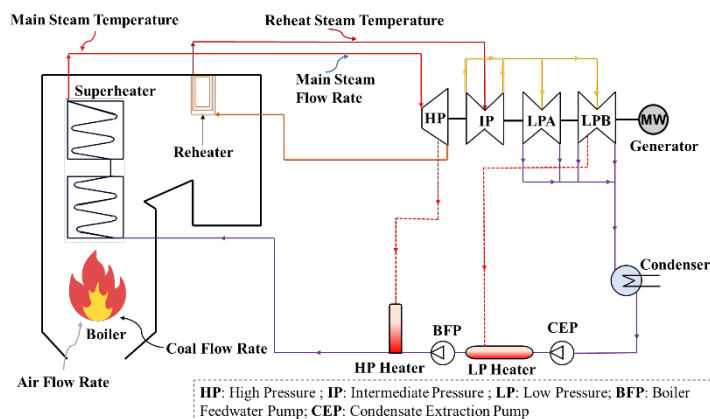


Figure 1. Schematic diagram of a coal power plant operation.

supercritical coal power plant. The selected input variables are strictly controlled within their operating ranges during ramp-up and ramp-down operation of the power plant. It is

important to mention here that the sustained power production from the power plant contributes to the frequency control of the nearby connected grid thereby security and safety of the grid is critical to ensure the smooth power production and transmission. Almost 2000 hourly averaged observations (i.e., almost three months) from 50 % to 100 % power generation capacity of the power plant are taken for all operating variables. The visualization of the input and output data-distribution space for the variables is presented in Figure 2. Continuous data-distribution of the input variables, i.e.,  $\dot{m}_f$ , MST, MSF, RHT as well as output variable (Power) can be observed for their operating ranges. The data-distribution of the variables is reasonably good that possesses information about the operating values for the variables under different power generation capacity and is essentially beneficial for the development of a well-performing machine learning model.

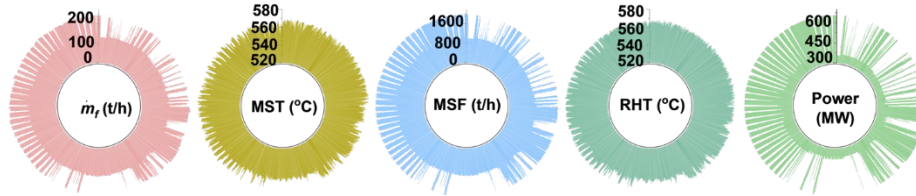


Figure 2. Data-visualization of the operation control parameters;  $\dot{m}_f$ , MST, MSF, RHT and power. Good data-distribution space is observed for the selected variables.

The operating ranges of the variables are significantly different which can affect the generalization and prediction efficacy of the machine learning model. Therefore, the data of all variables is transformed into an equal range. The mathematical expression for the data-transformation into [a,b] scale is given as:

$$X'_i = \frac{(X_i - X_{min})(b-a)}{X_{max} - X_{min}} + a \quad (1)$$

Here,  $X_i$  is the value of the original variable that is transformed into  $X'_i$ .  $X_{min}$  and  $X_{max}$  are the minimum and maximum value of  $X$ ;  $a$  and  $b$  are taken as -1 and 1 respectively that represents the scale of the data transformation for the variables.

### 3.0 Development of Artificial Neural Network Model

Artificial neural network, commonly known as multi-layered perceptron, is one amongst the powerful modelling algorithms widely deployed to approximate the non-linear functions constructed on high-dimensional input space (Gueddar and Dua 2012). The working of ANN is inspired by the functioning of human brain and the details can be found in the reported work (Muhammad Ashraf, Moeen Uddin et al. 2020, Ashraf, Uddin et al. 2021).

In this work, three layered shallow ANN architecture is constructed. The number of neurons in the hidden layer are varied from 4 to 10 (1 ~ 2.5 times of input layer neurons) to find an optimal configuration of ANN network. Tangent hyperbolic and linear activation function is applied on the hidden and output layer respectively. Levenberg-Marquardt algorithm is used for the parametric optimization of ANN architecture. The performance metrics built on coefficient of determination ( $R^2$ ), mean absolute error (MAE) and root mean square error (RMSE) are deployed for evaluating the training performance of the ANN. The mathematical expression of the performance metrics adopted in this work is given as:

$$R^2 = 1 - \frac{\sum_i^N (y_i - \hat{y}_i)^2}{\sum_i^N (y_i - \bar{y}_i)^2}, \quad 0 \leq R^2 \leq 1 \quad (2)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (3)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad (4)$$

here,  $y_i$ ,  $\hat{y}_i$  and  $\bar{y}_i$  are the actual value, model projected value, and mean of actual values, with  $i = 1, 2, 3, \dots, N$  being the number of observations.

Figure 3 shows the pictorial representation of the trained ANNs with different number of hidden layer neurons (from 4 to 10). The performance matrix constructed on  $R^2$ , MAE and RMSE is computed for the developed networks. The performance metrics corresponding to the trained ANNs are compared to find the ANN model having optimal hidden layer neurons. It is found that ANN built on nine neurons in the hidden layer has comparatively better values for performance metrics with  $R^2 = 0.99$ , MAE = 3.22 MW and RMSE = 4.43 MW. Therefore, ANN model with nine neurons in the hidden layer is deployed to undergo external validation test as described in the next section.

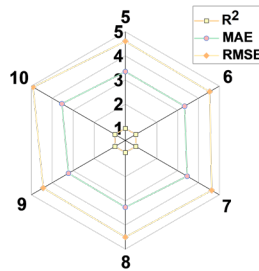


Figure 3. The training performance of ANN with different number of hidden layer neuron (5,6,...,10) is measured on  $R^2$ , MAE, and RMSE. ANN with nine neurons in the hidden layer has demonstrated comparatively better performance.

#### 4.0 Validation of the developed ANN model

The validation of the developed ANN model is a crucial step towards the applicability of the machine learning algorithms in the large-scale industrial complexes. In this regard, the external validation test is performed to check the efficacy of the developed ANN model towards the prediction of unseen operating conditions of the industrial complexes. Therefore, the developed ANN model is deployed for the external validation test to evaluate its generalization and modelling performance for predicting the power generation from the power plant. The external validation dataset consists of 134 randomly selected observations for the operating variables corresponding to 50% to 100% power generation capacity of the power plant. Subsequently, the external validation dataset is deployed to be predicted from the developed ANN model. The performance metrics of the ANN network for the external validation dataset is calculated:  $R^2 = 0.99$ , MAE = 2.35 MW and RMSE = 3.08 MW which are comparable with the results reported in literature (Ashraf, Uddin et al. 2022). The performance of the developed ANN network in the external validation test confirms the good generalization and modelling ability of the model, and thus the operational insights about the power plant operation using the ANN model can be developed.

#### 5.0 Partial-derivative based sensitivity analysis of ANN model

Partial-derivative based sensitivity analysis technique is utilized to evaluate the significance of the input variables on the power generation. The partial derivative-based sensitivity analysis computes the partial derivative of the output variable with respect to the input variable at each sample of the dataset. Thus, the explicit expression for the sensitivity analysis of the ANN model can provide robust diagnostic information about the variable significance on the output variable.

The mathematical expression for partial derivative-based sensitivity of input variable ( $X_p$ ) on the output variable ( $Y$ ) for the ANN model is given as (Nourani and Fard 2012):

$$\frac{\partial Y}{\partial X_p} = \sum_{h=1}^n W_2 (1 - M^2) W_1 \quad (5)$$

here,  $W_2$  is the weight matrix built on the weight connections from the hidden layer to output layer neuron;  $W_1$  is the weight connections from the particular input variable ( $X_p$ ) to the hidden layer neurons;  $(1 - M^2)$  is the partial derivative of the activation function applied on the hidden layer with respect to the summation computed at the hidden layer neuron.

The variance produced in the output variable with respect to the input variable is calculated using Eq. 5 and the procedure is repeated for the remaining input variables. Subsequently, the variance computed for all the input variables is normalized to account for the percentage significance of the input variable towards the prediction of output variable. Higher the value of percentage significance is for an input variable, the more significant the input variable is towards the output variable and vice versa. The mathematical expression of percentage significance is given as:

$$\text{Percentage Significance} = \frac{\sigma_{y_i|x_i}^2}{\sum_{i=1}^d \sigma_{y_i|x_i}^2} 100 \quad (6)$$

here,  $\sigma_{y_i|x_i}^2$  refers to the variance produced in output variable  $y_i$  with respect to the input variable  $x_i$ ;  $\sum_{i=1}^d \sigma_{y_i|x_i}^2$  is the summation of the variance produced in the output variable with respect to input variables;  $i = 1, \dots, d$  equals to number of input variables.

Figure 4 shows the percentage significance of the input variables on the power production from the coal power plant. MSF turns out to be the most significant variable contributing the percentage significance of 67. The second most sensitive input variable towards the power generation is  $\dot{m}_f$  followed by RHT and MST having a percentage significance value of 23.5, 7.05 and 2.34 respectively. The order of significance of the input variables on the power production is similar to the reported research in literature in

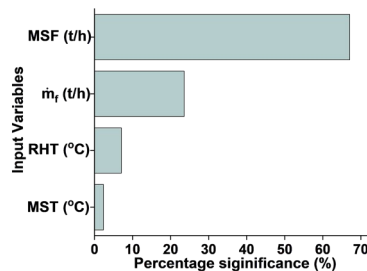


Figure 4. Percentage significance of the input variables for the power generation operation of coal power plant.

which response surface methodology was conducted to evaluate the variables' significance (Ashraf, Uddin et al. 2022). Furthermore, the significance order of the variables on the power production is also explainable as per the operational knowledge of the power plant.

## 6.0 Conclusion

In this work, ANN model is trained on the data taken from a 660 MW supercritical coal power plant to model the power production from 50% to 100% generation capacity. ANN model having nine neurons in the hidden layer is developed after rigorous hyperparameters tuning and the model has presented excellent performance metrics in the external validation test ( $R^2 = 0.99$ , MAE = 2.35 MW and RMSE = 3.08 MW). The partial derivative-based sensitivity analysis of ANN model is performed, and it is found that MSF is the most significant input variable towards the power production having a percentage significance value of 67. Whereas, percentage significance value for  $\dot{m}_f$ , RHT and MST is 23.5, 7.05 and 2.34 respectively. The effective utilization of ANN model for the complex and large-scale industrial power generation operations is presented in this study and the enterprise-level performance enhancement using machine learning based analytics would be investigated in the future studies to quantify the energy efficiency improvement and reduction in emissions footprint from the energy sector.

## Reference

- Ashraf, W. M., G. M. Uddin, H. A. Ahmad, M. A. Jamil, R. Tariq, M. W. Shahzad and V. Dua (2022). "Artificial intelligence enabled efficient power generation and emissions reduction underpinning net-zero goal from the coal-based power plants." *Energy Conversion and Management* **268**: 116025.
- Ashraf, W. M., G. M. Uddin, S. M. Arafat, J. Krzywanski and W. Xiaonan (2021). "Strategic-level performance enhancement of a 660 MWe supercritical power plant and emissions reduction by AI approach." *Energy Conversion and Management* **250**: 114913.
- Blackburn, L. D., J. F. Tuttle, K. Andersson, A. Fry and K. M. Powell (2022). "Development of novel dynamic machine learning-based optimization of a coal-fired power plant." *Computers & Chemical Engineering* **163**: 107848.
- Gueddar, T. and V. Dua (2012). "Novel model reduction techniques for refinery-wide energy optimisation." *Applied energy* **89**(1): 117-126.
- Kumar, R., K. Nikam and R. Jilte (2020). A simulation model to predict coal-fired power plant production rate using artificial neural network tool. *Applied Computer Vision and Image Processing*, Springer: 150-160.
- Li, Q. and G. Yao (2017). "Improved coal combustion optimization model based on load balance and coal qualities." *Energy* **132**: 204-212.
- Mohammadi, K., J. Immonen, L. D. Blackburn, J. F. Tuttle, K. Andersson and K. M. Powell (2022). "A review on the application of machine learning for combustion in power generation applications." *Reviews in Chemical Engineering*.
- Muhammad Ashraf, W., G. Moeen Uddin, A. Hassan Kamal, M. Haider Khan, A. A. Khan, H. Afroze Ahmad, F. Ahmed, N. Hafeez, R. Muhammad Zawar Sami and S. Muhammad Arafat (2020). "Optimization of a 660 MWe supercritical power plant performance—a case of Industry 4.0 in the data-driven operational management. Part 2. Power generation." *Energies* **13**(21): 5619.
- Nourani, V. and M. S. Fard (2012). "Sensitivity analysis of the artificial neural network outputs in simulation of the evaporation process at different climatologic regimes." *Advances in Engineering Software* **47**(1): 127-146.
- Safdarnejad, S. M., J. F. Tuttle and K. M. Powell (2019). "Dynamic modeling and optimization of a coal-fired utility boiler to forecast and minimize NOx and CO emissions simultaneously." *Computers & Chemical Engineering* **124**: 62-79.
- Tunckaya, Y. and E. Koklukaya (2015). "Comparative prediction analysis of 600 MWe coal-fired power plant production rate using statistical and neural-based models." *Journal of the Energy Institute* **88**(1): 11-18.

## Hydrogen production using renewable energy: solar PV and offshore wind power – An economic evaluation in Bahia.

Gustavo de S. dos Santos<sup>a</sup>, Chrislaine do B. Marinho<sup>a</sup>, Leonardo O. S. Santana<sup>a</sup>,  
Artur S. Bispo<sup>a</sup>, Fernando L. P. Pessoa<sup>a,b</sup>, José L. G. Almeida<sup>a</sup>, Ewerton E. S.  
Calixto<sup>a</sup>

<sup>a</sup> *Centro Universitário SENAI CIMATEC, Avenida Orlando Gomes, Salvador 41650-010, Brazil*

<sup>b</sup> *Universidade Federal do Rio de Janeiro, Av. Athos da Silveira Ramos, Rio de Janeiro-RJ 21941-909, Brazil*  
*gustavo.eqrj@gmail.com*

### Abstract

Green hydrogen is an interesting and sustainable alternative capable of diversifying the energy grid and reducing the dependency on fossil fuels, contributing to decarbonizing the global economy. Despite being widely applied on an industrial scale, hydrogen still faces major challenges, mainly due to the high cost incurred in its production, storage, transport, and the electricity required. Thus, the objective of this study was to evaluate the levelized cost of Hydrogen (LCOH) of offshore wind power and solar photovoltaic energy sources in the state of Bahia located in the northeast of Brazil. The two scenarios were created for a PEM electrolysis plant: 50 MW and 100 MW. As a result, for the LCOH, we got, as a result, 2.28 US\$/kg H<sub>2</sub> and 2.16 US\$/kg H<sub>2</sub> for solar PV, and 3.81 US\$/kg H<sub>2</sub> and 3.57 US\$/kg H<sub>2</sub> for offshore wind unit in the 50 MW and 100 MW systems, respectively. Despite the lower cost, the intermittency of renewable sources must be considered for the operation of solar PV and wind power plants. Offshore wind plants have the advantage of being installed near the main source for electrolysis – water – and of having a large space available, but with a trade-off on cost. The results show that a hybrid system uniting has the potential to be the most viable strategy.

**Keywords:** hydrogen production, water electrolysis, economic evaluation, renewable energy, PEM electrolysis.

### 1. Introduction

The development of sustainable energy generation requires technological changes, which revolve around three pillars: economy on production and demand, increased efficiency on production, and the use of clean energy sources to replace fossil fuels (Lund, 2007). In this sense, green hydrogen emerges as a viable and sustainable alternative capable of diversifying the energy matrix and reducing the dependence on fossil fuels, contributing to decarbonizing global economies. For this, by achieving global targets for reducing greenhouse gas (GHG) emissions and increasingly adopting the use of clean sources in the energy matrix, such as wind and solar photovoltaic (Nikolaidis e Poullikkasop, 2017).

Green hydrogen is produced through an electrolysis process, which uses electricity from a renewable source to split up the water molecule. Despite of already being widely applied on an industrial scale, this technology still faces major challenges, mainly due to the high cost related to its production, storage, transport, and the required electrical energy. With regard to energy and technological advances, wind and solar PV have increased their efficiency and capacity factor through the years and reduced the final price.

In Brazil, mainly in the Northeast Region, a variety of natural resources is available for the production of electricity with a great generation potential in using water sources, wind, solar PV, among others. In view of the global trend towards decarbonizing its economy, as well as diversifying its energy matrix, Brazil has made large investments and developed public

policies for the use of new and sustainable technologies, such as wind turbines and electrolysis on hydrogen production, and improved efficiency of the existing ones.

Bahia is the fifth largest state in terms of territory and has geographic and meteorological characteristics that favor the development of these renewable sources (Atlas solar, 2018; Atlas eólico, 2013). Furthermore, it has an industrial pool capable of meeting the required demands. Currently, the State of Bahia is the second generator of wind energy and eighth in solar PV (distributed generation), with an installed capacity of wind and solar PV of 6.46 GW and 621.4 MW, respectively (Abeeólica, 2022; Absolar, 2022). Its industrial complex of Camaçari has already started the project for producing Green Hydrogen, demonstrating the State's maturity in the face of technological advances aimed at these segments. In this sense, the present article proposes a comparative study between two sources of clean energy – offshore wind and solar photovoltaic – to identify which of these has greater economic attractiveness for producing green hydrogen in the state of Bahia. For this purpose, two scenarios were created for a hydrogen plant with different configurations of electricity supply (offshore wind and solar photovoltaic) and a capacity of 50 MW and 100 MW of the PEM electrolysis unit, simultaneously, the potential of the renewable energy generation and application was studied. The work assumes average demands and does not consider dynamic elements.

## 2. Methodology

### 2.1. Geographic localization

In the present work, geographic location was determined in two instances. One for the Solar PV Plant (SPVP) and the other for the Offshore Wind Power Plant (OWPP). As the scope of the study uses the state of Bahia as a case, the peculiarities of the region were considered so the modeling could represent the desired scenario and the evaluation of potential hydrogen production. The data and calculations assume average demands as a preliminary analysis for this study, if the results conclude the real potential considering all infrastructure necessary, dynamics elements should be implemented in future work.

The SPVP uses the data from *Bahia's Solar ATLAS* (2018), a document from the government of Bahia, in partnership with the SENAI CIMATEC, that shows and details the solar energy potential of the state. Protection and agriculture areas, forestry formation, lakes and reservoirs, and areas with a land inclination bigger than 3% were excluded. Thus, the useful and free space could be evaluated and compared with the results of the modeling calculations, being a parameter limitation. Hence, the SPVP localization considered the average Global Horizontal Irradiation (GHI) potential of all region and the area available for implementation of the solar panels. Further implementation of the plant, in this case, should consider areas of *Bahia's Solar ATLAS* with the availability of water in order to determine the exact location in the state.

In the other case, the OWPP was located in a region of the cost of Bahia that first satisfied the technical calculations, second the proximity with a port and infrastructure, and third the environmental sustainability. So, the best area chosen with aid of GLOBAL WIND ATLAS (GWP, 2022), was 13 km from Ilhéus city, which has proximity to a port, gives an average wind velocity in the year that satisfy the modeling, and avoids the Abrolhos Bank and Vitória-Trindade Chain, an important region with marine biodiversity, coral reefs, endemic and endangered marine species. It is important to note, however, that other areas could be evaluated with good potential for the installation of the OWPP, as the coast of Bahia extends for more than 1000 km.

### 2.2. Proposed scheme of modeling

In order to calculate the LCOH from the proposed scenario, this study considers the modeling according to Figure 1. The demand and cost of energy source, compression, storage, transport, and treated water (in the case of desalinization for OWPP) are a function of the power of the electrolysis plant, which presents the major demand for energy.

In the case of SPVP, it was considered that the solar panels and the entire solar energy production infrastructure are close to the electrolysis production area. However, for the OWPP, each turbine was considered to produce hydrogen coupled with the turbine using water from desalinization for the PEM electrolysis, therefore, hydrogen is transported

through a pipeline and main storage occurs onshore (Groenemans, 2022). Transport for SPVP was considered as the number of compressed hydrogen trucks necessary to handle the daily production of the electrolysis plant (Mints, 2006). On the other hand, transport for OWPP was considered as the use of a pipeline to the near harbor and assumed an existing infrastructure available to distribute the hydrogen (Groenemans, 2022).

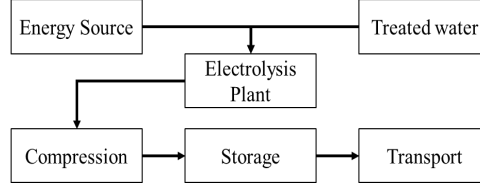


Figure 1: Scheme considered in the LCOH.

### 2.2.1. Solar

The model of the SPVP was developed, as aforementioned, as a function of the electrolysis plant. Thus, the energy that solar PV power should deliver needs to satisfy all the demands of the electrolysis unit and compression in this study. For this, the average GHI of Bahia was taken into account for the capacity of solar energy production (Atlas solar, 2018). A model of a photovoltaic monocrystalline module with 670 W, a dimension of 3.1 m<sup>2</sup>, and 21.5 % efficiency were considered to compose the plant.

The annual energy production was calculated according to Equation 1 (Hyjack, 2022):

$$\text{Solar PV production} \left( \frac{\text{kWh}}{\text{year}} \right) = \text{GHI} \left( \frac{\text{kWh}}{\text{m}^2 \cdot \text{year}} \right) * \text{Area}(\text{m}^2) * \text{Module Efficiency} \quad (1)$$

Then, using the tool Solver in Excel, the energy demand and solar PV production converged by alternating the area required for the unit. After this, the number of required PV modules was calculated so the costs could be evaluated.

### 2.2.2. Offshore Wind Power Plant

The Wind Farm modeling was structured first considering the potential of the selected region. For this, the annual average of wind velocity was considered using the Global Wind Atlas (GWP, 2022). It considered a height of 200 m for the tower, swept area of 38,000 m<sup>2</sup>, and blades with 107 m length, the air density used was 1.225 kg/m<sup>3</sup> (Groenemans, 2022). So, it was possible to calculate the power that the model of the wind turbine could generate in such conditions of operation using Equation 2 (Catrinus, 2017):

$$P(W) = \frac{1}{2} \rho A v^3 C_p \quad (2)$$

where, P is the power in W,  $\rho$  the air density in kg/m<sup>3</sup>, A the swept area in m<sup>2</sup>, v is the wind speed in m/s, and C<sub>p</sub> is the power coefficient, considered 0.65.

Since the OWPP provides all the energy, the power required accounts for the energy demand of the electrolysis unit, compression unit, and desalination unit.

### 2.3. PEM Electrolysis

The electrolysis unit data was obtained from the software “Size and Costing” in Hyjack (Hyjack, 2022). It compiles different suppliers to inform the energy consumption, hydrogen production, and efficiency as a function of the desired power of the electrolysis plant. The platform was created to assist projects of hydrogen and work as a regularly updated software package to help with analysis and information from specialized workers in the area.

Finally, to select the plant balance data, a useful life of 90,000 h of the pile was considered as a base criterion (Macedo, 2022). In this study, a 50 MW and 100 MW system was considered, with hydrogen output of 40 bar and 70 °C.

### 2.4. Compression

The compression unit considers the specific work necessary for a reciprocating compressor to compress hydrogen from the output pressure of the PEM unit to up 350 bar, the same pressure of storage. For this, Equation 3 is used (Hyjack, 2022):



$$\text{Specific Work} \left( \frac{J}{kg} \right) = \frac{\frac{\gamma}{\gamma-1} \frac{R \cdot T_1}{M} \left[ \left( \frac{P_2}{P_1} \right)^{\frac{\gamma-1}{\gamma}} - 1 \right]}{\text{isentropic efficiency}} \quad (3)$$

Where  $\gamma$  is isentropic coefficient assumed 1.4, R is the universal law constant,  $T_1$  the initial temperature,  $P_2$  and  $P_1$ , output pressure and inlet pressure, respectively, and M the molar mass of hydrogen. Isentropic efficiency, in turn, is calculated according to Equation 4 (Hyjack, 2022).

$$\text{Isentropic Efficiency} = -2.3082 * r^2 + 20.717 * r + 40.719 \quad (4)$$

The ratio between the discharge and suction pressure is  $r$ . After this, it is possible to calculate the power required for the system to compress the mass of hydrogen to the desired pressure. Equation 5, considers (Hyjack, 2022):

$$\text{Power (kW)} = \frac{Q * \left( \frac{\rho}{3600 * 1000} \right) * \text{Work} * (1 + \text{leaks})}{\text{Mechanical efficiency} * \text{Electrical efficiency}} \quad (5)$$

Q is the flow rate in kg/h,  $\rho$  the relative density obtained by data regression from Hyjack (2022), leaks assumed 3%, mechanical and electrical efficiencies assumed 79% and 95%, respectively.

### 2.5. Storage

Hydrogen storage was assumed in tanks of compressed gas of 350 bar with a capacity of 90 kg according to supplier information (FIBA, 2022). Furthermore, the quantity necessary to store the daily production was calculated according to the balance of mass of the system.

### 2.6. LCOH evaluation and economic assessment

In order to determine the comparison between the two strategies of hydrogen production, leveled cost of hydrogen was estimated according to the following Equation 6 (Macedo, 2022).

$$\text{LCOH} \left( \frac{\$}{kg} \right) = \frac{\text{Total Costs}(\$)}{\text{Total hydrogen production}(kg)} \quad (6)$$

The costs and data necessary to estimate total CAPEX and O&M are presented in Table 1:

	CAPEX	O&M (%CAPEX)	Notes	Reference
Total PEM	$C(\text{€}) = 1.7 * \text{Power} * (6046 * \text{Power}^{-0.2})$	4%	Power in kW	Hyjack (2022)
Replacement	N/A	20% Total PEM	N/A	Singlitico (2021)
Compressor	$C(\text{€}) = (75700 * \text{Power}^{-0.62}) * \text{Power}$	8%	Power in kW	Hyjack (2022)
Storage	$C(\text{€}) = \text{Cap} * N * (0.0015 * P^2 + 0.2521 * P + 448.54)$	5%	Cap = capacity of tank; N = number of tanks; P (bar)	Hyjack (2022)
Transport	$C(\$) = N * \text{CHT}$	-	N = number of trucks; CHT = cost of hydrogen truck	Mintz (2006)
SPVP	$C(\$) = N * \text{Pot} * \text{Cm}$	6%	N = number of modules; Pot = Power in Watt of each module; Cm = cost of each module (\$/W)	-
Desalination	$C(\text{€}) = 61200 * \left( \frac{\text{Water demand}(l/h)}{2000} \right)^{0.6}$	2,50%	N/A	Catrinus (2017)
Pipeline	$C(\text{M\$}) = A + 0.3555 * A + 0.57125 * A$	2%	$A = 13.3 * \left( \frac{D(km)}{16} \right)^{0.6}$	-
OWPP	$C(\$) = \text{Pot}_{req} * N * 3 * 10^6$	3%	$\text{Pot}_{req}$ = Power required in MW; N = number of wind turbine	McDonagh (2020)
Eolic DECEX	N/A	10%	N/A	McDonagh (2020)

**Table 1:** Economic data.

With regard to economic assumptions, it was estimated the lifetime of the plant for the SPVP and OWPP, of 25 and 30 years respectively (Gutiérrez, 2021; Singlitico 2021). Besides, the cash flow of each strategy was calculated, considering the first year of implementation of the plant, replacement of equipment necessary each 10 years, and the decommissioning at the end of lifetime expected. Therefore, after annualizing the results for the year 2022, it was possible to estimate the LCOH in the state of Bahia using the potential of renewable sources.

## 3. Results

As a result of the methodology applied, Figure 2 shows the LCOH calculated.

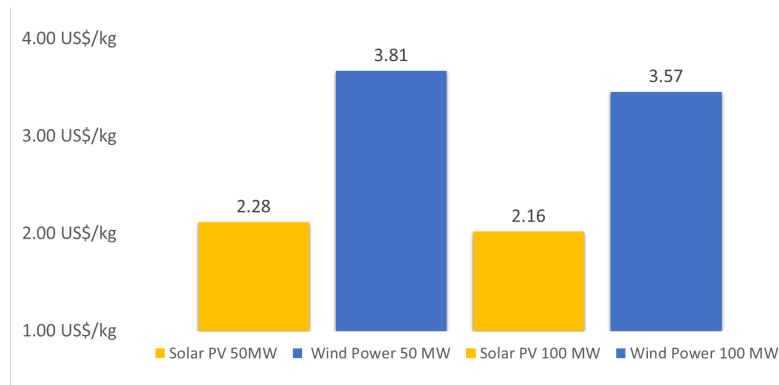


Figure 2: LCOH Solar PV and Offshore Wind Power 50 MW and 100 MW.

As observed, solar photovoltaic strategy shows the best performance in the levelized cost of hydrogen. However, it is important to note the advantages and disadvantages of each proposal. Solar PV power requires a larger area than offshore wind power, and this could be a constraint on the operation of a power plant. Moreover, offshore wind power is still in process of being more competitive in the next years, and despite the higher cost in the year analyzed, the technology has the advantage of a large area available at sea (McDonagh, 2020). Also, with the goal of the Brazilian government to expand decarbonization projects, political decisions, and regulatory marks can enable the implementation of strategy and reduce costs.

In addition, even Solar PV power having the lowest cost in this analysis, it is important to note the intermittent availability of this source during the day as the wind power (Minutillo, 2021). Hence, this should be a constraint in a project. In this case, a hybrid scheme with Wind Power should satisfy this problem that also affects the OWPP with the velocity varying along the day. In this way, both strategies could complement the lack of energy source during a cloudy day or a period with lower wind speed, for example.

#### 4. Conclusions

Following the study of previous years, the feasibility of renewable projects is becoming more real, especially in countries like Brazil where renewable sources are abundant. The LCOH calculated is already reaching the current price of steam methane reform which is the leading technology to produce hydrogen nowadays. Nevertheless, the intermittence of solar and wind energy is still a challenge. Therefore, in order to nearly fully decarbonize the energy source in a hydrogen production plant, the intermittence of renewable energy should be overcome. For example, a hybrid system of solar PV and wind power could minimize this challenge and can improve the level of dependency from grid energy for operation in a place that provides a good renewable potential.

#### 5. References

- Associação brasileira de energia eólica (Abeólica). InfoVento 28. São Paulo, nov., 2022. Disponível em: <<https://abeolica.org.br/energia-eolica/dados-abeolica/>>.
- Associação brasileira de energia solar fotovoltaica (ABSOLAR). Energia solar fotovoltaica no Brasil: infográfico Absolar. nov., 2022. Disponível em: <<https://www.absolar.org.br/mercado/infografico/>>.
- AWS Truepower, Camargo Schubert Engenheiros Associados, Fieb/Senai Cimatec. Atlas solar: Bahia. Curitiba: Camargo Schubert; Salvador: Secti; Seinfra: Cimatec/Senai. 76 p. 2018.
- Catrinus Jepma, Gert-Jan Kok, Malte Renz, Miralda V. Schot, Kees Wouters. Towards sustainable energy production on the North Sea - Green hydrogen production and CO2 storage: onshore or offshore? North Sea Energy, 2017.
- GWP - Global Wind Atlas. Energy data info. Available at: <[globalwindatlas.info/en](http://globalwindatlas.info/en)>. Access in: October, 2022.

- Groenemans, H., Saur, G., Mittelsteadt, C., Lattimer, J., & Xu, H. (2022). Techno-economic analysis of offshore wind PEM water electrolysis for H<sub>2</sub> production. *Current Opinion in Chemical Engineering*, 37, 100828. <https://doi.org/10.1016/j.coche.2022.100828>
- Gutiérrez-Martín, F., Amodio, L., & Pagano, M. (2021). Hydrogen production by water electrolysis and off-grid solar PV. *International Journal of Hydrogen Energy*, 46(57), 29038–29048. <https://doi.org/10.1016/j.ijhydene.2020.09.098>
- Lund H. Renewable energy strategies for sustainable development. *Energy* 2007;32(6):912–9.
- Macedo, S. F., & Peyerl, D. (2022). Prospects and economic feasibility analysis of wind and solar photovoltaic hybrid systems for hydrogen production and storage: A case study of the Brazilian electric power sector. *International Journal of Hydrogen Energy*, 47(19), 10460–10473. <https://doi.org/10.1016/j.ijhydene.2022.01.133>
- McDonagh, S., Ahmed, S., Desmond, C., & Murphy, J. D. (2020). Hydrogen from offshore wind: Investor perspective on the profitability of a hybrid system including for curtailment. *Applied Energy*, 265, 114732. <https://doi.org/10.1016/j.apenergy.2020.114732>
- Mintz, M., Gillette, J., Elgowainy, A., Paster, M., Ringer, M., Brown, D., & Li, J. (2006). Hydrogen Delivery Scenario Analysis Model for Hydrogen Distribution Options. *Transportation Research Record: Journal of the Transportation Research Board*, 1983, 114–120. <https://doi.org/10.3141/1983-16>
- Minutillo, M., Perna, A., Forcina, A., di Micco, S., & Jannelli, E. (2021). Analyzing the levelized cost of hydrogen in refueling stations with on-site hydrogen production via water electrolysis in the Italian scenario. *International Journal of Hydrogen Energy*, 46(26), 13667–13677. <https://doi.org/10.1016/j.ijhydene.2020.11.110>
- Nikolaïdis, Pavlos; Poullikkas, Andreas. Uma visão comparativa dos processos de produção de hidrogênio. *Revisões sobre energia renovável e sustentável*, v. 67, p. 597-611, 2017.
- Schubert, Camargo. *Atlas Eólico: Bahia*. Governo do Estado da Bahia, 2013.
- Singlítico, A., Østergaard, J., & Chatzivasileiadis, S. (2021). Onshore, offshore or in-turbine electrolysis? Techno-economic overview of alternative integration designs for green hydrogen production into Offshore Wind Power Hubs. *Renewable and Sustainable Energy Transition*, 1, 100005. <https://doi.org/10.1016/j.rset.2021.100005>.
- Sizing and Costing. Hyjack Hydrogen Tech Online. Available at: <hyjack.tech/components>. Access in: October, 2022.

# Resilience-aware multi-scale integration of distributed energy systems

Natasha J. Chrisandina<sup>a</sup>, Shivam Vedant<sup>b,c</sup>, Eleftherios Iakovou<sup>b,d</sup>, Efstratios N. Pistikopoulos<sup>a,b</sup>, Mahmoud M. El-Halwagi<sup>a,b,g</sup>

<sup>a</sup>*Artie McFerrin Department of Chemical Engineering, Texas A&M University, 3122 TAMU, 100 Spence St., College Station, TX 77843, USA*

<sup>b</sup>*Texas A&M Energy Institute, Texas A&M University, College Station, TX, 77843, USA*

<sup>c</sup>*Department of Multidisciplinary Engineering, Texas A&M University, College Station, USA*

<sup>d</sup>*Department of Engineering Technology and Industrial Distribution, Texas A&M University, College Station, USA*

<sup>e</sup>*Gas and Fuels Research Center, Texas A&M Engineering Experiment Station, College Station, USA*

*el-halwagi@tamu.edu*

## Abstract

To cost-competitively harness both renewable and traditional energy sources, the use of large-scale facilities should be considered along with distributed energy systems (DESSs), which are networks of small-scale energy production facilities and their associated suppliers, intermediaries, end users, and transportation links. Furthermore, to accurately capture the different operational scales in a DES and integrate long-term planning decisions alongside short-term scheduling decision, a multi-scale approach may be utilized. In this work, we present a multi-scale framework for the design and analysis of cost-competitive and resilient DESSs. To this end, the framework integrates: 1) facility location decisions to achieve specified threshold service level; 2) quantitative reliability analysis on individual process modules; and 3) considerations of tradeoff between multiple objectives such as cost, resilience, and sustainability. The framework is applied to a biomass value chain that uses various agricultural waste materials to produce a range of biofuels.

**Keywords:** distributed energy system, multi-scale engineering, resilience, supply chain, sustainability

## 1. Introduction

As the energy sector moves away from predominant use of fossil fuels to a more diverse suite of sources, two great challenges have emerged. First, new energy sources such as solar power, biomass, and wind display significant geo-temporal variability. Second, energy systems have been under great strain in the face of disruptions such as natural disasters, geopolitical conflicts, and global pandemics (Kakodkar et al., 2022). To simultaneously address both these challenges, distributed energy systems (DESSs) have been proposed as a promising pathway (Yang and You, 2018). To enable cost-competitive utilization of locally available energy sources, DESSs integrate small-scale modular energy production facilities across geographical locations with other key stakeholders along the supply chain, including traditional large-scale facilities.

In the current literature, there has been growing interest in DESSs and their capability to provide operational flexibility and resilience benefits in an uncertain world (Allen et al., 2019; Al-Fadhli et al., 2019). Utilizing modular facilities within a DES can also reduce

set-up and maintenance time, and provide flexibility in the form of mobile modular units (Baldea et al., 2017). However, a major question is how to balance the multiple objectives that need to be considered: cost-competitiveness, sustainability, and resilience. In this work, a multi-scale approach is employed to accurately capture the various spatio-temporal scales of operation prevalent in energy systems to allow discovery of potential synergies that can be leveraged to achieve win-win solutions. Building on this approach, we propose a framework that integrates planning and operational decisions – production facility design and facility location-allocation – to achieve a resilience-aware DES in a cost-competitive manner. The proposed methodology is demonstrated through a case study for planning and scheduling of a biomass value chain.

## 2. Methodology

### 2.1. Framework summary

In this work, a two-step procedure is presented wherein: 1) given the process flow diagram (PFD) for a modular unit, major process equipment selections are made to maximize the modular unit availability subject to a cost constraint, 2) a network is designed using the modular units from the previous stage to maximize the total system profit subject to a minimum reliability at every location. The framework is modeled as two mixed-integer programs (MIP) to represent each step independently.

#### 2.1.1. Nomenclature

The nomenclature used in the rest of this section is outlined below.

Table 1: Sets

Notation	Description	Notation	Description
$\mathcal{K}$	Set of modules	$\mathcal{F}$	Set of all feedstock resources
$\mathcal{L}$	Set of potential facility locations	$\mathcal{Y}$	Set of years in the planning horizon
$\mathcal{J}$	Set of source locations	$\mathcal{P}$	Set of all product resources
$\mathcal{J}$	Set of sink locations	$\mathcal{M}$	Set of months in one year
$\mathcal{V}$	Set of vehicles for transport	$\mathcal{U}$	Set of all utility resources
$\mathcal{E}$	Set of equipment	$\mathcal{B}$	Set of equipment types

Table 2: Parameters

Notation	Description	Notation	Description
$p_e$	CAPEX of equipment $e$	$cap_k$	Maximum capacity of module $k$
$p_k$	Annual CAPEX of module $k$	$qty_{b,k}$	Number of unit type $b$ used in module $k$
$p_k^{max}$	Maximum CAPEX for module $k$	$type_{e,b}$	1 if unit $e$ is of type $b$ , 0 otherwise
$capex_k$	CAPEX multiplier for module $k$	$life_k$	Lifetime of module $k$ (years)
$MTBF_e$	Maximum MTBF of equipment $e$	$m_{x,l}$	Distance between location $x$ and facility $l$ (where $x \in \mathcal{J}$ or $\mathcal{J}$ )
$MTTRes_e$	Minimum MTTRes of equipment $e$	$Re_l^{min}$	Minimum availability allowable in facility $l$
$conv_{f,k}$	Conversion rate of feedstock $f$ by module $k$	$conv_{p,k}$	Yield of product $p$ by module $k$

$\text{conv}_{u,k}$	Conversion rate of utility $u$ by module $k$	$r_p$	Revenue from product $p$
$c_f$	Cost of feedstock $f$	$sc_p$	Shipping cost of product $p$ per mile per ton
$c_u$	Cost of utility $u$	$sc_f$	Shipping cost of feedstock $f$ per mile per ton

Table 3: Variables

Notation	Description	Notation	Description
$p_k$	CAPEX of module $k$		<u>Binary variables</u>
$\text{qty}_{e,k}$	Number of unit $e$ used in module $k$	$y_{e,k}$	1 if unit $e$ is chosen for module $k$ , 0 otherwise
$Re_k$	Availability of module $k$	$y_{k,l}$	1 if module $k$ is chosen for location $l$ , 0 otherwise
$\text{prod}_{k,l}$	Actual production by module $k$ in facility $l$	$x_l$	1 if location $l$ is chosen for the network, 0 otherwise
$\text{qty}_{k,l}$	Number of module $k$ used in facility $l$		
$Re_l$	Availability of facility $l$		

## 2.2. Stage 1: Process module design

Modular unit reliability is estimated using the methodology outlined by Al-Douri et al. (2021). In this methodology, PFDs of the modular units are assumed to be known. The MTBF, which is the inverse of the failure rate, and MTTRes data for every equipment are derived from the 2015 edition of the OREDA Handbook for Onshore and Offshore Reliability Data. Modular unit availability, used here as a proxy metric for reliability, is then calculated as follows:

$$Re_k = \prod_B \left( \frac{\text{MTBF}_b}{\text{MTTRes}_b + \text{MTBF}_b} \right) \quad (1)$$

The availability of each modular unit is then used to modify profit calculation to express the fact that no revenue and operating expenses are applicable when the unit is non-operational.

### 2.2.1. Constraints

There are three groups of constraints that are needed in this stage of the model. The first group of constraints limits the number and type of units used per operation called for by the PFD:

$$\sum_{\varepsilon \text{ if type}_{e,b}=1} y_{e,k} \leq 1 \quad \forall k \text{ in } \mathcal{K} \quad (2)$$

$$\sum_{\varepsilon \text{ if type}_{e,b}=1} \text{qty}_{e,k} = \text{qty}_{b,k} \quad \forall k \text{ in } \mathcal{K} \quad (3)$$

The second group of constraints restricts the total CAPEX of the module constructed based on the maximum CAPEX allowable:

$$\frac{\text{capex}_k \times \sum_{\varepsilon} \text{qty}_{e,k} \times p_e}{\text{lifex}_k} \leq p_k^{\max} \quad \forall k \text{ in } \mathcal{K} \quad (4)$$

The third group of constraints calculates the aggregate availability of each module based on the availabilities of the units chosen for the module:

$$Re_k = \prod_{\mathcal{B}} \left( \sum_{\varepsilon \text{ if type}_{e,b}=1} y_{e,k} \times \left( \frac{\text{MTBF}_e}{\text{MTTRes}_e + \text{MTBF}_e} \right) \right)^{\text{qty}_{b,k}} \quad \forall k \text{ in } \mathcal{K} \quad (5)$$

### 2.2.2. Objective functions

The objective function in this stage is to maximize the total availability of every module being constructed:

$$\max \sum_{\mathcal{K}} Re_k \quad (6)$$

### 2.3. Stage 2: facility location-allocation

Mass balance constraints for this stage of the model is outlined in Chrisandina et al. (2022). Additional constraints to specify aggregate availability for each potential location are outlined below.

#### 2.3.1. Constraints

The availability of each location is calculated as an aggregate of the availability of all modules selected for that location, assuming that the modules run in parallel:

$$Re_l = 1 - \prod_{\mathcal{K}} (1 - Re_k)^{\text{qty}_{k,l}} \quad \forall l \text{ in } \mathcal{L} \quad (7)$$

The annual profit per location is calculated as follows, where the location availability is used to modify both the revenue and operating cost factors:

$$\text{profit}_l = Re_l (\text{rev}_l - \text{OPEX}_l) - \sum_{\mathcal{K}} y_{k,l} \times \text{CAPEX}_k, \quad \forall l \text{ in } \mathcal{L} \quad (8)$$

where the revenue, operating and capital costs are expressed below:

$$\text{rev}_l = \sum_{\mathcal{K}} \sum_{\mathcal{P}} r_p \times \text{conv}_{p,k} \times \text{prod}_{k,l} \times \text{qty}_{k,l} \quad \forall l \text{ in } \mathcal{L} \quad (9)$$

$$\text{OPEX}_l = \sum_{\mathcal{K}} \left( \begin{aligned} & \sum_{\mathcal{F}} c_f \times \text{conv}_{f,k} \times \text{qty}_{k,l} + \sum_{\mathcal{U}} c_u \times \text{conv}_{f,k} \times \text{qty}_{k,l} \\ & + \sum_{\mathcal{J}} \sum_{\mathcal{F}} sc_f \times m_{i,l} \times \text{conv}_{f,k} \times \text{prod}_{k,l} \\ & + \sum_{\mathcal{J}} \sum_{\mathcal{P}} sc_p \times m_{j,l} \times \text{conv}_{p,k} \times \text{prod}_{k,l} \end{aligned} \right) \quad \forall l \text{ in } \mathcal{L} \quad (10)$$

$$\text{CAPEX}_l = \sum_{\mathcal{K}} p_k \times \text{qty}_{k,l} \quad \forall l \text{ in } \mathcal{L} \quad (11)$$

#### 2.3.2. Objective function

The objective of this stage is to maximize the total profit obtained by the system, with consideration of location availability:

$$\max \sum_L profit_i \quad (12)$$

### 3. Case study

#### 3.1. Setup

The framework is applied towards the simultaneous design and schedule optimization of a new multi-city biomass value chain to meet a monthly demand for several value-added products. A set of microbiorefineries (MBRs) are available to produce different products, and the PFDs for each MBR are fixed. The case study considers three biomass resources (cassava, banana, and mango) and three products (biogas, methanol, and mango oil). Agricultural cities which supply each resource also act as potential locations for placing MBRs. The availability of the resources fluctuates monthly based on local harvest patterns, while the demand for products remains constant year-round.

#### 3.2. Results & discussion

In the first stage of the framework, MBRs are constructed given their PFDs and a selection of possible equipment units to purchase with associated costs, MTBF, and MTTRes values. It can be observed that purchasing more reliable units increases both the cost and availability of an MBR. However, there is a point of diminishing return (see Fig. 1) after which additional investment in costly but higher-performing units do not significantly increase aggregate MBR availability. MBR manufacturers can utilize this insight to decide on the level of investment into reliable units based on the desired performance. Furthermore, it is also possible to conduct sensitivity analysis on the major components within an MBR to determine the contribution of each equipment to overall reliability and direct targeted investments towards the most critical components.

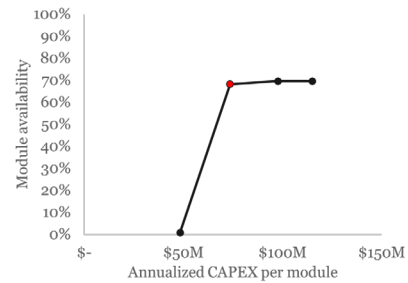


Figure 1: Pareto front for annualized CAPEX per process module vs module availability

In the second stage of the framework, a supply chain network using MBRs, designed in the previous stage, was constructed. Two separate scenarios were generated: one scenario using lower-cost and less reliable MBRs, and one scenario using higher-cost and more reliable MBRs. As shown in Fig. 2, higher profit is achievable when more reliable MBRs are used despite the higher capital cost per module installed in this scenario. This is due to the fact that fewer MBRs need to be installed at each location to meet the minimum availability requirement per location, leading to lower total capital costs. Furthermore, the installed MBRs

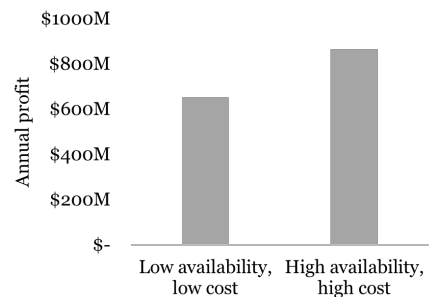


Figure 2: Comparison of total network profit based on the availability and cost of MBRs installed in each location



are operational for a longer proportion of the year which leads to higher revenue generating potential.

#### 4. Conclusion

DESS offer a potential pathway towards a more resilient energy system due to their flexibility in responding to fluctuating supply or demand patterns. Utilization of modular production units can also provide additional benefits since units can be designed to enhance overall system resilience. This work has presented a two-stage framework to integrate modular production facility design with supply chain network optimization. It was observed that designing modular units to maximize reliability contributes to increased network profit compared to purchasing cheaper but less reliable modular units. Depending on the risk tolerance and budget constraints of the decision-maker, the limits on modular unit cost and location reliability can be adjusted. In this work, modular units are considered as stationary during the planning horizon. However, modular units can be transported between locations to limit idle time if feedstock supply is unavailable at a particular location. To address this limitation, the framework can be extended to account for modular unit transport. Some idle time during transport needs to be considered, as well as the effect of transportation-related wear and tear on the performance and reliability of the units.

#### References

- R. Kakodkar, G. He, C.D. Demirhan, M. Arbabzadeh, S.G. Baratsas, S. Avraamidou, D. Mallapragada, I. Miller, R.C. Allen, E. Gençer, and E.N. Pistikopoulos, 2022, A review of analytical and optimization methodologies for transitions in multi-scale energy systems, *Renewable and Sustainable Energy Reviews*, 160, 112277.
- M. Yang, F. You, 2018, Modular methanol manufacturing from shale gas: Techno-economic and environmental analyses of conventional large-scale production versus small-scale distributed, modular processing, *AIChE Journal*, 64, 495-510.
- R. C. Allen, D. Allaire, and M.M. El-Halwagi, 2019, Capacity planning for modular and transportable infrastructure for shale gas production and processing, *Ind. Eng. Chem. Res.*, 58, 5887–5897.
- F.M. Al-Fadhli, H. Baaqeel, and M.M. El-Halwagi, 2019, Modular design of carbon-hydrogen-oxygen symbiosis networks over a time horizon with limited natural resources. *Chem. Eng. Process. Process Intensif.* 141, 107535.
- M. Baldea, T.F. Edgar, B.L. Stanley, and A.A. Kiss, 2017, Modular manufacturing processes: Status, challenges, and opportunities. *AIChE Journal*, 4262-4272.
- N.J. Chrisandina, S. Vedant, E. Iakovou, E.N. Pistikopoulos, and M.M. El-Halwagi, 2022, Multi-scale Integration for Enhanced Resilience of Sustainable Energy Supply Chains: Perspectives and Challenges, *Computers & Chemical Engineering*, 107891.
- OREDA, 2015, *Offshore and Onshore Reliability Data Handbook Vol. 1.*

# Techno-enviro-economic analysis of H<sub>2</sub> economy in China from H<sub>2</sub> production to utilization

Xiaodong Hong<sup>a\*</sup>, Zuwei Liao<sup>b\*</sup>, Yao Yang<sup>a,b</sup>, Jingdai Wang<sup>b</sup>, Yongrong Yang<sup>b</sup>

<sup>a</sup>*ZJU-Hangzhou Global Scientific and Technological Innovation Center, Hangzhou 311215, China*

<sup>b</sup>*State Key Laboratory of Chemical Engineering, College of Chemical and Biological Engineering, Zhejiang University, Hangzhou 310027, China*

\**hongxiaodong@zju.edu.cn*

## Abstract

China, the world's largest CO<sub>2</sub> emitter, has promised to achieve carbon neutrality before 2060. The hydrogen economy is one of the most promising alternatives to cut emissions. The planning of the hydrogen economy should consider the diversity of China, such as renewable energy distribution, industry distribution, etc. In this work, the impact of renewable energy availability on the production cost of hydrogen and the impact of hydrogen transport are explored on a national scale. A comprehensive techno-enviro-economic analysis of the hydrogen economy in China from its production to utilization is carried out. A mixed integer linear programming (MILP) model is developed to optimize the levelized cost of hydrogen (LCOH) and CO<sub>2</sub> emissions of hydrogen (CEOH) utilizing local solar, wind, and grid electricity. Key questions, such as import H<sub>2</sub> from where via which form for what application, are answered for each province. These are crucial and useful for decision-makers to plan hydrogen economy development.

**Keywords:** techno-enviro-economic, hydrogen economy, carbon neutrality.

## 1. Introduction

With the world's population growing, human energy demand keeps rising, which has caused climate change and environmental degradation and threatened the way of life. Over the past decades, the world has devoted significant efforts to developing renewable energy technologies. The renewables in the global power generation energy mix have been rising rapidly. However, the developments in renewable energy technologies also reveal the critical lack of transportable forms of energy has become the main bottleneck in a global shift toward renewables. Growing adoption of grid-connected photovoltaic systems caused in advertent stress on the electrical grid (Obi et al., 2016). A chemical form of energy storage, simply like fossil fuel storing energy over millions of years, is needed to compensate for the intermittent nature of solar and wind. This throws the spotlight onto the hydrogen economy (Majumdar et al., 2021), where hydrogen has dual properties of energy carrier and material carrier. Much research has been carried out to reveal the role of different hydrogen production technologies, hydrogen transport carriers, and hydrogen applications in a hydrogen-based low-carbon society. Staffell et al. (2019) review the challenges of the hydrogen economy, showing considerable improvements are still required for hydrogen to become truly competitive. Parkinson et al. (2019) evaluated 12 different hydrogen production techniques from technical, economic, and environmental aspects. The results show that the most cost-effective methods of decarbonization still utilize fossil feedstocks, and methane pyrolysis may be the most cost-effective short-term abatement

solution. The cost of renewable electrolytic routes is high, while they offer significantly higher emissions reduction. Hong et al. (2021) studied four hydrogen carriers and two hydrogen end-use scenarios. Results showed that liquid hydrogen is preferred for road transport application. Sunny et al. (2020) studied the regional transition of the heating sector from natural gas-based infrastructure to H<sub>2</sub> by a supply chain optimization model. They found that the synergistic deployment of autothermal reforming of methane and biomass gasification with CO<sub>2</sub> capture and storage (CCS) is critical in achieving cost-effective decarbonization.

The impact of renewable energy availability on the production cost of hydrogen, such as solar radiation and wind speed, has been studied. Al-Sharafi et al. (2017) investigated the potential of power generation and hydrogen production via solar and wind energy resources at five different locations in the Kingdom of Saudi Arabia, Toronto, and Sydney. Pan et al. (2020) studied the levelized cost of hydrogen in 31 provinces considering provincial differences in electricity prices, shares of renewables in power grids, and solar radiations in China. The cost varies from 31.5 ¥ per kg to 46.8 ¥ per kg. However, the impact of hydrogen transmission and delivery on the hydrogen cost and carbon footprint has not been explored. Relatively cheap hydrogen produced in renewable energy abundant regions can lead to higher landed costs higher than hydrogen produced locally, and an even higher carbon footprint depending on its supply chain.

In light of the aforementioned concerns, we first develop an optimization model of Hybrid Power-to-Hydrogen System, which can design a cost-minimal or emission-minimal system considering solar and wind resource availability and time-of-use grid power prices. Then, a systematical study and comparison of levelized cost of hydrogen (LCOH) and carbon emissions of hydrogen (CEOH) are presented for the 31 provinces of mainland China studied in Pan et al. (2020). H<sub>2</sub> via local production and provincial transport are all compared for each province.

## 2. Hybrid Power-to-Hydrogen System

A H<sub>2</sub> supply chain is constructed as shown in Figure 1, considering local production by a hybrid power-to-hydrogen system (HPtoH). The water electrolyzer (WE) is the main component, which adopts power from solar photovoltaic panels (PV), wind turbines (WT), utility-scale lithium batteries (BT), and power grids (GRID). In the study, proton exchange membrane (PEM) electrolysis is adopted, due to its higher efficiency and flexibility than alkaline electrolysis.

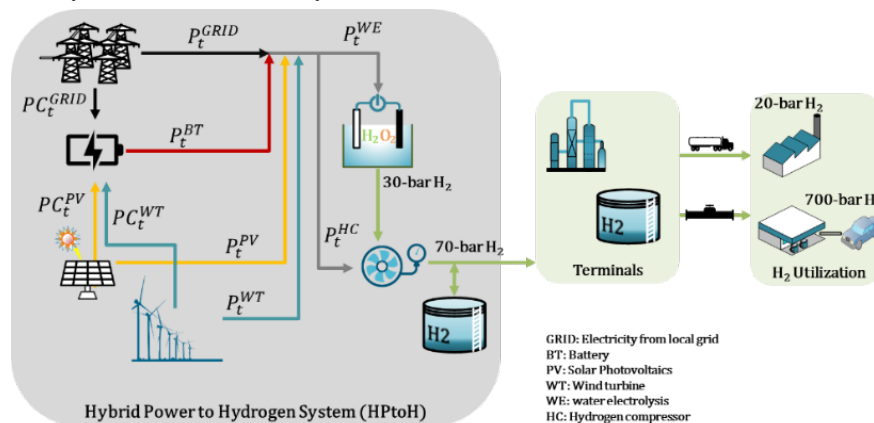


Figure 1. Illustration of Hybrid Power-to-Hydrogen System (HPtoH).

In Figure 1, it is assumed that 30-bar H<sub>2</sub> is produced, and then compressed to 70 bar for transport. Considering the intermittent nature of solar and wind, a utility-scale lithium battery storage and buffer storage tank (ST) are adopted to ensure the system could have steady daily output. Batteries can store the excess generated electricity of PV and WT while charging batteries during valley-time-price hours and discharging batteries during peak-time-price hours are also considered. After 70-bar H<sub>2</sub> is delivered to terminals, it will be further compressed to 540 bar for truck delivery as 540-bar H<sub>2</sub> (CH540) and liquefied for truck delivery as liquid hydrogen (LH<sub>2</sub>) (Hong et al., 2021). Besides, 70-bar H<sub>2</sub> can also be directly transported by pipelines.

### 3. Model Description

The objective of the developed model is to minimize the levelized cost of hydrogen (LCOH, ¥/kg), as shown by Eq. 1. LCOH is obtained by the annual output of the system and the annual cost. The annual cost consists of three parts, the capital cost of facilities (CAPEX), fixed operation and maintenance cost (FOM), and variable operation cost (VOP). CAPEX is obtained by the annualized factor, the unit capital investment cost, and the capacity of the facility (PV, WT, WE, BT, ST). The annualized factor is derived from the lifetime of the facility and the discount rate. The overall CO<sub>2</sub> emissions of hydrogen (CEOH) can be obtained by summing up emissions in the entire supply chain, including grid electricity consumption and truck fuel consumption. Note that, the CEOH can also be adopted as an objective.

$$LCOH = \frac{CAPEX+FOM+VOP}{\lambda \times 1000 \times 365} \quad (1)$$

The model also includes energy balance equations for PV, WT, WE, and BT, mass balance equations for WE and ST, constraints on daily H<sub>2</sub> output and energy availability, etc. The techno-economic parameters are obtained from the literature (Terlouw et al., 2022; Wisser et al., 2021; Feldman et al., 2021), as shown in Table 1.

Table 1. Techno-economic parameters.

Item	Unit	Value
Photovoltaic (PV) investment cost	¥/kW	5477.4
Fixed operation and maintenance (O&M) cost of PV	¥/kW-year	547.74
Wind turbine (WT) investment cost	¥/kW	10087.8
Fixed operation and maintenance (O&M) cost of WT	¥/kW-year	296.7
Water electrolyzer (WE) investment cost	¥/kW	9213.98
Fixed operation and maintenance (O&M) cost of WE	¥/kW-year	212.63
Battery (BT) investment cost	¥/kW	9522.0
Fixed operation and maintenance (O&M) cost of BT	¥/kW-year	69
Hydrogen storage (HS) investment cost	¥/kg H <sub>2</sub>	3588
Fixed operation and maintenance (O&M) cost of HS	¥/kg H <sub>2</sub> -year	35.8
Hydrogen compressor (HC) investment cost	¥/kW	27887.8
Fixed operation and maintenance (O&M) cost of HC	¥/kW-year	1115.5
Fixed operation and maintenance (O&M) cost of GRID	¥/kW-year	336~576
Conversion efficiency of water electrolyzer	-	60.5%
Discount rate	-	4%

#### 3.1. Renewable energy availability

In this study, hourly solar power data are collected from the PVWatts Calculator developed by National Renewable Energy Laboratory (NREL). A PV module with a

two-axis tracking system (dual-axis tracking system) is considered to calculate the power generation capacity. For power generation by wind turbines, the data of wind speed at 50 meters is obtained from NASA by API. The wind speed at 80 meters can be obtained by the following Eq. 2.

$$v = v_{ref} \times \left(\frac{H}{H_{ref}}\right)^\alpha \quad \forall t \in T \quad (2)$$

where  $v_{ref}$  and  $v$  represents the wind speed at the height of  $H_{ref}$  and  $H$  and  $\alpha$  is surface roughness (0.14). Based on the hourly wind speed at 80 meters, the turbine power can be obtained based on the turbine power curve.

### 3.2. Grid electricity

In China, industrial consumers can choose the single electricity price and the time-of-use price. In this study, the time-of-use price is adopted. The electricity prices are categorized into the rush, peak, flat, and valley time prices. The valley time price could be as low as only 40% of the flat price, while the peak time price could be 150% of the flat price. And the rush time price can be even 20% higher than the peak time price. The rush time price is mostly only available in specific hours in a few months with extremely cold or hot weather (e.g. July, August, December). With the time-of-use price and intermittent nature of solar and wind, the optimization model will find the optimal operating strategy of the HPtoH system, by minimizing LCOH and/or CEOH.

## 4. Results and Discussion

Provincial economic and environmental analysis of H<sub>2</sub> produced by the HPtoH system is presented in a case study. LCOH and CEOH of locally produced H<sub>2</sub> and delivered H<sub>2</sub> in each province are obtained. In each province, at least five locations are chosen for analysis. Solar and wind resource availability is obtained for each location, and production cost and emissions of H<sub>2</sub> are obtained via the HPtoH model (by Pyomo). The optimal system size, including water electrolysis, battery, H<sub>2</sub> storage, and compressor, is also determined. With the production cost of H<sub>2</sub>, the LCOH and CEOH can be calculated for various end users via different transport pathways.

Figure 2 shows the LCOH matrix, where the LCOH of 20-bar H<sub>2</sub> received by end-users is shown for each province. H<sub>2</sub> can be locally produced or by provincial transportation via the 70-bar pipeline. For the locally produced H<sub>2</sub>, we assume a transmission distance of 50 kilometers and a distribution distance of 15 km, from the producers to end-users. For the provincial transportation, distances between any two provinces are calculated. In Figure 2, each column represents the LCOH in each province. For example, the first column indicates the LCOH of 20-bar H<sub>2</sub> in Beijing. The First row is the one for the locally produced H<sub>2</sub>. The second one is from Tianjin. Only the LOCH lower than local production in each province is represented in the Figure. All blank blocks mean uneconomical supply chains. For the columns where most of the blocks are blank, local production is the most economical path, such as Inner Mongolia, Qinghai, and Xinjiang. On the contrary, these provinces tend to have fewer blank blocks, from the perspective of rows. Mongolia, Qinghai, Xinjiang, Ningxia, and Shanxi all have the potential to provide H<sub>2</sub> for many neighboring provinces.

It is worth mentioning that Figure 2 shows the case study where the LCOH is minimized. Thus, the HPtoH systems tend to use grid electricity if the valley time price is low. As a result, the CEOH of the produced H<sub>2</sub> is relatively high for most of the provinces, since the renewable energy share of the grid is still low. In another case, the CEOH is minimized for the HPtoH system. A similar plot of Figure 2 is obtained, which indicates the LOCH

varies from 40 ¥/kg H<sub>2</sub> to 66 ¥/kg H<sub>2</sub>. Given the length of the article, the results are not provided.

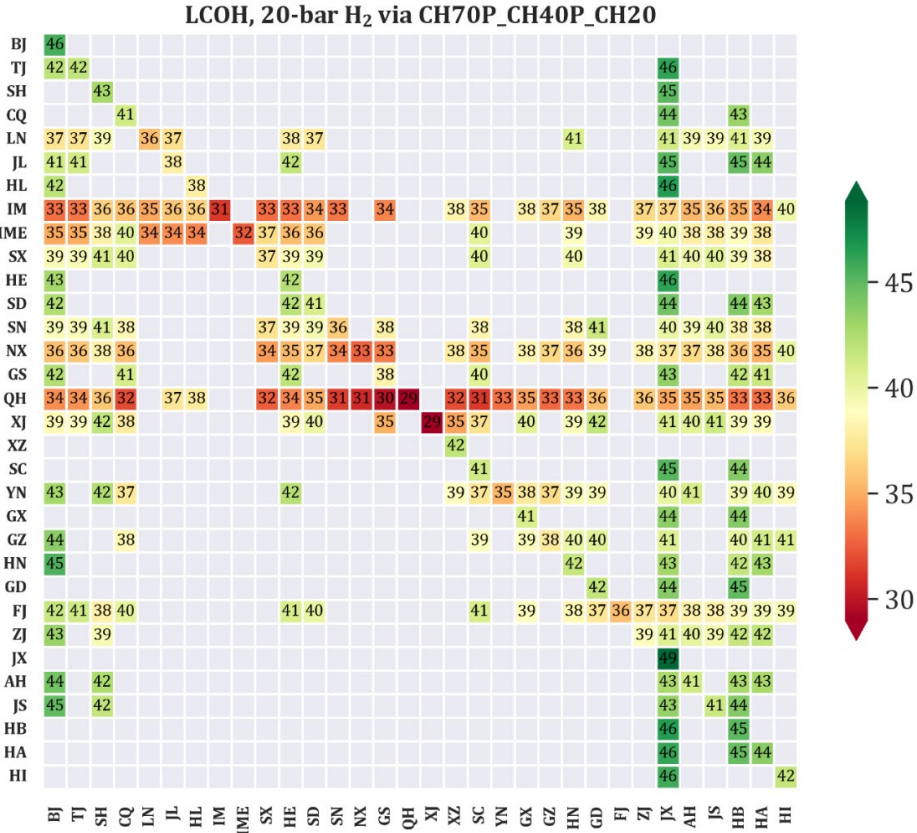


Figure 2. LCOH (¥/kg H<sub>2</sub>) of 20-bar H<sub>2</sub> received end users. (BJ: Beijing, TJ: Tianjin, SH: Shanghai, CQ: Chongqing, LN: Liaoning, JL: Jilin, HL: Heilongjiang, IM: Inner Mongolia West, IME: Inner Mongolia East, SX: Shanxi, HE: Hebei, SD: Shandong, SN: Shaanxi, NX: Ningxia, GS: Gansu, QH: Qinghai, XJ: Xinjiang, XZ: Tibet, SC: Sichuan, YN: Yunnan, GX: Guangxi, GZ: Guizhou, HN: Hunan, GD: Guangdong, FJ: Fujian, ZJ: Zhejiang, JX: Jiangxi, AH: Anhui, JS: Jiangsu, HB: Hubei, HA: Henan, HI: Hainan)

## 5. Conclusions

This study presents a comprehensive techno-enviro-economic analysis of hydrogen economy, to evaluate the competitiveness of hydrogen in each province of China. Hydrogen production, hydrogen delivery, and hydrogen utilization are considered simultaneously. A mixed integer linear programming (MILP) model is developed to optimize the levelized cost of hydrogen (LCOH) and CO<sub>2</sub> emissions of hydrogen (CEOH). The landed cost of H<sub>2</sub> from any other provinces in China and the carbon footprint are obtained for each province. Provinces in western China, including Inner Mongolia, Qinghai, Xinjiang, Ningxia, and Shanxi, have a higher potential to provide low-cost hydrogen for other provinces, as they have more abundant renewable energy sources. Many key questions still need to be answered with further research.

- Importing H<sub>2</sub> via liquid hydrogen, ammonia, and liquid organic carriers will be compared with local production
- High-pressure gas hydrogen (e.g. 100 bar, 300 bar) directly produced via PEM WE will be considered for other hydrogen applications, such as refueling stations.

- Off-shore wind will be considered for coastal provinces.
- Sensitivity analysis will be carried out to evaluate the impact of uncertainties.
- A multiperiod supply chain model will be developed for the planning of the hydrogen economy considering all provinces simultaneously.

### Acknowledgment

We gratefully acknowledge the financial support of the National Natural Science Foundation of China (U22A20415) and ZJU-Hangzhou Global Scientific and Technological Innovation Center.

### Nomenclature

LCOH	Levelized cost of hydrogen, ¥/ kg H <sub>2</sub>
CEOH	CO <sub>2</sub> emissions of hydrogen, kg CO <sub>2</sub> /kg H <sub>2</sub>
CAPEX	Annual capital cost of the hydrogen system, ¥/year
FOM	Annual fixed operation and maintenance cost of the hydrogen system, ¥/kW
VOP	Annual variable operation cost of the hydrogen system, ¥/kW
$\lambda$	Average daily H <sub>2</sub> production capacity, tonne per day
$v_{ref}/v$	Wind speed at the height of $H_{ref}/H$
$\alpha$	Surface roughness (0.14)

### References

- M. Obi, R. Bass, 2016, Trends and Challenges of Grid-Connected Photovoltaic Systems – A Review. *Renewable and Sustainable Energy Reviews*, 58, 1082–1094.
- A. Majumdar, J. M. Deutch, R. S. Prasher, T. P. Griffin, 2021, A Framework for a Hydrogen Economy. *Joule*, 5 (8), 1905–1908.
- I. Staffell, D. Scamman, A. Velazquez Abad, P. Balcombe, P. E. Dodds, P. Ekins, N. Shah, K. R. Ward, 2019, The Role of Hydrogen and Fuel Cells in the Global Energy System. *Energy & Environmental Science*, 12 (2), 463–491.
- B. Parkinson, P. Balcombe, J. F. Speirs, A. D. Hawkes, K. Hellgardt, 2019, Levelized Cost of CO<sub>2</sub> Mitigation from Hydrogen Production Routes. *Energy & Environmental Science*, 12 (1), 19–40.
- X. Hong, V. B. Thaore, I. A. Karimi, S. Farooq, X. Wang, A. K. Usadi, B. R. Chapman, R. A. Johnson, 2021, Techno-Enviro-Economic Analyses of Hydrogen Supply Chains with an ASEAN Case Study. *International Journal of Hydrogen Energy*, 46 (65), 32914–32928.
- N. Sunny, N. Mac Dowell, N. Shah, 2020, What Is Needed to Deliver Carbon-Neutral Heat Using Hydrogen and CCS? *Energy & Environmental Science*, 13 (11), 4204–4224.
- A. Al-Sharafi, A. Z. Sahin, T. Ayar, B. S. Yilbas, 2017, Techno-Economic Analysis and Optimization of Solar and Wind Energy Systems for Power Generation and Hydrogen Production in Saudi Arabia. *Renewable and Sustainable Energy Reviews*, 69, 33–49.
- G. Pan, W. Gu, Q. Hu, J. Wang, F. Teng, G. Strbac, 2021, Cost and Low-Carbon Competitiveness of Electrolytic Hydrogen in China. *Energy & Environmental Science*, 14 (9), 4868–4881.
- T. Terlouw, C. Bauer, R. McKenna, M. Mazzotti, 2022, Large-Scale Hydrogen Production via Water Electrolysis—a Techno-Economic and Environmental Assessment. *Energy & Environmental Science*, 15, 3583–3602.
- R. Wiser, M. Bolinger, B. Hoen, D. Millstein, J. Rand, G. Barbose, N. Darghouth, W. Gorman, S. Jeong, A. Mills, B. Paulos, 2021, Land-Based Wind Market Report: 2021 Edition. 87.
- D. Feldman, V. Ramasamy, R. Fu, A. Ramdas, J. Desai, R. Margolis, 2021, U.S. Solar Photovoltaic System and Energy Storage Cost Benchmark: Q1 2020. *Renewable Energy*, 120. PVWatts Calculator. <https://pvwatts.nrel.gov/> (accessed 2022-07-21).
- POWER | Data Access Viewer. <https://power.larc.nasa.gov/data-access-viewer/> (accessed 2022-07-29).

## Towards a novel concept for solid energy storage

Adrian Straub<sup>a</sup>, Bogdan Dorneanu<sup>a</sup>, Harvey Arellano-Garcia<sup>a</sup>

<sup>a</sup>*LS Prozess- und Anlagentechnik, Brandenburgische Technische Universität Cottbus-Senftenberg, Cottbus, D-03046, Germany*  
*arellano@b-tu.de*

### Abstract

In this contribution, the model-based development of a novel process concept for the storage and release of ammonia in solids is proposed. The concept is validated by means of the Aspen Plus® process simulator. As a promising prospect, Hexaaminenickel(II) chloride is selected. After a preparative stage, the process can cycle between the storage and release of energy. The process is split into a reaction and a separation section, in such a way that the same equipment is used for both storage and release steps. Sensitivity analysis and design parameter optimization are used to determine key process parameters. The operation ranges from standard conditions (25 °C and 1 atm) to temperatures not higher than 120 °C. Moreover, the simulation results show that it is possible to store over 50% of the base material in form of ammonia, equivalent to almost 10 wt.% hydrogen, placing the concept within the specific system targets set by the U.S. Department of Energy.

**Keywords:** Process design, Process modelling, Aspen Plus®

### 1. Introduction

Climate change and energy supply count amongst the most important issues of our time. The focus of many countries and companies has shifted from the use of fossil fuels as an energy source to renewable energies (Sterner, 2019). In 2021, 42.4 % of the German electricity came from renewable energy sources. With the increasing number of green energies, the variability of technologies based on wind or solar caused by changes in weather can impact the electricity grid. Achieving a reliable power delivery to satisfy changing supply and demand is crucial for a modern economy (Ausfelder et al., 2017). This calls for strategies to ensure stability of the grid during times of low power production or high demand. Apart from a sophisticated grid structure, the storage of energy is important. Whilst there are already established approaches, such as mechanical (pumped hydroelectric storage), thermal (sensible heat storage) or chemical (hydrogen in gaseous or liquefied form), most of them are geo-limited or require adverse operating conditions (*e.g.*, high pressure, high and low temperature, etc.) (Koohi-Fayegh, 2020). In recent years the interest in hydrogen has increased, with many nations striving towards a hydrogen economy. Figure 1 gives an overview of the different storage technologies currently available for hydrogen.

To determine the best suited material for hydrogen storage, certain properties, such as the gravimetric and volumetric hydrogen density, decomposition temperature, reusability and production cost, need to be evaluated and taken into account. To this end, the U.S. Department of Energy (DOE) has set targets for solid hydrogen storage used in cars, albeit these are not restricted to the automotive sector. Those main goals are (Rather, 2020):

- Gravimetric energy density: 1.5 kWh/kg system (4.5 wt.% hydrogen)
- Volumetric energy density: 1.0 kWh/L system (0.030 kg hydrogen/L)



- Cost per stored energy: \$10/kWh (\$333/kg stored hydrogen capacity)

In addition, the materials should be acceptable from the point of view of health and safety and the environment.

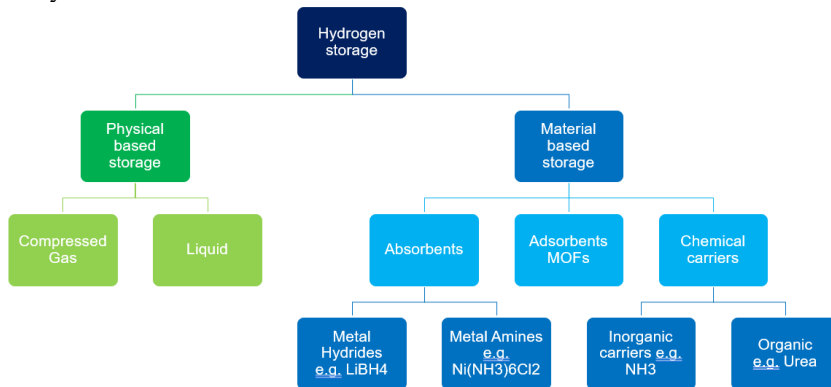


Figure 1: Overview of hydrogen storage technologies

In recent years the application of many different materials for solid energy storage has been investigated (Amiri and Shahbazian-Yassar, 2021). However, these new materials are lacking a process design concept to be applied at commercial scale. The focus of this work is the development of such a novel concept by means of the Aspen Plus® process simulator. It has been chosen over other process simulators due to its flexibility, extensive component and physical properties database, and advanced solids handling capabilities. These include an extensive library of physical properties (e.g., SOLIDS, ELECNRTL, etc.) and equipment models for solid processing (e.g., fluidized bed reactor, cyclone, classifiers, crystallizer, dryer, filter, centrifuge, etc.). Furthermore, it allows solids to take part in reactions, and detailed description of granular solids (e.g., particle size distribution, sub-streams of different solid phases, etc.). Another novel aspect of this work is the use the resources provided by Aspen Plus in the context of solid-solid separations and reactions beyond typical operations such as gasification.

## 2. Process design

### 2.1. Storage material

After investigating several options, an ammonia salt, Hexaaminenickel(II) chloride (HEX), is chosen as the energy storage material. A blue/violet solid, the component is an indirect hydrogen storage solution, capable of storing up to 44.1 wt.%  $\text{NH}_3$  (Kojima and Yamaguchi, 2020). Unlike the group of adsorbing materials, such as metal and covalent organic frameworks (MOF's/COF's) or intermetallic alloys and metal hydrides on the absorbing side, ammonia salts have received less attention in the pursuit of materials for solid energy storage. For the production of HEX, larger-scale production has not yet been proposed. The starting point is absorption by introducing  $\text{NH}_3$  gas to nickel chloride ( $\text{NiCl}_2$ ). HEX can additionally be prepared in liquid form by mixing aqueous solutions of  $\text{NiCl}_2$  and  $\text{NH}_4\text{Cl}$  with subsequent crystallisation (Leineweber and Jacobs, 2000). However, this reaction path is not chosen for the conceptual design discussed in this work, for it would require additional equipment like crystallisers and dryers, and, apart from being energy intensive, these units would not be used in other processing steps. When HEX is subjected to heat it releases  $\text{NH}_3$ , decomposing into Diaminenickel(II) chloride (DIA). Ultimately, with more heat added, DIA decomposes back to  $\text{NiCl}_2$ .

The reactions discussed above obey the following stoichiometry:

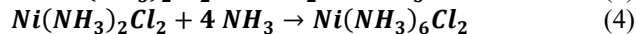
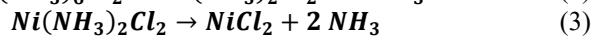
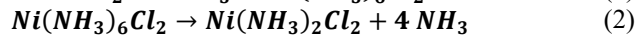


Table 1 shows the different reactions as well as their temperatures and kinetic parameters as they were determined by researchers at the Nagoya University of Japan (Kubota et al., 2014).

Table 1: Summary of the reaction steps, their reaction kinetics and conditions.

Step	Reaction equation	T [K]	k [-]	E <sub>A</sub> [kJ/mol]	n [-]
Preparation	(1)	323	$6.82 \times 10^5$	66.6	2.3
Desorption	(2)	383	$3.80 \times 10^4$	50.5	0
Complete desorption	(3)	413	NA	NA	NA
Absorption	(4)	303	$4.85 \times 10^0$	24.7	0

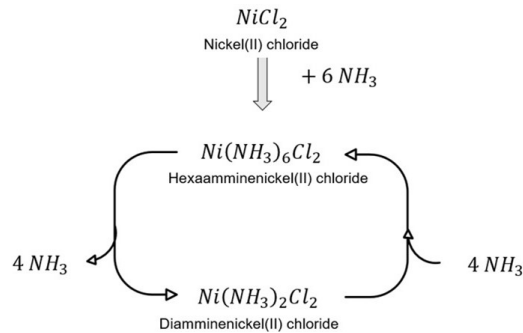


Figure 2: Concept of cycling between desorption (left) and absorption (right) to release and store ammonia after initial preparation from  $\text{NiCl}_2$ .

By cycling through the Absorption and Desorption steps, as shown in Figure 2,  $\text{NH}_3$  can be repeatedly released and stored. The Complete desorption step in Reaction (3) is not considered in the current work.

## 2.2. Process concept

The flowsheet for the proposed process concept based on the reaction pathways for the production and decomposition of HEX is illustrated in Figure 3.

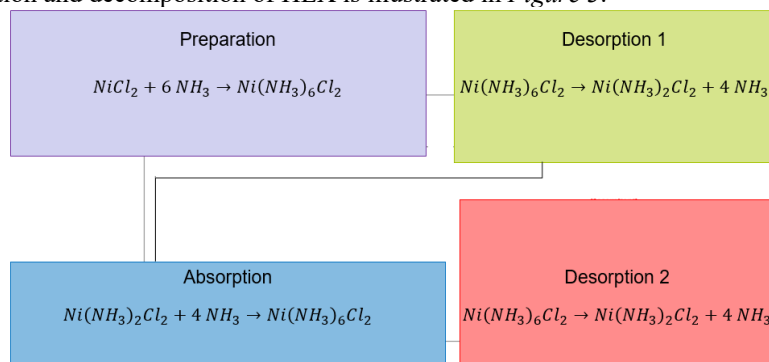
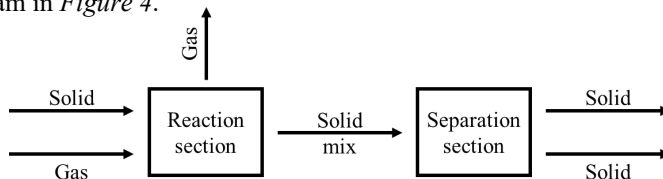


Figure 3: Design concept for storing and releasing  $\text{NH}_3$ .

The starting point is  $\text{NiCl}_2$ , assumed to be broadly available. Beginning with the initial Preparation step (purple), in which  $\text{NH}_3$  is stored as HEX, a first stage of releasing  $\text{NH}_3$ , in which DIA is produced – Desorption 1 (green) – is considered. The main storage and release cycle starts with an Absorption step (blue), in which  $\text{NH}_3$  reacts with DIA, and its release in the Desorption 2 step (red). In each of these four process steps considered, gas-solid reactions occur. In order to facilitate them, suitable equipment must be selected. To this end, the proposed concept considers, for each of the four steps described above, a Reaction section, in which solid and gas reactants come into contact, and a Separation section, in which solid-solid and solid-gas components are separated, as illustrated in the block diagram in *Figure 4*.



*Figure 4: Block diagram of the process concept stages.*

For the storage and release reactions, a fluidized bed reactor (FBR) is chosen. The liquid-like behaviour of the FBR eases the transport of solids and, due to its turbulent nature, eliminates the need for other mixing equipment. Furthermore, this turbulence favours heat transfer and leads to a uniform product (Reschetilowski, 2020). Since no conclusive particle size distribution (PSD) is available for the solid materials considered for this process (HEX and DIA, respectively), an assumption based on scanning electron microscope (SEM) imagery (Rejitha et al., 2011) is made.

Thus, it is assumed that all solid materials present in the system will have the same PSD. Moreover, to simplify further the process simulation, only one class of particles – of 50  $\mu\text{m}$  – is considered, meaning that all particles will have the same diameter. Hence, in the Separation section a separation of the solid materials based on density is proposed.

### 2.3. Process concept development procedure

The process concept design is developed in Aspen Plus v12. Initially, the storage of  $\text{NH}_3$  is simulated using a continuously stirred tank reactor (CSTR), with a volume determined for a lab size scale design. The starting point for the feed flowrates of reactants is selected using a sensitivity analysis. This was used in the selection of an efficient production target, under constraints of low corresponding feed flowrates of DIA and  $\text{NH}_3$ . Subsequently, these flowrates are minimized to still meet the selected production target. For the implementation of in Aspen Plus, some settings are already given by the process at hand, others are carefully picked and designed. This is done to ensure a compact reactor (in terms of size) for reduced material cost, as well as a low pressure drop to cut down on operational costs – while obtaining the maximum possible amount of the product.

Among the fixed settings, the possible desired and undesired reactions, the temperature of the favoured reaction (see Table 1), the fluidization behaviour and the PSD of the produced solids (assumed to be constant) need to be specified. Based on density and the PSD assumption, all solid materials are Geldart A solids (Geldart, 1973).

For determining the geometry of the FBR, an optimization is conducted with the objective of maximising the production of HEX, whilst minimizing the column. Since Aspen Plus cannot perform multiobjective optimizations, an equation is used to combine both targets using weights, with higher weighting assigned to the maximisation of the main product.

The reactor diameter, height and the solid discharge location are varied to find the best solution. To ensure a feasible result, some additional constraints are added. These include

defined targets for the HEX production, ratios for the FBR geometry, etc. To maintain a FBR typical shape, the height is to be bigger than the diameter.

In the Separation section, common equipment is used to perform the solid-solid and solid-gas separations. These processes are designed to be as sharp as possible, while maintaining industry standard flowrates and dimensions. The calculations are performed using the standard calculation methods of the respective unit (e.g., cyclone, sifter, etc.) implemented in Aspen Plus.

In contrast to the simulation of two phase separations of solids (gas-solid, liquid solid), the separation of multiple solids requires a different stream setup. Each solid component is assigned its own sub stream in order to facilitate different PSDs and enable the separation of the streams.

### 3. Results

The results are presented as rounded component units (CU) based on the molar flowrate of the initial reactant,  $\text{NiCl}_2$ . At the beginning of the Preparation stage, 1 CU of  $\text{NiCl}_2$  reacts with 4.6 CU ammonia, producing HEX and a small amount of DIA, following Reaction (1). In the following stage, Desorption 1, the HEX gets desorbed to DIA, following Reaction (2), and releasing 2.6 CU of  $\text{NH}_3$ . Subsequently, the DIA streams produced during the first and second stages are mixed together, resulting in a flowrate of 1.4 CU. The stream is then introduced in the Absorption stage, where it absorbs 4 CU of  $\text{NH}_3$ , producing HEX, following Reaction (4). In the final stage (Desorption 2), 1 CU of HEX releases 4 CU of ammonia, according to Reaction (2).

This marks the success of the proposed concept. Starting out from the broadly available  $\text{NiCl}_2$ , ammonia is stored and released in several stages. The temperature of each step is not exceeding  $120\text{ }^\circ\text{C}$  (393 K). The optimizations regarding the reactor geometry reveal the FBR to have a height – diameter ratio of 4.5, while the overall pressure drop of all reactors registers below 25 mbar. Thus, the simulated separation processes produced results that are within the range of the established guidelines.

In the resulting process simulation, the  $\text{NiCl}_2$  is able to store 52.50 % of its weight in ammonia. This is equivalent to 9.31 wt.% hydrogen, placing the concept within the guidelines (> 4.5 wt.% hydrogen) of the U.S. Department of Energy. Furthermore, the design is set up in such a way that the same equipment is used for both the Storage and Release steps, potentially reducing the equipment cost of the pilot plant.

The complete desorption to  $\text{NiCl}_2$  requires significantly higher heat compared to the desorption step of HEX to DIA. If only this part is to be considered, the storage capacity must then be based on the 1 CU of HEX entering the Desorption 2 stage. This results in 3.67 wt.% hydrogen being stored, fulfilling only 81.6 % of the target. However, the targets are set for the automotive industry. With the current reactor dimensions, its use in the automotive industry is not viable. The aim is to combine the proposed process in a power-to-ammonia (P2A) concept. Nonetheless, this integration of the solid storage with the utilization of green ammonia and ammonia to power processes (e.g., solid oxide fuel cell) is expected to result in additional losses, due to the conversion efficiency (Jeerh et al., 2021). Hence the DOE targets are formulated for the produced energy, which needs to be evaluated once all the various elements are integrated in the final process concept.

### 4. Conclusion

A novel process concept for solid energy storage has been introduced. In line with advances in hydrogen technology, ammonia as an indirect hydrogen carrier is used, considering safety aspects as well. As a storing compound, the metal ammine halide

Hexaamminenickel(II) chloride is chosen, due to its relatively high hydrogen density (7.81 wt.%) and kinetic data available in the public domain. The concept is developed around the cyclical storage (Absorption) and release (Desorption) of ammonia. For the gas solid reactions, a fluidized bed reactor is used. After the reaction, several separation steps of the different solids present in the reactor streams, as well as of solids and gas, are needed. The process design is set up to have four stages, each at a specific temperature, to favour the desired reaction. Extensive model-based analyses have been conducted in Aspen Plus to validate the proposed process concept. This has been shown to store an equivalent of 9.31 wt.% hydrogen. Furthermore, the operating conditions of the process are within a normal range, with temperatures below 120 °C, and the predominant pressure of one bar. Thus, no costly temperature resistant or pressure specific equipment is needed. In both the reaction and the separation sections, industry standard equipment can be used. This marks a step up from the analytical equipment used for investigating the storage materials and offers the possibility of scaling up in the later stage of process development. Improvements can be made for the simulation in Aspen Plus, as well as on the practical/experimental side of the process. Regarding the simulation further, optimization of the models can be done. In addition to fine-tuning flow rates and equipment in the case of the reactor, or using a custom model for the separation, switching to a dynamic simulation can bring further insight on the practical operation and stability of the system. Furthermore, information regarding the start-up and shut-down of the various cycles, as well as the control strategies can be investigated with the aid of a dynamic model that can be constructed in Aspen Plus Dynamics. Material investigations with a broader temperature and pressure ranges, either in a thermogravimetric analyser or a FBR, can improve on the reaction kinetics or could provide further knowledge on the process conditions.

## References

- A. Amiri and R. Shahbazian-Yassar, 2021, Recent progress of high-entropy materials for energy storage and conversion, *Journal of Materials Chemistry A* 9, pp. 782-823
- F. Ausfelder et al., 2017, Energy storage as part of a secure energy supply, *ChemBioEng Reviews* 4 (3), pp. 144-210
- D. Geldart, 1973, Types of gas fluidization, *Powder Technology* 7 (5), pp. 285-292
- G. Jeerh et al., Recent progress in ammonia fuel cells and their potential applications, *Journal of Materials Chemistry A* 9, pp. 727-752
- Y. Kojima and M. Yamaguchi, 2020, Ammonia storage materials for nitrogen recycling hydrogen and energy carriers, *International Journal of Hydrogen Energy* 45 (16), pp. 10233–10246
- S. Koochi-Fayegh, 2020, A review of energy storage types, applications and recent developments, *Journal of Energy Storage* 27, 101047
- M. Kubota et al., 2014, Absorption and desorption characteristics of NH<sub>3</sub> with metal chlorides for ammonia storage, *Journal of Chemical Engineering of Japan* 47 (7), pp. 542–548
- A. Leineweber and H. Jacobs, 2000, Preparation and crystal structures of Ni(NH<sub>3</sub>)<sub>2</sub>Cl<sub>2</sub> and of two modifications of Ni(NH<sub>3</sub>)<sub>2</sub>Br<sub>2</sub> and Ni(NH<sub>3</sub>)<sub>2</sub>I<sub>2</sub>, *Journal of Solid State Chemistry* 152 (2), pp. 381–387
- S.U. Rather, 2020, Preparation, characterization and hydrogen storage studies of carbon nanotubes and their composites: A review, *International Journal of Hydrogen Energy* 45, 4653-4672
- K. Rejitha et al., 2011, Thermal decomposition studies of [Ni(NH<sub>3</sub>)<sub>6</sub>]X<sub>2</sub> (X = Cl, Br) in the solid state using TG-MS and TR-XRD, *Journal of Thermal Analysis and Calorimetry* 103 (2), pp. 515–523
- W. Reschetilowski, 2020, *Handbuch Chemische Reaktoren*, Springer Spektrum Berlin, Heidelberg
- M. Sterner, 2019, *Handbook of energy storage: Demand, technologies, integration*, Springer-Verlag Berlin, Heidelberg

# Multiobjective optimization of distributed energy systems design through 3E (economic, environmental and exergy) analysis

Bogdan Dorneanu,<sup>a</sup> Sayeef Miah,<sup>b</sup> Evgenia Mechleri,<sup>b</sup> Harvey Arellano-Garcia,<sup>a</sup>

<sup>a</sup>*LS Prozess- und Anlagentechnik, Brandenburgische Technische Universität Cottbus-Senftenberg, Cottbus D-03046, Germany*

<sup>b</sup>*Department of Chemical and Process Engineering, University of Surrey, Guildford GU27XH, United Kingdom*

## Abstract

Distributed energy systems (DES) are promising alternative to conventional centralized generation, with multiple financial incentives in many parts of the world. Current approaches focus on the design optimization of a DES through economic and environmental cost minimization. However, these two criteria alone do not satisfy long-term sustainability priorities of the system. The novelty of this paper is the simultaneous investigation of economic, environmental and exergetic criteria in the modelling of DES through the two most commonly used solution methodologies for solving multi-objective optimization problems – the weighted sum and the epsilon-constraint methods. Out of the set of Pareto optimal solutions, a best-compromised solution is chosen using the fuzzy-based method. Numerical results reveal reduction of around 93% and 89-91% in environmental and primary exergy input, respectively.

**Keywords:** Multiobjective optimization, Distributed energy systems, Exergy, Mixed-integer linear programming, Fuzzy-based methods.

## 1. Introduction

Due to the commitments set through the Paris Agreement, many countries are implementing plans to reduce greenhouse gas emissions relative to the 1990 baseline. To facilitate this, the utilization of renewable energy systems (DES), as well as a shift from centralized to distributed generation using distributed energy systems plays a key role (Clarke et al., 2021). Thus, a lot of work has been focused in recent years on the development of models that can optimally design the layout of DES, which show significant savings in cost and a low carbon impact (Sidnell et al., 2021). However, considering only the economic and environmental cost of the DES does not satisfy the long-term sustainability priorities. The evaluation of the impact that exergy has on the DES must be considered, as well to make the rational use of available energy (Zhao et al., 2022).

Di Somma et al. (2015) first study the use of exergy analysis within DES, where a model for the optimal operational strategy of a DES is proposed, with a multi-objective whereby cost and exergy efficiency are considered. In terms of optimal DES layout, the focus is on minimization of annual operating and investment costs (Yang et al., 2017, Sidnell et al., 2021). Additionally, environmental cost, as carbon emissions or carbon tax, is taken as a second objective function alongside annual cost in a multi-objective approach (Jing et al., 2018, Karmellos & Mavrotas, 2019). The majority focus on two objective functions,

with a single study considering annual cost, CO<sub>2</sub> emissions and electrical system unavailability (Wouters et al., 2017). However, the three objective functions are never simultaneously solved.

The novelty of this paper is twofold, with the first being the investigation of exergy with DES design optimization through a multi-objective approach, whilst considering economic and environmental cost, thus simultaneously minimizing three objectives in the context of DES. The second is the use and comparison of two most commonly used methodologies for solving multi-objective optimization problems, namely the weighted sum (WSM) and the epsilon constraint method (ECM).

## 2. Methodology

Multiple technologies are available for the generation and use of energy, such as photovoltaic panels (PV), natural gas boilers, combined heat and power (CHP) units, or the national grid. A multiobjective optimization problem (MOP) is developed with the aim to optimally select the type and number of technologies within the DES, as well as the layout of the heating pipeline network, and the operation of the system.

The economic objective is equated to the total annual cost, made up of the annualized investment and operational costs, respectively. The environmental objective is equal to the total carbon emissions multiplied by the carbon tax. Finally, the exergetic objective is equal to the primary exergy input to the DES.

The resulting model is applied to a calendar year split into 24 different periods – 6 periods per day for 4 representative seasons.

### 2.1. Problem formulation

The structure of the DES with all the available technologies is shown in Figure 1. The three most common technologies are chosen to complement the national grid – CHPs, PVs and gas fired boilers.

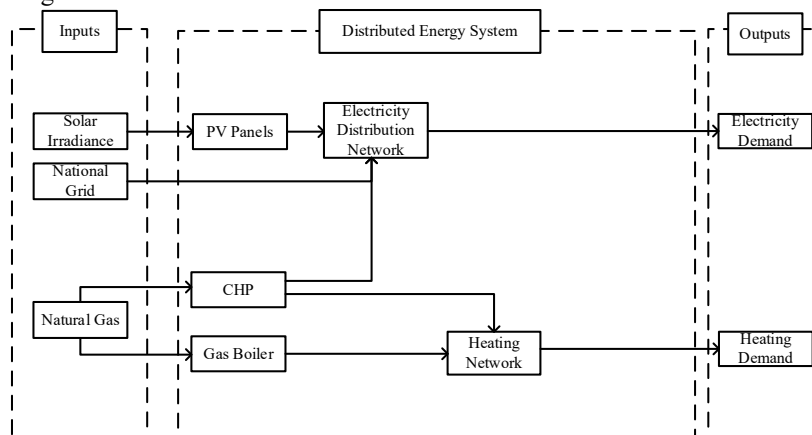


Figure 1: Energy flows of the DES

In detail, the national grid, PV panels and CHP units meet the customer electricity demand, while the heating demand is satisfied by gas fired boilers and CHP units.

The *economic objective* is the minimization of the total annualized cost of the DES,  $C_{ANNUAL}$ , which includes the investment cost,  $C_{INV}$ , the operational and fuel cost,  $C_{OP}$ , the annual cost for purchasing energy from the grid,  $C_{PUR}^{GRID}$ , as well as the income from selling energy back to the grid  $C_{SAL}^{GRID}$ , and the feed-in tariffs,  $FIT$ . All costs are in British Pounds (£).

$$C_{ANNUAL} = C_{INV} + C_{OP} + C_{PUR}^{GRID} - C_{SAL}^{GRID} - FIT^{PV} - FIT^{CHP} \quad (1)$$

The *environmental objective* is the minimization of environmental cost of the DES,  $C_{ANNUAL}$ , which contains carbon emissions arising from the combustion of natural gas within the CHP units and gas boilers, as well as the carbon emissions related to the national grid:

$$C_{ENV} = CT \cdot \left\{ \sum_{i,m,p} [E_{i,m,p}^{GRID} \cdot CI_{ELEC} \cdot day(m) \cdot hours(p) \cdot season(m)] + \sum_{i,m,p,k} \left[ \frac{(E_{i,m,p,k,SAL}^{CHP} + E_{i,m,p,k,SELF}^{CHP})}{\eta_k^{CHP}} \cdot CI_{GAS} \cdot day(m) \cdot hours(p) \cdot season(m) \right] + \sum_{i,m,p} \left[ \frac{H_{i,m,p}^B}{\eta^B} \cdot CI_{GAS} \cdot day(m) \cdot hours(p) \cdot season(m) \right] \right\} \quad (2)$$

Where  $CI_{ELEC}$  and  $CI_{gas}$  are the carbon intensities of the national grid and natural gas [ $kg_{CO_2}/kWh$ , respectively], while  $CT$  is the carbon tax [ $£/kg_{CO_2}$ ].

The *exergetic objective* is the maximization of the exergy efficiency of the DES,  $\varphi$ , given as the total exergy output,  $Ex^{out}$  [kWh], divided by the primary exergy input,  $Ex^{in}$  [kWh] (Di Somma et al., 2015):

$$\varphi = \frac{Ex^{out}}{Ex^{in}} \quad (3)$$

The total annual exergy output is the annual exergy required to meet the customer electricity and heating demand, as given by:

$$Ex^{out} = \sum_{i,m,p} [CLoad_{i,m,p}^{ELEC} \cdot day(m) \cdot hours(p) \cdot season(m)] + \sum_{i,m,p} [CLoad_{i,m,p}^{HEAT} \cdot F_{i,m,p}^{HEAT}] \quad (4)$$

Where  $CLoad$  refers to the customer demand at a given month and period, while  $F_{i,m,p}^{HEAT}$  is the Carnot factor.

The primary exergy input rate of the DES consists of the exergy from the national grid,  $Ex^{GRID}$  [kWh], natural gas,  $Ex^{NG}$  [kWh] and solar energy,  $Ex^{PV}$  [kWh], given by:

$$Ex^{in} = Ex^{GRID} + Ex^{NG} + Ex^{PV} \quad (5)$$

Constraints related to the energy balance, design and operation of the various technologies (PV, CHP, heating pipeline network), and interaction with the national grid as described in Clarke et al. (2021) are added to the model.

## 2.2. Optimization methods

Two of the most commonly used optimization methodologies for solving multiobjective linear problems (MOLP), the WSM and the ECM are utilized.

Applying the *WSM* results in a single objective function that is a linear combination of the three objective function in Eqs. (1)-(3):

$$Weighted_{obj} = \alpha \cdot \frac{C_{ANNUAL}}{C_{ANNUAL}^{REF}} + \beta \cdot \frac{C_{ENV}}{C_{ENV}^{REF}} + \gamma \cdot \frac{Ex^{in}}{Ex^{in,REF}} \quad (6)$$

Where  $\alpha$ ,  $\beta$  and  $\gamma$  are weighting factors that translate to the relative importance of each objective function; their sum should be equal to 1.



As all three objectives have different orders of magnitude, the value of each is divided by its respective reference value. The reference is equal to the value of the cost or exergy input when considering pure optimization for the corresponding objective term ( $\alpha = 1$ ,  $\beta = 1$ , or  $\gamma = 1$ , respectively). The Pareto frontier is obtained by varying all three weighting factors such that the sum of all weights remains equal to 1.

For the ECM, the improved version of the augmented version, AUGMECON2 is used. In AUGMECON2, a single function is chosen as the main objective to be optimized whilst the remaining are applied as constraints (Mavrotas & Florios, 2013). By varying the right-hand side of the constraints parametrically, the Pareto optimal solutions are found. For the MOLP presented here, the annual cost – Eq. (1) – is chosen as the main objective function, whilst Eq. (2)-(3) are applied as constraints.

The solution of the multiobjective problem results in a set of Pareto efficient solutions, all of which are said to be optimal. To choose the best solution out of the Pareto set, a fuzzy based method is used (Yalcin & Erginel, 2011).

### 2.3. Case study

As an illustrative example, the model is applied to a cluster of 5 houses in Bristol, UK. Figure 2 illustrates the system, with available technologies for each house, as well as the heating pipeline network.

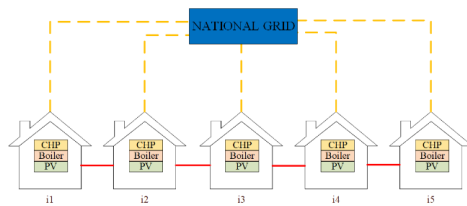


Figure 2: Layout of the DES

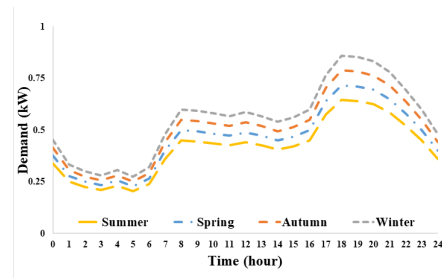


Figure 3: Electricity demand profile

Electricity usage for each house is a percentage of the UK average based on 250 houses and is illustrated in Figure 3. Heating demand is calculated using the heating degrees hour method applied to air temperature recorded by the Filton weather station (Met Office, 2006). Likewise, hourly solar irradiance is obtained from radiation observations recorded by the Liscombe weather station (Met Office, 2006).

Heating and electricity loads are distributed across the five houses (80%, 90%, 100%, 110% and 120%), such that the average of the loads is equal to the baseline.

## 3. Results and discussion

The problem formulated in Section 2 is implemented using the Branch and cut algorithm and the CPLEX solver in GAMS 25.1.2 on a PC with an AMD FX-8350 4 GHz processor and 8 GB of RAM. The full calendar year is represented as 6 hourly periods per day for 4 representative seasons: winter (December – February), spring (March – May), summer (June – August) and autumn (September – November).

In case of the WSM, Four scenarios are represented on the Pareto set (Figure 4a), namely optimization with high preference to annual cost ( $\alpha = 0.8, \beta = 0.1, \gamma = 0.1$ ), high preference to environmental cost ( $\alpha = 0.1, \beta = 0.8, \gamma = 0.1$ ), high preference to primary exergy input, and equal preference to all objectives ( $\alpha = 0.33, \beta = 0.33, \gamma = 0.33$ ). The total time taken to run all four scenarios to full optimality is 11 minutes.

Using AUGMECON2 at full optimality with all three objective functions active for the ECM returns a set of Pareto optimal solutions (Figure 4b) is 1 hour and 20 minutes with weakly optimal and repeated results avoided.

The best-compromised solutions obtained by applying the fuzzy based method are summarized in Table 1, while the layout of the resulting networks are illustrated in Figure 5. For comparative purposes, the above results are compared to a conventional design whereby there are no DER technologies within the neighborhood, nor does a heating pipeline network exist. Instead, the electricity demand is fully met by the national grid, while the heating demand is satisfied by gas-fired boilers.

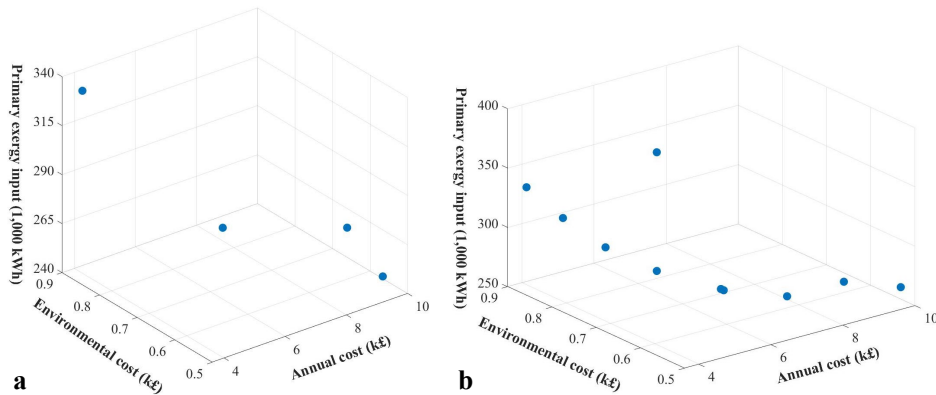


Figure 4: Pareto optimal set for (a) WSM and (b) ESM

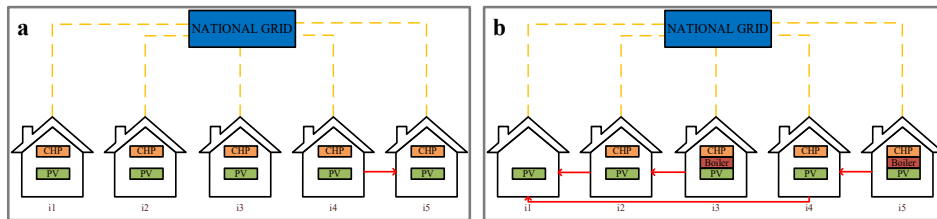


Figure 5: DES layout for (a) WSM and (b) ECM

Table 1: Best compromised solutions

Design	Economic cost (£)	Environmental cost (£)	Primary exergy input (kWh)
Conventional	1,237	7,699	2,481,000
WSM	10,930	519.3	213,000
ECM	9,836	552.8	287,000

In case of the WSM, the best-compromised solution reveals considerable reduction in both environmental cost (93%) and primary exergy input (91%). However, the annual cost increases by 784% due to the capital and operational cost of using the distributed

energy resources technologies. A similar behavior is observed by applying the ECM, as seen in Table 1.

#### 4. Conclusions

This paper investigates the impact of exergy on the design and operation of a DES via a multiobjective optimization approach that includes annual economic and environmental costs, and considering both the weighted sum and the epsilon-constraint methods for the solution of the resulting MOLP. A best-compromised solution is produced using the fuzzy-based method. The proposed procedure is applied on a case study, showing significant reductions (of around 90%) in both environmental and primary exergy input for both approaches. Furthermore, the results show that the use of exergy analysis is a powerful tool towards the rational use of fossil fuels.

Future work will consider the addition of more renewable generation technologies, such as wind turbines, biomass boilers, heat pumps or solar thermal collectors. Additionally, a microgrid, which enables the exchange of electricity between the dwellings, will be included in the network design. These will provide further understanding on the exergy analysis within DES.

The application of different methods for the selection of the best-compromised solutions, such as the min-max (Kparib et al., 2019) or the R-method (Rao et al., 2021) and the comparison with the fuzzy-based method in the context of DES design optimization will also be investigated.

#### References

- M. Di Somma et al., 2015, Operation optimization of a distributed energy system considering energy costs and exergy efficiency, *Energy Conversion and Management* 103, 739-751
- F. Clarke et al., 2021, Optimal design of heating and colling pipeline networks for residential distributed energy systems, *Energy* 235, 121430
- R. Jing et al., 2018, A multi-objective optimization and multicriteria evaluation integrated framewrok for distributed energy system optimal planning, *Energy Conversion and Management* 166, 445-462
- M. Karmellos, G. Mavrotas, 2019, Multi-objective optimization and comparison framework for the design of distributed energy systems, *Energy Conversion and Managemetn* 180, 473-495
- D.Y. Kparib et al., 2019, A min-max strategy to aid decision making in a bi-objective discrete optimization problem using an improved ant colony algorithm, *American Journal of Operations Research* 9, 161-174
- G. Mavrotas, K. Florios, 2013, An improved version of the augmented  $\varepsilon$ -constraint method (AUGMECON2) for finding the exact Pareto set in multi-objective integer programming problems, *Applied Mathematics and Computation* 219, 9652-9669
- Met Office, 2006, MIDAS: Global radiation observations and UK daily temperature data, NCAS British Atmospheric Data Centre
- R.V. Rao et al., 2021, Ranking of Pareto-optimal solutions and selecting the best solution in multi- and many-objective optimization problems using R-method, *Soft Computing Letters* 3, 100015
- T. Sidnell et al., 2021, Optimal design and operation of distributed energy resources systems for residential neighbourhoods, *Smart Energy* 4, 100049
- C. Wouters et al., 2017, A policy-based multi-objective optimization framework for residential distributed energy system design, *Renewable Energy and Environmental Sustainability* 2, 5
- G.D. Yalcin, N. Eginel, 2011, *Determinioning weights in multi-objective linear programming under fuzziness*, London
- H. Zhao et al., 2022, Economy-environment-energy performance evaluation of CCHP microgrid system: A hybrid multi-criteria decision-making method, *Energy* 240, 122830

# Calcium looping for CO<sub>2</sub> capture, H<sub>2</sub> and electricity coproduction in coal fired power plants

B. Basant Kumar Pillai, Ramsagar Vooradi, Sarath Babu Anne\*

*National Institute of Technology, Warangal, TG State - 506004, India*

*\*sarat@nitw.ac.in*

## Abstract

In recent years, global warming has become one of the most important concerns in the world. To achieve net zero carbon emissions by 2050, fossil fuel-based power plants must be integrated with suitable carbon capture and sequestration/utilization technologies. Currently, more attention is given to the looping technologies for post combustion CO<sub>2</sub> capture as it has the advantage of low efficiency penalty. In this work, a coal fired power plant configuration is integrated with double calcium looping system for CO<sub>2</sub> capture, production of electricity and hydrogen. The overall assessment of the developed configuration is carried out on the basis of energy, exergy and environmental parameters and compared with conventional coal power plant configuration. The thermal integration strategy adopted in the configuration of double calcium looping and gasification integrated power plant contributes an energy and exergy gain of 1.4% and 1.2% respectively as compared to conventional coal power plant configuration.

**Keywords:** Double calcium looping, coal fired power plant, carbon capture and sequestration, hydrogen production, energy and exergy analysis.

## 1. Introduction

Energy plays a very critical role in the socio-economic development of any country. At present, a major part of this energy demand is fulfilled by fossil fuels such as coal, oil, and natural gas. However, these fossil fuels are also a major source of Green House Gas (GHG) emissions (mostly CO<sub>2</sub>) that cause climate change. Among all the fossil fuels, coal is widely used around the world. India's coal consumption in 2018 contributes to around 45% of the country's total energy consumption and is the second largest coal consuming country after China. While coal is fully combusted for power generation, it can also be converted into H<sub>2</sub> through the gasification process. It has been reported that around 27% of all the hydrogen in the world is produced by coal gasification alone (Hydrogen from coal., 2021).

Researchers are currently working on developing different carbon capture technologies that can be integrated with coal-based power generation and hydrogen production systems. Second generation CO<sub>2</sub> capture technologies are being developed to further reduce the energy penalty and economic costs. Globally, calcium looping has been proven to be a viable method of capturing CO<sub>2</sub> in a number of successful demonstration projects (Hanak et al., 2015). It is one of the attractive sustainable 2<sup>nd</sup> generation CO<sub>2</sub> capture options due to its ability to retrofit with the existing plants with lower economic and energy penalties.

Over the last decade, most of the research in calcium looping has been focused mainly on enhancing its feasibility to capture CO<sub>2</sub>. Several intensification strategies have been proposed in the last decade for reducing calcium looping technology's energy penalty

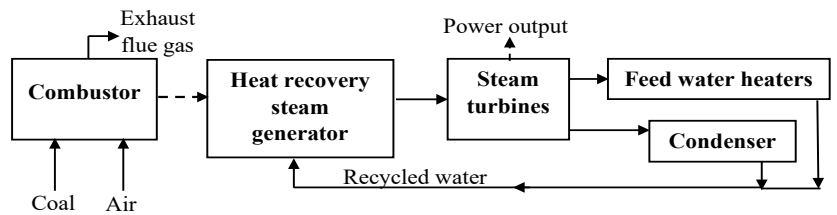
(Duan et al., 2016). Some studies have also been carried out to co-produce hydrogen along with CO<sub>2</sub> capture. Chen et al., 2011 proposed a calcium looping based novel reactor configuration involving a compact fluidized bed for CO<sub>2</sub> capture and hydrogen cogeneration. The parametric study revealed that the hydrogen production rate and purity were reduced when pressure was increased from 1 bar to 10 bar. Shaikh et al., 2022 proposed a biomass-based calcium looping configuration for power generation and hydrogen production to achieve a net zero or negative CO<sub>2</sub> emission. However, these configurations require pure O<sub>2</sub> to function and thus needs an Air Separation System (ASS) that consumes high energy. In this aspect, the Ca-Cu looping and double calcium looping configuration are some of the distinctive configurations that completely eliminates the use of ASS and thus have the potential to reduce the energy penalty significantly (Diego et al., 2016). An advantage of double calcium looping over Ca-Cu looping is its ability to operate with single sorbent (calcium) that is widely available at low cost in many countries. Over the years, several research studies have been conducted on the integration of carbon capture with hydrogen and electricity generation (Lin et al., 2002; Florin and Harris., 2008; Dean et al., 2011). However, there is still a scope to enhance performance through effective integration by utilizing double calcium looping schemes. Further, limited studies are available in the literature on overall energy, exergy, and environmental assessment with high ash coal.

Therefore, the main aim of the present research work is to propose a novel double calcium looping and coal gasification integrated power plant configuration (HDCFPP) for power and hydrogen coproduction along with CO<sub>2</sub> capture. A conventional coal fired power plant (CFPP) and double calcium looping integrated coal fired power plant (DCFPP) are also developed in this research work. This is done to demonstrate the performance of the proposed HDCFPP against DCFPP and CFPP configurations. Both configurations are evaluated based on energy, exergy and environmental parameters.

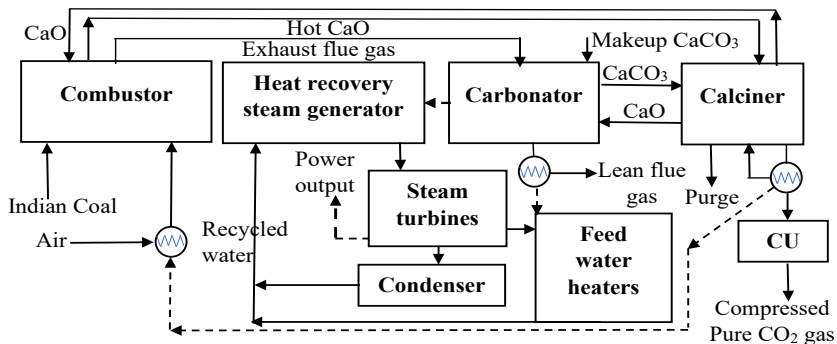
## 2. Methodology

A coal-fired power plant with operating conditions of 450 kg/s of live steam (at 590 °C, 300 bar) and a double calcium looping process are developed and validated against the literature models using aspenONE software (Reddy et al., 2013; Diego et al., 2016). Using the validated process models, a CFPP, DCFPP and HDCFPP configuration are synthesized as shown in Figure 1. Thermodynamic and environmental analyses are carried out to investigate the performance of the proposed HDCFPP against DCFPP and CFPP configurations. As a part of the study, the following characteristics of Indian high ash coal are considered: carbon 39.16%, ash 48.87%, oxygen 7.92%, hydrogen 2.76%, sulfur 0.51%, nitrogen 0.78 and high heating value 15.83 MJ/kg. The assumptions considered to simplify the analyses of this research work are:

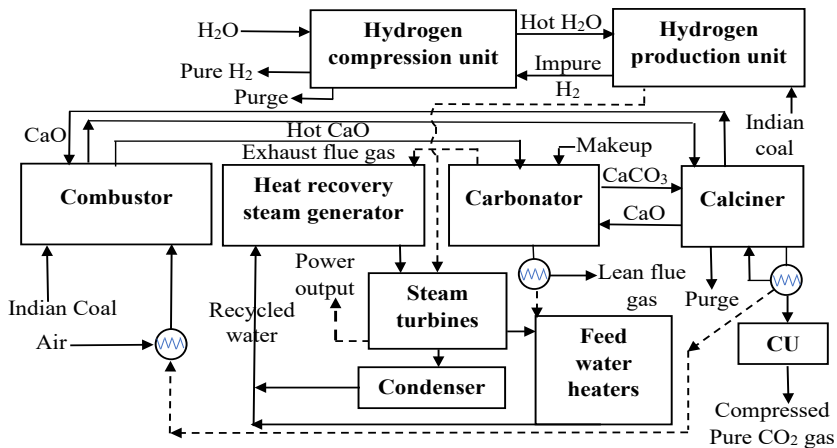
1. Process is operating in steady state.
2. Thermal loss and pressure drop in pipes (streams) are not considered.
3. Around 1.5% of input fuel energy is lost through convection and radiation in the combustor.
4. The auxiliary power consumption ( $W_{aux}$ ) is assumed to be 9% of the gross power output of the power plant that includes lighting and control systems, transformer losses, heating, ventilation and air conditioning systems etc.
5. The overall plant life of all the configurations are 25 years and operate for 7008 hours per year.



(A)



(B)



(C)

Figure 1: Layout of (A) CFPP, (B) DCFPP and (C) HDCFPP configurations

### 2.1. Coal fired power plant (CFPP)

The construction of a supercritical CFPP based on the Rankine cycle developed in aspenONE is shown in Figure 1A. The low-ash coal feed is replaced with Indian (high-ash) coal, and the operating parameters are considered in accordance with the Indian climatic conditions (Suresh et al., 2010).

### 2.2 Double calcium looping integrated coal fired power plant (DCFPP)

The double calcium looping process consists of two calcium loops instead of one in the conventional calcium looping process as shown in Figure 1B. The primary loop between the calciner and carbonator is used to separate CO<sub>2</sub> from the hot flue gas. The secondary loop between the combustor and calciner is used to transfer the required heat

energy from the combustor to the calciner to carry out the calcination reaction. This configuration helps to avoid the need of ASS for calcium looping system and ensures high CO<sub>2</sub> capture efficiency.

### 2.3 Hydrogen production and double calcium looping integrated CFPP (HDCFPP)

A calcium looping gasification-based hydrogen production unit (HPU) is integrated with double calcium looping CFPP for producing concentrated H<sub>2</sub> along with CO<sub>2</sub> capture and electricity generation as shown in Figure 1C.

### 2.4 Thermodynamic and environmental analysis

The proposed configurations are assessed based on thermodynamic and environmental parameters. The techno-economic analysis of calcium looping processes has been carried out by considering key performance indicators such as energy efficiency, exergy efficiency, CO<sub>2</sub> capture efficiency and specific CO<sub>2</sub> emissions. The plant efficiency  $\eta_{ele}$  on the basis of net electrical energy output is estimated using Equation 1. In this  $W_{aux}$ ,  $W_{ASS}$ ,  $W_{pump}$ ,  $W_{comp}$  and  $W_{gross}$  are sum of all auxiliary power consumptions, energy consumption by ASS, centrifugal pumps, air compressor and gross power output produced by the turbines respectively. The EXERGYMS property set provided by aspenONE is used to estimate exergy of all streams (Olaleye et al., 2015). Using this exergy data, the exergy efficiency  $\psi_n$  and exergy destruction ratio ( $\lambda_n$ ) of a unit is calculated with the help of Equations 2 and 3. In these equations,  $Ex_{in}$ ,  $(Ex_{dest})_n$  and  $(Ex_{out})_n$  are the exergy input, exergy destruction and exergy output flowrates for each single unit  $n$ , respectively and  $\Sigma Ex_{dest}$  represents the overall exergy destruction of the configuration. Specific CO<sub>2</sub> emission is assessed in terms of kg CO<sub>2</sub> emitted per MWh of net energy generation. CO<sub>2</sub> capture efficiency is calculated as the ratio of CO<sub>2</sub> captured and CO<sub>2</sub> generated during coal combustion and hydrogen production.

$$\eta_{ele} = \frac{W_{gross} - W_{comp} - W_{pump} - W_{ASS} - W_{aux}}{m_{fc} \times HHV_{fc}} \quad (1)$$

$$\psi_n = \frac{(Ex_{out})_n}{(Ex_{in})_n} = 1 - \frac{(Ex_{dest})_n}{(Ex_{in})_n} \quad (2)$$

$$\lambda_n = \frac{Ex_{dest,n}}{\Sigma Ex_{dest}} \quad (3)$$

## 3. Results and Discussion

In this section, the performance of the proposed DCFPP and HDCFPP configurations are analyzed and compared with the conventional CFPP on the basis of overall energy, exergy and environmental parametric indicators. Table 1 represents the overall comparative evaluation based on energy and exergy. The higher amount of input energy and exergy in the HDCFPP configuration is because of the additional coal used for the production of hydrogen. CFPP, DCFPP, and HDCFPP configurations are found to generate 515.3 MW, 507.8 MW, and 476.31 MW of net electrical power, respectively. The net energy and exergy efficiencies of DCFPP configuration are 35.4% and 32.4%. The DCFPP configuration has very low energy and exergy penalties as compared with CFPP configuration. Further, its performance is much better than the MEA integrated Indian coal fired power plant (Karmakar et al., 2013) having an energy efficiency of 28.7% only. The performance of the proposed HDCFPP is found to be thermodynamically better than DCFPP. The results revealed that the HDCFPP configuration has an energy and exergy gain of 1.4% and 1.2% respectively as compared to the CFPP due to electric power and hydrogen cogeneration.

CO<sub>2</sub> emissions reduction is also one of the primary objectives of this work, so an evaluation is performed based on environmental parameters such as CO<sub>2</sub> capture efficiency, specific CO<sub>2</sub> emissions, and lifetime CO<sub>2</sub> emissions. A CO<sub>2</sub> capture efficiency of 91.05% is used as a benchmark in DCFPP and HDCFPP for environmental analysis. From the environmental analysis, it is observed that the HDCFPP configuration has the lowest specific CO<sub>2</sub> emission of 78.57 kg/MWh as compared to DCFPP and CFPP configurations that are having emissions of 82.94 kg/MWh and 885.17 kg/MWh respectively. It is estimated that the lifetime CO<sub>2</sub> emissions of the CFPP configuration is around 7,99,16,195 Mg, whereas the DCFPP and HDCFPP configurations have considerably lower emissions of 73,79,424 Mg and 77,95,699 Mg, respectively.

Table 1 Plant performance of CFPP, DCFPP and HDCFPP based on energy and exergy analysis

Parameters	Energy Analysis			Exergy Analysis			
	CFPP	DCFPP	HDCFPP	Parameters	CFPP	DCFPP	HDCFPP
Input (MW)	1432	1432	1511	Input (MW)	1565	1565	1651
Gross power output (MW)	590	591	570				
Electricity consumption (MW)	75.25	83.85	94.06		75.25	83.85	94.06
Net electricity output (MW)	515.3	507.8	476.3		515.3	507.8	476.3
Energy from H <sub>2</sub> (MW)			89.18	Exergy from H <sub>2</sub> (MW)			87.99
Net electrical efficiency (%)	35.98	35.46	31.51	Net electrical efficiency (%)	32.92	32.44	28.84
Net energy efficiency (%)	35.98	35.46	37.41	Net exergy efficiency (%)	32.92	32.44	34.16
Energy efficiency penalty (%)		0.52	-1.43	Exergy efficiency penalty (%)		0.48	-1.24
Energy consumption by CU (MW)			55.3			52.75	55.3

The exergy analysis for individual units/modules is performed to investigate the thermal losses due to process irreversibilities. The analysis not only identifies the location of losses but also helps to determine the maximum scope of improvement possible in the system. Figure 2A shows the exergy efficiency of main units/modules and Figure 2B shows the improvement potential of the same. The DCaL represents a complete module of combustor, calciner and carbonator and SPGU represents the module of HRSG, steam turbines, condenser and FWH. The results also revealed that the air preheater (HE3), CU and HCU although holds quite low exergetic efficiency, the scope of exergetic improvement potential in these respective units is found to be very less.



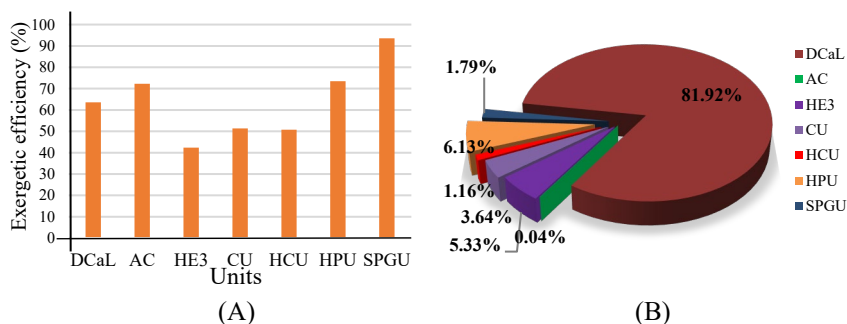


Figure 2: (A) Exergy efficiency (B) Improvement potential of HDCFPP configuration

#### 4. Conclusions

The conventional CFPP configuration is integrated with double calcium looping and gasification units for CO<sub>2</sub> capture and hydrogen production. The energy and exergy analyses reveal that the net plant efficiency of the proposed DCFPP configuration is higher than the conventional MEA integrated power plant reported in the literature. Extending this DCFPP for hydrogen production by integrating it further with calcium looping gasification process enhanced the process performance. The HDCFPP configuration has an energy and exergy gain of 1.4% and 1.2% respectively as compared to the CFPP due to electric power and hydrogen cogeneration. The proposed schemes can be adopted for sustainable electricity and hydrogen cogeneration in the current crisis due to global warming and can be used as a bridging measure until the green technology becomes a more feasible option in the future.

#### References

- VS. Reddy, SC. Kaushik, SK. Tyagi, 2013, Clean technologies and environmental policy, 15, 1, 133-145.
- SY. Lin, Y. Suzuki, H. Hatano, M. Harada, 2002, Energy conversion and management, 1, 43(9-12), 1283-1290.
- S. Karmakar, MV. Suresh, AK. Kolar, 2013, International journal of green energy, 10, 10, 1011-1025.
- S. Chen, D. Wang, Z. Xue, X. Sun, W. Xiang, 2011, International journal of hydrogen energy, 36, 8, 4887-4899.
- NH. Florin, AT. Harris, 2008, Chemical engineering science, 1, 63(2), 287-316.
- ME. Diego, B. Arias, JC. Abanades, 2016, Journal of cleaner production, 117, 110-121.
- L. Duan, T. Feng, S. Jia, X. Yu, 2016, Energy, 115, 942-953.
- IEA, Key World Energy Statistics, IEA, Paris. 2020, <https://www.iea.org/reports/key-world-energy-statistics-2020>.
- Hydrogen from coal, Coal Age, 2021, <https://www.coalage.com/features/hydrogen-from-coal/>.
- G. Fytianos, S. Ucar, A. Grimstvedt, A. Hyldbakk, HF. Svendsen, HK. Knuutila, 2016, International journal of greenhouse gas control, 46, 48-56.
- DP. Hanak, EJ. Anthony, V. Manovic, 2015, Energy and environmental science, 8, 8, 2199-2249.
- CC. Dean, J. Blamey, NH. Florin, MJ. Al-Jeboori, PS. Fennell, 2011, Chemical engineering research and design, 1, 89(6), 836-855.
- AR. Shaikh, Q. Wang, L. Han, Y. Feng, Z. Sharif, Z. Li, J. Cen, S. Kumar, 2022, Sustainability, 14, 4, 2189.
- AK. Olaleye, M. Wang, G. Kelsall, 2015, Fuel, 151.

## 3-E (Energy, Exergy & Ecological) analyses of Chemical Looping Air Separation integrated Oxy-Coal Combustion based Power plant

Agarwal Shruti Sumankumar, Venkata Suresh Patnaikuni, Praveen Kumar Bommineni

*National Institute of Technology Warangal, Warangal – 506004, India*  
*Email of the Corresponding Author: [pvsuresh@nitw.ac.in](mailto:pvsuresh@nitw.ac.in)*

### Abstract

Chemical looping Air Separation (CLAS) is an efficient and economic oxygen supply method for integration into oxy-fuel combustion plants for producing power with CO<sub>2</sub> capture. The present work deals with comprehensive analysis in terms of energy, exergy, and ecological parameters of the CLAS integrated oxy-coal plant using high ash coal and Indian environmental conditions. In this study, steady-state simulation of (i) conventional supercritical powerplant without capture (ii) oxy-coal combustion plant with conventional ASU, and (ii) CLAS integrated oxy-coal plant using manganese oxide as oxygen carrier, has been carried out. Comparative analysis for all three plants conducted based on 3-Es (energy, exergy, and ecological) shows that net efficiency of CLAS-based plant is higher by 7.05% and 8.36% respectively compared to conventional CFPP without capture and ASU based plant. The study concluded that CLAS integrated oxy-coal plant is energetically, energetically, and ecologically efficient compared to its counterparts.

**Keywords:** CO<sub>2</sub> capture; Chemical looping air separation; CLAS; Oxy-coal combustion

### 1. Introduction

Global warming caused by greenhouse gases is a global concern because of its detrimental impact on climate change. Emissions of CO<sub>2</sub> are increasing because of rising energy demands. Coal-fired power plants (CFPPs) are still significant sources of power generation in many countries and have become primary sources of CO<sub>2</sub> emissions. With almost a trillion tonnes of known reserves worldwide, coal is the most commonly used fossil fuel. Around 40% of world's electricity is produced using coal-based plants. The conventional CCS need to contribute to roughly 20% of emission reductions across all sectors to achieve a global reduction of 50% in CO<sub>2</sub> emissions by 2050. But the existing post combustion capture methods are expensive. Hence, there is need to create advanced, affordable CO<sub>2</sub> capture technology. Chemical looping, one of the recent technological options, can aid in adoption of low-emission technologies and provide advantages in applications involving production of fuels, chemicals, and power. Oxy-fuel combustion is one of the efficient ways to reduce CO<sub>2</sub> emission, but conventional air separation units (ASUs) are usually energy intensive and result in high energy penalties (Shah et al., 2012). Chemical looping air separation (CLAS) is a potential alternative method to separate oxygen from air. This process consists of an oxygen carrier that releases oxygen in reduction reactor and absorbs oxygen in oxidation reactor (Qing et al., 2020).

Some of the simulation and thermodynamic studies on CLAS are reviewed here. Manganese, cobalt, and copper oxides were found to be the most suited carriers for CLAS

procedure (Moghtaderi, 2010; Shah et al., 2012). Shah et al., 2013a studied the impact of flue gas pollutants on oxygen carriers in integrated CLAS and reported that addition of flue gas pollutants raises the operating temperature. However, this method can produce  $O_2$  at lower costs compared to cryogenic ASUs (Shah et al., 2013b). The production of electricity and  $CO_2$  from the oxy-fuel combustion of lignite was studied by Tagliaferri et al., 2018. Three oxy-fuel IGCC power plants were simulated by Shi et al., 2019 using various ASUs to provide power systems with almost no  $CO_2$  emissions. Aspen Plus software was used to simulate CLAS based on a dual fluidized bed with Mn-based oxygen carrier by Cao et al., 2019. Qing et al., 2020 determined the impact of recycling position for oxy-combustion power plants integrating CLAS.  $SiO_2$ ,  $ZrO_2$ ,  $TiO_2$ , and  $MgAl_2O_4$  were used as inert binders for Cu-based oxygen carriers by Wang et al., 2019 and determined stability of these composites in a packed bed facility.

As observed from the literature, thermodynamic evaluation of CFPPs based on high-ash coal and sophisticated steam boilers is scarce. A comprehensive study involving 3-E (energy, exergy, and ecological) assessments is not available for direct comparison of CLAS based CFPPs with SupC steam boiler parameters using coal, even though a few studies on energy and exergy analyses are available on CLAS as discussed above. Such a study is crucial while choosing the best technology in meeting increasing clean energy demand, especially in developing nations, where coal is the primary energy source. Main goal of this work is to present a comprehensive theoretical analysis of the viability of integrating CLAS in SupC coal-based power plants using high ash coal. Performance of the CLAS integrated oxy-coal plant is also compared against that of conventional CFPP without capture and oxy-coal plant with conventional ASU for oxygen supply.

## 2. Methodology

In this study, steady-state process simulation of three types of coal-fired power plants was carried out using aspenONE software – (i) conventional SupC coal-fired power plant without capture (ii) oxy-coal combustion plant with conventional ASU, and (iii) CLAS integrated oxy-coal fired power plant using manganese oxide as the oxygen carrier. The first two plants were adopted from Suresh et al., 2010 and Suresh et al., 2011, while flowsheet of the third plant was developed by integrating the CLAS into conventional oxy-coal combustion plant. The first two plants were simulated, validated and considered as reference plants in the current study for comparing performance of the synthesized CLAS integrated oxy-coal fired power plant. All the three plants were simulated for the same fuel energy input of 1628.91 MW based on HHV (equivalent to the coal input of 102.9 kg/s). Indian coal with high ash content having the following characteristics was considered as the fuel in this study - carbon 39.16%, ash 48.87%, oxygen 7.92%, hydrogen 2.76%, sulfur 0.51%, nitrogen 0.78 and high heating value 15.83 MJ/kg. The conventional fluids were defined by stream classes MIXED, whereas conventional and unconventional solids were defined by the stream classes CIPSD and NCPDS, respectively. For estimation of properties of solids, air, flue gases, steam, and water, Peng Robinson and Boston Mathias (PR-BM), Soave Redlich Kwong (RKS), ideal gas, and STEAMNBS were used. Comparative performance analysis for all three plants was conducted based on the 3-Es (energy, exergy, and ecological). Following is the brief description of each of the above plants.

### 2.1. Description of conventional SupC CFPP without $CO_2$ capture

A typical CFPP consisting of a combustor, feed water heaters (FWHs), condenser, steam turbine units, sizing and drying, and heat recovery steam generation

(HRSG) was considered with feed-water temperature of 280°C and steam parameters of high, intermediate and low-pressure steams of 242.2 bar/537°C, 42.0 bar/565°C, and 2.9 bar/215.6°C, respectively (Suresh et al., 2010). Three high-pressure FWHs and four low-pressure FWHs make up this plant. After being crushed and dried with air present, coal was put into the combustor chamber. In the HRSG unit, steam is produced and heat is recovered from the hot flue gas produced by the combustor. The model built in the present work was validated using operational data for 660MW plant reported by Suresh et al., 2010. No CO<sub>2</sub> capture was considered in this reference case.

### 2.2. Description of SupC oxy-coal combustion plant with conventional ASU

Sizing and drying unit, Air separation unit, combustor unit, HRSG, FWHs, condenser units, CO<sub>2</sub> compressor and storage units are the key parts of this plant. The number of high-pressure and low-pressure FWHs for these plants is the same as that of the traditional CFPPs. The steam parameters (temperature and pressure) are used as 290 bar/600°C, 61 bar/620°C, and 3 bar/211.3°C, respectively, for HP, IP, and LP steams as specified in Suresh et al., 2011. With a final feed-water temperature of 305°C, the facility also includes a single stage of reheating. The oxygen required for combustion is obtained from air separation unit which consists of two stage compression with intermediate cooling of air and then air goes through cryogenic column for separation of air. This plant is simulated and validated with the results of the 800 MW conventional SupC oxy-fuelled power plant reported by Suresh et al., 2011.

### 2.3 CLAS integrated oxy-coal fired power plant

This plant flowsheet is synthesized by replacing the conventional ASU in the oxy-coal combustion plant described in section 2.2 with CLAS by ensuring proper heat integration. Additionally, this flowsheet is connected with the CO<sub>2</sub> capture and compressor unit. Manganese oxide (Mn<sub>2</sub>O<sub>3</sub>) is used as the oxygen carrier for the CLAS process. CLAS unit contains heat exchanger (heatX and heater), two Gibbs reactors for oxidation and reduction reaction, two stream splitters (Ssplit), one mixer, one splitter and flash separator to remove water. Air and oxygen carrier first enter oxidation reactor and then after reaction (eq. 1), it goes to solid separator where the reduced air is removed and goes through the heat exchanger to heat the inlet air. The solid goes to reduction reactor and reduces (eq. 2) in the presence of steam. Solid is then separated in the separator and again recycled whereas oxygen with water stream goes through flash separator to remove water. Pure oxygen is then mixed with recycled flue gas before it enters the combustor unit.

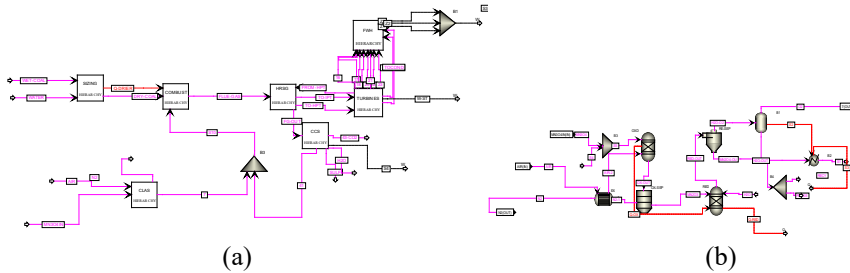
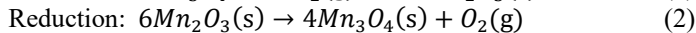
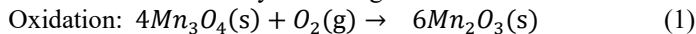


Fig. 1. Aspen plus flow diagram for (a) the CLAS integrated coal-fired power plant and (b) detailed flowsheet of CLAS hierarchy block

### 2.4 3-E analyses

Assessment of the environmental impact is crucial for any power production system's evaluation, in addition to the thermodynamic analysis. Hence, in the current

study, the 3-E (energy, exergy and ecological) analyses were carried out to examine the performance of all the cases. Below is a brief presentation of the performance indicators for each of these analyses (Surywanshi et al, 2019).

#### 2.4.1 Energy analysis

The gross energy efficiency ( $\eta_{Gross}$ ) and the net energy efficiency ( $\eta_{Net}$ ) are calculated by the ratio of the plant's gross or net power output ( $W_{Gross}$  or  $W_{Net}$ ) to the energy input (which is the product of coal mass flow rate ( $\dot{m}_{coal}$ ) and HHV of the coal)

$$\eta_{Gross} = \frac{W_{Gross}}{\dot{m}_{coal} \times HHV \text{ of the coal}} \quad (3) \quad \eta_{Net} = \frac{W_{Net}}{\dot{m}_{coal} \times HHV \text{ of the coal}} \quad (4)$$

#### 2.4.2 Exergy analysis

To measure the exergy flow rate of specific streams, EXERGYFL property set is added to aspenONE. Exergy destruction and efficiency of each unit in the plant are calculated using eq. 5 & 6 apart from calculating net exergy efficiency of overall plant (eq.7)

$$Ex_d = \sum Ex_{in} - \sum Ex_{out} \quad (5) \quad \varepsilon_{unit} = \frac{\sum Ex_{out}}{\sum Ex_{in}} = 1 - \left( \frac{Ex_d}{\sum Ex_{in}} \right) \quad (6)$$

$$\varepsilon_{Net} = \frac{\text{Net power output}}{\dot{m}_{coal} \times \text{Specific exergy of the coal}} \quad (7)$$

#### 2.4.3 Ecological analysis

An yearly  $CO_2$  emission as the product of 7000 operating hours per unit of fuel ( $\varepsilon_f^a CO_2$ ) is calculated using eq. 8 as ratio of the yearly energy intake to the power plant in ( $E_{chf}^a$  in  $MW.h$ ) and the annual  $CO_2$  emissions from plant into the atmosphere ( $\dot{m}_{CO_2}^a$  emit in kg)

$$\varepsilon_f^a CO_2 = \frac{(\dot{m}_{CO_2}^a \text{ emit})}{(3.6 \times E_{chf}^a)} \quad (8) \quad CO_2 \text{ capture efficiency} = \frac{(\dot{m}_{CO_2} \text{ captured})}{(\dot{m}_{CO_2} \text{ prior to capture})} \times 100 \quad (9)$$

### 3. Results and Discussion

Parametric analysis was conducted to fix operating parameters for the synthesized CLAS integrated Oxy-coal plant, while other plants were analysed with conditions given in the reference plants (Suresh et al., 2010, 2011). Operating pressure of oxidation and reduction reactors was maintained at 1.01325 bar, and other parameters of the CLAS integrated plant obtained from sensitivity analysis were as follows: air flow rate – 632.4 kg/s; reduced carrier flow rate – 4090 kg/s, temperatures of reduction & oxidation reactors as 850°C & 640°C; steam flow rate, temperature and pressure– 1260 kg/s, 980°C, 1.01325 bar.

#### 3.1 Energy analysis

The three plants were simulated for constant fuel energy input of 1628.91 MW (based on HHV). The power output, power consumption, and gross & net plant efficiency for each plant were calculated and are summarised in Table 1. Gross power output of CLAS integrated plant was more than that of other two plants. Improved steam parameters at superheater and reheater's outlet increased steam turbine output of ASU based and CLAS based oxy-coal plants compared to that of the conventional CFPP, which boosted up the gross power production. The net power consumption in ASU based plant was almost 3 times higher than that of conventional CFPP, while it was 1.68 times the traditional CFPP's power consumption in case of CLAS integrated plant. Subsequently, the net efficiency of all the three plants was calculated. The net energy efficiency of CLAS integrated plant was found to be higher than that of the traditional CFPP and ASU based plant by 7.05% and 8.36% , while an energy penalty of 1.3% was observed in ASU based plant compared to traditional CFPP because of addition of cryogenic air separation unit.

3.2 Exergy analysis

CLAS based oxy-coal power plant was found to have higher net exergy efficiency (calculated using eq. 7) compared to the other two plants. The exergy efficiency of ASU based Oxy-coal plant was slightly lower than that of the conventional plant because of the greater exergy destruction in the CO<sub>2</sub> compressor and air compressor (5.52 bar in ASU vs. 1.04 bar in conventional CFPP) (not present in conventional CFPP). All the three plants were subjected to component-wise exergy analysis in order to identify the units that experience the most exergy loss or destruction. The % contribution to exergy destruction by significant units of CLAS based oxy-coal plant was shown in Figure 2. Due to higher operating conditions (T and P) and heat transfer processes, the boiler section accounted for more energy destruction than other units. Table 2 lists the exergy and energy losses, as well as the improvement potential (IP) of key units of the plant. Boiler with highest exergy destruction contributed to 81% of total exergy destruction of the plant, while other units showed less improvement potential even though their energy losses were high (for e.g., condensers, HPFWH, ST, CLAS etc. as can be seen from Table 3). These results showed that boiler and CLAS are the units to be focused to further improve overall plant efficiency.

Table 1. Energy analysis of conventional, oxy-coal and CLAS based oxy-coal plants

Parameter	Conv. SupC CFPP without capture	ASU based oxy-coal plant	CLAS integrated oxy-coal plant
Fuel input energy (MW)	1628.91	1628.91	1628.91
Steam Turbine output (MW)	659.53	702.77	716.06
Auxiliary loss (MW)	49.46	52.71	53.70
Gross power output (MW)	610.06	650.06	662.35
Net power consumption (MW)	81.08	259.27	136.66
Net power output (MW)	464.53	443.50	579.40
Gross electric efficiency (%)	37.45	39.91	40.66
Net electric efficiency (%)	28.52	27.21	35.57

Table 2. Energy loss, exergy destruction and exergetic improvement potential for the CLAS integrated SupC oxy-coal plant.

Unit/block	Energy loss (MW)	Exergy destruction (MW)	IP (MW)
Boiler	19.89	899.60	308.35
Steam Turbine	204.56	97.11	4.99
Condenser	808.20	21.67	13.07
LPFWH	171.82	39.45	16.17
HPFWH	211.30	14.62	0.42
Deaerator	26.05	21.37	5.45
Pumps	24.67	2.49	0.07
Compressors	36.11	12.47	2.70
CLAS	137.95	312.13	95.90

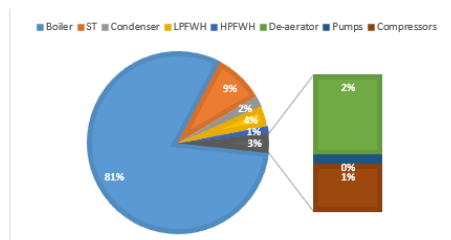


Fig. 2. Exergy destruction % contribution of key units of CLAS integrated oxy-coal plant

### 3.3 Ecological analysis

Apart from energy and exergy analyses, ecological analysis was carried out to see the effect of these plants on the environment. Table 3 provides comparison of different ecological parameters for all the three plants. Because of the absence of capture unit, the conventional CFPP displayed about  $3.27 \times 10^9$  kg annual CO<sub>2</sub> emissions into the atmosphere, while the other two plants are able to capture all the CO<sub>2</sub> in the flue gas. However, the net electric output per kg of CO<sub>2</sub> captured was found to be higher for CLAS integrated plant compared to ASU based plant because of its higher energy efficiency.

Table 3. Comparative ecological analysis of the three plants

Parameter	Conv. SupC CFPP without capture	ASU based oxy-coal plant	CLAS integrated oxy-coal plant
CO <sub>2</sub> emission (kg/s)	129.78	~0.00	~0.00
CO <sub>2</sub> captured (kg/s)	0.00	129.79	129.75
CO <sub>2</sub> capture efficiency (%)	0.00	~100	~100
Specific CO <sub>2</sub> emission (kg/MWh)	1005.77	0.01	0.00
Annual CO <sub>2</sub> emission (kg)	$3.27 \times 10^9$	~0.00	~0.00
Net electric output per kg of CO <sub>2</sub> captured (MW/kg)	-	3.41	4.47

## 4. Conclusions

The study demonstrated the feasibility of integrating chemical looping air separation in oxy-coal power plant utilizing high ash Indian coal for generating power with CO<sub>2</sub> capture. The performance of the integrated plant was assessed using 3-E (energy, exergy, and ecological) analyses and compared against that of conventional SupC CFPP and ASU based oxy-coal plant. The key conclusions of the study are: The net efficiency of the Mn<sub>3</sub>O<sub>4</sub>-based CLAS integrated SupC oxy-coal plant was higher by 7.05% and 8.36% compared to that of the conventional SupC CFPP and oxy-coal SupC plant with ASU. Exergy analysis revealed the highest exergy destruction rate and improvement potentials occurred in the boiler unit followed by CLAS reactors in case of the integrated plant. CLAS integrated oxy-coal power plant would eliminate the annual CO<sub>2</sub> release of  $3.27 \times 10^9$  kg from a 716 MW power plant (based on coal input). Based on the 3-E analyses study, it can be concluded that the CLAS based oxy-coal combustion fired power plant is energetically, exergetically and ecologically efficient.

## References

- K. Shah et al., 2012, Energy Fuels, 26, 4, 2038-2045.
- M. Qing et al., 2020, Energy, 206, 118136.
- B. Moghtaderi, 2010, Energy Fuels, 24, 1, 190-198.
- K. Shah et al., 2013a, Fuel, 103, 932-942.
- K. Shah et al., 2013b, Fuel, 107, 356-370.
- C. Tagliaferri et al., 2018, Chemical Engineering Research and Design, 131, 686-698.
- B. Shi et al., 2019, Energy Conversion and Management, 195, 290-301.
- Y. Cao et al., 2019, Energy Conversion and Management, 196, 286-295.
- M. Qing et al., 2020, Energy, 206, 118136.
- K. Wang et al., 2019, Powder Technology, 343, 40-48.
- MVJJ. Suresh et al., 2010, International Journal of Energy Research, 34, 716-735.
- MVJJ. Suresh et al., 2011, Journal of Engineering for Gas Turbines and Power, 133, 063001-1.
- G. Surywanshi et al., 2019, Energy Conversion and Management, 200, 112050.

# Reliability analysis for robust power-to-X design and storage sizing under renewable uncertainty

Jeongdong Kim<sup>1</sup>, Meng Qi<sup>1</sup>, Il Moon<sup>\*</sup>

<sup>a</sup> Department of Chemical and Biomolecular Engineering, Yonsei University, Seoul 03722, Republic of Korea

*1* These authors contributed equally to this study

## Abstract

During the operation of power-to-X system (PtX), the renewable profile strongly fluctuated according to its geographical and seasonal factors. Under the uncertainty of the renewable profile, the economic performance and renewable power penetration of the PtX can be enhanced by regulating two variables in design level; LCES size, renewable power plant scale. In this study, we firstly proposed the overall framework for data-driven-based reliability analysis of the PtX. First, using the pre-trained generator network in the GAN, the noise vector can be converted into the annual renewable profile. Given the generated renewable profiles, the scheduling model calculates the two performances; levelized cost of X production (LCOX), renewable power penetration. Integration of two formulated model enables us to infer the distribution of the performance via uncertainty propagation. During the propagation, the reliability analysis calculates the probability of failure (POF), which is defined as the constraint violation of renewable penetration. Conclusively, the overall framework can predict the LCOX distribution and calculate the POF of the system with constrained renewable penetration.

**Keywords:** Power-to-X, chemical energy storage, data-driven approach, power generation uncertainty

## 1. Problem statement

As shown in Fig. 1, the PtX process is composed of following unit: an electrolyzer, an intermediate hydrogen storage, and an electrified “X” production unit

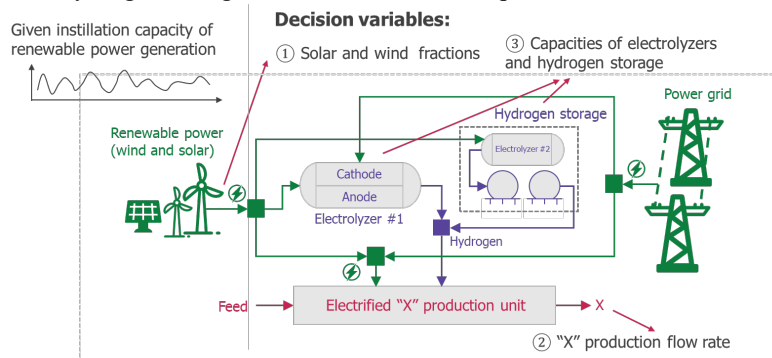


Fig. 1 Decision variables in PtX system.

Under the uncertainty of the renewable profile, the economic performance and renewable power penetration of the PtX can be enhanced by regulating design variables: installation



capacity of renewable power generation, solar and wind fractions, “X” production capacity, and hydrogen storage capacities.

Sustaining economic feasibility of the PtX system in design level, minimization of the production cost (LCOX) is important issue. Additionally, for decarbonization, the consideration of the constraining the grid energy penetration rate (GD) is also crucial factor in design level, and low GD rate prevent the additional carbon emission accompanied by the grid power utilization. In this situation, it is important to investigate the impact of the design variables on the performance while considering the uncertainty of time-series renewable profile.

Considering the uncertainty of time-series in PtX system was conducted via two main model: data-driven renewable profile generator, mathematical power allocation model. In case of the first model, the adversarial generative network (GAN) is firstly pretrained using the real historical data of wind and solar profile. After the training, Using the generator network in the GAN, the Gaussian noise vector can be converted into the annual renewable profile which follows the distribution of real historical data. Given the generated renewable profiles, the formulated scheduling model determines the hourly allocation of the renewable and grid power of the PtX in one-year period, and calculates the two performances; Levelized cost of X production (LCOX), annually averaged renewable power penetration. The linkage between the generator network and process scheduling model enables us to infer the probability density distribution of the PtX performance via uncertainty propagation of renewable profile. During the propagation, the reliability analysis calculates the probability of failure (POF), which is defined as the constraint violation of renewable penetration.

## 2. Methodology

### 2.1. Data-driven renewable profile generator

Due to the nature of intermittency and temporal dynamics of renewable, the model-based approach is not feasible option for uncertainty modeling of the renewable (Pinson et al., 2009; Lee et al., 2016). Thus, in this study, the profile generator is derived from the training GAN with real historical data of wind and solar (Chen et al., 2018).

The objective function  $V(G, D)$  of GAN training is defined as following:

$$\min_{\theta(G)} \max_{\theta(D)} V(G, D) = \underbrace{\mathbb{E}_{\tilde{x} \sim P_g} [D(\tilde{x})] - \mathbb{E}_{x \sim P_r} [D(x)]}_{\text{Critic loss}} + \lambda \underbrace{\mathbb{E}_{\hat{x} \sim P_{\hat{x}}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2]}_{\text{Gradient penalty}} \quad (1)$$

where  $\theta$  represents parameter,  $\tilde{x}$  is fake sampled from the noise  $P_g$ ,  $x$  is real sampled from the historical distribution  $P_r$ ,  $\lambda$  is gradient penalty coefficient,  $\hat{x}$  is the sampled from the uniform distribution  $P_{\hat{x}}$  formed by the  $P_g$  and  $P_r$

For GAN training, the NASA POWER Data, which provides application programming interface, is used to download the temporal data of wind speed and solar irradiance in certain time period and geographical site. In this study, the Jeju Island, Korea is considered as the geographical site. The noise vector of 1,000 dimensionality was considered, real data was sampled from the batch size of 25, and the training was conducted by 1,000 epochs. The training result of wind speed and solar irradiance based GAN were shown in Fig. 2.

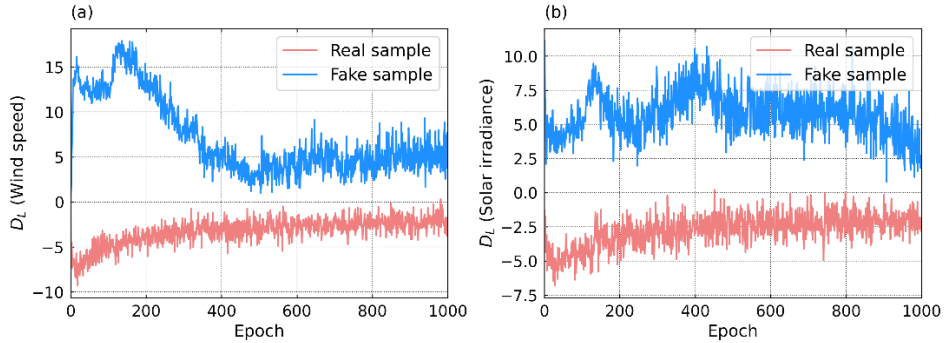


Fig. 2 Training result of GAN, (a) wind speed data, (b) solar irradiance data.

### 2.2. Formulating power allocation model

First, the wind and solar data are converted into power generation outputs by calculating the capacity factors for solar PV and wind turbines.

Given a installation capacity, the wind ( $P_W(t)$ ) and solar ( $P_S(t)$ ) power profiles are first defined at an hourly resolution. Then the total renewable power profile is calculated by  $P_{ren}(t) = P_W(t)f_W + P_S(t)f_S$ , where  $f_W$  and  $f_S$  are the fractions of wind and solar power. Assuming the production flowrate of “X” is  $F_x$ , the mismatch of power between the renewable and PtX power load,  $P_{mis}(t)$ , can be defined as:

$$P_{mis}(t) = P_W(t)f_W + P_S(t)f_S - P_{ptx} - P_{H2} \quad (9)$$

$$P_{ptx} = SP_{ptx}F_x \quad (10)$$

$$P_{H2} = F_{H2}F_x \quad (11)$$

$$F_{H2} = SP_{H2}\gamma_{H2} \quad (12)$$

where  $P_{ptx}$  and  $P_{H2}$  are power consumptions of the process and the electrolyzer, respectively, kW.  $SP_{ptx}$  and  $SP_{H2}$  are specific power consumptions of the “X” production process and the electrolyzer, respectively.  $F_{H2}$  is the required hydrogen flowrate, and  $\gamma_{H2}$  is the specific ratio of the required hydrogen flowrate to the “X” product flowrate.

Based on the mismatch profile of power, the mathematical model of power allocation can be formulated.

According to the formulated model, the time series profile of renewable and grid power, and storage profiles can be calculated as shown in Fig. 3.

Conclusively, given the generated scenario and design variables, the allocate power profile can be converted in to the important performance indicator: production cost, grid penetration, and thermal efficiency of the PtX system.

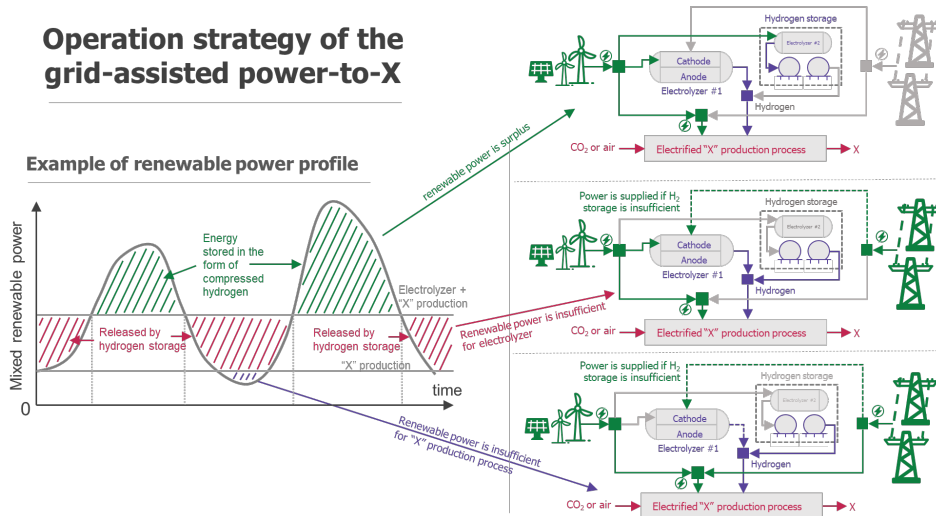


Fig. 3 Flexible operation strategy of the grid-assisted PtX process.

2.3. Uncertainty propagation model: PtX evaluation model

Using the formulated generator and power allocation model, the PtX evaluation model can be constructed as shown in Fig. 4. In the model, the noise vector firstly sampled from the multivariate standard normal distribution via the Monte-Carlo Sampling (MCS). The batch of noise sample transfer to the generator network, and finally converted in to batch dataset of annual profile of wind speed and solar irradiance. The mixture of renewable profile transfer to the power allocation model (operation model), and the model solves the power allocation problems via the mathematical approach, and converted into the production cost and grid penetration of each generated renewable profiles. Conclusively, using the PtX evaluation model, the uncertainty propagation of the complex renewable uncertainty can be simplified as the uncertainty of the noise with MCS. As a result, the PtX model can generate the distribution of the cost and penetration under the given design variables.

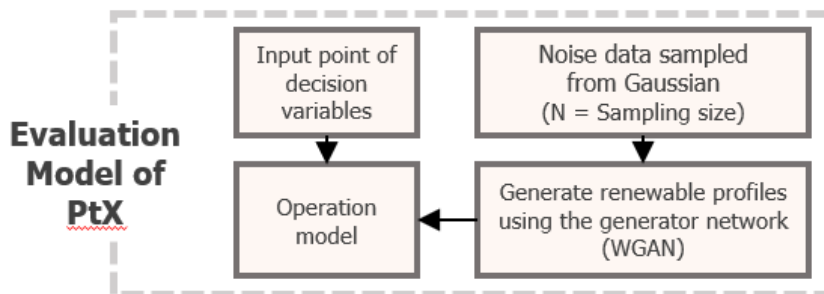


Fig. 4 proposed PtX evaluation model

3. Results and discussion

To test constructed PtX model, we conducted the uncertainty propagation in three different design variables, and the results are shown in Figs. 5 to 7.

As shown in the distribution, increasing the wind fraction effectively reduce the variance of the both production cost and grid penetration rate. Meanwhile, the range of production costs itself dramatically increased. The case study proves that the determined design variables has high impact on the probabilistic distribution of the PtX performance. Thus, using the constructed PtX model, the PtX design problem can be expanded into the robust optimization problem in future work.

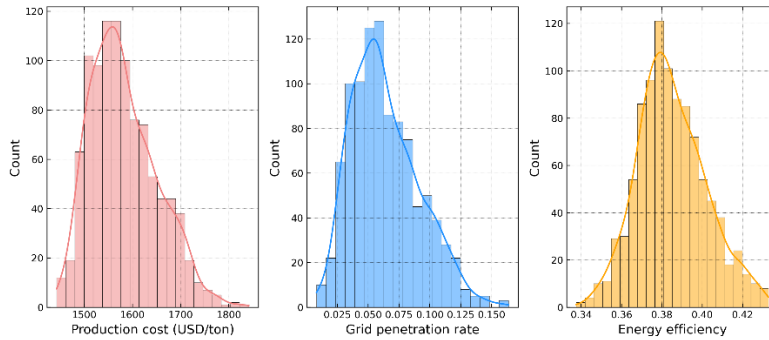


Fig. 4 Uncertainty propagation at wind fraction of 0.64, “X” flowrate of 1905, and hydrogen storage size of 38848

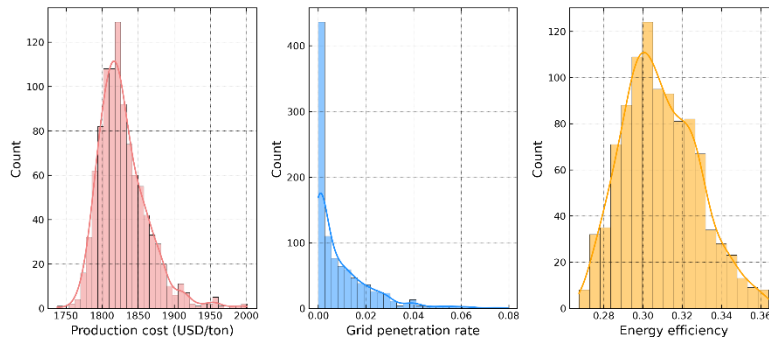


Fig. 5 proposed PtX evaluation model at wind fraction of 0.47, “X” flowrate of 1312, and hydrogen storage size of 40477

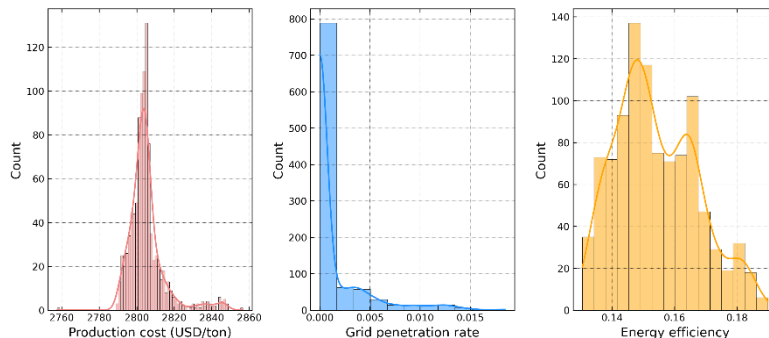


Fig. 6 proposed PtX evaluation model at wind fraction of 0.99 “X” flowrate of 716 and hydrogen storage size of 39656

**References**

- Zinson P, Madsen H, Nielsen HA, Papaefthymiou G, Klöckl B. From probabilistic forecasts to statistical scenarios of short-term wind power production. *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology*. 2009;12:51-62.
- Y. Brown, Lee D, Baldick R. Load and wind power scenario generation through the generalized dynamic factor model. *IEEE Transactions on power Systems*. 2016;32:400-10.
- Meng Y, Wang Y, Kirschen D, Zhang B. Model-free renewable scenario generation using generative adversarial networks. *IEEE Transactions on Power Systems*. 2018;33:3265-75

# Systematic Development of Strategies for the Decarbonization of Process Utility Systems

Julia Jiménez-Romero,<sup>a</sup> Adisa Azapagic,<sup>a</sup> Robin Smith,<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, The University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom*

## Abstract

This paper describes an optimization-based tool for developing a strategy for the design of sustainable industrial energy systems that can be constrained to reflect the evolution of the design and operation through time. The model determines the best utility system structure, energy mix, and seasonal operation for reducing CO<sub>2</sub> emissions while meeting seasonal demand for electricity, heat, and cooling. Furthermore, the framework includes the selection of the steam main operating conditions (pressure and temperature) as a design variable to exploit of the close interrelationship between the site processes and the utility system at multiple temperature levels, ensuring the design of efficient and cost-effective process utility systems.

The resulting multiperiod mixed integer nonlinear model is addressed by an extended version of the decomposition algorithm BEELINE, and a  $\epsilon$ -constraint approach is included to consider both economic and environmental performance. As a result, provide the decision-maker with several solutions for developing a sustainable pathway from current systems to future ones. The application to a case study from the literature shows that the framework could provide cost-effective guidelines for the energy transition of process utility systems.

**Keywords:** process utility systems, energy systems design, hybrid energy systems, multi-objective optimization, nonconvex mixed integer problem model.

## 1. Introduction

Most industrial facilities meet their heat and power needs with on-site utility systems. On the one hand, process utility systems are the largest consumers of industrial primary energy and the largest source of industrial CO<sub>2</sub> emissions, especially when powered by fossil fuels. On the other hand, process utility systems are an efficient solution for energy-efficient generation and distribution, site-wide heat recovery, and increasing energy supply reliability and adaptability. For this reason, both academia and industry are focused on the optimal design and operation of these types of systems. However, achieving a ‘optimal’ design is not straightforward, particularly for large-scale facilities, as the design is dependent on a number of factors, including the availability of energy conversion technologies, time-dependent energy demand and supply, utility prices (which can vary significantly from industry to industry), and the emission factor of technologies and sources. This has highlighted the need for specific decision-making tools to address this problem.

Many authors in the literature have addressed the optimization of energy system design and operation. Ganschinetz (2021) study provides a comprehensive literature review on this topic. Recent studies have emphasized energy security and environmental impacts, but energy system design and optimization are still driven by techno-economic objectives.

Environmental impacts are often only used as benchmarks or converted into economic terms, penalizing fossil fuel consumption and/or emissions. Although this strategy may lead to lower-carbon technologies, it does not allow the analysis of trade-offs between economic and environmental performance, which may overlook cost-effective options.

Notably, due to the mathematical complexity of maximizing energy supply efficiency by simultaneous consideration of energy integration and switching of fuel and/or technology, previous research has either neglected site-wide heat recovery opportunities or simplified the utility system configuration to the point where only a few utility components with fixed efficiencies have been considered. Moreover, critical practical issues such as equipment part-load performance and steam sensible heat effect are usually overlooked, resulting in misleading targets and thus incorrect analysis of the utility system design and its associated emissions (Jiménez-Romero, Azapagic, & Smith, 2023).

For this reason, the objective of this work is to provide a decision support framework capable of considering a wide range of technology options, its configuration and operation, while ensuring a sustainable energy efficient site. The model, in particular, considers not only the optimal configuration of the utility system but also the heat integration potential, through the optimization of steam mains operating conditions (temperature and pressure) to exploit site wide heat recovery and power generation trade-offs, ensuring that the design is energy efficient.

## 2. Problem statement

An on-site utility system supplies heat and power to several independent plants in the same industrial cluster. The utility system can generate steam from excess heat from one plant to use in another to improve site energy efficiency. Since heat can be recovered/used at a wide temperature range, the best steam distribution system operating conditions (temperature and pressure) must be chosen to balance fuel consumption and power generation.

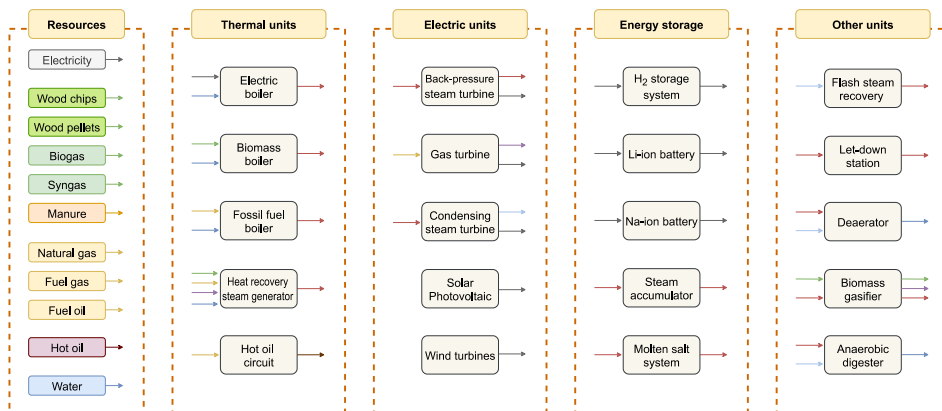


Figure 1. Schematic of the utility components and resources

To meet energy demands, the framework considers a broad technology portfolio (electrode boilers, biomass boilers, gas turbines, wind turbines, steam turbines, solar photovoltaic, among others) as well as electric and thermal energy storage units, as shown in Figure 1. Other utility components include deaerators, let-down stations, furnaces, and

flash steam recovery (FSR). Furthermore, the design framework takes operational issues (such as part-load performance and limits, start-up constraints) and practical constraints into account (i.e. steam temperature limitations, steam sensible and latent heat).

The optimal configuration and operation of the process utility system is determined by two criteria: total annualized costs and global warming potential (CO<sub>2</sub> equivalent emissions) (Vaskan, Guillén-Gosálbez, Turkay, & Jiménez, 2014).

### 3. Optimization Approach

The resulting optimization problem is a nonconvex mixed integer nonlinear programming (MINLP). The proposed model incorporates three sources of difficulty in terms of:

- Variable type. Decision variables for equipment selection and operation/activation as well as continuous variables for equipment size and load, steam mass flowrates and enthalpies
- Constraint type. Nonlinearities result from equipment cost functions (economy of scale), environmental impact qualification, and the selection of steam main operating condition (the integration of steam temperature/enthalpy as a design variable, introducing bilinear terms in the energy balance and some equipment performances as well as the need of functions to estimate the steam properties)
- Objective function. The economic and environmental performance are conflictive objectives.

Consequently, the problem could be computationally intensive when applied to practical size problems. For this reason, the following solutions have been adopted to approach the problem.

#### 3.1. Dealing with non-linearity

Non linearities caused by equipment costs functions and qualification of environmental impact are handled by piecewise affine approximations of the functions. On the other hand, non-linearities introduced by steam temperature as a design variable (and the calculation of steam properties) are addressed by developing a linear approximation for the accurate definition of superheated conditions.

In order to solve this, the BEELINE strategy presented in Jimenez-Romero, Azapagic, and Smith (2022) strategy comprises a bilevel decomposition of the problem, where the master problem is a relaxed version of the mixed-integer linear program (MILP). In the MILP problem the value of the binary variables are defined, to then re-optimize the continuous variables considering the non-linearities in a non-convex linear (NLP) program model.

#### 3.2. Dealing with the objective function

##### 3.2.1. Economic objective—minimizing annualized total cost

The economic objective is to minimize the total annual costs (TAC), including the total annualized investment cost ( $C^{\text{inv}}$ ), maintenance cost ( $C^{\text{main}}$ ), operating cost ( $C^{\text{op}}$ ) and start-up cost ( $C^{\text{start}}$ ).

$$\min \text{TAC} = C^{\text{inv}} + C^{\text{main}} + C^{\text{op}} + C^{\text{start}} \quad (1)$$



### 3.2.2. Environmental objective—minimizing GWP (CO<sub>2</sub>-eq·y<sup>-1</sup>)

The environmental objective is to minimize the total annual greenhouse emissions (TAGWP). The TAGWP comprises the emissions generated by (i) the resources (E<sup>re</sup>), (ii) installation and decommissioning of equipment (E<sup>eq</sup>), and (iii) operation (and waste disposal) of the utility system (E<sup>op</sup>) – including purchased electricity –, as expressed by Eq (2).

$$\min \text{TAGWP} = E^{\text{re}} + E^{\text{eq}} + E^{\text{op}} \quad (2)$$

To determine the Pareto-optimal curve and define the trade-off between economic and environmental performance, the  $\varepsilon$ -constraint approach (Mavrotas, 2009) is employed. The  $\varepsilon$ -constraint approach begins by establishing minimum and maximum values for the objective functions TAC and TAGWP, which are determined using the lexicographic optimization. Once defined the upper and lower bounds of the objective functions, the rest of Pareto-optimal points are obtained by solving the proposed model with the objective of minimizing TAC, while TAGWP is converted to a parametrized equality constraint as shown in Eq. (3)

$$\min: \text{TAC} + \delta \left( \frac{s_2}{r_2} \right) \quad (3)$$

$$\text{TAGWP} + s_2 = \varepsilon_2$$

$$r_2 = \text{TAGWP}_{\max} - \text{TAGWP}_{\min}$$

$\delta$  is a small number (on the order of  $[10^{-6}, 10^{-3}]$ ),  $s_2$  is a nonnegative slack variable for the objective function TAGWP.  $r_2$  is a parameter, which is equal to the difference between the upper and lower bound of TAGWP. To obtain the set of Pareto solutions the parameter  $\varepsilon_2$  is used for asses different interval points between the minimum and maximum emissions.

## 4. Case study and results

The proposed approach was applied to case study of a five-plant petrochemical complex by Sun, Doyle, and Smith (2015). To take into account energy demand variability over time, the production profile across the year is assumed to be similar to that provided for the petrochemical plant in Bungener, Hackl, Van Eetvelde, Harvey, and Marechal (2015) study. A comparison with optimized designs with fixed steam main conditions (noted as reference design) will be performed, to demonstrate the benefit of the methodology. The operating conditions of steam mains described by Sun et al. (2015) are used as reference. Table 1 summarizes the main findings.

Table 1. Main Results of Test Case Optimization

Parameters	Economic objective		Environmental objective	
	Reference design	Optimized design	Reference design	Optimized design
TAC [m€ y <sup>-1</sup> ]	70.89	59.17 (-16.5%)	294.28	177.65 (-39.6%)
TACWP [kt CO <sub>2</sub> -eq y <sup>-1</sup> ]	524.58	384.68 (-26.7%)	22.40	11.38 (-49.2%)

As shown in Table 1, optimizing the system configuration while also considering the optimal steam main conditions results in significant reduction not only in terms of TAC but also in terms of GHG emissions (TAGWP). This is because taking advantage of the trade-offs between heat integration and power generation allows for a reduction in overall energy demand. In turn, this can enhance the results of decarbonization when low-carbon technology is implemented.

Figure 2 shows the Pareto curve for the bi-objective optimization problem. Each point corresponds to one sub-problem solved to minimize total costs subject to a certain upper bound of GHG emissions. Figure 3 depicts how the selection and configuration of energy conversion technology varies from the current optimal design (point 0) as GHG emission limits become more restrictive.

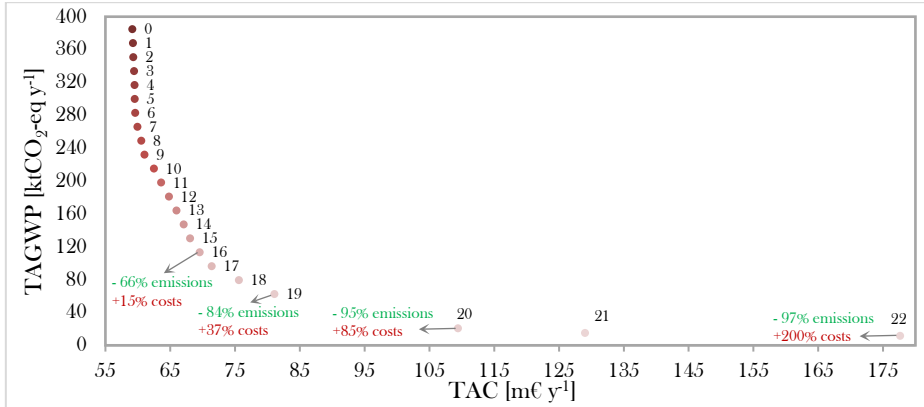


Figure 2 Pareto front: Total annualized emissions (TAGWP) vs total annualized costs (TAC)

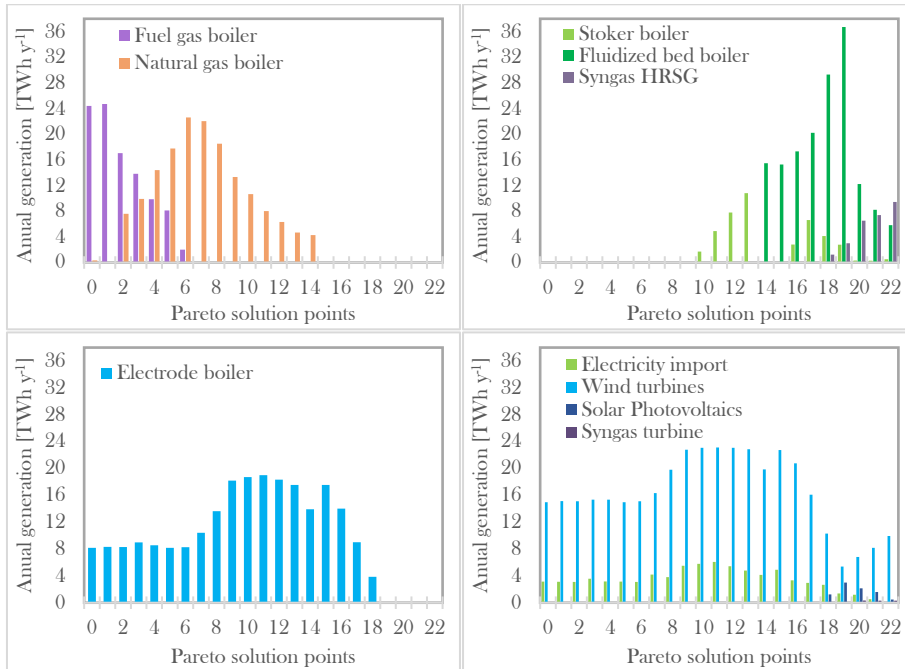


Figure 3 Pareto front: Energy conversion technology at each design point

It is observed that a combination of fossil fuel and low carbon technology can reduce GHG emissions by up to 66%. However, if further decarbonization is required, a complete transition to renewable energy sources is required. It is worth noting that for decarbonization levels above 84%, electricity import is largely avoided, resulting in the remaining demand being met with syngas generated on-site (via gasification), nearly doubling operational costs (in comparison to a conventional fossil fuel-based design) while primarily affecting capital and maintenance costs (due to gasification equipment investment).

## 6. Conclusions

A bi-objective multiperiod optimization model was introduced to consider the environmental impacts of industrial utility systems while providing cost-effective solutions. The model takes into account the selection of the steam main operating conditions in order to take advantage of the close interrelationship between the site processes and the utility system at various temperature levels. The application to the design of a utility system for a petrochemical complex demonstrates that the proposed approach is a highly promising tool for the design of efficient and cost-effective sustainable utility systems. The optimization of steam main conditions can improve the results obtained from decarbonization through a better energy utilization, leading to lower fuel consumption as well as smaller equipment sizes.

Future development of the methodology will address the integration of further energy sources and technologies.

## Acknowledgments

This work has been funded by the Norwegian Research Council through HighEFF (project no. 257632). The authors gratefully acknowledge the financial support from the Research Council of Norway and user partners of HighEFF. Also, a special gratitude to the National Secretariat for Higher Education Science, Technology and Innovation of Ecuador (SENESCYT) for its support.

## References

- Bungener, S., Hackl, R., Van Eetvelde, G., Harvey, S., & Marechal, F. (2015). Multi-period analysis of heat integration measures in industrial clusters. *Energy*, *93*, 220-234.
- Ganschinietz, D. C. (2021). Design of on-site energy conversion systems for manufacturing companies – A concept-centric research framework. *Journal of Cleaner Production*, *310*, 127258.
- Jimenez-Romero, J., Azapagic, A., & Smith, R. (2022). *BEELINE: Bilevel decomposition algorithm for synthesis of industrial energy systems*. Department of Chemical Engineering, University of Manchester, Manchester.
- Jiménez-Romero, J., Azapagic, A., & Smith, R. (2023). Style: A new optimization model for Synthesis of utility systems with steam level placement. *Computers & Chemical Engineering*, *170*, 108060.
- Mavrotas, G. (2009). Effective implementation of the  $\epsilon$ -constraint method in Multi-Objective Mathematical Programming problems. *Applied Mathematics and Computation*, *213*(2), 455-465.
- Sun, L., Doyle, S., & Smith, R. (2015). Heat recovery and power targeting in utility systems. *Energy*, *84*, 196-206.
- Vaskan, P., Guillén-Gosálbez, G., Turkay, M., & Jiménez, L. (2014). Multiobjective Optimization of Utility Plants under Several Environmental Indicators Using an MILP-Based Dimensionality Reduction Approach. *Industrial & Engineering Chemistry Research*, *53*(50), 19559-19572.

# Energy and Exergy analysis of Biomass Direct Chemical Looping Combustion for CO<sub>2</sub> capture and Utilization

Shailesh Singh Sikarwar, Ramsagar Vooradi, Venkata Suresh Patnaikuni

*National Institute of Technology, Warangal, TG State - 506004, India*  
*sarat@nitw.ac.in*

## Abstract

Since a substantial increase in the usage of fossil fuels, particularly in recent past, has resulted in depletion of fossil fuels and the deteriorating state of the environment. Recently, electricity generation from renewable energy sources, such as wind, solar, and biomass, have garnered more attention. To minimize the emissions of CO<sub>2</sub>, it is crucial to use biomass on a large scale in energy-related applications. In recent years, chemical looping combustion of biomass gained attention as it is a carbon negative technology which can capture CO<sub>2</sub> with minimal energy penalty and a potential alternative to the conventional processes. In addition, hydrogen can also be produced using three reactor chemical looping. In this work, a novel cogeneration configuration is proposed for generation of electricity, hydrogen and methane. The proposed integrated system consists of a biomass direct chemical looping unit, organic Rankine cycle, water separation and compression unit, and CO<sub>2</sub> & H<sub>2</sub> utilization unit. To increase process energy efficiency, a sensitivity analysis is performed to optimize important design parameters such as oxygen carrier flow rate, operating pressure and operating temperature of air, fuel and steam reactors. The performance of the BDCL power plant is compared with conventional and CLC integrated biomass power plants in terms of energy and exergy. Results showed that BDCL plant has higher energy and exergy efficiency compared with the other two.

**Keywords:** Biomass direct chemical looping, H<sub>2</sub> production, CO<sub>2</sub> utilization, energy and exergy analyses, Organic Rankine cycle.

## 1. Introduction

Renewable energy sources must be adopted by all countries, since fossil fuel usage is a major concern both internationally and domestically. The renewable energy utilization can be further increased by encouraging the use of locally available renewable energy sources in the way as it covers the needs of the customer. With approximately 6% of the world's energy supply and a share of 55% of all renewable energy, modern bioenergy is the most significant source of renewable energy worldwide. Between 2010 and 2021, the use of contemporary bioenergy rose by an average of 7% year and is now on the rise (IEA., 2022). In 2019, 22% of India's total energy supply was derived from renewable sources. Moreover, biomass accounts for approximately 85% of all renewable energy (IEA., 2021). There has been an assessment of the economic, environmental, and energy sustainability of different conventional and non-conventional energy sources in recent years. The power production from biomass is more environmentally sustainable than the coal, while both options score similar exergetic sustainability (Stougie et al., 2018). In the conventional biomass fired power production (BFPP), the biomass burned directly with air in the combustor and the flue gases contain significant amounts of CO<sub>2</sub> and N<sub>2</sub>.

This scheme generally referred as carbon neutral power production. It has become increasingly relevant to integrate CO<sub>2</sub> capture technology into BFPP power plants to make this technology a carbon negative. Chemical Looping Combustion (CLC) is one of the technologies being explored on a wide scale. A typical CLC system comprises of air reactor and the fuel reactor in loop as shown in. The metal oxide (oxygen carrier) circulates in between these two reactors and facilitates the transfer of oxygen from air reactor to fuel reactor, which avoids the direct involvement of air in combustion process. The resulting flue gasses contains only CO<sub>2</sub> and steam. The studies reported in the literature on CLC integrated biomass-fired power plants (BFPP) mainly focused on integration of CLC technology with steam Rankine cycle for power production with in-situ CO<sub>2</sub> capture (Adanez et al., 2018). The integration of CLC with the low capacity BFPP which uses the organic Rankine cycle (ORC) is not widely studied in the literature (Coppola and Scala, 2021). Integration of CLC technology with low capacity BFPP for power production from biomass, makes the process as carbon negative, as well as a sustainable alternative for developing grid-independent plants. Further, by exploring the different cogeneration options along with improving the overall efficiency of the process, the needs of the communities can be met in a sustainable manner. Novel energy systems based on (i) biomass gasification to produce the useful commodities of electricity, heating and hydrogen (Ishaq and Dincer, 2019) (ii) polygenerational system driven by biomass and solar resources to produce useful products of power, hydrogen, hot water for fish farming and cooling (Ishaq et al., 2018) have high energy and exergy efficiencies and shown the distinct advantages of cogeneration in meeting the essential community needs. The Biomass Direct Chemical Looping (BDCL) system is a very efficient option when aiming for co-production of power and hydrogen using CLC technology (Cormos, 2015). This study aims at synthesis of novel BDCL configuration by integrating three reactor CLC unit, ORC unit and H<sub>2</sub> & CO<sub>2</sub> units for power and hydrogen cogeneration. Furthermore, the produced hydrogen can be utilized for the production of methane. The overall assessment of the developed configurations is carried out on the basis of energy, exergy and environmental parameters and compared with conventional biomass power plant configuration.

## 2. Methodology

In this study a CLC integrated BFPP configuration is developed and validated against the literature model (Sikarwar et al., 2020). The two reactors in the CLC loop of the reference model are replaced with three reactors for electricity and hydrogen cogeneration. Finally, the resulting BDCL configuration is integrated with Methane synthesis unit for deriving multigeneration process configuration. Steady state simulations and thermodynamic analyses of the above mentioned configurations are performed in Aspen Plus V10. All of these plant configurations are simulated for sugarcane bagasse with the following characteristics: carbon 44.48%, ash 2.09%, oxygen 40.69%, hydrogen 6.057%, sulfur 0.047%, nitrogen 0.19% and high heating value 16.45 MJ/kg. Fe<sub>2</sub>O<sub>3</sub> with 30 wt.% Al<sub>2</sub>O<sub>3</sub> support is used as oxygen carrier in all cases.

### 2.1 Description of CLC integrated BFPP for power production

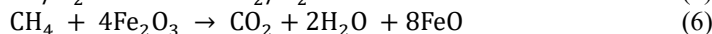
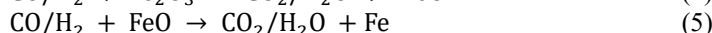
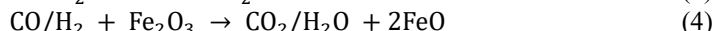
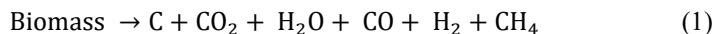
In this study, CLC integrated BFPP is considered as a reference case. The plant had already been simulated and detailed analysis of the plant flowsheet can be found in Sikarwar et al., 2020. This plant is mainly composed of two reactors (air and fuel reactors) which act as combustor unit, ORC unit to produce power at small scale, gas

turbine and compression unit. Both air and fuel reactors are simulated using RGibbs reactor.

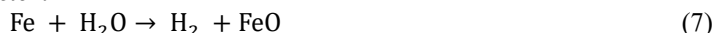
### 2.2 Description of BDCL plant for H<sub>2</sub> and power co-production

The proposed BDCL process plant utilizes three reactors (air, fuel and steam reactors) instead of two reactors used in CLC integrated BFPP. In this case, both fuel and steam reactors are simulated as moving bed reactor (multi-stage RGibbs reactor) while air reactor is simulated as single-stage RGibbs reactor. Multi-stage RGibbs reactor is used because it helps in emulating the moving bed conditions. The biomass is combusted in fuel reactor using the oxygen carrier (Fe<sub>2</sub>O<sub>3</sub> with 30 wt.% Al<sub>2</sub>O<sub>3</sub> support) where Fe<sub>2</sub>O<sub>3</sub> is reduced to FeO/Fe. Then, reduced FeO/Fe is partially oxidized to Fe<sub>3</sub>O<sub>4</sub>/FeO using steam in steam reactor. Partially oxidized Fe<sub>3</sub>O<sub>4</sub>/FeO is fully oxidized to Fe<sub>2</sub>O<sub>3</sub> in air reactor and sent back to fuel reactor thus forming a loop. The chemical reactions involved in all three reactors are given below (Li et al., 2010).

#### **Fuel reactor:**



#### **Steam reactor:**



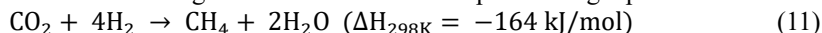
#### **Air reactor:**



Hot gaseous streams of CO<sub>2</sub>-rich flue gas produced from fuel reactor and H<sub>2</sub>-rich gas produced from steam reactor, are sent directly to evaporator in ORC unit to extract the heat. After extracting the heat from CO<sub>2</sub>-rich flue gas and H<sub>2</sub>-rich gas, they are sent to water separation and compression unit where pure CO<sub>2</sub> and H<sub>2</sub> gases are compressed to sequestration-ready pressures. Water separated from CO<sub>2</sub>-rich flue gas and H<sub>2</sub>-rich gas streams, is converted to steam using extra heat available from these two streams. This steam is recycled back to steam reactor for H<sub>2</sub> production.

### 2.3 Description of BDCL plant with Methane synthesis

The proposed BDCL plant is further integrated with methanation unit for methane production as shown in Fig.1. In this configuration the CO<sub>2</sub> and H<sub>2</sub> produced from BDCL plant are sent to methanation reactor. In the reactor, CO<sub>2</sub> enters at 30 bar pressure and 68 °C temperature and H<sub>2</sub> enters at 30 bar pressure and 121 °C temperature. In the presence of Ni-catalyst, both of them react to produce Methane. Methanation reactor is simulated as single-stage RGibbs reactor where Sabatier reaction (Eq. 11) takes place at 385 °C (Rönsch et al., 2016). Sabatier reaction is exothermic in nature which supplies the heat required to maintain high temperature in methanation reactor. The product stream of methane along with steam is sent to a separator to get pure methane.



## 2.4 Performance Evaluation

The overall energy and exergy efficiencies of all plant configurations are calculated using the equations (12) and (13) respectively (Suresh et al., 2010). Here,  $m_{biomass}$  is the mass flow rate of biomass ( $kg/s$ ) and  $HHV$  is the higher heating value of the biomass ( $MJ/kg$ ).

$$\eta_{Gross\ or\ net} = \frac{Gross\ or\ net\ work\ output}{m_{biomass} \times H.H.V.\ of\ the\ biomass} \quad (12)$$

$$\varepsilon_{Gross\ or\ net} = \frac{Gross\ or\ net\ work\ output}{m_{biomass} \times Specific\ exergy\ of\ the\ biomass} \quad (13)$$

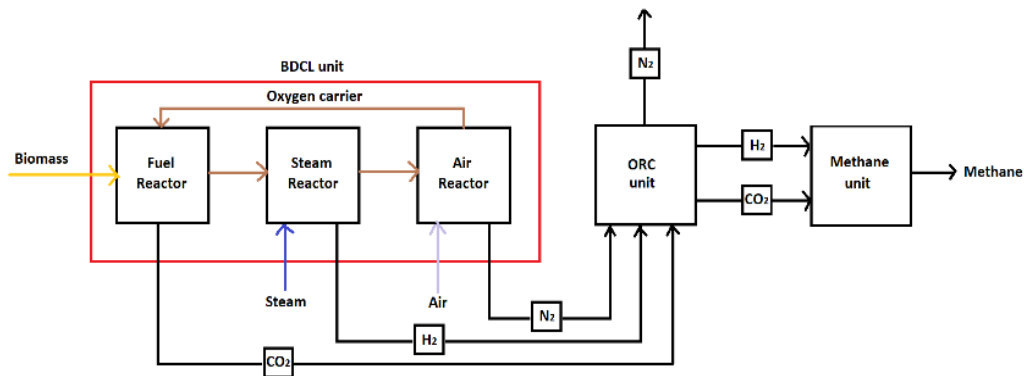


Fig.1: Schematic flow diagram of BDCL plant with Methane synthesis

## 3. Results and discussion

### 3.1 BDCL plant for $H_2$ and power co-production

The BDCL plant configuration is simulated for the biomass flow rate of 0.139 kg/s under steady state condition. The BDCL plant's three main operational parameters: oxygen carrier to biomass ratio, the operating pressure and operating temperatures of the three reactors have been optimized using a sensitivity analysis. The sensitivity analysis results are presented in Figs. 2-5. As operating pressure increases, the energy and exergy efficiencies are exponentially increased up to 15 bar pressure. When the reactor temperature reaches 1150 °C, the maximum energy and exergy efficiencies are observed, but after this extra heat is required to maintain the temperature. Since the reactions in steam and fuel reactors are endothermic, more amount of heat need to be supplied to maintain their temperatures. However, decreasing the temperatures results in low hydrogen production. From the sensitivity analysis, the optimal operating conditions for different units are identified as: loop operating pressure: 15 bar, Air reactor temperature: 1150 °C, Steam reactor temperature: 750 °C and Fuel reactor temperature: 800 °C. Subsequently, the performance of BDCL plant has been compared with conventional and CLC integrated BFPP configurations. In Fig. 6, the exergy

efficiency of individual units is illustrated for both plants. From the results presented in Table 1, it is observed that the proposed BDCL plant has high energy and exergy efficiencies.

### 3.2 BDCL plant with Methane synthesis

Synthesis of methane results in loss of energy due to utilization of all of the produced hydrogen from BDCL plant. While producing 62.08 kg/hr of methane from the above BDCL plant, the net overall energy and exergy efficiencies are found to be 13.65% and 11.74%, respectively. The stoichiometric ratio of H<sub>2</sub> to CO<sub>2</sub> is 4:1 for the synthesis of methane. Hence, the methane synthesis plant utilizes 21% of CO<sub>2</sub> and 100% of H<sub>2</sub> produced from the BDCL plant.

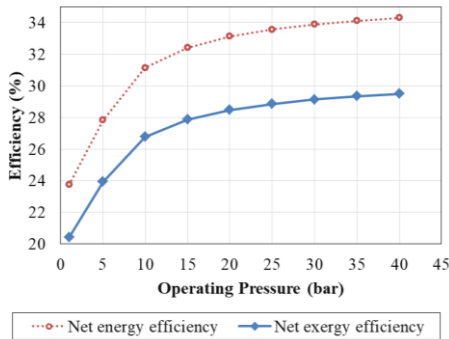


Fig.2. Effect of operating pressure on overall energy and exergy efficiencies.

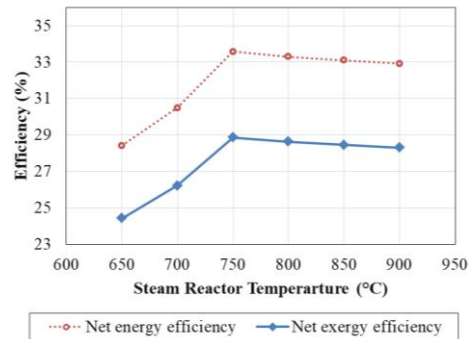


Fig.4. Effect of SR Temperature on overall energy and exergy efficiencies.

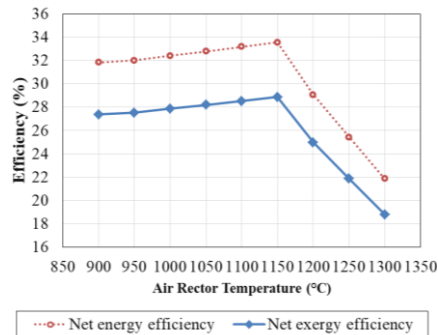


Fig.3. Effect of AR Temperature on overall energy and exergy efficiencies.

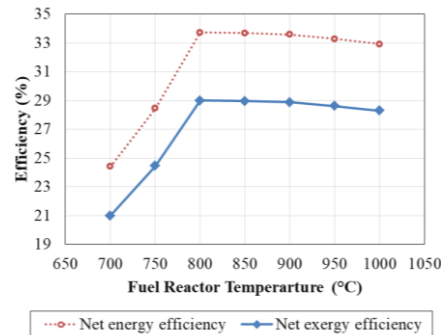


Fig.5. Effect of FR Temperature on overall energy and exergy efficiencies.

Table 1: Performance comparison based on energy and exergy efficiencies.

Case	Net overall energy efficiency (%)	Net overall exergy efficiency (%)
Conventional-BFPP	11.98	10.30
CLC integrated BFPP	21.59	18.56
BDCL plant	33.72	28.99



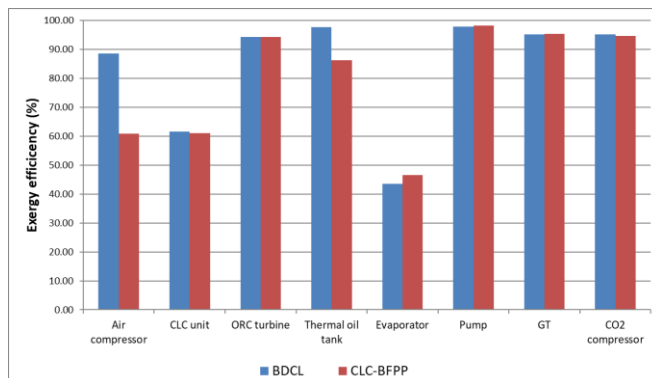


Fig.6: The exergy efficiency of individual units

#### 4. Conclusions

In this work, a novel three-reactor BDCL plant is simulated using Aspen Plus V10 for power and hydrogen co-production. The performance of the BDCL plant is compared with the conventional-BFPP and CLC-integrated BFPP energetically and exergetically. The simulation results showed that net overall energy efficiency of the BDCL plant is higher than the conventional-BFPP and CLC-integrated BFPP by 21.74% and 12.13%, respectively. In addition, net overall exergy efficiency of the BDCL plant is higher than the conventional-BFPP and CLC-integrated BFPP by 18.69% and 10.43%, respectively. Furthermore, a CO<sub>2</sub> and H<sub>2</sub> utilization strategy i.e. methane synthesis is integrated with the BDCL plant. Integration of methane synthesis results in loss of net overall energy efficiency and net overall exergy efficiency by 20.07% and 17.25%, respectively. However, 21% of captured CO<sub>2</sub> is effectively utilized in methane synthesis.

#### References

- IEA (2022), Bioenergy, IEA, Paris <https://www.iea.org/reports/bioenergy>, License: CC BY 4.0
- IEA (2021) World Energy Balances and Renewables Information
- L. Stougie, N. Giustozzi, H. Kooi, A. Stoppato, 2018, International journal of energy research, 42, 2916-2926
- J. Adánez, A. Abad, T. Mendiara, P. Gayán, LF. de Diego, F. García-Labiano, 2018, Progress in energy and combustion science, 65,6-66
- Coppola, F. Scala, 2021, Energy & fuels, 35, 23, 19248–19265
- H. Ishaq, I. Dincer, 2019, Energy conversion and management, 196, 395-409
- H. Ishaq, I. Dincer, GF. Naterer, 2018, International journal of hydrogen energy, 43, 52, 23148-23160
- Cormos, 2015, Fuel processing technology, 137, 16-23
- SS. Sikarwar, GD. Surywanshi, VS. Patnaikuni. M. Kakunuri, R. Vooradi, 2020, Renewable energy, 155, 931-949
- F. Li, L. Zeng, LS. Fan, 2010, Fuel, 89, 12, 3773-3784
- S. Rönsch, J. Schneider, S. Matthischke, M. Schlüter, M. Götz, J. Lefebvre, P. Prabhakaran, S. Bajohr, 2016, Fuel, 166, 276-296
- MVJJ. Suresh, KS. Reddy, AK. Kolar, 2010, International journal of energy research, 34, 8, 716-735

# End-effect mitigation in multi-period stochastic programming of energy storage operations

Teemu J. Ikonen,<sup>a</sup> Dongho Han,<sup>b</sup> Jay H. Lee,<sup>b</sup> Iiro Harjunkoski,<sup>a,c</sup>

<sup>a</sup>*Aalto University, Department of Chemical and Metallurgical Engineering, PO Box 16100, 00076 Aalto, Finland*

<sup>b</sup>*Korea Advanced Institute of Science and Technology, Department of Chemical and Biomolecular Engineering, 291 Daehak-ro, Yuseong-gu, Daejeon, 34141, Republic of Korea*

<sup>c</sup>*Hitachi Energy Research, Havellandstr. 10-14, 68309 Mannheim, Germany*

## Abstract

Energy storage units offer vital balancing power for energy systems with increasing amount of variable renewable energy sources. The operation of such systems can be optimized by stochastic programming, which anticipates the uncertainty related to the variable energy generation. However, these optimization problems can only be formulated for optimization horizons of a limited length (e.g., 24 or 48 h), due to the rapidly increasing problem size. In this work, we propose the resulting end-effect to be mitigated by valuation of the terminal stored energy level based on an electricity price forecast. We present results on a hybrid energy system, consisting of photovoltaic power generation and an energy storage unit, which trades electricity in the day-ahead market.

**Keywords:** stochastic programming, energy storage, solar energy, electricity market

## 1. Introduction

Energy storage units are essential components of a modern electricity grid, as they can mitigate the undesired variability in many renewable energy sources, such as wind and solar. However, the variable renewable energy sources (VRES) are also inherently uncertain, which needs to be considered in the optimal operation of the energy storage units (Weitzel and Glock, 2018). In the literature, many of the proposed optimization models are based on stochastic programming, i.e., a framework to formulate optimization problems under uncertainty by optimizing the expected outcome (Birge and Louveaux, 2011). Stochastic programming is based on scenarios of realized uncertainty, each of which has dedicated optimization variables. Thus, stochastic programming problems can only be solved for limited horizon lengths, as the optimization problems quickly become intractable due to the increasing number of variables.

The finiteness of the optimization horizon causes a phenomenon, referred to as the end-effect (Grinold, 1983). The end-effect causes the energy storage system to be drained empty at the end of the current optimization horizon if no value is given for the terminal energy level (see, e.g., Singh and Knueven, 2021). Such decisions can be highly sub-optimal over long time horizons. Often used methods to handle the end-effect are i) to enforce the terminal inventory (i.e., in our case, the level of stored energy) to equal the starting inventory or ii) to use a rolling horizon method, such that the current optimization horizon spans beyond the next re-optimization time. Shin et al. (2017) mitigate the end-effect of the energy storage system in wind power-based energy grid systems using a value function learned from reinforcement learning. Han and Lee (2021) propose a

method to determine the optimal design and operation strategy of energy grid systems based on stochastic programming while assuming that the terminal level of stored energy is equal to the starting level. Dong and Maravelias (2021) propose multi-material terminal inventory constraints, which are applicable to online production scheduling. In this work, we propose the end-effect to be mitigated by valuating the terminal stored energy level based on an electricity price forecast for the following day.

## 2. Methods

The studied energy system trades electricity in the day-ahead market and consists of a photovoltaic power station and an energy storage unit (Figure 1). In this section, we describe the used stochastic programming model, the methods to mitigate the end-effect on this energy system, and the generation of scenarios.

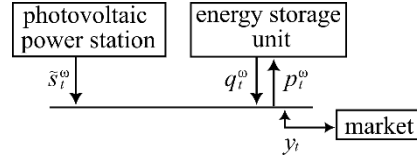


Figure 1: A hybrid energy system.

### 2.1. Stochastic programming model

The objective of the stochastic programming model is to maximize the expected sum of i) the profit during the optimization horizon and ii) the value of stored energy at the end of the horizon. The first stage decision variables, made before the start of each period, are the promised power to be traded,  $y_t$ , at hour  $t \in T$ . A positive value of  $y_t$  means selling electricity, and a negative value buying electricity. The second stage decision variables, for a given scenario  $\omega \in \Omega$ , are the amount of stored energy  $x_t^\omega$ , the power of charging the storage unit  $p_t^\omega$ , the power of discharging the storage unit  $q_t^\omega$ , and the photovoltaic power  $\tilde{s}_t^\omega$  at hour  $t \in T$  (see Figure 1). Further,  $\Delta_t^{+, \omega}$  and  $\Delta_t^{-, \omega}$  are the positive and negative deviation variables from the promised power at hour  $t \in T$  (the former means surplus in the delivered energy). We denote the electricity price by  $R_t$ , the time interval by  $\tau$ , the cost of discharging the storage system by  $C_d$  (which represents the aging of the storage system due to cycling), the value of stored energy at the end of the horizon by  $V$ , and the probability of scenario  $\omega$  by  $\rho_\omega$ . The penalty of deviating from the promised energy is linearly proportional to  $R_t$  and  $\Delta_t^{+, \omega}$  or  $\Delta_t^{-, \omega}$  with coefficients  $\gamma^+$  and  $\gamma^-$ , respectively (Ding et al., 2014). The model is defined as follows:

$$\max \sum_{\omega \in \Omega} \rho_\omega \left\{ \tau \sum_{t \in T} [R_t (\tilde{s}_t^\omega + q_t^\omega - p_t^\omega) - C_d q_t^\omega - \gamma^+ R_t \Delta_t^{+, \omega} - \gamma^- R_t \Delta_t^{-, \omega}] + V x_{|T|+1}^\omega \right\} \quad (1)$$

$$\Delta_t^{+, \omega} \geq \tilde{s}_t^\omega + q_t^\omega - p_t^\omega - y_t, \quad t \in T, \omega \in \Omega \quad (1)$$

$$\Delta_t^{-, \omega} \geq -(\tilde{s}_t^\omega + q_t^\omega - p_t^\omega - y_t), \quad t \in T, \omega \in \Omega \quad (2)$$

$$x_{t+1}^\omega = x_t^\omega + \eta_c \tau p_t^\omega - \frac{\tau}{\eta_d} q_t^\omega, \quad t \in T, \omega \in \Omega \quad (3)$$

$$x_1^\omega = x_0 \quad \omega \in \Omega \quad (4)$$

$$0 \leq p_t^\omega \leq \bar{P}, \quad t \in T, \omega \in \Omega \quad (5)$$

$$0 \leq q_t^\omega \leq \bar{Q}, \quad t \in T, \omega \in \Omega \quad (6)$$

$$\underline{X} \leq x_t^\omega \leq \bar{X}, \quad t \in \{1, \dots, |T| + 1\}, \omega \in \Omega \quad (7)$$

$$0 \leq \tilde{s}_t^\omega \leq S_t^\omega, \quad t \in T, \omega \in \Omega \quad (8)$$

$$\Delta_t^{+, \omega}, \Delta_t^{-, \omega} \geq 0, \quad t \in T, \omega \in \Omega$$

$$y_t \geq -\bar{P}, \quad t \in T$$

Constraints 1 and 2 ensure that (at each hour  $t \in T$  in scenario  $\omega \in \Omega$ ) the surplus and lack of power is tracked by variables  $\Delta_t^{+, \omega}$  and  $\Delta_t^{-, \omega}$ , respectively. Constraints 3 track the amount of stored energy, where  $\eta_c, \eta_d \in [0, 1)$  are the efficiencies of charging and discharging, respectively. Constraints 4 define the amount of stored energy  $x_0$  at the start of the optimization horizon. Constraints 5 and 6 specify the upper bounds,  $\bar{P}$  and  $\bar{Q}$ , for charging and discharging the energy storage unit, respectively. Accordingly, constraints 7 specify the lower and upper bounds,  $\underline{X}$  and  $\bar{X}$ , for the stored energy. Here, it is worth noticing index  $t$  being over set  $\{1, \dots, |T| + 1\}$ , instead of  $T$ . Variable  $x_{|T|+1}^\omega$  tracks the stored energy level after the last hour of the period (it appears also in the objective function). Constraints 8 facilitate curtailment of the photovoltaic power  $\xi_t^\omega$ , where  $S_t^\omega$  is the maximum available photovoltaic power at hour  $t$  in scenario  $\omega$  (see Section 2.3).

### 2.2. End-effect mitigation

The focus of this work is on valuation of stored energy level at the end of an optimization horizon, in order to mitigate the end-effect. In the above-described stochastic programming model, the valuation is controlled via the parameter  $V$ , which is the multiplier of the terminal storage level  $x_{|T|+1}^\omega$  in different scenarios  $\omega \in \Omega$ . We determine parameter  $V$  based on the forecasted electricity price obtained by the Lasso Estimated AutoRegressive (LEAR) model by Lago et al. (2021), which we have modified to also forecast electricity prices beyond the day-ahead. We test setting  $V$  to the minimum, mean, and maximum forecasted electricity price on the period following the current optimization horizon of 24 hours (see Table 1). We also test two reference methods. The first is a myopic method that does not mitigate the end-effect at all, referred to as ‘no mitigation’. The second is a rolling horizon method, in which the optimization horizon is set to 48 hours.

Table 1: Studied methods to mitigate the end-effect.

method	horizon length $ T $	Valuation of terminal energy level $V$
min price-based valuation	24	$V$ is the predicted <i>minimum</i> electricity price during the next period
mean price-based valuation	24	$V$ is the predicted <i>mean</i> electricity price during the next period
max price-based valuation	24	$V$ is the predicted <i>maximum</i> electricity price during the next period
no mitigation	24	$V = 0$
rolling horizon	48	$V = 0$

### 2.3. Scenario generation

The described energy system is highly affected by the uncertainty in the hourly solar irradiance. Solcast is a state-of-the-art model for solar irradiance forecasting (Bright, 2019). Its forecasts are based on satellite images from five geostationary weather satellites, as well as models for aerosol and vapor concentrations. In addition to the mean of the forecast, Solcast also yields estimates of the forecast uncertainty, represented by the 10<sup>th</sup> and 90<sup>th</sup> percentiles. Figure 2 shows a seven-day probabilistic forecast of the global horizontal irradiance (GHI) in Bavaria, Germany.

Based on these probabilistic forecasts, we generate the scenarios of the global solar irradiance (GHI) as follows. First, we fit a four-parameter beta distribution to the predicted mean and the 10<sup>th</sup> and 90<sup>th</sup> percentiles at each time point  $t \in T$ . Second, we generate the scenarios of the hourly GHI  $S_{\text{GHI}, t}^\omega, t \in T$  by the statistical method proposed by Pinson et al. (2009), which takes into account the temporal dependency of the realized

solar power generation. The dependency is determined based on historical data of past forecasts and realizations. We further assume that the photovoltaic panels are installed horizontally. The solar power generation scenarios are then

$$S_t^\omega = \eta_s A S_{\text{GHI},t}^\omega \quad t \in T, \omega \in \Omega \quad (9)$$

where  $S_{\text{GHI},t}^\omega$  is the GHI value obtained by the above-described procedure,  $A$  is the total area of the photovoltaic panels, and  $\eta_s$  is their efficiency.

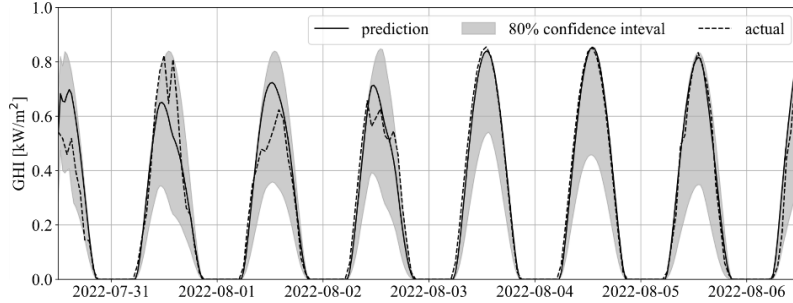


Figure 2: Probabilistic forecast of global horizontal irradiance (GHI) in Bavaria, Germany (49.158°N, 11.433°E) (Solcast, 2019).

### 3. Results

We test the methods, described in the previous section, on a 91-day time window from July 31 to October 29, 2022. The forecasted and actual GHI are obtained from Solcast (Solcast, 2019). The dependency in the realized GHI at the different time points in the forecast horizon (with respect to the forecasts) is modeled from the data recorded from June 20 to July 30, 2022. We use electricity market data from the German Federal Network Agency (<https://www.smard.de/en>), transmission system operator Tennet (<https://netztransparenz.tennet.eu/>), and ENTSO-E transparency platform (<https://transparency.entsoe.eu/>). The results are generated on a laptop with an Intel i5-8365U processing unit. The stochastic programming models are solved by Gurobi 9.1.0.

We investigate the same time window with two configurations of the electricity storage unit (Table 2). The discharge cost of  $C_d = 0.04434$  EUR/kWh is calculated for a lithium-ion battery, having an estimated acquisition cost of \$132/kWh (Bloomberg, 2021) (we use a currency conversion rate of \$1 = 0.96 EUR), expected lifetime of 3000 cycles, and operation and maintenance costs of  $2.1 \times 10^{-3}$  EUR/kWh (Zakeri and Syri, 2015). We set the number of scenarios  $N = 100$ , the time interval  $\tau = 1$  h, and photovoltaic panels to have an area of  $A = 4000$  m<sup>2</sup> and efficiency of  $\eta_s = 0.25$ . We set the deviation penalty coefficients  $\gamma^+ = \gamma^- = 10$ , aiming to represent a situation where trading at the intraday market is not possible. The charge and discharge efficiencies are  $\eta_c = \eta_d = 0.9$ . At the beginning of the first 24 h period, the energy storage unit is at the minimum charge level,  $x_0 = \underline{X}$ . For the sake of a fair comparison, the last period is optimized using the ‘no mitigation’ method.

Table 3 lists the average daily profits, obtained using the five methods to mitigate the end-effect. For a reference, we have also included a method, referred to as *perfect foresight*, which has access to the actual GHI and electricity prices, and uses a horizon length of 72 hours. The value of the stored energy is not considered in the profit. In both

configurations, the min price-based valuation yields higher average profits than those based on mean and maximum electricity prices. The result is reasonable because the energy system could also buy energy during the early morning hours when the minimum electricity price often occurs. The min price-based valuation yields 14.2 and 18.3% higher profits than using no end-effect mitigation. Figure 4 visualizes the storage operations during first two days using these methods. In contrast to no mitigation, the energy system optimized with the min price-based valuation avoids buying electricity in the morning of Day 2 (Figure 4b and Figure 4d). The average profits of the min price-based valuation are 2.0 and 5.2% higher than those of the rolling horizon method. Figure 3 shows the cumulative profits of the two configurations when using different end-effect mitigation methods.

Table 2: Parameters of the energy storage system.

parameter	Configuration 1	Configuration 2
maximum charge power $\bar{P}$	500 kW	1000 kW
maximum discharge power $\bar{Q}$	500 kW	1000 kW
maximum charge level $\bar{X}$	2000 kWh	4000 kWh
minimum charge level $\underline{X}$	400 kWh	800 kWh

Table 3: Numerical results of the end-effect mitigation methods.

end-effect mitigation	average daily profit [EUR]	
	Configuration 1	Configuration 2
min price-based valuation	1098.89	1268.06
mean price-based valuation	1096.20	1261.50
max price-based valuation	1066.10	1173.45
no mitigation	961.94	1071.84
rolling horizon	1077.84	1205.65
perfect foresight	1424.82	1628.08

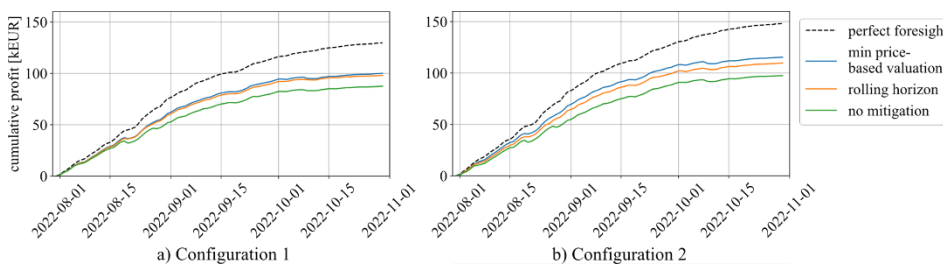


Figure 3: Cumulative profit of the energy system with the two different configurations.

#### 4. Conclusions

This paper proposes the end-effect in multi-period stochastic programming of energy storage operations to be mitigated by valuation of the terminal energy storage level. On the two studied energy system configurations, the valuation method based on minimum forecasted price yields 14.2 and 18.3% higher profits than the myopic method, often used in the literature. In comparison to the rolling horizon method, the main benefit is that the proposed method enables the use of shorter optimization horizons without a reduction in the long-term profit.

**Acknowledgment:** Financial support from the Academy of Finland is gratefully acknowledged (project RELOOP, decision number 330388).

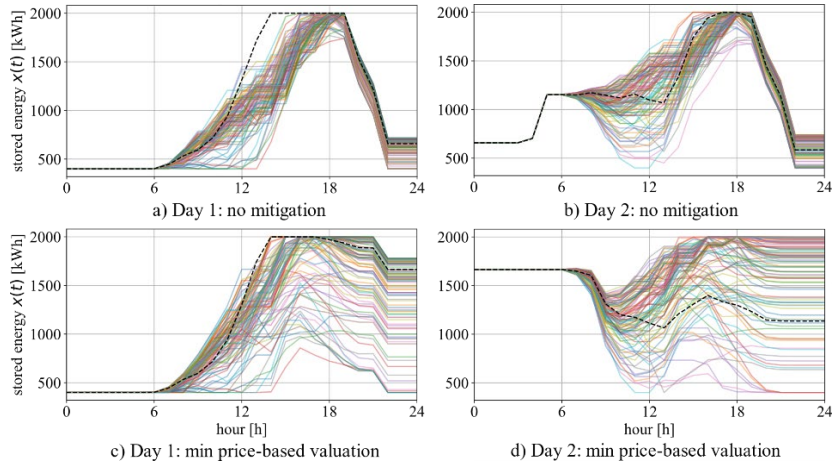


Figure 4: Stored energy  $x(t)$  during the first two days for Configuration 1 when using ‘no mitigation’ or ‘min price-based valuation’. Continuous lines are the scenarios. The dashed line is the realized path.

## References

- J. R. Birge and F. Louveaux, 2011, Introduction to stochastic programming, Springer Science & Business Media.
- J. M. Bright, 2019, Solcast: Validation of a satellite-derived solar irradiance dataset, *Solar Energy*, 189:435–449.
- Solcast, 2019. Global solar irradiance data and PV system power output data. URL <https://solcast.com/>
- H. Ding, Z. Hu, & Y. Song, 2014. Rolling optimization of wind farm and energy storage system in electricity markets. *IEEE Transactions on Power Systems*, 30(5):2676–2684.
- Y. Dong and C. T. Maravelias, 2021, Terminal inventory level constraints for online production scheduling, *European Journal of Operational Research*, 295(1):102–117.
- R. C. Grinold, 1983, Model building techniques for the correction of end effects in multistage convex programs, *Operations Research*, 31(3):407–431.
- D. Han and J. H. Lee, 2021, Two-stage stochastic programming formulation for optimal design and operation of multi-microgrid system using data-based modeling of renewable energy sources, *Applied Energy*, 291:116830.
- J. Lago, G. Marcjasz, B. De Schutter, and R. Weron, 2021, Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark, *Applied Energy*, 293:116983.
- P. Pinson, H. Madsen, H. A. Nielsen, G. Papaefthymiou, and B. Klöckl, 2009, From probabilistic forecasts to statistical scenarios of short-term wind power production. *Wind Energy*, 12(1):51–62.
- J. Shin, J. H. Lee, and M. J. Realff, 2017. Operational planning and optimal sizing of microgrid considering multi-scale wind uncertainty. *Applied energy*, 195:616–633.
- B. Singh and B. Knueven, 2021. Lagrangian relaxation based heuristics for a chance-constrained optimization model of a hybrid solar-battery storage system. *Journal of Global Optimization*, 80(4):965–989.
- T. Weitzel and C. H. Glock, 2018, Energy management for stationary electric energy storage systems: A systematic literature review. *European Journal of Operational Research*, 264(2):582–606.
- Bloomberg, 2021. <https://about.bnef.com/blog/battery-pack-prices-fall-to-an-average-of-132-kwh-but-rising-commodity-prices-start-to-bite/> [accessed on June 15, 2022], 2021.
- B. Zakeri and S. Syri, 2015. Electrical energy storage systems: A comparative life cycle cost analysis. *Renewable and sustainable energy reviews*, 42: 569–596.

# How to accurately fast-track sorbent selection for post-combustion CO<sub>2</sub> capture? A comparative assessment of data-driven and simplified physical models for screening sorbents

Sai Gokul Subraveti,<sup>a,\*</sup> Luca Riboldi,<sup>a</sup> Hao Yang Xu,<sup>b</sup> Yannick Jooss,<sup>a</sup> Simon Roussanaly,<sup>a</sup> Leif Erik Andersson,<sup>a</sup> Rahul Anantharaman<sup>a</sup>

<sup>a</sup>*SINTEF Energy Research, 7019 Trondheim, Norway*

<sup>b</sup>*Department of Chemical and Materials Engineering, University of Alberta, Edmonton, T6G 1H9, Canada*

\**sai.gokul.subraveti@sintef.no*

## Abstract

The recent discovery of a multitude of hypothetical materials for CO<sub>2</sub> capture applications necessitated the development of reliable computational models to aid the quest for better-performing sorbents. Given the computational challenges associated with existing detailed adsorption process design and optimization frameworks, two types of screening methodologies based on computationally inexpensive models, namely, data-driven and simplified physical models, have been proposed in the literature. This study compares these two screening methodologies for their effectiveness in identifying best-performing sorbents from a set of 369 metal-organic frameworks (MOFs). The results showed that almost 60% of the MOFs in the top 20 best-performing materials ranked by each of these approaches were found to be common. The validation of these results against detailed process simulation and optimization-based screening approach is currently underway.

**Keywords:** adsorption; metal-organic frameworks; post-combustion CO<sub>2</sub> capture; machine learning; modelling and optimization.

## 1. Introduction

Among several CO<sub>2</sub> capture technologies considered for post-combustion CO<sub>2</sub> capture, solid adsorbents are seen as a promising alternative to traditional liquid solvents for the separation of CO<sub>2</sub> from flue gases. These adsorbents are typically deployed in pressure/vacuum swing adsorption (PVSA) or temperature swing adsorption (TSA) processes. The choice of the adsorbent plays a critical role in determining the separation performance of PVSA or TSA processes [1][2]. Conventionally, better-performing adsorbents are identified through experimentation and testing of a few handfuls of adsorbents as means to understand their performance in the real process [3]. However, this approach is challenged by the recent advent of highly tunable adsorbents, such as metal-organic frameworks (MOFs) for CO<sub>2</sub> capture applications, resulting in thousands of potential hypothetical adsorbents [3]. As the experimental evaluation of a multitude of adsorbents is practically impossible, computational screening of the adsorbents has been considered, where process simulations and optimizations based on adsorbent properties are carried out to evaluate the process-scale performance [2][4]. This approach is computationally expensive and time-consuming [2][4][5], which makes it computationally inadequate to handle large databases of adsorbents.



Different approaches have been proposed to reduce the computational costs of existing simulation and optimization tools. One approach is the development of simplified physical models obtained through a simpler description of the process which can be solved in seconds [6][7]. These models proved able to provide reasonable estimations of the process performance. The other approach is the development of surrogate or data-driven models built based on statistical methods that act as faster approximations of process metrics [5]. With groundbreaking advances in machine learning, novel approaches that incorporate physics into surrogate models are also developed to reliably represent physical processes [8].

The goal of this study is to compare the performance of data-driven and simplified physical modeling approaches in rapidly screening databases of adsorbents based on a techno-economic assessment for post-combustion CO<sub>2</sub> capture applications. For this analysis, a set of 369 MOFs from the CoREMOf 2014 database provided by Leperi et al. [4] is used and the material performance is assessed using a four-step PVSA cycle with light product pressurization [1].

## 2. Screening methodologies

### 2.1. Data-driven model-based optimization framework

This methodology utilizes data-driven models built based on machine learning principles as a faster approximation for calculating process performance metrics. The data-driven models are coupled with the cost model and the non-dominated sorting genetic algorithm (NSGA – II) to optimize each material for the minimum cost of CO<sub>2</sub> avoided. The data-driven models used in this study are artificial neural networks (ANNs) based on the machine-assisted adsorption process learning and emulation (MAPLE) framework [5].

The inputs to the MAPLE model are process features consisting of process operating conditions: adsorption step duration, vacuum pump flow rates, column size, high pressure, intermediate pressure, low pressure, and feed composition; and adsorbent features comprising dual-site Langmuir isotherm parameters of CO<sub>2</sub> and N<sub>2</sub>, and particle morphology. Individual ANN models were trained for each output, namely, step durations, purity, recovery, and energy consumption. These predicted quantities form inputs to the cost model to calculate the cost of the CO<sub>2</sub> avoided based on the approach presented in Subraveti et al. [1]. The neural network architecture comprises a feed-forward fully connected network with one input layer including 19 process and adsorbent features, three hidden layers with 10-15 neurons, and an output layer with one output. A tanh activation function was used in the hidden layers and a linear activation was used for the output layer. Around 9000 unique combinations of the input variables generated using the Latin hypercube sampling along with the corresponding outputs were used as samples in the training of the neural networks. Note that the outputs were previously obtained by simulating the detailed adsorption process model until the cyclic-steady state condition [1][9]. The neural networks were trained using Bayesian regularization with the back-propagation algorithm ‘*trainbr*’ in MATLAB 2022a [5][9].

### 2.2. Simplified physical model-based optimization framework

The other approach for the screening of adsorbent materials involves utilizing models that describe the physical phenomena occurring in an adsorption process but introducing simplifications to decrease the computational effort. The larger the simplification level applied, the lower the computational effort. Conversely, the expected accuracy of the

models is expected to decrease. For this work, two simplified models are proposed. The first one termed the modified equilibrium model (MEM), relies on the key assumption of local equilibrium, i.e., CO<sub>2</sub> gets instantaneously adsorbed onto the adsorbent materials. The model is an extension of an approach presented in the literature [6]. The main modifications with respect to its original formulation include a different routine to solve the adsorption step and the possibility to simulate a 4-step cycle [10]. The second model, termed the reduced-order kinetic model (ROKM), attempts to go beyond the equilibrium assumption by introducing a methodology to implicitly solve the linear driving force (LDF) approximation and, hence, account for mass transfer resistances. The methodology builds on a set of simplifying assumptions, therefore a degree of inaccuracy in capturing the kinetics effects is expected.

The simplified physical models are coupled with the same cost model as in the MAPLE approach and a Bayesian optimization (BO) algorithm to optimize each material for the minimum cost of CO<sub>2</sub> avoided. The BO algorithm was developed in-house and tested for the optimization of PSA processes, showing a good balance between computational time and reliability [11]. Simplified physical models might not directly provide all necessary inputs to the cost model owing to the inherent assumptions made while developing these models. For example, the MEM model cannot provide the step durations that are critical for estimating cost. In such cases, relevant and consistent assumptions were made for all the materials to enable the integration of the techno-economic analysis framework for these simplified models. The MEM model needs more assumptions compared to the ROKM model for cost evaluation.

### 3. Results and discussion

The two methodologies are compared in their ability to reliably screen adsorbents for post-combustion CO<sub>2</sub> capture. Table 1 briefly summarizes the merits and demerits of each screening model in evaluating the adsorbent performance. In this study, a dry flue gas with CO<sub>2</sub>/N<sub>2</sub> binary mixture is separated using a four-step PVSA process, a widely studied process that has been demonstrated at the pilot scale [12]. The cycle consists of adsorption (ADS) step, a co-current blowdown (BLO) step, a counter-current evacuation (EVAC), and a light-product pressurization (LPP) step.

Table 1: Merits and demerits of both types of screening models for rapid evaluation of adsorbent performance.

Screening model type	Strengths	Limitations
Data-driven models	<ul style="list-style-type: none"> <li>• Very fast computations</li> <li>• Embeds all physical phenomena from the detailed model</li> <li>• Predictions represent the real process performance</li> </ul>	<ul style="list-style-type: none"> <li>• Requires computational efforts to generate data for training</li> <li>• Black-box model – Applicability within the training range</li> </ul>
Simplified physical models	<ul style="list-style-type: none"> <li>• Interpretability through the simplified description of physics</li> <li>• Easy to develop</li> <li>• Entails wider model applicability</li> </ul>	<ul style="list-style-type: none"> <li>• Simulations may not represent the real process performance</li> <li>• May lead to convergence failures and false optima in optimizations</li> </ul>

The performance of each MOF is assessed based on its techno-economic performance in the four-step PVSA cycle. The metric used for ranking the materials was the minimum CO<sub>2</sub> avoided cost obtained after optimizing the process operating conditions for each material. It is worth reiterating that both the surrogate and the simplified physical models

predict the process performance indicators which are later used as inputs to the cost model within the optimization framework. As mentioned earlier, the cost model employed herein is based on Subraveti et al. [1] for both approaches.

The screening of 369 MOFs was individually carried out for MAPLE, MEM, and ROKM, and the top 20 best-performing materials in terms of cost from each of these methodologies are reported in Fig. 1 for three different CO<sub>2</sub> compositions in the flue gas, namely, 7.5%, 13%, and 20%. The top 20 MOFs from the MAPLE-based screening were compared with MEM and ROKM approaches, and the common MOFs are highlighted in green. For the 7.5% CO<sub>2</sub> composition case, 8 out of 20 MOFs were featured in both MAPLE and MEM screening methodologies. On the other hand, 12 out of 20 were found

MOF ID			MOF ID			MOF ID		
MAPLE	MEM	ROKM	MAPLE	MEM	ROKM	MAPLE	MEM	ROKM
CUGLTM01	NIKDAM	UBUROY	CUGLTM01	NIKDAM	QEYWUN	MEHPAQ	QEYWUN	MEHPAQ
MEHPAQ	CUGLTM01	QEYWUN	ARIBOS	MEHPAQ	MEHPAQ	CUGLTM01	UBUROY	CUGLTM02
CUGLTM02	CUGLTM02	MEHPAQ	MEHPAQ	CUGLTM02	CUGLTM01	CUGLTM02	MEHPAQ	QEYWUN
CUGLTM	UBUROY	QUFFED	NELWIL	CUGLTM	ARIBOS	BEPPOD	CUGLTM01	CUGLTM
HIDBON	MEHPAQ	QUGNOV	CUGLTM02	QEYWUN	UBUROY	NELWIL	CUGLTM02	UBUROY
NELWIL	FURREE	CUGLTM02	UMELUU	CUGLTM01	CUGLTM02	TIVFEK	NIKDAM	ARIBOS
QUFFED	PURQEZ	NICJUG	HIDBON	IWEPUW	LUMZUO	HIDBON	CUGLTM	IWEPUW
RIYDEJ	CUGLTM	ARIBOS	NICJUG	LUMZUO	CUGLTM	UMELUU	PURQEZ	UMELUU
UMELUU	QEYWUN	LUMZUO	BEPPOD	ARIBOS	QUGNOV	NAKLIW	ZNGLUD01	NICJUG
QAVDEW	CATART03	NIKDAM	QUFFED	FURREE	QUFFED	LUMZUO	ARIBOS	LUMZUO
WAJHIA	VODPOU	IWEPUW	CUGLTM	XACZEH	VODPOU	ARIBOS	RIYDEJ	TIVFAG
ARIBOS	ZNGLUD	UGEPEB	RIYDEJ	ZNGLUD	TIVFAG	CUGLTM	ZNGLUD	WUDNAK
NICJUG	XACZEH	VODPOU	TIVFEK	QUGNOV	NICJUG	QUGNOV	XACZEH	QUGNOV
TIVFEK	ESUMOU	WAJHIA	NINHOH	PURQEZ	IWEPUW	NICJUG	LUMZUO	QUFFED
BOWQAG	NICJUG	UMELUU	QUGNOV	RIYDEJ	UMELUU	VODPOU	QUGNOV	TIVFEK
NIRDOI	CDLGLU02	PURQEZ	WAJHIA	QUFFED	TIVFEK	CAYDOX06	NICJUG	YERJEL
TIVFAG	NICZAA01	TIVFEK	BOWQAG	CDLGLU02	XACZEH	OKILEA	VODPOU	IFENOY02
VODPOU	ZNGLUD01	WUDNAK	QEYWUN	VODPOU	WAJHIA	KENJEU	IWEPUW	VODPOU
QEYWUN	IFENOY02	TIVFAG	IFENOY02	ZNGLUD01	WUDNAK	ICEGED	FURREE	XACZEH
QUGNOV	RIYDEJ	IFENOY02	CEHWOD	ESUMOU	UGEPEB	NINHOH	ESUMOU	WAJHIA

(a) (b) (c)

Figure 1: Top 20 best-performing MOFs obtained using MAPLE, MEM, and ROKM screening methodologies for (a) 7.5% (b) 13%, and (c) 20% CO<sub>2</sub> compositions in the flue gas. The common materials found in MAPLE, MEM, and ROKM are highlighted in green. The orange highlighted materials in MEM and ROKM were found in the Top 50 MOFs from MAPLE screening. The red ones were not found in the Top 50. Note that the MOFs are denoted with their index number in the list.

to be common for both MAPLE and ROKM. The MOFs highlighted in orange indicate that they appeared in the top 50 of the MAPLE-based ranking. This indicates that some of the top 20 MOFs from MEM and ROKM methodologies were also good-performing MOFs in the MAPLE-based screening. The red-shaded materials were not found in the top 50 of the MAPLE-based ranking. It is worth mentioning that the percentage differences in the minimum CO<sub>2</sub> avoided cost between the top-ranked and the 50<sup>th</sup>-ranked MOF in the MAPLE-based ranking for 7.5%, 13%, and 20% CO<sub>2</sub> composition cases are 94%, 39%, and 31%, respectively. Similarly, the analysis is extended to 13% and 20% CO<sub>2</sub> composition cases, and the common materials between MAPLE and ROKM approaches were found to be more than the matched materials between MAPLE and MEM approaches. The addition of simplified mass transfer kinetics in ROKM compared to MEM, which was only based on the equilibrium-based description of physics, could possibly be the reason for the improvement in the number of common materials. Note that the MAPLE model was trained on the data generated by the detailed process model that described the complete physics of adsorption column dynamics.

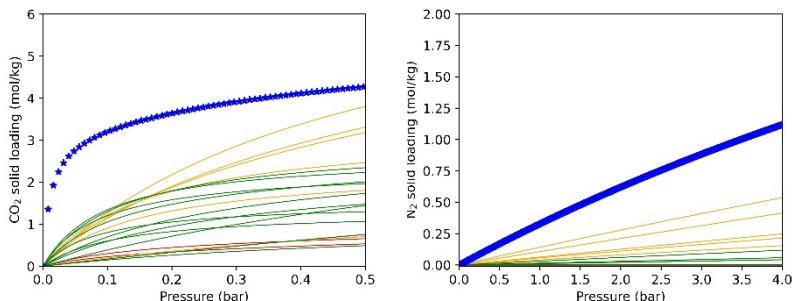


Figure 2: CO<sub>2</sub> and N<sub>2</sub> isotherms of the materials in the top 20 of the ROKM approach for the case of 13% CO<sub>2</sub> composition, out of which 12 were found in the MAPLE-based top 20 (green), 6 MOFs in the MAPLE-based top 50 (orange), and the remaining ones (red). Blue markers represent CO<sub>2</sub> and N<sub>2</sub> isotherms on zeolite 13X.

The CO<sub>2</sub> and N<sub>2</sub> isotherms of the top 20 MOFs from the ROKM-based screening approach for the case of 13% CO<sub>2</sub> composition are illustrated in Fig. 2. For comparison, CO<sub>2</sub> and N<sub>2</sub> isotherms of zeolite 13X are also shown as a reference. In Fig. 2, the green lines indicate CO<sub>2</sub> and N<sub>2</sub> isotherms of those MOFs that were found in the top 20 of both ROKM- and MAPLE-based screening. The orange lines represent the CO<sub>2</sub> and N<sub>2</sub> isotherms of the MOFs in the top 20 of the ROKM-based ranking that were also found in the top 50 of the MAPLE-based ranking. The red line isotherms are those MOFs in the top 20 of the ROKM-based ranking that were not found in the top 50 of the MAPLE-based ranking. It is interesting to notice that the best-performing materials have fairly linear CO<sub>2</sub> isotherms and very low N<sub>2</sub> adsorption. This observation remains consistent with several previous studies [1][13][14]. Finally, these common MOFs in the top 20 along with their normalized costs are tabulated in Table 2. It must be stressed that the main objective of the models presented, especially the simplified physical models, is to rank adsorbents rather than provide precise cost figures. Hence, this comparative analysis focused on the relative performances among the adsorbents.

Table 2: List of top-performing MOFs common in all three methodologies with normalized minimum CO<sub>2</sub> avoided cost for 13% CO<sub>2</sub> composition case. Note that the normalized CO<sub>2</sub> avoided cost was obtained by scaling the min. CO<sub>2</sub> avoided costs between the CO<sub>2</sub> avoided costs of the top-ranked MOF and the 20th-ranked MOF in each category.

MOF	Normalized minimum CO <sub>2</sub> avoided cost (-)		
	MAPLE model	MEM model	ROKM model
CUGLTM01	0.00	0.47	0.28
ARIBOS	0.05	0.57	0.31
MEHPAQ	0.11	0.14	0.18
CUGLTM02	0.39	0.16	0.37
QUFFED	0.73	0.81	0.53
CUGLTM	0.74	0.17	0.49
QUGNOV	0.79	0.71	0.53
QEYWUN	0.88	0.40	0.00

#### 4. Conclusions

Two types of computationally inexpensive modelling and optimization frameworks are assessed to enable rapid screening of adsorbents for post-combustion CO<sub>2</sub> capture using a four-step PVSA cycle. On the one hand, the data-driven-based MAPLE model was coupled with the cost model and NSGA – II optimizer to minimize the CO<sub>2</sub> avoided cost for the set of 369 MOFs to evaluate their techno-economic performance. On the other

hand, simplified physical models, MEM and ROKM, both with simpler descriptions of the physics of adsorption columns, were combined with Bayesian optimization to rank the set of MOFs in terms of minimum CO<sub>2</sub> avoided cost. The results showed that almost 60% of the MOFs in the top 20 best-performing materials ranked by each of these approaches were found to be common. The validation of these screening approaches against the detailed process simulation and optimization approach is currently ongoing.

## Acknowledgments

The authors are supported by the PrISMa project (No 299659) that is funded through the ACT Programme (Accelerating CCS Technologies, Horizon 2020 Project No 294766), and it receives financial contributions from BEIS, NERC, and EPSRC (UK), RCN (Norway), SFOE (Switzerland) and US-DOE (USA). The authors thank Professor Arvind Rajendran of the University of Alberta for graciously providing us with the data needed for the MAPLE model.

## References

1. S. G. Subraveti, S. Roussanaly, R. Anantharaman, L. Riboldi, A. Rajendran, 2021, Techno-economic assessment of optimised vacuum swing adsorption for post-combustion CO<sub>2</sub> capture from steam-methane reformer flue gas, *Sep. Purif. Technol.*, 256, 117832.
2. T. D. Burns, K. N. Pai, S. G. Subraveti, S. P. Collins, M. Krykunov, A. Rajendran, T. K. Woo, 2020, Prediction of MOF performance in vacuum swing adsorption systems for postcombustion CO<sub>2</sub> capture based on integrated molecular simulations, process optimizations, and machine learning models, *Environ. Sci. Technol.*, 54 (7), 4536-4544.
3. A. H. Farmahini, S. Krishnamurthy, D. Friedrich, S. Brandani, L. Sarkisov, 2021, Performance-based screening of porous materials for carbon capture, *Chem. Rev.*, 121 (17), 10666-10741.
4. K. T. Leperi, Y. G. Chung, F. You, R. Q. Snurr, 2019, Development of a general evaluation metric for rapid screening of adsorbent materials for postcombustion CO<sub>2</sub> capture, *ACS Sustain. Chem. Eng.*, 7 (13), 11529-11539.
5. K. N. Pai, V. Prasad, A. Rajendran, 2020, Generalized, adsorbent-agnostic, artificial neural network framework for rapid simulation, optimization, and adsorbent screening of adsorption processes, *Ind. Eng. Chem. Res.*, 59 (38), 16730-16740.
6. B. Maring, P. Webley, 2013, A new simplified pressure/vacuum swing adsorption model for rapid adsorbent screening for CO<sub>2</sub> capture applications. *Int. J. Greenh. Gas Control*, 15, 16-31.
7. V. S. Balashankar, A. K. Rajagopalan, R. de Pauw, A. M. Avila, A. Rajendran, 2019, Analysis of a batch adsorbent analogue for rapid screening of adsorbents for postcombustion CO<sub>2</sub> capture, *Ind. Eng. Chem. Res.*, 58 (8), 3314-3328.
8. S. G. Subraveti, Z. Li, V. Prasad, A. Rajendran, 2022, Physics-based neural networks for simulation and synthesis of cyclic adsorption processes, *Ind. Eng. Chem. Res.*, 61 (11), 4095-4113.
9. S. G. Subraveti, Z. Li, V. Prasad, A. Rajendran, 2019, Machine learning-based multiobjective optimization of pressure swing adsorption, *Ind. Eng. Chem. Res.*, 58 (44), 20412-20422.
10. L. Riboldi, C. Charalambous, E. Moubarak, R. Anantharaman, S. Roussanaly, C. Fu, B. Smit, J. Young, M. van der Spek, E. Sanchez-Fernandez, D. Ongari, S. Majumdar, E. Garcia-D'Arce, V. Kulakova, S. Garcia, 2020, Advanced methodology for screening of novel adsorption materials for cost-efficient CO<sub>2</sub> capture. *SSRN Electronic Journal*.
11. L. E. Andersson, J. Schilling, L. Riboldi, A. Bardow, R. Anantharaman, 2022, Bayesian Optimization for techno-economic analysis of pressure swing adsorption processes. *Comput Aided Chem Eng*, 51:1441-6.
12. S. Krishnamurthy, V. R. Rao, S. Guntuka, P. Sharratt, R. Haghpanah, A. Rajendran, M. Amanullah, I. A. Karimi, S. Farooq, 2014, CO<sub>2</sub> capture from dry flue gas by vacuum swing adsorption: A pilot plant study, *AIChE J.*, 60, 1830-1842.
13. A. K. Rajagopalan, A. Rajendran, 2018, The effect of nitrogen adsorption on vacuum swing adsorption based post-combustion CO<sub>2</sub> capture, *Int. J. Greenh. Gas Control*, 78, 437-447.
14. S. G. Subraveti, S. Roussanaly, R. Anantharaman, L. Riboldi, A. Rajendran, 2022, How much can novel solid sorbents reduce the cost of post-combustion CO<sub>2</sub> capture? A techno-economic investigation on the cost limits of pressure-vacuum swing adsorption, *Appl. Energy*, 306, A, 117955.

# A framework for decision-making to encourage utilization of residential distributed energy systems in Brazil

Ana Paula Alves Amorim,<sup>a</sup> Bogdan Dorneanu,<sup>b</sup> Karen Valverde Pontes,<sup>a</sup>  
Harvey Arellano-Garcia<sup>b</sup>

<sup>a</sup>*Programa de Pós Graduação em Engenharia Industrial – PEI, Universidade Federal de Bahia, Salvador, 40210-630, Brazil*

<sup>b</sup>*LS Prozess- und Anlagentechnik, Brandenburgische Technische Universität Cottbus-Senftenberg, Cottbus, D-03046, Germany*

arellano@b-tu.de

## Abstract

The Distributed Energy Systems (DES) or microgrid arose from the need to reduce greenhouse gases (GHG) emitted into the atmosphere by burning fossil fuels to generate energy. Reduction of energy losses, reconfiguration of the protection system and reduction of costs, and optimizing the configuration of these systems is recommended. Despite new research in literature, there is still a lack of optimization models that address the Brazilian reality. Therefore, the objective of this work is to introduce a decision-making framework for the design and operation of residential DES that takes into account the particularities of Brazil, based on mixed-integer programming models. The applicability of the framework is tested on a case study of a residential DES of 5 houses, located in Salvador, and used to compare scenarios pre- and post-COVID-19. The results show significant reduction in total annual cost and GHG emissions *versus* the base case without DES. This indicates that, although the country has a mostly “clean” energy matrix due to the use of hydroelectric plants, DES can enable improvement in residential electricity generation.

**Keywords:** Distributed energy systems, Microgrid, Mixed-integer non-linear programming, Net metering.

## 1. Introduction

Currently, energy demand is generally provided by fossil fuels, which release large amounts of suspended particles and greenhouse gas (GHG), leading to global warming and large-scale shifts in climate patterns (Mahmoudan et al., 2022). Distributed energy systems (DES) can improve environmental sustainability and provide an efficient way to achieve the objective of reducing dependence on fossil fuels (Clarke et al., 2021). The adoption of distributed generation from renewable sources is still understated in Brazil, compared to other countries in the world. This can be attributed to regulatory challenges and high acquisition costs of renewable generation technologies. Furthermore, the Brazilian energy matrix is strongly influenced by significant availability of renewable resources (Lima et al., 2020). For the use of this high renewable energy potential that Brazil has, public policies and projects that aim to further encourage the installation of microgrids in the country are needed. Thus, the present work aims to introduce a decision-making framework for the design and operation of residential DES based on mixed-integer programming models that take into consideration the particularities of Brazil. The

models optimize the cost of the design and utilization of a microgrid with integrated pipelines to supply hot water and biogas besides electricity. The time horizon for operating the DES is split into time periods based on clustering methods. Furthermore, the efficiency of the various technologies to be installed is considered to vary in time, and several different options are included for the renewable incentive policy, which results in a mixed-integer nonlinear programming (MINLP) model, *versus* the commonly used mixed-integer linear programming (MILP) models in literature (Sidnell et al., 2021).

## 2. Problem definition

The aim of the optimization problem is to scale the system at a lower cost in order to meet the electrical, hot water, air cooling, and gas for cooking demand of a residential network in Salvador (Brazil) using three energy sources: solar, wind and biogas. The technologies that can be installed in each household to meet the electricity demand are photovoltaic (PV) arrays, wind turbines (WT), biogas generators (BG), and batteries (BT). To meet the hot water demand, electric showers (ES), gas heaters (GH), biogas heaters (BH), and thermosolar collectors (SC) are used. To meet the air cooling demand, air conditioners (AC) are used, while a biodigester (BD) will satisfy both the biogas demand for the BH and cooking for the whole network. Furthermore, the electricity generated from renewable resources is used for self-consumption, transferred to another household, fed into the central grid, or stored in electric batteries. To meet demand, in addition to the renewable energy generated, electricity can be purchased from the central grid. Regarding hot water production, this can be used for self-consumption, transferred to another house, or it can be produced in SC, if available. Moreover, the demand for cooking gas can be supplied by natural gas from the grid or by biogas produced in the BD.

To improve on previous work (Sidnell et al., 2021a, Clarke et al., 2021), the time horizon is divided in time periods by means of a systematic clustering method that takes into account the profiles of electrical demand, solar irradiation and wind speed of the site, and the central grid electricity rate that varies during the day. The division is done for each pre- and post-COVID scenarios. Thus, for the pre-COVID scenario three seasons are obtained (*m1*: January, February, March, April, *m2*: May, June, July, August, September, *m3*: October, November, December), with eight periods during the day (*p1*: 12am-3am, *p2*: 3am-6am, *p3*: 6am-8am, *p4*: 8am-4pm, *p5*: 4pm-6pm, *p6*: 6pm-9pm, *p7*: 9pm-10pm, *p8*: 10pm-12am), while for the post-COVID scenario the same three seasons are obtained, but with only five periods (*p1*: 12am-7am, *p2*: 7am-4pm, *p3*: 4pm-6pm, *p4*: 6pm-9pm, *p5*: 9pm-12am). The irradiation and wind speed data used are found in World Bank Group (2022) and DTU (2022). Based on the approximation of the depreciation time of the components of the energy sources used, the useful life of the project is considered to be 20 years. For the interest rate, a value of 0.07 is used, according to the financial cost rate for renewable energy adopted by BNDES (*Banco Nacional do Desenvolvimento Econômico e Social* – The National Bank for Economic and Social Development). The purchase and sale tariff for the central grid is based on the amount adopted by COELBA (*Companhia de Eletricidade do Estado da Bahia* – Electricity Company of Bahia). Because no studies are available on the post-COVID average energy demand, they are estimated in relation to the pre-COVID data. For the electricity demand, this value is 25% higher, on the assumption of continuous behavior during the day, due to the remote work adopted (Macedo, 2021). For the hot water demand, an increase of 15% was considered; for the cooling demand, an increase of 50%, while for cooking gas, an increase of 25% is assumed. In order to determine the best incentive policy to be adopted for distributed generation, the Net Metering (NEM) and Feed in Tariff (FIT) models were considered.

### 3. Mathematical model

The model is implemented as a mixed-integer nonlinear programming (MINLP) problem and solved in the GAMS language. The model is solved on a desktop PC with Intel Core i9-9900K CPU @ 3.60 GHz and 32.0 GB of RAM, using the BARON solver, to full optimality.

#### 3.1. Model formulation

The objective function to be minimized is given by the total annual investment and annual operation cost, described by:

$$C_{TOTAL} = C_{INV} + C_{OM} + C_{BUY}^{GRID} + C^{NG} + C^{ENV} - NEM - FIT \quad (1)$$

where  $C_{INV}$  is the total investment,  $C_{OM}$  is the operating and maintenance cost,  $C_{BUY}^{GRID}$  is the total cost of electricity purchased from the grid,  $C^{NG}$  is the cost to purchase natural gas from the grid,  $C^{ENV}$  is the environmental costs, NEM is the credit revenue obtained by using the Net Metering policy, FIT is the amount received for adopting a Feed in Tariff. All these costs and revenues are measured in Brazilian Reals (R\$).

The investment is the sum of all the costs invested to acquire each technology:

$$C_{INV} = C_{OM}^{PV} + C_{OM}^{BG} + C_{OM}^{WT} + C_{OM}^{BT} + C_{OM}^{BD} + C_{OM}^{PPBD} + C_{OM}^{MG} + C_{OM}^{TS} + C_{OM}^{SC} + C_{OM}^{GH} + C_{OM}^{BH} + C_{OM}^{ES} + C_{OM}^{AC} \quad (2)$$

where the superscript OM refers to the maintenance and operation cost. The fixed and variable operating and maintenance costs for each technology are given by the nominal equipment capacity and their operating cost.

The cost of buying power from the central grid is given by:

$$C_{BUY}^{GRID} = \sum_{i,m,p} (E_{i,m,p}^{GRID} \cdot P_{m,p}^{Elec} \cdot day(m) \cdot hours(p) \cdot season(m)) \quad (3)$$

Here,  $E_{i,m,p}^{GRID}$  is the power purchased from the central grid [kW], while  $P_{m,p}^{Elec}$  is the price of electricity purchased during each time period [R\$/kWh], the subscripts  $i$  represent the houses, while the season is indicated by the index  $m$  and the time period by the index  $p$ .

The cost of buying natural gas from the grid is determined from:

$$C^{NG} = \frac{P^{NG}}{q^{NG}} \cdot \sum_{i,m,p} (Use_{i,m,p}^{NG} \cdot day(m) \cdot hours(p) \cdot season(m)) + \sum_i Y_i^{NG} \cdot C_{Fix}^{NG} \cdot season(m) \quad (4)$$

This cost depends on variable and fixed factors. The variable cost is determined by the volume of gas purchased from the grid,  $Use_{i,m,p}^{NG}$  [kW], the price of gas,  $P^{NG}$  [R\$/m<sup>3</sup>], the calorific capacity of the natural gas,  $q^{NG}$  [kW/m<sup>3</sup>], and the number of hours that the gas has been used. The fixed costs,  $C_{Fix}^{NG}$  [R\$] is related to the monthly amount calculated only when natural gas is purchased from the grid. The variable  $Y_i^{NG}$  identifies whether natural gas is used or not from the grid.

The environmental costs are calculated from the carbon tax per amount of CO<sub>2</sub> emitted for each technology considered:

$$C^{ENV} = CO_2^{GRID} + CO_2^{NG} + CO_2^{WT} + CO_2^{PV} + CO_2^{SC} + CO_2^{BIO} \quad (5)$$

The credits obtained by each incentive policy considered (NEM and FIT, respectively) depend on the total energy sold to the grid and its selling rate. In the case of NEM, it corresponds to the electricity rate adopted by the central grid. For the FIT, the selling rate varies for each technology used, as illustrated in the following equation :

$$FIT = \sum_{i,m,p} (E_{i,m,p}^{SALE,PV} \cdot P^{FIT,PV} + E_{i,m,p}^{SALE,WT} \cdot P^{FIT,WT} + E_{i,m,p}^{SALE,BG} \cdot P^{FIT,BG}) \cdot day(m) \cdot hours(p) \cdot season(m) \cdot Y^{FIT} \quad (6)$$

Where  $E_{i,m,p}^{SALE,Tech}$  is the amount of the energy fed into the central grid by the technology  $Tech$  (kW),  $P^{FIT,Tech}$  is the rate of the sale for technology  $Tech$  (\$/kW), with  $Tech = PV, WT, BG$ , and  $Y^{FIT}$  is a binary variable for deciding if the policy is being used or not.



### 3.2. Design and operational constraints

The constraints found in Sidnell et al. (2021a) and Clarke et al. (2021) are adopted for the operation of the microgrid, the pipeline, and the battery, while the operation of the photovoltaic panel and solar collector are described based on Karamov et al. (2021), while for the wind turbines, the constraints are defined as described in Pallabazzer (2003). For the sizing of the biodigester, Araujo (2017) and Otim et al. (2006) are used. Additionally, the model considers that the efficiency of the technologies installed in dwellings is not constant, but varies with the local conditions. For instance, in the case of the PV panels, the efficiency is defined as (Karamov et al., 2021):

$$ne_{m,p}^{PV} = nee^{PV} \cdot [(1 - \beta_s) \cdot T_{m,p}^{PV/SC} - 48] \quad (7)$$

where  $nee^{PV}$  is the nominal energy efficiency of the PV panel,  $\beta_s$  is the temperature coefficient for silicon PV panels, and  $T_{m,p}^{PV/SC}$  is the operating temperature of the PV converter.

## 4. Case study

The model presented in the previous section is applied to a microgrid of 5 houses located in Salvador da Bahia (Brazil) for two different scenarios: pre- and post-COVID. The results of these scenarios are illustrated in Figure 1 (a) and (b), respectively.

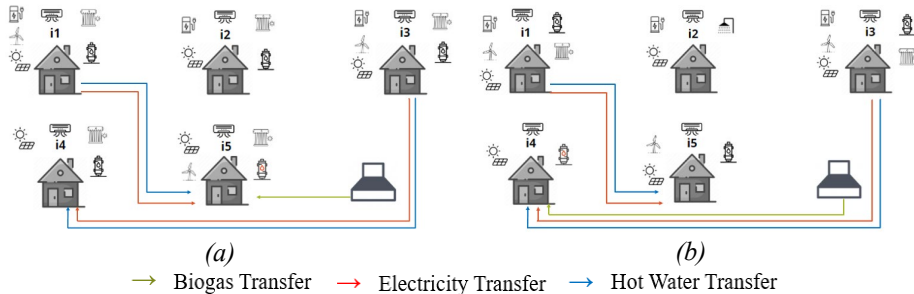


Figure 1: DES network for the scenario: (a) pre-COVID-19, (b) post-COVID-19

For the pre-COVID scenario, PV panels (for electricity generation) and SC (for hot water) are installed in every home, due to the fact that the period of sunlight in the region is distributed over more than one period ( $p3$  to  $p5$ ), which means the best use of sunlight. In addition, for the electricity generation, WT are considered in houses  $i1$ ,  $i3$  and  $i5$ . Thus, house  $i1$  produces more than what was necessary to meet its own demand, enabling energy sharing with house  $i5$ . A similar behavior is observed between houses  $i3$  and  $i4$ . Batteries are installed in houses  $i1$ ,  $i2$  and  $i3$ . It is inferred that it is more efficient to install smaller batteries in  $i1$  and  $i3$  that share energy with the houses of higher demand ( $i4$  and  $i5$ ) than to install a larger battery in these houses. For house  $i2$ , due to having only PVs to meet the demand, the battery is used at times when energy from the central grid is more expensive (e.g., peak hour periods  $p5$ - $p7$  in the pre-, or  $p3$ - $p5$  in the post-COVID scenario, respectively), as illustrated in Figure 2 for both pre- and post-COVID scenarios.

For hot water generation, besides the SC, a BH is installed in house  $i5$  and a natural GH in houses  $i1$ ,  $i2$ ,  $i3$  and  $i4$ , respectively. This is due to the limitation of biogas production, considering it is better to use it in the dwelling with higher demand ( $i5$ ). Since natural gas has a continuous supply, it is used in the other houses to meet the rest of the demand that the SC could not meet. In addition, house  $i5$  receives hot water from house  $i1$ , and house  $i4$ , which has the highest demand, receives hot water from house  $i3$ . Furthermore, AC is used to meet the cooling demand in all the houses. Finally, a BD is installed in the network

and biogas is distributed to house *i5* to meet the gas demand for cooking and hot water generation.

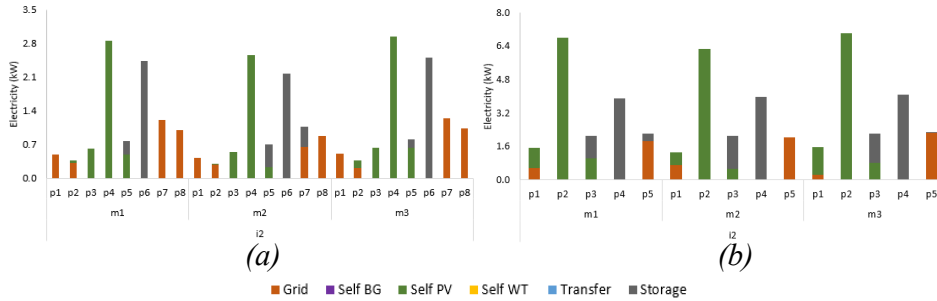


Figure 2 - Profile of resources used to meet the electrical demand in the house *i2* for the scenario: (a) pre-COVID-19, (b) post-COVID-19

The sharing of electricity and hot water is designed to go from the house of lower to the one with higher demand, to ensure the same equipment can serve more than one house. Due to a limitation in the capacity, the opposite direction of sharing is not possible, as this will require equipment above the upper bound of the capacity.

In the post-COVID scenario, PV panels are installed in all residences, but only houses *i1*, *i3* and *i5* adopt the WT. Because houses *i4* and *i5* have a higher electricity demand, the production of houses *i1* and *i3* is shared with these houses, as illustrated in Figure 1. Furthermore, as in the pre-COVID scenario, the battery is installed in houses *i1*, *i2* and *i3*, with lower demand, which share energy with the houses of higher demand (*i4* and *i5*). Although the same type of equipment is installed for electricity generation in both scenarios, their capacity is different due to differences in the distribution of the demand between the pre- and post-COVID scenario. To meet the hot water demand in houses *i1* and *i3*, natural GHs and SCs are installed. Since these houses also share hot water with houses *i5* and *i4*, respectively, and have a lower hot water demand, the model chooses the most economically feasible option. This is to install the two technologies to meet the demands of the 4 houses than to install, for example, a bigger equipment in houses *i4* and *i5*.

Moreover, a BH is installed to supply the rest of the hot water demand in house *i4* and the natural GH in house *i5*. In house *i2* the ES is sufficient to meet all its hot water demand. Furthermore, AC is used to meet the cooling demand in all the houses. A BD is installed in the network and connected to house *i4*, to supply part of the gas demand for cooking. This can be explained by the limited biogas production and the lower demand of house *i4* compared to house *i5*, for example.

In terms of renewable incentives policy, for all scenarios, the option selected is net metering. This is the current policy adopted in the country and confirms that, for the current conditions of technology and local electricity prices, it continues to be the most advantageous.

Table 1: Costs and revenues for all scenarios (all costs are in R\$)

Scenario		Total	Investment	O&M	Environmental	Grid imports	Incentives
Pre-COVID	Baseline	160,210.26	990.15	84,672.04	2,015.82	72,532.24	-
	5Houses	28,956.95	24,450.86	38,539.39	1,428.47	3,445.71	38,907.48
Post-COVID	Baseline	249,889.14	1,398.02	109,351.27	3,145.50	135,994.34	-
	5Houses	31,596.10	28,056.46	80,756.17	2,562.79	7,741.56	87,520.89

Table 1 shows the costs and revenues for all scenarios. In relation to the pre-COVID baseline, there is an 82% reduction in total cost and a 29% reduction in environmental cost. In addition, R\$38,907.48 in energy credits are generated. For the post-COVID scenario, a similar reduction is achieved in terms of total cost (87% versus the baseline), while lower savings (19%) are achieved for the environmental cost. However, for this case, higher energy credits (R\$87,520.89) are generated due to the higher resource capacity installed compared to the pre-COVID design.

Furthermore, despite a 25% increase in electricity cost, 50% increase in gas cost, and the increase in demand mentioned above, the optimizer is able to find a configuration with an increase of only 9% in total cost for the post-COVID scenario.

## 5. Conclusion

This paper presents the development of a framework to support decisions in terms of residential DES policies applied to Brazil. The proposed framework considers the division of time periods using clustering methods, variable efficiency of the installed technologies, several options for the selection of renewable incentive policies and is able to attest the potential high economic and environmental gains that can be achieved with the adoption of renewable energy resources, for various scenarios considered. The results show that the approach is able to provide information regarding the best incentive policy for the country and the viability of using biogas, solar and wind energy to meet the electrical, hot water, air cooling and gas for cooking demands. Furthermore, significant reduction of economic and environmental costs is achieved, despite the clean energy matrix of Brazil. For future work, a larger set of houses in order to observe the behavior of the microgrid sizing in other configurations, uncertainty in generation or demand, or variable energy pricing will be included.

## References

- A.P.C. Araújo, 2017, Produção de biogás a partir de resíduos orgânicos utilizando biodigestor anaeróbico, Universidade Federal de Uberlândia, pp. 42
- F. Clarke et al., 2021, Optimal design of heating and cooling pipeline networks for residential distributed energy systems, *Energy* 235, 121430
- DTU Wind Energy, 2022, Global Wind Atlas. Available in: <https://globalwindatlas.info/> (Accessed: Mar. 15, 2021)
- D.N. Karamov et al., 2021. Optimization of isolated power systems with renewables and storage batteries based on nonlinear Volterra models for the specially protected natural area of lake Baikal, *Journal of Physics: Conference Series* 1847(1), 012037
- M.A. Lima et al., 2020, Renewable energy in reducing greenhouse gas emissions: Reaching the goals of the Paris agreement in Brazil, *Environmental Development* 33, 100504
- A. Mahmoudan et al., 2022, A geothermal and solar-based multigeneration system integrate with a TEG unit: Development, 3E analyses, and multi-objective optimization, *Applied Energy* 308, 118399
- G. Otim et al., 2006, Design of Biogas Plant for Rural Households in Uganda (Case Study Apac District), *Second International Conference on Advances in Engineering and Technology*, pp. 544-550
- R. Pallabazzer, 2003, Parametric analysis of wind siting efficiency, *Journal of Wind Engineering and Industrial Aerodynamics* 91(11), pp.1329–1352
- T. Sidnell et al., 2021, Optimal design and operation of distributed energy resources systems for residential neighbourhoods, *Smart Energy* 4, 100049
- World Bank Group, 2022, Global Solar Atlas, Available in: <https://globalsolaratlas.info/map?c=-13.859414,-42.220459,7&s=-12.972442,-38.441162&m=site&pv=small,0,12,1> (Accessed: Mar. 15, 2021)

# Hydrogen infrastructure planning for heat decarbonisation in Great Britain

Margarita E. Efthymiadou, Vassilis M. Charitopoulos and Lazaros G. Papageorgiou

*The Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, University College London (UCL), Torrington Place, London WC1E 7JE, UK*

*[l.papageorgiou@ucl.ac.uk](mailto:l.papageorgiou@ucl.ac.uk)*

## Abstract

Towards Net-Zero emissions goal in 2050, the UK has set a greenhouse gas emissions target to reduce its environmental footprint. Consequently, the exploitation of alternative low-carbon pathways and energy carriers for the heat sector are required. In this work, we propose a multi-period, spatially-explicit mixed-integer linear programming (MILP) optimisation framework for hydrogen infrastructure to meet hydrogen heating demand in Great Britain. The mathematical model aims to minimise the total cost accounting for investment and operational decisions. The proposed optimisation framework includes dual temporal resolution: 5-year steps 2035- 2050 and typical days with hourly resolution. To enhance the computational performance of the studied model, we developed a hierarchical solution approach that results in near-optimal solutions while reducing the solution time by an order of magnitude.

**Keywords:** Net-Zero; Heat Decarbonisation; Hydrogen Infrastructure; Energy systems modelling; MILP model;

## 1. Introduction

In 2021, UK greenhouse gas emissions reached 447 MtCO<sub>2</sub>e (BEIS, 2022), with target reduction of 68% from 2019 to 2035 (BEIS, 2020). The domestic heating sector is responsible for 14% of UK emissions mainly due to the wide use of natural gas boilers (McDowall & Britchfield, 2021). The pathway to energy decarbonisation requires environmentally friendly policies, which support the introduction of low-carbon energy carriers and end-use technologies. Regarding the heat sector, the UK government has made several commitments to achieve heat decarbonisation allocating over £1 billion to support them (HM Government, 2021). In this context, BEIS (2021) forecasted that hydrogen demand will increase significantly by early 2030 (suggesting production capacity of 7-20 GW by 2035). Among other alternatives, hydrogen play a key role in the future energy mix either with its direct use or as an efficient storage carrier for renewable electricity production. In both cases, new infrastructure networks connecting supply to demand are required and thus, a novel modelling tool for hydrogen supply chain is essential to investigate design decisions and what-if analysis scenarios for hydrogen infrastructure.

The supply chain optimisation of hydrogen has received considerable attention from the PSE community over the last decade. One of the first infrastructure planning optimisation models was developed by Hugo *et al.* (2005) using a multi-objective optimisation approach. An integrated MILP spatially-explicit framework for hydrogen transportation demand infrastructure design was proposed by Almansoori and Shah (2009). Based on

the aforementioned work, Guillen *et al.* (2010) formulated a bi-criterion model for simultaneous minimisation of system cost and environmental impact of the system. Agnolucci *et al.* (2013) introduced SHIPmod, which included Carbon Capture and Storage (CCS) system in the spatially-explicit multi-period MILP model. SHIPmod was extended by Moreno-Benito *et al.* (2016) adding hydrogen pipelines for regional transmission. The role of oxygen as a by-product of hydrogen production for transportation fuel was examined by Ogumerem *et al.* (2018) using a multi-objective approach. Sunny *et al.* (2020) and Samsatli and Samsatli (2019) investigated the role of hydrogen in heat decarbonisation. He *et al.* (2021) proposed a snapshot spatial model which determines the least-cost hydrogen planning for multiple end-uses. Stochastic approaches for hydrogen supply chains have been studied by Camara *et al.* (2019) and Ochoa-Bique *et al.* (2021). This work focuses on the development of an optimisation-based framework for hydrogen infrastructure planning while its applicability is demonstrated through a case study for residential hydrogen heat demand in Great Britain.

## 2. Problem Statement

Overall, the problem investigated in this work can be stated as follows:

**Given:** (i) H<sub>2</sub> heating demand and renewables availability in each region, year, cluster and hour, (ii) capital and operating costs for production technologies, storage sites, hydrogen and CO<sub>2</sub> pipelines and road transportation modes, (iii) minimum and maximum capacity as well as ramp rates and lifetime of production plants and storage sites, (iv) minimum and maximum flowrate in pipelines, (v) capacity of H<sub>2</sub> caverns and CO<sub>2</sub> reservoirs, (vi) H<sub>2</sub> import price, (vii) carbon tax and capture rates for CO<sub>2</sub> emissions as well as CO<sub>2</sub> emission targets

**Determine the optimal:** (i) location and capacity of production technologies, storage sites and renewable farms, (ii) H<sub>2</sub> production and storage rate in each region, year, cluster and hour, (iii) H<sub>2</sub> and CO<sub>2</sub> transmission investments between regions, (iv) H<sub>2</sub> and CO<sub>2</sub> flowrates between regions in year, cluster and hour, (v) electricity generation of renewables, (vi) H<sub>2</sub> import rates in each year, cluster and hour

**So as** to minimise the total system cost subject to emission target.

## 3. Mathematical Formulation and Hierarchical Solution Approach

The problem for the cost optimal hydrogen infrastructure design is formulated as a multi-period mixed integer linear programming (MILP) model.

The objective of the model is the minimisation of the total system cost ( $TC$ ) as described by the Eq. (1).

$$\min TC = PCC + SCC + TCC + POC + SOC + TOC + CEC + IIC + ReC \quad (1)$$

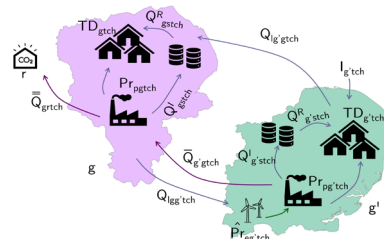
The total cost comprised of production capital cost ( $PCC$ ), storage capital cost ( $SCC$ ), transportation capital cost ( $TCC$ ), production operating cost ( $POC$ ), storage operating cost ( $SOC$ ), transportation operating cost ( $TOC$ ), carbon emissions cost ( $CEC$ ), international import cost ( $IIC$ ) and the renewable cost ( $ReC$ ).

The hydrogen energy balance can be described by the Eq. (2).

$$\begin{aligned} & \sum_{p \in P} Pr_{pgtch} + \sum_{l \in \{Road\}} \sum_{g' \in N_{gg'}^r} Q_{lg'gtch} + \sum_{l \in \{Pipe\}} \sum_{g' \in N_{gg'}^p} Q_{lg'gtch} + \sum_{s \in GS_{gs}} Q_{gstch}^R + I_{gtch} \\ & = \sum_{l \in \{Road\}} \sum_{g' \in N_{gg'}^r} Q_{lgg'tch} + \sum_{l \in \{Pipe\}} \sum_{g' \in N_{gg'}^p} Q_{lgg'tch} + \sum_{s \in GS_{gs}} Q_{gstch}^I + TD_{gtch} \end{aligned} \quad (2)$$

$$\forall g \in G, t \in T, c \in C, h \in H$$

More specifically, in each region  $g$ , year  $t$ , cluster  $c$  and hour  $h$ , the total production rate of production technologies  $p$  ( $Pr_{pgtch}$ ), the flowrates of all transportation modes  $l$  ( $Q_{lg'gtch}$ ) to region  $g$ , the rejected hydrogen from storage site  $s$  ( $Q_{gstch}^R$ ) and the imported hydrogen ( $I_{gtch}$ ) are equal to the flowrates ( $Q_{lgg'tch}$ ) from region  $g$ , the injected hydrogen to storage sites  $s$  ( $Q_{gstch}^I$ ) and the total demand ( $TD_{gtch}$ ).  $GS_{gs}$



**Figure 1:** Visual representation of energy and mass balance between two regions.  $GS_{gs}$  stands for the regions where storage sites can be commissioned and  $N_{gg'}^r$  and  $N_{gg'}^p$  stand for the neighboring regions for road and pipeline connections respectively.

Due to number of technologies and high spatio-temporal resolution, the resulting optimisation problem is usually computationally intensive. Typical computational times for solving the monolithic model exceed 15h. Reduction of CPU time could be achieved by developing hierarchical solution procedures. In this context, a hierarchical approach has been developed in this work, which includes the steps as presented in Table 1.

**Table 1:** Solution procedure

<b>Step 1</b>	Solve the model without considering pipeline investment decisions
<b>Step 2</b>	Fix production and storage investment decisions (integer variables) for all time periods
<b>Step 3</b>	Solve reduced model

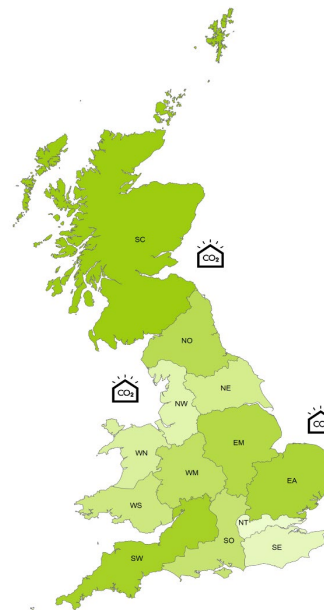
#### 4. Case Study

The investigated case study demonstrates the applicability of the multi-period spatially-explicit MILP for hydrogen infrastructure optimisation for domestic heating in Great Britain (GB) from 2035 to 2050. GB is divided into 13 regions according to local gas distribution zones (LDZ) of the incumbent natural gas network. The temporal resolution of the model is dual considering 5-year time steps for investment decisions and representative days with hourly resolution for operational decisions, such as hydrogen production and storage rates.

This case study includes hydrogen production, storage and transmission technologies and a carbon capture and storage (CCS) system. Moreover, the electricity required for water electrolysis is generated from renewable sources including wind and solar energy.

In the context of model size reduction, the total heat demand and renewable technologies availability hourly data are clustered using clustering techniques. K-Medoids clustering is used for representative days selection according to Charitopoulos *et al.* (2022).

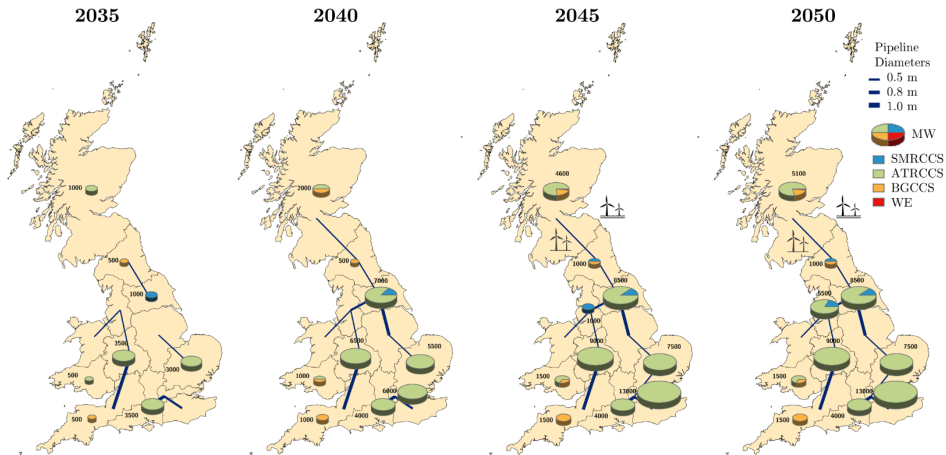
The system-wide hydrogen demand and emission targets for residential heat are obtained from National Grid ESO (2021) System Transformation scenario.



**Figure 2:** Regional representation of Great Britain.

## 5. Results & Discussion

The model is implemented in GAMS 38.2.1 and solved with Gurobi 9.1.2 with optimality gap 5%, using a Dell workstation with Intel® Core™ i9-10980XE CPU @ 3.00 GHz and 128 GB RAM. The multi-period model results provide the optimal cost evolution of hydrogen infrastructure in the UK to meet hydrogen heat demand.



**Figure 3:** Production capacity and hydrogen pipelines expansion maps from 2035 to 2050.

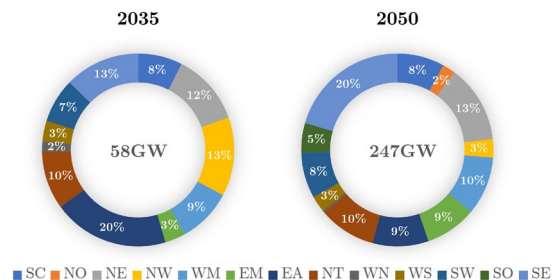
In Fig. 3, the production capacity expansion maps are illustrated from 2035 to 2050. In 2035, a total of 13.5 GW production capacity is commissioned of ATR, SMR and biomass gasification technologies. The total capacity is increased to 33.5 GW in 2040 and 56.6 GW in 2050. Reforming technologies are mostly installed due to their lower production costs in comparison with the other technologies. Moreover, 0.1 GW of water electrolysis for green hydrogen production is commissioned in 2045 in Scotland while the electricity required for electrolysis is generated from wind onshore and offshore farms.

The regional transmission of hydrogen takes place through a pipeline network. As shown in Fig. 3, in 2035 there is a partial installation of the hydrogen transmission network. Over the years, the hydrogen network is expanded connecting most neighboring regions.

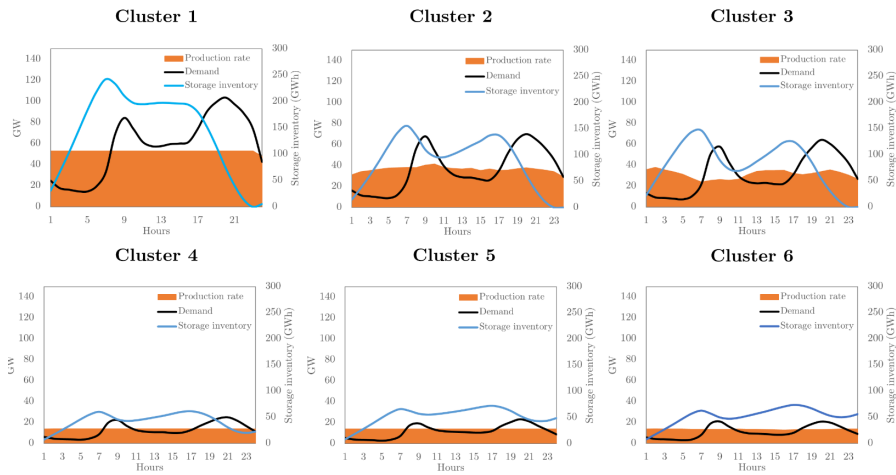
Storage is a key element in hydrogen infrastructure strategy to meet peak heat demand.

Fig. 4 illustrates the evolution of regional storage capacity from 2035 to 2050. Medium pressure storage vessels are installed for hydrogen storage in the majority of GB regions. In 2050, a total of 247 GW is commissioned with 50GW installed in EA and 32 GW in NE.

Fig. 5 illustrates the production rate, storage inventory and demand in each cluster in 2050. Cluster 1 is the peak day of the year in which hydrogen heating demand reaches 100 GW. As indicated, hydrogen production rate during the day is almost steady as there is a significant installation of gas reforming and biomass gasification units which have low ramp rates.



**Figure 4:** Hydrogen storage capacity in GB regions from 2035 to 2050.



**Figure 5:** Demand, production rate and storage inventory profiles in each cluster in 2050.

The resulting model includes 299,304 equations, 178,162 continuous variables and 776 discrete variables. As it is computationally intensive, the hierarchical approach described in Section 3 is implemented to enhance solution performance. Table 2 summarises the computational statistics for each solution scheme.

**Table 2:** Computational performance of monolithic and hierarchical approaches.

Approach	Monolithic	Hierarchical			
		Step 1	Step 2	Step 1	Step 2
<b>Optimality gap limit (%)</b>	5	5	5	1	1
<b>Optimality gap (%)</b>	4.74	1.00	2.04	1.00	0.88
<b>CPU time (h)</b>	15.48	0.83	0.20	0.87	3.55
<b>Objective value (£b)</b>	71.1	71.4		71.1	

It can be noticed in Table 2 that hierarchical strategy leads to a feasible solution while the computational time is reduced up to 93% without significant compromise of the solution quality. For 1% optimality gap limit, the hierarchical approach can result in the same objective value with the monolithic strategy while it improves CPU time by 72%.

### 6. Concluding Remarks

In this work, an optimisation-based framework has been proposed to facilitate the investigation of design and operating decisions in hydrogen infrastructure in the UK. Reforming technologies with CCS constitute the most cost-effective low carbon alternative for hydrogen production. Hydrogen transmission is based on a pipeline network connecting most neighboring regions in GB. This work also focuses on a hierarchical solution procedure to deal with the combinatorial complexity of the resulting model. A hierarchical approach has been developed to tackle efficiently the large-scale optimisation problem and enables greater model complexity while it provides a near-optimal solution reducing computational time up to 93%.

Future research includes sector coupling of heat and power sector to provide a more holistic approach for heat decarbonisation future decisions. Moreover, the introduction of uncertainty in hydrogen supply chains is a future step for a more risk averse infrastructure strategy in parallel with the exploration of new solution approaches and decomposition techniques to deal with the high computational times.



## Acknowledgements

Authors gratefully acknowledge the financial support from Engineering and Physical Sciences Research Council (EPSRC) under the project EP/T022930/1.

## References

- Agnolucci, P., Akgul, O., McDowall, W., Papageorgiou, L. G., 2013, The importance of economies of scale, transport costs and demand patterns in optimising hydrogen fuelling infrastructure: An exploration with SHIPMod (Spatial hydrogen infrastructure planning model), *Int. J. Hydrog. Energy*, 38, 11189–11201.
- Almansoori, A., Shah, N., 2009, Design and operation of a future hydrogen supply chain: multi-period model, *Int. J. Hydrog. Energy* 34, 7883–7897.
- Câmara, D., Pinto-Varela, T., Barbósa-Povoa, A. P., 2019, Multi-objective optimization approach to design and planning hydrogen supply chain under uncertainty: A Portugal study case, *Comput. Aided Chem. Eng.*, 46, 1309–1314.
- Charitopoulos, V. M., Fajardy, M., Chyong, C. K., Reiner, D. M., 2022, The case of 100% electrification of domestic heat in Great Britain, Faculty of Economics, University of Cambridge.
- Department for Business, Energy & Industrial Strategy (BEIS), 2020, Policies for the Sixth Carbon Budget and Net Zero, [www.theccc.org.uk/wp-content/uploads/2020/12/Policies-for-the-Sixth-Carbon-Budget-and-Net-Zero.pdf](http://www.theccc.org.uk/wp-content/uploads/2020/12/Policies-for-the-Sixth-Carbon-Budget-and-Net-Zero.pdf).
- Department for Business, Energy & Industrial Strategy (BEIS), 2021, UK Hydrogen Strategy, [www.gov.uk/government/publications/uk-hydrogen-strategy](http://www.gov.uk/government/publications/uk-hydrogen-strategy).
- Department for Business Energy & Industrial Strategy (BEIS), 2022, Progress in reducing emissions 2022 Report to Parliament, [www.theccc.org.uk/publication/2022-progress-report-to-parliament](http://www.theccc.org.uk/publication/2022-progress-report-to-parliament).
- Guillén-Gosálbez, G., Mele, F. D., Grossmann, I. E., 2010, A bi-criterion optimization approach for the design and planning of hydrogen supply chains for vehicle use, *AIChE J.*, 56, 650–667.
- He, G., Mallapragada, D. S., Bose, A., Heuberger, C. F., Gencer, E., 2021, Hydrogen supply chain planning with flexible transmission and storage scheduling. *IEEE Trans. Sustain. Energy*, 12, 1730–1740.
- HM Government, 2021, Net Zero Strategy: Build Back Greener, [www.gov.uk/government/publications/net-zero-strategy](http://www.gov.uk/government/publications/net-zero-strategy).
- Hugo, A., Rutter, P., Pistikopoulos, E.N., Amorelli, A., Zoia, G., 2005, Hydrogen infrastructure strategic planning using multi-objective optimization, *Int. J. Hydrog. Energy*, 30, 1523–1534.
- Moreno-Benito, M., Agnolucci, P., Papageorgiou, L. G., 2017, Towards a sustainable hydrogen economy: Optimisation-based framework for hydrogen infrastructure development, *Comput. Chem. Eng.*, 102, 110–127.
- National Grid ESO, 2021, Future Energy Scenarios, [www.nationalgrideso.com/future-energy/future-energy-scenarios/archive](http://www.nationalgrideso.com/future-energy/future-energy-scenarios/archive).
- Ochoa Bique, A., Maia, L. K. K., Grossmann, I. E., Zondervan, E., 2021, Design of hydrogen supply chains under demand uncertainty - A case study of passenger transport in Germany, *Phys. Sci. Rev.*, 6, 000010151520200052.
- Ogumerem, G. S., Kim, C., Kesisoglou, I., Diangelakis, N. A., Pistikopoulos, E. N., 2018, A multi-objective optimization for the design and operation of a hydrogen network for transportation fuel, *Chem. Eng. Res. Des.*, 131, 279–292.
- Samsatli, S., Samsatli, N. J., 2019, The role of renewable hydrogen and inter-seasonal storage in decarbonising heat – Comprehensive optimisation of future renewable energy value chains. *Appl. Energy*, 233–234, 854–893.
- Sunny, N., Mac Dowell, N., Shah, N., 2020, What is needed to deliver carbon-neutral heat using hydrogen and CCS?, *Energy Environ. Sci.*, 13, 4204–4224.
- W. McDowall, C. Britchfield, 2022, Decarbonising at home, [www.instituteforgovernment.org.uk/publications/decarbonising-heating](http://www.instituteforgovernment.org.uk/publications/decarbonising-heating).

# Optimal Concentrated Solar Plant (CSP) location accounting for social and environmental impact: A three-location study in Spain

José A. Luceño Sánchez<sup>a</sup>, Mariano Martín<sup>b</sup>, Sandro Macchietto<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, Imperial College London South Kensington Campus, London SW7 2AZ, UK*

<sup>b</sup>*Departamento de Ingeniería Química y Textil. Universidad de Salamanca. Pza. Caídos 1-5, 37008 Salamanca (Spain)*  
*j.luceno-sanchez@imperial.ac.uk*

## Abstract

A mathematical framework is presented for sustainable energy production via multiple concentrated solar power (CSP) plants, accounting for geographical, technical, economic, water-energy tradeoff, and social aspects. The model considers variables related to the possible location of facilities (e.g., direct normal irradiance, ambient pressure, humidity, temperature, land, and water availability), technical and economic aspects (e.g., power production, cooling technology, plant and equipment size and costs, or water consumption), and the social impact of a facility (in the form of development, unemployment, and population ratios). The problem is formulated as a multiperiod MILP in which several locations are considered. The location, design, and operation of each facility are chosen to optimize various techno-social-environmental-economic objectives. Results for a representative, three-location case study (Spain) demonstrate the approach flexibility and the importance of social and environmental aspects in the optimal solution.

**Keywords:** Concentrated Solar Plants; energy sustainability tradeoffs; optimal location; mathematical optimization; MILP.

## 1. Introduction

Renewable energy production is essential to the green transition while meeting rising energy consumption related to population growth (Scheffran et al. 2020). Solar thermal energy offers an interesting choice as solar radiation is available in any place. However, various locations present variable amount of radiation, even in the same country, and daily and seasonal variations directly affect power production (Lozano Santamaria et al. 2021). Furthermore, other important variables depend on location, such as water availability for cooling systems (Guerras and Martín 2020), which strongly affects the facility viability and sustainability (Hamiche et al. 2016), and the social impact associated with the construction of the plant (Heras and Martín 2020). Choosing the best locations for Concentrated Solar Power (CSP) plants has an important impact not only on production but also on the society and the environment, for example, on water depletion, as water is lost by evaporation in cooling towers. A detailed procedure has not been presented in the literature for the optimal CSP location selection considering simultaneously costs, technological options, social impacts, and water-energy tradeoffs. In this work, such a mathematical model is developed. Potential locations in various

regions are assumed, characterized by monthly average data, to study the combined effect of techno-economic and social aspects with a sustainable scope.

## 2. Problem formulation

This work aims to optimize the location of CSP facilities for renewable energy production, accounting for aspects such as social concerns and energy-water tradeoffs. A mathematical model, divided into 5 modules, considers a number  $n_{fac}$  of CSP facilities producing electricity in  $Loc$  locations over a horizon,  $TD$  divided into  $t$  periods:

**-Facility surrogate model:** the equipment in a CSP plant is grouped into sections (Figure 1). Shared variables such as mass and heat flowrates are calculated using a surrogate model based on energy and material balances as presented in previous work (Martín and Martín 2013). The equipment in each section is costed using correlations from Martín and Grossmann, 2022.

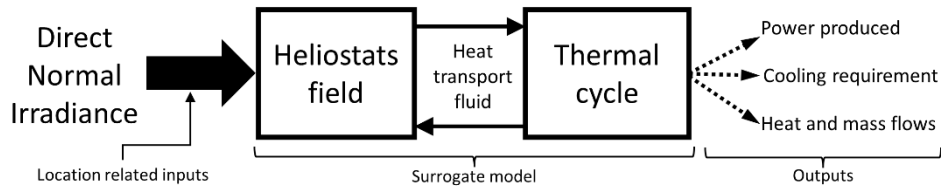


Figure 1. Scheme of CSP plant model.

**-Meeting electricity demand:** each facility produces a net electricity power,  $W_{net,t,l}$  [kW], and cumulatively plants must meet an overall demand  $W_{Dem,t}$  [kW]. Part of this demand may be supplied by other sources ( $W_{add,t}$  [kW]), at an additional cost (Eq. 1).

$$W_{Dem,t} \leq \sum_{l=1}^{Loc} (W_{net,t,l}) + W_{add,t} \quad \forall t \in TD \quad (1)$$

**-Location selection:** the selection of facility locations and cooling technology for each is formulated using binary variables (Eqs. 2-3), where  $y_{ff,l}$  defines the plant existence, and  $y_{DC,l}$  and  $y_{WC,l}$  define the choice of dry-cooling or wet-cooling, respectively.

$$\sum_{l=1}^{Loc} y_{ff,l} = n_{fac} \quad (2)$$

$$y_{DC,l} + y_{WC,l} = y_{ff,l} \quad \forall l \in Loc \quad (3)$$

The number of facilities  $n_{fac}$  should be lower than a pre-established maximum, while meeting constraints on the total investment and the existence of plants in specific locations (if any). Furthermore, the existence of a facility in a location  $l$  generating  $W_{net,t,l}$  defines the corresponding water consumption (for wet-cooling) or power consumption  $W_{consum,t,l}$  [kW] (for dry-cooling) and related costs. All these variables are determined applying a BigM formulation in each equation, as shown in Eq. 4 for power consumption as function of the total power produced  $W_{gen,t,l}$  [kW] =  $W_{net,t,l} + W_{consum,t,l}$ :

$$W_{consum,t,l} \geq W_{gen,t,l} \cdot 0.05 - (1 - y_{DC,l}) \cdot BigM \quad \forall l \in Loc, \forall t \in TD \quad (4)$$

**-Water index:** water consumption  $Water_{consum,t,l}$  [L/kWh], related to the use of wet-cooling technologies, is computed according to literature correlations considering location weather conditions (Guerras and Martín 2020). A relative location water impact  $Walp_l$  is defined considering a location water scarcity. A water index  $WI$  [€] is calculated as in Eq. 5, where  $R_{L \rightarrow \epsilon}$  [€/L] is the water cost and  $OP_t$  [h] is the monthly operation time.

$$WI = \sum_{t=1}^{month} \sum_{l=1}^{Loc} [Walp_l \cdot W_{gen,t,l} \cdot OP_t \cdot Water_{consum,t,l} \cdot R_{L \rightarrow \epsilon}] \quad (5)$$

**-Social indices:** the social impact of a facility in a region  $l$  is computed considering as inputs the Gross Domestic Product ( $GDP_l$ ), the facility cost ( $Cost_{fac,l}$ ) and design power, the Unemployment Ratio ( $UV_l$ ), the jobs created per MW and salary per job, and the

Relative Population ( $RP_l$ ), as found in the literature (Heras and Martín 2020). GDP, UV and RP are scaled to their maximum values, hence are dimensionless in the range [0-1]. The sum for all regions gives overall economic, unemployment and population indices ( $Social_{DR}$ ,  $Social_{UR}$  and  $Social_{UR}$ ), all in [€]. A combined social index (SI) [€] is defined as the sum of these indices, each with suitable weight parameters (P), as seen in Eq. 6:  

$$SI = P_{DR} \cdot (Social_{DR}) + P_{UR} \cdot (Social_{UR}) + P_{RP} \cdot (Social_{RP}) \quad (6)$$

### 3. Solution procedure and Case Study

The problem is formulated as a multi-period MILP optimization, where the size and cost of a facility are evaluated at the same time as other important indices (such as social and water indices). Three representative locations in Spain (Cadiz [south], Cuenca [middle], and Bizkaia [north]) are defined as possible CSP plant locations from which the optimal one can be chosen. A year of operation is considered, with a monthly time discretization for the conceptual evaluation of the locations. The specified electricity power demand profile (Table 1) is a small fraction of Spanish non-renewable electricity production. Data for the three locations are shown in Table 2. Their direct normal irradiance (DNI) and sun hours data (as monthly averages) were obtained from an EU database (PVGIS 2022).

**Table 1.** Power demand, as monthly averages [MW] over 1 year ( $W_{Dem,t}$  for  $t=[1, \dots, 12]$ ).

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
$W_{Dem,t}, MW$	1.18	0.95	0.86	0.86	0.96	1.21	1.44	1.36	1.30	1.26	1.28	1.18

**Table 2.** Location-related parameters. GDP, UV, RP and WaIp are dimensionless.

	Province [km <sup>2</sup> ]	Land Cost [€/m <sup>2</sup> ]	GDP	UV	RP	WaIp
<b>Cadiz</b>	7,436	0.20	0.52	2.29	0.79	0.80
<b>Cuenca</b>	17,140	0.06	0.41	0.83	0.99	0.93
<b>Bizkaia</b>	2,217	0.14	0.16	0.44	0.36	1.00

Two case studies are presented: 1) a stand-alone study for each location, aiming at characterizing the optimal solution if a CSP were built there to maximize the production, using Eq. 7 as objective function, and establishing a benchmark for the results obtained later; and 2) a combined study of possible plants in all locations to meet demand, with 4 different objective functions (Eqs. 7-10). Equations 8-10 represent the minimization of facilities cost, the maximization of social impact, and the combined minimization of all aspects, respectively. In Eq. 10 (scaled to M€), the term in the first square brackets is the cost related to the power surplus and its management, the one in the second brackets is the annualized cost of facilities considering 20 years as the years of operation, and the term in the third brackets is the cost associated to any additional energy purchased to meet the demand of Table 1. A value  $AddCost_{kWh \rightarrow \epsilon}$  (0.10 €/kWh) is attributed to any power surplus, and a cost penalty  $Penalty_{kWh \rightarrow \epsilon}$  (5 €/kWh) for any underproduction ( $W_{add,t}$ ).

$$\min\{-\sum_{t=1}^{month} [\sum_{l=1}^{Loc} (W_{net,t,l}) + W_{add,t}]\} \quad (7)$$

$$\min\{\sum_{l=1}^{Loc} Cost_{fac,l}\} \quad (8)$$

$$\min\{-SI\} \quad (9)$$

$$\min\{-SI + WI - [\sum_{t=1}^{month} (W_{Dem,t} - \sum_{l=1}^{Loc} W_{neat,t,l}) \cdot OP \cdot AddCost_{kWh \rightarrow \epsilon}] + [(\sum_{l=1}^{Loc} Cost_{fac,l})/20] + [(\sum_{t=1}^{month} W_{add,t} \cdot OP \cdot Penalty_{kWh \rightarrow \epsilon})]\} \quad (10)$$

The following assumptions are made: 1) no cost or losses for transport of electricity; 2) the maximum plant area per location is 3% of a region area; 3) unless specified, the max

total investment is 30M€; 3) the area of a single heliostat is 120 m<sup>2</sup>; 4) the dry-cooling cost is related to air coolers exchanger area; 5) the value of weight parameter  $P_{DR}$  is 1.

## 4. Results

### 4.1. Case study 1: CSP plant in individual locations

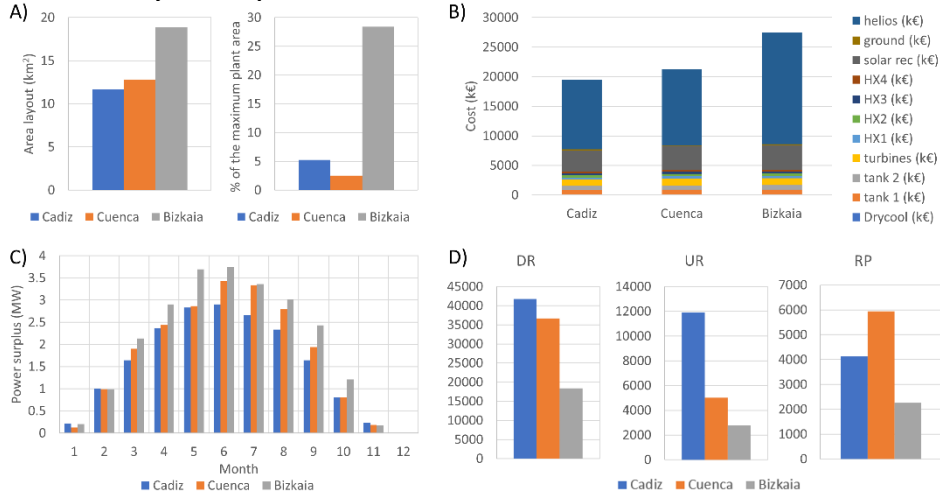


Figure 2.- Main results of Case Study 1: A) land occupation (in km<sup>2</sup> and as % of region land); B) facility cost; C) power surplus; D) Social impacts (relative to current).

Here, we wish to establish the characteristics and impact of individual CSP plants in the three locations, each built to meet the demand in Table 1. The main results of Case study 1 are collected in Figure 2, with dry-cooling systems selected for comparative purposes. Cadiz requires the lowest surface (~12 km<sup>2</sup>, Figure 2.A), due to its higher values of DNI. Cuenca, with a slightly higher plant area, presents a lower fraction of the occupied surface as it is a larger region. Bizkaia (in the north) needs the largest surface, which would occupy 28% of the surface of the region. The occupied surface is related to the facility cost, which increases with the number of heliostats (Figure 2.B). The results show that, in terms of plant surface and cost alone, the best locations are in the south of the country. Cuenca, the largest province, might be preferred considering possible surface limitations and reduction in costs related to ground and heliostats field.

Regarding power production, in all locations the CSP plant is sized to meet the electricity demand in December, the most demanding month as DNI and solar hours are the lowest. The resulting larger design flowrates (i.e., heat transport and storage, thermal cycle, and cooling systems sizes) lead to substantial excess production in other months (Figure 2.C). As Bizkaia presents lower DNI values than the other regions, and the number of sun hours is also lower, the facility requires larger equipment to overcome the demand in December, giving it the largest power surplus during the rest of the year (28% higher than Cadiz) (Figure 2.C). The CSP facilities of Cuenca and especially Cadiz show reduced overproduction during the majority of the year. Figure 2.D plots the three social indices for plants in each location. Cadiz may be viewed as the best option considering the large impacts on development and employment (i.e., large differences relative to the current state in the DR and UR indices). Cuenca is the most favorable choice in terms of repopulation concerns, due to its lowest RP value. Additionally, Bizkaia seems to be the worst location for social impact as it presents the smallest improvement in all the indices.

4.2. Case 2: CSP plants in multiple locations and social-environmental concerns

Here, CSP plants may be built in one or more of the three locations, with combined electricity production to meet the demand in Table 1. Results are collected in Table 3, in which: Scenario 1 is the maximum power production (Eq. 7); Scenario 2 is the minimum cost (Eq. 8); Scenario 3 is the minimum total facility cost if the demand is 10 times the presented in Table 1, an unlimited budget is available, and a minimum investment of 20 M€ must be made in Bizkaia; Scenario 4 is the maximum social impact solution (Eq. 9) with  $P_{UR} = P_{RP} = 0.5$ ; Scenario 5 is again the social impact optimization but considering different social priorities ( $P_{UR} = 0.5$  and  $P_{RP} = 50$ ); Scenario 6 is the combined water-energy-social problem (Eq. 10) with the social priorities of Scenario 4; and Scenario 7 is the same problem with the social priorities of Scenario 5.

Table 3. Main results - Case Study 2.

Variable	Location	Scenario						
		1	2	3	4	5	6	7
Number of heliostats	Cadiz	142,761	92,495	877,085	146,075	0	97,363	0
	Cuenca	0	0	0	0	129,106	0	129,106
	Bizkaia	0	0	81,370	0	0	0	0
Power Design [MW]	Cadiz	6.35	4.11	39.00	6.50	0	4.33	0
	Cuenca	0	0	0	0	6.07	0	6.07
	Bizkaia	0	0	2.70	0	0	0	0
Facility cost [M\$]	Cadiz	30.00	20.34	168.86	30.00	0	20.71	0
	Cuenca	0	0	0	0	30.00	0	30.00
	Bizkaia	0	0	20.00	0	0	0	0
SI	Cadiz	19.67	13.30	111.36	19.70	0	13.57	0
	Cuenca	0	0	0	0	74.89	0	74.89
	Bizkaia	0	0	3.98	0	0	0	0
WI	Cadiz	0.1515	0.0982	0.9309	0	0	0	0
	Cuenca	0	0	0	0	0	0	0
	Bizkaia	0	0	0	0	0	0	0

The first two scenarios show the maximum and the minimum design capacities that could meet the demand, respectively; in both scenarios, only one facility is chosen in Cadiz with wet-cooling ( $WI > 0$ ), because it does not have associated power consumption. These results are in line with the analysis in the previous section. In scenario 3 (demand 10 times that in Table 1, infinite budget, and a minimum of 20M€ to be invested in Bizkaia), results show two facilities (Cadiz+Bizkaia) as the optimal (but very expensive) solution, with Cadiz facility larger than Bizkaia one, due to value of DNI. In Scenario 4, Cadiz is again selected, probably related to the large impact of unemployment (UR value). In Scenario 4, increased production always increases the economic development ratio, DR, up to when some other constraint is encountered (here, on max capital cost); this results in a larger facility than in Scenario 1, but with dry-cooling; dry-cooling is chosen instead of wet-cooling because of the higher production required to meet the power consumption, which also affects UR and RP values. With higher social priority assigned to repopulation than unemployment in Scenario 5, a single plant is located in Cuenca; this shows that

social policies have an important effect on the optimal decision. In Scenario 6, Eq. 10 is employed with the social priority of Scenario 4 to evaluate the combined effect of all variables. Cadiz is chosen as the optimal location, presenting a solution similar to Scenario 2. This result shows the influence of environmental and social impacts: dry-cooling is chosen to reduce WI and increase design power, which also increases social impact (as seen in Scenarios 4-5) compared to previous ones, balancing with cost reduction and production surplus. Scenario 7, with the same social priority as Scenario 5, presents the same optimal solution as Scenario 6.

## 5. Conclusions

The proposed general model for the optimal location of multiple CSP facilities, to reduce fossil-based energy dependency, incorporates geographical, technical, economic, and social components. The plant model considers technical design variables for sizing key sections and calculating water consumption. The social impact model reflects key social variables in a location. Technical, social, and environmental variables may be used individually or in combination as objective functions or within constraints. The model was addressed by a multiperiod MILP approach, and its generality and flexibility were demonstrated for a small but representative 3-location case study in Spain.

Results suggest that, based on meeting electricity demand or investment cost, the most suitable CSP facility location is Cadiz, due to the high DNI values in the south of Spain. If social indices are also taken into account, in particular, high priority is given to depopulation, the optimal location changes from Cadiz to Cuenca, with dry-cooling used instead of wet-cooling. These results show the promise of the model to be used to improve decisions about future facility distribution across countries, including social aspects.

## Acknowledgments

The authors acknowledge Imperial College London and University of Salamanca for access to scientific sources and databases. The authors acknowledge the Margarita Salas grant, co-funded by the European NextGenerationEU Fund, Spanish “Plan de Recuperación, Transformación y Resiliencia” Fund, Spanish Ministry of Universities, and University of Salamanca (“Ayudas para la recualificación del Sistema Universitario español 2021-2022”), to JA Luceño Sánchez.

## References

- Guerras, L. S.; Martín, M. (2020). On the Water Footprint in Power Production: Sustainable Design of Wet Cooling Towers. *Appl. Energy*, 263, 114620.
- Hamiche, A.M.; Stambouli, A.B.; Flazi, S. (2016). A review of the water-energy nexus. *Renew. Sust. Energy Rev.* 65, 319-331.
- Heras, J.; Martín, M. (2020). Social Issues in the Energy Transition: Effect on the Design of the New Power System. *Appl. Energy*, 278, 115654.
- Lozano Santamaria, F.; Luceño, J.A.; Martín, M.; Macchietto, S. (2021). Optimal Operation and Cleaning Scheduling of Air Coolers in Concentrated Solar Power Plants. *Computers and Chemical Engineering* 150, 107312.
- Martín, L.; Martín, M. (2013). Optimal year-round operation of a concentrated solar energy plant in the south of Europe. *Appl. Therm. Eng.* 59, 627-633.
- Martín, M.; Grossmann, I.E. (2022). Mathematical modeling for renewable process design, in: *Sustainable Design for Renewable Processes*. Elsevier, pp. 35–100.
- PVGIS, Photovoltaic Geographical Information System. Interactive tools. Last accessed Oct 2022. [https://re.jrc.ec.europa.eu/pvg\\_tools/en/](https://re.jrc.ec.europa.eu/pvg_tools/en/)
- Scheffran, J.; Felkers, M.; Froese, R. (2020) Economic Growth and the Global Energy Demand. In *Green Energy to Sustainability*; Wiley, pp 1–44.

# Inventory data generation for prospective lifecycle design through process modeling of energy recovery from waste plastics

Shoma Fujii<sup>a</sup>, Yuichiro Kanematsu<sup>b</sup>, Yasunori Kikuchi<sup>abc</sup>

<sup>a</sup>*Institute for Future Initiatives, The University of Tokyo, 113-8654 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan*

<sup>b</sup>*Presidential endowed chair for “Platinum society”, The University of Tokyo, 113-8656 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan*

<sup>c</sup>*Department of chemical system engineering, The University of Tokyo, 113-8656 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan*  
*shoma.fujii@ifi.u-tokyo.ac.jp*

## Abstract

A novel energy recovery processes for waste plastics has been proposed, which consists of surface treatment to boilers at waste incineration facilities to increase operating rates and waste heat recovery to generate cold heat. In order to propose the best combination of technologies and to conduct a prospective life cycle assessment that feeds back into technology development, the technologies under development were reflected in a process flow model. Preventing fouling of the boiler tubes leads to less frequent maintenance and longer operating hours. The increase in operating hours results in a slight increase in total annual power generation, but if the annual waste input remains constant, the daily waste throughput will decrease, resulting in lower power output and lower boiler exhaust gas temperatures. Waste heat recovery to generate cold heat in absorption chillers can be done either onsite or offsite, and hot water, steam, and hot air were considered as heat recovery media. The results show that the off-site heat supply is reduced by half to one-fourth compared to the on-site heat supply.

**Keywords:** circular economy, thermal energy storage, cold heat, power generation

## 1. Introduction

Plastics recycling is attracting attention as a way to realize the circular economy (Schwarz *et al.*, 2021). Plastics circulation should be implemented in a cascade of advanced sorting, material recycling, chemical recycling, and finally energy recovery. A computer-aided process engineering that reflects the emerging technical information supports the system design of plastics circulation by optimal combination of these processes from the lifecycle viewpoints (Volk *et al.*, 2022). Each recycling process has various problems, and measures are being considered to solve them. As for the energy recovery process, the operating rate has decreased due to fouling of the boiler tubes and the heat utilization rate has decreased due to the lack of nearby heat demand. A new high-efficient energy recovery process (NEDO, 2022) that combines improvement of operation rate through surface treatment on boiler tubes (Naganuma *et al.*, 2022) and cold heat generation (Isojima *et al.*, 2020) for logistics demands through waste heat recovery was proposed. The objective of this study is to develop a process flow model of above novel



energy recovery process for waste plastics circulation, and to generate an inventory for life cycle assessment for designing a best mix of plastics circulation strategy.

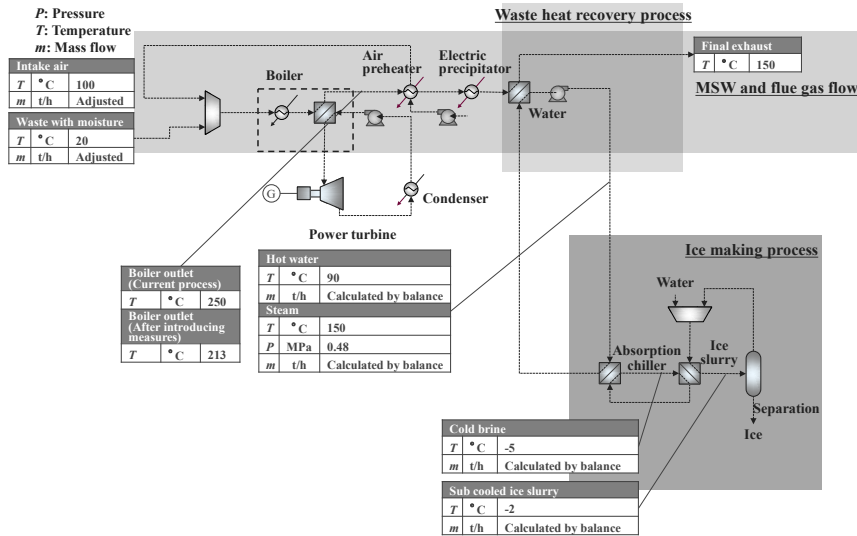
## 2. Material and methodology

### 2.1. Process modeling

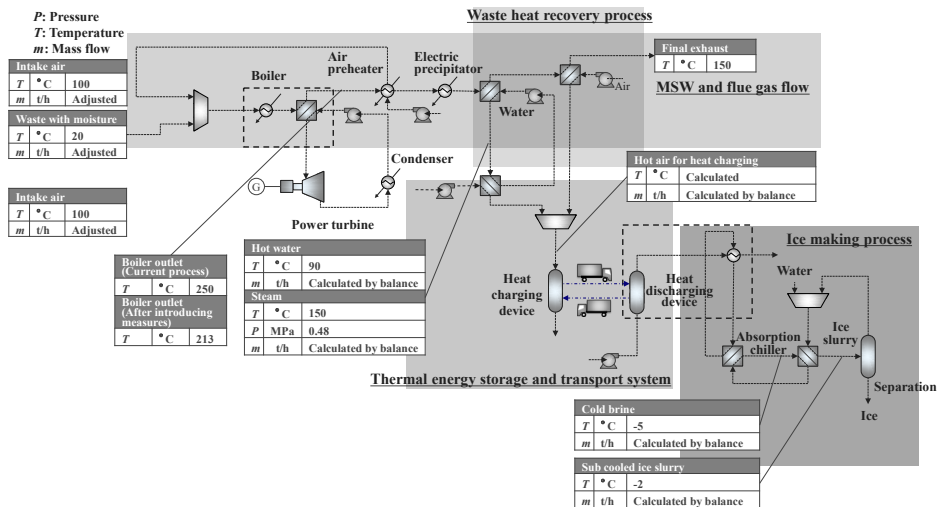
Figure 1 shows the combined process of waste incineration and energy recovery. Water-containing municipal solid waste (MSW) is incinerated in a boiler, the water supplied to the boiler is boiled, and the steam generated is introduced to power turbine to generate electricity. For more efficient energy recovery from waste incineration plant, two measures are being considered: one on the high-temperature side to improve power generation and operating rate, and the other on the low-temperature side to recover waste heat. For high-temperature side, when wastes containing plastics are incinerated, gases containing chlorine and sulfur, as well as ash with low melting points, adhere to the boiler tubes leading decrease in heat transfer coefficient. Thermal spray treatment of boiler tube surfaces can reduce ash adhesion and increase plant operation rate (Naganuma et al., 2022). For low-temperature side, two types of heat recovery methods for boiler exhaust gas are considered in this study: on-site and off-site cold heat recovery (Kimura et al., 2022). In the on-site cold heat generation process (Fig 1 (a)), the heat capacity of the exhaust gas is recovered as water or steam in a heat exchanger and used directly as a heat source in an absorption chiller to generate sub-zero brine, then sub-zero ice slurry is generated under sub cooling condition in an ice generator by heat exchanging with cold brine, and cold ice is produced through a releasing subcooling and a separation process. In the off-site cold heat generation process (Fig 1 (b)), a thermal energy storage and transport system is required. Adsorbent is employed as a heat storage medium, and heat is charged by drying it with air heated by waste heat, transported by truck to a remote cold demand, and supplied to the heat source of an absorption chiller by adsorbing ambient moist air. There are two methods to recover waste heat. Assuming that the heat capacity of the exhaust gas of boiler is recovered by the outside air through a heat exchanger, the hot air generated is directly used to dry the adsorbent (i.e., heat charging). The other method of waste heat recovery is to generate hot water or steam by heat exchange with exhaust gas of boiler as in the on-site heat recovery process. After hot water or steam generation, hot air is generated through a heat exchange process. This process requires two heat exchangers (one to recover waste heat to generate hot water or steam and the other to generate hot air), but the generated hot water or steam can also be used directly as a heat source for an absorption chiller onsite, thus achieving a combination of on-site and off-site heat recovery.

The system is assumed to be retrofitted to a typical waste incineration plant with an energy recovery process. First, to model MSW incineration with only a typical power generation process, without heat recovery process and power generation enhancement measures, the heat exchange process in the boiler is modeled by the specific heat and heating value of the combustion gas and the heat capacity of the feed water. The moisture content of the input waste was assumed to be 45%. The unknown parameters of mass flow of waste input and intake air were adjusted until the target values of 250 °C boiler outlet temperature and 10% oxygen concentration in the final flue gas were reached, and the UA (overall heat transfer coefficient) value of the boiler was calculated. The energy balance of the heat exchanger was divided into the phase change section and other sections, and the phase change section was divided into 10 sections and the UA of each section was calculated using the logarithmic mean temperature difference method.

Next, a case was assumed in which fouling is prevented by surface treatment to the boiler tubes, thereby extending the operating period. Retrofit of the existing boiler was assumed by fixing the calculated UA values. Since the annual waste input to the boiler remains the same before and after the introduction of the energy recovery process, the hourly waste input would decrease if the boiler tubes were surface treated for an extended operating period of one year. The mass flow rate of steam for power generation was recalculated to maintain the UA value under the updated waste input conditions, and the amount of power generation and cold heat recovery were evaluated.



(a) On-site heat recovery process



(b) Off-site heat recovery process

Figure 1 Process flow of waste incineration plant with heat recovery process

## 2.2. Case study

As a case study, a case at a steam flow rate of 25 t/h in a boiler was evaluated. It is assumed that the boiler will be operated for 9 months per year, with an extended operating period of 1 month due to surface treatment to prevent fouling of the boiler tubes. The heat transfer coefficient was assumed to be constant during the operation period, but it is possible to simulate the effect of decrease in the heat transfer coefficient by changing the UA.

In the waste heat recovery process, it is assumed that the heat is recovered as saturated steam at 0.48 MPa (saturated temperature of 150 °C) for recovery with steam case, and as hot water at atmospheric pressure of 90 °C for recovery with hot water case. The minimum temperature approach of the heat exchanger for the production of hot air for thermal energy storage system was assumed to be 5 °C. The absorption chiller is a two-stage refrigerator capable of producing sub-zero brine, and it is assumed to produce brine at -5 °C based on the literature, with a COP (Coefficient of performance) of 0.2 considering waste heat input (Kimura et al., 2022). The auxiliary power of the absorption chiller was assumed to be 10% of the cooling capacity. In the ice slurry production, the ice fraction in the ice slurry was assumed to be 30%, and the auxiliary power of the absorption chiller was assumed to be 20% of the cooling capacity. For thermal energy storage system, the equilibrium adsorption water uptake before (4%) and after adsorption (23%) was calculated using adsorption isotherms (Miyahara et al., 2020) and the heat storage density was calculated from the difference between the equilibrium adsorption water uptake and the heat of adsorption (Miyahara et al., 2020). The gas temperature at the outlet of the heat charging device was set to 60 °C, and 80% of the difference from the inlet temperature was assumed to be stored. Heat discharging amount was assumed to be 80% of the heat storage density (Fujii et al., 2022). The pressure drop in both the heat charging and discharging devices was assumed to be 1 kPa, and the efficiency of the blower that injects hot air into the devices was assumed to be 70%. These parameters are subject to sensitivity analysis in the feedback to design based on life cycle thinking in future work.

As evaluation indicators, the increase in power generation, the amount of cold heat generation, the amount of ice slurry produced, and the additional energy input (power for ice slurry production, heat storage, and transportation) are evaluated. These indicators can be incorporated into the next stage of this study, a prospective life cycle assessment for design.

## 3. Result and discussion

The results of power and cold heat generation, auxiliary energy consumption of all processes were shown in Fig. 2. As a result, the boiler air excess ratio was about 1.7. The daily waste input resulted in 250 t/d. The power output was about 2,100 kW, resulting in an annual output of 12,200 MWh. The daily waste input was then reduced to 222 t/d when the operating time was extended by one month by preventing fouling of the boiler tubes. Adjusting the steam flow rate to fix the UA value of the boiler reduced the power output to 1,950 kW, but the extended operation period resulted in a slightly higher annual power output of 12,840 MWh. In addition, the temperature of exhaust gas at the boiler outlet was reduced from 250 °C to 214 °C. Despite the decrease in the daily waste input, the UA value of the boiler was kept at the previous value for retrofit, which resulted in a large heat transfer area relative to the heat exchange rate and lower temperature of boiler outlet. In the heat recovery part, the difference between steam and hot water as a heat source to the absorption chiller was not taken into account in the on-site case, resulting in a cold

heat production of approximately 213 kW in both cases. On the other hand, in the off-site case, the heat source that can be input to the absorption chiller at the demand site is smaller than in the on-site case, because the heat capacity of the exhaust gas of the carrier gas cannot be recovered at the heat charging and discharging devices. When the exhaust gas from the waste incineration was directly exchanged for hot air to the heat charging device, the ice-making capacity was 124 kW. If the heat is recovered with steam or hot water, an additional heat exchange process with air for heat storage is required, which further reduces the ice-making capacity compared to the case where the heat is recovered directly from the exhaust gas with hot air. The ice-making capacity is 110 kW when heat is recovered with steam, and in the case of heat recovery with hot water, the amount of heat that can be transported is further reduced because the temperature of hot air obtained through heat exchange is about 60 °C lower than that of steam case, resulting in an ice-making capacity of 64 kW.

These results can be reflected in a prospective life cycle assessment as inventory data to provide feedback for each technology development. Since the UA values of the boiler and heat exchanger are also quantified, the heat transfer area can be calculated by giving the heat transfer coefficient, and the manufacturing and disposal environmental loads of each device can be estimated.

In this case, the measures taken to prevent fouling of the boiler tubes did not result in a significant increase in power generation. However, the heat transfer coefficient decreases with time from the start of operation in the case of the non-surface treatment, so the power output of 2,100 kW in the case of the non-surface treatment gradually decreases until maintenance is performed. The annual power output may be significantly increased by the surface treatment to boiler tubes. Taking into account this change in heat transfer coefficient based on experimental data is the next step.

In order to preemptively evaluate technologies that are still under development, uncertain technology information is reflected in the process modeling, and it is important to be able to make comparisons in each process and to know the approximate contribution of the new process rather than absolute values.

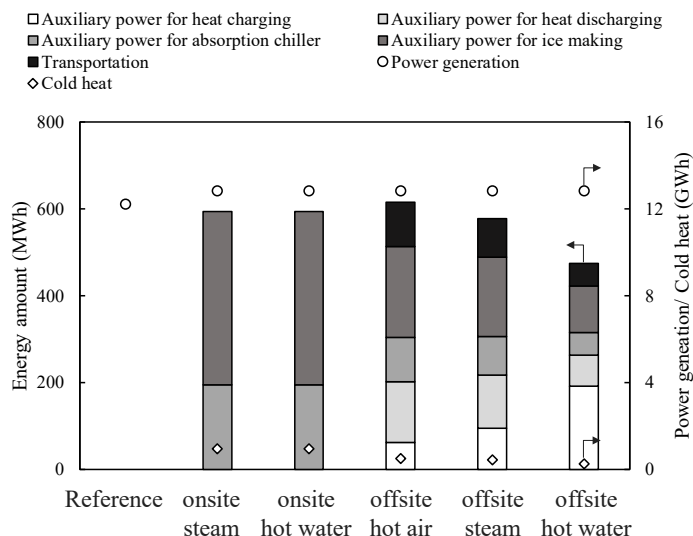


Figure 2 Results of power and cold heat generation, auxiliary energy consumption of each process (Process without any measures, onsite, offsite and hot air)

#### 4. Conclusion

This study modeled process flow of waste incineration plant including waste plastics to generate inventory data of novel energy recovery process combined with surface treatment to boiler tubes and waste heat recovery to generate cold heat via absorption chiller, as one of the measures of waste plastics circulation strategy. As a result, total power generation could be slightly increased by extending operation period due to prevention of boiler tubes fouling. Waste heat could produce cold heat for both onsite and offsite demand, but in the offsite case, the amount of cold heat generation is reduced by one-fourth compared to the on-site heat supply.

#### Acknowledgement

This work was supported by e New Energy and Industrial Technology Development Organization (NEDO, Grant number JPNP20012), JSPS KAKENHI Grant-in-Aid for Early-Career Scientists (Grant Number 22K18061), JST COI-NEXT (Grant Number JPMJPF2003), JST PRESTO (Grant Number JPMJPR2278) and the Environment Research and Technology Development Fund (Grant Number JPMEERF20213R01) of the Environmental Restoration and Conservation Agency of Japan. The activities of the Presidential Endowed Chair for “Platinum Society” at the University of Tokyo are supported by the KAITEKI Institute Incorporated, Mitsui Fudosan Corporation, Shin-Etsu Chemical Co., ORIX Corporation, Sekisui House, Ltd., East Japan Railway Company, and Toyota Tsusho Corporation.

#### References

- S. Fujii, T. Nakagaki, Y. Kanematsu, Y. Kikuchi, 2022, Prospective life cycle assessment for designing mobile thermal energy storage system utilizing zeolite, *Journal of Cleaner Production*, 365, 132592
- S. Isojima, H. Noda, M. Taninno, 2020, Crystallization Temperature of the Working Fluid LiBr - H<sub>2</sub>O - 1-Propanol for Absorption Refrigerator to Generate Below-Zero Brine (in Japanese), *Proceedings of the JSRAE Annual Conference, 2020*, ROMBUNNO.E213
- T. Kimura, H. Hatano, H. Noda, T. Inada, M. Tanino, H. Naganuma, M. Hotta, I. Naruse, 2022, High-efficient energy recovery of waste plastics Part 2- Effective utilization of low temperature waste heat (in Japanese), *Kagaku-sochi*, 64, 1, 48-54
- H. Miyahara, M. Suzuki, S. Matsuda, K. Morimoto, K. Mampuku, Y. Kawakami, H. Nawa, K. Yamauchi, K. Matsunaga, M. Tanino, 2020, Development of Open-type Adsorption Thermal Storage Heat Pump System Applying HAS-Clay Part 2 Hydration Heat Caused by Water Vapor Adsorption on Low-Temperature Regenerative Heat Storage Material (in Japanese), *Transactions of the Society of Heating, Air-conditioning and Sanitary Engineers of Japan*, 285, 1-8
- H. Naganuma, I. Naruse, M. Hotta, H. Hatano, H. Noda, T. Inada, M. Tanino, 2022, High-efficient energy recovery of waste plastics Part 1- Effective utilization of high temperature waste heat (in Japanese), *Kagaku-sochi*, 64, 1, 43-47
- NEDO New Energy and Industrial Technology Development Organization, 2022, Innovative Plastic Resource Circulation Process Technology Development, [https://www.nedo.go.jp/english/activities/activities\\_ZZJP\\_100179.html](https://www.nedo.go.jp/english/activities/activities_ZZJP_100179.html) (Access Sep-27, 2022)
- A.E. Schwarz, T.N. Lighthart, D. Godoi Bizarro, P. De Wild, B. Vreugdenhil, T. van Harmelen, 2021, Plastic recycling in a circular economy; determining environmental
- R.Volk,C. Stallkamp J. J. Steins, S. P. Yogish, R. C. Müller, D. Stapf, F. Schultmann, 2021, Techno-economic assessment and comparison of different plastic recycling pathways, *Waste Management*, 25, 5, 1318-1337

## Process intensification in a fixed bed reactor for a small-scale process in the stranded assets

Adrian R. Irhamna<sup>a</sup> and George M. Bollas<sup>a</sup>

<sup>a</sup> *Department of Chemical and Biomolecular Engineering, UTC Institute for Advanced Systems Engineering, University of Connecticut, 159 Discovery Dr, Storrs, CT, 06269, USA.*

### Abstract

The implementation of small-scale processes on stranded fossil fuel assets could alleviate the loss of valuable energy resources and the environmental damage caused by leaks of greenhouse gases. However, the current state of conventional chemical processes is not economical on a small scale. This study explores intensified fixed-bed reactor designs for small-scale hydrogen production and methane emissions treatment in petroleum-stranded assets. The reactor uses the chemical-looping principle, as it is packed with a metal oxygen carrier capable of performing redox and catalytic reactions. The reactor was designed following a dynamic optimization and a reactor intensification strategy. Performance metrics of each optimally designed reactor were then evaluated at operating conditions that were optimized in terms of manipulated variables and scheduling. The reforming reactor produces hydrogen from stranded natural gas with H<sub>2</sub>/CO ratio of 2.15 at the reforming stage, and >98% nitrogen stream at the oxidation stage. On the other hand, the flare emissions treatment reactor converts all the methane emissions in the flare exhaust stream without supply of external heat.

**Keywords:** Small-scale process, fixed-bed reactor, process intensification, dynamic modeling and optimization.

### 1. Background

Stranded assets are small fossil fuel reservoirs that cannot be processed due to their remote location, distribution challenges, small capacity, or the associated regulatory uncertainty. Specifically, stranded natural gas is, typically, flared or vented during oil recovery. Moreover, as the world transitions away from greenhouse gas-emitting energy production, more stranded assets will be generated and could yield losses of up to 1.4 trillion USD (Semieniuk et al., 2022). Abandoned assets can, also, cause environmental damage due to leakages of greenhouse gases, such as methane, generated by an improper shutdown of production and distribution systems. Approximately an average methane emission factor of 6.0 g/h was reported for more than 4 million abandoned oil and gas wells in the U.S. and Canada (Williams et al., 2021). Therefore, stakeholders consider small-scale process alternatives for stranded fossil fuel assets to minimize future financial implications due to greenhouse gas emissions and misuse of resources. The challenge, however, is that the current state of the art in small-scale process systems engineering is not economically viable and lacks efficiency.

This study explores the use of a fixed-bed reactor for small-scale applications in stranded assets utilizing the chemical looping principle and dynamic optimization as a process intensification strategy (Baharudin et al., 2021; Tan et al., 2021). A fixed-bed reactor is selected due to its simplicity, flexibility, and potential modularity. As the oxygen carrier does not circulate, the redox process is determined by switching between alternating fuel and air feed streams. Using a fixed-bed reactor also eliminates the potential for attrition and breakage of oxygen carriers, a common problem in fluidization

reactors. As a result, the life cycle of the oxygen carrier is extended, and maintenance costs are reduced. The heart of the proposed process concept lies in the ability of the oxygen carrier to promote gas-solid reactions, assist the catalytic reactions, store the heat generated during the oxidation stage, and release it during the reduction/reforming stages. In this work, two reactor designs are proposed: one to handle methane emissions from oil & gas flaring systems; and one to produce hydrogen and nitrogen for liquid ammonia production from stranded natural gas. We study this concept with high-fidelity dynamic models of the reactors, for which kinetic and transport parameters were validated previously.

## 2. Methodology

### 2.1. Reactor model and process description

A 1D heterogeneous fixed bed reactor model was developed to account for the mass, energy, and momentum balances of the solid and fluid phases. The details of the equations used for reactor modeling are documented in (Han et al., 2013; Han & Bollas, 2016). Kinetic correlations, extensively studied by (Han et al., 2016a, 2016b), were used to predict reactions occurring over Ni-based oxygen carriers. The chemical reactions and kinetic expressions are documented in (Irhamna & Bollas, 2023). The model explores the potential small-scale utilization of stranded gas assets specifically for improving flaring activities and in-situ reforming of stranded natural gas. Fig. 1 shows the process flow diagrams for two applications.

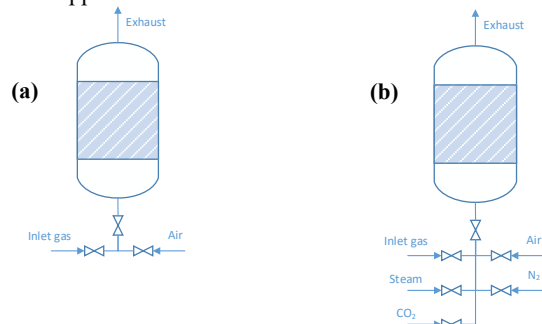


Figure 1: Process diagram of reactor used in this study: (a) flaring; and (b) reforming applications

In flaring applications, the reactor is used as a cleaning unit at the exit of the flare, ensuring no methane is released into the atmosphere. The reactor takes the exhaust gas from the flaring unit as feed and converts the unreacted methane to carbon dioxide to exploit the significant difference in the greenhouse gas potential of these two gases. Based on the process experienced by the oxygen carrier, there are two main stages: reduction and oxidation. The reduction stage occurs as the oxygen carrier reacts with methane and other reducing gases via gas-solid reactions. The reactions occur using the lattice oxygen of the oxygen carrier and heat stored in the bed from the previous stage of exothermic oxidation. As the reduction continues, oxygen in the bed eventually becomes scarce, and the temperature drops, leading to lower methane conversion. A brief period of air flow is introduced to regenerate oxygen and increase the bed temperature, which is coined as the oxidation stage in the reactor. The airflow stops when the bed has sufficient oxygen and heat for methane oxidation. The reactor then switches back to the reduction stage.

In the reforming application, the reactor is designed to convert natural gas from the stranded wellhead to syngas. The reactor takes natural gas from the wellhead as the main feed and operates in three stages: reduction, reforming, and oxidation. Initially, the oxygen carrier is reduced by the feed gas during the reduction stage to provide the catalyst

for reforming natural gas. Then reforming commences as the reactor feed switches to steam and CO<sub>2</sub>. Overall, the proposed scheme accomplishes tri-reforming of methane in one reactor and three stages. Both the reduction and reforming stages leverage the heat stored in the bed from the preceding oxidation cycle. As reforming progresses, the bed temperature drops, resulting in lower fuel conversion. When this happens, the feed gas is replaced by air, commencing the oxidation stage. During the oxidation stage, the bed temperature increases due to heat generated by oxygen carrier oxidation. The stage switches back to reduction when the sufficient heat, required for the following cycle, is stored in the bed.

We used typical flared gas flowrate reported in (Sherrick et al., 2011) and emission factors reported in (E & P Forum, 1994) to estimate flared gas composition for flaring application. Natural gas composition reported in (Kidnay et al., 2020) was used as feedstock for reforming application. Table 1 presents the reactor dimensions and operating conditions used for each flaring and reforming application, which were the result of dynamic optimization discussed in the following section. Dynamic simulation and optimization of each reactor design was conducted in the equation-oriented modeling language gPROMS (Process Systems Enterprise, 2021).

Table 1: Operating conditions of the fixed bed reactor for flaring and reforming applications

Reactor Parameters	Flaring		Reforming		
Reactor diameter (m)	10		0.6		
Reactor length (m)	1.5		5		
Operating pressure (bar)	1		1		
Inlet gas composition (% mole)	RED	OX	RED	REFORM	OX
- CH <sub>4</sub>	0.0030	0.0024	0.071	0.1609	-
- H <sub>2</sub> O	0.1810	0.1448	-	0.5802	-
- CO	0.0005	0.0004	-	-	-
- CO <sub>2</sub>	0.0910	0.0728	0.429	0.2498	-
- N <sub>2</sub>	0.7237	0.7390	0.4995	-	0.7805
- O <sub>2</sub>	-	0.0401	-	-	0.209
- Ar	-	-	-	-	0.009
Temperature of inlet gas (°C)	650		600		
Temperature of inlet air (°C)	307		600		
Volumetric flowrate (m <sup>3</sup> /s)	40	54	0.375	0.525	0.525
Ni percentage in oxygen carriers (%)	16		8.0		
Duration of each stages (s)	200	60	210	190	190

## 2.2. Dynamic Reactor Optimization

Reactor intensification is accomplished by simultaneously optimizing reactor admissible inputs and scheduling. The objective function, design variables, and optimization formulation for the flaring and reforming applications are shown in Table 2. The dynamic optimization problem was solved in gPROMS using control vector parametrization and single shooting (Process Systems Enterprise, 2021).

In Table 2,  $\mathbf{f}$  is the set of differential-algebraic equations (DAEs) representing the reactor model;  $\mathbf{x}$  is the vector of state variables (i.e., mass, temperature, and pressure);  $\mathbf{u}$  is the piecewise constant functions of the admissible inputs;  $\boldsymbol{\theta}$  is the system parameters



including kinetic constants, describing the reactivity of oxygen carrier;  $F$  is the molar flowrate of each species;  $T$  is the bed temperature in  $z$  direction;  $P$  is the total pressure of the reactor in  $z$  direction.

Table 2: Dynamic optimization problem formulation for the two reactor applications studied.

Flaring	Reforming
Objective function, $\gamma$ $\gamma = \int_{t_0}^{t_f} \frac{F_{CO_2,out}}{F_{CH_4,in}} dt$	Objective function, $\gamma$ $\gamma = \int_{t_f}^{t_f} \frac{F_{H_2,out}}{F_{CH_4,in}} dt$
Design Variables, $\phi$ <ul style="list-style-type: none"> <li>• Feed air temperature</li> <li>• Air-to-Flared gas Ratio</li> <li>• Reduction &amp; oxidation time interval</li> <li>• Oxygen carrier loading</li> </ul>	Design Variables, $\phi$ <ul style="list-style-type: none"> <li>• Temperature of the feed gas</li> <li>• Reduction, reforming, &amp; oxidation interval</li> <li>• Oxygen carrier loading</li> </ul>
Optimization Problem $\max_{\phi \in \Phi} \gamma$ s.t. $f(\dot{\mathbf{x}}(t), \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}, t) = 0,$ $\mathbf{f}_0(\dot{\mathbf{x}}(t_0), \mathbf{x}(t_0), \mathbf{u}(t_0), \boldsymbol{\theta}, t_0) = 0,$ $T(t_i, z) - 1100 \text{ }^\circ\text{C} \leq 0, \forall t \in [t_0, t_f],$ $700^\circ\text{C} - \frac{1}{L} \int_{z=0}^{z=L} T(t_i, z) dz \leq 0, \forall t \in [t_0, t_f],$ $100 \text{ }^\circ\text{C} - T_{air} \leq 0,$ $T_{air} - 427 \text{ }^\circ\text{C} \leq 0,$ $\Delta P(t)/P(t, z = 0) - 15\% \leq 0, \forall t \in [t_0, t_f],$	Optimization Problem $\max_{\phi \in \Phi} \gamma$ s.t. $f(\dot{\mathbf{x}}(t), \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}, t) = 0,$ $\mathbf{f}_0(\dot{\mathbf{x}}(t_0), \mathbf{x}(t_0), \mathbf{u}(t_0), \boldsymbol{\theta}, t_0) = 0,$ $95\% - X_{CH_4} \leq 0, \forall t \in [t_0, t_f],$ $T(t_i, z) - 1100 \text{ }^\circ\text{C} \leq 0, \forall t \in [t_0, t_f],$ $700^\circ\text{C} - \frac{1}{L} \int_{z=0}^{z=L} T(t_i, z) dz \leq 0, \forall t \in [t_0, t_f],$

### 3. Result and Analysis

Fig. 2 shows the optimal reactor performance for the flaring applications at a cyclic steady state. Near-complete methane conversion ( $> 98\%$ ) is achieved at the reduction and oxidation stages. Fig. 2(b) confirms that methane is practically undetectable at the reactor exit. The flue gas has a constant exit temperature of  $690 \text{ }^\circ\text{C}$  despite the cyclic steady state. Fig. 2(c) shows that the temperature gradients exist only in the first 20% of the reactor length. The heat front is pushed towards the end of the reactor, creating a relatively uniform temperature distribution near the reactor exit. Fig. 2(d) shows that the NiO form of the oxygen carrier dominates, allowing for sufficient oxygen to oxidize the methane feed. The Ni-NiO reaction occurs at the first 20% of the reactor length, aligned with the temperature gradients.

Fig. 3 shows the optimal performance of the reactor for the reforming application at a cyclic steady state. Fig. 3(a) shows that high methane conversion ( $> 95\%$ ) is achieved at the reduction and reforming stages. At reduction, methane conversion occurs in the first 20-40% of the reactor length, while at reforming, the conversion occurs further down the reactor, between 40-80% of the reactor length. Fig. 3(b) shows that the reactor produces different gases in each stage. Hydrogen produced during the reforming stage accounts for 30% of the exit gas, with a  $H_2/CO$  ratio of 2.15. Moreover, during oxidation, the reactor produces an exit stream with a very high nitrogen concentration. The production of nitrogen and hydrogen in each reactor stage provides an opportunity for small-scale ammonia production (Burrows et al., 2021; Burrows & Bollas, 2022). Despite the different types and compositions of the exit gas, the reactor exit temperature is relatively constant, with minor fluctuations around  $820 \text{ }^\circ\text{C}$ . Fig. 3(c) shows that temperature gradients occur in the first 70% of the reactor length. The bed temperature increases during oxidation due to the exothermic reactions of this stage. The heat front is then pushed and stored at about 60% of the reactor length. Afterward, the reactor front

undergoes a temperature drop during reduction and reforming, as determined by the endothermic reactions of these stages. Despite these temperature gradients, the temperature bed is relatively uniform near the end of the reactor. Fig. 3(d) shows that Ni-NiO reactions mainly occur at the first 70% of reactor length at the oxidation and reduction stages. Ni form dominates the oxygen carrier bed during reforming, providing sufficient catalytic activity for the reforming reactions.

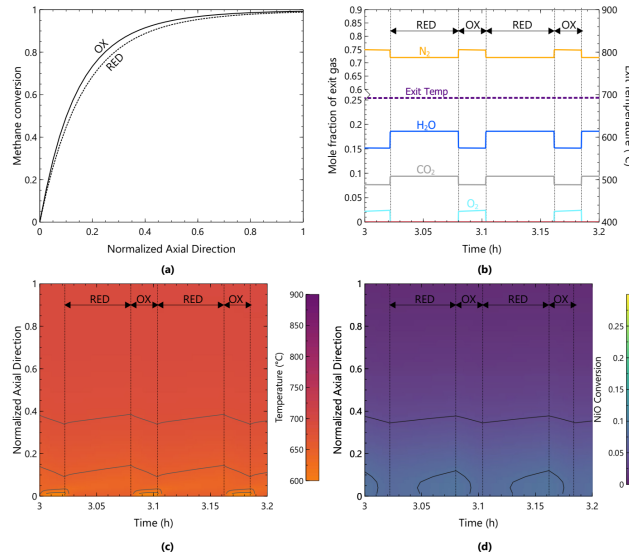


Figure 2: Reactor performance of optimized reactor design and control at cyclic steady state for the flaring application: (a) methane conversion at the end of Reduction and Oxidation stages; (b) exit mole fraction and temperature of the the reactor; (c) temperature contour plot of the reactor bed; and (d) NiO conversion contour of reactor bed – conversion of 0 refers to NiO and conversion of 1 refers to Ni. RED = Reduction stage; OX = Oxidation stage.

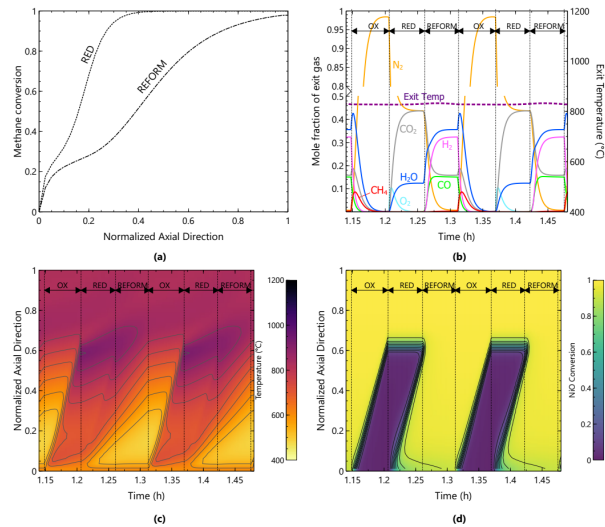


Figure 3: Reactor performance of optimized reactor design and control at cyclic steady state for the reforming application: (a) methane conversion at the end of Reduction and Reforming stages; (b) exit mole fraction and temperature of the reactor; (c) temperature contour plot of reactor bed; and (d) NiO conversion contour of reactor bed.

(d) NiO conversion contour of reactor bed – conversion of 0 refers to NiO and conversion of 1 refers to Ni. RED = Reduction stage; REFORM = Reforming stage; OX = Oxidation stage.

#### 4. Conclusions

Conceptual small-scale methane processing options were studied for lean methane from natural gas flaring and hydrogen production from stranded gas. In the flaring application, the intensified reactor converts more than 98% of the methane in the flue gas. In the reforming application, the intensified reactor converts more than 95% methane and produces syngas with a H<sub>2</sub>/CO ratio of 2.15. The reactor also produces a stream rich in N<sub>2</sub> during oxidation, which could enable small-scale liquid ammonia synthesis.

#### References

- Baharudin, L., Indera L, A. A., Watson, M. J., & Yip, A. C. K. (2021). Process intensification in multifunctional reactors: A review of multi-functionality by catalytic structures, internals, operating modes, and unit integrations. *Chemical Engineering and Processing - Process Intensification*, 168(July), 108561. <https://doi.org/10.1016/j.cep.2021.108561>
- Burrows, L., & Bollas, G. M. (2022). Stability Assessment of Small-Scale Distributed Ammonia Production Systems. *Industrial and Engineering Chemistry Research*. <https://doi.org/10.1021/acs.iecr.2c00631>
- Burrows, L., Gao, P. X., & Bollas, G. M. (2021). Thermodynamic feasibility analysis of distributed chemical looping ammonia synthesis. *Chemical Engineering Journal*, 426(May), 131421. <https://doi.org/10.1016/j.cej.2021.131421>
- E & P Forum. (1994). *Methods for Estimating Atmospheric Emissions from E&P Operations. The Oil Industry International Exploration & Production Forum Report No 2.59/197*.
- Han, L., & Bollas, G. M. (2016). Chemical-looping combustion in a reverse-flow fixed bed reactor. *Energy*, 102, 669–681. <https://doi.org/10.1016/j.energy.2016.02.057>
- Han, L., Zhou, Z., & Bollas, G. M. (2013). Heterogeneous modeling of chemical-looping combustion. Part 1: Reactor model. *Chemical Engineering Science*, 104, 233–249. <https://doi.org/10.1016/j.ces.2013.09.021>
- Han, L., Zhou, Z., & Bollas, G. M. (2016a). Model-Based Analysis of Chemical-Looping Combustion Experiments. Part I: Structural Identifiability of Kinetic Models for NiO Reduction. *AIChE Journal*, 62(7), 2419–2431. <https://doi.org/10.1002/aic>
- Han, L., Zhou, Z., & Bollas, G. M. (2016b). Model-Based Analysis of Chemical-Looping Combustion Experiments. Part II: Optimal Design of CH<sub>4</sub>-NiO Reduction Experiments. *AIChE Journal*, 62(7), 2432–2446. <https://doi.org/10.1002/aic>
- Irhamna, A. R., & Bollas, G. M. (2023). Intensified reactor for lean methane emissions treatment. *AIChE Journal*, 1–40. <https://doi.org/10.1002/aic.18040>
- Kidnay, A. J. ., Parrish, W. R. ., & McCartney, D. G. (2020). *Fundamentals of Natural Gas Processing*. CRC Press, Taylor & Francis Group. <https://www.ptonline.com/articles/how-to-get-better-mfi-results>
- Process Systems Enterprise. (2021). *gPROMS*. [www.psenterprise.com/products/gproms](http://www.psenterprise.com/products/gproms)
- Semieniuk, G., Holden, P. B., Mercure, J. F., Salas, P., Pollitt, H., Jobson, K., Vercoulen, P., Chewpreecha, U., Edwards, N. R., & Viñuales, J. E. (2022). Stranded fossil-fuel assets translate to major losses for investors in advanced economies. *Nature Climate Change*, 12(6), 532–538. <https://doi.org/10.1038/s41558-022-01356-y>
- Sherrick, J. J., Mueller, G. R., & Loos, K. R. (2011). *Texas Commission on Environmental Quality Work Assignment 5 Draft Flare Waste Gas Flow Rate and Composition*.
- Tan, J., Ji, Y. N., Deng, W. S., & Su, Y. F. (2021). Process intensification in gas/liquid/solid reaction in trickle bed reactors: A review. *Petroleum Science*, 18(4), 1203–1218. <https://doi.org/10.1016/j.petsci.2021.07.007>
- Williams, J. P., Regehr, A., & Kang, M. (2021). Methane Emissions from Abandoned Oil and Gas Wells in Canada and the United States. *Environmental Science and Technology*, 55(1), 563–570. <https://doi.org/10.1021/acs.est.0c04265>

# Optimal capacity planning integrating ammonia storage for power and heat decarbonisation

Georgios L. Bounitsis, Vassilis M. Charitopoulos

*The Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, University College London, Torrington Place, London WC1E 7JE, UK*  
[v.charitopoulos@ucl.ac.uk](mailto:v.charitopoulos@ucl.ac.uk)

## Abstract

Deep decarbonisation strategies for achieving the ambitious net-zero target pledged by the UK require thorough investigation of sector coupling dependencies. With heat being the single biggest energy consumer in the UK, pathways including dense energy carriers, which can store the excessive intermittent renewable energy, constitute a focal point of interest. In this work the role of ammonia as an energy vector for long-term storage in Great Britain's (GB) power system planning is evaluated. A linear programming spatially explicit snapshot model is formulated to set a fine-grained temporal resolution and capture seasonality of renewable sources. Decisions on capacity planning and operation of the system for a target year are optimised. Moreover, both power and heat demands are optimised as the heat demand mix is determined. The proposed model is applied on a case study of GB's power system in 2040 and the role of ammonia is investigated.

**Keywords:** Ammonia; Sector coupling; Heat decarbonisation; Energy systems planning; Hydrogen.

## 1. Introduction

### 1.1. Motivation

UK has been one of the leading countries in the race towards Net Zero. The fulfilment of its goals renders the integration of renewable sources in the UK's energy systems imperative. However, increased volatility within the electricity grid and the short-term duration of battery storage do not enable an efficient exploitation of the excessive renewable energy. Resorting to chemical energy storage options, hydrogen ( $H_2$ ), which can be also renewable-based produced, constitutes a challenging dense energy carrier (DEC) for long-term storage due to safety and storage costs aspects. Hence, ammonia ( $NH_3$ ) has emerged as a promising alternative to store the hydrogen-based energy. As Haber-Bosch synthesis allows exclusively renewable-based  $NH_3$  production research studies have focused on its role (Valera-Medina & Bañares-Alcantara, 2021).

### 1.2. Literature review

Ammonia's renewable based production constitutes a focal point of research recently, as its benefits for the energy systems can be twofold: (a) its long-term storage and transportation are inexpensive and safe compared to other alternatives, and (b) the stored energy can be efficiently contributed back to energy systems at different time or geographical regions (Valera-Medina & Bañares-Alcantara, 2021). Therefore, critical research studies on power systems scheduling and planning have recently included both renewable-based ammonia production and ammonia-to-power pathways in their considerations (Ikäheimo et al., 2018). Palys & Daoutidis (2020) proposed a capacity planning and scheduling model in order to minimise the levelised cost of ammonia in a

specific location. Recently, technoeconomic analyses have additionally demonstrated encouraging results on the forecasted levelised cost of electricity from renewable ammonia (Cesaro et al., 2021). Regarding supply chain optimisation, a recent study indicates the impact of renewable ammonia as mean of satisfying demand in the ammonia supply chain (Allman et al., 2017). Finally, ammonia is also considered for energy carriers' supply chain optimisation (Tso et al., 2019).

Here we aim to present a spatially explicit linear programming (LP) model for the strategic planning and operation of GB's power sector including energy carriers' production. The proposed framework aims to assess the role of ammonia as energy vector in GB system's decarbonisation. To this end, a case study focusing on GB's optimal planning and operation in year 2040 is investigated. The remainder of the article is organised as follows: in Section 2 the superstructure of the problem and mathematical developments are outlined. In Section 3, the detailed case study and results are discussed. Finally, future directions are proposed in Section 4.

## 2. Optimisation of integrated production and scheduling of power sector

### 2.1. Superstructure of the problem

Electricity, hydrogen and ammonia systems' planning are integrated in the investigated problem, as sector coupling considerations have been proven beneficial for energy planning (He et al., 2021). The proposed snapshot model is capable to evaluate the mix of technologies and the operation of the sectors in a future year while considering hourly time resolution. Moreover, cases for decarbonisation pathways are investigated while satisfying an optimised region-specific electricity demand. Concurrently the energy mix for the residential heat demand satisfaction is determined. Decisions on the latter are taken into consideration due to its high energy requirements and consequent carbon intensity (Charitopoulos et al., 2022).

In this work, energy planning problem delves into GB's power system using a high-resolution spatial analysis, as this is considered to be imperative on energy planning of whole country's power system (Koltaklis et al., 2014). Thus, GB is divided in 13 regions ( $g \in G$ ) according to the Local Distribution Zones (LDZ) of the gas network (Fig. 1). Regarding decisions on the GB's infrastructure, data on existing technologies' ( $j \in J$ ) generation and storage capacities as well as transmission capacities are considered. Moreover, bidirectional interconnection capacities to third parties ( $i \in I$ ) for electricity supply are set. Regarding operational decisions, the power system is optimised in a certain future time period employing a full year hourly time discretisation ( $h \in H$ ). In this context, energy carriers, i.e.,  $H_2$  and  $NH_3$ , are included introducing alternative electricity generation ways, energy storage and transportation alternatives towards decarbonisation. In particular,  $NH_3$  pathways can be considered using  $H_2$  and electricity as initial flows to Haber-Bosch (HB) process in order to produce  $NH_3$ . This energy can be stored in liquid  $NH_3$  storage tanks and be contributed back to the system for electricity generation in  $NH_3$  gas turbines (GT). In this work, HB process is assumed to include the accompanied air separation units (ASU) for the necessary nitrogen ( $N_2$ ) production. In real world, this process can be also exclusively renewable based. The superstructure of the network for the problem is envisaged in Fig. 2. Overall, the resources ( $a \in A = \{Elec, H_2, NH_3\}$ ) are depicted in circles and can be considered as carriers of energy. Storage and interregional transmission/transportation capabilities are also

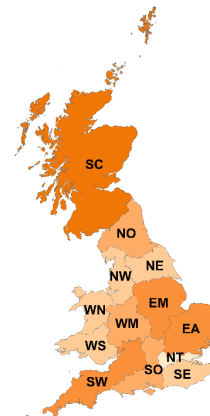


Figure 1: LDZ of Great Britain.

available for them. Finally, region-specific hourly electricity and heat demands profiles are satisfied, while concurrently optimising electricity demand (and its peak) and determining the breakdown of heat demand satisfaction.

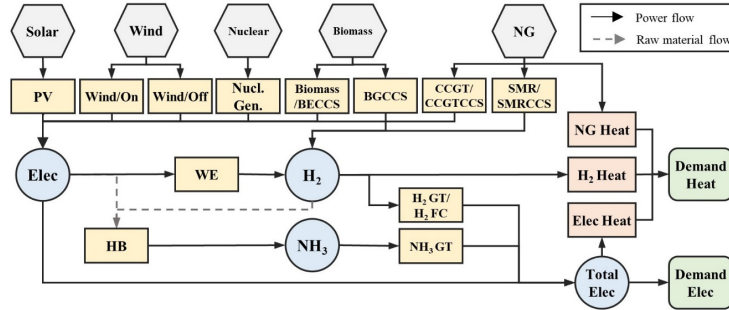


Figure 2: Superstructure of power and heat supply chain.

## 2.2. Mathematical modelling

An LP spatio-temporal model is developed to address the integrated problem. Power and heat hourly demand profiles for a future year ( $DEM_{gh}^{elec}$  and  $DEM_{gh}^{heat}$  respectively) are used as input for the model. The breakdown of heat demand is calculated by the model:

$$DEM_{gh}^{heat} = \frac{Q_{gh}^{elec}}{COP_{gh}} + \frac{Q_{gh}^{H2}}{\eta^{H2,boiler}} + \frac{Q_{gh}^{NG}}{\eta^{gas,boiler}} \quad \forall g \in G, h \in H \quad (1)$$

Efficiency of the  $H_2$  and natural gas (NG) boilers ( $\eta$ ) is set equal to 90% and the coefficient of performance ( $COP_{gh}$ ) of heat pumps, which consume electricity, is calculated based on the real-world temperatures data (ranges from 1.6 to 3.5). Capturing power requirement from heat electrification ( $Q_{gh}^{elec}$ ) the real electricity demand ( $D_{gh}$ ) is:

$$D_{gh} = (1 + DL) \cdot (DEM_{gh}^{elec} + Q_{gh}^{elec}) \quad \forall g \in G, h \in H \quad (2)$$

Accounting for electricity's distribution losses is achieved through factor  $DL$ , which is set equal to 6.5% according to calculations from historical data (DUKES, 2022). Concerning the resources balances, generation/production or consumption of the resource ( $P_{gjh}$ ), storage charging ( $CH_{agjh}$ ) and discharging ( $DC_{agjh}$ ) and bidirectional transmission/transportation ( $TR_{argg'h}$ ) (among losses) for the satisfaction of demands are taken into consideration. As far as the electricity, bidirectional interconnection to third countries ( $PIM_{igh}$ ), renewables curtailment ( $LC_{gh}$ ) and load shedding ( $LS_{gh}$ ) can be also considered. The spatial explicit mass balance is given by Eq. (3):

$$\begin{aligned} & \sum_{j \in PR_{aj}} P_{gjh} \cdot (1 - PL_j) + \sum_{j \in ST_{aj}} DC_{agjh} + \sum_{i \in IG_{ig}} PIM_{igh} \cdot (1 - L_{ig}^{int}) \Big|_{a=\{elec\}} \\ & + \sum_{(r,g') \in TG_{argg'}} TR_{argg'h} \cdot (1 - LD_{g'g} \cdot L_{ar}^{tr}) = D_{gh}|_{a=\{el.\}} + Q_{gh}^{H2}|_{a=\{H2\}} \\ & + \sum_{j \in ST_{aj}} CH_{agjh} + \sum_{j \in CON_{aj}} \frac{P_{gjh}}{\eta_j^c} + \sum_{(r,g') \in TG_{argg'}} TR_{argg'h} + (LC_{gh} - LS_{gh}) \Big|_{a=\{elec\}} \\ & \forall a \in A, g \in G, h \in H \end{aligned} \quad (3)$$

Crucial aspects for the planning are also enforced by capacity constraints. For the sake of stability of the power supply chain, the system-wide electricity peak demand ( $D^{peak}$ ) must be met by the installed capacities ( $Cap_{gj}$ ) (Eq. 5).  $D^{peak}$ , which is optimised by the

model (Eq. 4), is additionally multiplied by a security margin factor ( $RM$ ). Moreover, de-rating factors ( $DF_j$ ) are considered as the technologies are not assumed capable to constantly operate at 100% of their maximum capacity (National Grid ESO, 2022a).

$$D^{peak} \geq \sum_{g \in G} D_{gh} \quad \forall h \in H \quad (4)$$

$$\sum_{g \in G, j \in (ST_{elec,j} \cup PR_{elec,j})} DF_j \cdot Cap_{gj} \geq D^{peak} \cdot (1 + RM) \quad (5)$$

Decisions on capacity investments are bounded by building ratios and land availability data (Heuberger et al., 2017; Samsatli et al., 2016). For each sector separate carbon budgets are in place which cumulatively are in line with the GB's Climate Change Committee (CCC) net-zero trajectory. In summary, the proposed LP model aims at the minimisation of the total system cost ( $TSC$ ). Terms regarding the capital cost (using capacity recovery factors), variable operational costs (including storage, fuels, curtailment, value of loss load), interconnection costs and carbon taxes are included.

$$\begin{aligned} & \text{minimise} \quad TSC = CAPEX + OPEX + CINT + CCT \quad (6) \\ & \text{subject to:} \quad \begin{cases} \text{Capacity bounds and expansion constraints} \\ \text{Energy balances and Technologies operation constraints} \\ \text{Resources storage and transmission constraints} \\ \text{Carbon emission constraints} \end{cases} \end{aligned}$$

### 3. Case study: Power and heat decarbonisation in the UK in 2040

#### 3.1. Data preparation

The proposed model is tested for the planning of GB's energy system in 2040 using 2015 data and installed infrastructure as basis. Profiles of electricity demands can be derived by historical data of National Grid (National Grid, 2022). Additional profiles of heat demand are based on hourly natural gas consumption data obtained from the UK gas distribution companies. Both demands are calculated by projecting the data of the base year to the predicted demand values by National Grid Future Energy Scenarios (FES). In particular, "System Transformation" scenario is employed (National Grid ESO, 2022b). Finally, carbon goals for year 2040 are calculated based on data by the UK's Sixth Carbon Budget report (Climate Change Committee, 2020). Predictions on the related fuels' prices and carbon taxes are also reported in FES. Regarding renewables availability we use Renewables.Ninja platform (Pfenninger & Staffell, 2016). The model is calibrated using historical data (of base year 2015) to introduce corrective impact to availability and interconnection prices data (DUKES, 2022; ENTSO-E, 2022). The capacity allocation of the 2015 UK system can be found on DUKES statistics and techno-economic parameters predictions till year 2040 can be retrieved from recent reports (BEIS, 2018; IEA, 2021).

#### 3.2. Results and discussion

First of all, electricity demand in 2040 is optimised through the model and heat electrification decisions play a crucial role on this. In Fig. 3 the breakdown of heat demand mix in 2040 is given. The consequent increase on electricity demand is accompanied with decisions on investments on electricity generation technologies and particularly in renewable technologies. The total capacity, load factors and capacity allocation for year 2040 are determined by the model and are visualised in Figs. 4-5. Sizes in Fig. 4 are representative of the relative total generation.

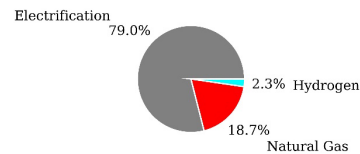


Figure 3: Heat demand mix in 2040.

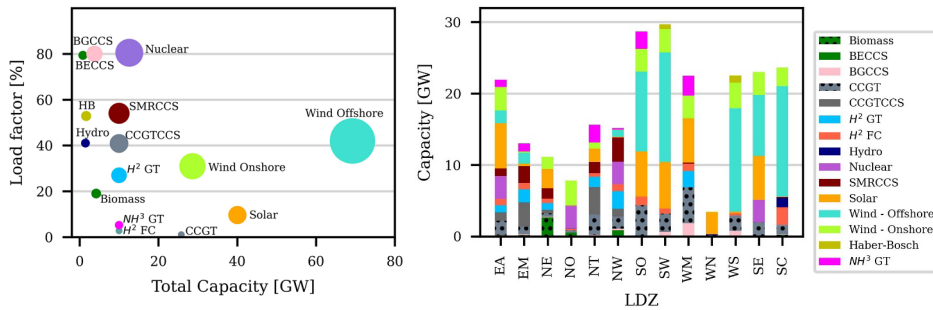


Figure 4: Load factors in 2040.

Figure 5: Capacity allocation in the UK in 2040.

Renewable technologies do not achieve high load factors because of their intermittent nature and the seasonality of renewable sources. In contrast, nuclear generation can provide high loads constantly. Furthermore, biomass-fueled technologies (BECCS & BGCCS) display high load factors and also contribute to Net Zero goal via carbon capture and storage (CCS). The remaining thermal technologies operate with lower load factors because of the decarbonisation regulations. However, their installed capacities are vital for system's security considerations. Focusing on renewables and DEC, full year profiles of generation and production are envisaged in Fig. 6 (colors as in Fig. 5). Although during summer small generation rates from thermal or DEC-fueled technologies are observed, during winter months the latter seem to contribute significantly. During summer the higher solar irradiation assist to meet demand. Regarding DEC production, even though hydrogen production displays seasonality, notable production is executed through whole year. This is not true for ammonia, as Haber-Bosch is demonstrated to yield further during summer.

Table 1: Total Storage Capacity [GWh].

Grid-level	73.7
Hydrogen	100.6
Ammonia	4,939

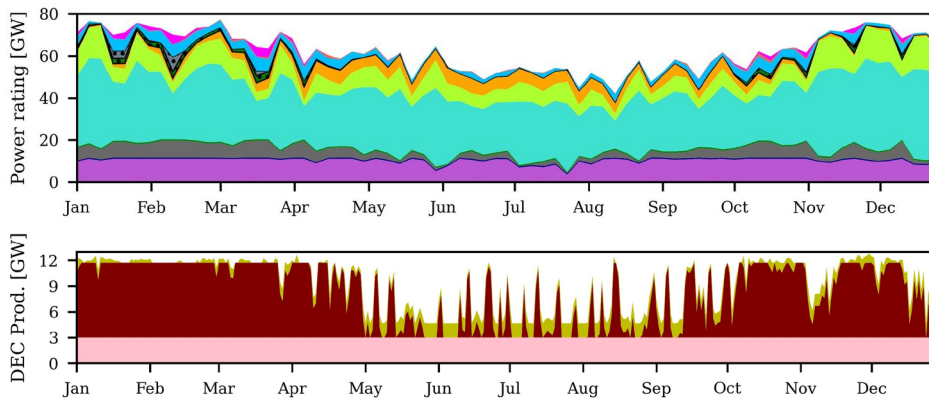


Figure 6: Electricity generation and DEC production profiles in GB in 2040.

Results from the optimisation model demonstrate the role of ammonia as interseasonal energy carrier. Insights from the installed storage capacities in Table 1 also indicate this. Furthermore, stored ammonia's amount is visualised in Fig. 7. Ammonia storage is consumed during the colder months contributing to power supply chain. Finally, the carbon intensity of the system is drastically enhanced as the calculated goal from CCC for 7.87 MtCO<sub>2</sub> for the considered system in 2040 is achieved. In parallel, the levelised cost of electricity is calculated to £65.1 per MWh of electricity supplied. This is competitive to third countries' leading to net exports of 6.6 TWh for the whole year.



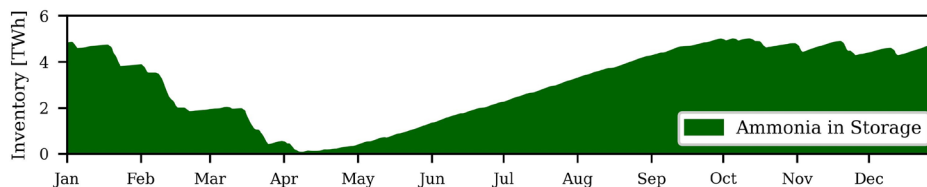


Figure 7: Inventory level of stored ammonia in GB in 2040.

#### 4. Conclusions

The proposed spatially explicit snapshot model's results for the capacity planning and operation of the power system demonstrate the significant role of ammonia as energy vector for UK's decarbonisation. In particular, ammonia offers alternative pathways for power generation and interseasonal energy storage. Future work is directed on the planning and decarbonisation of power systems accounting for stochastic uncertainty.

**Acknowledgements:** Financial support from the EPSRC (under project EP/T022930/1) is gratefully acknowledged.

#### References

- A. Allman, P. Daoutidis, D. Tiffany, S. Kelley, 2017. A framework for ammonia supply chain optimization incorporating conventional and renewable generation. *AIChE J.*, 63, 4390–4402.
- BEIS, 2018. Hydrogen Supply Chain: Evidence base UK. London: Department for Business, Energy & Industrial Strategy.
- Z. Cesaro, M. Ives, R. Nayak-Luke, M. Mason, R. Bañares-Alcántara, 2021. Ammonia to power: Forecasting the levelized cost of electricity from green ammonia in large-scale power plants. *Appl. Energy*, 282, 116009.
- V.M. Charitopoulos, M. Fajardy, C.K. Chyong, D.M. Reiner, 2022. Cambridge Working Papers in Economics. The case of 100% electrification of domestic heat in Great Britain
- Climate Change Committee, 2020. The Sixth Carbon Budget. The UK's path to Net Zero.
- DUKES, 2022. Digest of UK Energy Statistics (DUKES) 2022 – National Statistics.
- ENTSO-E, 2022. European Network of Transmission System Operators for Electricity.
- G. He, D.S. Mallapragada, A. Bose, C.F. Heuberger, E. Gençer, 2021. Sector coupling via hydrogen to lower the cost of energy system decarbonization. *Energy Environ. Sci.* 14, 4635–4646.
- C.F. Heuberger, E.S. Rubin, I. Staffell, N. Shah, N. Mac Dowell, 2017. Power capacity expansion planning considering endogenous technology cost learning. *Appl. Energy* 204, 831–845.
- IEA, 2021. International Energy Agency. Ammonia Technology Roadmap.
- J. Ikäheimo, J. Kiviluoma, R. Weiss, H. Holttinen, 2018. Power-to-ammonia in future North European 100 % renewable power and heat system. *Int. J. Hydrog. Energ.*, 43, 17295–17308.
- N.E. Koltsaklis, A.S. Dagoumas, G.M. Kopanos, E.N. Pistikopoulos, M.C. Georgiadis, 2014. A spatial multi-period long-term energy planning model: A case study of the Greek power system. *Appl. Energy* 115, 456–482.
- National Grid, 2022. Historical Data.
- National Grid ESO, 2022a. Electricity Market Reform. Capacity Market.
- National Grid ESO, 2022b. Future Energy Scenarios.
- M. J. Palys, P. Daoutidis, 2020. Using hydrogen and ammonia for renewable energy storage: A geographically comprehensive techno-economic study. *Comput. Chem. Eng.*, 136.
- S. Pfenninger, I. Staffell, 2016. Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy* 114, 1251–1265.
- Samsatli, S., Staffell, I., Samsatli, N.J., 2016. Optimal design and operation of integrated wind-hydrogen-electricity networks for decarbonising the domestic transport sector in Great Britain. *Int. J. Hydrogen Energy* 41, 447–475.
- W.W. Tso, C.D. Demirhan, S. Lee, H. Song, J.B. Powell, E.N. Pistikopoulos, 2019. Energy Carrier Supply Chain Optimization: A Texas Case Study. *Comput. Aided Chem. Eng.* 47
- A. Valera-Medina, R. Bañares-Alcantara, 2021. Techno-economic challenges of green ammonia as energy vector. Academic Press.

# European Union's biomass availability for Sustainable Aviation Fuel production and potential GHG emissions reduction in the aviation sector: An analysis using GIS tools for 2030

Sivaramakrishnan Chandrasekaran,<sup>a</sup> Nora B. Salah,<sup>a</sup> John A. Posada,<sup>a</sup>

<sup>a</sup>*Delft University of Technology, Mekelweg 5, Delft and 2628 CD, the Netherlands*

## Abstract

To be the first carbon-neutral continent by 2050, the European Union (EU) should decarbonize the aviation sector. According to the ReFuel initiative, sustainable aviation fuels (SAFs) are crucial in reducing carbon emissions from the sector. The Clean Sky 2 program by the EU commission, shortlisted four promising technologies - hydro-processed esters and fatty acids (HEFA), Fischer-Tropsch (FT), fast pyrolysis (FP), and alcohol-to-jet (ATJ) for the production of SAFs from bio-based sources. This study addresses the potential of these four technologies to reduce net and total greenhouse gas (GHG) emissions in the aviation sector. With a focus on mapping feedstock availability in 33 European countries for meeting the national demand in 2030. The investigation identified the best pathway combinations for each country, having the highest GHG emissions reduction while satisfying fuel demand when considering different degrees of biomass competition. Without any political and economic barriers to SAF production and biomass competition, we estimated a sufficient biomass supply exists to support the European SAF demand across all forecasted scenarios in 2030.

**Keywords:** Sustainable aviation fuels, GIS tools, Greenhouse gases emission, biomass conversion

## 1. Introduction

Nowadays, there is greater accessibility to air travel, with the majority of flights being used for passenger transport rather than freight. Accordingly, aviation contributes significantly to greenhouse gas (GHG) emissions. By 2050, the GHG emissions from the aviation sector are predicted to be 3.1 billion tons globally. Which is a sharp rise from 0.78 billion tons in 2015 (Doliente et al. 2020). To reduce the impact of these emissions on climate change and also to become the first carbon-neutral continent in the world, the European Union (EU) is taking different initiatives such as the EU Green deal and ReFuelEU aviation initiatives (Soone 2022; Zachmann, Tagliapietra, and Claeys 2019). Aligning with the same goals, CleanSky2 initiated TRANSCEND (Technology Review of Alternative and Novel Sources of Clean Energy with Next-generation Drivetrains). TRANSCEND investigated the potential of drop-in sustainable aviation fuels (SAFs), as a possible means of reducing aviation's influence on the environment. Drop-in SAFs are a new class of aviation fuel blended with regular fossil jet fuel and used in existing aircraft engines without significant changes. It is believed that bio-based drop-in fuels can significantly mitigate the GHG impacts on the aviation sector. TRANSCEND identified four promising technologies for the production of SAFs from bio-based sources. These

include the hydro processed esters and fatty acids (HEFA), Fischer–Tropsch (FT), fast pyrolysis (FP), and alcohol-to-jet (ATJ) technologies. This study contributes to this field by addressing the potential of these technologies to reduce net and total GHG emissions in the aviation sector based on biomass availability (type and amount) in different European countries while considering local demand for different forecasted scenarios (Muijden et al. 2021).

The objective of this study is to estimate the total supply of residual biomass available in Europe in 2030. Further, to evaluate potential SAF availability and related net reductions in GHG emission for various SAF conversion pathways. Finally, the results are shown in a comprehensive map indicating the suitable conversion pathway that produces the maximal net GHG emission reduction for each country. The research questions focused in this study are

- Is there enough biomass available to meet European SAF demand in 2030?
- What is the best combination of biomass type and conversion technology leading to the maximum net GHG emission reduction for each country?

## 2. Scope and Boundary

This study uses a mixed-method approach to address the mentioned research questions, where quantitative findings are validated and expanded using data gathered from peer-reviewed articles.

### 2.1. Spatial and temporal Boundary

The geographical scope of this study is limited to 33 European countries. The 33 European countries consist of EU27 + Serbia, Ukraine, Bosnia, Albania, Moldova, and Macedonia. With respect to the temporal scale, this study is limited to the medium-term year of 2030. This was chosen to align with the policy targets in the sector and the maturity of various conversion pathways. “Conversion pathway” refers to a technology and feedstock combination for analysis purposes.

### 2.2. Conversion Pathways

In terms of conversion pathways, technologies considered in this investigation are hydro-processed esters and fatty acids (HEFA), Fischer–Tropsch (FT), fast pyrolysis (FP), and alcohol-to-jet (ATJ) technologies. Likewise, this study focuses on residual biomass as a SAF source because it requires no additional land, consumes lower water and has less commercial value than energy crops. Also, this conform with EU’s commitment to using non-food feedstock as biomass sources for SAF. The three types of residual biomass considered in this report are agricultural residues (AR), forestry residues (FR) and municipal solid waste (MSW).

Agriculture residues consists of lignocellulosic by-products, harvesting, logging, and post-harvest operations (e.g. wood processing, crushing, and milling). Examples of agricultural residues include cereal straw, rice straw and corn stover. Similarly, forest residues are by-products of timber-harvesting activities. Examples of forestry leftovers include unprocessed fragments of cut trees (such as stumps and branches), wood pulp, and sawdust and cutter shavings. Lastly, municipal solid waste is defined as waste collected from a municipality and commonly referred to as trash or garbage (e.g. food processing waste and industry and commercial processing waste). In this study, we assesses the biomass availability (in tonnes) of 36 feedstock types in 33 European

countries. Of the 36 feedstock types, 18 can be categorised as forest residue, 15 as agricultural residue, and 3 as municipal solid waste.

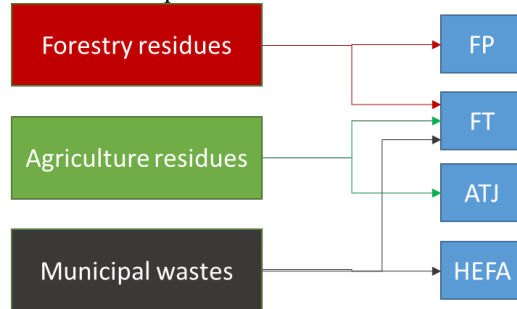


Figure 1: Schematic overview of the combinations of feedstock categories and conversion technologies considered.

The possible conversion pathways are shown in Figure 1. The selection is based on two criteria. The primary criterion is the compatibility between feedstock type and conversion technology. This compatibility depends on the type of biomass sources the technologies can process. The compatibility of feedstock type with certain technologies is obtained from the literature. For example, FT can process almost any carbon-rich material; therefore, it matches all three feedstock categories. The second criterion is whether literature data is readily available on process performance and GHG emissions intensity specific to the aviation sector. For example, the combination of FP and MSW shows promising results for the road sector but was excluded from this study due to missing data on process performance for SAF quality. Thus, only conversion pathways analyzed in previous studies in terms of SAF yield and associated GHG emissions specific to the aviation sector are included. In total, this study examined 85 conversion pathways for the 2030 scenario.

### 2.3. Geographic Information System

Biomass and GHG reduction potentials are mapped using ArcGIS Pro. ArcGIS is a geographic information system (GIS) that captures, stores and displays data related to spatial coordinates. It aids the user in understanding spatial patterns and relationships using layered maps. A demographic feature layer containing all EU countries provided by the ArcGIS database was used as a base map. Data was then added to the base map to illustrate its spatial distribution at a national level. Since this study uses a country-specific approach, the NUTS0 level is the most suitable to illustrate the results.

## 3. Scenario Evaluation

### 3.1. Estimation of potential SAF production

#### 3.1.1. Biomass availability and SAF demand

The data on biomass availability for 2030 by type and country is collected using the S2BIOM tool (Dees et al. 2017). With specification of regional level, year and feedstock type, this tool generate data for the amount of biomass available (in kilo tonnes dry mass) per year and potential type combination as a map. For this study, the technical potential is considered. According to the S2BIOM tool, this potential illustrates the maximum lignocellulosic biomass volume available, without any competing uses and with sustainable constraints. For the SAF demand in 2030, data for each country was provided

by the Royal Netherlands Aerospace Centre (NLR). Figure 2 illustrates the spatial distribution of biomass and national SAF demand.

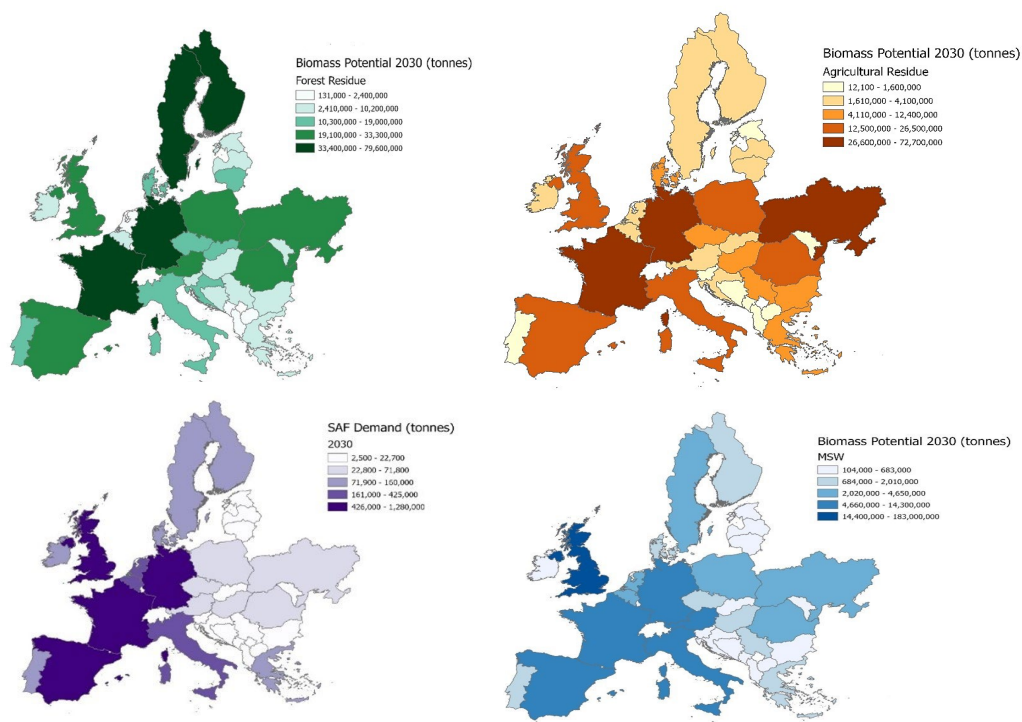


Figure 2: Spatial distribution of forestry residues (top left), Agriculture residues (top right), municipal solid waste (bottom right), and national SAF demand (bottom left).

### 3.1.2. Potential fuel output (PFO)

Potential fuel output can be calculated based on the potential SAF supply and national SAF demand. In order to estimate SAF supply, quantitative conversion yields for certain conversion pathways and biomass availability were taken into account. The conversion yields were obtained from literature data (Capaz et al. 2021; Santos et al. 2018). Due to data limitation for European context, some conversion pathway yields were obtained for the United States and assumed to be the same for the EU. Table 1 shows, example PFO of the most promising conversion pathways for UK in 2030 with 10% SAF blending policy.

$$PFO = \frac{\text{Potential SAF supply}}{\text{National SAF demand}} * 100\% \quad (1)$$

Table 1: PFO for some representative conversion pathways for UK in 2030 with 10% SAF blending

Technology	Feedstock	Biomass availability (in million tons)	SAF supply (in kilotonnes)	Demand (in kilotonnes)	PFO (in %)
FT	Thinnings from nonconifer trees	2,1	471	1283	37
FT	Thinnings from conifer trees	1,1	250	1283	19
FT	Oil seed rape straw	2,5	508	1283	40
FT	Sunflower straw	0,0	1	1283	0
ATJ	Oil seed rape straw	2,5	661	1283	51
ATJ	Sunflower straw	0,0	2	1283	0
FP	Thinnings from nonconifer trees	2,1	578	1283	45
FP	Thinnings from conifer trees	1,1	306	1283	24
HEFA	UCO	171,6	142445	1283	11102

### 3.2. Net GHG emissions and biomass competing uses

Based on environmental life cycle analysis, net, and total GHG emissions were calculated for each conversion pathway. A well-to-wake (WtWa) approach was followed in this investigation. We relied on studies that utilized the Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation (GREET) model to assess SAF conversion pathways' medium- and long-term GHG emission performance (Jong 2018). However, GREET uses the US as its geographic coverage of the input data. Therefore, the actual values for GHG emissions in Europe could be different. For the WtWa analysis, the CO<sub>2</sub> emissions from SAF combustion are considered part of the biogenic carbon cycle. The average WtWa emission intensity for typical jet fuel used in the United States (87.5 g CO<sub>2</sub> eq MJ<sup>-1</sup>) is utilized as the fossil fuel baseline.

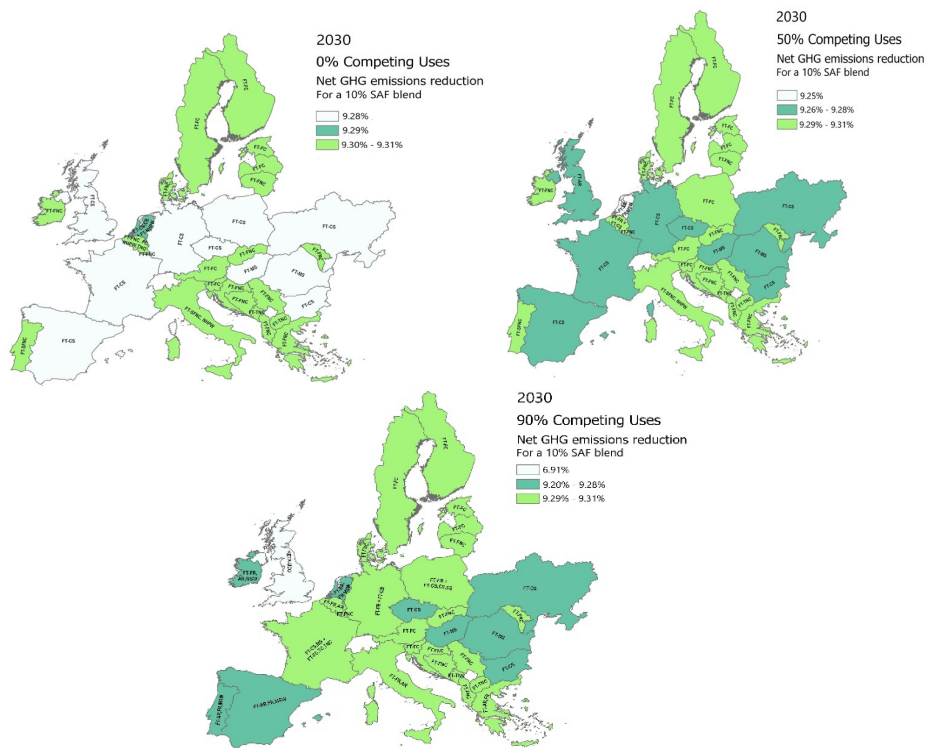


Figure 3: The net GHG emissions reduction potential (in %) based on 10% SAF blend for each European country in 2030 for different biomass availabilities.

Effect of biomass competing sources (0%, 50%, and 90%) was also analyzed as feedstock competition is expected in the future. Figure 3 illustrates the net GHG emissions reduction potential for the considered 33 European countries while accounting for biomass competition. The potential net GHG emissions reduction ranges between 9.25% and 9.31% across all biomass availability levels for most countries. The maximum total GHG emissions reduction potentials collectively for Europe in the 2030 scenario was found to be 21 Mt (0% competing uses), 19 Mt (50% competing uses), 15 Mt (90% competing uses). The difference in values is attributed to the selected feedstock category. For example, FT-AR pathway has slightly higher net GHG emissions (6.3g CO<sub>2</sub>/MJ SAF) than FT-FR (6g CO<sub>2</sub>/MJ SAF). The UK is the only country where biomass supply limits

potential net GHG emissions reduction. From Figure 3, to cover the UK's SAF demand for 90% competing uses, the conversion pathway is changed to HEFA-UCO, with a significantly lower net GHG emissions reduction potential (6.91%). Overall, the results indicate at least a 9% net reduction in GHG can be achieved for the 33 European countries, even when 90% competing uses are considered.

#### 4. Conclusions

In this research, we initially assessed if there is sufficient biomass to supply the SAF demand of Europe; and secondly, identified the best conversion pathways that lead to the highest GHG emissions reductions for each country. For a scenario in 2030 with a 10% blending of SAF with fossil jet fuel, we also investigated the impact of biomass competition with biomass uses varying from 0%, 50%, and 90%. Collectively, the 33 European countries can produce sufficient SAF to meet the total demand in 2030. This was the case for all the competing percentages. One key learning from this study was that conversion pathways with maximum yield need not always perform best in terms of GHG reduction. For example, in the UK, HEFA-UCO offers the highest PFO indicating higher yield, however, FT-FR had the highest GHG emission reduction while satisfying the demand. The conversion pathway that leads to the highest GHG emissions reduction across the EU for 2030 is FT – FR/AR/FR+AR, based on biomass availability. The maximum total GHG emissions reduction potentials collectively for Europe in the 2030 scenario is 21 Mt (0% competing uses), 19 Mt (50% competing uses), and 15 Mt (90% competing uses). Based on the analysis conveyed, it can be concluded that bio-based drop-in fuels can go a long way towards mitigating the impact of the aviation sector. However, there is a lot of room for refinement as a follow-up. Expanding the temporal scale to 2050 and assuming different blending percentages can spotlight some key areas of concern. Therefore, it is believed that the future SAF supply and the associated GHG emission reductions will be strongly impacted by policy incentives, the pace of technology development, and exact biomass supply and demand.

#### References

- Capaz, Rafael S. et al. 2021. "Mitigating Carbon Emissions through Sustainable Aviation Fuels: Costs and Potential." *Biofuels, Bioproducts and Biorefining* 15(2): 502–24.
- Dees, Matthias et al. 2017. "A Spatial Data Base on Sustainable Biomass Cost- Supply of Lignocellulosic Biomass in Europe - Methods & Data Sources. Project Report. S2BIOM – a Project Funded under the European Union 7th Framework Programme for Research. Grant Agreement N°608622." <https://zenodo.org/record/1478483>.
- Doliente, Stephen S. et al. 2020. "Bio-Aviation Fuel: A Comprehensive Review and Analysis of the Supply Chain Components." *Frontiers in Energy Research* 8(July): 1–38.
- Jong, Sierk De. 2018. "Green Horizons." Utrecht University.
- Muijden, J et al. 2021. "Final Results Alternative Energy and Propulsion Technology Literature Study Deliverable D1.1 of the TRANSCEND Project." (March).
- Santos, Catarina I. et al. 2018. "Integrated 1st and 2nd Generation Sugarcane Bio-Refinery for Jet Fuel Production in Brazil: Techno-Economic and Greenhouse Gas Emissions Assessment." *Renewable Energy* 129: 733–47. <https://doi.org/10.1016/j.renene.2017.05.011>.
- Soone, J. 2022. "ReFuelEU Aviation Initiative: Sustainable Aviation Fuels and the Fit for 55 Package ." *Think Tank | European Parliament* (June). [https://www.europarl.europa.eu/thinktank/en/document/EPRS\\_BRI\(2022\)698900](https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2022)698900).
- Zachmann, Georg, Simone Tagliapietra, and Grégory Claeys. 2019. "How to Make the European Green Deal Work Grégory." *European Commission* 13(9). <https://www.jstor.org/stable/resrep28626>.

# Statistical modeling of electrodes manufacture for CO<sub>2</sub> electroreduction to value-added products.

Jose Antonio Abarca, Guillermo Díaz-Sainz, Iván Merino-Garcia, Jonathan Albo, Angel Irabien.

*Departamento de Ingenierías Química y Biomolecular, Universidad de Cantabria, ETSIIT, Avenida de los Castros s/n, 39005, Santander, Spain.*

## Abstract

CO<sub>2</sub> electrocatalytic reduction to value-added products stands out as one of the most favorable utilization pathways for captured CO<sub>2</sub> to reduce emissions and mitigate climate change. However, current electrode fabrication techniques limit the implementation of the process at an industrial scale, as they are neither reproducible nor automatized. In this work, an automatic spray pyrolysis technique is studied by assessing the effect of different fabrication variables, such as (i) spraying nozzle height, (ii) distance between steps, and (iii) ink flowrate, from a statistical point of view, developing a linear regression model and a neural network-based predictive model that can forecast the behavior of the electrodes fabricated under different conditions. These statistical models are developed to advance from a rudimentary to an automatized electrode fabrication method, considering this as a first step towards establishing an electrode manufacturing protocol based on machine learning.

**Keywords:** CO<sub>2</sub> electroreduction, Spray pyrolysis technique, Gas Diffusion Electrode, Statistical modelling, Factorial experimental design.

## 1. Introduction

Climate change and global warming are part of the main challenges that society is facing nowadays. Several strategies have been proposed to minimize carbon dioxide (CO<sub>2</sub>) emissions in the energy-intensive industrial sectors [1], such as (i) the improvement of energy efficiency in industrial processes, (ii) the development of innovative sources of energy, or (iii) carbon capture, storage, and utilization (CCSU) approaches. The conversion of CO<sub>2</sub> to different chemical products is one of the most promising utilization pathways from both economic and environmental points of view [2]. The electrochemical reduction of CO<sub>2</sub> (CO<sub>2</sub>ER) to value-added products emerges as the most suitable for an industrial-level scale-up. Furthermore, this technology allows for storing energy from intermittent and renewable sources in the form of chemical bonds [3]. However, the current technological readiness level of the ER technology is under development stages, between 3 and 5 [4]. Therefore, research efforts are needed to bring ER process closer to industrial scale.

One of the key elements of CO<sub>2</sub>ER systems are the electrodes where CO<sub>2</sub> reduction reaction occurs. Among the several electrode configurations, gas diffusion electrodes (GDEs), which imply gas phase reagents, promote the contact between the reactive gas, the catalyst, and the electrolyte. The main advantages of these electrodes are that they keep high concentrations of CO<sub>2</sub> in the reaction area, and allow the operation at high current densities [7,8]. The GDE is composed of three main layers, (i) carbonaceous



support, (ii) a microporous layer (MPL), and (iii) a catalyst layer. The porous structure of the GDE allows a constant flow of gas to the electrocatalyst surface, where contacts the liquid electrolyte to produce the reduction reaction to obtain value-added products. In previous works, electrodes were fabricated by manual airbrushing [5–7], which is a rudimentary technique with high variability in the manufacturing process. The relative standard deviations (RSD) reached for formate rates were as high as 15%, which clearly supports the need to develop reproducible and dedicated GDE fabrication techniques. For this purpose, the spray pyrolysis technique is selected as the most suitable due to its simplicity of operation and automatization. The catalytic inks are atomized into small particles or drops that are deposited onto the support and placed over a heated surface [8]. The automatized operation allows the control of several fabrication variables that may affect electrode morphology [9].

This work aims to analyze, from a statistical point of view, the fabrication of Bi carbon-supported Gas Diffusion Electrodes (Bi/C GDE) for CO<sub>2</sub> electroreduction to formate, developing a linear regression and neural network model to predict the influence of the different fabrication variables in process performance.

## 2. Methodology

### 2.1. Spray Pyrolysis Fabrication Technique

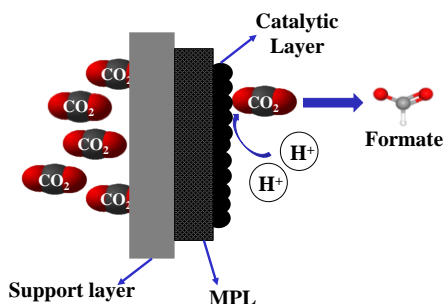


Figure 1. Scheme of the fabricated Bi/C GDE

The fabrication of the electrodes has been carried out using an ND-SP Mini Ultrasonic Spray Coater (Nadetech Innovations), whose functioning is based on an automatic spray technique that incorporates an ultrasonic reservoir that keeps the precursor solution dispersed. The GDEs fabricated, as illustrated in Figure 1, are composed of carbonaceous support, (Toray Carbon Paper TGP-H-60), an MPL based on a 3 wt.% dispersion of Carbon Black (Vulcan XC- 72R), and

PTFE (60 % wt., Sigma Aldrich) (40/60), and the catalytic ink 3 wt.% in isopropanol (Fischer Chemicals) with Bi carbon-supported nanoparticles as catalysts and Nafion D-521 (5% wt., Alfa Aesar) as a binder (70/30). The geometric area of the electrode is 10 cm<sup>2</sup>, with an MPL load of 2 mg cm<sup>-2</sup> and a catalytic loading of 0.75 mg cm<sup>-2</sup>, like previous studies in the research group [6].

### 2.2. Experimental set-up

The electrodes manufactured are tested in a filter-press electrochemical reactor [6] as working electrodes. 0.5 M KCl (PanReac, AppliChem) + 0.45 M KHCO<sub>3</sub> (PanReac, AppliChem) and 1 M KOH (85 % pellets, Pan Reac AppliChem), are used as catholyte and anolyte, respectively. The flow rates per geometric surface area for both electrolytes are 0.57 mL min<sup>-1</sup> cm<sup>-2</sup> and the pure CO<sub>2</sub> stream flow rate is 200 mL min<sup>-1</sup>. The formate concentration obtained in the output stream is measured by Ionic Chromatography (Dionex ICS 1100) [9].

### 2.3. Experimental design

The variables proposed are, spraying nozzle height, X<sub>1</sub>, the current density applied, X<sub>2</sub> as an operational variable, the distance between steps, X<sub>3</sub>, and ink flowrate, X<sub>4</sub>. For each different fabrication condition, the performance of the process to produce formate is

studied. For this purpose, a factorial experimental design 2<sup>4</sup>+8 is proposed following the approximations reported in our group [10–12]. These four variables have two levels, except for the spraying nozzle height, which has a central point. The value of the different levels of these variables is collected in Table 1.

Table 1. Variables and different levels were assessed in the factorial experimental design.

Variable		+	-	O
Nozzle height (mm)	X <sub>1</sub>	15	35	25
Current density (mA cm <sup>-2</sup> )	X <sub>2</sub>	200	90	-
Step distance (mm)	X <sub>3</sub>	1	3	-
Ink flowrate (mL h <sup>-1</sup> )	X <sub>4</sub>	20	10	-

The response variables considered are the average formate rate (r) and the energy consumption (CE), shown in Eq. 1 and Eq. 2, respectively.

$$\text{Formate rate} \left( \frac{\text{mol}}{\text{m}^2 \text{ s}} \right) = \frac{C_{\text{Formate}}}{A} \quad (1)$$

$$\text{Energy Consumption} \left( \frac{\text{kWh}}{\text{kmol}} \right) = \frac{Q \cdot V}{C_{\text{Formate}}} \quad (2)$$

Where C<sub>formate</sub> is the molar flowrate of formate obtained in the output stream of the electrochemical reactor (mol s<sup>-1</sup>), A is the geometric surface area of the electrode (m<sup>2</sup>), Q is the total charge supplied to the electrochemical cell (A), and V is the absolute cell potential (V).

The approximation to study the influence of the different variables in the electrode performance is to get a linear regression model. Furthermore, to avoid the negative effect of the absolute values in the regression model, the results of the output variables (average formate rate, and energy consumption) are normalized between [-1, 1]. The neural network design is carried out with the software Neural Designer (Artificial Intelligence Techniques, Ltd).

### 3. Results

Table 2 shows the statistical analysis of the factorial design 2<sup>4</sup>+8, only considering the linear interaction of the variables. For this analysis, the results of the factorial design are expressed in a linear regression model, obtaining the coefficients for each variable within a 95 % confidence level.

$$Y = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \beta_3 \cdot X_3 + \beta_4 \cdot X_4 \quad (3)$$

Equation 3 represents the linear regression model taken as the base for the development of the model to predict the behavior of the electrode under different fabrication and operation conditions.

It can be observed that only two variables have a significant effect on formate rates, the spraying nozzle height (X<sub>1</sub>) and the current density (X<sub>2</sub>). To determine the significance of a variable, as found before [12-14], the value of the coefficient (β) must be higher than 0.05, and simultaneously, the probability lower than 0.05. In the case of the nozzle height, the coefficient has a negative value (-0.065), which indicates that larger distances of the nozzle from the substrate surface during the electrode fabrication promote higher formate

rate values. This could be attributed to an improved catalyst distribution over the electrode surface, enhancing the surface-active area available for CO<sub>2</sub> reduction. In the case of current density, the effect is even more significant, as its coefficient has a higher value (0.769), which indicates that increasing the current supplied to the electrochemical reactor promotes formate rates in the output stream, as expected from previous results [6].

Table 2. Linear regression analysis of the factorial experimental design.

	Interception		X <sub>1</sub> , Nozzle height (mm)		X <sub>2</sub> , Current density (mA cm <sup>-2</sup> )		
	β <sub>0</sub>	β <sub>1</sub>	Error	Probability	β <sub>2</sub>	Error	Probability
$\bar{r}$	-0.093 ±0.154	-0.065 ±0.154	0.026	0.000	0.769 ±0.820	0.029	0.000
$\overline{CE}$	-0.157 ±0.237	0.051 ±0.121	0.034	0.000	0.675 ±0.756	0.038	0.000
	X <sub>3</sub> , Step distance (mm)			X <sub>4</sub> , Ink flowrate (mL min <sup>-1</sup> )		Adjustment	
	β <sub>3</sub>	Error	Probability	β <sub>4</sub>	Error	Probability	R <sup>2</sup>
$\bar{r}$	0.014 ±0.075	0.029	0.634	0.034 ±0.103	0.032	0.297	0.968
$\overline{CE}$	0.001 ±0.081	0.038	0.988	-0.026 ±0.117	0.043	0.546	0.929

On the other hand, for energy consumption, as it is for formate rates, only the nozzle height (X<sub>1</sub>) and the current density (X<sub>2</sub>) applied have a significant effect on the response variable. For both variables, the coefficients are positive, therefore, a smaller distance of the nozzle from the substrate (0.051) and a higher current density supplied to the cell (0.675) increase the energy consumption per molar unit of formate. In the case of the nozzle height, reducing the distances from the substrate causes the non-uniform deposition of the ink. This may induce the formation of agglomerates of the catalyst nanoparticles, worsening the surface conductivity of the electrode. The effect observed with current density makes sense, as for higher current densities it is necessary to supply more energy to reach the set conditions.

Furthermore, the values obtained for the interception indicate that the linear regression model is sensitive to changes produced in the variables assessed. In addition, the adjustment reached for both linear regression models is satisfactory (R<sup>2</sup>>0.92). However, simplified models can be proposed disregarding the variables that do not have enough significance, resulting in a worse adjustment, but at the same time, facilitating the understanding of the models.

The simplified linear regression models proposed for formate rate and energy consumption are shown below, where only the significant variables (nozzle height and current density) are taken into count:

$$\bar{r} = -0.093 - 0.065 \cdot X_1 + 0.769 \cdot X_2 \quad (4)$$

$$\overline{CE} = -0.157 + 0.051 \cdot X_1 + 0.675 \cdot X_2 \quad (5)$$

The spraying nozzle height curvature is analyzed. For this purpose, formate rate values for the different variable levels [-1 (35 mm), 0 (25 mm), 1 (15 mm)] are depicted in Figure 2.

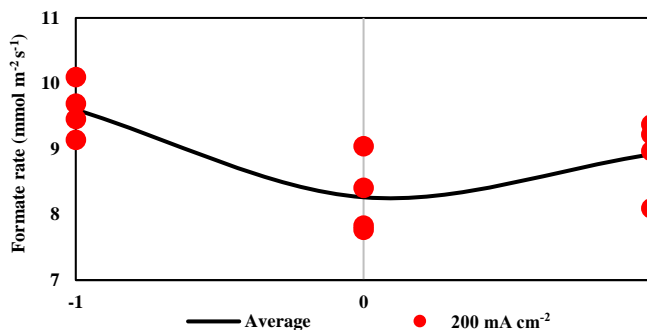


Figure 2. Spraying nozzle height curvature analysis as a function of formate rate values.

The data shows a non-linear function with a minimum at the central point (25 mm). This indicates that the optimization of this variable is much more complicated than the simplified linear regression model.

Based on the results obtained in linear regression, it can be seen that the variables in the model are only significant from a mathematical point of view. Thus, a predictive model is proposed using more advanced modeling methods, such as neural networks. In this case, the relevance of the selected input variables is assessed when constructing the neural network, which is formed by different neurons depending on its function: i) the scaling layer to eliminate the influence of the absolute value of the variables, ii) the perceptron layers, in which the model is able to learn the input correlations, iii) the unscaling layers, and iv) the bounding layers, to obtain the predicted output value. As observed in Figure 4, the input variables that have a significant effect in the model for the formate rate (Figure 4a) are the same as the linear regression, while for the energy consumption (Figure 4b), all input variables are considered significant.

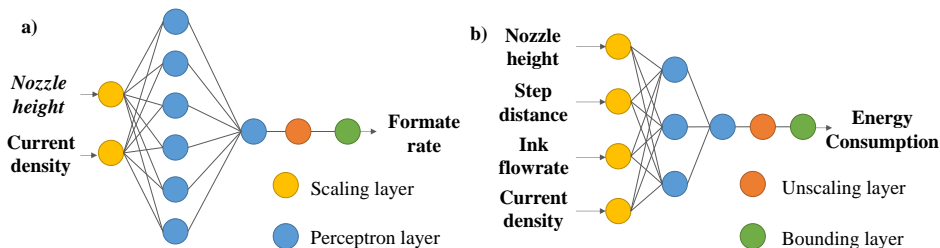


Figure 3. Neural networks for the predictive models: a) Formate rate, b) Energy consumption.

The models obtained from the neural networking have adjustment values ( $R^2$ ) of 0.97 and 0.94 for the formate rate and energy consumption, respectively. The information extracted from the first model confirms the linear regression findings: the nozzle height is the manufacturing variable that most influences the performance of the electrode in terms of formate rate. However, this predictive model determines that all manufacturing variables affect the physicochemical structure of the electrodes, and, as this affects their electrical resistance, it impacts on energy consumption.

#### 4. Conclusions

A statistical model that relates fabrication and operational variables with electrode performance for CO<sub>2</sub> ER to formate has been proposed. A factorial experimental design

2<sup>4</sup>+8 has been proposed, obtaining two different linear regression models, with a determination coefficient ( $R^2$ ) over 0.9, establishing the effect of each variable and their interactions on the response variables selected: (i) the average formate rate and (ii) energy consumption. Although not all variables and interactions are significant from a statistical point of view, two of them stand up as the most relevant ones, the nozzle height and the current density applied. Additionally, a neural network predictive model is proposed, confirming the information obtained for the formate rate linear regression model. However, for the energy consumption predictive model, all the fabrication variables are considered, in contrast to the linear regression model. These results are relevant to understand the behavior of the electrodes fabricated by the automatized spray pyrolysis technique and are useful as a first step for applying data science to the electrochemical process. As a result, future work will focus on working with neural networking to bring artificial intelligence and machine learning to this process, aiming at further developing a useful tool to predict the performance of electrodes manufactured at any scale, which may save time and experimental efforts.

## References

- [1] C. Hepburn, E. Adlen, J. Beddington, E.A. Carter, S. Fuss, N. Mac Dowell, 2019, The technological and economic prospects for CO<sub>2</sub> utilization and removal. *Nature*. 575, 7781, 87–97.
- [2] R. Chauvy, N. Meunier, D. Thomas, and G. de Weireld, 2019, Selecting emerging CO<sub>2</sub> utilization products for short- to mid-term deployment. *Applied Energy*. 236, 662–680.
- [3] A. Irabien, M. Alvarez-Guerra, J. Albo, and A. Dominguez-Ramos, 2018, C.A. Martínez-Huitle, M.A. Rodrigo, O. Scialdone (eds.) *Electrochemical Conversion of CO<sub>2</sub> to Value-Added Products.*, pp. 29–59.
- [4] M. Rumayor, A. Dominguez-Ramos, P. Perez, and A. Irabien, 2019, A techno-economic evaluation approach to the electrochemical reduction of CO<sub>2</sub> for formic acid manufacture. *Journal of CO<sub>2</sub> Utilization*. 34, 490–499.
- [5] G. Díaz-Sainz, M. Alvarez-Guerra, and A. Irabien, 2021, Continuous electroreduction of CO<sub>2</sub> towards formate in gas-phase operation at high current densities with an anion exchange membrane. *Journal of CO<sub>2</sub> Utilization*. 56, 101822.
- [6] G. Díaz-Sainz, M. Alvarez-Guerra, J. Solla-Gullón, L. García-Cruz, V. Montiel, and A. Irabien, 2019, CO<sub>2</sub> electroreduction to formate: Continuous single-pass operation in a filter-press reactor at high current densities using Bi gas diffusion electrodes. *Journal of CO<sub>2</sub> Utilization*. 34, 12–19.
- [7] G. Díaz-Sainz, M. Alvarez-Guerra, B. Ávila-Bolívar, J. Solla-Gullón, V. Montiel, and A. Irabien, 2021, Improving trade-offs in the figures of merit of gas-phase single-pass continuous CO<sub>2</sub> electrocatalytic reduction to formate. *Chemical Engineering Journal*. 405, 126965.
- [8] P.S. Maldar, M.A. Gaikwad, A.A. Mane, S.S. Nikam, S.P. Desai, S.D. Giri, 2017, Fabrication of Cu<sub>2</sub>CoSnS<sub>4</sub> thin films by a facile spray pyrolysis for photovoltaic application. *Solar Energy*. 158, 89–99.
- [9] N. Kumari, M.A. Haider, U. Anjum, and S. Basu, 2020, Identifying operating mechanism in the electrochemical reduction of CO<sub>2</sub> on thin-film praseodymium-doped ceria electrodes. *Ionics*. 26, 11, 5673–5684.
- [10] M. Alvarez-Guerra, S. Quintanilla, and A. Irabien, 2012, Conversion of carbon dioxide into formate using a continuous electrochemical reduction process in a lead cathode. *Chemical Engineering Journal*. 207–208, 278–284.
- [11] M. Alvarez-Guerra, A. del Castillo, and A. Irabien, 2014, Continuous electrochemical reduction of carbon dioxide into formate using a tin cathode: Comparison with lead cathode. *Chemical Engineering Research and Design*. 92, 4, 692–701.
- [12] A. Del Castillo, M. Alvarez-Guerra, and A. Irabien, 2014, Continuous electroreduction of CO<sub>2</sub> to formate using Sn gas diffusion electrodes. *AIChE Journal*. 60, 10, 3557–3564.

# Comparative exergoeconomic analysis of two Organic Rankine Cycle (ORC) configurations

Ladislao E. Méndez-Cruz,<sup>a</sup> Miguel A. Gutiérrez-Limón,<sup>b</sup> Raúl Lugo-Leyte,<sup>c</sup>  
Mauricio Sales-Cruz,<sup>a,\*</sup>

<sup>a</sup>Universidad Autónoma Metropolitana – Cuajimalpa, Av. Vasco de Quiroga No. 4871, Mexico City 05348, Mexico. \*[asales@cua.uam.mx](mailto:asales@cua.uam.mx)

<sup>b</sup>Universidad Autónoma Metropolitana – Azcapotzalco, Av. San Pablo No. 180, Mexico City 02200, Mexico.

<sup>c</sup>Universidad Autónoma Metropolitana – Iztapalapa, Av. Ferrocarril San Rafael Atlixco No. 186, Mexico City 09340, Mexico.

## Abstract

Currently, power generation systems that use residual thermal sources are being studied, such as the Organic Rankine Cycle (ORC) technology. This paper presents a comparative exergoeconomic analysis between a simple ORC and one with a preheater. Both configurations work with R134a as refrigerant, under the same waste heat source. The results show a 47.31% reduction in the exergoeconomic cost of residue formation when operating with the ORC with preheater. In addition, the unit production cost is 0.049 USD/kWh with the simple ORC, while it is reduced to 0.046 USD/kWh with the ORC configuration with preheater.

**Keywords:** Organic Rankine Cycle, exergoeconomic analysis, process configurations.

## 1. Introduction

The best use of fossil resources for power generation is through the employment of technologies that require residual gases as a thermal source. In this sense, the Organic Rankine Cycle (ORC) corresponds to this type of power cycles that use low-temperature thermal sources, to generate power from 10 kW to 10 MW (Nondy and Gogoi, 2021). There are various ORC configurations that have been studied from an exergoeconomic point (thermodynamic and economic) that aims to evaluate, diagnose, improve, and optimize this type of energy systems (Torres et al., 2002). Due to this, the exergetic cost theory is a methodology that is widely used and is based on the productive purpose and the formation of the residue (Torres and Valero, 2021).

The main configuration of the ORC consists of a turbine (T), an electric generator (EG), a condenser (COND), a condensate pump (P) and a heat recovery steam generator (HRSG), the latter composed of an economizer (EC), an evaporator (EV) and a superheater (SH). Two configurations are considered to carry out a comparative exergoeconomic study: a simple ORC and an ORC with preheater (PH), as shown in Figure 1. Both configurations operate with R134a as organic working fluid. The purpose of this work is to evaluate the behavior of the exergoeconomic cost of residue formation ( $\Pi_R$ ), and the unit exergoeconomic cost ( $c_p$ ) of power generation, for the same exergoeconomic cost of the external resource ( $\Pi_c$ ). Therefore, it is intended to analyze the contribution of the equipment in the cost for the residue formation generated by a configuration of the given system.

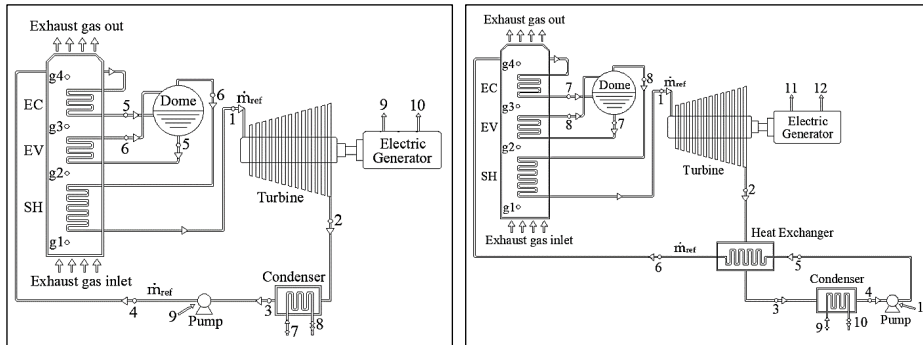


Figure 1. ORC plant: (a) simple, (b) with HE

## 2. Methodology

The framework begins with the construction of a productive structure, which defines which are the productive components that contribute to the formation of the product and the residue of the system. For this, the exergy flows of the resource and the product of the productive components are identified (Lugo-Mendez et al., 2020). From the configurations shown in Figure 1, their productive structures are built, which are shown in Figure 2. Both productive structures describe the product that comes from the environment (0) and is the system resource that corresponds to the exergy change of the exhaust gases ( $g_1 - g_4$ ), which is distributed into the HRSG components (SH, EV and EC) through node 1. The union of the products of these components is mainly used for the generation of power in the T and EG, which is distributed both for the generation of the useful product of the system, and in the power supplied to the P, of both configurations. Likewise, the products of the components of SH, EV, EC, P and HE (the latter for the second configuration) that enter node 2, contribute to the formation of the residue in COND in different proportions.

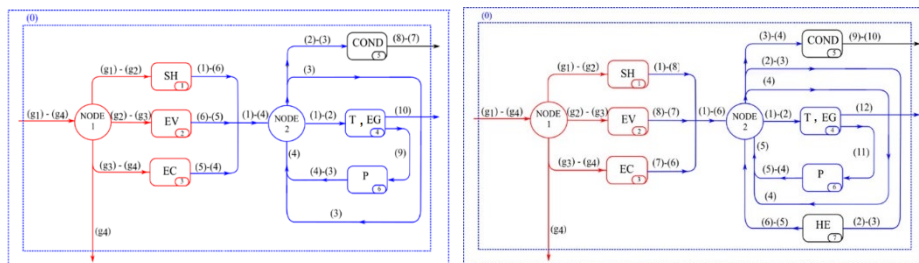


Figure 2. ORC productive structure: (a) simple, (b) with HE

Table 1 presents the exergy fluxes of the resource (fuel,  $F$ ), product ( $P$ ) and residue ( $R$ ) of the ORC components of both configurations, which are based on the definition of the resource product  $\dot{F} = \dot{P} + \dot{I}$ . Where, the irreversibility flux ( $\dot{I}$ ) generated by each component of the system is determined by the difference between the exergy flux of the resource ( $\dot{F}$ ) and the product ( $\dot{P}$ ).

The exergetic and exergoeconomic cost functions in the ORC economic model are based on the F-P-R representation of the system, where it relates the exergoeconomic variables of the system to external resources (Cuadra and Capilla, 2000). According to this representation, the exergetic fluxes are initially determined and later the exergetic costs

of the resource, product and irreversibility of the productive components of the system, as well as the exergetic cost for the formation of the residue, all of them based on the external resource. and the distribution coefficients of the product and the residue. Table 2 presents the matrix expressions that determine the exergetic flows and the exergetic costs of the product, resource and residue of the productive components, which are used for both ORC configurations.

Table 1. Exergy fluxes of fuel, product and residue.

Components of ORC	$\dot{F}$ (kW)	$\dot{P}$ (kW)	$\dot{R}$ (kW)	Components of ORC with HE	$\dot{F}$ (kW)	$\dot{P}$ (kW)	$\dot{R}$ (kW)
HRSG	$\dot{E}_{g1} - \dot{E}_{g4}$	$\dot{E}_1 - \dot{E}_4$	-	HRSG	$\dot{E}_{g1} - \dot{E}_{g4}$	$\dot{E}_1 - \dot{E}_6$	-
T	$\dot{E}_1 - \dot{E}_2$	$\dot{W}_T$	-	T	$\dot{E}_1 - \dot{E}_2$	$\dot{W}_T$	-
COND	$\dot{E}_2 - \dot{E}_3$	-	$\dot{E}_8 - \dot{E}_7$	COND	$\dot{E}_3 - \dot{E}_4$	-	$\dot{E}_9 - \dot{E}_{10}$
P	$\dot{W}_B$	$\dot{E}_4 - \dot{E}_3$	-	P	$\dot{W}_B$	$\dot{E}_5 - \dot{E}_4$	-

Table 2. Expressions of the exergetic flows and costs of the productive components for both ORC configurations.

Exergetic flows	Exergetic cost
$P = \langle P   F_c \text{ where } \langle P   = (K_D - \langle FP \rangle)^{-1}$	$P^* = \langle P^*   F_c \text{ where } \langle P^*   = (U_D - \langle FP \rangle - \langle RP \rangle)^{-1}$
$F = \langle F   F_c \text{ where } \langle F   = K_D \langle P  $	$F^* = F_c + \langle FP \rangle P^*$
$I = \langle I   F_c \text{ where } \langle I   = (K_D - U_D) \langle P  $	$R^* = \langle RP \rangle P^*$
$R = \langle R   F_c \text{ where } \langle R   = \langle RP \rangle \langle P  $	

The exergetoeconomic costs of the product as a function of external resources are determined from the balance of costs

$$\Pi_p = Z + \Pi_F \quad (1)$$

Where, Z represents non-energy factors of production, that is, factors of amortization, maintenance, operation, etc. On the other hand, the exergetoeconomic cost of the product ( $\Pi_p$ ) and of the resource ( $\Pi_F$ ), depending on the external resources and the distribution coefficients are

$$\Pi_p = (U_D - \langle FP \rangle - \langle RP \rangle)^{-1} \Pi_e \quad ; \quad \Pi_F = \Pi_e + \langle FP \rangle \Pi_p \quad (2)$$

The exergetoeconomic cost for the residue formation is given by

$$\Pi_R = \langle RP \rangle \Pi_p \quad (3)$$

The exergetoeconomic unit cost of the resource and of the product of the  $i$ th component exhibits the monetary amount to produce one kWh, and is represented as follows

$$c_{F,i} = \frac{\Pi_{F,i}}{F_i} \quad ; \quad c_{P,i} = \frac{\Pi_{P,i}}{P_i} \quad (4)$$

The exergetoeconomic cost of external resources ( $\Pi_e$ ) includes the investment and maintenance costs of the system,  $\dot{Z}_i$ , and the economic unit cost of fuels,  $\dot{E}_{0i}$ ; where, the cost of the equipment per unit of time ( $\dot{Z}_i$ ) is a function of the cost of the equipment, the annual amortization factor ( $a$ ) the maintenance factor ( $\phi$ ), and the number of operating hours per year ( $N$ ).



$$\Pi_e = \left[ \dot{Z}_i + c_i \dot{E}_{0i} \right]_{i=1, \dots, n} \quad (5)$$

$$\dot{Z}_i = \frac{\alpha \varphi}{N} Z_i; \text{ where } a = \frac{i(1+i)^n}{(1+i)^n - 1} \quad (6)$$

Table 3 shows the correlations used to determine the costs of the equipment that make up both the ORC configurations and the economic parameters of the system (Nondy and Gogoi, 2021).

Table 3. Equipment cost and economic parameters of the ORC for both configurations.

Equipment cost (USD)	Economic parameters	
$Z_{SH} = 309.143(A_{SH}) + 231.915$	Annual hours, $N$ (h)	7446
$Z_{EV} = 309.143(A_{EV}) + 231.915$	Maintenance factor, $\varphi$ (-)	1.06
$Z_{EC} = 309.143(A_{EC}) + 231.915$	Interest rate, $i$ (%)	12
$Z_T = 6000(\dot{W}_T)^{0.7}$	Years of life, $n$ (years)	20
$Z_P = 3540(\dot{W}_P)^{0.71}$		
$Z_{HE} = 1.3(190 + 310A_{HE})$		
$Z_{COND} = 1773(\dot{m}_{ref})$		

The heat exchange areas of each component of the HRSG (this is, SH, EV, EC and HE) are determined as follows

$$A_i = \frac{\dot{Q}_i}{U_i DTML} \quad (7)$$

Where the total heat transfer coefficients are obtained from Mohammadi et al. (2018). Finally, Table 4 presents the operating conditions of the ORC for both configurations. The conditions of the residual thermal source correspond to the exhaust gases that come from the hybrid combined cycle plant in the Valley of Mexico (CFE, 2005).

Table 4. ORC operating conditions for both configurations.

Environment		Organic fluid		Waste heat source	
$P_{atm}$ , (bar)	0.78	$T_1$ , (°C)	150	$T_{g1}$ , (°C)	185
$T_{amb}$ , (°C)	10 - 40	$\Delta T_{cond}$ , (°C)	15	$T_{g4}$ , (°C)	100
		$\eta_{sit}$ , (-)	0.90	$\dot{m}_{eg}$ (kg <sub>eg</sub> /s)	297.1
		$\eta_{sib}$ , (-)	0.80	$c_{peg}$ , (kJ/kg K)	1.13

### 3. Results

The simulation of the process was carried out using the Python and Julia software. Figures 3 show the exergoeconomic costs of the turbine product as a function of the total area of the HRSG, for different evaporation and ambient temperatures. It can be seen, an increase in the exergoeconomic cost and the total area in the HRSG with the increase in the evaporation temperature. Figure 3(a) shows an average increase in the area of 20 m<sup>2</sup> when going from an ambient temperature of 25 to 40°C. In addition, the exergoeconomic cost that oscillates between 224 and 238 USD/h, represents an average increase of 4.77%

when going from an evaporation temperature of 80 to 95°C for the different ambient temperatures. Figure 3(b) shows that the area variation for the same ambient temperature range occurs at temperatures above 90°C, with increments of 13 m<sup>2</sup>. On the other hand, for evaporation temperatures lower than 85°C, there is no significant increase in the area, that is, the area remains almost constant with the variation of the different environmental temperatures. Likewise, the exergoeconomic cost of the turbine product has an increase when adding the HE to the ORC; however, this exergoeconomic cost increases only by 3.81% with respect to the results shown in the ORC configuration when going from an evaporation temperature of 80 to 95°C.

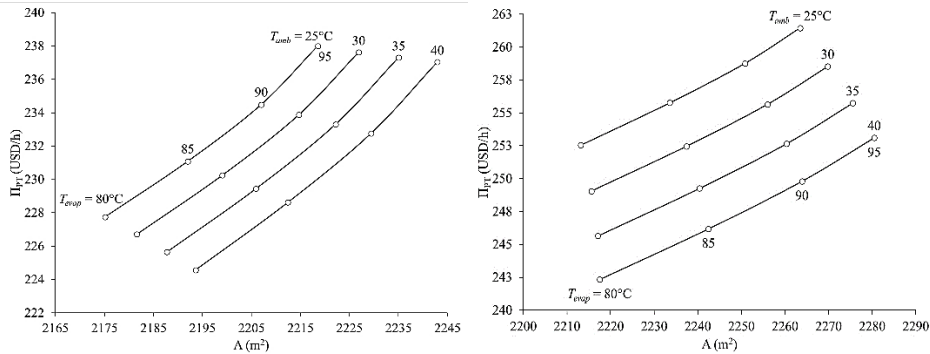


Figure 3. Exergoeconomic cost of the turbine product as a function of heat exchange area for (a) the simple ORC, and (b) the ORC with HE.

Figure 4 shows the exergoeconomic cost of the turbine product and power in both configurations, as a function of evaporation and ambient temperatures. Both configurations show an increase in cost when the evaporation temperature increases from 80 to 95°C; for the simple ORC shown in Figure 4(a), this increase is 4.77%; while in the case of the ORC with HE, it presents an increase of 3.82%. In addition, both configurations also show that, under the condition of the minimum ambient temperature and the highest evaporation temperature, the highest power and the highest cost are presented. The ORC configuration shows the highest power of 4,491.87 kW at a cost of 238 USD/h. On the other hand, the ORC with HE shown in Figure 4(b) shows that the power can be increased by 825.30 kW under the same operating conditions and the cost reflects an increase of 23.44 USD/h.

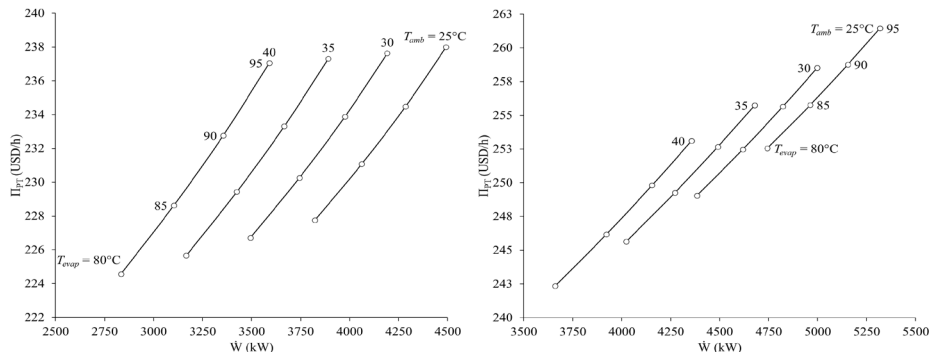


Figure 4. Exergoeconomic cost of the turbine product and power output for (a) the simple ORC, and (b) the ORC with HE.

Finally, Table 5 shows the exergoeconomic cost for the residue formation and the unit exergoeconomic cost of the components of both configurations, for the lowest ambient temperature and the highest evaporation temperature. The cost associated with the external resource (fuel) that enters the HRSG of the two configurations has the same value of 167 USD/h; however, the total cost shows a difference between the simple ORC and ORC with HE due to the increased cost of HE. On the other hand, when adding this component, the total exergoeconomic cost for the residue formation reflects a decrease of 47.31% and, in addition, the unit exergoeconomic cost of the functional product of the system also decreases from 0.049 USD/kWh to 0.046 USD/kWh.

Table 5. Exergoeconomic costs for the residue formation and unit exergoeconomic cost.

	ORC				ORC with HE			
	$\Pi_F$ (USD/h)	Z (USD/h)	$\Pi_R$ (USD/h)	$c_P$ (USD/kWh)	$\Pi_F$ (USD/h)	Z (USD/h)	$\Pi_R$ (USD/h)	$c_P$ (USD/kWh)
(1) SH	64.22	3.54	25.79	0.031	75.20	4.21	15.64	0.027
(2) EV	40.67	2.82	16.55	0.034	46.41	3.42	9.81	0.028
(3) EC	62.11	5.12	25.59	0.042	45.39	4.08	9.74	0.029
(4) T	0.00	38.14	0.00	0.049	0.00	54.23	0.00	0.046
(6) P	0.00	3.85	8.14	0.102	0.00	5.50	4.88	0.087
(7) HE	-		-	-	0.00	3.74	7.38	0.067
Total	167.00	53.47	76.06	-	167.00	71.44	40.07	-

#### 4. Conclusions

Based on the exergoeconomic analysis carried out in the simple ORC and ORC with HE, it is concluded that the unit exergoeconomic cost of the functional product of the system can be reduced by taking advantage of the thermal current that leaves the turbine through a HE. In addition, the exergoeconomic cost for the formation of the residue also reflects a decrease compared to the ORC, both configurations with the same exergoeconomic cost of the fuel associated with the external resource.

#### References

- CFE - Comisión Federal de Electricidad, 2005, Prontuario de la Central Termoeléctrica de Ciclo Híbrido del Valle de México, Mexico.
- C. T. Cuadra, A. V. Capilla, 2000, Termoeconomía, Universidad de Zaragoza, España.
- H. Lugo-Méndez, S. Castro-Hernández, M. Salazar-Pereyra, J. Valencia-López, E.V. Torres-González, R. Lugo-Leyte, 2020, Residue Cost Formation of a High Bypass Turbofan Engine, *Applied Sciences* 10 (24), 9060.
- M. Mohammadi, M.H. Ashouri, M. Ahmadi, M. Bidi, M. Sadeghzadeh, T. Ming, 2018, Thermoeconomic analysis and multiobjective optimization of a combined gas turbine, steam, and organic Rankine cycle, *Energy Science and Engineering*, 6 (5), 506–522.
- J. Nondy, T. K. Gogoi, 2021, Exergoeconomic investigation and multi-objective optimization of different ORC configurations or waste heat recovery, *Energy Conversion and Management*, 245, 114593.
- C. Torres, A. Valero, 2021, The exergy cost theory revisited, *Energies*, 14, 1594.
- C. Torres, A. Valero, L. Serra, J. Royo, 2002, Structural theory and thermoeconomic diagnosis: Part I. On malfunction and dysfunction analysis, *Energy Conversion and Management*, 43 (9–12), 1503–1518.

## Hybrid polymeric membrane – chemical absorption system for pre-combustion CO<sub>2</sub> capture

Maytham Alabid<sup>a</sup>, Nela Slavu<sup>a,b</sup>, Marius Sandru<sup>c</sup>, Cristian Dinca<sup>a,b</sup>

<sup>a</sup>Faculty of Energy, University Politehnica of Bucharest, Splaiul Independenței, 060042 Bucharest, Romania

<sup>b</sup>Academy of Romanian Scientists, Ilfov 3, 050044 Bucharest, Romania

<sup>c</sup>SINTEF Materials and Chemistry, Sem Saelands vei 2A, 7465, Trondheim, Norway

### Abstract

The purpose of the article is to develop CO<sub>2</sub> capture technologies to face the increasing decarbonization requirements of the industrial sector. In this context, air or O<sub>2</sub> have been used as oxidizing agents for gasification process with the usage of a hybrid configuration of polymer membranes and chemical absorption process that has been technically and economically analyzed. The role of polymer membranes is to pre-concentrate the CO<sub>2</sub> stream in order to separate it in the absorption column. The chemical absorption technology analyzed is based on the use of MEA 30 wt.% solvent. The proposed hybrid solution was integrated pre-combustion in a BIGCC system equipped with a combined cycle. The study modeled energy processes using CHEMCAD version 8.1 software and polymer membranes developed in the CO<sub>2</sub> Hybrid research project. The LCOE indicator demonstrated CO<sub>2</sub> capture to be remarkably higher when utilizing air at 142.5 €/MWh versus 86 €/MWh for the O<sub>2</sub> case.

**Keywords:** Polymeric membrane, Gasification, air separation, chemical absorption.

### 1. Introduction

Energy necessities (particularly electricity energy) are constantly increasing due to the raise in the world population and the evolution of living levels [1]. Presently, a coal-fired power plant is the most energy producer in the world and it is the main CO<sub>2</sub> emission that directly influences climate change and causes an increase in the atmosphere temperature [2]. In this context, different solutions have been studied to match energy demands with low/zero carbon dioxide emissions, such as utilizing biomass, considered a CO<sub>2</sub>-neutral fuel [3,4]. The primary benefit of biomass usage is that it utilizes and absorbs carbon dioxide through its outgrowth, similar to the amount generated during the combustion process [5]. In a combined cycle with integrated gasification, biomass is used efficiently to generate electricity (BIGCC technology). In such a cycle, any type of biomass can be harnessed [6]. Integrating carbon capture systems (CCS) in a combined cycle that used biomass to produce energy can be considered an optimistic technology to achieve carbon mitigation targets and reduce climate change issues [7]. The polymeric membrane is currently used with chemical absorption to capture carbon dioxide emissions with high purity [8-10]. The study aims to evaluate membrane and chemical absorption performance with the gasification process that used either air or oxygen as an oxidizing agent.

### 2. Methodology

In this paper, two cases have been considered as follows:

**Case A:** BIGCC based on air as a gasification agent with hybrid CO<sub>2</sub> capture technology using membrane and chemical absorption technology;

**Case B:** BIGCC based on O<sub>2</sub> as a gasification agent with CO<sub>2</sub> capture technology using membrane technology.

Both technologies will be compared with the variants without CO<sub>2</sub> capture technology.

### 2.1. Gasification process

In this study, air and oxygen have been both used as oxidizing agents, where in the case of O<sub>2</sub>, a compressor was harnessed in the line of O<sub>2</sub> inlet with 5 bar to increase the syngas pressure introduced into the membrane technology, the pressure being the maximum accepted for the membrane technology developed in the University Politehnica of Bucharest laboratory. The energy required to separate 1 ton of O<sub>2</sub> from N<sub>2</sub> is about 540 kJ/kg [11]. The biomass feed was a poplar bark with a LHV (lower heating value) of 18,532.8 kJ/kg, table (1) below demonstrates the elemental composition of the waste in weight concentration (wt.%) [12].

**Table 1.** Biomass waste base characteristics.

Biomass	Main composition [wt.%]						
	C	H	S	O	N	A	W
Poplar bark	48.02	6.00	0.09	35.21	0.27	2.02	8.40

The gasification process is recognized by the equivalent ratio (ER, in mol/mol) and cold gas ratio (CGR), represented in formulas (1,2). F<sub>1</sub>, in kg/h, demonstrates the real air/O<sub>2</sub> flow inserted into the gasification reactor, while F<sub>2</sub>, in kg/h, presents the stoichiometric air/O<sub>2</sub> stream needed for the complete combustion of each combustible substance.

$$ER = \frac{F_1}{F_2}, \quad (1)$$

$$CGR = \frac{S_1 \times LHV_1}{S_2 \times LHV_2} \times 100\%, \quad (2)$$

While S<sub>1</sub> is the syngas stream rate, in kg/h; LHV<sub>1</sub> the syngas lower heating value, in kJ/kg; S<sub>2</sub> the biomass stream rate, in kg/h; LHV<sub>2</sub> the biomass lower heating value, in kJ/kg.

### 2.2. Membrane process

In the current research, one stage of membrane with 3,000 GPU [10], a vacuum pump and a compressor has been used for capturing CO<sub>2</sub> emissions from syngas before its utilization in a combined cycle. Thus, membrane technology has the role of pre-concentrate CO<sub>2</sub> in the permeate stream. The membrane consists of 2 layers: a PSF 50 K, the support polymer, and the CA enzyme, the polymer solution [13]. The membrane module used was a spiral wound with a flow style of counter-current, the permeability of N<sub>2</sub> was assumed 20 GPU [8]. In case A, a compressor was needed to increase the syngas pressure in order to increase the membrane efficiency for CO<sub>2</sub> separation before the chemical absorption unit. However, in Case B, no compressor was used before the membrane unit due to the syngas pressure that comes from gasification process (5 bar). Therefore, the chemical absorption unit is not integrated due to the high purity, more than 99.9%.

### 2.3. Chemical absorption process

In the chemical absorption process (CAP) a chemical solvent is used for capturing the CO<sub>2</sub> from different gaseous streams in the absorption column before being regenerated in the stripper column [14,15]. CAP technology is only used in case A due to the low CO<sub>2</sub> content in the syngas. The chemical solvent utilized was monoethanolamine in a 30% weight concentration. To reduce the chemical solvent flow and heat required for its regeneration, the syngas stream was treated in the first step by membrane technology.

### 2.4. Combined cycle technology

The syngas obtained after the membrane process (the retentate stream) was compressed at 17 bar and introduced to the combined cycle for electricity generation. The steam obtained from the back-pressure steam turbine (5 bar) was used either to provide heat energy for solvent regeneration (case A) or to cover the WGR (water gas shift reactor) requirements for case B, see figure (1). In case A, the steam required for WGR was generated from HRSG (heat recovery steam generator), where 500 kg/h of natural gas was used.

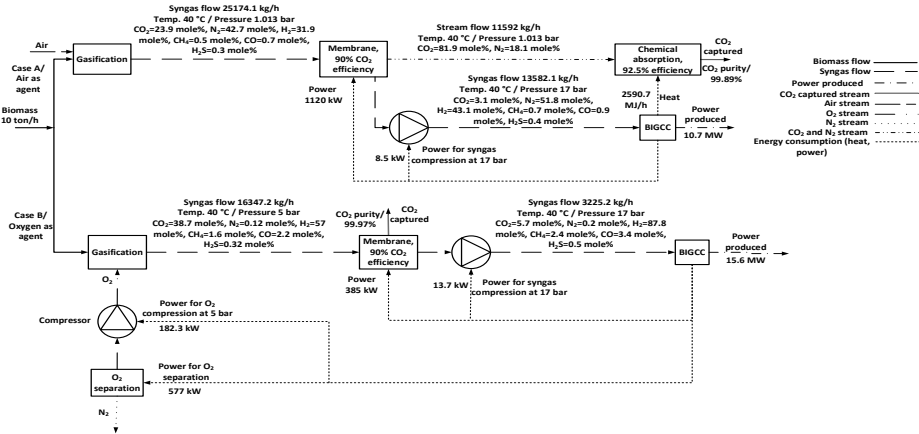


Figure 1. Scheme of pre-combustion carbon capture in BIGCC with main results.

### 2.5. Economic assumptions

Table (2) shows the main assumptions considered in the current research.

Table 2. Main assumptions for economic assessment [16].

Indicators	Main data
Factor of availability, %	85
Electricity cost, €/MWh	200
CO <sub>2</sub> fee, €/t	82 [17]
Working hours, h/year	75% of 8,760
Membrane module cost, €/m <sup>2</sup>	50
Life duration of membrane, years	5
Vacuum pump / CO <sub>2</sub> pump price, €/kW	1,300 / 1,350
CO <sub>2</sub> compressor price, €/kW	1,800
Replacement membrane price, €/m <sup>2</sup>	10
Cost of Labor, €/h	15
Absorber / Stripper part, M€	27.7 / 31.4
Pump of rich-lean amine with heat exchanger, M€	6.89
Gas / steam turbine cost, M€	6 / 3
Condenser / Gasification unit cost, M€	39 / 45 [18]
Water gas shift / Ash treatment cost, M€	21 / 16 [18]
Heat recovery steam generator, M€	3 [18]
CO <sub>2</sub> compression compressor, M€	11.7
Compressor inter-stage coolers and separators, M€	0.87

Equation (3) shows the total CO<sub>2</sub> emissions, while equation (4) presents CO<sub>2</sub> emission factor, FCO<sub>2</sub>, for cases with and without CO<sub>2</sub> capture (in kg/MWh).

$$CO_{2wc} = CO_{2CG} - CO_{2A} - CO_{2BA} + CO_{2NG}, \quad (3)$$

Where CO<sub>2wc</sub> is the total CO<sub>2</sub> emissions; CO<sub>2CG</sub> is the CO<sub>2</sub> generated in the gasification and combustion process; CO<sub>2A</sub> is CO<sub>2</sub> emissions captured by membrane and CAP; CO<sub>2BA</sub> is CO<sub>2</sub> emissions required for biomass growth; CO<sub>2NG</sub> is CO<sub>2</sub> produced from natural gas usage (case A), all parameters are in kg/h. When no capture technology integrated, CO<sub>2woutc</sub>, CO<sub>2A</sub> is zero. The CO<sub>2BA</sub> is the sum of CO<sub>2</sub> generated in the gasification

process and CO<sub>2</sub> generated in the combustion process according to CO that exists in syngas. Regarding the value used, CO<sub>2</sub> in syngas is 23.9 mole% of the syngas flow, and the CO<sub>2</sub> produced from CO exits in the syngas is 276.7 kg/h (as it is in figure 1), therefore, CO<sub>2</sub> generated after syngas and combustion process is 6,293.3 kg/h, which represents the amount of CO<sub>2</sub> absorbed during biomass growth (photosynthesis process). Considering that the biomass is carbon neutral, all the CO<sub>2</sub> emissions generated by the carbon from the biomass composition were considered neutral as well.

The amount of CO<sub>2</sub> absorbed by biomass growth is lower compared to the CO<sub>2</sub> generated from gasification and combustion process as a result of CO<sub>2</sub> emissions from the preparation and transport of biomass to the power plant which have not been considered to simplify the research.

$$FCO_2 = \frac{CO_2 \text{ wc or } CO_2 \text{ woutc}}{\text{Total electricity generation}}, \quad (4)$$

To evaluate the techno-economic effect of CO<sub>2</sub> removal process, the coming factors have been suggested [19].

Levelised cost of electricity (LCOE), in €/MWh, can be computed with the relation (5).

$$LCOE = \frac{CAPEX+OPEX}{W_{net}}, \quad (5)$$

$W_{net}$  is the fine electric energy produced (MWh) by each technology with and without CO<sub>2</sub> capture; CAPEX represents capital expenditure while OPEX indicates operational expenditure of the project, (in €).

Net present value (NPV) was computed as follows, (in €).

$$NPV = \sum_{i=1}^{n_f} \frac{IN_i - C_i - A_i}{(1+r)^i} - \sum_{i=1}^{n_r} I_i \cdot (1+r)^i, \quad (6)$$

$IN_i$  is the real bonus (€/year);  $C_i$  the price of operation and maintenance (€/year);  $A_i$  rate of payback loan (€/year);  $I_i$  real investment cost (€/year),  $r$  is the rate of discount, which is 8% in energy section.

Discounted payback period (DPP) was computed by formula (7), (in years).

$$NPV = \sum_{i=1}^{DPP} \frac{IN_i - C_i - I_i}{(1+r)^i}, \quad (7)$$

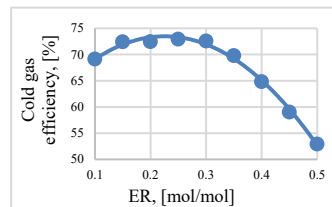
The profitability index (IP) was founded by the next equation, where IA is the deduct investment.

$$IP = \frac{NPV+IA}{IA} \quad (8)$$

### 3. Results and Discussions

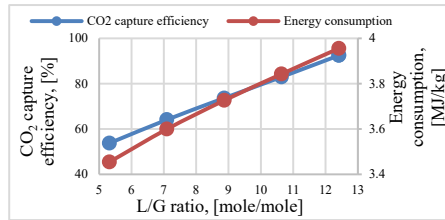
The gasification and capture systems were simulated through CHEMCAD software of 8.1 version while the combined cycle was modeled in excel. In the gasification process, the equivalent ratio was considered 25% for both air and O<sub>2</sub>. As an example, in figure (2) the variation of the CGE according to ER where O<sub>2</sub> was used as an oxidizing agent. Membrane unit either capture most of CO<sub>2</sub> content in syngas with lower purity and send the recovered flow (mainly CO<sub>2</sub>

and N<sub>2</sub>) to the chemical absorption (case A) or capture the CO<sub>2</sub> stream at 99.97% purity (case B) due to the low N<sub>2</sub> content in the syngas obtained from gasification process. Chemical absorption usage in case A helps to raise the concentration of CO<sub>2</sub> removal from membrane at a specific capture rate. Figures (3, 4) show that CO<sub>2</sub> capture efficiency of chemical absorption continually rise with a higher L/G ratio, where an extra chemical solvent is inserted into the process. Taking into account the low CO<sub>2</sub> content of the syngas

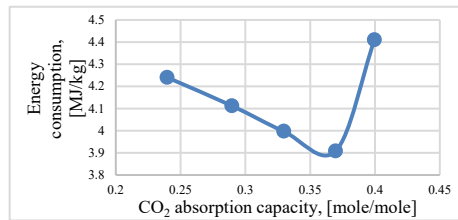


**Figure 2.** CGE variation according to ER for O<sub>2</sub> utilization.

stream, the L/G ratio is greater than 12 kg solvent/kg syngas for a minimum capture efficiency of 90%. The thermal energy consumption required to regenerate the chemical solvent was about 3.9 MJ/kg for a CO<sub>2</sub> absorption capacity of 0.36 mole CO<sub>2</sub>/mole solvent. The CO<sub>2</sub> emissions for biomass growth in case A are slightly lower than emissions of case B because of the difference in hydrocarbon concentrations in the syngas of each case. The CO<sub>2</sub> emissions generated from natural gas usage were not treated in CCS, therefore, CO<sub>2</sub> emission factor for case A is higher than the other case. The global efficiency in case A reduced by CCS utilization, where both electrical and thermal energy was consumed for the membrane and solvent regeneration process. On the other hand, the global efficiency in case B wasn't highly influenced by using CCS due to the low power required for the membrane vacuum pump (385 kW). Table (3) shows the optimum results of each case simulated.



**Figure 3.** Effect of L/G ratio on CO<sub>2</sub> capture efficiency and energy consumption.



**Figure 4.** CO<sub>2</sub> absorption capacity influence on energy consumption.

**Table 3.** The essential results regarding the four cases examined.

Cases	Case A		Case B	
	No CO <sub>2</sub> capture	With CO <sub>2</sub> capture	No CO <sub>2</sub> capture	With CO <sub>2</sub> capture
Biomass flow, ton/h	10	10	10	10
Heat flow with syngas, kW	24,979	24,979	8,698	8,698
CO <sub>2</sub> emissions for biomass growth, kg/h	6,293.3		6,891	
CO <sub>2</sub> emissions from natural gas usage, kg/h	No use	1,380.3	No use	No use
Total CO <sub>2</sub> emissions, kg/h	6,293.3	2,009.6	6,891	689.1
Membrane surface area, m <sup>2</sup>	-	25,000	-	2,500
Compressor pressure, bar	-	2.2	-	No use
Vacuum pump pressure, bar	-	0.25	-	0.25
L/G ratio for Chemical absorption, mole/mole	-	12.4	-	No use
Solvent flow, kmol/h	-	350	-	No use
Energy consumption for chemical absorption, MJ/kg	-	3.9	-	No use
Gas turbine power, kW	8,968.2	8,968.2	12,234.4	12,234.4
Steam turbine power, kW	2,896.2	2,896.2	4,559.1	4,559.1
Global efficiency, %	42.0	38.0	42.3	41.3
CO <sub>2</sub> factor, kg/MWh	-	-435.5	-	-877.4
NPV, M€	9,152	67,411	32,470	124,830
DPP, years	17.73	8.17	8.07	4.5
Profitability index	1.18	2.15	1.62	3.87
LCOE, €/MWh	157	142.5	135.3	86.1

In case A, the gasification process occurred with atmospheric pressure, so no electricity was required to compress the stoichiometric air stream. The syngas produced has a high flow rate, therefore, a compressor was integrated before the membrane stage to raise the efficiency of CO<sub>2</sub> removal and that increased the electricity needed. Despite the fact that CAP helped raise the purity of CO<sub>2</sub> recovery, it increased CAPEX cost and mitigated the total combined cycle power generation due to the heat energy required. On the other hand, the usage of O<sub>2</sub> as an oxidizing agent has the advantage of no CAP utilization in the process due to the high purity (99.9%) achieved after the membrane stage, which enhanced the reduction of CAPEX cost in case B. In contrast to case A, electrical energy



was needed for O<sub>2</sub> separation which can be considered a power consumption from the combined cycle. However, the integration of a 5-bar compressor in the O<sub>2</sub> stream (in gasification) generated the syngas at high pressure where no need to provide high electricity for the membrane stage as in the air case. Consequently, the electricity produced in case B is greater than that produced in case A.

### Acknowledgments

The research has been granted by the UEFISCDI within the International Project NO Grants 2014–2021, under project contract no. 13/2020.

### References

1. Gaye A , 2007/ 2008, Human development report, fighting climate change: human solidarity in a divided world. Access to Energy and Human Development.
2. IEA. World energy outlook. Paris: International Energy Agency; 2020.
3. Slavu, N., Badea, A. and Dinca, C., 2022. Technical and Economical Assessment of CO<sub>2</sub> Capture-Based Ammonia Aqueous. Processes, 10(5), p.859.
4. Slavu, N., Dinca, C. and Banica, C.K., 2021, October. Production of H<sub>2</sub>-rich Syngas from Biomass Gasification with CO<sub>2</sub> Capture Technology. In 2021 10th International Conference on Energy and Environment (CIEM) (pp. 1-5). IEEE.
5. Dinca, C.R.I.S.T.I.A.N., Badea, A.D.R.I.A.N., Marculescu, C.O.S.M.I.N. and Gheorghe, C.O.R.A., 2009. Environmental analysis of biomass combustion process. In Proceedings of the 3rd WSEAS Int. Conf. on Renewable Energy Sources (pp. 234-238).
6. Cormos, A.M., Dinca, C., Petrescu, L., Chisalita, D.A., Szima, S. and Cormos, C.C., 2018. Carbon capture and utilisation technologies applied to energy conversion systems and other energy-intensive industrial applications. Fuel, 211, pp.883-890.
7. Rhodes, J.S. and Keith, D.W., 2005. Engineering economic analysis of biomass IGCC with carbon capture and storage. Biomass and Bioenergy, 29(6), pp.440-450.
8. Cormos, C.C., Petrescu, L., Cormos, A.M. and Dinca, C., 2021. Assessment of Hybrid Solvent—Membrane Configurations for Post-Combustion CO<sub>2</sub> Capture for Super-Critical Power Plants. Energies, 14(16), p.5017.
9. Alabid, M., Cormos, C.C. and Dinca, C., 2022. Critical Assessment of Membrane Technology Integration in a Coal-Fired Power Plant. Membranes, 12(9), p.904.
10. Sandru, M., Sandru, E.M., Ingram, W.F., Deng, J., Stenstad, P.M., Deng, L. and Spontak, R.J., 2022. An integrated materials approach to ultrapermeable and ultrasensitive CO<sub>2</sub> polymer membranes. Science, 376(6588), pp.90-94.
11. Banaszkiwicz, T., Chorowski, M. and Gizicki, W., 2014. Comparative analysis of oxygen production for oxy-combustion application. Energy Procedia, 51, pp.127-134.
12. Vassilev, S.V., Baxter, D., Andersen, L.K. and Vassileva, C.G., 2010. An overview of the chemical composition of biomass. Fuel, 89(5), pp.913-933.
13. CO<sub>2</sub> hybrid project [online website] [http://co2hybrid.upb.ro/index\\_ro.html](http://co2hybrid.upb.ro/index_ro.html).
14. Wang, M., Laval, A., Stephenson, P., Sidders, J. and Ramshaw, C., 2011. Post-combustion CO<sub>2</sub> capture with chemical absorption: A state-of-the-art review. Chemical engineering research and design, 89(9), pp.1609-1624.
15. Koronaki, I.P., Prentza, L. and Papaefthimiou, V., 2015. Modeling of CO<sub>2</sub> capture via chemical absorption processes— An extensive literature review. Renewable and Sustainable Energy Reviews, 50, pp.547-566.
16. Chemengonline. [www.chemengonline.com](http://www.chemengonline.com), 2022. (Accessed 01 June 2022). <https://www.chemengonline.com/pci-home>
17. Trading economics [online website] <https://tradingeconomics.com/commodity/carbon> (Accessed 29 November 2022).
18. Kramer, S., 2003. Gasification plant cost and performance optimization. Nexant Inc.(US).
19. Cormos, C.C. and Dinca, C., 2021. Techno-economic and environmental implications of decarbonization process applied for Romanian fossil-based power generation sector. Energy, 220, p.119734.

# Electrification of ethylene production: exploring the potential of flexible operation

Julia L. Tiggeloven,<sup>a</sup> André P.C. Faaij,<sup>a,b</sup> Gert Jan Kramer<sup>a</sup> and Matteo Gazzani<sup>\*a</sup>

<sup>a</sup> *Copernicus Institute of Sustainable Development, Utrecht University, The Netherlands*

<sup>b</sup> *TNO Energy Transition, the Netherlands*

*\* Corresponding author*

## Abstract

Electrification of the ethylene production process has potential for reducing hard-to-abate CO<sub>2</sub> emissions, however, this comes at the expense of increased electricity consumption. Flexible operation of electric naphtha crackers in response to fluctuating electricity prices can compensate for the increased operating costs, depending on the boundary conditions. However, given the level of technological complexity, it remains uncertain to what extent flexible operation of the ethylene production process is feasible. In this work we use mathematical optimization to investigate the added value of various levels of flexible operation of the process in a low-CO<sub>2</sub> energy system. We find that (i) highly variable electricity prices promote flexible operation and (ii) providing flexibility in low-CO<sub>2</sub> energy systems by ramping the production or using electricity storage is essential.

**Keywords:** electrification, flexibility, optimization.

## 1. Introduction

Electrification of naphtha steam cracking is a potential alternative to reduce emissions in the ethylene production process. Accordingly, the electricity consumption and operating cost of the process are expected to increase significantly, leading to a new balance between investment and operating costs. Due to the high investment costs and complexity of conventional steam crackers, plants are operated at maximum capacity. However, this prevents exploiting the large variation in electricity price in time, especially when and where undispatchable renewables are substantially installed. Flexible operation of electric crackers in response to electricity prices might be able to compensate for the increased operating costs, opening up new opportunities and challenges in the chemical industry. This study aims at further investigating this topic by optimizing the full-year, hourly operation and system design of a simplified ethylene production process, where either conventional or electric steam cracking of naphtha is used to meet an assumed ethylene demand. More specifically, this work uses optimization to assess the system behavior under multiple scenarios, modifying hourly-resolved electricity price profiles, carbon intensity of electricity, and the extent of operational flexibility of the process. Section 2 describes the modeling framework and the case study input data. Section 3 shows the results.

## 2. Method

### 2.1. Modeling framework

The mathematical problem formulation in this work is derived from an existing optimization framework developed by Gabrielli et al. [1], [2]. The decision variables that

are optimized by the framework include design variables (i.e. selection and size of technologies) and operational variables (i.e. energy and material flows and storage levels). The framework uses hourly resolved input data on weather conditions, prices and demand data together with a set of available technologies and the corresponding cost and performance coefficients. The objective function of the problem is the total annualized system cost,  $J$ , that is the sum of the technology cost,  $J_c$ , and the operating cost,  $J_o$ . The annualized technology cost is defined as

$$J_c = \sum_{i \in \mathcal{M}} (1 + \psi_i)(\lambda_i S_i + \zeta_i) a_i$$

where  $\lambda_i$  and  $\zeta_i$  represent the size-dependent and size-independent cost parameters of technology  $i$  and the annuity factor  $a_i$  is used to compute the annualized capital costs. The maintenance cost of the technology is included as fraction of the annual capital costs  $\psi_i$ . The operating cost of the system is determined by the annual amount of electricity import, which is expressed as

$$J_o = \sum_{t=1}^T (u_{e,t} U_{e,t})$$

where  $U_{e,t}$  is the electricity import and  $u_{e,t}$  the hourly electricity price at hour  $t$ . The total emissions include indirect emissions from imported electricity and direct emissions from the ethylene production process and are calculated as

$$e_{CO_2} = \sum_{t=1}^T \left[ \varepsilon_{elec} U_{elec,t} + \sum_{i \in \mathcal{M}} \epsilon_i F_{naphtha,i,t} \right]$$

where  $\varepsilon_e$  is the CO<sub>2</sub> intensity of the electricity grid and  $\epsilon_i$  is the emission factor of the technology per unit of input.

For each hour  $t \in T$  and each carrier  $j \in \mathcal{N}$ , the material and energy balances are formulated as:

$$\sum_{i \in \mathcal{M}} (U_{j,i,t} + P_{j,i,t} - V_{j,i,t} - F_{j,i,t}) - D_{j,t} = 0$$

where  $U_{j,i,t}$  is import,  $P_{j,i,t}$  is the production,  $V_{j,i,t}$  is export,  $F_{j,i,t}$  is consumption and  $D_{j,t}$  is the demand. The conversion of the ethylene production plant is described as

$$P_{j',t} = \alpha_{j'} F_{naphtha,t}$$

and

$$F_{j'',t} = \beta_{j''} F_{naphtha,t}$$

where  $\alpha_{j'}$  and  $\beta_{j''}$  represent the conversion efficiency per kg of naphtha input of the outputs  $j' \in J' \subset J$  and the ratio between inputs  $j'' \in J'' \subset J$ , respectively. Inputs and outputs of the process include naphtha, methane, electricity and steam. The operation range of the plant is expressed as

$$\gamma_{naphtha} S_{plant} \leq F_{naphtha,t} \leq S_{plant}$$

where  $\gamma_{naphtha}$  is the minimum feasible operating point of the plant as fraction of the installed capacity. For a detailed description of the modeling of storage, PV and offshore wind turbines the reader is referred to Gabrielli et al. [2] and to Weimann et al. [3].

## 2.2. System description

The ethylene production process consist of three main components: pyrolysis (i.e. the cracker), compression and separation. In the cracker furnace, naphtha is preheated and vaporized with superheated steam before it is cracked into light olefins (i.e. ethylene) in the gaseous state. In the conventional cracker the required process heat is supplied by the

combustion of a mixture of recovered process gasses, which in this work is assumed to consist of only methane. The exact gas composition depends on the cracker design and on the overall process type; a typical composition is 18% H<sub>2</sub>, 81% CH<sub>4</sub> and 1% residues (i.e. CO and C<sub>2</sub>H<sub>4</sub>) [4]. For electric cracking we assume that indirect resistance heating is used. In both cases the heat in the cracked gas leaving the furnace is recovered by producing high-pressure steam. In the conventional plant design, the high-pressure steam is superheated by hot flue gas in the furnace and used to drive the compressors. As the electric cracker does not produce hot flue gas and therefore no superheated steam, it is assumed that electric compressors are used. After compression and further cooling and drying, the compressed gas enters the separation train where the different products are separated through distillation, refrigeration and extraction.

The modeling framework is used to optimize the design and operation of the ethylene production process for different scenarios. In all scenarios we consider a constant hourly ethylene demand of around 1000 kt ethylene/year, which is typical for ethylene production plants [5]. In the case of the electric cracker, the operation of the plant can be ramped up and down and ethylene can be stored in tanks to maintain the production stable.

### 2.3. Input data

The input and output performance parameters of the conventional and electric ethylene production process are obtained from [5], [6] and shown in Table 1.

Table 1: Conversion parameters of the conventional and electric ethylene production process.

	$\beta_{\text{methane}}$ [kWh]	$\beta_{\text{elec}}$ [kWh]	$\beta_{\text{steam}}$ [kWh]	$\alpha_{\text{ethylene}}$ [kg]	$\alpha_{\text{methane}}$ [kWh]	$\alpha_{\text{steam}}$ [kWh]	$\epsilon_i$
<b>Conventional plant</b>	1.868	0.083	0.923	0.303	2.349	0.841	0.600
<b>Electric plant</b>	-	1.942	0.492	0.303	2.349	0.542	-

The cost parameters  $\lambda_i$ ,  $\zeta_i$  and  $\psi_i$  of all technologies are shown in Table 2.

Table 2: Cost data of all technologies.

	Type	Unit	$\lambda_i$ [€/unit]	$\zeta_i$ [M€]	$\psi_i$ [%]	Source
<b>Conventional plant</b>	C <sub>2</sub> H <sub>4</sub> production	kg <sub>C<sub>2</sub>H<sub>4</sub></sub>	4531	350	0.04	[5]
<b>Electric plant</b>	C <sub>2</sub> H <sub>4</sub> production	kg <sub>C<sub>2</sub>H<sub>4</sub></sub>	4433	358	0.02	[5]
<b>Ethylene tank</b>	Storage	kg <sub>C<sub>2</sub>H<sub>4</sub></sub>	18.24	-	0.02	[7]
<b>Li-ion battery</b>	Storage	kWh	128.8	-	0.01	[8]
<b>9.5 MW turbine</b>	Offshore wind	turbine	17.8 · 10 <sup>6</sup>	-	0.02	[9]
<b>Ground mounted PV</b>	PV	m <sup>2</sup>	140.9	-	0.01	[9]

The hourly resolved solar irradiations and wind speeds are obtained from the Dutch meteorology institute (KNMI) and the Dutch Offshore Wind Atlas (DOWA), respectively. It is assumed that the plant is located in the Netherlands near the port of Rotterdam to accommodate a direct connection to offshore wind parks. Hourly resolved spot electricity prices are obtained from ENTSO-E for the year 2019 in the Netherlands.

## 3. Results

In the first analysis we determine the CO<sub>2</sub> reduction potential of electrification of the conventional process and analyze the new balance between investment and operating costs that result from electrification. Next, we investigate the potential of flexibility in electric cracking and its sensitivity towards different economic input parameters. Finally,

we assess the effect of the minimum feasible operating range on the optimal system design and operation, including when batteries and renewables are adopted along with the electric cracker. An overview of the investigated cases is shown in Table 3.

Table 3: Available technologies in the four investigated cases.

	Technologies available	Electricity import
Reference	Conventional cracker, Ethylene tank	Yes
Case A	Electric cracker, Ethylene tank	Yes
Case B	Electric cracker, Ethylene tank, Li-ion battery	Yes
Case C	Electric cracker, Ethylene tank, Li-ion battery, offshore wind, PV	No

### 3.1. Environmental potential of electrification

The performance of the electric cracker in terms of total direct and indirect CO<sub>2</sub> emissions is compared to that of the conventional cracker in four scenarios in which electricity is bought from the grid with different projected CO<sub>2</sub> intensities (Table 4). In this analysis we only include emissions related to the heat provision of the process, while the CO<sub>2</sub> content derived from the product end-of-life is ignored.

Table 4: Overview of the different projected CO<sub>2</sub> intensities and RES production rates of the four scenarios with the corresponding emission reduction potential.

Year	Carbon intensity [kg CO <sub>2</sub> /kWh]	Share of RES production [%]	Reduction potential for electric [%]
2019	0.369 [10]	22.0	0
2025	0.208 [10]	54.7	36.1
2030	0.094 [10]	72.0	71.1
2050	0	100	100

Not surprisingly, the results show that the emission reduction potential of electrification is strongly determined by the carbon intensity of electricity. For a carbon intensity of 0.369 kgCO<sub>2</sub>/kWh, which is similar to today's value, there is no CO<sub>2</sub> reduction potential of electrification. Based on the projected carbon intensity, a 71% reduction of the total energy related emissions can be achieved in 2030. This potential is relatively small considering the investments, however, it is worth noting that cracker emissions are hard to abate otherwise.

### 3.2. Added value of flexible operation

Based on literature data, the difference in investment costs between a conventional and an electric ethylene production plant is around 2%. However, considering the spot electricity prices of 2019, the import costs of electric cracking account for 73.5% of the total annual costs, compared to 10.5% for the conventional plant. Although the cost parameters are highly uncertain, it is unlikely that the conventional plant would be more expensive than the electric alternative, unless high carbon prices are present. Moreover, the low share of electricity costs of the conventional plant makes flexible operation in response to electricity prices irrelevant.

As opposed to conventional ethylene production, the large share of electricity costs in electric steam cracking can lead to significant cost savings when plants are operated flexibly in response to electricity price. However, adapting the plant to flexible operation requires a more challenging cracker design, ethylene storage and production capacity to shift the production in time, thus increasing the investment costs. We analyze this trade-off under different economic boundary conditions by comparing the steady state operation design with a system design where flexibility in the operation is optimized.

More specifically, we change the value of the (i) size-dependent and (ii) size-independent cost parameters of the plant and the (iii) mean and (iv) standard deviation of the electricity price and evaluate how they affect the value of flexibility (VOF) of the system. The VOF is here defined as

$$VOF = J^{Steady\ state} - J^{Flexible}$$

where  $J^{Steady\ state}$  is the total annual cost of the steady state operation design and  $J^{Flexible}$  is the total annual cost for the optimal design where flexibility is allowed. The result of this sensitivity analysis is shown in Figure 1 and indicate that an electricity price with a high mean and standard deviation results in a high VOF, as well as a low value of the size dependent investment costs parameter. Furthermore, the slope of the dashed green line, representing the VOF for a high mean (+50%), implies that the effect of variability in the electricity price on the VOF is greater for a high mean.

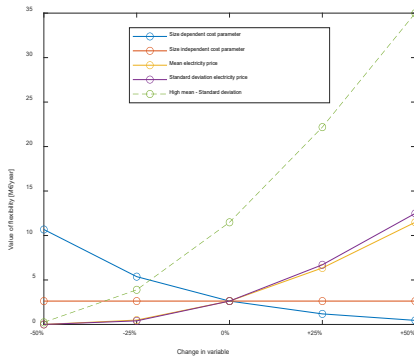


Figure 1: Result of the sensitivity analysis for the electricity price and investment cost parameters.

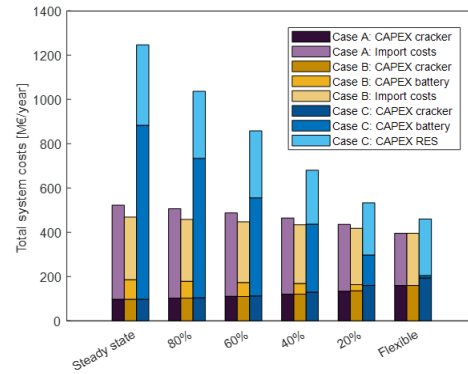


Figure 2: Total system costs for optimization of Case A, B & C considering different operational ranges.

### 3.3. Dynamic operation constraints

The previous section shows the optimal system design for a system in which the operational limits of the cracker, such as operational range, are not constrained. This is however unrealistic. As the feasible operational range of electric crackers is uncertain, we assess the effect of the minimum load of the electric steam cracker on the optimal design and operation and the total costs for the three scenarios in Table 4. The aim is to (i) compare different levels of flexible cracker operation with using a battery for flexibility and (ii) analyze how the system performs when grid import is not allowed and the production process is directly integrated with RES. The total system costs of these cases for an electricity price profile with a high mean and variability is shown in Figure 2.

The results show that the investment costs  $J_c$  increase and the total system costs  $J$  decrease with flexibility, caused by the substantial electricity cost savings in cases A and B. In case C, the savings are a result of the smaller required RES and battery capacity. For example, operating the cracker in case C with a minimum feasible operating point ( $\gamma_{naphta}$ ) of 80% increases the cracker size by 2.21% compared to steady state operation of case C, but decreases the required battery and offshore wind capacity by 19.8% and 16.8%, respectively, leading to an overall cost decrease of 16.8%. Allowing the cracker in case A to operate with a minimum operating point of 20% of the installed capacity can lead to an overall cost reduction of 16.5% compare to steady state operation without battery. Including a battery to a steady state cracker operation (Case A vs. Case B) decreases the

costs by 10.3%, however, it requires a battery capacity of around 4 GWh. Steady state cracker operation for case C requires a battery capacity of 10 GWh, which is 40 times larger than the largest current battery installed and has a volume of 85000 m<sup>3</sup> [8].

#### 4. Conclusion

In this work we used mathematical optimization to investigate the added value of electrification and flexible operation of the ethylene production process in a low-CO<sub>2</sub> energy system. We quantified the relation between CO<sub>2</sub> reduction potential and the CO<sub>2</sub> intensity of the grid, showing that electrification of the ethylene production process only makes sense for a fully decarbonized electricity grid. However, the high share of low-CO<sub>2</sub> electricity will lead to a more variable production of electricity, which will be reflected in the price. Our results indicate that operating an electric cracker flexibly in response to these variable prices can lead to significant costs savings compared to steady state operation. Moreover, battery storage shows to be an economically viable alternative or complementary to provide flexibility in the production process. However, the storage capacity that is required in the investigated scenarios result in a significantly large volume, especially when directly integrating the process with offshore wind and PV. We conclude that system optimization to identify the optimal combination and sizes of technologies and flexible operation of the plant is essential to enable the full decarbonization of the ethylene production process.

#### References

- [1] P. Gabrielli, M. Gazzani, and M. Mazzotti, "Electrochemical conversion technologies for optimal design of decentralized multi-energy systems: Modeling framework and technology assessment," *Appl Energy*, vol. 221, pp. 557–575, Jul. 2018, doi: 10.1016/J.APENERGY.2018.03.149.
- [2] P. Gabrielli, M. Gazzani, E. Martelli, and M. Mazzotti, "Optimal design of multi-energy systems with seasonal storage," *Appl Energy*, vol. 219, pp. 408–424, Jun. 2018, doi: 10.1016/J.APENERGY.2017.07.142.
- [3] L. Weimann, P. Gabrielli, A. Boldrini, G. J. Kramer, and M. Gazzani, "Optimal hydrogen production in a wind-dominated zero-emission energy system," *Advances in Applied Energy*, vol. 3, p. 100032, Aug. 2021, doi: 10.1016/J.ADAPEN.2021.100032.
- [4] V. Spallina, I. C. Velarde, J. A. M. Jimenez, H. R. Godini, F. Gallucci, and M. van Sint Annaland, "Techno-economic assessment of different routes for olefins production through the oxidative coupling of methane (OCM): Advances in benchmark technologies," *Energy Convers Manag*, vol. 154, pp. 244–261, Dec. 2017, doi: 10.1016/J.ENCONMAN.2017.10.061.
- [5] H. Zimmerman and R. Walzi, "Ullman's Encyclopedia of Industrial Chemistry," *Wiley Online*, vol. 15, pp. 9–17, 2007.
- [6] T. Ren, M. Patel, and K. Blok, "Olefins from conventional and heavy feedstocks: Energy use in steam cracking and alternative processes," *Energy*, vol. 31, no. 4, pp. 425–451, 2006, doi: 10.1016/J.ENERGY.2005.04.001.
- [7] "Analysis Synthesis and Design of Chemical Processes 5th Edition | Richard Turton | West Virginia University." <https://richardturton.faculty.wvu.edu/publications/analysis-synthesis-and-design-of-chemical-processes-5th-edition> (accessed Jun. 10, 2022).
- [8] DEA, "Technology Data-Energy storage," 2022. Accessed: Nov. 14, 2022. [Online]. Available: <http://www.ens.dk/teknologikatalog>
- [9] DEA, "Technology Data-Energy Plants for Electricity and District heating generation," 2022. Accessed: Nov. 14, 2022. [Online]. Available: <http://www.ens.dk/teknologikatalog>
- [10] P. Planbureau voor de Leefomgeving, "Klimaat- en Energieverkenning 2021," 2021. Accessed: Dec. 02, 2022. [Online]. Available: [www.pbl.nl/kev](http://www.pbl.nl/kev)

# An MINLP model for the optimal design of CO<sub>2</sub> transportation infrastructure in industrial clusters

Jude O. Ejeh,<sup>a</sup> Sergey B. Martynov,<sup>b</sup> and Solomon F. Brown,<sup>a\*</sup>

<sup>a</sup>*Department of Chemical & Biological Engineering, The University of Sheffield, Mappin Street, Sheffield, S1 3JD, United Kingdom*

<sup>b</sup>*Department of Chemical Engineering, University College London, London, WC1E 7JE, United Kingdom*

\**s.f.brown@sheffield.ac.uk*

## Abstract

Carbon capture, utilisation and storage (CCUS) is still the most promising decarbonisation route for carbon-intensive sectors. In the context of industrial clusters, although benefits exist from economies of scale and shared transportation infrastructure, challenges remain relating to planning and investing in large-scale implementations. This involves the optimal sizing of the transportation infrastructure to avoid underestimating the transported amounts that could lead to stranded emitters or overestimating the capacity that would translate to financial losses due to unjustified capital expenditures. In this work, a mixed integer non-linear programming (MINLP) model for the design of carbon dioxide (CO<sub>2</sub>) transportation infrastructure in industrial clusters is proposed. The optimisation model considers a set of CO<sub>2</sub> emitters, available potential CO<sub>2</sub> transport and storage/utilisation options, etc. to obtain the detailed infrastructure design. The model assumes a transportation mode solely by pipelines, with steady-state fluid flow throughout the period of consideration. With an overall (capital and operating) cost minimisation objective, globally optimal network design results are obtained for candidate emitters and network routes within industrial clusters subject to constraints relating to storage balances, nominal pipe selection, and pipeline/fluid flow equations.

**Keywords:** CCUS industrial clusters; CO<sub>2</sub> transport infrastructure; Optimisation; MINLP.

## 1. Background

Carbon capture, utilisation and storage (CCUS) remains a promising pathway for decarbonising carbon-intensive industrial sectors and bridging the technology gap for a future net-zero society (Bui et al., 2018; Hasan et al., 2022). This technology consists of key elements such as the carbon dioxide (CO<sub>2</sub>) sources, capture technologies, conditioning and re-conditioning (compression, purification, etc.), transportation, utilisation, and storage technology and infrastructure.

Capture technologies include all chemical and physicochemical processes that take as input flue gases from emitting sources to produce a CO<sub>2</sub>-rich stream. Examples are all pre-, post- and oxy-fuel combustion processes, the most popular being post-combustion capture using Monoethanolamine (MEA) as a solvent. Conditioning technologies prepare the CO<sub>2</sub>-rich stream for a suitable mode of transport - via pipelines, ships, motor and/or rail. Amongst available modes of transport, pipelines are still the most popular and widely



researched option owing to their cost-effectiveness over long distances, larger scales and a wide range of conditions (Ejeh et al., 2022). CO<sub>2</sub> captured from emitting sources, if not utilised on-site, may be transported to other points of utilisation and/or storage sites for sequestration. Each of these CCUS elements has received (and still receives) a considerable amount of research interest, and significant progress has been made over the past years. However, barriers still exist in both small- and large-scale deployments. One significant barrier to the large-scale deployment of CCUS technologies is cost (Hasan et al., 2022). Although in the context of industrial clusters, benefits may be realised from economies of scale and shared capture and/or transport infrastructure, challenges still exist in the optimal planning and investment decisions for large-scale projects (Ejeh et al., 2022). CO<sub>2</sub> transportation has been known to take a significant portion of overall costs in relation to other elements of the CCUS supply chain. Hence, substantial research efforts have focused on the optimal planning and design of CO<sub>2</sub> transport infrastructure.

Optimal pipeline transport infrastructure design goes beyond source-sink matching and involves determining the pipeline size (inner/outer diameters), pipeline material, pipeline network route, specifications, target transportation conditions, location, and size of compression/ purification stations (D'Amore et al., 2021). Pipeline diameters are selected from industrial nominal pipe options based on proven correlations available in the literature (Mechleri et al., 2017), with the most popular model objective being the total capital and operating costs for the design. A widely adopted optimisation model for CO<sub>2</sub> transportation infrastructure is as described by Zhang et al. (2018) who used the cost model proposed in Serpa et al. (2011). The latter authors adopted a piecewise linear approximation of CO<sub>2</sub> cost estimates combined with other publicly available estimates from large natural gas projects. Similar approaches have been adopted by several authors as outlined in Ejeh et al. (2022) but these estimates have been noted to grossly underestimate actual cost values for CO<sub>2</sub> transport owing to the difference in fluid properties (Kim et al., 2018). Although it may become more difficult to solve, it then becomes beneficial to adopt more rigorous models that consider unique CO<sub>2</sub> flow properties in costing transport infrastructure designs as adopted, for example, by Mechleri et al. (2017).

An optimisation model for CO<sub>2</sub> transportation infrastructure design via pipelines is thus proposed that determines the optimal pipeline size and material from industrial nominal sizes for target transportation conditions. It builds on the mixed integer non-linear programming (MINLP) model proposed by Mechleri et al. (2017) that considers rigorous hydraulic equations for CO<sub>2</sub> flow in pipelines to minimise the total capital and operating costs of the infrastructure. Our model determines the globally optimal cost for a network of emitters within industrial clusters whilst accounting for utilisation amongst cluster members.

In the rest of the paper, the optimisation problem is clearly described in Section 2, and our proposed MINLP model is presented in Section 3. The model is then applied to a case study in Section 4, with findings discussed and some conclusions drawn in Section 5.

## 2. Problem description

Figure 1 gives a pictorial representation of a typical CO<sub>2</sub> transportation network across multiple sources and sinks. The proposed model thus solves the following optimisation problem.

*Given* a set of emitter nodes ( $n \in N$ ) with known locations and CO<sub>2</sub> emission ( $Q_{nt}^{em}$ ) and/or utilisation rates ( $Q_{nt}^{um}$ ); a set of transport routes represented by the arcs ( $(n, n') \in A$ )

between nodes to final storage points, minimum and maximum pipeline operating conditions and CO<sub>2</sub> fluid properties.

Determine the optimal pipeline diameters (internal and outer) from nominal sizes, pipeline material, and the number of pipeline segments to minimise the capital and operating costs for transporting emitted CO<sub>2</sub> to pre-specified sequestration sites.

It is assumed that the CO<sub>2</sub> transportation network operates at steady state.

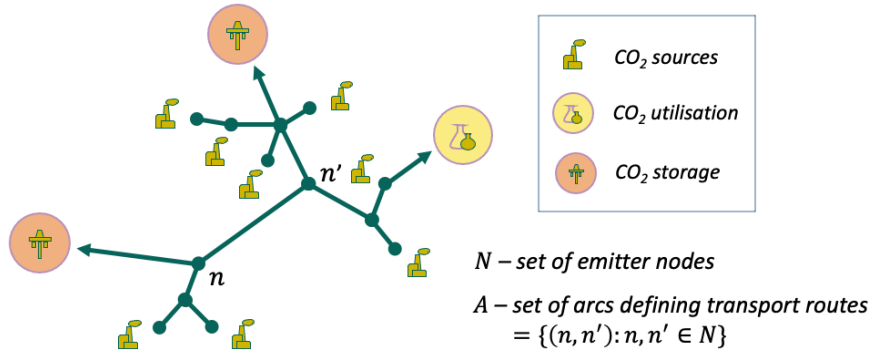


Figure 1: Multi-source, multi-sink transportation network

### 3. Mathematical formulation

The objective function (eq. (1)) for the proposed model is the sum of the capital ( $TC^C$ ) and operating ( $TC^O$ ) cost of the transportation infrastructure subject to the constraints outlined in eqs. (2) - (16). The capital cost is as expressed in Mechleri et al. (2017) and the operating cost is taken as 3% of the capital cost (IEA-GHG, 2005).

$$\min TC^C + TC^O \quad (1)$$

The mass balance equation in eq. (2) ensures that the flow of CO<sub>2</sub> ( $Q_{nn'}^{fm}$ ) to and from a node  $n$  balances the amount being emitted ( $Q_{nt}^{em}$ ) and utilised ( $Q_{nt}^{um}$ ) at the node.  $N_{nn'}^p$  in eq. (3) is an integer variable representing the number of pipeline segments for arc  $(n, n') \in A$ .

$$\sum_{n':(n',n) \in A} Q_{n'n}^{fm} + Q_{nt}^{em} = \sum_{n':(n,n') \in A} Q_{nn'}^{fm} + Q_{nt}^{um} \quad \forall n \in N \quad (2)$$

$$Q_{nn'}^{fm} = N_{nn'}^p Q_{nn'}^{f_{m'}} \quad \forall (n, n') \in A \quad (3)$$

The Darcy-Weisbach-based hydraulic equation is given in eq. (4), where  $D_{nn'}^{pi}$ ,  $f_{nn'}^D$ ,  $\rho^0$  and  $\Delta P_{nn'}$  represent the internal diameter of the pipe, the Darcy friction factor, the CO<sub>2</sub> density and the pressure drop in flow respectively. The pressure drop is defined by eq. (5) as the difference between the desired output pressure ( $P_n^o$ ) from node  $n$  and the terminal pressure ( $P_{n'}^i$ ) for flow into node  $n'$ .

$$D_{nn'}^{pi} = \left( \frac{8 \cdot f_{nn'}^D \cdot (Q_{nn'}^{f_{m'}})^2}{\pi^2 \rho^0 \cdot \frac{\Delta P_{nn'}}{L_{nn'}}} \right)^{\frac{1}{5}} \quad \forall (n, n') \in A \quad (4)$$

$$\Delta P_{nn'} = P_n^o - P_{n'}^i, \quad \forall (n, n') \in A \quad (5)$$

The Darcy friction factor is calculated using Swamee–Jain's approximation of the Colebrook–White equation (Mechleri et al., 2017) (eq. (6)).  $\xi$  is the pipe roughness and Reynold's number ( $Re_{nn'}$ ) is calculated using eq. (7).

$$f_{nn'}^D = \frac{1}{\left( \log_{10} \left( \frac{\xi/D_{nn'}^{pi}}{3.7} + \frac{5.74}{Re_{nn'}} \right) \right)^2} \quad \forall (n, n') \in A \quad (6)$$

$$Re_{nn'} = \frac{4Q_{nn'}^{fm'}}{\mu \pi D_{nn'}^{pi}} \quad \forall (n, n') \in A \quad (7)$$

The required pipe wall thickness ( $t_{nn'}^{wc}$ ) calculation (eq. (8)) considers the longitudinal joint factor ( $\kappa^E$ ), outer diameter ( $D_{nn'}^{po}$ ), specified minimum yield stress ( $\tau^p$ ), design factor ( $\kappa^{DF}$ ) of the pipe and its material, as well as the maximum operating pressure ( $P^{MOP}$ ) for the pipe. Internal and outer diameters, as well as the pipe wall thickness, are matched with calculated values and the set of nominal pipe sizes ( $d$ ) according to eqs. (9) - (13).  $B_{nn'd}^d$  is a binary variable which takes a value of 1 if the nominal pipe size  $d$  is selected for arc  $(n, n')$ .

$$t_{nn'}^{wc} = \frac{P^{MOP} \cdot D_{nn'}^{po}}{2\tau^p \kappa^E \kappa^{DF}} \quad \forall (n, n') \in A \quad (8)$$

$$t_{nn'}^w \geq t_{nn'}^{wc} \quad \forall (n, n') \in A \quad (9)$$

$$D_{nn'}^{po} = D_{nn'}^{pi} + 2 \cdot t_{nn'}^w \quad \forall (n, n') \in A \quad (10)$$

$$D_{nn'}^{po} = \sum_d D_d^{so} B_{nn'd}^d \quad \forall (n, n') \in A \quad (11)$$

$$D_{nn'}^{pi} = \sum_d D_d^{si} B_{nn'd}^d \quad \forall (n, n') \in A \quad (12)$$

$$\sum_d B_{nn'd}^d = 1 \quad \forall (n, n') \in A \quad (13)$$

To prevent problems of erosion within the pipe during flow, the erosional velocity ( $v_{nn'}^e$ ) is calculated using eq. (14) and the CO<sub>2</sub> flow velocity ( $v_{nn'}^f$ ) constrained by it (eq. (16)).

$$v_{nn'}^e = \sqrt{\frac{8\tau^w}{\rho^o f_{nn'}^D}} \quad \forall (n, n') \in A \quad (14)$$

$$v_{nn'}^f = \frac{4Q_{nn'}^{fm'}}{\pi \rho^o D_{nn'}^{po.2}} \quad \forall (n, n') \in A \quad (15)$$

$$v_{nn'}^f \leq v_{nn'}^e \quad \forall (n, n') \in A \quad (16)$$

#### 4. Case study

The proposed model outlined in section 3 was applied to a case study shown in Figure 2. It consisted of 5 CO<sub>2</sub> emitters and 4 arcs defining the transport route between each node. Each of these 5 nodes represents large CO<sub>2</sub> emitters in Belgium and The Netherlands and is chosen as part of the investigations of the Horizon 2020 [C<sup>4</sup>U project](#). C<sup>4</sup>U is a holistic interdisciplinary project addressing all the essential elements required for the optimal

integration of CO<sub>2</sub> capture in the iron and steel industry as part of the CCUS chain. It includes a demonstration of highly efficient CO<sub>2</sub> capture technologies for an iron and steel plant and detailed consideration of the safety, environmental, societal, policy and business aspects for successful incorporation into CCUS clusters. For this study, each of nodes 0 - 4 were assumed to have 0.52Mt, 1.23Mt, 1.08Mt, 0.77Mt and 0.48Mt of CO<sub>2</sub> being transported per year respectively at a maximum pipeline operating pressure of 110 bar. The model was solved using Pyomo with BARON (Kilinc & Sahinidis, 2018) solver on an Intel Xeon E-2146G with 32GB and 12 threads running on Windows 10.



Figure 2: Case Study showing emission nodes and transport route

Table 1 shows the results obtained from the solution of the model. The internal and outer diameters of each pipe along the route were determined and matched to standard nominal pipe sizes whilst meeting constraints of the pipe wall thickness (eq. (10)) and the erosional velocity (eq. (16)). The number of parallel pipeline segments was also determined for each arc to meet the material transport requirements at a minimal cost. Model solutions were obtained in 78 seconds with a relative gap of 0%, showing a modest computational performance.

Table 1: Optimal pipeline sizes and costs

$n$	$n'$	$L_{nn'}$ (km)	$D_{nn'}^{pi}$ (m)	$D_{nn'}^{po}$ (m)	$N_{nn'}^p$	$TC^C$ (M€)	$TC^O$ (M€)
0	1	5.7	0.1091	0.1143	1	2.74	0.08
1	2	10.0	0.2091	0.2191	1	3.50	0.10
2	3	8.1	0.2091	0.2191	1	3.25	0.10
3	4	13.7	0.2604	0.273	1	4.33	0.13

## 5. Conclusion

In this work, an MINLP model for the optimal sizing of infrastructure for CO<sub>2</sub> transportation via pipelines is presented. Our proposed model adopted rigorous hydraulic

equations to accurately describe CO<sub>2</sub> flow within pipelines in an optimisation model, as opposed to using cost estimates based on existing natural gas projects which may underestimate final costs. The resulting model is an extension of the MINLP proposed by Mechleri et al. (2017) and seeks to minimise the total capital and operating costs of pipelines required for a multi-source multi-sink network. It further accounts for CO<sub>2</sub> utilisation at nodes, and multiple pipe segments, whilst being solved to global optimality. The proposed model was applied to a case study with 5 nodes and 4 arcs obtaining the optimal internal and outer pipeline diameters, number of segments and associated capital and operating costs in modest computational times. Further work will include extending the model for multi-period and multi-modal transport considerations.

## 6. Acknowledgement

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 884418. The work reflects only the authors' views and the European Union is not liable for any use that may be made of the information contained therein.

## References

- Bui, M., Adjiman, C. S., Bardow, A., Anthony, E. J., Boston, A., Brown, S., Fennell, P. S., Fuss, S., Galindo, A., Hackett, L. A., Hallett, J. P., Herzog, H. J., Jackson, G., Kemper, J., Krevor, S., Maitland, G. C., Matuszewski, M., Metcalfe, I. S., Petit, C., Puxty, G., Reimer, J., Reiner, D. M., Rubin, E. S., Scott, S. A., Shah, N., Smit, B., Trusler, J. P., Webley, P., Wilcox, J., Mac Dowell, N., 2018. Carbon capture and storage (CCS): The way forward. *Energy and Environmental Science* 11, 1062–1176.
- D'Amore, F., Romano, M. C., Bezzo, F., 2021. Optimal design of European supply chains for carbon capture and storage from industrial emission sources including pipe and ship transport. *International Journal of Greenhouse Gas Control* 109, 103372.
- Ejeh, J. O., Yousef, A. Z., Bugryniec, P., Wilkes, M. D., Porter, R. T., Martynov, S., Mahgerefteh, H., Brown, S., 2022. Perspectives on Future CCUS Infrastructure Design. <http://dx.doi.org/10.2139/ssrn.4271511>.
- Hasan, M. F., Zantye, M. S., Kazi, M.-K., 2022. Challenges and opportunities in carbon capture, utilization and storage: A process systems engineering perspective. *Computers & Chemical Engineering* 166, 107925.
- IEA-GHG, 2005. Building the cost curves for CO<sub>2</sub> storage: European sector. Tech. rep.
- Kilinc, M. and N. V. Sahinidis, 2018. Exploiting integrality in the global optimization of mixed-integer nonlinear programming problems in BARON, *Optimization Methods and Software*, 33, 540-562.
- Kim, C., Kim, K., Kim, J., Ahmed, U., Han, C., 2018. Practical deployment of pipelines for the CCS network in critical conditions using MINLP modelling and optimization: A case study of South Korea. *International Journal of Greenhouse Gas Control* 73, 79–94.
- Mechleri, E., Brown, S., Fennell, P. S., Mac Dowell, N., 2017. CO<sub>2</sub> capture and storage (CCS) cost reduction via infrastructure right-sizing. *Chemical Engineering Research and Design* 119, 130–139.
- Serpa, J., Morbee, J., Tzimas, E., 2011. Technical and economic characteristics of a CO<sub>2</sub> transmission pipeline infrastructure. Tech. rep., JRC62502.
- Zhang, S., Liu, L., Zhang, L., Zhuang, Y., Du, J., 2018. An optimization model for carbon capture utilization and storage supply chain: A case study in Northeastern China. *Applied Energy* 231, 194–206.

# Data Driven Leak Detection in a Real Heat Exchanger in an Oil Refinery

Aslı Yasmal<sup>a,b</sup>, Gizem Kuşoğlu Kaya<sup>a</sup>, Emirhan Oktay<sup>a</sup>, Ceylan Çölmekci<sup>a</sup>,  
Erdal Uzunlar<sup>b</sup>

<sup>a</sup> *Turkish Petroleum Refinery, Körfez, Kocaeli, 41780, Türkiye*

<sup>b</sup> *Department of Chemical Engineering, Izmir Institute of Technology, Urla, Izmir,  
35430, Türkiye*

*erdaluzunlar@iyte.edu.tr*

## Abstract

This study focuses on implementation of a data-based leak detection method in a heat exchanger in a petroleum refinery. We have studied on the two real leakage cases in a heat exchanger in Izmit TUPRAS Refinery. Leaks are one of the major problems that occur in operations. The autoencoder (AE) method is implemented for leak detection. Reconstruction error is used as the leak indicator. In case of leakage, the reconstruction value is expected to increase. For both cases examined, the reconstruction error is found to be around 1-5 under normal operating conditions. On the other hand, reconstruction error is observed to change between 10 and 60 under the conditions with leakage. Besides, the AE is able to indicate the start of one leakage case before the process engineers noticed it.

**Keywords:** Oil and gas industry, leak detection, heat exchanger, data-based, autoencoder

## 1. Introduction

The oil and gas industry converts crude oil into valuable products and plays crucial role in global economy (Asghari and Rakhshanikia, 2013). This industry consists of many complex equipment and pipeline systems which are important to production and transportation of products and by-products (Gossen and Velichkina, 2006). Continuous, safe, efficient and environmentally sustainable production is sought for. However, faults (or anomalies) such as equipment failures, human errors or leaks sometimes disrupt the production process (Ohtani, 2020). Leaks constitute a major portion of faults, making leak detection an important requirement in the production process.

In the literature, leak detection methods are grouped under three main categories. These are direct, indirect and inferential methods (Zaman *et al.*, 2020). Inferential methods generally show the probability of the leak, so generally these methods are not preferred for leak detection. Direct methods involve hardware-based methods, which require the installation of external sensors to detect leaks in pipeline. Indirect methods involve software-based methods that are classified as data-driven and model-based methods. In addition to these methods, there are biological methods that make use of the senses of humans and experienced animals (Mujtaba *et al.*, 2020). These methods are generally used for the detection of leaks in pipelines.

The methods used in the literature to detect leaks in heat exchangers are limited and mostly suitable for offline detection. As an online method optimal exponential moving average (EMA) method was proposed by Panday et al. (2021), who studied leaks at a thermal power plant. The performances of Kalman filter (optimal EMA) and simple EMA filter were compared. It was found that Kalman filter can successfully represent the actual data trend. A long short-term memory (LSTM) method was proposed by Guillen et al. (2020), who studied leaks at a nuclear power plant. They could successfully detect the leaks in the heat exchanger. In another study, Habbi et al. (2009) studied a pilot heat exchanger and proposed A Takegi-Sugeno (TS) fuzzy model-based approach to model heat exchanger. They focused on the residuals that are calculated by using model estimated values and actual measurements. Leaks in the heat exchanger could be detected based on the residuals. In addition, they tested the relationship between leak magnitude and leak detection capability of the model. It was concluded that detectability increased with the increasing leak volume.

In this study, leaks in a heat exchanger which is located in the Integrated Unicracking Unit of İzmit TUPRAS Refinery is studied. The autoencoder method is implemented at two different leakage cases, and detection of leaks are investigated.

## 2. Materials and methods

### 2.1. The Process Unit

Figure 1 shows the simplified process flow diagram of the Integrated Unicracking Unit under study. The feed of the unit consists of Coker Naphtha (CN), Heavy Coker Gas Oil (HCGO), Light Coker Gas Oil (LCGO), Heavy Vacuum Gas Oil (HVGO), Light Vacuum Gas Oil (LVGO). The unit consists of two parts: a reactor part and a separation part. The streams coming from reactor part are collected in a fractionator column and separated into the desired products. Coker Naphtha (CN), Heavy Coker Gas Oil (HCGO), Light Coker Gas Oil (LCGO), Heavy Vacuum Gas Oil (HVGO), Light Vacuum Gas Oil (LVGO) come from different units and compose the feed of Unicracking unit. The main purpose of this unit is to produce kerosene and diesel in accordance with certain specifications. The heat exchanger, at which leakage occurred several times, is located in the part where separation steps take place.

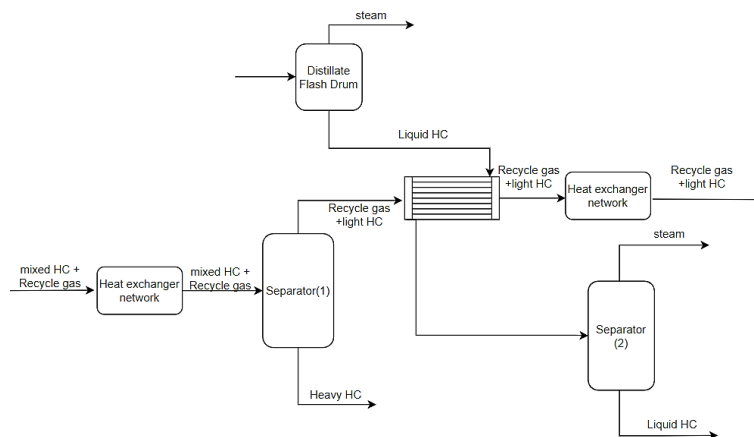


Figure 1: The scheme of the unit

The heat exchanger under study is the shell-and-tube, counter flow-type heat exchanger. Liquid hydrocarbon (HC) passes from the shell side, recycle gas and liquid HC passes from the tube side. If a leak in a heat exchanger is not noticed in time, it causes a sudden and dangerous shutdown of the process unit. Sudden interruption of the unit feed causes coking on the catalyst. At the same time, the sudden increase in temperature during startup of the unit also causes approximately 3-4 month decrease in catalyst lifespan. When the long-term effect is examined, catalyst replacement means an extra 10-day downtime for the unit. Also, the downtime of a particular process unit affects other process units in the refinery. Liquid HC passes from the shell side contains H<sub>2</sub>S. It causes the deformation on surface of the tubes. Because the pressure of the tube side higher than shell side, leaks occur from tube to the shell. In cooperation with the process engineers of the unit, it is determined that the leaks were detected by the opening of the valve controlling the separator (2). The variables that may be related with the valve opening (%) value are determined. These are inlet and outlet temperatures of the shell and tube parts in the heat exchanger (°C), the flowrate of the unit (kg/h), the pressure of the separator (kg/cm<sup>2</sup>).

### 2.2. Autoencoder (AE)

The proposed method for leak detection is autoencoder (AE) method. AE is an artificial neural network technique, which is generally used to reduce dimensionality of the dataset, extract features, denoise and recognize images. Five important hyperparameters of AE are encoder, decoder, bottleneck, loss function and epoch number. Figure 2 shows a basic representation of AE. Each block shown in Figure 2 consists of hidden layers. Each layer consists of nodes. In the encoder part, the sent data is compressed and fed to the bottleneck where the data is compressed. Compressed data stored in the bottleneck is then sent to the decoder where the data is reconstructed.

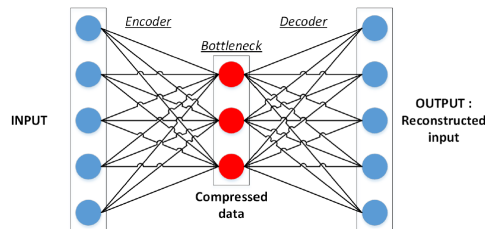


Figure 2. The schematic structure of autoencoder (Sublime *et al.*, 2019).

Encoding and decoding networks are mathematically represented in the Eq. (1-2). Here,  $W$  is the weight matrix in the encoder part,  $W'$  is the weight matrix in the decoder part,  $b$ , and  $b'$  are bias vectors in the encoder and decoder parts, respectively,  $z$  is the bottleneck dimension,  $\sigma$  is the activation function in the encoder part and  $\sigma'$  is the activation function in the decoder part, and  $x'$  is the output.

$$z = \sigma(Wx + b) \tag{1}$$

$$x' = \sigma'(W'z + b') \tag{2}$$

Reconstruction error (RE) is the metric to evaluate the deviation of the reconstructed data from the actual data. RE is used as the outlier score (An and Cho, 2015). While this error



is small for normal data samples, it is expected to be high for data with anomaly. RE is given is Eq. (3) as

$$\text{Reconstruction Error} = \sqrt{\sum_{i=1}^n (x_i - x'_i)^2} \quad (3)$$

where n refers to number of data samples,  $x_i$  is the original data sample and  $x'_i$  is the reconstructed data sample.

### 2.3. Methodology

The procedure of AE is given below as:

- Data is taken from TUPRAS Historian Database.
- The data is split as training and test dataset (training set: normally operated dataset without anomalies; test set: the dataset which has anomalies).
- Model hyperparameters are determined and the model is trained.
- Test dataset is fed to the model as an input.
- Reconstruction error is calculated using x and x'.

## 3. Results and Discussion

The AE network of this study is based on a literature study where AE was used for clustering of time series data (Tavakoli et al., 2020). In this structure, there are 3 layers in the encoder part with number of nodes descending in the order of 100, 50, 20. The bottleneck contains as many nodes as the number of variables (see Section 2.1). The decoder part is symmetrical to the encoder part with the ascending number of nodes. The AE is implemented in MATLAB using Deep Learning Toolbox. A 2-months dataset is chosen as the training set. The selection of the training set is done in collaboration with process engineers. During training, an MSE (mean square error) value of  $3.76 \times 10^{-6}$  is achieved with an epoch number of 37.

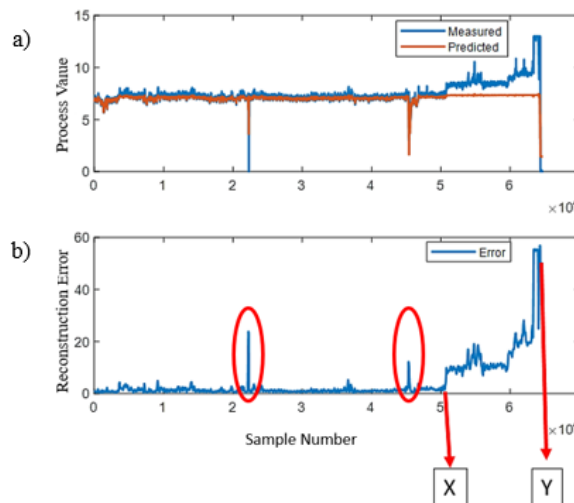


Figure 3. AE results for Case 1, (a) measured and predicted process values, (b) reconstruction error: circle on the left: flowrate changes, circle on the right: pump

replacement, point X: leakage starting point as detected by AE, point Y: the point leak noticed by process engineers.

The results for the test dataset are given in Figure 3. Figure 3a shows the measured and predicted process values and Figure 3b shows the reconstruction error. In the Figure 3b, the red circle on the left is due to a flowrate change in the process (deliberate or due to flowrate changes in upstream processes), and the red circle on the right is due to a pump replacement, point X is the moment at which the leak started as detected by AE, and point Y is the moment at which the leak was noticed by the process engineers and the unit was shut down. As seen in Figure 3b, the reconstruction error is very small, between a value of 1-5, under normal operating conditions without any flowrate changes, equipment replacements, or leaks. Reconstruction error temporarily increases with equipment changes and flowrate variations and diminishes afterwards. However, after some period of time (at point X), the reconstruction error is seen to increase steadily to a higher value than that of temporary changes (reconstruction value of 55-60 versus 20-30). Thus, point X seems to indicate the starting moment of the leakage as detected by AE. With the AE method, the leak could be detected almost 6 days ago before the process engineers took action for shut down the process.

Figure 4 shows the AE results for Case 2, a similar procedure is followed as Case 1. Since the shutdowns occur very frequently in the unit, a 1-month dataset is determined as the training dataset. Since a shorter training set interval is chosen for Case 2 compared to Case 1 and long-term flowrate variations are observed for Case 2, an MSE value of 0.0045 is obtained after 500 epochs in the training step. The obtained results for the test dataset are given in Figure 4. The effect of flowrate changes in the unit is clearly reflected in the predicted values and reconstruction error. No indication of leak is observed before process engineers noticed the leak for Case 2 unlike Case 1. In Case 2, the leak seems to show a non-steady behavior unlike Case 1 where the leak seems to increase steadily over time. This might be due to different leak types occurring in the cases. In both Case 1 and Case 2, leaks are observed to increase the reconstruction error, as expected.

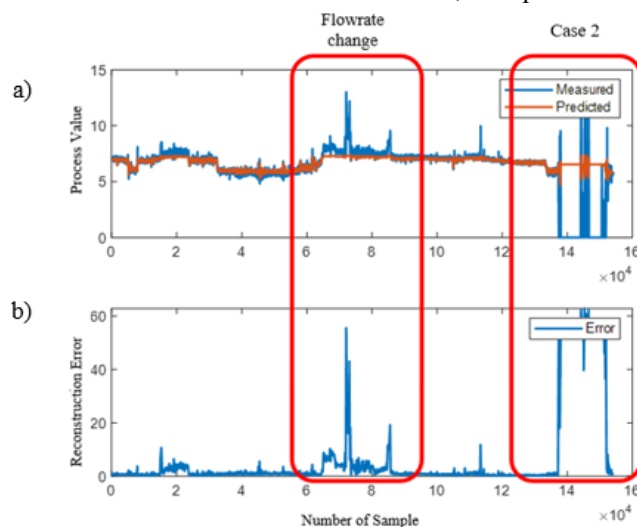


Figure 4. AE results for Case 2, (a) measured and predicted process values, (b) reconstruction error: ellipse on the left: flowrate changes, ellipse on the right: leakage case.

#### 4. Conclusion

In this study, we investigate the software-based detection of leaks in heat exchangers using real process data. The cases are located in the kerosene and diesel production unit of the TUPRAS Izmit Refinery. Two different leakage cases occurring in the same heat exchanger are examined. For both cases, AE method enables the detection of leaks in the unit. For one of the cases, it is possible to predict the leak 6 days in advance with the AE method, whereas an early prediction of leak is not possible in the other case. This might be due to different types of leaks, i.e., some leaks may start with small volumetric flowrates and progress slowly enabling early detection by AE, and some leaks may occur instantly at high volumetric flowrate hindering early detection by AE. These results seem to indicate that the AE method can be used as a pre-emptive method for leaks with the low onset volumetric flowrate, and as an alarm for leaks with instant high volumetric flowrate.

#### References

- A. Habbi, M. Kinnaert, M. Zelmat, 2009, A complete procedure for leak detection and diagnosis in a complex heat exchanger using data-driven fuzzy models', *ISA Transaction*, 48, 3 , 354-361.
- A. Tavakoli, S.S. Namini, M. A. Khanghah, F. M.Soltani, A. S. Namin, 2020, An Autoencoder-Based Deep Learning Approach for Clustering Time Series Data, *SN Applied Science* 2(5), 1-25.
- D. P. Guillen, N. Anderson, C. Krome, R. Boza, L. M. Griffel, J. Zouabe, and A. Y. Al Rashdan, 2020, A RELAP5-3D/LSTM Model for the Analysis of Drywell Cooling Fan Failure, *Progress in Nuclear Energy* 130, 103540.
- D. Zaman, M. K. Tiwari, A. K. Gupta, D. Sen, 2019, A review of leakage detection strategies for pressurised pipeline in steady-state, *Engineering Failure Analysis*, 109, 104264.
- J. An, S. Cho, 2015, Variational Autoencoder based Anomaly Detection using Reconstruction Probability, *Signal & Image Processing : An International Journal*, 2(4), 13–25
- J. Sublime and E. Kalinicheva, 2019, Automatic Post-Disaster Damage Mapping Using Deep-Learning Techniques for Change Detection: Case Study of the Tohoku Tsunami, *Remote Sensing*, 11(9), 1123
- L. P. Gossen, and L.M. Velichkina, 2016, 'Environmental problems of the oil-and-gas industry (review)', *Petroleum Chemistry*, 46(2), 67–72.
- M. Asghari, and M. A. Rakhshanikia, 2013, 'Technology in Oil Industry, Significance and Challenges', *Procedia - Social and Behavioral Sciences*, 75, 264–271.
- M. Mujtaba, T.A. Lemma, S.A.A Taqvi, T.N.Ofei, 2020, 'Leak detection in gas mixture pipelines under transient conditions using hammerstein model and adaptive thresholds', *Processes*, 8(4).
- R. Panday, N. Indrawan, L. J. Shadle, R. W. Vesel, 2021, Leak Detection in a Subcritical Boiler, 185, 116371.
- T. Ohtani, 2020, Application of AI to Oil Refineries and Petrochemical Plants', 63(1), 7–10.

# Optimization of flash-separation based CO<sub>2</sub> purification units

Lorenzo Sala<sup>a</sup>, Syed Ali Zaryab<sup>a</sup>, Paolo Chiesa<sup>a</sup>, Matteo Romano<sup>a</sup>, Emanuele Martelli<sup>a\*</sup>

<sup>a</sup>*Department of Energy, Politecnico di Milano, via Lambruschini 4, 20156, Milano, Italy*  
*\*emanuele.martelli@polimi.it*

## Abstract

Several CO<sub>2</sub> capture and storage technologies, such as oxycombustion systems, membranes, pressure/temperature swing adsorption systems separate a stream of CO<sub>2</sub> which does not meet the tight composition specifications required by the storage site. Thus, the captured CO<sub>2</sub> must be processed by a CO<sub>2</sub> purification unit (CPU). This paper will focus on flash separation-based CPUs where the partial liquification of the pressurized feed gas takes place by temperature reduction at constant pressure. In this paper two schemes of the CPUs are considered, one is a conventional flash separation-based CPUs and the other is a novel design with increased recovery of the CPU. The specific cost (€/ton<sub>CO2</sub>) of these CPUs has been optimized for different levels of carbon tax and captured CO<sub>2</sub> purity using SCR, a recently published surrogate-based global-search algorithm.

**Keywords:** CPU, Global Optimization, CCS, process optimization, CO<sub>2</sub> purification

## 1. Introduction

CCS (CO<sub>2</sub> capture and storage) technologies have been proposed as an attractive method to capture the anthropogenic CO<sub>2</sub> released from industrial sources to the atmosphere and mitigating the effect of climate change. This captured CO<sub>2</sub> has to be compressed, transported and stored as per guidelines based on environmental regulations, technical, economical and/or safety reasons [1]. All this can be achieved with a CPU (CO<sub>2</sub> compression and purification unit) [2] which is used to remove impurities from the feed gas and obtain high purity CO<sub>2</sub>.

In a CPU the phase separation takes place due to the difference in volatility between CO<sub>2</sub> and other non-condensable gases. Based on the method used for separation of CO<sub>2</sub> the CPU can be divided into two types:

- 1- Flash Separation based CPUs: In these CPUs the partial liquification of the pressurized feed gas takes place by temperature reduction at constant pressure. These CPUs do not have any complex separation equipment and are less costly.
- 2- Stripping/Distillation-based CPUs: In these CPUs a stripper or distillation column (integrated with flash separation) is used to increase the purity of CO<sub>2</sub> in the feed gas stream and remove unwanted contaminants. This method provides a high purity CO<sub>2</sub>.

## 2. CPU Configuration

This paper will focus on two different types of Flash separation-based CPUs. The first CPU configuration (CPU 1) is taken from the IEAGHG report ref. [3]. This configuration

includes two cryogenic multi-flow heat exchangers, two flash drums for phase separation and two valves which are used to provide auto refrigeration by throttling the purified  $\text{CO}_2$ .

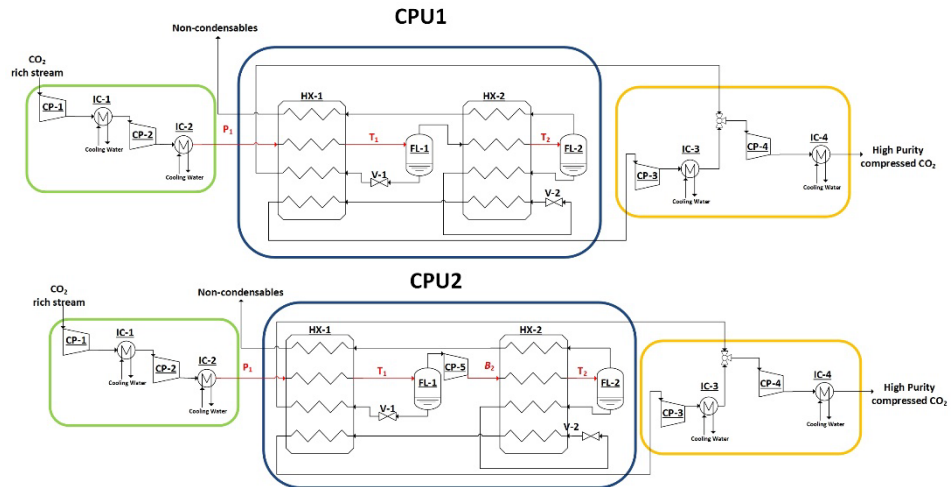


Fig. 1 Flowsheet of the two CPUs selected for the study. On the top there is a conventional flash-based CPU taken from [3] denoted as CPU1. On the bottom there is a novel flash-based CPU, denoted as CPU2.

The flowsheet of CPU1 is shown in Fig. 1, here the  $\text{CO}_2$  rich feed gas stream is initially cooled and compressed by a multi-stage intercooled compressor. This compressed gas is then sent to the first heat exchanger where it is cooled down to cryogenic temperatures against the evaporating streams, superheating  $\text{CO}_2$  streams, and the waste stream. This cooled feed is then sent to the first flash where it is separated into liquid and vapor phase. The liquid phase is almost pure  $\text{CO}_2$  while the vapor phase contains some  $\text{CO}_2$  along with all the lighter components (waste product) present in the gas stream. This vapor phase is sent to the second heat exchanger where it is cooled to a lower temperature and is then sent to a second flash. The vapor from the second flash is vented: this vent stream contains impurities and a small part of the  $\text{CO}_2$ . This vent stream is sent back through the two heat exchangers where it provides cooling to the  $\text{CO}_2$  rich streams entering the heat exchangers. This stream is then released into the atmosphere. The liquid phase from the first flash is throttled through a valve, heated, and then sent for compression. While the liquid phase coming from the second flash is heated, throttled through a valve, and then passed through the two heat exchangers. Throttling of the second liquid stream results in a very low temperature (close to the solidification temperature of  $\text{CO}_2$ ) so that it can be used as a refrigerator in the second heat exchanger. This second stream is then compressed. The two streams are then mixed and finally compressed for transportation and storage.

The main parameters of this system are the pressure and the temperature of the two flash separators because they influence the composition of the outlet liquid, the gas flows, and the cooling requirement. The pressure of the first flash can be optimized by changing the pressure at the discharge of the first compressor chain. Pressure of the second flash is fixed and equal to that of the first flash minus the pressure drops introduced into the exchanger. Temperatures are determined by the energy balance on the heat exchangers and therefore are determined by the inlet currents, whose flow rates, compositions, pressure, and temperature are functions of the pressure at which the flash separators operate, and the pressure drops introduced by the valves.

Based on these considerations, the scheme of CPU1 can be modified, to improve its performance, by decoupling the pressures of the two flash separators. This can be achieved by adding a compressor between the two flash separators. This modified CPU scheme is shown in Fig. 1 which will be called CPU2 in this paper.

### 3. Modelling Methodology

The models of the CPUs have been developed in Aspen Plus V10, a commercial sequential modular process simulation software using the sequential convergence mode. In order to perform an accurate evaluation of the performance of the CPU, it was crucial to adopt a reliable Equation of State model (EoS) which can predict the vapor-liquid equilibria of multicomponent systems. Therefore, experimental data of multi components systems containing N<sub>2</sub> and CO<sub>2</sub> at different temperatures and pressures available in refs. [4] [5] [6] were compared with the results of different EoS in Aspen Plus to determine which method best approximates the experimental results. This analysis showed that the Peng-Robinson EoS with Van der Waals mixing rules provides the best approximation as the mean square error between this EoS and the experimental data is the lowest. Thus, Peng-Robinson equation of state has been adopted for the evaluation of the properties of all the streams involved in the simulations.

### 4. Optimization Methodology

In this work the specific cost (€/ton<sub>CO2</sub>) of these CPUs has been optimized using SCR, a recently published optimization algorithm [7]. SCR is a surrogate-based derivative-free global optimization algorithm which works well for process engineering problems and provides the optimum in limited number of functional evaluations compared to other algorithms [7]. SCR uses latin hypercube sampling (LHS) to sample a limited number of points in the design space where the black box function is evaluated. Using the results of these sample points SCR creates a kriging based surrogate model of the objective function and its constraints. SCR then calls CMA-ES to find the global optimum of the constrained surrogate model. Within CMA-ES the surrogate models of the objective function and the constraints are combined using quadratic penalty approach. Once the global optimum of the surrogate model is found, SCR calls the local search algorithm RQLIF [8] which uses a limited number of black box functional evaluation, to find the local minimum of the real function, around the region where the global minimum of the surrogate model was found. After this the surrogate models are updated with the points that RQLIF tested along with some new sample points found by LHS. The SCR algorithm stops when for two subsequent iterations RQLIF finds the same optimum within a certain tolerance range or when the maximum number of black box functional evaluations is reached.

For the optimization, four key performance indicators (KPIs) have been selected for the CPUs. These KPIs include Specific cost (SC), Recovery, Purity, and O<sub>2</sub> concentration in compressed CO<sub>2</sub>. The optimization variables for the CPU1 includes the pressure at the exit of the compression step ( $P_1$ ) and the temperatures of the two flash ( $T_1$  &  $T_2$ ) (shown in red in Fig. 1). The CPU2 has the same optimization variables with the addition of the pressure ratio of the new compressor ( $B_2$ ). The optimization activity is repeated considering three different feed streams coming from (i) the Allam cycle [9] (ii) from a combined cycle with a membrane-based CO<sub>2</sub> capture plant [10] & (iii) a cement plant with VPSA capture process [11]. The objective function is to minimize the specific cost of the CPU (€/ton<sub>CO2</sub>) at different values of carbon tax (from 100-300 €/ton<sub>CO2</sub>) and different constraint on purity of the captured CO<sub>2</sub> (95%-97%). Each optimization run took 2.5 hours corresponding to 500 runs of the flowsheet simulation.

Table 1: Flowrates and molar concentration of the three types of feed gas considered in this analysis

Source	Pressure [bar]	Flowrate [kg/s]	Molar concentration			
			CO <sub>2</sub>	N <sub>2</sub>	O <sub>2</sub>	Ar
Cement plant with VPSA [11]	1	27.29	90.00%	8.70%	1.30%	0.00%
Combined cycle with membrane [10]	1	41.15	88.32%	6.66%	5.02%	0.00%
Allam cycle [9]	30	44.37	98.04%	1.18%	0.21%	0.57%

## 5. Results

The results of the optimization at different tax level and at different constraints of purity and different input feeds are shown in the Table 2 for both CPUs. Furthermore, the graphs in the Fig. 2 show the distribution of costs at different carbon tax level for different input streams for the selected CPUs. The Total annual cost (TAC) is divided into five main categories, (i) the cost of CO<sub>2</sub> emissions (Tax Cost), (ii) the electricity cost (El Cost), (iii) the cost of the compressors at inlet (CPR in, shown in green in Fig. 1), (iv) the cost of the cryogenic separation unit (Coldbox, shown in blue in Fig. 1), and (v) the cost of the final compressor (CPR fin, shown in yellow in Fig. 1).

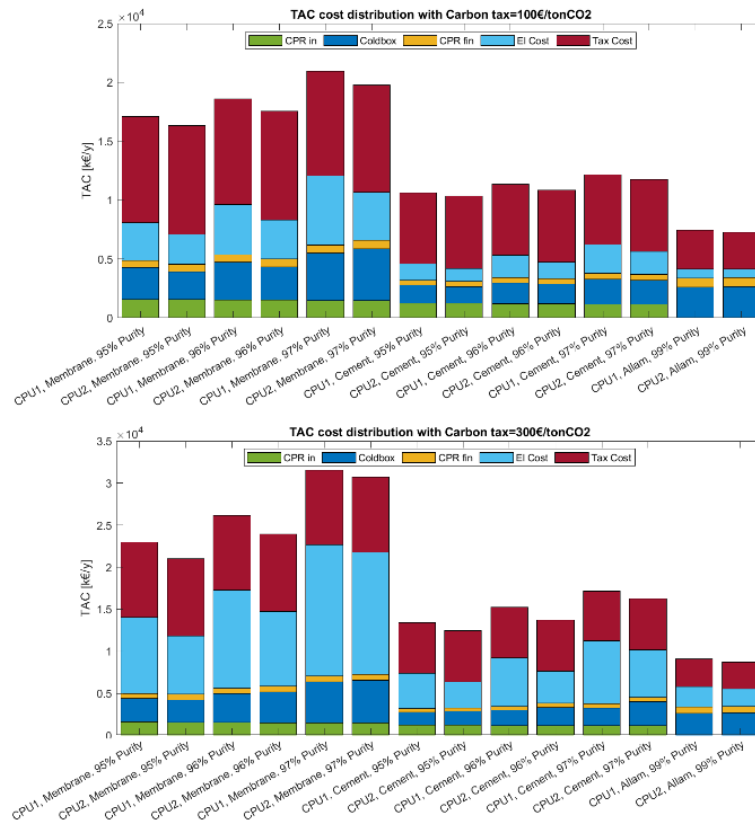


Fig. 2: The distribution of costs for different input streams for the selected CPUs at different carbon tax

The results show that CPU2 has the lowest yearly cost and specific cost in all cases compared to CPU 1. The low cost of the CPU2 is because of its high recovery, which results in a low cost of emissions (Tax cost) as seen in Fig. 2. Thus, it can be stated that

the novel design of the CPU2 is generally beneficial for the overall cost reduction of the system.

From the Table 2, it can be seen that the increase in carbon tax does not have any effect on the recovery of the system, this is because for Flash separation-based CPUs the recovery is dependent on the constraint of CO<sub>2</sub> purity at the exit of the CPU. In general, the results of the optimization have shown that all the optimal solutions for any carbon tax feature two important elements: 1- the purity of the liquid stream of the two flashes are the same and equal to the constraint 2- the optimal temperature of the second flash is set to the minimum possible value of -50°C. Thus, for CPU1 the pressure of the two flashes is automatically set by the vapor liquid equilibrium of the second flash at -50°C and the target purity level. Similarly, the temperature of the first flash is set by the vapor liquid equilibrium of the feed gas at the pressure required by the second flash and by the target purity. So these values of temperature and pressure of the two flashes determine the recovery and are independent of the carbon tax. This is evident from the results of Table 2. Instead for CPU 2 the presence of the intermediate compressors allows to change the pressures of the two flashes making it possible to vary the recovery independently from the purity although for a very limited range.

Table 2: Key performance indicators along with optimized variables for both CPUs at different feed compositions for different cases

Membrane		CPU 1							CPU 2						
Tax	Cons. Purity	P1 (bar)	T1 (°C)	T2 (°C)	Specific Cost (€/ton <sub>CO2</sub> )	Recovery	Purity	P1 (bar)	T1 (°C)	T2 (°C)	B2	Specific Cost (€/ton <sub>CO2</sub> )	Recovery	Purity	
100	>95%	34.36	-32.34	-50.0	16.18	97.04	95.00	30.73	-40.00	-50.0	1.52	15.37	97.69	95.00	
300	>95%	37.21	-25.18	-50.0	21.71	97.23	95.04	32.00	-34.73	-50.0	1.68	19.75	97.88	95.00	
100	>96%	29.21	-33.79	-50.0	17.79	96.07	96.00	26.46	-39.10	-50.0	1.45	16.63	96.95	96.00	
300	>96%	32.13	-26.39	-50.0	24.93	96.43	96.00	27.78	-32.72	-50.0	1.59	22.62	97.27	96.00	
100	>97%	24.19	-34.51	-50.0	20.34	94.60	97.04	23.87	-34.72	-50.0	1.14	18.90	96.24	96.81	
300	>97%	26.88	-27.06	-50.0	30.43	95.23	97.00	25.15	-30.35	-50.0	1.14	29.57	95.55	97.00	
<b>Cement</b>															
100	>95%	40.67	-31.26	-50.0	14.77	98.08	95.00	38.57	-34.76	-50.0	1.71	14.31	98.55	95.00	
300	>95%	40.82	-30.88	-50.0	18.66	98.08	95.00	40.43	-27.51	-50.0	1.80	17.28	98.59	95.00	
100	>96%	34.63	-32.54	-50.0	15.87	97.40	96.00	33.11	-34.00	-50.0	1.62	15.11	98.11	96.00	
300	>96%	34.31	-33.45	-50.0	21.34	97.38	96.00	34.09	-26.05	-50.0	1.92	19.10	98.27	96.00	
100	>97%	30.32	-29.14	-50.0	17.20	96.61	97.00	27.68	-33.46	-50.0	1.45	16.45	97.36	97.00	
300	>97%	30.24	-29.35	-50.0	24.24	96.60	97.00	29.39	-25.42	-50.0	1.43	22.76	97.43	97.05	
<b>Allam</b>															
100	>99%	20.86	-25.02	-50.0	5.94	99.37	99.00	22.34	-21.64	-50.0	1.01	5.80	99.43	99.00	
300	>99%	20.83	-24.92	-50.0	7.25	99.36	99.01	22.14	-21.44	-50.0	1.07	6.95	99.46	99.00	

## 6. Conclusions & Future works

In this paper optimization and comparison of a conventional Flash separation-based CPU (CPU1) and its novel modified version (CPU2) was carried out. It was seen that CPU2 always has a higher recovery and a lower specific cost for different feed inputs and different constraints on the output purity of the captured CO<sub>2</sub>. As shown in Fig. 3, the novel design has a cost that is 3% to 9% lower than the CPU1 in all cases no matter what feed gas is being used.

As mentioned earlier a conventional flash-based CPU is a cost-effective way to compress and purify CO<sub>2</sub> rich feed gas. This type of CPU can provide good purity feed gas at a



lower cost and can be used in applications where there is not a strict purity requirement. The modified version of this CPU (CPU2) shown in this paper helps in further reducing the yearly cost of the CPU and increasing the recovery that can be achieved.

On the other hand, there are certain applications which require a CO<sub>2</sub> stream with O<sub>2</sub> concentration below 100 ppm and in such cases flash-based CPUs cannot meet such requirement. In this case, the more expensive distillation-based CPU must be used. In the future, the same optimization methodology of this paper would be applied on distillation-based CPUs.

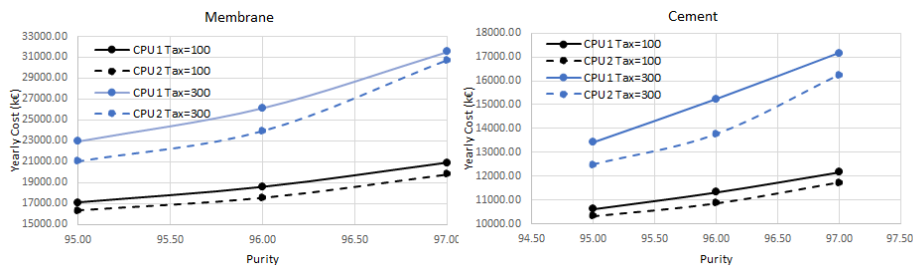


Fig. 3: Yearly cost comparison of CPU1 and CPU2

## References

- [1] E. d. Visser, C. Hendriks and et. al., "Dynamis CO<sub>2</sub> quality recommendations," *Int J Greenhouse Gas Control*, no. 2, pp. 478-484, 2008.
- [2] J. Yan, M. Anheden and et. al., "Conceptual development of flue gas cleaning for CO<sub>2</sub> capture from coal-fired oxyfuel combustion power plant," in *8th International Conference on Greenhouse Gas Control Technologies*, Trondheim, 2006.
- [3] IEAGHG, "Rotating Equipment for Carbon Dioxide Capture and Storage," 2011.
- [4] S. Lasala, P. Chiesa, R. Privat and J. Jaubert, "VLE properties of CO<sub>2</sub>-based binary systems containing N<sub>2</sub>, O<sub>2</sub> and Ar: experimental measurements and modelling results with advanced cubic equations of state, Fluid Phase Equilibria".
- [5] T. A. Al-Sahhaf et. al., "Liquid + Vapor Equilibria in the N<sub>2</sub> + CO<sub>2</sub> + CH<sub>4</sub> System," *Industrial & Engineering Chemistry Fundamentals*, vol. 22, no. 4, pp. 372-380, 1983.
- [6] G. H. Zenner and L. I. Dana, "Liquid-vapor equilibrium compositions of carbon dioxide - oxygen - nitrogen mixtures," in *Chem. Eng. Prog., Symp.*
- [7] S. A. Zaryab, A. Manno and E. Martelli, "SCR: A Novel Surrogate-Based Global Optimization Algorithm for Constrained Black-Box Problems," in *32nd European Symposium on Computer Aided Process Engineering*, Toulouse, 2022.
- [8] A. Manno, E. Amaldi, F. Casella and E. Martelli., "A local search method for costly black-box problems and its application to CSP plant start-up optimization refinement," *Optimization and Engineering*, 2020.
- [9] S. A. Zaryab, R. Scaccabarozzi and E. Martelli, "Advanced part-load control strategies for the Allam cycle," *Applied Thermal Engineering*, no. 168, 2020.
- [10] D. Turi et. al., "CO<sub>2</sub> capture from natural gas combined cycles by CO<sub>2</sub> selective membranes," in *International Journal of Greenhouse Gas Control*, 2017.
- [11] M. Voldsund et. al., "D4.6 CEMCAP comparative techno-economic analysis of CO<sub>2</sub> capture in cement plants," DOI: 10.5281/zenodo.2597090, 2019.

# Big data analysis of solar energy fluctuation characteristics and integration of wind-photovoltaic to hydrogen system

Yaoxi Chen, Siyu Yang, Yu Qian\*

*School of Chemical Engineering, South China University of Technology, Guangzhou  
510640, PR China*

Email address: [ceyuqian@scut.edu.cn](mailto:ceyuqian@scut.edu.cn)

## Abstract

Solar energy is a kind of sustainable energy; however, its random and intermittent fluctuation characteristics restrict its large-scale and high permeable applications. In this paper, the frequency spectrum analysis and filtering analysis are applied to reveal that both wind power and photovoltaic power have daily (24 h) and annual (8760 h) fluctuation cycles and the phase differences between their fluctuation period, which establish the scientific basis for energy complementarity to suppress the fluctuation. The data analysis of several regions in north and northwest China shows that the coupling of the wind power and the photovoltaic power alleviates the individual energy fluctuation. Both the unit cost of hydrogen production and the CO<sub>2</sub> emission intensity of the large-scale stable wind-photovoltaic power to hydrogen production and supply system are significantly lower than that of the non-complementary coupled solar hydrogen production system.

**Keywords:** solar energy; fluctuation period; wind-photovoltaic energy coupling; hydrogen production.

## 1. Introduction

In the context of the global transition to clean and low-carbon energy, renewable energy sources such as wind and light have great potential to compensate for the decline in coal power (Lu et al., 2009; Kabir et al., 2018). Solar energy is broadly any energy produced by the Sun, including wind energy from the atmosphere and solar energy with radiation and heat from the Sun, with strong randomness and intermittency. This random and intermittent fluctuation limits the large-scale and high-proportion application of wind and solar energy. When conventional grids are connected to a high proportion of wind and solar power, they often have to curtail electricity (Bird et al., 2016), or use large-capacity energy storage devices (Leonard et al., 2020), which have great technical obstacles in terms of safety, reliability and economy (Shaker et al., 2016; Ueckerdt et al., 2015).

In recent years, research on the complementarity of wind and solar energy has attracted attention. Wind energy and solar energy complementarity refers to the joint output of wind energy and solar energy to seek to smooth out the volatility of energy supply, improve the reliability of the system, and reduce the cost of power generation. The better the degree of complementarity between wind and solar, the higher the stability of the system output. From the perspective of data correlation, some researchers proposed to

evaluate the degree of complementarity between wind and solar by correlation coefficient (Ren et al., 2019; Tong et al., 2021). The closer the correlation coefficient is to -1, the better the complementarity of the energy in the place. However, no one has proposed a new evaluation method for the degree of complementarity of wind and solar energy from the perspective of the fluctuation law of wind energy and solar energy.

In view of the above problems, this paper explores the scientific laws of fluctuation changes in wind and solar energy. From the perspective of fluctuation periodicity, a new evaluation method for the complementarity of wind and solar energy was proposed, using data analytics to predict the phase difference between the two energies due to intermittence; A wind-solar coupling hydrogen production system to produce a stable and reliable low-carbon hydrogen source was designed; And the reasonable capacity configuration of each unit of wind energy and solar energy electrolysis hydrogen production system was considered. The works were performed in order to form an important technological innovation to promote the complementary utilization of large-scale wind and solar.

## **2. Wind energy and solar energy basic data collection and analysis**

### *2.1. Data collection of wind and solar energy*

The data in this article was taken from the World Meteorological Organization (WMO) database and the National Aeronautics and Space Administration (NASA) reanalysis database (The Modern-Era Retrospective analysis for Research and Applications Version 2, MERRA2) (Zhang et al., 2018). Studies have shown that these two databases are reliable in terms of wind and solar data. In this paper, the data of solar irradiance intensity and hourly wind speed of 50m from the surface were obtained from the NASA-MERRA2 meteorological database for 11 years from 2010 to 2020. Select the areas where large-scale energy bases in northern China have or are suitable for the construction of large-scale wind power stations and photovoltaic farms, and focus on the areas of Tongliao, Baotou, Gansu Jiuquan and Xinjiang Zhundong in Inner Mongolia.

### *2.2. Analysis of periodic fluctuation characteristics of wind energy and solar energy*

In order to analyze the fluctuation characteristics of wind power and photovoltaics, the spectrum analysis of wind power and photovoltaic power data was carried out by fast Fourier transform, and the frequency of each decomposition wave was obtained, and the frequency was converted to the corresponding time in order to reflect the time scale, and the spectrogram is shown in Figure 1. The spectrum analysis results show that the fluctuations of wind energy and solar energy have two cycles: day and night (24 h) and annual (8760 h).

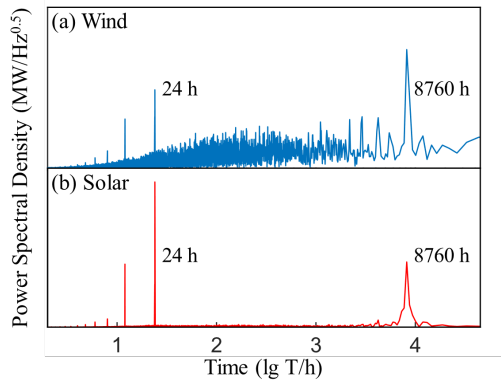


Figure 1: Frequency characteristics of wind speed and solar irradiance intensity fluctuation

Using the filter analysis method, specific frequency waves can be filtered and screened. From Figure 2, the 24 h fluctuation period of wind and solar irradiance intensity can be clearly seen, and it shows the characteristics of complementary strength.

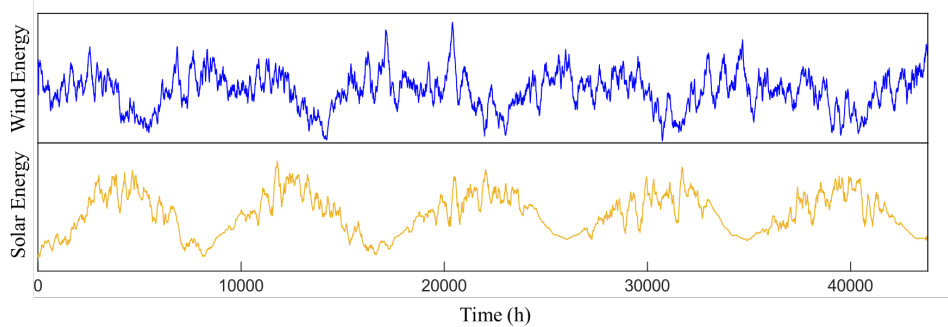


Figure 2: Wind and solar fluctuation curves with 8760 h periodic characteristics after filtering

### 2.3. Feasibility analysis of complementary coupling of wind energy and solar energy

Spectrum analysis shows that the fluctuations of wind energy and solar energy have significant periods of 8760 h and 24 h, but the fluctuation amplitude and phase of wind energy and solar energy may be different. Theoretically, the superposition of two energies can obtain the following different effects: (1) the phase of the solar energy and wind energy fluctuation cycle is the same ( $=$ ), the two fluctuation peaks and peaks are superimposed, and the two cannot cancel each other; (2) The two fluctuations have a certain phase difference ( $-$ ), partially complementary, and the two-energy coupling output will flatten and complementary, which can partially flatten the volatility. (3) If the annual fluctuation phase difference between wind and solar is 6 months, and the difference between day and night fluctuation phase is 12 hours, then on the annual and daily time scale, the peaks and troughs of wind and solar fluctuations are completely

consistent, which can be regarded as the best complementarity between the two on the annual and daily time scales.

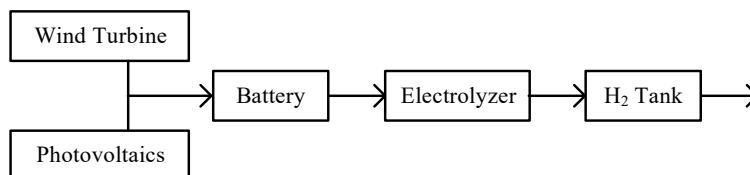
Spectrum analysis, fundamental extraction and fitting of wind energy and solar energy data in Tongliao, Baotou, Jiuquan and Zhundong can be analyzed, and the initial phase and phase difference of wind energy and solar energy in various places can be analyzed. The preliminary phase data analysis in Table 1 shows that different places have certain complementarities of scenery and solar, but they are not the same.

**Table 1: The initial phases of 8760 h and 24 h cycles of wind energy and solar energy in four locations**

Location	Annual Cycles (8760 h)			Daily Cycles (24 h)		
	Wind Power Initial Phase	Photovoltaic Power Initial Phase	Phase Difference	Wind Power Initial Phase	Photovoltaic Power Initial Phase	Phase Difference
Tongliao	December 10	March 22	3 Months 12 Days	14.25	6.21	8 Hours 4 Minutes
Baotou	December 26	March 21	2 Months 24 Days	10.24	6.09	4 Hours 14 Minutes
Jiuquan	February 3	March 23	1 Month 18 Days	6.12	6.48	36 Minutes
Zhundong	October 25	March 27	5 Months 1 Day	22.51	6.33	7 Hours 42 Minutes

### 3. Integration of wind-photovoltaic to hydrogen production system

Based on theoretical analysis and demonstration, this paper selects Zhundong with a more suitable degree of wind-solar energy complementarity for wind-solar complementary coupling, and proposes an integrated system for hydrogen production from wind-solar complementary coupled energy, including wind turbine, photovoltaics, battery, electrolyzer and H<sub>2</sub> tank, as shown in Figure 3.



**Figure 3: Topological structure of the Wind-PV to Hydrogen System**

**3.1. Specifications for integrated systems**

In this paper, the output values and the input values are the power supply or hydrogen flow rate; the average output is the average of output values in T period. Volatility is often used to evaluate the stability of a system or unit output. Volatility within the T period is defined as the average of relative deviations of the output values and the average output. The closer to zero, the less volatility.

From the perspective of the stability and safety of the operation of the integrated system, the power supply fluctuation of the battery should be less than 45%; Alkaline electrolyzers are subjected to input current fluctuations ranging from 15% to 100% of rated power; The hydrogen supply fluctuation rate of hydrogen storage equipment should be less than 10%. In order to achieve this series of key technical indicators, it is necessary to rationally allocate system capacity.

**3.2. System capacity configuration and hydrogen supply capacity**

Based on the wind energy and solar energy coupling of Xinjiang Zhundong, the hydrogen production was set to 7500 t/y, and the capacity configuration of the wind - photovoltaic to hydrogen system was carried out, and the data results are shown in Table 2.

**Table 2: The capacity configuration of wind-photovoltaic to hydrogen production system**

	PV to H <sub>2</sub>	Wind Power to H <sub>2</sub>	Wind - PV to H <sub>2</sub>
Wind Turbines Capacity (MW)	-	234	130
Photovoltaics Capacity (MW)	320	-	137
Battery Storage Capacity (MWh)	4225	7745	1250
Hydrogen Electrolyzer Capacity (m <sup>3</sup> /h)	13000	13000	13000
Hydrogen Storage Tank Volume (m <sup>3</sup> )	28790	11800	4170
Annual Hydrogen Production (t)	7500	7500	7500

After the capacity of the device was reasonably configured, the fluctuation of hydrogen production and supply of the system is within the specified range, as shown in Figure 4. At this time, the integrated system of hydrogen production from wind-solar complementary coupled energy can stably supply H<sub>2</sub> per 7500 t/y. The high-purity hydrogen produced can be used as fuel for hydrogen fuel cell buses. One hydrogen bus travels an average of 200 km per day and consumes 14 kgH<sub>2</sub>. The largest hydrogen refueling station currently has a daily injection capacity of 4.8 t. The hydrogen supply capacity of the coupled hydrogen production system can be used for at least 4 hydrogen refueling stations to refuel 1368 hydrogen fuel cell buses. According to the latest statistics, in 2019, 4411 buses in Urumqi, Xinjiang Uygur Autonomous Region, that is,

hydrogen fuel cell buses supported by hydrogen supply are enough to replace 31% of buses in Urumqi, Xinjiang.

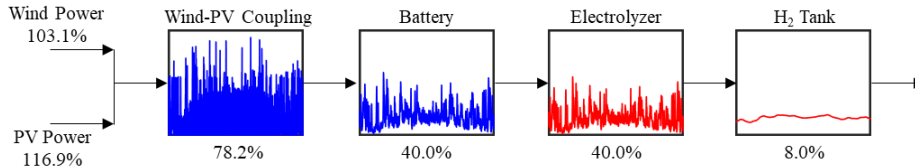


Figure 4: Variation of volatility  $\delta$  of the integrated system

#### 4. Techno-economic and environmental performance analysis

In this paper, the life cycle assessment method is used to evaluate and analyze the carbon emission performance of the wind-solar coupled hydrogen production system. Life cycle assessment is an analytical tool that evaluates the environmental impact of a product, process or activity throughout its life cycle, from the collection and processing of raw materials to the production, transportation, sale, use, maintenance, recycling and final disposal of the product (Dennison et al., 1999). The lifecycle boundaries of the system are shown in Figure 5.

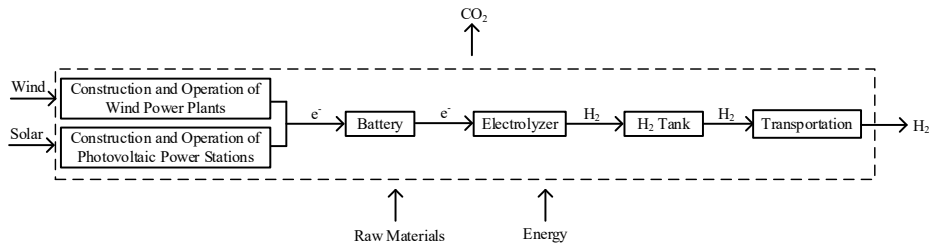


Figure 5: The life cycle boundary of wind-photovoltaic to hydrogen production system

The  $\text{CO}_2$  emission intensity and unit hydrogen production cost of photovoltaic to hydrogen production system, wind power to hydrogen production system and wind-solar coupling to hydrogen production system are compared. Among the three types of systems, the unit hydrogen production cost of photovoltaic hydrogen production system is the highest, which is 95.6 CNY/kg $\text{H}_2$ ; The unit hydrogen production cost of wind power hydrogen production system is the next second, which is 77.2 CNY/kg $\text{H}_2$ . The wind-solar coupled hydrogen production system has the lowest unit hydrogen production cost of 25.5 CNY/kg $\text{H}_2$ , which is significantly lower than the average cost of individual energy hydrogen production systems. This shows that the coupling and complementarity of wind energy and solar energy can smooth out the fluctuation of some power generation, greatly reduce the capacity configuration of batteries and hydrogen storage tanks required by the system, and thus reduce the hydrogen production cost of the wind and solar coupling system. The  $\text{CO}_2$  emission intensity of

photovoltaic hydrogen production and wind power hydrogen production was 3.91 kgCO<sub>2</sub>/kgH<sub>2</sub> and 2.21 kgCO<sub>2</sub>/kgH<sub>2</sub>, respectively. The CO<sub>2</sub> emission intensity of the wind-solar coupled hydrogen production system was 2.34 kgCO<sub>2</sub>/kgH<sub>2</sub>, which was significantly lower than the average CO<sub>2</sub> emission intensity of the separate energy hydrogen production system.

## 5. Conclusions

(1) Wind energy and solar energy have a fluctuation cycle of years and days. There is an optimal complementary phase difference on the complementary time scale, with an annual period of 6 months and a daily period of 12 hours. The analysis of case data from many places in northern China shows that the fluctuation characteristics of wind energy and solar energy in various places are different, but there are certain complementary characteristics.

(2) Design a wind-photovoltaic to hydrogen production system for large-scale hydrogen production, select battery packs and hydrogen storage tanks with appropriate capacity, and reduce the flow fluctuation rate of hydrogen supply to less than 10% to meet the requirements of stable hydrogen supply. The design system delivers enough hydrogen to allow nearly one-third of the public transport capacity of a large and medium-sized city to be replaced with hydrogen fuel cell buses.

(3) Case studies show that the hydrogen production cost of wind-photovoltaic to hydrogen production system in this paper is 25.5 CNY/kg H<sub>2</sub>, and the emission intensity is 2.34 kg CO<sub>2</sub>/kg H<sub>2</sub>, which is basically the same as that of the wind power to hydrogen production system, and the CO<sub>2</sub> emission intensity of the wind energy and solar energy coupling system is basically the same, and the obvious decrease is compared with the photovoltaic power to hydrogen production system.

## References

- Bird L, Lew D, Milligan M, et al. 2016, Wind and solar energy curtailment: a review of international experience. *Renewable and Sustainable Energy Reviews*, 65, 577-586.
- Dennison F J, Azapagic A, Clift R, et al. 1999, Life cycle assessment: comparing strategic options for the mains infrastructure-Part I. *Water Science and Technology*, 39, 315-319.
- Kabir E, Kumar P, Kumar S, et al. 2018, Solar energy: potential and future prospects. *Renewable and Sustainable Energy Reviews*, 82, 894-900.
- Leonard M D, Michaelides E E, Michaelides D N. 2020, Energy storage needs for the substitution of fossil fuel power plants with renewables. *Renewable Energy*, 145, 951-962.
- Lu X, McElroy M B, Kiviluoma J. 2009, Global potential for windgenerated electricity. *PNAS*, 106, 27, 10933-10938.
- Ren G R, Wan J, Liu J F, et al. 2019, Spatial and temporal assessments of complementarity for renewable energy resources in China. *Energy*, 177, 262-275.
- Shaker H, Zareipour H, Wood D. 2016, Impacts of large-scale wind and solar power integration on California's net electrical load. *Renewable and Sustainable Energy Reviews*, 58, 761-774.
- Tong D, Farnham D J, Duan L, et al. 2021, Geophysical constraints on the reliability of solar and wind power worldwide. *Nature Communications*, 12, 6146.
- Ueckerdt F, Brecha R, Luderer G. 2015, Analyzing major challenges of wind and solar variability in power systems. *Renewable Energy*, 81, 1-10.
- Zhang H, Cao Y, Zhang Y, et al. 2018, Quantitative synergy assessment of regional wind-solar energy resources based on MERRA reanalysis data. *Applied Energy*, 216, 172-182.





# Dynamic Modeling and Control of CO<sub>2</sub> Capture Systems with a Biphasic Solvent

Panagiotis Kazepidis<sup>a,b</sup>, Athanasios I. Papadopoulos<sup>a</sup>, Panos Seferlis<sup>a,b</sup>

<sup>a</sup>*Chemical Process & Energy Resources Institute (C.P.E.R.I.), Center for Research and Technology Hellas (CE.R.T.H.), P.O. Box 60361, 57001, Thessaloniki, Greece*

<sup>b</sup>*Department of Mechanical Engineering, Aristotle University of Thessaloniki, P.O. Box 484, 54124 Thessaloniki*  
*seferlis@auth.gr*

## Abstract

Biphasic solvents are receiving increased attention in post-combustion CO<sub>2</sub> capture systems as they reduce the regeneration energy requirements compared to conventional, single-phase solvents. The appearance of a second liquid phase in biphasic solvents results in a complex dynamic behavior of the capture process that has never been investigated. We develop for the first time a dynamic model for the novel biphasic solvent S<sub>1</sub>N/DMCA (N-Cyclohexyl-1,3-Propanediamine/ N,N -Dimethylcyclohexylamine) to investigate its behavior compared to the conventional MEA (monoethanolamine). The model accounts for both the mass and energy holdup in the liquid phase, whereas a control scheme with two proportional-integral (PI) controllers is used to compare the dynamic performance of the two solvents. For 10% increase in the total flue gas flow, the biphasic solvent is more robust, due to 2.5% and 3% lower change in the reboiler energy demand and solvent replenishing requirement than MEA. The required steam cost is also reduced by 3% compared to MEA.

**Keywords:** CO<sub>2</sub> capture, process control, phase-change solvents.

## 1. Introduction

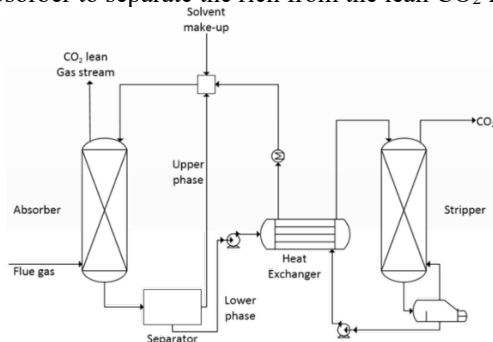
Solvent-based absorption/ desorption processes represent a mature technology for post-combustion CO<sub>2</sub> capture, but their wide deployment is prohibited by increased regeneration energy requirements that result in high costs. Such challenges are approached through the development of process schemes that enhance the capture and regeneration driving forces (Bui et al., 2018), of control schemes that enable efficient operation (Pátron and Ricardez-Sandoval, 2022) and of advanced solvents (Papadopoulos et al., 2021). In the latter case, phase-change or biphasic solvents represent a promising technology for the reduction of regeneration energy and capture process economics. Biphasic solvents form a second liquid phase upon reaction with CO<sub>2</sub> at the prevailing process conditions. Only the liquid phase with the higher content in CO<sub>2</sub> is directed to the stripper for solvent regeneration, enabling up to 51% lower energy requirements and 47% lower cost for the capture process than the conventional MEA (monoethanolamine) solvent (Kazepidis et al., 2021). These findings are important, but they have been attained through steady-state capture process optimization. The variability that is observed in industrial systems has detrimental effects on the capture process operation, hence a systematic assessment of the dynamic behavior and the subsequent design of efficient

control schemes are essential (Salvinder et al., 2019). However, such approaches have not yet been implemented in biphasic capture systems.

In the area of CO<sub>2</sub> capture, there are several control approaches that have been proposed for conventional (non-phase change) solvents. There are works that consider the reference MEA solvent in approaches such as proportional-integral-derivative (PID) control (Luu et al., 2015), model-predictive control (MPC) (Jung et al., 2021), linear MPC (Jung et al., 2020), non-linear MPC (Pátron and Ricardez-Sandoval, 2020), and advanced approaches that combine real-time optimization, non-linear MPC and rolling horizon estimation (Pátron and Ricardez-Sandoval, 2022). Fewer dynamic modeling approaches have been proposed that consider different solvents such as MEA, DEA (Diethanolamine), (Gáspár and Cormos, 2021; Damartzis et al., 2018) MDEA (Methyl-diethanolamine), AMP (2-Amino-2-methyl-1-propanol) (Gáspár and Cormos, 2021) and MPA (3-Amino-1-propanol) (Damartzis et al., 2018). Despite the growing interest in developing optimal control strategies for conventional solvent-based capture systems, the dynamics of CO<sub>2</sub> capture processes involving biphasic solvents remain unexplored. The presence of two liquid phases in the absorption column and the subsequent phase separator increase the complexity of the process unit interactions and of the dynamic behavior of the system. The aim of this work is to study the dynamic behavior of a biphasic CO<sub>2</sub> capture system and to develop a control strategy for the economic operation of the capture system under variable operating conditions.

## 2. Models and methods

The biphasic solvent mixture of S<sub>1</sub>N/DMCA (N-Cyclohexyl-1,3-Propanediamine/ N,N -Dimethylcyclohexylamine) proposed by Papadopoulos et al. (2021) is considered in this work with the capture process flowsheet shown in Figure 1. The key difference compared to a conventional capture process is that a liquid-liquid phase separator (decanter) is deployed after the absorber to separate the rich from the lean CO<sub>2</sub> liquid phases.



**Figure 1:** Basic Flowsheet diagram of VLLLE process for post combustion CO<sub>2</sub> capture with the S<sub>1</sub>N/DMCA mixture.

The developed dynamic model considers vapor-liquid-liquid equilibrium (VLLLE) along with material and energy balances for each phase. An orthogonal collocation on finite elements (OCFE) technique is employed to account for the variation of the concentration and temperature profiles along the height of the columns. The OCFE technique enables the formation of a compact in size, dynamic process model which facilitates greatly the numerical solution with good control of the approximation error (Damartzis et al., 2018). The OCFE technique considers the heights of the absorption and desorption columns as continuous variables and uses selected points defined as the roots of orthogonal Legendre polynomials called collocation points, in which mass and energy balances are exactly

satisfied. Lagrange interpolation polynomials are used within the finite elements. Dividing the height of the columns into finite elements allows problems with steep profile changes to be accurately represented. This is very important during dynamic transitions where steep fronts in concentration and temperature may move along the columns.

The equations describing the mass balance of the liquid phases in a three-phase dynamic model are the following (Damartzis and Seferlis, 2010):

$$\frac{dm_i^I(s_j)}{dt} = \tilde{L}_t^I(s_j - 1) - \tilde{L}_t^I(s_j) + \varphi_{i,j}^I(s_j) \quad i = 1, \dots, NC, \quad j = 1, \dots, n \quad (1)$$

$$\frac{dm_i^{II}(s_j)}{dt} = \tilde{L}_t^{II}(s_j - 1) - \tilde{L}_t^{II}(s_j) + \varphi_{i,j}^{II}(s_j) \quad i = 1, \dots, NC, \quad j = 1, \dots, n \quad (2)$$

where  $s_j$  is a column position coordinate,  $L_t$  is the liquid flow of component  $i$  in one of the liquid phases ( $I, II$ ) and  $\varphi$  is the holdup. Similarly, the energy balance is formulated as follows:

$$\begin{aligned} \frac{dU(s_j)}{dt} = & \tilde{L}_t^I(s_j - 1)\tilde{H}^{LI}(s_j - 1) + \tilde{L}_t^{II}(s_j - 1)\tilde{H}^{LII}(s_j - 1) - \tilde{L}_t^I(s_j)\tilde{H}^{LI}(s_j) - \\ & \tilde{L}_t^{II}(s_j)\tilde{H}^{LII}(s_j) + \tilde{G}_t(s_j + 1)\tilde{H}^G(s_j + 1) - \tilde{G}_t(s_j)\tilde{H}^G(s_j) + Q(s_j) + \\ & \varphi_{t,j}^I(s_j)\tilde{H}^{LI}(s_j) + \varphi_{t,j}^{II}(s_j)\tilde{H}^{LII}(s_j), \quad j = 1, \dots, n \end{aligned} \quad (3)$$

where  $L_t$  is the total liquid flow in one of the liquid phases ( $I, II$ ),  $G_t$  is the total vapor flow,  $H$  is the enthalpy of each phase,  $Q$  is the heat losses and  $\varphi_t$  is the total holdup. The dynamic behavior is introduced into the three-phase model by the mass holdup term of the two liquid phases and is calculated by the following equation:

$$\frac{dm_i^I(s_j)}{dt} + \frac{dm_i^{II}(s_j)}{dt} = \frac{\rho_{mix}^I(s_j)a^{intV}(s_j)\delta^I \tilde{L}_t^I(s_j)}{MW_{mix}^I(s_j) \tilde{L}_t^I(s_j)} + \frac{\rho_{mix}^{II}(s_j)a^{intV}(s_j)\delta^{II} \tilde{L}_t^{II}(s_j)}{MW_{mix}^{II}(s_j) \tilde{L}_t^{II}(s_j)} \quad (4)$$

where  $\rho$  is the density of each liquid phase,  $a^{int}$  is the liquid phase interphase,  $V$  is the control volume in every collocation point,  $\delta$  is the film thickness of each liquid phase on the packing and  $MW_{mix}$  is the molecular weight of each phase. The control volume is calculated as the volume contained between two collocation points on the separation columns, centered at the location of the collocation points. In the case of the phase separator and reboiler, this volume is equal to the bottom control volume of the absorption and regeneration column, respectively. Note that the vapor phase mass balance of Eq. (1) and (2) remains the same as in the steady-state VLE model because it is assumed that there is no holdup of the gas phase. Finally, the mass density of each liquid phase is calculated using the composition.

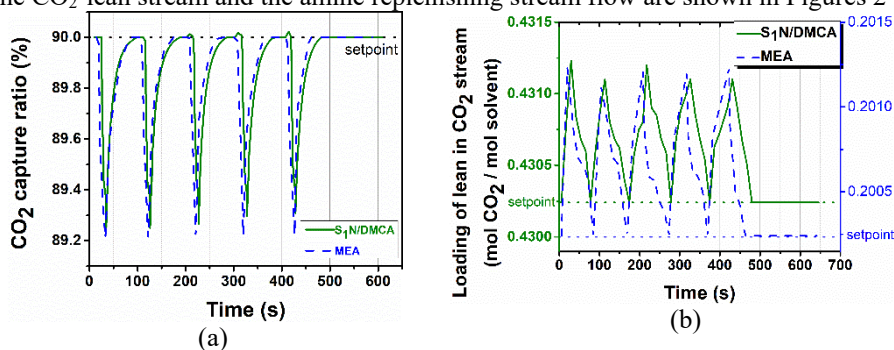
### 3. Implementation

The proposed dynamic model is implemented on the novel biphasic mixture of S<sub>1</sub>N/DMCA (Papadopoulos et al., 2021). This mixture forms two liquid phases directly upon reaction with CO<sub>2</sub>, it has high CO<sub>2</sub> capture capacity and releases the captured CO<sub>2</sub> at lower temperature than MEA. The flue gas stream used in the implementation is based on a quicklime plant and consists of 12 vol.% CO<sub>2</sub>, has a temperature of 298.15 K and a total flow equal to 320.22 mol/s. To evaluate the performance of the S<sub>1</sub>N/DMCA mixture a 10% increase in the total flow of the system is imposed in the dynamic simulation. A control scheme is designed for the capture system with MEA and S<sub>1</sub>N/DMCA under variable operating conditions. It consists of two PI control loops that aim to keep the CO<sub>2</sub> capture ratio and the loading (mol CO<sub>2</sub> / mol amine) of the CO<sub>2</sub>-lean stream at their

reference values. The CO<sub>2</sub> capture ratio is controlled by the manipulation of the reboiler thermal duty, whereas the loading of the CO<sub>2</sub>-lean stream is controlled by the manipulation of the solvent replenishing (make-up) stream flow. Disturbances are imposed on the system through five successive step changes in the total flue gas flow, each representing a 2 % increase of the nominal operating point. The disturbances begin at 20 s and are repeated every 100 s. The tuning parameters of both controllers for the two cases are calculated by the response curve method using a Ziegler-Nichols correlation. For the CO<sub>2</sub> capture ratio control loop we used  $K_c = 1.147$ ,  $\tau_i = 0.12$  and  $K_c = 1.151$ ,  $\tau_i = 0.22$  for the MEA and S<sub>1</sub>N/DMCA cases. For the lean loading control loop we used  $K_c = 1.25$ ,  $\tau_i = 1.5$  and  $K_c = 0.83$ ,  $\tau_i = 1.1$  for the MEA and S<sub>1</sub>N/DMCA cases. The setpoint values for the lean stream loading are 0.20 and 0.43 mol CO<sub>2</sub> / mol amine for the MEA and S<sub>1</sub>N/DMCA cases, respectively, as obtained from the steady-state optimization. In both cases the CO<sub>2</sub> capture ratio is set to 90 %. The values of  $a^{int}$  and  $\delta$  are considered constant throughout the process ( $a^{int} = 250 \text{ m}^2 \cdot \text{m}^{-3}$  and  $\delta = 0.0005 \text{ m}$ , Rocha et al., 1993). The system of differential-algebraic equations was solved using the solver implemented in DASSL (Petzold, 1982).

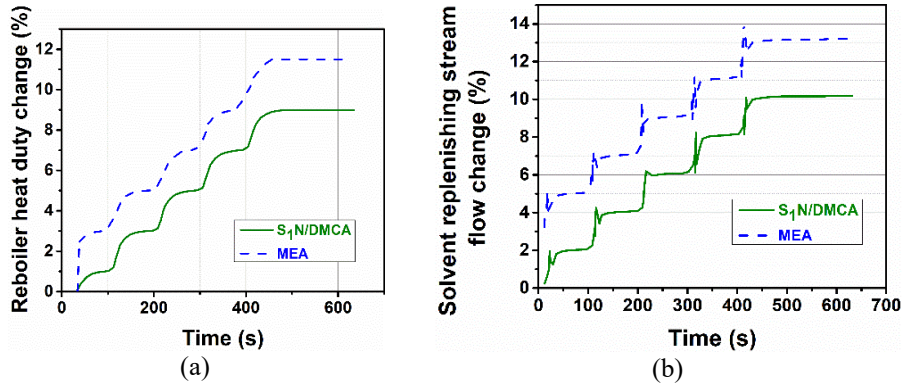
#### 4. Results and discussion

The dynamic response of the CO<sub>2</sub> capture ratio, the reboiler thermal duty, the loading of the CO<sub>2</sub>-lean stream and the amine replenishing stream flow are shown in Figures 2-4.



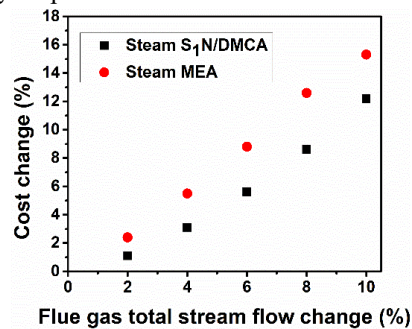
**Figure 2:** Closed loop response of a) the CO<sub>2</sub> capture ratio, and b) the CO<sub>2</sub>-lean stream loading.

According to the results of Figures 2 and 3 for S<sub>1</sub>N/DMCA, an increase of about 10% in the amine replenishing stream flow and 9% in the reboiler heat duty are sufficient to compensate the effects of the studied disturbance on the controlled variables. S<sub>1</sub>N/DMCA has a clear advantage compared to MEA, as it requires less energy to compensate for the effect of the disturbance, as shown in Figure 3 (a). This is mainly attributed to the formation of the second liquid phase in S<sub>1</sub>N/DMCA, which reduces the solvent flow rate by 70% compared to MEA (Kazepidis et al., 2021). In the case of MEA, the response is faster by 9% (settling time is smaller by 9 s) and the amine replenishing stream flow increases by 13%, while the energy supplied to the reboiler increases by 11.5% (Figure 3). This response is expected as MEA has been reported to have faster kinetics than most biphasic solvents (Zhang, 2013). Also, the system appears to be quite responsive to changes in the reboiler heat duty in both solvent cases, because it has only a 9% change from the optimum state value for the S<sub>1</sub>N/DMCA mixture and 11.5% for the MEA case (Figure 3a). In contrast, the amine replenishing stream changes more than 10% to be able to compensate for the disturbance in the case of the S<sub>1</sub>N/DMCA mixture and 13% in the case of MEA (Figure 3b).



**Figure 3:** Closed loop performance of a) the reboiler heat duty and b) the solvent replenishing stream flow.

An important parameter is the economic effect of the disturbance in the system. Cost calculations in this study were based on Kazepidis et al. (2021). Figure 4 shows the steam costs change with respect to the total flue gas flow change. While energy needs increase by about 2% with each step change, the steam cost has a higher rate of increase. Thus, the total increase of steam cost after all five step changes is 12.2% for S<sub>1</sub>N/DMCA which is greater than the change in reboiler energy needs (9%). In the case of MEA, the increase in the steam cost is equal to 15.3%. Regarding the solvent replenishing stream effects on economic performance, S<sub>1</sub>N/DMCA has 6 times lower solvent replenishing requirements, i.e., 0.011 mol/s for S<sub>1</sub>N/DMCA and 0.066 mol/s for MEA (Kazepidis et al., 2021). This is due to the overall lower flows of S<sub>1</sub>N/DMCA to attain the desired capture rate and to its lower volatility compared to MEA. Under disturbance, the percentage change in the flowrate of the replenishing stream of S<sub>1</sub>N/DMCA is lower by 3 percentage points than that of MEA. Assuming that the solvent price of S<sub>1</sub>N/DMCA is comparable to that of MEA, the latter is greatly outperformed in this indicator.



**Figure 4:** Steam cost change with respect to the flue gas total stream flow change.

## 5. Conclusions

The process is controlled efficiently by the proposed scheme, as the CO<sub>2</sub> capture ratio is reduced less than 1% from the nominal point, under the influence of disturbances. MEA has a faster response, but is also more sensitive to the disturbance by having 11.5% and 13% increase in the reboiler demands and solvent replenishing stream. S<sub>1</sub>N/DMCA has 9% slower response but is more robust as it has only 9% and 10% increase in the reboiler demands and solvent replenishing stream. The biphasic solvent has a better economic

performance than MEA, due to 3% lower steam cost increase and 6 times lower solvent replenishing needs. These new findings suggest that biphasic solvents enhance the reliability and stability of solvent-based CO<sub>2</sub> capture systems, making it a feasible option for widespread use. The dynamic performance of such systems is essential as variability in the flue gas conditions is large and the capture targets may vary for optimal operation in most industrial applications.

### Acknowledgements

This research has been financed by the European Union and Greek National Funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call Research – Create – Innovate, Project T2EAK-01911

### References

- M. Bui, C.S. Adjiman, A. Bardow, E.J. Anthony, A. Boston, S. Brown, P.S. Fennell, S. Fuss, A. Galindo, L.A. Hackett, and J.P. Hallett, 2018, Carbon capture and storage (CCS): the way forward, *Energy & Environmental Science*, 11, 5, pp. 1062-1176.
- G.D. Patrón, and L. Ricardez-Sandoval, 2022, An integrated real-time optimization, control, and estimation scheme for post-combustion CO<sub>2</sub> capture, *Applied Energy*, 308, p. 118302.
- A.I. Papadopoulos, F.A. Pedromo, F. Tzirakis, G. Shavaliyeva, I. Tsivintzelis, P. Kazepidis, E. Nessi, S. Papadokostantakis, P. Seferlis, A. Galindo, G. Jackson, C.S. Adjiman, 2021, Molecular engineering of sustainable phase-change solvents: From digital design to scaling-up for CO<sub>2</sub> capture, *Chemical Engineering Journal*, 420, p. 127624.
- P. Kazepidis, A.I. Papadopoulos, F. Tzirakis, and P. Seferlis, 2021, Optimum design of industrial post-combustion CO<sub>2</sub> capture processes using phase-change solvents, *Chemical Engineering Research and Design*, 175, pp. 209-222.
- K.M.S. Salvinder, H. Zabiri, S.A. Taqvi, M. Ramasamy, F. Isa, N.E.M. Rozali, H. Suleman, A. Maulud and A.M. Shariff, 2019, An overview on control strategies for CO<sub>2</sub> capture using absorption/stripping system. *Chemical Engineering Research and Design*, 147, pp. 319-337.
- M.T. Luu, N.A. Manaf, and A. Abbas, 2015, Dynamic modelling and control strategies for flexible operation of amine-based post-combustion CO<sub>2</sub> capture systems, *International Journal of Greenhouse Gas Control*, 39, pp. 377-389.
- H. Jung, S. Heo, and J.H. Lee, 2021, Model predictive control for amine-based CO<sub>2</sub> capture process with advanced flash stripper, *Control Engineering Practice*, 114, p. 104885.
- H. Jung, D. Im, S. Heo, B. Kim, and J.H. Lee, 2020, Dynamic analysis and linear model predictive control for operational flexibility of post-combustion CO<sub>2</sub> capture processes. *Computers & Chemical Engineering*, 140, p. 106968.
- G.D. Patrón, and L. Ricardez-Sandoval, 2020, A robust nonlinear model predictive controller for a post-combustion CO<sub>2</sub> capture absorber unit, *Fuel*, 265, p. 116932.
- J. Gáspár, A.-M. Cormos, 2012, Dynamic modeling and absorption capacity assessment of CO<sub>2</sub> capture process, *International Journal of Greenhouse Gas Control*, 8, pp. 45–55.
- T. Damartzis, A.I. Papadopoulos and P. Seferlis, 2018, Solvent effects on design with operability considerations in post-combustion CO<sub>2</sub> capture plants, *Chemical Engineering Research and Design*, 131, pp. 414-429.
- T. Damartzis and P. Seferlis, 2010, Optimal design of staged three-phase reactive distillation columns using nonequilibrium and orthogonal collocation models, *Industrial and Engineering Chemistry Research*, 49, 7, pp. 3275-3285.
- J.A. Rocha, J.L. Bravo and J.F. Fair, 1993, Distillation columns containing structured packings: A comprehensive model for their performance. 1. hydraulic models, *Industrial & Engineering Chemistry Research*, 32, pp. 641-651.
- L. Petzold, 1982, A description of DASSL: a differential/algebraic equation system solver, *Proc. IMACS World Congress*, pp. 430-432.
- J. Zhang, 2013, Study on CO<sub>2</sub> capture using thermomorphic biphasic solvents with energy-efficient regeneration, PhD. Thesis, Technical University Dortmund, Dortmund.

# Development of integrated liquid air energy storage systems based on air separation units using waste energy from power plant: A case study of South Korea

Jaerak Ko, Seolji Nam, Soonho Hwangbo\*

*Department of Chemical Engineering, Gyeongsang National University 501, Jinjudaero, Jinju-si, Gyeongsangnam-do, 52828 Korea*

*\*s.hwangbo@gnu.ac.kr*

## Abstracts

Power consumption has been rising gradually as technology has developed. It is necessary that power generation using renewable energy that does rarely emit carbon. An air separation unit that separates air into nitrogen, oxygen, and argon is used in a variety of industries, including steel mills and chemical plants, which demand pure gas. In order to feed pure gas to the industrial complex utilizing stored air and produce power at the same time, the air separation unit is combined with a liquid air energy storage system. This study used a process simulation program to develop a model aiming to integrate an air separation unit and a liquid air energy storage system using waste energy produced by a liquefied natural gas power plant. According to a Korean case study, the leveled cost of electricity of the integrated process is \$320/MWh, which is about \$51/MWh less than the leveled cost of electricity of the current liquid air energy storage system. Through this study, it is expected that carbon emissions can be significantly decreased by supplying the Korea petrochemical complex with the pure gas it needs to meet its 5.6% mandated renewable energy supply power need.

**Keywords:** Liquid air energy storage system, Process integration, Techno-economic and environmental assessment, Utilizing waste energy

## 1. Introduction

Power consumption in the industry has been rising rapidly as technology has developed. Fossil fuel-generating systems, such as coal-fired power generation, emit a large amount of carbon dioxide into the air (Chehade and Dincer, 2021). Global carbon emission reductions were discussed during the 2015 Paris Climate Agreement. As a result, the development of a carbon-free power generation system using renewable energy is necessary to replace the current power generating system that emits a small quantity of carbon. Since the production of renewable energy is unpredictable and irregular, it is challenging to consistently produce energy in response to demand (Nam et al., 2020). The excess energy is stored by integrating an energy storage system (ESS), and the stored energy may later be used if it needs to make up for the limitations of the current renewable energy. Typically, the general ESSs take the form of lithium-ion batteries. However, a considerable quantity of energy cannot be stored due to the high cost of system installation and the limited capacity of energy storage (Koohi Fayegh, 2020). As a next-generation ESS to overcome this, liquid air energy storage (LAES) systems are gaining appeal. LAES systems show high energy density regardless of the installation location (Borri et al., 2021). A 20MW LAES system has been substantiated by Highview power,



and several similar studies are being conducted (Rarely, 2017; Morgan et al., 2015). The challenge is that setting up LAES system is expensive and difficult to apply in reality. However, some options, such as the use of liquefied natural gas (LNG) cold heat or solar heat, are currently being intensively researched in attempt to address these drawbacks (Qi et al., 2020). This study aims to increase economic feasibility by integrating an air separation unit (ASU) process to separate air into nitrogen, oxygen, and argon, release energy, supply power to petrochemical complex, and supply industrial gas to each customer for overcoming the shortcomings of the recent LAES system. Utilizing waste heat and LNG cold energy from LNG power plants near to petrochemical complex also contributed to improving energy efficiency. Levelized cost of electricity (LCOE) was estimated in a case study of South Korea, and the contribution of Korea's carbon neutral policy was determined through the calculation of the required supply of renewable energy.

## 2. Method

### 2.1. Design of ASES system

Figure 1 shows the power and industrial gas supply network in integration with the LNG power plant, the petrochemical complex, and an air separation energy storage (ASES) system. The ASES system consists of a charging process and discharging process. During charging, power is sourced from low price power grid, and ASU is used to separate and liquefy the air. Charging process combines an air separation step in which gaseous air is liquefied at cryogenic temperature and separated into nitrogen, oxygen, and argon with a LAES system that uses waste heat from power plants to vaporize it before altering it into electricity through turbines. Charging process that uses LNG cold energy to get energy of air separated at a cryogenic temperature includes a compressor, a cold box, a low pressure column, a high pressure column, and an argon column for air separation. During discharging, liquid air that was stored when power demand was at its peak is vaporized to produce electricity through turbines. Through a waste heat source produced in a power plant, discharging process evaporates liquid air that is separated and stored, and turbines subsequently use the energy to produce electricity. Then, pipelines are used to deliver nitrogen, oxygen, and argon to demand sources located within the petrochemical complex. Discharging process consists of cryogenic storage tanks, heat exchangers for vaporization, and turbines.

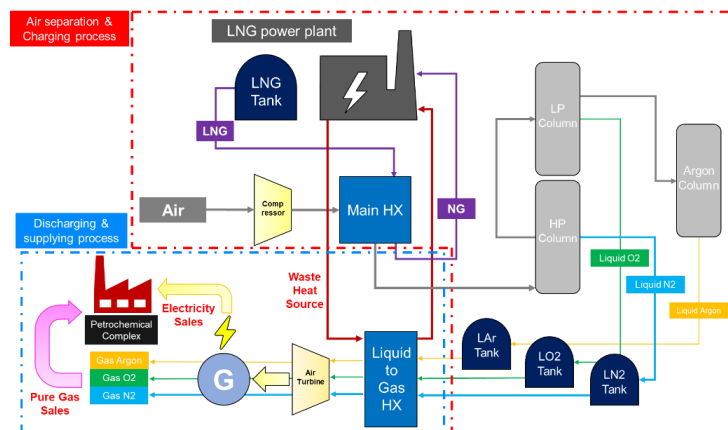


Figure 1. Conceptual design of ASES system for energy charging and discharging via utilizing cold energy and waste heat from LNG power plant.

Table 1. General assumptions which are used in process simulation.

Parameter	Value	Unit
Number of stages of HP column	65	N/A
Number of stages of LP column	80	N/A
Number of stages of Ar column	70	N/A
Expansion ratio of the air turbine	0.3	N/A
Isentropic efficiency of air turbine	90	%
Isentropic efficiency of compressor	98	%
Isentropic efficiency of pumps	70	%
Temperature of dry air	295.0	K
Temperature of inlet LNG	110.5	K
Temperature of outlet LNG	288.8	K

Table 1 is presented the important parameters used in simulation. Number of stages of HP, LP and Ar column are 65, 80, 70. The expansion ratio of the air turbine is 0.3 and Isentropic efficiency of the air turbine, compressor, pumps are 90%, 98% and 70%, which were obtained from current studies on LAES systems. Tray by tray modelling was used to replicate the distillation column. Peng Robinson's equation of state was used for simulation of this system.

### 2.2. Economic analysis

Economic analysis method for the ASES system was choose total revenue requirement (TRR) method based on The Electric Power Research Institute (EPRI) used method (Ko et al., 2018). By calculating direct cost (DC), indirect cost (IC), and other outlays (OO) based on purchased equipment cost (PEC), this method determines the total capital investment (TCI). The cost of revenue all outlays and collecting them for system operation is the recovery cost per year. TRR is made up of the total of carrying charges and expenses. According to TCI, carrying charges are estimated as sums that include capital recovery, return on equity, return on debt, income taxes, other taxes and insurance. To evaluate the economy, the following economic indicators were used: 1) The annual inflation rate was 2.5% 2) An 8% discount rate 3) The equity to debt ratio was 50%. 4) The interest rate was 8%. 5) The 30-year lifespan of the plant 6) The system operated for 8,000 hours (Zakeri et al., 2015).

## 3. A case study

### 3.1. ASES system in Daesan petrochemical complex

Daesan petrochemical complex is in Seosan-si, Chungcheongnam-do, Korea, with 10 large petrochemical companies. The ASES system is in Dangjin, Chungcheongnam-do, Korea, and the facility capacity air into the system at a flow rate of 84,375 kg/h. Figure 2(A) shows the site for the construction of the ASES system in integration with the Dangjin LNG power plant. Figure 2(B) shows a pipeline for supplying nitrogen, oxygen, and argon to the Daesan petrochemical complex. The length of the pipeline from the ASES system to the Daesan petrochemical complex is about 15 km.

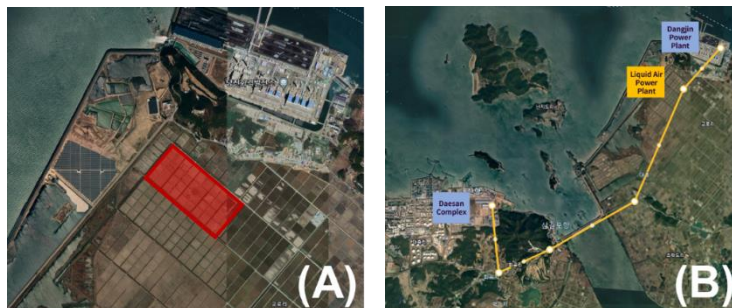


Figure 2. (A) A land decision of ASES system. (B) A pipeline connected from ASES system to chemical plant in Daesan petrochemical complex.

## 4. Results and discussion

### 4.1. Simulation results of the ASES system

Figure 3 shows the process simulation results of the ASES system in the case study. AVEVA Pro/ii simulation is used for process simulation. The mole fractions for the major streams of three main processes are displayed, and the outputs are as follows. The air was separated into nitrogen, oxygen, and argon through the HP, LP, and Argon columns, respectively. Stream A is an inflow of air from the outside and consists of a general composition of air. Stream B is compressed air from main HX and is similarly the composition of general air. Stream C is the stream where high purity nitrogen flows into the turbine, a 0.9995 mole fraction of  $N_2$ . Stream D is the result of the Argon column, a 0.99 mole fraction of Ar. Stream E is the result of oxygen flow into the turbine, and a 0.99 mole fraction of  $O_2$ . Stream F is the result of  $N_2$  purge gas and a 0.994 mole fraction of  $N_2$ .

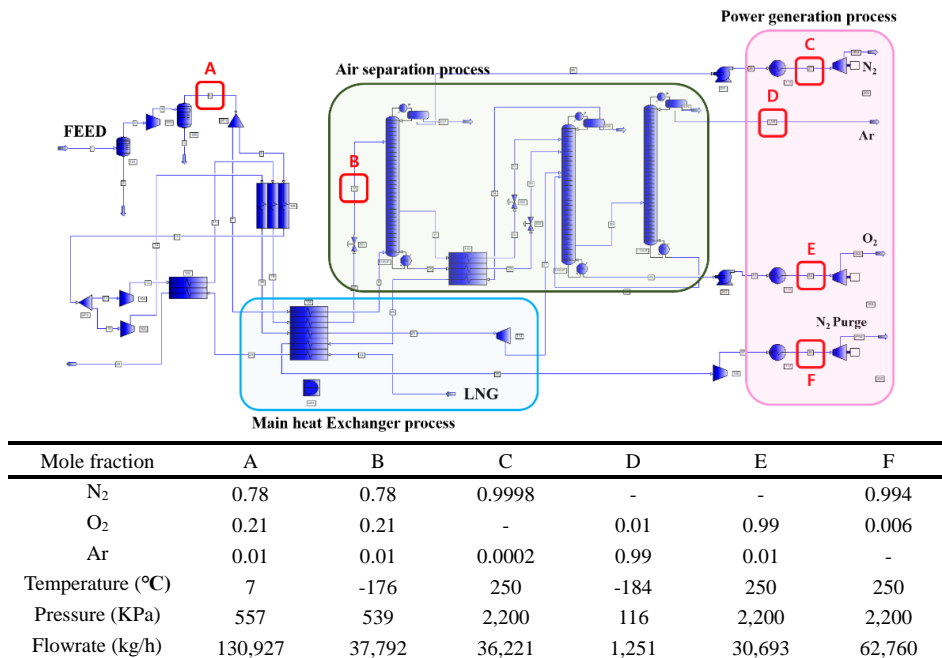


Figure 3. Simulation results of the air separation unit with power generation systems using waste energy.

By simulating ASES system with PRO/II simulation, we can estimate how much electricity is generated during the process of generating power. Through waste heat from power plant, the liquefied air that is stored in the tank heats up to 250°C and generates energy via turbine. 3,224 kW of electricity are produced by using N<sub>2</sub> (Figure 3(stream C)). 2,403 kW of electricity is generated by using oxygen as an electricity generator (Figure 3(stream E)). Nitrogen purge gas will be used to generate 5,584 kW of electricity (Figure 3(stream F)). Thus, we assumed the discharging time per day is 8 hours and the total capacity of generated power is 11,211 kW, the total generated power per day is approximately 90 MWh.

#### 4.2. Economic analysis

To estimate the TCI of the ASES system, we make several economic assumptions. PEC is used to estimate direct cost, indirect cost, and other costs. TCI is the sum of the three sections. TCI was estimated using parameter values derived from experience and appeared to US\$ 167,398,686,952. TRR and LCOE were estimated using this data. Figure 4 shows the comparison using the results of previous research to compare LCOE with other ESS (Zakeri et al., 2015). The carrying cost ratio determined in this study was used to add capital recovery, return on equity, return on debt, income taxes, other taxes and insurance, which were not taken into account when calculating LCOE in other studies, to other ESS. The results showed that using cold energy and waste heat and selling the by-products resulted in a cheaper LCOE than the current LAES system of US\$ 51/MWh.

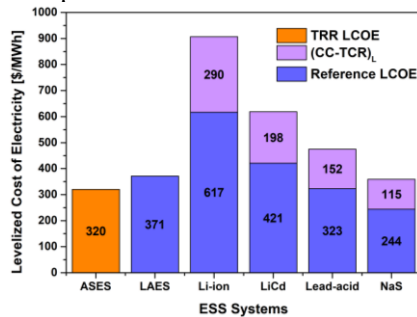


Figure 4. LCOE comparison results of ASES and other ESS systems.

#### 4.3. Environment assessment

In this study, we estimated the percentage of overall renewable energy that must be supplied to Daesan petrochemical complex that can be supplied through this process. The Daesan petrochemical complex will require 1,564,035 MWh/yr of renewable energy, and the ASES system has the capacity to provide an additional 88,000 MWh/yr of electricity. 345.96 kg of CO<sub>2</sub> are released into the atmosphere for each 1 MWh of electricity produced using sub-bituminous coal (the most of using fuel for generating in Korea) as a fossil fuel. So, this system can significantly reduce 30,444 ton of CO<sub>2</sub> of year and equates to 5.6% of the necessary supply of renewable energy. Specific environmental assessment of the proposed energy system will be implemented using life cycle assessment as future perspectives.

## 5. Conclusions

This study aims to 1) develop integrated energy systems that enable generating electrical energy from LAES and producing industrial utilities from ASU and 2) conduct techno-economic assessment to find out the feasibility of the developed model by applying to Daesan petrochemical complex. Waste heat and cold heat generated by LNG power plants were used for air vaporization and liquefaction to boost energy efficiency. To reduce power costs, liquefied air is stored during the late-night hours when energy use is low, and electricity is generated utilizing the air stored during the power peak period to supply power to the petrochemical complex. For the case study in Korea, economic feasibility may be awarded a good rating based on an in-depth economic evaluation that takes into account the carrying cost, and the economic evaluation reveals a reduced equivalent power cost of US\$ 51/MWh lower equivalent power generating cost than the present LAES system. It is anticipated that it would be employed as a novel energy storage system for sustainable power generation in Korea.

## References

- Borri, E., Tafone, A., Romagnoli, A., & Comodi, G., 2021, A review on liquid air energy storage: History, state of the art and recent developments. *Renewable and Sustainable Energy Reviews*, 137, 110572.
- Chehade, G., & Dincer, I., 2021, Progress in green ammonia production as potential carbon-free fuel. *Fuel*, 299, 120845.
- Ko, A., Park, S., Kim, J. Y., & Cha, J. M., 2018, Development and reliability optimization of economic analysis module for power generation system from industrial waste heat recovery. *Journal of Energy Engineering*, 27(4), 50-63.
- Koohi-Fayegh, S., & Rosen, M. A., 2020, A review of energy storage types, applications and recent developments. *Journal of Energy Storage*, 27, 101047.
- Morgan, R., Nelmes, S., Gibson, E., & Brett, G., 2015, Liquid air energy storage—analysis and first results from a pilot scale demonstration plant. *Applied energy*, 137, 845-853.
- Nam, K., Hwangbo, S., & Yoo, C., 2020, A deep learning-based forecasting model for renewable energy scenarios to guide sustainable energy policy: A case study of Korea. *Renewable and Sustainable Energy Reviews*, 122, 109725.
- Qi, M., Park, J., Kim, J., Lee, I., & Moon, I., 2020, Advanced integration of LNG regasification power plant with liquid air energy storage: Enhancements in flexibility, safety, and power generation. *Applied energy*, 269, 115049.
- Zakeri, B., & Syri, S., 2015, Electrical energy storage systems: A comparative life cycle cost analysis. *Renewable and sustainable energy reviews*, 42, 569-596.

# A systematic method for performing pinch analysis of the Liquid Air Energy Storage (LAES) process

Vuppanapalli Chaitanya<sup>a</sup>, S. Narasimhan<sup>a\*</sup>, G. Venkatarathnam<sup>b</sup>

<sup>a</sup> *Department of Chemical Engineering, Indian Institute of Technology Madras, Chennai 600036, India*

<sup>b</sup> *Department of Mechanical Engineering, Indian Institute of Technology Madras, Chennai 600036, India*

\*Corresponding author at Department of Chemical Engineering, Indian Institute of Technology Madras, India; E-mail address: [naras@iitm.ac.in](mailto:naras@iitm.ac.in)

## Abstract

In process design, the insights obtained by applying pinch technology have played an important role in maximizing energy efficiency through energy integration [1]. Pinch technology exclusively addresses temperature variations in process streams resulting from indirect heat exchange. The analysis does not take into account temperature changes in process streams induced by changes in stream pressure.

In this work, a systematic procedure for applying pinch analysis to the Liquid Air Energy Storage (LAES) process is proposed. In this process, air undergoes significant pressure changes, which results in phase changes as well as wide variations in the specific heat capacity. Since the air temperature varies from well above ambient to well below ambient conditions, multiple minimum approach temperature specifications have to be imposed. A parameterized version of the Grand Composite Curves (GCCs) is proposed for pinch analysis that takes into account for all of these special features. The parameterized GCCs are used to identify the feasible design space for a LAES process.

**Keywords:** Pinch analysis, Pressure change, Liquid air energy storage, Grand Composite Curves, Feasible design space.

## 1 INTRODUCTION

The world is rapidly shifting to the use of renewable energy sources to replace fossil fuels for power generation. However, the intermittent nature of renewable energy and the significant difference between its availability and demand time periods are two challenges that need to be tackled. The use of energy storage technologies is one approach to address the above challenges. Liquid Air Energy Storage (LAES) is a thermal energy storage system that has attracted considerable attention recently due to its advantages, such as its high energy density, long life span, no geographical constraints, and environmental friendliness [2]. The main limitation of LAES systems is their low round-trip efficiency (RTE) of around 50–60% for large-scale systems. Several different process designs have been proposed by different researchers to improve the RTE. For example, a part of the cold utility requirements can be provided by cold air or by integrating an LNG regasification process with the LAES process [2]. The proposed designs have to be simulated and optimized to determine the maximum possible RTE gains. In process

design, the insights obtained by applying pinch technology have played an important role in maximizing energy efficiency through energy integration [1]. Using pinch technology, it is possible to derive bounds on the minimum utility requirements prior to design. It also provides useful information on the utility design (the quantity and temperature levels at which external utilities can be provided) and how best to integrate compressors and turbines with the thermal process. However, pinch technology only deals with temperature changes in process streams caused by heat exchange. Temperature change in a process stream induced by a change in the stream pressure is not considered in the analysis. Aspelund et al. [3] proposed an expanded pinch analysis based on exergy analysis for taking into account pressure-induced temperature changes. Such a technique is required to analyze processes such as LAES, air separation, and natural gas liquefaction. The LAES process poses the following additional challenges for pinch analysis:

- Use of a non-global minimum temperature approach in the LAES process.
- A large operating temperature range (100-600 K) results in significant variation in the fluid properties, specifically heat capacity.
- The charging and the discharging sections are thermally coupled with thermal energy storage media, which forces thermal energy storage media to operate in a cyclic mode.

In this work, we propose a systematic approach for applying pinch analysis that deals with all of the above challenges. Based on a parametric version of the Grand Composite Curve (GCC), we quantitatively identify the feasible operating regimes and minimum utility requirements. The GCC can potentially be used for identifying alternative choices of cold storage utilities and the potential for integrating refrigeration cycles with the process.

## 2 PINCH ANALYSIS OF LAES PROCESS

### 2.1 Process description

A typical standalone LAES process is shown in Fig. 1. It includes a charging section (air liquefaction) and discharging section (power recovery). In the charging section, power from a renewable energy source is used to liquefy air and store it in insulated storage tanks. The heat generated from air compression is recovered into a high-temperature energy storage system (thermal oil). When energy is required, liquid air from the storage tank is heated and expanded in the discharging section to generate power. The cold energy of the liquid air is recovered and stored in a low-temperature energy storage system (propane and methanol) to be utilized in the next charging cycle.

### 2.2 Process streams for pinch analysis

In the LAES process, air is compressed and later expanded to produce liquid air. The change in pressure results in phase changes and wide variations in the specific heat capacity. Furthermore, the fraction of liquid air produced depends on the process conditions. Thus, it is not possible to construct a stream table and apply the standard pinch analysis of this process. In order to construct the stream table, the operating pressures ( $P_C$  and  $P_D$ ) of the charging and discharging sections have to be chosen. The fraction of air that is liquefied ( $Y$ ) must also be chosen because this determines the vapor flow rate (stream 7) from the phase separator. The choice of these operating parameters have a significant effect on the utility requirements, and have to be parametrically varied

to study their impact. The time periods of operation of the discharge and charging sections are generally specified based on the application requirements, and are assumed to be fixed. Given these specifications, the stream table can be constructed for unit inlet air flow rate chosen as a basis. Any process stream that undergoes pressure changes is treated as multiple segments by considering the segments before and after pressure change as separate streams in the analysis. For example, in Fig. 1, segments from the inlet air between states 1-2, 3-4, and 6-6a (compression and expansion operations) are deleted, and the remaining segments 2-3, 4-5, and 7-8 (heat exchanger operations) are included as distinct streams. For different outlet pressure of compressors, parametric GCC are constructed that are further used in the analysis. The thermal oil that recovers the heat of compression and cold utilities methanol and propane are also included as process streams since they are operated in a cyclic mode. Propane and methanol are assumed to be liquids operating between their normal boiling point and triple point temperatures, while thermal oil is assumed to operate between ambient temperature and its maximum allowable temperature of 603 K. It should be noted that the energy recovered or supplied during the charging cycle by these utilities must match exactly with the energy supplied/recovered during the discharge cycle (to avoid the use of other external utilities). Since the operating time periods for the charging and discharge cycles are assumed to be fixed, the following constraints on the flow rate of methanol and propane can be derived.

$$Y = \frac{\dot{m}_{11}}{\dot{m}_9} = \frac{\dot{m}_{6C1}}{\dot{m}_{3C1}} = \frac{\dot{m}_{6C2}}{\dot{m}_{3C2}} = \frac{t_c}{t_d} \quad (1)$$

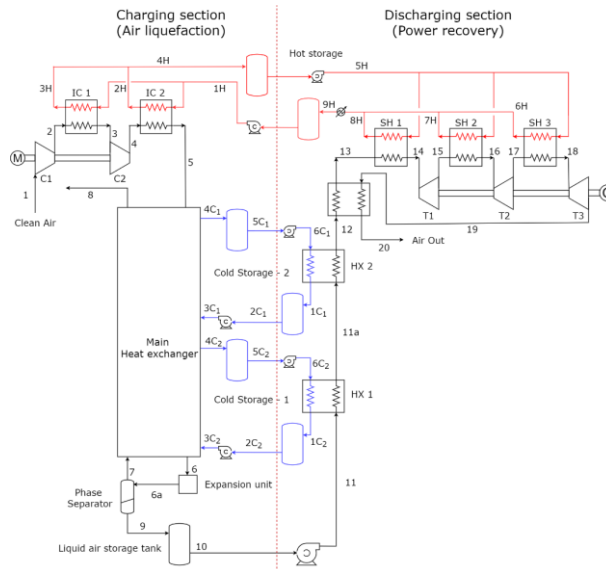


Figure 1: Typical process configuration of LAES

From the above equations, the flow rates of all streams can be determined for the given choice of liquid air yield. The specific enthalpies of the streams depend both on the temperature and pressure. Based on the chosen pressures ( $P_c$  and  $P_D$ ), a thermodynamic property estimation package is used to evaluate the specific enthalpy over the entire range of operating range. The GCC for the charging section can now be constructed for every choice of the operating variables  $P_c$ ,  $P_D$ , and  $Y$ . We refer to this GCC curve as a



parameterized GCC, since it depends on the chosen values of the operating variables. By analysing the parameterized GCCs over a wide range of the three operating parameters the feasible operating regime and maximum RTE that can be obtained, prior to design. The stream table constructed for the choice of  $P_C = 100$  bar,  $P_D = 150$  bar, and  $Y = 0.7$  is shown in Table 1. All thermodynamic properties are computed using the Peng Robinson equation of state. The specific heat capacities of all streams (air, propane, and methanol) change significantly over the operating temperature range. Therefore, the overall temperature range is divided into differential intervals of  $1^\circ\text{C}$  for the purpose of constructing the GCC.

Furthermore, a single minimum temperature approach (MTA) may not be appropriate, since the process streams operate well above and well below ambient conditions. It has been determined that the lower the minimum temperature approach below ambient temperature, the lower the temperature to which air can be cooled before being expanded in the cryo-turbine. This leads to a higher yield of liquid air and improved performance in the process. An assumption has been made that the MTA will be a minimum of 10 degrees above ambient temperature and 5 degrees below ambient temperature. The method presented by Robin Smith [1] is used to deal with the dual MTA specification.

### 3 RESULTS AND DISCUSSION

The stream table for a specific combination of  $P_C$ ,  $P_D$ , and  $Y$  is constructed and shown in Table 1. The GCC plot for this combination of operating parameters is presented in Fig. 2a. From Fig 2a, the pinch temperature is observed to be at 194 K, and the minimum external cold utility ( $CU_{\min}$ ) required by the LAES process is 17 kW. The  $CU_{\min}$  requirement has to be met by external utility below the pinch temperature, which implies that external refrigerants are required. If the utility requirements cannot be met using only ambient air (or water), then we consider the corresponding operating condition to be inadmissible or infeasible (since additional energy has to be spent to generate the external utility). The GCC plot for the same liquid yield, but a different combination of discharge and charging pressures ( $P_C = 100$  bar,  $P_D = 80$  bar) is shown in Fig. 2b. It can be observed from the GCC that the process is self-sufficient and does not require either hot or cold utility. Thus, the process operating condition is feasible. The GCC plot for the process operating at  $P_C = 100$  bar,  $P_D = 60$  bar) is shown in Fig. 2c. It can be observed from Fig. 2c that the process is not pinched and requires a minimum external heating utility ( $HU_{\min}$ ) of 7.5 kW. However, from the GCC, it can also be inferred that the entire  $HU_{\min}$  can be supplied using air available at 300K. Thus, this operating conditions is also called feasible. In order to identify the feasible operating regime, the analysis is performed for different operating conditions. The entire regime of operating conditions is divided into admissible and inadmissible regimes.

The feasible operating regimes identified using pinch analysis for the LAES process described in the preceding section are presented in Fig. 3, for different values of liquid air yield. Corresponding to any given liquid air yield, an operating point in the infeasible region has a GCC similar to the one shown in Fig. 2a, with the pinch temperature being below the ambient temperature. An operating point on the boundary separating the feasible and infeasible regimes has a GCC similar to the one shown in Fig. 2b. Under these operating conditions, the process does not require external utilities and can be

considered to be optimum operating points in terms of utility requirements. An operating point within the feasible region has a GCC similar to Fig. 2c, and requires external HU which can be met by ambient air. For these operating points, excess heat is generated during compression in the charging cycle than what can be stored by the thermal oil and utilized during the discharge cycle, and the excess heat is discharged to the ambient leading to sub-optimal efficiencies.

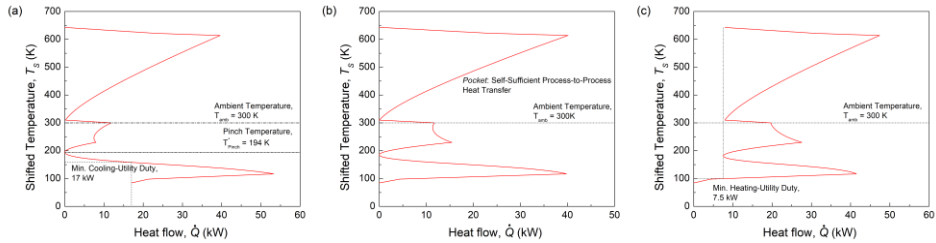


Figure 2: Grand Composite Curve (a) for the data in Table 1 with  $P_C = 100$  bar,  $P_D = 150$  bar and  $Y = 0.7$ , (b)  $P_C = 100$  bar,  $P_D = 80$  bar and  $Y = 0.7$  and (c)  $P_C = 100$  bar,  $P_D = 60$  bar and  $Y = 0.7$

Table 1: Process stream data for operating condition  $P_C = 100$  bar,  $P_D = 150$  bar and  $Y = 0.7$

Stream $i$	$T_i^{Supply}$ (K)	$T_i^{Target}$ (K)	$[\dot{m}C_{P_i}^{Min}, \dot{m}C_{P_i}^{Max}]$ (kW/K)	$\Delta H_i$ (kW)
2 → 3	621.00	302.00	[1.02, 1.05]	329.60
4 → 6	626.00	104.05	[1.07, 2.82]	707.59
7 → 8	79.15	295.00	[0.20, 0.21]	43.71
3C <sub>1</sub> → 4C <sub>1</sub>	225.00	295.00	[1.13, 1.15]	80.01
3C <sub>2</sub> → 4C <sub>2</sub>	93.00	225.00	[1.55, 1.64]	209.25
1H → 4H	300.00	603.72	[1.69, 2.84]	684.56

It can be observed from Fig. 3c that for the desired liquid air yield of 0.7 and  $P_D$  of 120 bar. The minimum required value of charging section operating pressure ( $P_C$ ) is 120 bar to make the process feasible. Similarly, the maximum possible liquid air yield for the process operating with  $P_C = 120$  bar and  $P_D = 60$  bar is 0.8. Any further desired liquid air yield, demands utility that cannot be met with ambient air. The round trip efficiencies corresponding to different operating points on the boundaries separating the feasible and operating regimes for different yields do vary, and can be compared to identify the best operating point that gives the maximum RTE, after which the process design and heat integration for the corresponding operating condition can be carried out.

A study conducted to examine the effect of the minimum temperature approach on the feasible operating region has revealed that a decrease in the MTA leads to an increase in the size of the feasible operating region.

## 4 CONCLUSIONS

A systematic procedure for applying pinch analysis to a LAES process is proposed which considers the impact of temperature and phase changes in process streams induced by pressure change, nonlinear dependence of specific heat on temperature, and multiple

minimum approach temperature specifications. The proposed procedure can be used to identify the feasible design space and other options for improving overall energy efficiency. The proposed procedure can potentially be applied to analyze other processes, such as air separation, natural gas liquefaction, and the use of vapor recompression for energy integration in distillation columns, where process streams undergo significant pressure change.

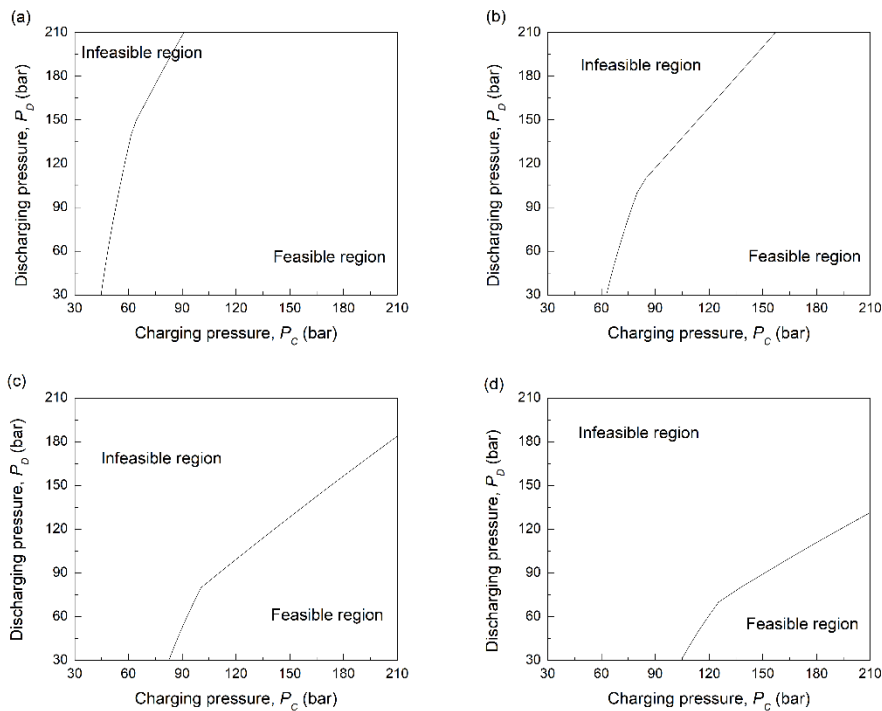


Figure 3: Influence of liquid air yield on the feasible operating region (a)  $Y = 0.5$ , (b)  $Y = 0.6$ , (c)  $Y = 0.7$ , and (d)  $Y = 0.8$ .

## 5 REFERENCES

- [1] Smith R. Chemical Process Design and Integration. vol. 83. 2005. <https://doi.org/10.1205/cherd.br.0509>.
- [2] She X, Zhang T, Cong L, Peng X, Li C, Luo Y, et al. Flexible integration of liquid air energy storage with liquefied natural gas regasification for power generation enhancement. Appl Energy 2019;251:113355. <https://doi.org/10.1016/j.apenergy.2019.113355>.
- [3] Aspelund A, Berstad DO, Gundersen T. An Extended Pinch Analysis and Design procedure utilizing pressure based exergy for subambient cooling. Appl Therm Eng 2007;27:2633–49. <https://doi.org/10.1016/j.applthermaleng.2007.04.017>.

# Climate-zone-based Techno-Economic Assessment of Different Carbon Capture Technologies for Coal-Fired Power Plants

Mohammadamin Zarei,<sup>a</sup> Ha-Jun Yoon,<sup>a</sup> Chul-Jin Lee,<sup>a,b</sup>

*<sup>a</sup>School of Chemical Engineering and Materials Science, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul, Republic of Korea*

*<sup>b</sup>Department of Intelligent Energy and Industry, Chung-Ang University, 84 Heukseokro, Dongjak-gu, Seoul, Republic of Korea*

## Abstract

As fossil fuels are consumed and proven reserves exist, reducing carbon emissions is crucial to ensuring global energy security. It is imperative to reduce CO<sub>2</sub> emissions from coal-fired power plants, which can be done most effectively through capturing and storing CO<sub>2</sub> (CCS). Contrary to other aspects of power generation, CCS technologies have not been investigated in terms of their sensitivity to climate zones. Having that information helps decision-makers minimize carbon emissions by making long-term plans. Therefore, in this study, a techno-economic assessment (TEA) model was developed for coal-fired power plants using post- and oxy-fuel combustion CCS technologies. A comparison of minimum and maximum values reveals that rising altitude, as reflected in changes in ambient pressure, decreases avoided CO<sub>2</sub> costs by 8% for post-combustion technology and 3% for oxy-combustion technology. It also reveals an increase in temperature reduces post-combustion costs by 10.5% and oxy-fuel costs by 5.8%.

**Keywords:** Coal-fired power plant, carbon capture and storage, climate.

## 1. Introduction

The global mean temperature has risen by 1 °C due to anthropogenic greenhouse gas emissions. About 0.3 °C of the global average temperature increase is attributed to CO<sub>2</sub> produced by coal combustion. Globally, coal-fired power plants generate 38% of electricity and 30% of carbon dioxide emissions. It is impossible for conventional fossil fuel-fired power plants to achieve environmental goals even after efficiency improvements and process optimization. World power generation continues to be dominated by fossil fuels despite considerable investment and falling renewable energy prices, particularly in developing countries. Consequently, fossil fuels will remain a key source of energy for the foreseeable future. To use a large amount of energy from fossil fuels, efforts are being made to reduce greenhouse gas emissions (Yadav and Mondal, 2022).

Using carbon capture and storage (CCS) technologies has demonstrated great potential in reducing greenhouse gas emissions while utilizing fossil fuels to generate electricity in response to the Paris Agreement, which aims to peak carbon emissions, neutralize carbon emissions, and achieve net zero emissions. The CCS technology combines a variety of integrated technologies to capture huge amounts of CO<sub>2</sub> from industrial sources such as power plants (Wei, Manovic and Hanak, 2020). The process can be divided into three

stages: capture, transportation, and storage. In the current state of CO<sub>2</sub> capture technology, there are primarily three types of systems, which are post-combustion, pre-combustion, and oxy-fuel combustion systems. The pre-combustion capture process involves the separation of fossil fuel carbon before combustion occurs. When oxygen is used instead of regular air as an oxidant in combustion, it is referred to as oxy-fuel combustion. The term "post-combustion capture" refers to capturing and storing CO<sub>2</sub> as a result of a conventional power station's exhaust gas (Ye *et al.*, 2019).

Developing a comprehensive understanding of how climate zones affect CCS efficiency is one of the challenges in technical evaluation of these technologies. Consequently, parametric studies in CCS should be performed by changing operating conditions such as temperature, pressure, and relative humidity. There are several studies on the components of power plants. Bataineh and Khaleel investigated the effect of ambient conditions on gas turbines in combined cycle power plants (Bataineh and Khaleel, 2020). Similarly, Zhai and Rubin examined local ambient conditions as a key factor influencing cooling system performance, cost, and water use (Zhai and Rubin, 2010). The influences of the meteorological data, including ambient temperature and DNI on solar power plants, are examined by Wang *et al.* (Wang *et al.*, 2019). However, to the best of our knowledge, there are limited studies that have focused on the effect of ambient parameters on CCS. Rim *et al.* worked on changes in temperature, humidity, and atmospheric pressure to set optimal process conditions for direct air capture of CO<sub>2</sub> (Rim *et al.*, 2022). A study on changes in CCS technologies in different climate zones has not been conducted, even though other aspects of power generation have investigated their sensitivity to climate zones. Therefore, this study aims to develop a techno-economic assessment (TEA) model of a coal-fired power plant using two methods of post- and oxy-fuel combustion CCS technology to measure the effect of climate parameters including ambient air pressure, relative humidity, and temperature on CCS technologies.

## 2. Methodology

### 2.1. Description of technical methods

As a result of the above mentioned research gap in CCS-related studies, this work focuses on two main types of CCS technology in coal-fired power plants as Figure 1 shows: 1) post-combustion which is the most cost-effective technology option (Ye *et al.*, 2019) and 2) oxy-fuel combustion: which has the high capacity to absorb CO<sub>2</sub> (Alalwan and Alminshid, 2021). The Integrated Environmental Control Model (IECM), developed by Carnegie Mellon University, is used to assess the cost of carbon capture in power plants (CMU, 2021). Through IECM, fossil fuel-fired power systems with CCS can be systematically analyzed in terms of their performance, emissions, and costs. To determine the range of climate-related factors, Köppen climate classification is used to analyze the various climate zones (Cui, Liang and Wang, 2021).

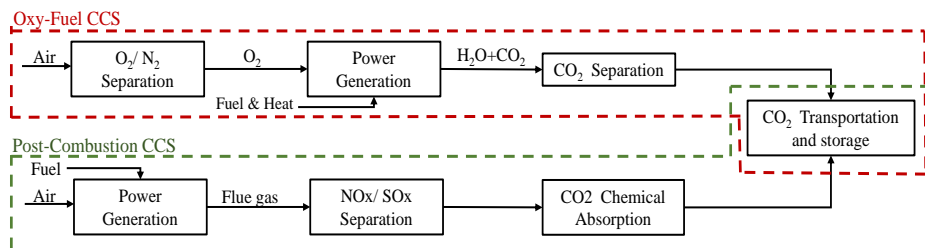


Figure 1. Integrated flow diagram of CCS technologies

2.2. The cost analysis framework

650 MW gross power coal-fired subcritical power plant capable of carbon capture is considered for the purpose of comparing costs between CCS options in different climate zones. A schematic of the cost-estimating methodology is presented in Figure 2. The methodology begins with defining the cases. This step determines ambient conditions, combustion process, CCS process, system capacity, and fuel type. In the second part of this analysis, unit prices such as fuel and electricity prices, financial parameters such as interest rates, taxes, and operating schedules such as load factor are used to calculate performance. Finally, cost measures such as Levelised cost of electricity (LCOE), cost of CO<sub>2</sub> captured, and cost of CO<sub>2</sub> avoided are defined to make the cost evaluation as follows:

$$LCOE = \frac{TCR \cdot FCF + FOM}{MW \cdot CF \cdot 8760} + VOM + HR \cdot FC, \tag{1}$$

$$C_{avoided} = \frac{C_{ccs} - C_{ref}}{E_{ref} - E_{ccs}}, \tag{2}$$

$$C_{captured} = \frac{C_{ccs} - C_{ref}}{E_{captured}}, \tag{3}$$

where TCR is total capital requirement, FCF is fixed charge factor, FOM is fixed operating and maintenance cost(O&M), VOM is variable O&M excluding fuel, HR refers to the heat rate of the powerplant, the cost of fuel per unit of energy is represented by FC, CF refers to capacity factor, and the plant's net power output is represented by MW. Using equations (2) and (3), a comparison is made between the plant using CCS and the reference plant without CCS(Keivani and Gungor, 2022).  $C_{ccs}$  and  $C_{ref}$  are power generation cost of CCS powerplant and reference powerplant, respectively.  $E_{ref}$  and  $E_{ccs}$  are carbon emission intensity of the reference and CCS powerplant. Carbon emission intensity of the CCS power plant before the capturing process is represented by  $E_{captured}$ .

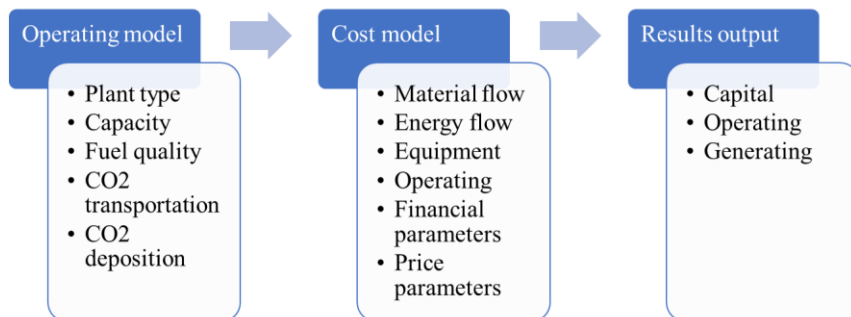


Figure 2. System model. structure of the CCS cost module.

The main parameters of the model include the following four modules: operating parameters of the CCS and reference power plant, parameters of fuel quality, sulfur dioxide treatment system parameters, and CO<sub>2</sub> capture system parameters. This study selects a commercial subcritical power plant with an output power of nearly 650 MW. In the case of the absorption-based CO<sub>2</sub> capture process, the processes that use MEA for the post-combustion unit (Rubin, Rao and Berkenpas, 2007) and High-sulfur configuration for the oxy-fuel combustion (IECM, 2021). The main design assumptions including main parameters and relevant technical and economic considerations used for design cases are reported in Tables 1 and 2. In this work, reference powerplant is located in Denver.

Table 1. Basic assumption for CCS technologies

	Post-combustion control	Oxy-fuel combustion
<b>NO<sub>x</sub> removal</b>	Hot-Side SCR	
<b>Particulate Removal</b>	Cold-Side ESP	Fabric filter
<b>SO<sub>x</sub> removal</b>	FGD	FGD
<b>CO<sub>2</sub> capture</b>	Amine system(MEA)	High-sulfur (>1.5%)
<b>Storage</b>	Geological Reservoir	
<b>Transportation</b>	Pipeline	

Table 2. Key assumptions for the economic evaluation

Parameter	Value	Unit
<b>Plant capacity</b>	650	MW
<b>Power plant type</b>	Subcritical	-
<b>Construction period</b>	3	years
<b>Capacity factor</b>	75	%
<b>Economic lifetime</b>	30	years
<b>Method</b>	Constant money value	
<b>Discount rate</b>	7	%
<b>Inflation</b>	Constant money value	
<b>Carbon tax</b>	0	\$
<b>Fuel type</b>	Bituminous	-
<b>Fuel cost</b>	49.87	\$/tonne
<b>Electricity cost</b>	43.84	\$/MWh
<b>Operating labor rate</b>	34.65	\$/hour
<b>Number of Operating Shifts</b>	4.75	shift/day

### 3. Results and Conclusion

A simulation showing the operating results of a newly constructed power plant that includes CCS and its reference power plant can be seen in figures 3 and 4. Figure 3 presents the effect of three main climate indicators on post-combustion CCS in coal-fired

power plant. It shows performance characteristics that distinctly depend on relative humidity. However, as temperature increases, cost of CO<sub>2</sub> avoided and added cost of CCS increase by 10.44% and 10.45%, respectively. Likewise, changing ambient pressure by 0.083 bar to 1.01 bar which represents changing 1700 meters in altitude decreases cost of CO<sub>2</sub> avoided and added cost of CCS by 8.39% and 8.15%, respectively. Figure 4 represents the effect of same climate indicators on oxy-fuel combustion. Same as Figure 3, it can be seen relative humidity does not affect CCS performances. However, ambient temperature and pressure increase cost by almost 5.8% and 3.5%. In the post-combustion process CO<sub>2</sub> is diluted in nitrogen that is existed in air and changing temperature highly effected separation costs. To avoid dilution, in oxy-fuel combustion process pure oxygen is used to yield more concentrated CO<sub>2</sub>. Therefore, it can be seen these indicators have a greater impact on post-combustion than oxy-fuel combustion. Consequently, it makes changes in LCOE. Increasing temperature and pressure in post-combustion plants increase LCOE by 5% and 3.6%. Moreover, increasing temperature and pressure in oxy-fuel combustion plants increase LCOE by less than 3%. It can be concluded humidity does not affect as much as temperature and pressure. It can be concluded that ambient temperature and pressure have an effect on CCS cost and on the other hand, it is independent of ambient relative humidity.

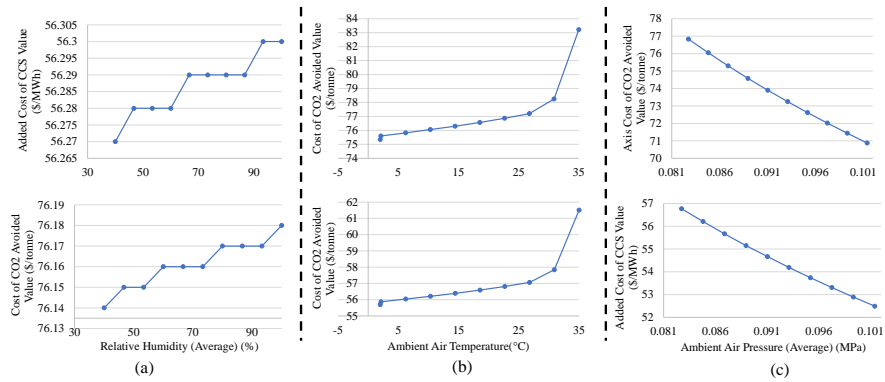


Figure 3 Impact of sensitivity factors on post-combustion performances: (a) relative humidity (b); ambient air; temperature.

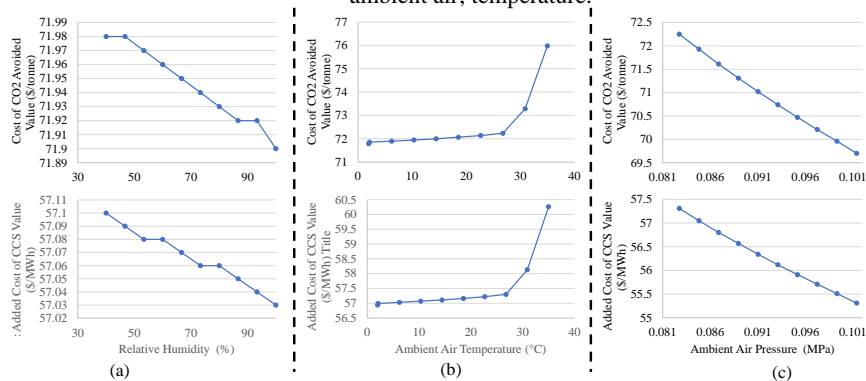


Figure 4 Impact of sensitivity factors on oxy-combustion performances: (a) relative humidity (b); ambient air; temperature.



## References

- Alalwan, H.A. and Alminshid, A.H. (2021) 'CO<sub>2</sub> capturing methods: Chemical looping combustion (CLC) as a promising technique', *Science of The Total Environment*, 788, p. 147850. Available at: <https://doi.org/https://doi.org/10.1016/j.scitotenv.2021.147850>.
- Bataineh, K. and Khaleel, B.A. (2020) 'Thermodynamic analysis of a combined cycle power plant located in Jordan: A case study', *Archives of Thermodynamics*, 41(1), pp. 95–123.
- IECM. (2021). Integrated Environmental Control Model (IECM) Version 11.5 (Carnegie Mellon University). <https://www.cmu.edu/epp/iecm/index.html>.
- Cui, D., Liang, S. and Wang, D. (2021) 'Observed and projected changes in global climate zones based on Köppen climate classification', *WIREs Climate Change*, 12(3), p. e701. Available at: <https://doi.org/https://doi.org/10.1002/wcc.701>.
- IECM (2021) *Amine-based Post-Combustion CO<sub>2</sub> Capture*. Pittsburgh. Available at: [https://www.cmu.edu/epp/iecm/documentation/2019Jan\\_IECM\\_Amine-based\\_CO2\\_Capture.pdf](https://www.cmu.edu/epp/iecm/documentation/2019Jan_IECM_Amine-based_CO2_Capture.pdf).
- Keivani, B. and Gungor, A. (2022) 'Techno-economic assessment of coal and torrefied biomass co-combustion: A case study of oxy-combustion carbon capture power plants in Turkey', *Journal of CO<sub>2</sub> Utilization*, 62, p. 102103. Available at: <https://doi.org/https://doi.org/10.1016/j.jcou.2022.102103>.
- Rim, G. et al. (2022) 'Sub-Ambient Temperature Direct Air Capture of CO<sub>2</sub> using Amine-Impregnated MIL-101(Cr) Enables Ambient Temperature CO<sub>2</sub> Recovery', *JACS Au*, 2(2), pp. 380–393. Available at: <https://doi.org/10.1021/jacsau.1c00414>.
- Rubin, E.S., Rao, A.B. and Berkenpas, M.B. (2007) *Development and application of optimal design capability for coal gasification systems*. Carnegie-Mellon University.
- Wang, A. et al. (2019) 'Thermodynamic and economic analyses of a parabolic trough concentrating solar power plant under off-design conditions', *Applied Thermal Engineering*, 156, pp. 340–350. Available at: <https://doi.org/https://doi.org/10.1016/j.applthermaleng.2019.04.062>.
- Wei, X., Manovic, V. and Hanak, D.P. (2020) 'Techno-economic assessment of coal- or biomass-fired oxy-combustion power plants with supercritical carbon dioxide cycle', *Energy Conversion and Management*, 221, p. 113143. Available at: <https://doi.org/https://doi.org/10.1016/j.enconman.2020.113143>.
- Yadav, S. and Mondal, S.S. (2022) 'A review on the progress and prospects of oxy-fuel carbon capture and sequestration (CCS) technology', *Fuel*, 308, p. 122057. Available at: <https://doi.org/https://doi.org/10.1016/j.fuel.2021.122057>.
- Ye, B. et al. (2019) 'Technical and economic analysis of amine-based carbon capture and sequestration at coal-fired power plants', *Journal of Cleaner Production*, 222, pp. 476–487. Available at: <https://doi.org/https://doi.org/10.1016/j.jclepro.2019.03.050>.
- Zhai, H. and Rubin, E.S. (2010) 'Performance and cost of wet and dry cooling systems for pulverized coal power plants with and without carbon capture and storage', *Energy Policy*, 38(10), pp. 5653–5660. Available at: <https://doi.org/https://doi.org/10.1016/j.enpol.2010.05.013>.

# Numerical tool for dynamic simulation of anaerobic digesters including an air-inflated double membrane gasholder

Giberto M. Yuki Junior,<sup>a, b</sup> Sabine Sochard,<sup>a</sup> Elio Dinuccio,<sup>b</sup> Frédéric Marias<sup>a</sup>

<sup>a</sup>*Universite de Pau et des Pays de l'Adour, E2S UPPA, LaTEP, Rue Jules Ferry, Pau 64000, France*

<sup>b</sup>*Department of Agriculture, Forestry and Food Science, University of Turin, Largo Paolo Braccini 2, Grugliasco 10095, Italy*

## Abstract

Air-inflated double membrane gasholders are often used in anaerobic digestion plants to store the biogas produced. In this work, a dynamic model based on conservation equations and the ideal gas law for such type of gas storage was proposed. Simulations during charging and discharging steps of the gas storage led to a stable pressure profile in both cases, but with a higher value through the first period. The model was also used to compare the operation at a full load with one at a medium load, which demonstrates that reducing the gas storage level can avoid overpressure events caused by temperature variations.

**Keywords:** biogas, double membrane, gasholder, model.

## 1. Introduction

In biogas plants, the gas produced is often stored inside the digester's headspace. The air-inflated double membrane gasholder is one of the most used structures to store biogas (Stur et al., 2022). It consists of two polymeric membranes delimiting two compartments over the digester: an outer one filled with air and an inner one filled with biogas. The external membrane is kept inflated by a continuous injection of air using a blower. One benefit of such structure is that the pressure can be controlled regardless of the storage level (Kube, 2018). Moreover, several digesters and post-digesters can have their headspaces interconnected (Kube, 2018). Hence, the flowrate between each unit can be manipulated by controlling the pressure in each reactor (Kube, 2018).

The double membrane is also equipped with a pressure relief valve, which is a safety device to avoid over and underpressure episodes. This equipment works by releasing biogas into the atmosphere in case of overpressure or by injecting air directly into the biogas storage compartment in case of underpressure (Stur et al., 2022). Although such valve is a safety equipment, a monitoring in a biogas plant in Germany showed that biogas release events were common, with CH<sub>4</sub> losses reaching up to 1.8% of annual production (Reinelt and Liebetrau, 2020). These events were linked to biogas temperature variations and lack of staff training to operate this kind of gasholder (Reinelt and Liebetrau, 2020). Such losses, however, could have been reduced by automatic gas level control (Reinelt and Liebetrau, 2020) or by operating at a lower storage level (Reinelt and Liebetrau, 2020; Stur et al., 2022).

To the best of authors' knowledge, the double membrane gasholder transient behavior has not yet been modelled despite of its implications on biogas plant operation. Indeed, dynamic modelling approaches in the biogas field have been concentrated on the

bioreactions (Emebu et al., 2022) or the digester's thermal behavior (Hreiz et al., 2017; Vilms Pedersen et al., 2020). Thus, in this work an approach to describe the transient operation of a double membrane gasholder is presented. Therefore, the model developed will be useful to improve the simulation of biogas plants and it will be helpful in evaluating control strategies for the gas storage management and to train staff operating such industrial plants. In future work, the digester and the double membrane models will be included in a global model for a multi-energy platform containing methanation, water electrolysis and biogas upgrading units.

## 2. Model description

### 2.1. General digester model

The digester model was developed by integrating different approaches to describe the bioreactions, the reactor thermal behavior and the mass transfer to the gas phase. The digestate phase was supposed well mixed and, thus, a 0-D approach was used. A mass based version of ADM1 was employed to predict the biogas production (Weinrich and Nelles, 2021). However, some small changes were implemented into the biokinetic model. For instance, the H<sub>2</sub>S production by sulfate-reducing bacteria were considered (Pokorna-Krayzelova et al., 2017) and the stoichiometry of the cell death reaction was rebalanced so that the atomic balances were fulfilled.

On the other hand, the thermal behavior was estimated by considering the heat losses to the environment through the digester boundaries using empirical correlations for the heat transfer resistances (Hreiz et al., 2017; Incropera et al., 2007; Vilms Pedersen et al., 2020). Finally, the mass transfer between the liquid and the gas phase was evaluated by estimating the mass transfer coefficient from the digestate properties and agitation mobile configuration (Garcia-Ochoa and Gomez, 2004) or using the mass and heat transfer analogy (Incropera et al., 2007).

### 2.2. Double membrane model

The behavior of the air inflated gasholder was described by applying mass and energy balances on the air (A) and the biogas (B) layers as shown in Figure 1 and Eqs. (1)-(4). In Eq. (1), the air molar flowrate through the blower,  $F_A^{in}$ , was calculated as a function of the relative air pressure,  $P_A^g$ , using a polynomial function based on the curves used in Kube's work (Kube, 2018), while the outlet flow was estimated using a valve equation, as shown in Eq. (5) and Figure 1 (a). Eq. (2) describes the mass balance on the biogas side, where  $F_B^{GL}$  represents the biogas production in the digester and  $F_B^{out}$  the current molar flowrate through the valve. Finally, the temperatures of each gas phase were estimated using Eqs. (3) and (4), which were obtained from the energy balances, assuming ideal gas behavior and constant values for the heat capacities. Moreover, the thermal inertia of the walls and the cover in contact with each phase were included in the energy balance in the form of  $m_{wall}Cp_{wall}$  in both energy balances. This approach was used to avoid overestimating the internal temperature variations with the external conditions. The heat transfer terms,  $Q$ , were calculated considering the convection inside and outside the digester, and the radiation exchange with the external environment. The state variables in each gas compartment are summarized in Figure 1 (b).

$$\frac{dn_A}{dt} = F_A^{in}(P_A^g) - F_A^{out}(P_A^g) \quad (1)$$

$$\frac{dn_B}{dt} = F_B^{GL} - F_B^{out}(P_B^g) \quad (2)$$

$$\frac{dT_A}{dt} = \frac{F_A^{in} C p_A (T_A^{in} - T_A) + RT_A (F_A^{in} - F_A^{out}) + Q_A}{n_A C v_A + m_{wall,A} C p_{wall,A}} \quad (3)$$

$$\frac{dT_B}{dt} = \frac{F_B^{GL} C p_B (T_B^L - T_B) + RT_B (F_B^{GL} - F_B^{out}) + Q_B}{n_B C v_B + m_{wall,B} C p_{wall,B}} \quad (4)$$

$$\dot{V} = CV\sqrt{P^g} \quad (5)$$

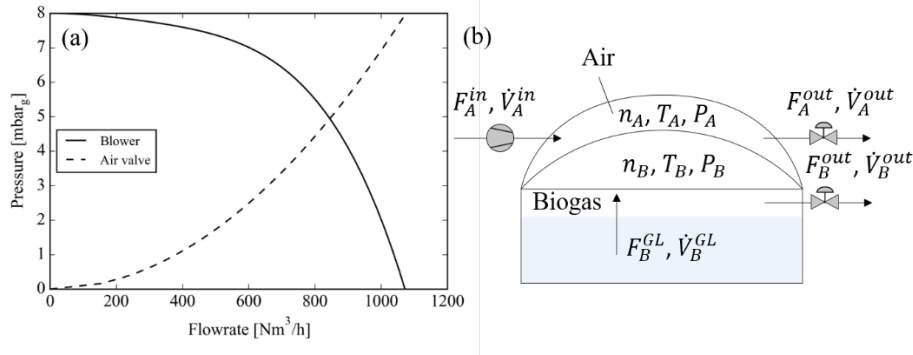


Figure 1 (a) Air blower and valve curves (b) Digester scheme showing the gas flows and state variables in the gas phases

### 2.2.1. Mobile regime

When the air volume was not at its threshold values, it was assumed that the inner membrane could freely move. All forces related to the membrane weight or its displacement were neglected. Thus, the pressures in both sides were supposed to be equal. The outer membrane, nevertheless, was considered as a rigid wall. Therefore, the total volume, delimited by this outer cover and the digester walls, could be considered constant as shown in Eq. (6), where  $V_L$  refers to the digestate volume.

$$V_{total} = V_A + V_B + V_L \quad (6)$$

Using Eq. (6), the equality of pressures and the ideal gas law, Eq. (7) could be obtained. Therefore,  $V_A$  can be directly evaluated from the molar quantities and temperatures.  $V_B$ ,  $P_A$  and  $P_B$  were then calculated using Eq. (6) and the ideal gas law.

$$V_A = (V_{total} - V_L) \frac{n_A T_A}{n_A T_A + n_B T_B} \quad (7)$$

### 2.2.2. Isochoric regime

When the biogas compartment is at its maximum level, any further increase in gas stored would lead to an increase in pressure, which would be sustained by the inner membrane and the attachment structure on digester's walls. On the opposite, at minimum air volume, a decrease in pressure would create an overpressure on the air side. In this situation, the pressure difference would be sustained by a central support structure. In both cases, it was assumed that the air compartment becomes an isochoric one. Therefore,  $P_A$  was calculated by setting  $V_A$  to its minimum or maximum value and employing the ideal gas law.

### 2.3. Biogas valve control

The biogas outlet stream was controlled through the valve constant,  $CV_B$ , which was evaluated as a function of  $\Delta\dot{V}_B$ , which represents a material accumulation rate in the biogas storage, as shown in Eq. (8), where the flowrates are in  $\text{Nm}^3\text{h}^{-1}$  and the relative pressure in mbar.

$$\dot{V}_B^{out} = \dot{V}_B^{GL} - \Delta\dot{V}_B = CV_B \sqrt{P_B^g} \quad (8)$$

Although a pressure relief valve has not been considered in the digester model, a maximum pressure event was dealt by manipulating  $\Delta\dot{V}_B$ . For instance, if  $P_B$  reaches its maximum allowed value,  $P_{B,max}$ ,  $\Delta\dot{V}_B$  can be set to a negative value until  $P_B$  decreases to the same level as  $P_A$ . Such measure forces an additional flowrate through the biogas valve, which represents the biogas that would be lost through the pressure relief valve.

## 3. Case studies

The simulations were carried out considering a digester with 24 m of diameter. The heights of the cylindrical part and the outer membrane were both set at 8 m. This yielded a total volume,  $V_{total}$ , of  $5697 \text{ m}^3$ . The thresholds limits applied to the air compartment was of  $375 \text{ m}^3$  when at full biogas load and of  $1965 \text{ m}^3$  at minimum load. The biogas volume limits depend also on the liquid volume. During the simulations,  $V_L$  was kept approximately at  $3400 \text{ m}^3$ . Therefore,  $V_{B,min}$  and  $V_{B,max}$  were around  $332$  and  $1922 \text{ m}^3$  respectively. The biogas production was computed considering a substrate feeding done once a day and an organic loading rate (OLR) at  $2.8 \text{ kg VS/m}^3\text{d}$ .

### 3.1. Double membrane characteristic curve

Stur et al. (2022) proposed a protocol to determine the maximum and minimum capacities of a double membrane gasholder. It consists in filling the gasholder from a technical empty condition until full load and then discharging it until the empty point. The full point is characterized by a biogas pressure surge and the empty one by a sudden pressure reduction (Stur et al., 2022).

To simulate this procedure,  $P_{B,max}$  was set at 7 mbar and the filling phase was defined by  $\Delta\dot{V}_B$  at  $180 \text{ Nm}^3\text{h}^{-1}$  and at  $-180 \text{ Nm}^3\text{h}^{-1}$  during the discharging period. The full load was defined as when  $P_B$  reached  $P_{B,max}$ , from that point forward  $\Delta\dot{V}_B$  was set to 0 for a one-hour period before starting the discharging phase. Figure 2 exhibits the simulation results for biogas pressure and volume. Even though a quantitative comparison is not possible, these results follow the same trend as the experimental values obtained by Stur et al. (2022). Indeed, the charging and discharging phases are characterized by two pressure plateaus, with the first one at 6.1 mbar and the latter at 3.7 mbar. Such behavior was previously explained by Kube (2018) and depends on the blower and valve curves. For instance, the pressure is higher when the storage is being filled because, at the air side, the air inlet flow is lower than the outlet when the pressure is above the operating point (Figure 1 (a)).

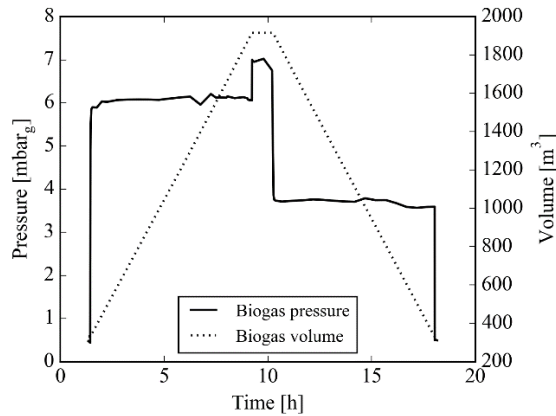


Figure 2 Double membrane's characteristic curve as computed by the model

### 3.2. High and medium load operations

Biogas temperature variations have been linked to overpressure events in double membrane gasholders (Reinelt and Liebetrau, 2020; Stur et al., 2022). Reducing the operating gas storage level has been pointed out as a strategy to mitigate biogas losses through the pressure relief valve (Stur et al., 2022). To test such situation, two simulations were carried out: one at high storage load, with biogas stored volume close to its maximum value, and another one at medium load. In both cases,  $\Delta\dot{V}_B$  was set at 0, except in case of overpressure, when  $\Delta\dot{V}_B$  was set at  $-100 \text{ Nm}^3\text{h}^{-1}$ . The simulations were carried out for three days using the weather data (wind speed, solar irradiance and external temperature) from 16<sup>th</sup> to 18<sup>th</sup> august 2009 in Pau, France ( $43^\circ 17' \text{N}$ ,  $24' \text{W}$ ) (Huld et al., 2012). The results of such simulations are shown in Figure 3. During the high load operation, pressure peaks were detected during the three days. However,  $P_{B,max}$  was reached only on the last two days, when  $T_B$  had its highest intraday increase. This also led to additional biogas outlet flow as shown in Figure 3 (b), which represents the losses that would have occurred through a pressure relief valve. On the other hand, operating with a lower storage level completely avoided any pressure surge, as the biogas volume fluctuations were within the threshold limits.

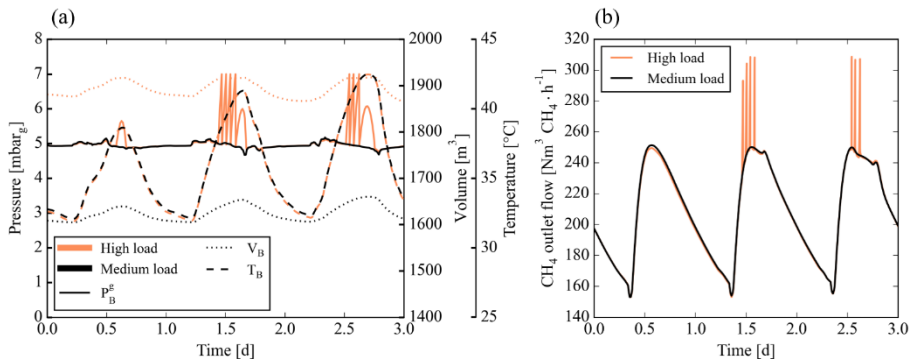


Figure 3 (a) Biogas volume, temperature and relative pressure profiles for the high and the medium load tests (b) Methane flowrate through the outlet valve during high and medium load operations

#### 4. Conclusion

A dynamic model for an air inflated double membrane gasholder was developed and integrated to a digester model. The gas storage model was able to simulate the characteristic curve of such kind of gasholder, with results qualitatively comparable to experimental values in the literature. Besides that, a strategy to reduce overpressure events could be tested. The next steps for this work are to validate the full digester model with measures from industrial biogas plants and to integrate it with other units, such as methanation and electrolysis reactors and biogas upgrading units.

#### References

- S. Emebu, J. Pecha, D. Janáčová, 2022, Review on anaerobic digestion models: Model classification & elaboration of process phenomena, *Renewable and Sustainable Energy Reviews* 160, 112288.
- F. Garcia-Ochoa, E. Gomez, 2004, Theoretical prediction of gas–liquid mass transfer coefficient, specific area and hold-up in sparged stirred tanks, *Chemical Engineering Science* 59, 2489–2501.
- R. Hreiz, N. Adouani, Y. Jannot, M.-N. Pons, 2017, Modeling and simulation of heat transfer phenomena in a semi-buried anaerobic digester, *Chemical Engineering Research and Design* 119, 101–116.
- T. Huld, R. Müller, A. Gambardella, 2012, A new solar radiation database for estimating PV performance in Europe and Africa, *Solar Energy* 86, 1803–1815.
- F.P. Incropera, D.P. DeWitt, T.L. Bergman, A.S. Lavine, 2007, *Fundamentals of Heat and Mass Transfer*, 6th edition. ed, John Wiley & Sons Ltd, New York.
- J. Kube, 2018, Management of Gas Storages in Biogas Plants, *Chem. Eng. Technol.* 41, 702–710.
- L. Pokorna-Krayzelova, K.E. Mampaey, T.P.W. Vannecke, J. Bartacek, P. Jenicek, E.I.P. Volcke, 2017, Model-based optimization of microaeration for biogas desulfurization in UASB reactors, *Biochemical Engineering Journal* 125, 171–179.
- T. Reinelt, J. Liebetrau, 2020, Monitoring and Mitigation of Methane Emissions from Pressure Relief Valves of a Biogas Plant, *Chem. Eng. Technol.* 43, 7–18.
- M. Stur, M. Pohl, C. Krebs, E. Mauky, 2022, Characterisation of biogas storages: influences and comparison of methods, *Agricultural Engineering Bd.* 77 Nr. 1 (2022).
- S. Vilms Pedersen, J. Martí-Herrero, A.K. Singh, S.G. Sommer, S.D. Hafner, 2020, Management and design of biogas digesters: A non-calibrated heat transfer model, *Bioresource Technology* 296, 122264.
- S. Weinrich, M. Nelles, 2021, Systematic simplification of the Anaerobic Digestion Model No. 1 (ADM1) – Model development and stoichiometric analysis, *Bioresource Technology* 333, 125124.

# An iterated double auction model for peer-to-peer electricity trading

Timothy Huty,<sup>a</sup> Prof Solomon Brown<sup>a</sup>

<sup>a</sup> *Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom, s.f.brown@sheffield.ac.uk*

## Abstract

Peer-to-peer (P2P) energy trading, whereby customers can trade with one another rather than the energy supplier only, has the potential to save money for consumers whilst also incentivising more efficient and environmentally beneficial behaviour. In this work, an agent-based simulation model for a local P2P market is developed. The market is implemented as an iterative double auction, which resembles wholesale electricity trading and gives participants a high degree of freedom to trade using arbitrary strategies. Strategies are developed to enable flexible devices (EV chargers) to participate profitably in the market. A case study is presented based on future scenarios for a primary substation in the north of England. It is found that the P2P market can achieve significant consumer savings and is also effective in ameliorating the reverse flow of solar power through the substation.

**Keywords:** peer-to-peer, energy, continuous double auction, agent-based, decarbonisation

## 1. Introduction

Peer-to-peer (P2P) energy trading enables energy customers to trade with one another, rather than with the energy supplying company only [1]. A key benefit of P2P is that it provides the incentive for participants to smartly coordinate their devices; for instance, this could involve the timing of electric vehicle (EV) charging to coincide with the solar generation surplus of a peer. This can bring financial savings for energy consumers and generators, greater energy independence for localities, and a reduction in greenhouse gas emissions [2]. It is also possible for P2P trading to relieve the burden on the electricity distribution grid, and here we consider the ability to avoid large reverse flows of energy at substations.

In this work we present a model for P2P energy trading via an iterated double auction. Double auction is chosen as it represents a very simple, generic and adaptable mode of trading [3]. The model is developed in the multi-paradigm simulation software AnyLogic [4], employing the agent-based modelling philosophy; agents participating in the market purely pursue their own self-interest. The model includes the three key demands of heat, power and transportation, which is key to real-world applicability. Strategies are developed to enable flexible devices including EV chargers and heat pumps to participate profitably in the market.



## 2. Method

### 2.1 Overview

In this work a local P2P electricity market is modelled, using AnyLogic simulation software in tandem with MILP optimisation. An iterated double auction is adopted as the market mechanism for P2P electricity trading - this system may be considered comparable to the continuous trading that takes place in wholesale electricity markets. In this work the trading is not envisioned to be truly continuous, but to take place in discrete rounds; this ensures that the speed of information propagation does not unfairly advantage any particular participant [5]. A schematic overview of the market construction and household bidding strategies is given in Figure 1.

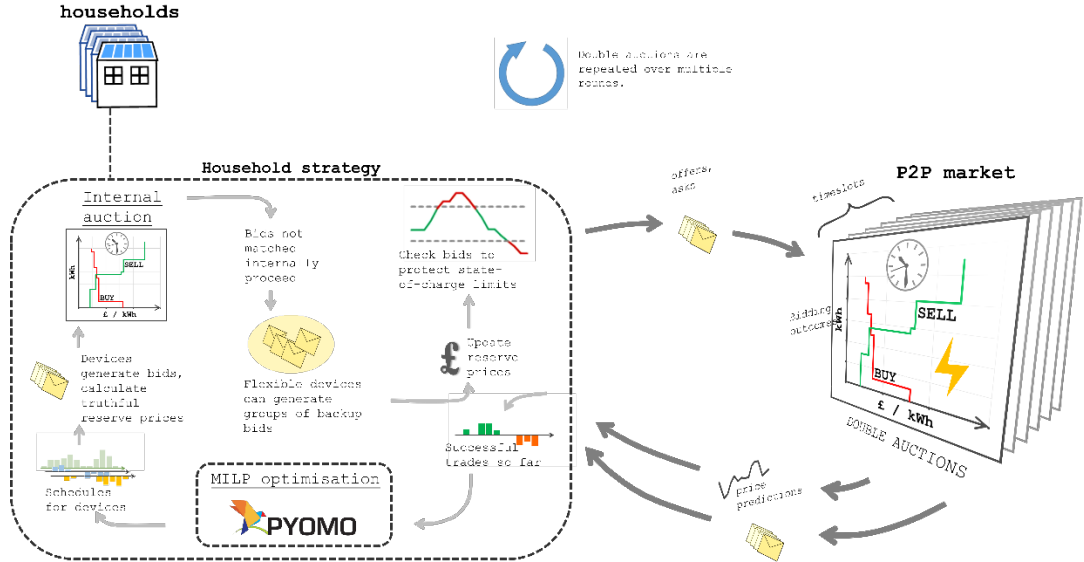
In each round of the double auction, bids to buy and sell electricity are collected from auction participants, a clearing price is determined and trades are executed. Each timeslot of the day ahead has its own auction; timeslots do not clear their auctions simultaneously in each round, but in sequence; this enables auction participants to adjust bids in other timeslots as each timeslot is cleared – an ability which is important for flexible devices. Participants gradually acquire their desired energy trades as the rounds proceed. It is assumed that participants do not sell / buy back energy that they have bought / sold in a previous round.

### 2.2 Formation of bidding strategy

The initial price prediction at the start of trading is equal to the mid-market rate midway between retail price  $p_{grid\_retail}$  and feed-in price  $p_{grid\_FI}$ . Subsequent price predictions at each timeslot are the weighted mean of the two most recent clearing prices. Full strategy formation by auction participants occurs at random every few rounds, with MILP optimisation of device schedules carried out in response to latest price predictions. Energy that has already been bought or sold is assigned to devices via an internal double auction at each timeslot, with bids not matched by the internal auction proceeding to the external P2P auction. Bids to (dis)charge the EV battery are first checked to ensure they cannot lead to infringements of state-of-charge limits; any bids that have to be trimmed are restored in future rounds if possible.

### 2.3 Truthful limit prices and submitted reserve prices

Bidders need to assign a value to the energy they are seeking to trade, i.e. limit prices – see Table 1. Generally, these are related to the grid retail price  $p_{grid\_retail}$  and feed-in tariff  $p_{grid\_FI}$ . Bids to charge the EV for essential travel are separated from those for arbitrage. Where an EV battery is used as energy storage, limit prices to charge (discharge) the battery may depend on the prices of any already secured trades to discharge (charge). Note the definitions of variables  $c_{V2X}$  (degradation cost per kWh discharged),  $\eta_{st}$  (DC round-trip efficiency) and  $\eta_{inv}$  (inverter efficiency). Heat pumps use a version of ‘bid-as-predicted’, never bidding above the predicted price  $\tilde{p}_{buy}$  plus a small increment  $\delta$ .



**Figure 1.** Schematic overview of the P2P double auction market for electricity sharing, showing the main steps in strategy formation.

**Table 1.** Truthful reserve prices (i.e. limit prices) assumed for different applications.

	Category	Truthful reserve price (£/kWh)
<b>Offers to buy power</b>	INFLEXIBLE_LOAD, EV_ESSENTIAL	$p_{grid\_retail}$
	ASHP	$\min(\tilde{p}_{buy} + \delta, p_{grid\_retail})$
	EV_ARBITRAGE (EV charging for later discharge to the house or to the grid; or for storage into the next day)	<p>For an amount corresponding to the EV_V2G bids that have been matched (internally or externally) at an average value of <math>p_{V2X}</math>:</p> $(p_{V2X} - c_{V2X}) \cdot \eta_{inv}^2 \cdot \eta_{st}$ <p>For a further amount not exceeding 10% of battery capacity in each auction round, where <math>\tilde{p}_{V2X}</math> is the <i>predicted</i> average value of corresponding EV_V2X.</p> $(\tilde{p}_{V2X} - c_{V2X}) \cdot \eta_{inv}^2 \cdot \eta_{st}$
<b>Asks to sell power</b>	PV_EXPORT	$p_{grid\_FI}$
	EV_V2G (Sale of EV power to the grid)	<p>For an amount corresponding to the EV_ARBITRAGE bids that have been matched (internally or externally) at an average value of <math>p_{ARB}</math>:</p> $\frac{p_{ARB}}{\eta_{inv}^2 \cdot \eta_{st}} + c_{V2X}$ <p>For a further amount not exceeding 10% of battery capacity in each auction round, where <math>\tilde{p}_{ARB}</math> is the average predicted price of the EV_ARBITRAGE bids not yet matched.</p> $\frac{\tilde{p}_{ARB}}{\eta_{inv}^2 \cdot \eta_{st}} + c_{V2X}$

Some auction participants submit their truthful limit prices  $p_{tr}$  with their bids, as per Table 1. This is termed an ‘aggressive’ strategy, since it maximises the chance of making a trade. Other participants are ‘zero-intelligence’ (Z.I.) bidders. Z.I. bidders submit a reserve price uniformly distributed between their truthful reserve price and an upper or lower bound price, as shown in Equation 1 (for buyers) and Equation 2 (for sellers); this can enable additional value to be extracted, at the cost of decreased probability of trading.

$$p_{res} \sim U(p_{grid\_FI}, p_{tr}) \quad \text{Eqn. 1}$$

$$p_{res} \sim U(p_{tr}, p_{grid\_retail}) \quad \text{Eqn. 2}$$

#### 2.4 Case study

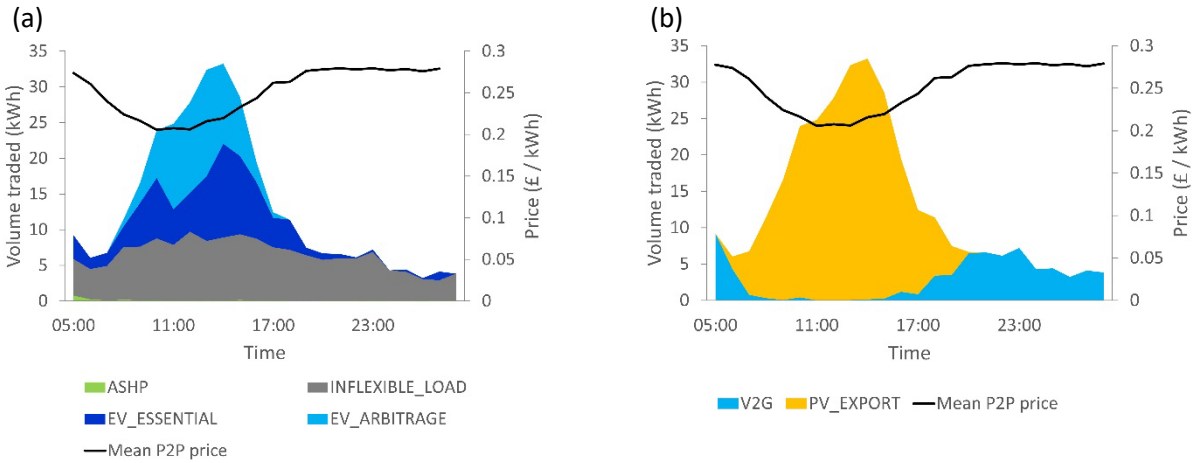
For the basis of a case study, we consider a primary substation located in Doncaster, UK at 53.5392°N 1.0443°W. Modelling by Northern Powergrid and Element Energy [6], [7] predicts that this substation will need to support 59 heat pumps, 98 EVs, and 160 kW<sub>p</sub> of embedded PV generation per 100 customers by 2050. Avoidance of excessive reverse flow of PV generation is our focus here.

Solar generation is modelled as reported in [8] using climate data from UKECN [9]. EV usage is modelled as in [10] where this model is based on data from the UK National Travel Survey [11]. Inflexible load profiles are taken from [12]. Heat demand is based on building archetypes from the CREST demand model [13]. 50 houses are modelled, since this is generally enough to capture diversity of demand; results can then be scaled to the actual size of the substation demand. The standard grid retail tariff is £0.285 / kWh, with feed-in tariff of £0.075 / kWh [14], [15].

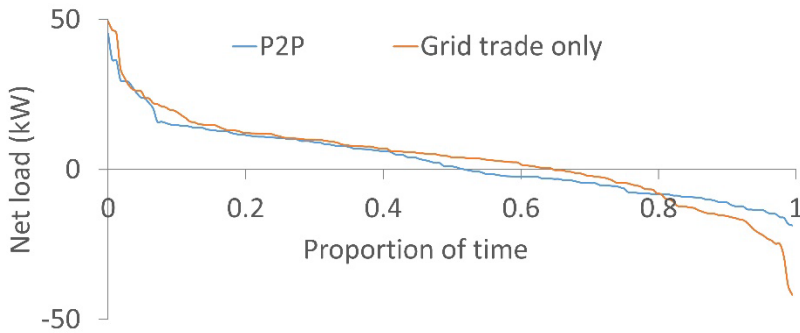
### 3. Results and discussion

We present results from a summer week with relatively high irradiance, which would typically lead to substantial reverse flow of power through the substation. Figure 2 illustrates how the P2P market operates over this week, showing the volumes of successful offers to buy and asks to sell, averaged as a diurnal profile. Also shown for comparison is the average P2P price by time of day. Overall turnover of electricity is 2.40 MWh representing £552 of transactions. Clearly, the main feature is the purchase of PV power in order to charge EVs in the middle of the day. Availability of solar power causes a trough in the price profile, although demand for this generation ensures that the P2P price stays well above the feed-in tariff. It is also interesting to note that small amounts of power for ‘essential’ EV charging are procured from the V2G power of neighbours’ cars.

Figure 3 shows the impact of the P2P trading on the load duration for the substation. Maximum demand at the substation is not significantly reduced. However, reverse power flow is markedly reduced, with the peak reverse flow cut by 48%. Overall imports and exports of grid electricity in energy terms are also reduced, as shown in Table 2. The P2P market also saves £3.29 over the week for the average customer, when compared with grid trade.



**Figure 2.** P2P electricity trades matched during the summer week with 2050 technology penetrations. (a) offers (b) asks. Shown are the volumes transacted as a daily profile averaged across the week, with the average P2P price for comparison.



**Figure 3.** Load duration curves at the grid connection for the simulated week.

**Table 2.** Summary metrics for summer week.

	Grid trade	P2P
Total exports (kWh)	713	601 (-16%)
Total imports (kWh)	1198	995 (-17%)
Peak exports (kW)	35.4	18.6 (-48%)
Peak imports (kW)	45.7	45.2 (-0.9%)
Shared power (kWh)	754	1890 (+ 151%)
Average net bill (£)	14.96	11.67 (-22%)

It is worth noting that around half of the P2P deals for EV charging are for arbitrage, where cheap power imported to the EV battery is discharged later (either to the home ‘V2H’ or to the grid ‘V2G’) when power is more expensive. This is clearly an important factor in enabling the P2P to increase PV self-consumption (which we also demonstrated in [10]). The cost of V2H / V2G in terms of battery degradation, represented here by  $c_{V2X}$ , may be a deciding factor in whether this technology actually becomes widespread. Certain

improvements are suggested to augment the applicability and power of the market model in the future:

- Allow trading more than one day in advance
- Incorporate forecasting errors for load and generation
- More trading strategies, e.g. adaptive-aggressive
- Allow non-physical trading
- More sophisticated price prediction

#### 4. Conclusions

A double auction model for P2P trading of electricity was developed, along with strategies for the engagement of flexible devices (particularly EV batteries) with this market. The market incentivises efficient behaviour of customers without imposing undue restrictions. Prices across the day respond to market forces, driven by the availability of generation. The P2P market reduces the burden on a substation operating with future penetrations of renewable technologies, with a significant reduction (-48%) in reverse power flow through the substation in sunny weather. Although a corresponding reduction in peak demand was not demonstrated, overall demand for imported electricity was reduced.

#### 5. Acknowledgements

This research makes use of data from the UKECN and UK Power Networks.

#### References

- [1] Y. Zhou, J. Wu, and C. Long, "Evaluation of peer-to-peer energy sharing mechanisms based on a multiagent simulation framework," *Appl. Energy*, vol. 222, no. November 2017, pp. 993–1022, 2018, doi: 10.1016/j.apenergy.2018.02.089.
- [2] Y. Liu, K. Zuo, X. (Amy) Liu, J. Liu, and J. M. Kennedy, "Dynamic pricing for decentralized energy trading in micro-grids," *Appl. Energy*, vol. 228, no. June, pp. 689–699, 2018, doi: 10.1016/j.apenergy.2018.06.124.
- [3] K. Chen, J. Lin, and Y. Song, "Trading strategy optimization for a prosumer in continuous double auction-based peer-to-peer market: A prediction-integration model," *Appl. Energy*, vol. 242, pp. 1121–1133, May 2019, doi: 10.1016/j.apenergy.2019.03.094.
- [4] The AnyLogic Company, "AnyLogic," 2019. <https://www.anylogic.com/>.
- [5] W. El-Baz, P. Tzscheutschler, and U. Wagner, "Integration of energy markets in microgrids: A double-sided auction with device-oriented bidding strategies," *Appl. Energy*, vol. 241, no. November 2018, pp. 625–639, 2019, doi: 10.1016/j.apenergy.2019.02.049.
- [6] Northern Powergrid, "Powering network data - Distribution Future Energy Scenarios," *northernpowergrid.com*, 2022. <https://www.northernpowergrid.com/network-data>.
- [7] Northern Powergrid, "Distribution Future Energy Scenarios," 2021. [Online]. Available: <https://www.northernpowergrid.com/asset/1/document/5836.pdf>.
- [8] T. D. Hutty, S. Dong, and S. Brown, "Suitability of energy storage with reversible solid oxide cells for microgrid applications," *Energy Convers. Manag.*, vol. 226, p. 113499, Dec. 2020, doi: 10.1016/j.enconman.2020.113499.
- [9] S. Rennie and J. Adamson, "UK Environmental Change Network (ECN) meteorology data: 1991-2015," *Centre for Ecology & Hydrology (Natural Environment Research Council)*, 2017. <https://doi.org/10.5285/fc9bcd1c-e3fc-4c5a-b569-2fe62d40f2f5>.
- [10] T. D. Hutty, A. Pena-Bello, S. Dong, D. Parra, R. Rothman, and S. Brown, "Peer-to-peer electricity trading as an enabler of increased PV and EV ownership," *Energy Convers. Manag.*, vol. 245, p. 114634, Oct. 2021, doi: 10.1016/j.enconman.2021.114634.
- [11] UK Government, "National Travel Survey," Jul. 2020.
- [12] UK Power Networks, "SmartMeter Energy Consumption Data in London Households," 2015.
- [13] E. McKenna and M. Thomson, "High-resolution stochastic integrated thermal–electrical domestic demand model," *Appl. Energy*, vol. 165, pp. 445–461, Mar. 2016, doi: 10.1016/j.apenergy.2015.12.089.
- [14] Ofgem, "Check if the energy price cap affects you," *ofgem.gov.uk*, 2022. <https://www.ofgem.gov.uk/check-if-energy-price-cap-affects-you>.
- [15] Solar Energy UK, "Smart Export Guarantee League Table," *solarenergyuk.org*, 2022.

# Comparing operational strategies for alkaline electrolysis systems considering a probabilistic wind power distribution

Lucas Cammann,<sup>a</sup> Johannes Jäschke<sup>a</sup>

<sup>a</sup> *Department of Chemical Engineering, Norwegian University of Science and Technology, NTNU, N-7491 Trondheim, Norway*

[johannes.jaschke@ntnu.no](mailto:johannes.jaschke@ntnu.no)

## Abstract

Coupling alkaline electrolysis processes with renewable energy sources requires rethinking electrolyzer operating practices that today are based on the assumption of constant power supply. Varying wind-power loads give rise to contrasting trade-offs in terms of safe and efficient operation. In this work, we study how keeping the lye flow rate and the pressure in the electrolyzer at their upper and lower bounds, respectively, compares with the policy of always adjusting these values optimally.

To further illustrate the implications of non-constant power supply, the strategies are analyzed considering a uniform, and a non-uniform power probability distribution. We show that currently employed strategies (e.g., high-pressure) may perform poorly when applied to systems subject to power profiles typical for renewable energy systems.

**Keywords:** Hydrogen, Renewable Energies, Optimization, Process Operation

## 1. Introduction

Increasing environmental awareness has led to a renewed interest in using Hydrogen (H<sub>2</sub>) as an energy vector for renewable electricity through electrolysis of water. Currently, most of the worldwide H<sub>2</sub> is produced by steam reforming of methane, with water electrolysis accounting for only 4% of the global production (Kumar and Himabindu, 2019). With rising CO<sub>2</sub> taxes and anticipated cost saving potentials for water electrolysis in the range of up to 54% by means of scale-up and R&D (Schmidt et al., 2017), this number is expected to grow considerably in the upcoming years. Nonetheless, the question of how to operate water electrolysis in conjunction with flexible power systems is largely unanswered, which is especially true for alkaline electrolysis according to Brauns and Turek (Brauns and Turek, 2022).

One fundamental consideration for the operation of alkaline electrolysis systems with renewable energy sources is how to safely reconcile the requirements for flexibility and process efficiency. While high pressures and lye flowrates improve the efficiency of the process, they also lead to an increased ratio of Hydrogen-to-Oxygen (*HTO*) in the effluent gas streams. For continuous operation, this metric may not exceed 50% of the lower explosion limit, which becomes particularly difficult in low load scenarios where gas production is reduced (Brauns and Turek, 2020). This work provides more insight into these trade-offs by expanding previous modelling and optimization efforts with further relevant mechanistic effects, such as bubble coverage and pressurization, introduced in

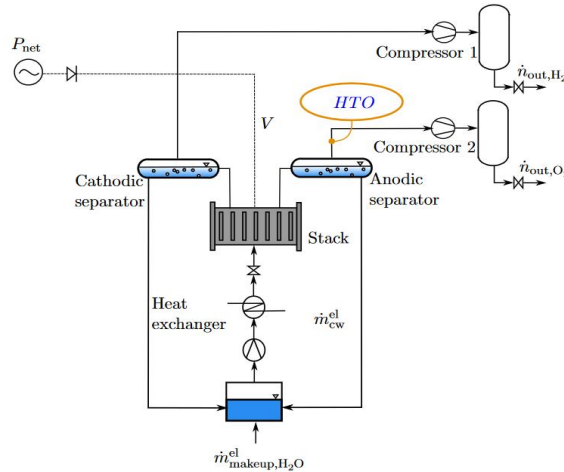
Section 2. The weighted production rate is proposed as comparative metrics and used to analyze the influence of setpoint strategies and power probabilities on the process performance in Section 3. Section 4 lastly discusses implications of the obtained results.

## 2. Model

The model developed in this work is based on previous work of Rizwan et al. (Rizwan et al., 2021), and further adapted to account for effects of pressurization and bubble development on the process operation. This section focuses on the expansion of said model, for a more in-depth description of the unit operations and parameter values the reader is referred to the original publication.

### 2.1. Plant description

Figure 1 shows the process flow diagram of the electrolysis process considered in this work. At the core of the process lies the alkaline water electrolyzer, which is fed by a liquid pump with the required lye (30 wt. % KOH) and which is connected to an AC/DC converter and a renewable power source. In the electrolyzer, the product gases  $H_2$  and  $O_2$  are formed at the cathode and anode, respectively, and must be separated from the liquid lye in two gravity settlers (separators). The gaseous phase is drawn from the top and compressed to the respective storage tanks, while the liquid lye of both separators is recycled to the buffer tank. Here they are mixed with a water make-up stream which replenishes the  $H_2O$  expended in the electrolysis reactions. The lye is then pumped through the heat-exchanger before being fed back to the electrolyzer.



**Figure 1:** Process flowsheet for alkaline water electrolysis.

### 2.2. Electrolyzer model

The electrolyzer model accounts for mass and energy balances as well as the current-voltage relationship. For the latter, the effective current density  $I_{\text{eff}}$  is introduced which depends on the bubble coverage at the electrode.

#### 2.2.1. Mass and energy balance

The energy balance for the electrolyzer is set up as in Ulleberg, (2003)

$$0 = \dot{m}_{\text{lye,in}}^{\text{el}} C p_{\text{lye}} (T_{\text{in}}^{\text{el}} - T^{\text{el}}) + n_c (U_k - U_{\text{tn}}) - A_s \left[ h_c (T^{\text{el}} - T_a) + \sigma \epsilon (T^{\text{el}4} - T_a^4) \right]. \quad (1)$$

Here, the term  $A_s$  denotes the active area w.r.t. radiation and convection,  $h_c$  the convective heat transfer coefficient in  $\text{W m}^{-2} \text{K}^{-1}$ ,  $\sigma$  the Stefan-Boltzmann constant in  $\text{W m}^{-2} \text{K}^{-4}$  and  $\epsilon$  and  $n_c$  the dimensionless emissivity constant and cell number, respectively. Component

balances are set up for the components H<sub>2</sub> and O<sub>2</sub> for the cathode and anode respectively, exemplarily shown below for the cathode

$$0 = \dot{n}_{in,H_2} + \dot{n}_{r,H_2} - \dot{n}_{diff,H_2} - \dot{n}_{out,H_2}^{cat} \quad (2)$$

$$0 = \dot{n}_{in,O_2} + \dot{n}_{diff,O_2} - \dot{n}_{out,O_2}^{cat} \quad (3)$$

The subscripts *in* and *out* distinguish in- and outgoing flows, while the subscripts *r* and *diff* describe reactive and diffusive fluxes, the latter two being calculated as follows for an arbitrary component *j*:

$$\dot{n}_{r,j} = \frac{I A^{el}}{z F} \quad \text{and} \quad \dot{n}_{diff,j} = \frac{p_{sys} S_j D_j A^{el}}{L_m} \quad (4)$$

In Eq. 4,  $A^{el}$  and  $L_m$  are the electrolytically active surface area and the thickness of the membrane, respectively, which are taken to be 598 m<sup>2</sup> and 0.005 m. The calculation of the solubility constant and the diffusion coefficient is further detailed in (Haug et al., 2017). The overall mass balance is closed by enforcing that the sum of in- and outgoing mass be equal.

### 2.2.2. Current-voltage relationship

The current-voltage relationship is assumed to follow the semi-empirical relationship proposed by Ulleberg (2003)

$$U = U_{rev} + (r_1 + r_2 T^{el}) I + \text{slog} \left( \left( t_1 + \frac{t_2}{T^{el}} + \frac{t_3}{T^{el^2}} \right) I + 1 \right) \quad (5)$$

Here,  $I$  is the current density in A m<sup>-2</sup>, while the terms  $r$ ,  $s$  and  $t$  are ohmic resistance and overvoltage parameters that need to be fitted or taken from literature. Eq. 5 does not account for influences of the pressure or the lye flowrate on the current-voltage relationship. The more important effect of the bubble coverage  $\theta$  on the effective current density can be modelled as follows:

$$I_{eff} = I \times (1 - \theta) \quad (6)$$

The bubble coverage itself is found by calculating the *stagnant bubble coverage*  $\theta_0$  (Vogt and Balzer, 2005) according to

$$\theta_0 = \left[ \frac{I}{I_{\theta \rightarrow 1}} \right]^m \quad (7)$$

which is corrected for by the lye flowrate (Eigeldinger and Vogt, 2000) through

$$\theta = \frac{\theta_0}{[1 + (K \times v)^2]^2} \quad (8)$$

In Eq. 7 the exponent  $m$ , the limiting current density  $I_{\theta \rightarrow 1}$  and  $K$  are fitting parameters, while  $v$  is the superficial velocity of the electrolyte in the electrolyzer. The value of  $m$  is set to 0.3, while the value of  $I_{\theta \rightarrow 1}$  is set to 7500 A m<sup>-2</sup> to indicate a low to medium bubble development. According to (Eigeldinger and Vogt, 2000), a value of  $K$  of 8 s m<sup>-1</sup> is representative for most experimental observations and is therefore also used in this work. The electrolyte velocity is calculated by means of continuity as

$$v = \frac{\dot{m}_{lye,in}^{el}}{\rho_{lye} \times A_{lyeflow}} \quad (9)$$

in which the area through which the lye flows  $A_{lyeflow}$  is calculated to be 3.63 m<sup>2</sup>, based on publicly available data.

### 2.3. Separator model

The molar component balances for the separators are set up similarly to the electrolyzer and shown exemplarily for the cathodic separator



$$\begin{aligned}
0 &= \dot{n}_{\text{out,H}_2}^{\text{cat}} - \dot{n}_{\text{out,H}_2}^{\text{sep,cat}} - \dot{n}_{\text{in,H}_2}^{\text{buffer}} = \dot{n}_{\text{out,H}_2}^{\text{cat}} - \dot{n}_{\text{out,H}_2}^{\text{sep,cat}} - 0.5\dot{V}_{\text{lye}}S_{\text{O}_2,\text{lye}}p_{\text{sys}}/\eta_{\text{sep}} \\
0 &= \dot{n}_{\text{out,O}_2}^{\text{cat}} - \dot{n}_{\text{out,O}_2}^{\text{sep,cat}}.
\end{aligned} \tag{10}$$

It is herein assumed that accumulation in the electrolyte is only relevant for the main gaseous species entering the anodic or cathodic separator. The solubility constants  $S_{i,\text{lye}}$  are calculated considering the mass fraction of the lye  $w_{\text{lye}}$  (Haug et al., 2017)

$$S_{i,\text{lye}} = \frac{S_{i,\text{H}_2\text{O}}}{10^{K_i w_{\text{lye}}}}, \quad \text{with } S_{i,\text{H}_2\text{O}} = \frac{\rho_{\text{H}_2\text{O}}}{M_{\text{H}_2\text{O}} p_{\text{atm}} H_{i,\text{H}_2\text{O}}}. \tag{11}$$

Here,  $K_i$  is the Setchenov constant and  $H_{i,\text{H}_2\text{O}}$  the Henry coefficient of species  $i$  in water, the values and correlations for both being taken from Haug et al.

The *HTO* is measured at the outlet of the anodic separator as

$$\text{HTO} = \frac{\dot{n}_{\text{out,H}_2}^{\text{sep,an}}}{\dot{n}_{\text{out,O}_2}^{\text{sep,an}}}. \tag{12}$$

#### 2.4. Power distribution modelling and weighted production rate

We consider two different power availability probability functions. These probability functions account for the uncertainty in the supply of renewable energy for the process, which is assumed to stem exclusively from wind power. In the first case, all power levels are assumed to be equally likely (i.e., a uniform probability distribution). In the second case, the power is assumed to follow the probability density function for the wind velocity, which is commonly modeled as a Weibull distribution:

$$f(\mathbf{x}, \lambda, k) = \begin{cases} \frac{\lambda}{k} \left(\frac{\mathbf{x}}{\lambda}\right)^{k-1} e^{-(\mathbf{x}/\lambda)^k}, & \mathbf{x} \geq 0 \\ \mathbf{x}, & \mathbf{x} < 0 \end{cases} \tag{13}$$

in which  $k$  is called the shape parameter,  $\lambda$  the scale parameter and  $\mathbf{x}$  the probabilistic variable. The generated power is calculated from the obtained wind profile according to

$$P_{\text{wind}} = \frac{1}{2} \times \rho_{\text{air}} \times A_{\text{blade}} \times v_{\text{wind}}^3, \tag{14}$$

in which  $A_{\text{blade}}$  is scaled such that the delivered power is equal or less to 2 MW in 95% of cases, based on wind speed data of the month April for the years 1991 to 2000 in Zhuhire, China (Yingni et al., 2006). To gauge how efficiently an operating concept can handle flexible power supplies, a new comparative metrics is proposed, the weighted production rate  $\overline{m}_{\text{out,H}_2}$ . This figure convolutes the absolute production rate at a certain power input with the probability of this power occurring,

$$\overline{m}_{\text{out,H}_2} = \Pr(P_{\text{net}}) \times \dot{m}_{\text{out,H}_2}. \tag{15}$$

The probability  $\Pr(P_{\text{net}})$  is computed numerically from the probability density function  $f(P_{\text{net}})$ , herein taken to follow either a uniform, or a Weibull distribution.

#### 2.5. Setpoint strategies

In this work, different fixed setpoint strategies for the lye flowrate and the pressure are compared in the face of varying power distributions. Strategy 1 entails to operate both flexibly (optimal), while Strategies 2 and 3 operate both at their upper and lower limit, respectively. Table 1 summarizes the values that are used at the basis of these strategies.

**Table 1:** Overview of setpoint strategies.

	Unit	Strategy 1	Strategy 2	Strategy 3
$p$	bar	Flexible	20	10
$\dot{m}_{\text{lye}}$	kg s <sup>-1</sup>	Flexible	30	15

### 2.6. Optimization problem

The model developed in this work is optimized for values of the available Power input  $P_{net}$  in the range of 15 kW to 2 MW according to the following formulation:

$$\begin{aligned}
 \max \quad & \dot{m}_{out,H_2} && (16) \\
 \text{s.t.} \quad & \dot{m}_{lye,min} \leq \dot{m}_{lye,k}^j \leq \dot{m}_{lye,max} && \forall k \in [in, out], j \in [el, buff] \\
 & -20 \text{ }^\circ\text{C} \leq \Delta T_{in}^{el} \leq 20 \text{ }^\circ\text{C} \\
 & 10 \text{ }^\circ\text{C} \leq T^{el} \leq 80 \text{ }^\circ\text{C} \\
 & p_{min} \leq p_{sys} \leq p_{max} \\
 & P_{comp} + P_{el} \leq P_{net} \\
 & \dot{m}_{cw}^{el} \leq 26.6 \text{ kg s}^{-1} \\
 & HTO \leq 2\%
 \end{aligned}$$

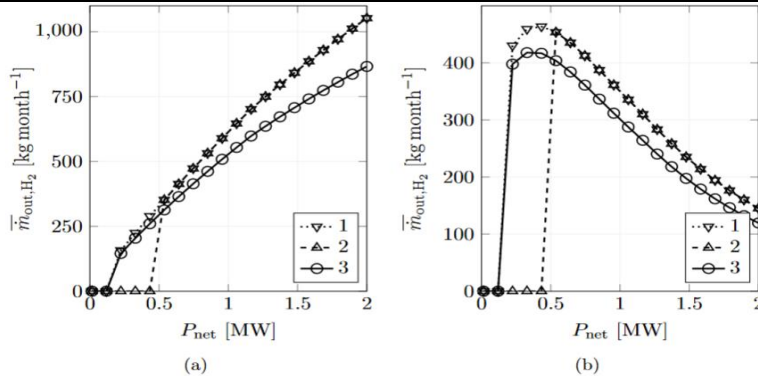
For strategies 2 and 3,  $p_{min}$  and  $p_{max}$  and  $\dot{m}_{lye,min}$  and  $\dot{m}_{lye,max}$  are set to the respective constant values indicated in Table 1, while these setpoints represent the upper and lower bounds for strategy 1. The optimization problem is solved in the programming language Julia v. 1.7.2. with the NLP solver IPOPT. In scenarios where the outlined constraints cannot be satisfied the plant is assumed to be shut down, and the production rate is equated to zero.

### 3. Results

Figure 2 shows the weighted production rate for strategies 1 to 3 for both considered probability distributions and different values of  $P_{net}$ . The cumulative monthly production rates are shown in Tab. 2 and used as comparative metric. Generally, these values are lower for the Weibull distribution due to it being right skewed. Fully flexible operation (strategy 1) maximizes Hydrogen production for both probability distributions. Considering a uniform probability distribution, the second most efficient setpoint strategy involves operating lye flow rate and pressure at their upper limit (strategy 2), with the mean production rate being 11% higher than lower limit operation (strategy 3).

**Table 2:** Total monthly production rate for strategies 1 to 3 and both probability distributions.

	Distribution	Unit	Strategy 1	Strategy 2	Strategy 3
$\dot{m}_{out,H_2}$	Uniform	kg month <sup>-1</sup>	11,617	10,946	9,862
$\dot{m}_{out,H_2}$	Weibull	kg month <sup>-1</sup>	5,715	4,362	4,977



**Figure 2:** Weighted production rate considering (a) uniform, and (b) Weibull power distribution.

This however reverses under consideration of the Weibull power availability distribution, where the operation at the lower limit of the pressure and the lye flowrate achieves a 14% higher mean production rate than the operation at the respective upper limits. The reason for this can be seen in Fig. 2, which shows that strategy 3 is inoperable at lower power inputs due to the violation of the *HTO*-constraint, leading to a production rate of zero up until ~0.5 MW. Operation in this low load region has a higher probability according to the Weibull distribution, leading to it being weighted more strongly than higher power inputs. The loss of Hydrogen production due to plant shutdown at lower power inputs in strategy 3 can therefore not be compensated for by the efficiency gain at high power inputs.

#### 4. Conclusion and Outlook

Current paradigm in the field of alkaline electrolysis is to operate with high pressures and lye flowrates in order to reduce both operating and investment costs, as well as to improve the efficiency of the process.

The herein presented results however highlight the need to critically assess the transferability of these design notions to the envisioned flexible operation of alkaline water electrolysis. For example, operating at lower powers and lye flowrates might allow for higher Hydrogen production rates in the face of more frequent low power scenarios as compared to the high setpoint strategy. While the optimal operational strategy remains subject to various other considerations, more detailed modelling and optimization efforts are required to elucidate improved strategies for flexible operation. For instance, different lye recycling strategies, such as separate lye loops with recombination phases, might be used to overcome operational limits in low power scenarios. More importantly, such strategies must be evaluated in a dynamic context to account for the intermittencies occurring on different timescales through the supply of renewable energy.

#### References

- S. Shiva Kumar; V. Himabindu, 2019, Hydrogen production by PEM water electrolysis – A review, *Materials Science for Energy Technologies* 2.3, pp. 442–454.
- J. Brauns and T. Turek, 2022, Experimental evaluation of dynamic operating concepts for alkaline water electrolyzers powered by renewable energy, *Electrochimica Acta*, 404 p. 139715
- O. Schmidt et al., 2017, Future cost and performance of water electrolysis: An expert elicitation study, *International Journal of Hydrogen Energy* 42.52, pp. 30470–30492
- J. Brauns, T. Turek, 2020, Alkaline Water Electrolysis Powered by Renewable Energy: A Review., *Processes*, Volume 8(2), 248
- M. Rizwan, V. Alstad, J. Jäschke, 2021, Design considerations for industrial water electrolyzer plants, *International Journal of Hydrogen Energy*, 46.75, pp. 37120–37136
- H. Vogt, R.J. Balzer, 2005, The bubble coverage of gas-evolving electrodes in stagnant electrolytes, *Electrochimica Acta*, 50.10, pp. 2073-2079
- J. Eigeldinger, H. Vogt, 2000, The bubble coverage of gas-evolving electrodes in a flowing electrolyte, *Electrochimica Acta*, 45.27, pp. 4449-4456
- Ø. Ulleberg, 2003, Modelling of advanced alkaline electrolyzers: a system simulation approach, *Hydrogen Energy*, 28, pp. 21-33
- P. Haug et al., 2017, Process modelling of an alkaline water electrolyzer, *International Journal of Hydrogen Energy*, 42.24, pp. 15689-15707
- J. Yingni et al., 2006, Wind Potential Assessment Using the Weibull Model at the Inner Mongolia of China, *Energy Exploration & Exploitation*, 24.3, pp. 211-221

# Conceptual design of pre-treatment units for co-electrolysis of CO<sub>2</sub> and water

Josephine Vos<sup>a</sup>, Andrea Ramirez<sup>a</sup>, Mar Pérez-Fortes<sup>a</sup>

<sup>a</sup>*Department of Engineering Systems and Services, Faculty of Technology, Policy and Management, Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands.*

## Abstract

Reaching our climate goals will require urgent advancements in the development of fossil-free technologies. Solid-oxide electrolysis (SOE) at high-temperature is a promising candidate for combining CO<sub>2</sub> utilization and renewable electricity use. Explorative techno-economic analyses are being performed to understand the full plant design requirements for integrated SOE systems. However, there is still a lack of understanding of the potential impact that the pre-treatment of CO<sub>2</sub> will have on the overall design and economics of a SOE-based system. To address this knowledge gap, as a first step, the process model of the pre-treatment units needed to purify CO<sub>2</sub> from a bioethanol plant is developed in Aspen Plus in the current work. Based on the preliminary results of this paper, the equipment costs mainly stem from the units related to the removal of sulfur (~65%) and alcohols (~32%). The energy costs are almost entirely related to the cryogenic distillation step required for the removal of non-condensable gases (~96%).

**Keywords:** solid-oxide electrolysis, co-electrolysis, syngas, pre-treatment, ex-ante techno-economic analysis, low carbon technologies

## 1. Introduction

Net-zero CO<sub>2</sub> emissions must be reached by 2050 to reduce the disastrous effects of climate change (IEA, 2020). Carbon dioxide electrochemical reduction (CO<sub>2</sub>ER) has the potential to mitigate CO<sub>2</sub> emissions as it combines CO<sub>2</sub> utilization and renewable (intermittent) electricity for the synthesis of chemicals. A specific type of CO<sub>2</sub>ER technology is the solid-oxide electrolyzer (SOE). Co-electrolysis of CO<sub>2</sub> and water in a SOE has a technology readiness level (TRL) between 5-7, and it can be used to generate syngas (Zheng et al., 2017). To reach higher TRLs and eventually large-scale commercial deployment, explorative techno-economic analyses (TEAs) are underway to better understand the full plant design requirements (for example, Gao et al., 2020; Herz et al., 2018).

SOE TEAs assume that future CO<sub>2</sub>-based plants will be operated using CO<sub>2</sub> from industrial point sources or direct air capture (DAC). It is also often assumed that this CO<sub>2</sub> stream is pure and available at the gate of the SOE plant. However, this may not be the case. The CO<sub>2</sub> purity level of several industrial sources can vary between ~30-96 mol% (Verma et al., 2019). Some SOE TEAs mention that purifying the CO<sub>2</sub> will be needed before entering the electrolyzer as experimental studies have indicated that even a slight concentration of impurities (for instance, sulfur or heavy metals) can alter the catalytic performance significantly or deactivate the catalyst entirely (Gao et al., 2020; Kibria et

al., 2019). However, neither the target purity levels nor the required types of purification steps have been specified or modeled in SOE TEAs literature.

Therefore, there is currently a lack of understanding of the potential impact of pre-treatment units on the design and economics of SOE plants. To address this knowledge gap, in this study an ex-ante TEA of the pre-treatment units for a CO<sub>2</sub> stream from a bioethanol plant for co-electrolysis via SOE to produce syngas is performed. The aim of this paper is therefore twofold: 1) to propose a conceptual design for pre-treatment units of CO<sub>2</sub> for future SOE plants, and 2) to evaluate the plant performance. In the upcoming section, the scope, modeling, and evaluation steps are described. In the subsequent sections, the preliminary results are discussed.

## 2. Methodology

### 2.1. Scope

The current techno-economic analysis consists of a ‘gate-to-gate’ ex-ante assessment. Figure 1a presents an overview of the overall SOE-based plant, and the current system boundaries. The electricity is considered to be continuous and acquired from the grid.

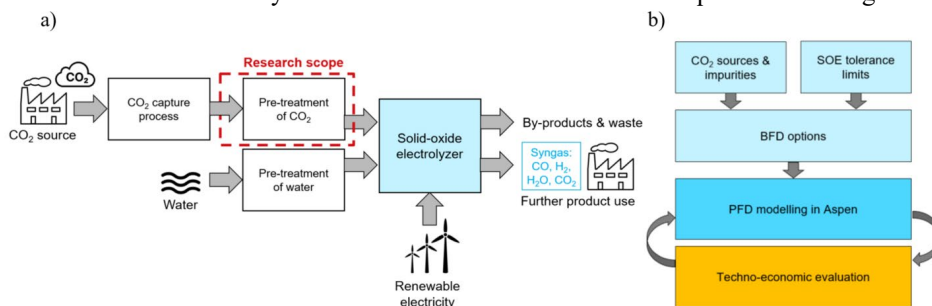


Figure 1: a) System boundary definition, focus on pre-treatment of CO<sub>2</sub>. b) Methodological steps for modeling and evaluation (BFD: block flow diagram, PFD: process flow diagram).

The methodological framework for this research follows the approach for analyzing future ‘N-of-a-kind’ (NOAK) plants (Roussanaly et al., 2021). Since this research involves an ex-ante analysis of a low TRL technology, this means that the entire analysis is based on a ‘what-if’ approach where the SOE is treated as if it was already commercially available today. It is assumed that the performance at industrial scale is the same as currently demonstrated in lab conditions. The model considers a CO<sub>2</sub>-based plant located in Northwestern Europe and installed in 2019. The scale of the plant corresponds to the production of 1 tonne of pure CO<sub>2</sub> per year (8000 hours).

### 2.2. Data collection, modeling & evaluation

The methodology followed for the data collection, modeling and evaluation of the pre-treatment configuration for future SOE plants is illustrated in Figure 1b. A targeted literature review was performed to collect information regarding (i) possible CO<sub>2</sub> sources and their compositions, and (ii) the electrolyzer’s tolerance levels per contaminant (considering a SOE Ni/YSZ cermet fuel electrode). Assumptions were made based on existing literature for solid-oxide fuel cells and water electrolysis, and information from experts in the field. Next, a screening of potential cleaning technologies for the removal of contaminants was done based on heuristics, temperature and pressure considerations to create a block flow diagram (BFD). For the current bioethanol case, the pilot plant

study from McKaskle et al., the research on purification units from Quevedo et al., and a CO<sub>2</sub> purification patent by Gupta et al. served as a starting point (Gupta, 1999; McKaskle, 2018; Quevedo, 2021). The subsequent process flow diagram (PFD) was created in Aspen Plus. The model was validated, unit-by-unit, using data from literature. The Aspen model was used to generate data on the equipment dimensions, mass and energy balances. Finally, key performance indicators (KPIs), i.e. CO<sub>2</sub> losses during the process, energy needs, and bare equipment costs, were calculated.

### 3. Results & discussion

The main impurities in the selected bio-based CO<sub>2</sub> stream are non-condensable gases (NCDS), such as oxygen and nitrogen, as well as sulfur components, water, alcohols, and other hydrocarbons (see Table 1). The investigation into the target purity levels for SOE revealed that even trace levels of sulfur components are likely to negatively affect the degradation rate of the electrodes (Caliandro, 2018; Rinaldi, 2019). The other identified contaminants from the bioethanol plant are expected to be less problematic for the SOE, but experimental data on the effect of these components is still limited in the current literature, for instance, for alcohols. In the current work, it is considered a conservative goal to reach at least ppb-level of the contaminants that are less known.

Table 1: Overview composition of the CO<sub>2</sub> stream from a bioethanol plant (i.e., the fermentation off-gas from an ethanol plant using corn as the biomass source) including major and trace impurities alongside tolerance limits for SOE (Caliandro, 2018; McKaskle, 2018; National Academies of Sciences, Engineering, 2022; Rinaldi, 2019; Yokokawa et al., 2008).

Composition	Bioethanol plant	Unit used	Degradation type	Tolerable amount	Unit used
CO <sub>2</sub>	90	%	C-poisoning		
CO	1	ppm	C-poisoning		
H <sub>2</sub> O	5	ppm			
CH <sub>4</sub>	3	ppm	C-poisoning		
SO <sub>x</sub>	1	ppm	S-poisoning	<2	ppm
NO <sub>x</sub>	1	ppm			
O <sub>2</sub>	100	ppm		<5	%
H <sub>2</sub> S	1	ppm	S-poisoning	~0.05	ppm
N <sub>2</sub>	98768	ppm			
Ar	-	ppm		~100	μmol/mol
Heavy metals	-	ppm	Metals	-	ppb-level
Cl	-	ppm	Cl-poisoning	<5	ppm
Alcohols	3-950	ppm			
Other hydrocarbons	1	ppm	C-poisoning	~2	μmol/mol
Aromatics (benzene, toluene, xylene)	3	ppm			
Carbonyl sulfide	1	ppm	S-poisoning	<2	ppm
Dimethylsulfide	1	ppm	S-poisoning	<2	ppm
Ethers	1	ppm			
Ketons	1	ppm			
Mercaptans	1	ppm			

Figure 2 shows the design of the preliminary CO<sub>2</sub> purification train for the removal of the contaminants listed in Table 2. In total, four different purification sections are needed to achieve this level of purity, namely (i) alcohols removal, (ii) sulfur removal, (iii) moisture removal, and (iv) non-condensable gases removal. The global property method Redlich-Kwong-Soave (RKS-BM) was selected in Aspen Plus, which is suitable for gas processes.

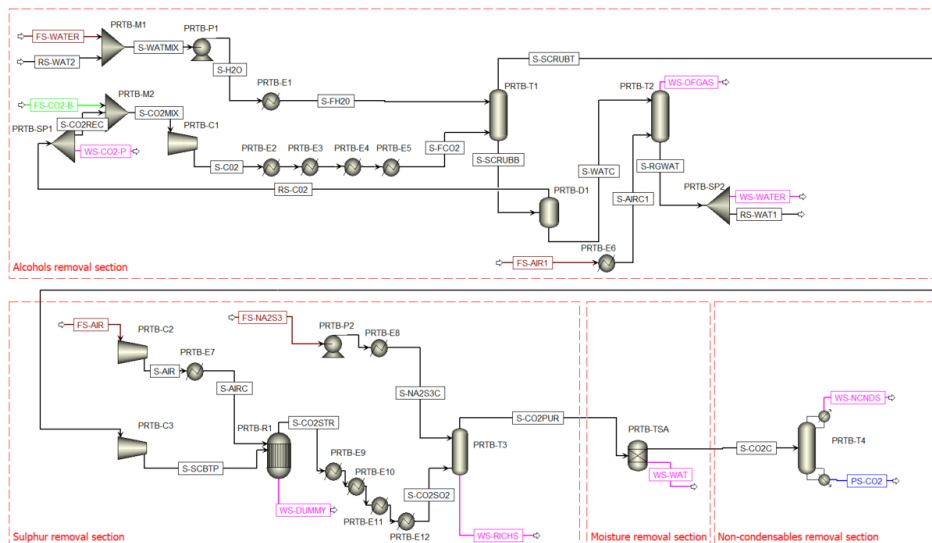


Figure 2: Aspen PFD for the purification of a CO<sub>2</sub> stream from a bioethanol plant to reach SOE purity levels.

The CO<sub>2</sub> from the bio-ethanol plant enters the process at 38°C and 0.9 atm (McKaskle, 2018) and is conditioned to 20°C and 10 bar before being sent to a pressurized water scrubbing tower (RadFrac column – PRTB-T1) for the removal of alcohols (Abu Seman & Harun, 2019). The water scrubbing tower removes the traces of alcohols as well as ~25% of the CO<sub>2</sub>, therefore the removed CO<sub>2</sub> is separated from the scrubbing water via a flash drum (Flash2 – PRTB-D1) and is recycled back (RS-CO<sub>2</sub>). The scrubbing water is regenerated using a counter-current hot air stream in a second tower (RadFrac column – PRTB-T2). Next, the CO<sub>2</sub> stream is mixed with air in a catalytic oxidation unit (REquil – PRTB-R1) operating at 400°C and 22 bar to remove the hydrogen sulfide and COS by converting them into sulfur dioxide and water (Gupta, 1999). The sulfur dioxide is removed by using an aqueous solution of Na<sub>2</sub>SO<sub>3</sub> in an absorption tower (RadFrac – PRTB-T3) at 38°C and 22 bar. Then, the remaining moisture is removed using a temperature swing adsorption (TSA) unit consisting of 2-4 columns (Gupta, 1999). The TSA was modeled in Aspen Plus using a separator block (i.e., black-box model for which the energy and economic performance are not yet evaluated in the current work). Finally, the non-condensable gases are removed via cryogenic distillation (RadFrac column – PRTB-T4) at 22 bar.

Table 2 summarizes the preliminary techno-economic results of the CO<sub>2</sub> purification train from Figure 2. The CO<sub>2</sub> from the bioethanol source was purified to 99.9984 wt% with water, carbon monoxide, methane, and argon as trace impurities. The purity level is in agreement with the tolerance limits for SOE (as shown in Table 1) and matches the specifications of food-grade CO<sub>2</sub> (i.e., 99.9% (Parker Hannifin Ltd, 2022)). During the purification process, about ~21% of the CO<sub>2</sub> inlet stream is lost as gaseous emissions despite recycling efforts. Possibilities for further improvement of CO<sub>2</sub> recycling should be investigated to improve the process in the future.

In terms of cost, the total bare equipment costs of the purification train are almost 3M€. The analysis of the different purification sections reveals that the equipment costs related to the removal of sulfur (~65%) and alcohols (~32%) are responsible for most of the total

equipment cost. The energy costs are almost entirely related to the removals of non-condensable gases (~96%) due to the cryogenic distillation step.

Table 2: Energy and economic preliminary results for the purification of the bioethanol CO<sub>2</sub> stream. The results are stated per tonne of clean CO<sub>2</sub> (VHPS: very high pressure steam at 100 bar, HPS: high pressure steam at 51 bar, MPS: medium pressure steam at 21 bar, LPS: low pressure steam at 5.5 bar, LLPS: low low pressure steam at 3.9 bar, CW: cooling water ~25-40°C, CHILLED: chilling water ~5-7.5°C, R50: refrigerant methane ~-162°C, NA: energy and equipment costs not available for the TSA unit, NCDS refers to non-condensable elements such as nitrogen, argon and oxygen).

Section	Energy needs									Energy costs	Bare equipment cost
	VHPS	HPS	MPS	LPS	LLPS	CW	CHILLED	R50	Electricity		
Unit	GJ/t	GJ/t	GJ/t	GJ/t	GJ/t	GJ/t	GJ/t	GJ/t	kWh/t	€/t	k€
Alcohols	-	-	-0.06	-0.10	0.30	0.23	0.04	-	115.29	10.80	930
Sulfur	-	1.00	-0.31	-0.31	0.75	0.76	-	-	554.29	48.95	1875
Moisture	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NCDS	-	-	-	-	2.80	-	-	3.80	-	1618.11	101
<b>Total</b>	-	<b>1.00</b>	<b>-0.38</b>	<b>-0.41</b>	<b>3.86</b>	<b>0.99</b>	<b>0.04</b>	<b>3.80</b>	<b>669.58</b>	<b>1677.86</b>	<b>2907</b>

#### 4. Conclusion & future work

A preliminary ex-ante techno-economic assessment of the pre-treatment units for purifying CO<sub>2</sub> from a bioethanol plant has been performed, as a first step towards the modeling of a complete SOE-based plant for co-electrolysis of CO<sub>2</sub> and water to produce syngas. This study aimed to propose an initial PFD and to analyze selected KPIs. Required CO<sub>2</sub> purity levels for SOE were identified based mostly on solid oxide fuel cells and water electrolysis literature. Four different purification sections were proposed to purify the CO<sub>2</sub> stream resulting in a CO<sub>2</sub> purity grade similar to food-grade CO<sub>2</sub>. The removal of sulfur (~65%) and alcohols (~32%) were the largest contributors to the overall equipment costs. The removal of non-condensable gases via cryogenic distillation was the main contributor to the overall energy costs (~96%). In future work, trade-offs will be further investigated by modeling the TSA unit using RadFrac columns and the possibilities for recycling flows, heat integration, and waste valorization will be further considered. Also, the purification of the water inlets stream will be analyzed.

#### Acknowledgements

This research receives funding from the project "Addressing the multiscale challenge of CO<sub>2</sub> electrochemical reduction", NWO ECCM tenure track grant (project number ECCM.TT.009), and "Sustainable design of multiscale CO<sub>2</sub> electrochemical conversion", MVI grant (ECCM.TT.MVITU.006).

#### References

- Abu Seman, N., & Harun, N. (2019). Simulation of pressurized water scrubbing process for biogas purification using Aspen Plus. *IOP Conference Series: Materials Science and Engineering*, 702(1). <https://doi.org/10.1088/1757-899X/702/1/012040>
- Caliandro, P. (2018). *Identification of Solid Oxide Cell Elementary Processes by Electrochemical Impedance Spectroscopy PAR*. PhD thesis number 8389 ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE.



- Gao, N., Quiroz-Arita, C., Diaz, L. A., & Lister, T. E. (2020). Intensified co-electrolysis process for syngas production from captured CO<sub>2</sub>. *Journal of CO<sub>2</sub> Utilization*, 43(May 2020), 101365. <https://doi.org/10.1016/j.jcou.2020.101365>
- Gupta, H. &. (1999). *Patent CO<sub>2</sub> purification system*. 15(2), 1–23. US patent number: 99108085.4.
- Herz, G., Reichelt, E., & Jahn, M. (2018). Techno-economic analysis of a co-electrolysis-based synthesis process for the production of hydrocarbons. *Applied Energy*, 215(February), 309–320. <https://doi.org/10.1016/j.apenergy.2018.02.007>
- IEA. (2020). Energy Technology Perspectives. *Energy Technology Perspectives*. <https://doi.org/10.1787/9789264109834-en>
- Kibria, G., Edwards, J. P., Gabardo, C. M., Dinh, C., Seifitokaldani, A., Sinton, D., & Sargent, E. H. (2019). *Electrochemical CO<sub>2</sub> Reduction into Chemical Feedstocks : From Mechanistic Electrocatalysis Models to System Design*. 1807166, 1–24. <https://doi.org/10.1002/adma.201807166>
- McKaskle. (2018). Evaluation of Carbon Dioxide Capture Options from Ethanol Plants. *University of Illinois, Circular 5*, ILLINOIS STATE GEOLOGICAL SURVEY Prairie Research Institute University of Illinois at Urbana-Champaign, <http://www.isgs.illinois.edu>.
- National Academies of Sciences, Engineering, and M. (2022). *Carbon Dioxide Utilization Markets and Infrastructure : Status and Opportunities : A First Report ( 2022 )*. <https://doi.org/10.17226/26703>
- Parker Hannifin Ltd. (2022). *CO<sub>2</sub> polishing for the sparkling beverage industry*, Accessed via [https://www.parker.com/literature/donnick%20hunter%20Industrial%20Division/Static%20Literature%20Files%20PDFs/Beverage/174004466\\_EN.PDF](https://www.parker.com/literature/donnick%20hunter%20Industrial%20Division/Static%20Literature%20Files%20PDFs/Beverage/174004466_EN.PDF).
- Quevedo, 2021. *Techno-economic and environmental comparative assessment of two renewable methanol production routes*, TU Delft master thesis, accessible via <http://resolver.tudelft.nl/uuid:923d22dd-b711-46ff-9fc3-a3ba695e31bd>.
- Rinaldi, G. (2019). *Long-term evolution of solid oxide fuel and electrolysis cell 3-D microstructure*. PhD thesis number 296. ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE. <http://infoscience.epfl.ch/record/265161>
- Roussanaly, S., Rubin, E. S., Van Der Spek, M., Booras, G., Berghout, N., Fout, T., Garcia, M., Gardarsdottir, S., Nair Kuncheekanna, V., Matuszewski, M., McCoy, S., Morgan, J., Mohd Nazir, S., & Ramirez, A. (2021). *Towards improved guidelines for cost evaluation of carbon capture and storage*.
- Verma, S., Lu, S., & Kenis, P. J. A. (2019). Co-electrolysis of CO<sub>2</sub> and glycerol as a pathway to carbon chemicals with improved technoeconomics due to low electricity consumption. *Nature Energy*, 4(6), 466–474. <https://doi.org/10.1038/s41560-019-0374-6>
- Yokokawa, H., Tu, H., Iwanschitz, B., & Mai, A. (2008). Fundamental mechanisms limiting solid oxide fuel cell durability. *Journal of Power Sources*, 182(2), 400–412. <https://doi.org/10.1016/j.jpowsour.2008.02.016>
- Zheng, Y., Wang, J., Yu, B., Zhang, W., Chen, J., Qiao, J., & Zhang, J. (2017). A review of high temperature co-electrolysis of H<sub>2</sub>O and CO<sub>2</sub> to produce sustainable fuels using solid oxide electrolysis cells (SOECs): Advanced materials and technology. *Chemical Society Reviews*, 46(5), 1427–1463. <https://doi.org/10.1039/c6cs00403b>

# A Process Design, Intensification, and Modularization Approach for Membrane-Assisted Reaction Systems

Yuhe Tian,<sup>a\*</sup> Ayooluwa Akintola,<sup>a</sup> Benjamin Akoh<sup>a</sup>

<sup>a</sup> *Department of Chemical and Biomedical Engineering, West Virginia University, Morgantown, West Virginia, United States*

## Abstract

We present a generalized process modeling and synthesis approach to design optimal modular intensified processes, with particular focus on membrane-assisted reaction systems. The approach is built on the phenomena-based Generalized Modular Representation Framework (GMF), which uses abstract mass/heat exchange modules to model diverse conventional and intensified chemical processes in a bottom-up manner. Gibbs free energy-based driving force constraints are developed to intensify the mass/heat transfer in these modules toward thermodynamic limits, and to dictate the optimal transfer format without pre-postulation of tasks (e.g., liquid-vapor separation, reaction, membrane permeation). Orthogonal Collocation on Finite Elements (OCFE) is applied to model the spatial distribution information within each module. The GMF/OCFE representation also enables the module sizing to be self-adjustable via process synthesis optimization, thus allowing for automatic selection of scaling up or numbering up while retaining computational compactness in the combinatorial space. A case study on hydrogen production via methane steam reforming is presented to showcase the proposed approach.

**Keywords:** Process Synthesis, Modular Process Intensification, Membrane Reactor.

## 1. Introduction

A central research question of modular chemical process intensification (MCPI) lies in how to exploit the synergy of multi-functional physicochemical phenomena at the optimal temporal and spatial scale to substantially improve process efficiency and sustainability (Stankiewicz et al., 2019). As a representative MCPI technology, membrane reactors integrate chemical reaction and membrane separation within a single unit to overcome reaction equilibrium limitation via selective mass transfer (Bernardo and Drioli, 2021). The intrinsic modularity of these systems also opens up the potential for distributed stranded resource utilization and small-scale on-site production (Sjardin et al., 2006).

To rigorously describe the multi-physics interactions in membrane reactors (e.g., reaction, sorption, diffusion, permeation), one-dimensional or two-dimensional first principles models are typically developed for simulation pertaining to pre-postulated reactor configuration and membrane characteristics (Karagöz et al., 2019). The resulting large-scale process-specific partial differential-algebraic equations (PDAE) pose challenges to synthesize optimal membrane reactor designs with structural variations or to incorporate systems-level process integration and intensification. Several recent works have applied phenomena building block-based process representation and synthesis approaches to design membrane-assisted reaction systems (Tula et al., 2019). This offers the potential to rapidly screen and quantitatively evaluate different membrane reactor structures and

unit/flowsheet intensification schemes at an early design stage (Monjur et al., 2021; Bishop and Lima, 2021; Babı et al., 2014). However, key open research questions remain on: (i) how to systematically identify the optimal reaction, separation, or hybrid schemes toward ultimate process performance limits, and (ii) how to simultaneously account for the impact of scaling up versus numbering up.

To address these challenges, in this work, we introduce a process intensification synthesis approach for membrane-assisted reaction systems built on the Generalized Modular Representation Framework (GMF) coupled with Orthogonal Collocation on Finite Elements (OCFE). The remainder of this paper is structured as follows. Section 2 presents the physics and mathematics fundamentals for GMF/OCFE modeling of membrane-assisted systems. Section 3 demonstrates the proposed approach on hydrogen production via methane steam reforming. Section 4 discusses ongoing work and future directions.

## 2. Process Intensification Synthesis Approach

### 2.1. An Overview on GMF

As illustrated in Fig. 1, Generalized Modular Representation Framework (Pistikopoulos and Tian, 2022) uses two sets of phenomena modules to capture the vast design space of unit operations and flowsheets from a lower-aggregated level:

- Mass/Heat exchange module – mass and heat transfer take place between the liquid and/or vapor streams into the module in the form of reaction, separation, reactive separation, reactive separation, etc.
- Heat exchange module – heat transfer takes place between the process and/or utility streams in the form of heating, cooling, heat integration, etc.

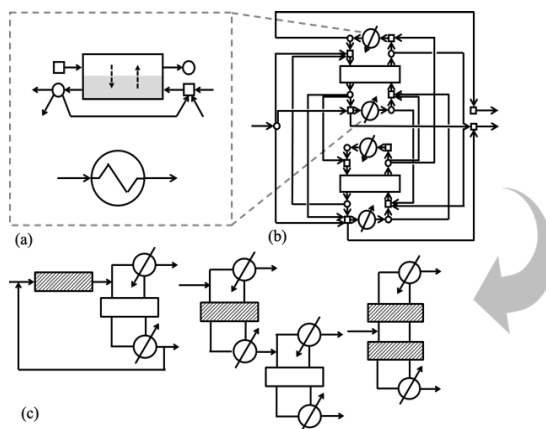


Figure 1: A schematic of GMF synthesis approach.

To characterize the mass/heat transfer feasibility in GMF modules (Fig. 1a), a novel driving force constraints formulation has been developed for conventional mass transfer process systems (e.g., reactors, absorption, distillation), based on total Gibbs free energy change ( $d(nG)_{T,P}^{tot} = \sum_i \mu_i dn_i \leq 0$ ). This enables to: (i) Intensify, and bound, mass transfer in each module against the generalized thermodynamic limits, and (ii) Automatically explore the synergistic or exclusive interactions of separation and reaction phenomena as illustrated in Tian and Pistikopoulos (2018). Physical and computational compactness can thus be attained via a single type of mass/heat exchange module while distilling the multi-functional synergy from thermodynamic fundamentals, instead of enumerating an extensive series of reaction, separation, and reactive separation modules. A superstructure network (Fig. 1b) is then constructed to encapsulate all plausible structural connections of mass/heat exchange modules, heat exchange modules, and process streams. The overall GMF synthesis model is formulated as a mixed-integer nonlinear programming (MINLP) model. Optimal process design solutions, represented using GMF modules (Fig. 1c), can be generated without pre-specifying conventional or intensified unit/flowsheet configurations to innovate the designs with maximal structural flexibility and multi-functional integrity. These designs will be further translated and validated with unit operation-based processes. A number of intensified process systems

have been studied with GMF such as dividing wall columns, reactive distillation, solvent selection, etc. (Tian et al., 2022; Tian and Pistikopoulos, 2018)

In what follows, we augment GMF representation and modeling for modular intensified membrane-assisted reaction systems. The novelty of this work is two-fold: (i) Extend the unified Gibbs free energy-based driving force constraints to incorporate the role of membrane selective mass transfer, (ii) Develop a GMF/OCFE modeling strategy for compact, adjustable, and modularized representation of spatially distributed processes.

### 2.2. Representation of Membrane-Assisted Systems

In this work, we focus on gas phase reaction systems modeled with one-dimensional axial distribution. Reaction kinetics and membrane performance data (e.g., permeability) are regarded as given information.

Fig. 2 depicts how pure reaction systems and membrane-assisted systems can be both captured via the generalized GMF mass/heat exchange modules. Two gas-phase inlet streams are considered to enter the module (i.e.,  $f^{SI}, f^{VI}$ ) in correspondence to two outlet streams (i.e.,  $f^{SO}, f^{VO}$ ). The flow directions can be either co-current or counter-current, as to be showcased in the case study. As per the role of membranes:

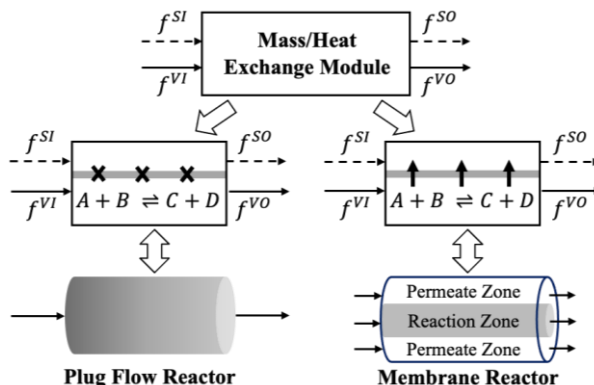


Figure 2: GMF representation for (membrane-assisted) reaction systems.

if transverse flows exist within the module, the mass/heat exchange module represents a membrane reactor with selective mass transfer across the membrane; otherwise, the module represents a plug flow reactor. Auxiliary binary variables (1-0) can be employed to denote the existence (or not) of the membrane. A key research question lies in how to identify when the use of membrane is preferred and vice versa.

### 2.3. Extension of Driving Force Constraints

We address the above question by revisiting the driving force constraints based on total Gibbs free energy change as a function of chemical potentials. Note that mass transfer across membrane is also driven by the difference of chemical potentials between reaction zone and permeate zone. In this regard, we can formulate the driving force constraints for the membrane-assisted systems in a unified form as the systems with conventional mass transfer. More specifically, Eq. 1 defines the driving force constraints which characterize the feasibility of reaction in reaction zone. Eq. 2 are introduced as membrane feasibility constraints to model the process taking place in the mass/heat exchange module. Chemical potential  $\mu_i^V$  serves as the joint node to connect the reaction phenomena and membrane permeation phenomena. If membrane is selected in module ( $y_{mbr} = 1$ ), transverse mass flux is activated in Eq. 2 which will impact driving force constraints; if membrane is not selected ( $y_{mbr} = 0$ ), transverse flux is set to 0 and only reaction occurs in module. It is also worth highlighting that, the selection (or not) of the membrane is dictated by thermodynamic rules of the physical process without pre-specification.

$$d(nG)_{R,P}^{tot} = \sum_i \mu_i^V dn_i^V \leq 0 \quad (1) \quad 0 \leq J_i = \lambda_i A (\mu_i^S - \mu_i^V) \leq J_{max} y_{mbr} \quad (2)$$

where  $\mu_i^V$  ( $\mu_i^S$ ) is chemical potential in reaction (permeate) zone,  $\lambda_i$  is permeability,  $A$  is membrane area,  $J_i$  is transverse mass flux,  $y_{mbr}$  is binary variable denoting membrane selection or not.

#### 2.4. Spatial Modeling via Orthogonal Collocation on Finite Elements

To model the axial distribution characteristics in (membrane-assisted) reaction systems, one approach is to utilize a number of modules in series as shown in Fig. 3a (Monjur et al., 2021; Bishop and Lima, 2021). However, it is a critical yet open question on how to determine the minimum essential number of phenomena building blocks required for representation, which affects both the combinatorial complexity and solvability of the resulting superstructure network and the mathematical model. Given this, we propose to extract spatial distribution information from intra-module. As shown

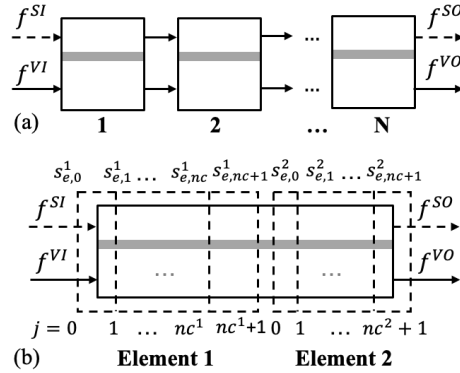


Figure 3: Spatial modeling: (a) Modules-in-series, (b) A single GMF module discretized with OCFE.

in Fig. 3b, each GMF mass/heat exchange module is discretized via Orthogonal Collocation on Finite Elements. The roots of Orthogonal Hahn Polynomials are used to determine the collocation point locations. Component flowrates within the module are approximated using OCFE and major physical modeling constraints are enforced at the collocation points (e.g., mass balances, driving force constraints, reaction kinetics). An indicative list of GMF modeling constraints is provided in Table 1. The following advantages are offered by GMF/OCFE representation: (i) Only continuous variables are introduced to describe the intra-module phenomena, without expanding the combinatorial space using binary variables, (ii) Element partition can be self-adaptive to the underlying faster/slower process changes to enhance modeling accuracy without adding on computational complexity, (iii) The sizing of module can thus be optimized for scaling up or scaling down without compromising modeling accuracy.

Table 1: Mathematical model for GMF/OCFE synthesis – an indicative list

Collocation point locations	$s_{e,j}^l - s_j(M_{e,l}) = 0, \quad e \in E, l \in L_e, j = 1, \dots, nc_{e,l}$
Weighting functions	$W_j(s) - \prod_{z=0, z \neq j}^{nc_{e,l}} \frac{s - s_{e,z}^l}{s_{e,j}^l - s_{e,z}^l} = 0, e \in E, l \in L_e, j = 0, \dots, nc_{e,l}$
Component mass balances	$f c_c^V(s_{e,j}^l) - \sum_{j'=1}^{nc_{e,l}} W_{j'}(s_{e,j}^l + \Delta M) f c_c^V(s_{e,j'}) - J_{e,l,j,c} + \sum_k \nu_c r_{e,l,j,k} \Delta Q_{cat} = 0$ $f c_c^S(s_{e,j}^l) - \sum_{j'=1}^{nc_{e,l}} W_{j'}(s_{e,j}^l + \Delta M) f c_c^S(s_{e,j'}) + J_{e,l,j,c} = 0$
Molar fraction summations	$\sum_{c \in C} f c_c^V(s_{e,j}^l) / f^V(s_{e,j}^l) - 1 = 0, \sum_{c \in C} f c_c^S(s_{e,j}^l) / f^S(s_{e,j}^l) - 1 = 0$
Continuity constraints for adjacent finite elements	$\sum_{c \in C} \left( \sum_{j=0}^{nc_{e,l-1}} W_j(M_{e,l-1}) f c_c^V(s_{e,j}^{l-1}) \right) - f^V(s_{e,0}^l) = 0$ $\sum_{j=0}^{nc_{e,l-1}} W_j(M_{e,l-1}) f c_c^V(s_{e,j}^{l-1}) - f c_c^V(s_{e,0}^l) = 0$
Membrane feasibility constraints: Eq. 4, Driving force constraints: Eq. 5	

where subscripts  $e$  is index for modules,  $l$  is for elements,  $j$  is for collocation points,  $c$  is for components. Superscripts  $V$  and  $S$  respectively refer to reaction and permeate side.  $s$  is variable for collocation point location,  $M$  is element length,  $W$  is Lagrange polynomial weighting function,  $f c$  is component mass flowrate,  $f$  is total mass flowrate.

### 3. Case Study: Hydrogen Production via Steam Reforming

#### 3.1. Problem Statement

In this section, we investigate the design of a lab-scale low pressure methane steam reforming (MSR) membrane reactor adapted from Iulianelli et al. (2010). The reactor is 145 mm in length with 20 mm internal diameter, using hydrogen selective Pd-Ag

membrane. The reaction utilizes Ni-based catalyst, following the kinetic laws in Xu and Froment (1989). The kinetic efficiency parameters with or without permeation are provided in Iulianelli et al. (2010). The feed gas flowrate is 2.390 mol/h, comprising 0.03 mol/mol CH<sub>4</sub>, 0.07 mol/mol H<sub>2</sub>O, and 0.90 mol/mol N<sub>2</sub>. The sweep gas flowrate is 0.132 mol/h with pure N<sub>2</sub>. Isothermal operation is considered at 723.15 K.

### 3.2. GMF/OCFE for Membrane Reactor Simulation

Simulation studies are first performed to test the representation efficacy of GMF/OCFE for the above membrane reactor system. Fig. 4a presents the reaction side H<sub>2</sub> molar fraction profiles in a co-current flow scheme, respectively simulated with: (i) high-fidelity ordinary differential equations (ODE), (ii) a single GMF/OCFE module decomposed with 2 elements and 4 collocation points per element, (iii) 5 GMF mass/heat exchange modules in series (with equivalent module length), and (iv) 10 GMF modules in series. As can be noted, GMF/OCFE provides better modeling accuracy compared to modules in series, while requiring the use of less simulation points both in the continuous space and the combinatorial space (i.e., 8 collocation points in total vs. 10 GMF modules). The ability to assign different element lengths also contributes to the GMF/OCFE representation enhancement. In this case, Element 1 is 0.050 m in length using 4 collocation points to describe the faster process at membrane reactor entrance. Element 2 is 0.140 m in length using another 4 collocation points to capture the rest slower process. GMF/OCFE is also tested on a counter-current flow scheme using an elevated feed gas flowrate to achieve 60% CH<sub>4</sub> conversion and 80% H<sub>2</sub> recovery. The H<sub>2</sub> molar fraction profiles at permeate side are shown in Fig. 4b, in which GMF/OCFE proves sufficient modeling accuracy compared to simulation using 20 modules in series.

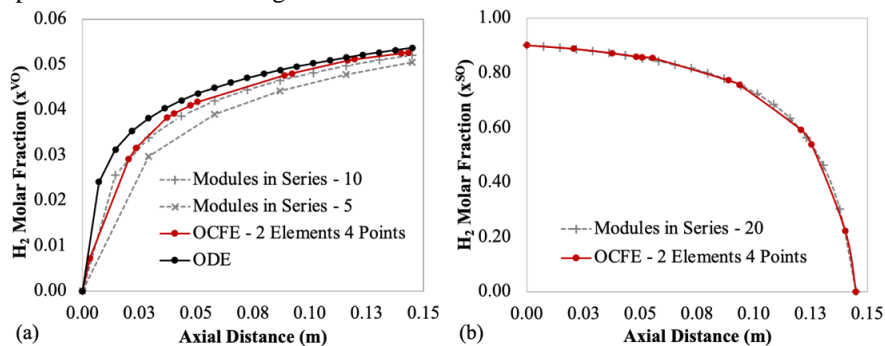


Figure 4: GMF/OCFE simulation for MSR membrane reactor:

(a) Co-current, H<sub>2</sub> profiles at reaction side, (b) Counter-current, H<sub>2</sub> profiles at permeate side.

### 3.3. Design Optimization of Membrane-Assisted Reaction Systems

The design optimization of a membrane-assisted reaction system is considered with at least 75% CH<sub>4</sub> conversion. Two optimization objectives are respectively studied to demonstrate the applicability and versatility of GMF/OCFE synthesis: (i) Maximize H<sub>2</sub> recovery, (ii) Minimize membrane area. A maximum of 4 GMF modules are available for use, each discretized using 2 elements and 4 collocation points per element. As a proof-of-concept example, the structural connections of modules are restricted as in series, without opening up the full superstructure network as shown in Fig. 1 (which encapsulates recycle streams, etc.) The selection or not of transverse mass transfer through membrane is to be determined by the process synthesis problem without pre-postulation. Fig. 5a shows the optimal design with maximal H<sub>2</sub> recovery, which features a membrane reactor

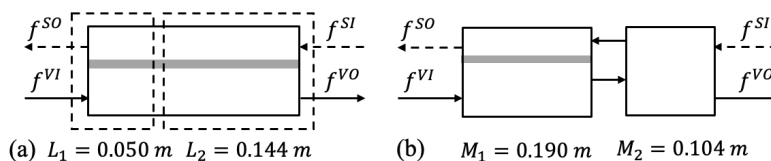


Figure 5: GMF/OCFE optimization: (a) Maximize  $H_2$  recovery, (b) Minimize membrane area.

of 0.195 m in length. Fig. 5b shows the optimal design with minimal membrane area, which features a sequential structure of membrane reactor (0.190 m) followed by a packed bed reactor (0.104 m). By comparing these two design alternatives, the packed bed reactor in Fig. 5b is used to replace the 0.004 m membrane reactor in Fig. 5a to reduce membrane area but attaining the required methane conversion level. For benchmark, a 0.20 meters long packed bed reactor is simulated which reaches only 16%  $CH_4$  conversion.

#### 4. Conclusion

In this paper, we have presented a process synthesis approach for membrane-assisted reaction systems using Generalized Modular Representation Framework. Ongoing work is addressing: (i) full superstructure-based synthesis to generate better and potentially more innovative design solutions, and (ii) the incorporation of life cycle assessment into optimization formulation to simultaneously account for sustainability considerations.

#### 5. Acknowledgement

The authors acknowledge financial support from the Chemical and Biomedical Engineering Department at West Virginia University.

#### References

- D. K. Babi, P. Lutze, J. M. Woodley, R. Gani. (2014). A process synthesis-intensification framework for the development of sustainable membrane-based operations. *Chem. Eng. Process.*, 86, 173-195.
- P. Bernardo, E. Drioli, 2021. Membrane Engineering for a Sustainable Production of Ethylene. *Fuel Process. Technol.*, 212, 106624.
- B. A. Bishop, F. V. Lima. (2021). Novel module-based membrane reactor design approach for improved operability performance. *Membranes*, 11(2), 157.
- A. Iulianelli, G. Manzolini, M. De Falco, S. Campanari, T. Longo, S. Liguori, A. Basile, 2010.  $H_2$  Production by Low Pressure Methane Steam Reforming in a Pd–Ag Membrane Reactor over a Ni-based Catalyst: Experimental and Modeling. *Int. J. Hydrog. Energy*, 35(20), 11514-11524.
- S. Karagöz, T. T. Tsotsis, V. I. Manousiouthakis. (2019). Multi-Scale Modeling and Simulation of a Novel Membrane Reactor/Adsorptive Reactor Process. *Chem. Eng. Process.*, 137, 148-158.
- M. S. Monjur, S. E. Demirel, J. Li, M. M. F. Hasan. (2021). SPICE\_MARS: A Process Synthesis Framework for Membrane-Assisted Reactive Separations. *Ind. Eng. Chem. Res.*, 60, 7635-7655.
- E. N. Pistikopoulos, Y. Tian. (2022). *Synthesis and Operability Strategies for Computer-Aided Modular Process Intensification*. Elsevier.
- M. Sjardin, K. J. Damen, A. P. Faaij. (2006) Techno-Economic Prospects of Small-Scale Membrane Reactors in a Future Hydrogen-Fuelled Transportation Sector. *Energy*, 31(14), 2523-2555.
- A. Stankiewicz, T. Van Gerven, G. Stefanidis. (2019). *The Fundamentals of Process Intensification*.
- Y. Tian, V. Meduri, R. Bindlish, E. N. Pistikopoulos. (2022). A Process Intensification Synthesis Framework for Design of Dividing Wall Column Systems. *Comput. Chem. Eng.*, 160, 107679.
- Y. Tian, E. N. Pistikopoulos. (2018). *Synthesis of Operable Process Intensification Systems – Steady-State Design with Safety and Operability Considerations*. *Ind. Eng. Chem. Res.*, 58(15), 6049-6068.
- A. K. Tula, M. R. Eden, R. Gani. (2019). Computer-Aided Process Intensification: Challenges, Trends and Opportunities. *AIChE J.*, 66(1).
- J. Xu, G. F. Froment. (1989). Methane Steam Reforming, Methanation and Water-Gas Shift: I. Intrinsic Kinetics. *AIChE J.*, 35(1), 88-96.

# Multiscale design and analysis of templated zeolite for Li-O<sub>2</sub> battery with improved discharge capacity

Khizar Hayat<sup>a,b</sup>, Daniel Bahamon<sup>a,b</sup>, Lourdes F. Vega<sup>a,b</sup>, Ahmed AlHajaj<sup>a,b</sup>

<sup>a</sup>Research and Innovation Center on CO<sub>2</sub> and Hydrogen (RICH Center), Khalifa University, P.O. Box 127788, United Arab Emirates.

<sup>b</sup>Department of Chemical Engineering, Khalifa University, P.O. Box 127788, Abu Dhabi, United Arab Emirates

## Abstract

The commercialization of practical Li-O<sub>2</sub> batteries is restricted by the lower discharge capacity induced by poor mass transport through the cathode structure. In view of that, we have designed four (RHO-, FAU-, MFI-, and BEA-ZTCs) hierarchical zeolite-templated carbon electrodes. We have utilized a multiscale modelling framework to assess the performance of developed structures by integrating cluster and cell-level structural domains. At the cluster level, a reactive forcefield molecular dynamic study is carried out to explore the species (Li<sup>+</sup>, O<sub>2</sub>, and DMSO) transport through porous structures. The results revealed that hierarchical structures, particularly RHO-ZTC, have superior mass transport. The subsequent cell level 2-D continuum modelling study resulted in the improved discharge capacities of hierarchical structures. Specifically, hierarchical RHO-ZTC showed nearly three times higher capacity than conventional SP carbon. This improved mass transport is attributed to the mesopores in the systems acting as oxygen transport tunnels and product storage reservoirs.

**Keywords:** ReaxFF-MD, Li-O<sub>2</sub> battery, Multiscale modelling, 2-D continuum model

## 1. Introduction

Among the myriad electrochemical storage technologies, the Li-O<sub>2</sub> battery is being considered as a promising power source for next-generation electric vehicles. However, its viable commercialization is still the dream due to lower discharge capacity which depends on the mass transport through porous electrode structure (Hayat *et al.*, 2022). Therefore, improving the discharge capacity of Li-O<sub>2</sub> battery (LOB) have become critically acclaimed topic among the scientific community. In that perspective, the designing of novel hierarchical LOB electrode with optimized structural characteristics (porosity, pore size distribution, surface area, and tortuosity) could be crucial in facilitated mass transport which ultimately may leads to high discharge capacity (Hayat *et al.*, 2021). The mass transfer through porous structures can be, essentially, enhanced through various routes (Schneider *et al.*, 2016) such as crystal size reduction, incorporating large number of micropores, and introducing mesopores along with micropores as additional storage reservoir and mass transport carriageways. From the Li-O<sub>2</sub> storage technology scenario, the third route could be a promising way of developing novel hierarchical structures.

In that sense, our group put significant efforts in developing hierarchical amorphous carbon cathodes (with combined micro- and mesopores) (Elabyouki *et al.*, 2019 and Hayat *et al.*, 2022). Despite promising improvement in the storage capacity of Li-O<sub>2</sub> battery, the molecular understanding of coupled mass transport and electrochemical reactions inside micro- and mesopores still needs better understanding for prolonged discharge operation and enhanced discharge capacity. Furthermore, the impact of



structural properties of hierarchical carbon frameworks on the mass transport (at molecular level) and discharge capacity (at cell level) require comprehensive assessment for designing future generation Li-O<sub>2</sub> batteries.

Therefore, in this contribution, we draw attention to the multiscale design of hierarchical zeolite templated (ZTC) carbon cathode structures. Our integrated multiscale study is comprised of two levels including molecular level, and cell level. The molecular level includes reactive forcefield molecular dynamic study (reaxFF-MD) of the screened hierarchical ZTCs. From reaxff-MD, we are intended to get the key performance indicators such as species diffusivities (Li<sup>+</sup>, O<sub>2</sub>, and electrolyte), main discharge products, pore volume, pore size distribution, carbon mass loadings, and surface area. The cell level computational work aims at performing process simulations of the main hierarchical ZTC structures considered in this study (using data from molecular simulations) and then evaluate their performance based of discharge capacity and cell voltage.

## 2. Methodology

To obtain physical insight into the transport properties within the hierarchical ZTCs cathodes and their impact on cell discharge capacities, we have constructed an integrated multiscale modelling (MSM) framework, as illustrated in Figure 1. The developed MSM is spanned over cluster and cell scales. The former scale focus on the molecular investigations of transport properties using reactive molecular dynamic simulation. The main key performance indicators that are obtained from reaxFF-MD study include effective-self diffusivities, molecular density number profile, and discharge product clusters formation. While cell scale model (i.e., 2-D continuum model) is developed to screen the hierarchical ZTC electrodes using their corresponding transport characteristics from cluster scale simulation. The cell scale modelling involves the investigation of surface area evolution, electrochemical kinetics, porosity evolution, and discharge capacity and specific energy density calculation during discharging of Li-O<sub>2</sub> battery cell. The detail of both modelling studies is explained in following sections.

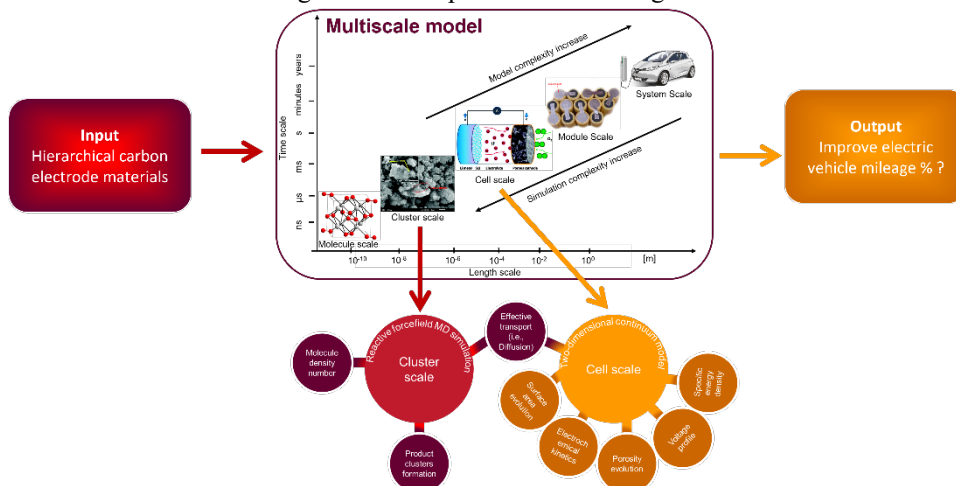


Figure 1. Multiscale modelling framework for Li-O<sub>2</sub> batteries

### 2.1. Reactive forcefield molecular dynamic simulation

Molecular dynamic simulation strategy involves the designing and characterizing hierarchical structures and performing reactive forcefield molecular dynamic simulations.

We have employed zeolite-templated carbons (ZTCs) as parent structures for the development of hierarchical electrodes for Li-O<sub>2</sub> battery (Braun *et al.*, 2018). Based on structural analysis of 47 parent ZTCs (using Poreblazer package), we have selected four microporous zeolite templated carbons (RHO-, FAU-, MFI-, and BEA-) which are further utilized for the generation of hierarchical ZTCs having both micro- and mesopores. Moreover, to create hierarchical frameworks from the parent structures, a slit type mesopore with typical size of 25 Å is introduced into the centre of the intrinsic supercells (Elabyouki *et al.*, 2019). The four generated hierarchical ZTCs composed of slit mesopore sandwiched between microporous walls with different thicknesses such as 17.3, 18.0, 14.0, and 12.8 Å for RHO-, MFI-, FAU-, and BEA-ZTCs, respectively.

For simulation box preparation, firstly we packed the hierarchical ZTC electrodes by randomly placing dimethyl sulfoxide (DMSO i.e., liquid non-aqueous electrolyte solvent) molecules. The number of DMSO molecules in each structure are adjusted to mimic the bulk electrolyte density (i.e., 1.09 g/cm<sup>3</sup>) (Lebel *et al.*, 1962). Next, to prepare 1 M electrolyte solution of LiPF<sub>6</sub>, we have added Li<sup>+</sup> and PF<sub>6</sub><sup>-</sup> salt-ions to the simulation cell. Further, different number of gaseous O<sub>2</sub> (100, 200, 300, and 500) molecules were included into the electrolyte solution to obtain various temporal configurations of hierarchical zeolite-templated carbon electrodes. Here, the main objective of adding different number of oxygen molecules is to demonstrate the discharge mechanism stages Li-O<sub>2</sub> cell, with 100 oxygen molecules being initial stage and 500 oxygen molecules being final step of the discharge process. For molecular interactions calculation, we have utilized reactive forcefield containing the parameter set for the C, H, O, Li, P, F, and S (Islam *et al.*, 2014) elements. Besides, to accurately predict the formation of Li<sub>2</sub>O<sub>2</sub>, the forcefield parameter values for Li and O were replaced with the one developed for Li-O<sub>2</sub> systems (Hearn *et al.*, 2020). Unlike the empirical classical forcefield, the reactive forcefield (reaxFF) is a bond order (BO) based method which permits the bond breakage and formation during the molecular dynamic computations (Raju *et al.*, 2014). In reaxFF, the total system energy is the sum of individual component energies such as bond, lone pair, torsional, valence, over-coordination, under-coordination, van der Waals, and coulombic energies. During reaxFF-MD simulations, bond order values are calculated at every time step based on interatomic bond distances, consequently acting as a component of prime importance for all types of bonded interactions such as torsional and valence interactions. When the bond breakage happens, all the bonded energy contributions are diminish leaving behind only long-range non-bonded interactions such as van der Waals and coulombic interactions, normally present among the non-bonded pairs of atoms. The reaxff-MD simulations are performed in LAMMPS package involving following steps; simulation cell optimization, equilibration at 10 K (for 10 ps) followed by heating from 10 – 300 K (for 10 ps), equilibration at 300 K (for 10 ps), and production run at target temperature i.e., 300 K (for 1.5 ns). A value of 0.1 fs is used for the time step and Nose Hoover thermostat is taken for simulation temperature control. For post-processing and visualization of data, we have employed OVITO software. For species diffusion calculation, the Einstein's correlation (Islam *et al.*, 2015) is utilized which computed self-diffusivity values based on the mean square displacement (MSD) of individual molecules as follows.

$$D_x = \frac{1}{2} \lim_{\Delta t \rightarrow \infty} \frac{\text{MSD}_x(t+\Delta t) - \text{MSD}_x(t)}{\Delta t} \quad (D_x, \text{Diffusivity along preferential direction}) \quad (1)$$

Here, the self-diffusivity values of Li<sup>+</sup> and O<sub>2</sub> are based on the total number of respective molecules present in any form such as O<sub>2</sub><sup>2-</sup>, LiO, Li<sub>2</sub>O, LiO<sub>2</sub>, and Li<sub>2</sub>O<sub>2</sub>.

## 2.2. Cell scale 2-D continuum model

To test the cell performance of hierarchical ZTC electrodes, we have developed the two-dimensional continuum model. The detail of the model equations, initial conditions, boundary conditions, and assumptions is given in supplementary material of our previously published work (Hayat *et al.*, 2021). The reactive forcefield study is integrated with the 2-D model through modifying the effective diffusivity equations which will mimic the molecular transport through hierarchical ZTC cathodes. The curve fitting analysis is utilized to fit the effective diffusivity correlations data with that obtained from reaxFF-MD study. The underlying algorithm for curve fitting involves the generalized reduced gradient method which finds the optimal solution based on the least square error. The mentioned method depends on the initial condition; therefore, we have used the curve fitting method with multi-start in which it automatically finds the most suitable initial condition. After developing 2-D model by including modified effective diffusivity correlations, we have imported the model equations into COMSOL Multiphysics software to perform equation-based simulation of Li-O<sub>2</sub> battery cell. During the discharging process, at the airside, oxygen dissolves into the non-aqueous liquid electrolyte (flowing through cathode pores network) and reacts with Li<sup>+</sup> subsequently generating the solid Li<sub>2</sub>O<sub>2</sub> discharge product. The stop condition is applied when the solid cathode potential on the air-side boundary becomes equal to 2.1 V. For all the variables, the convergence criteria are set to 10<sup>-4</sup>.

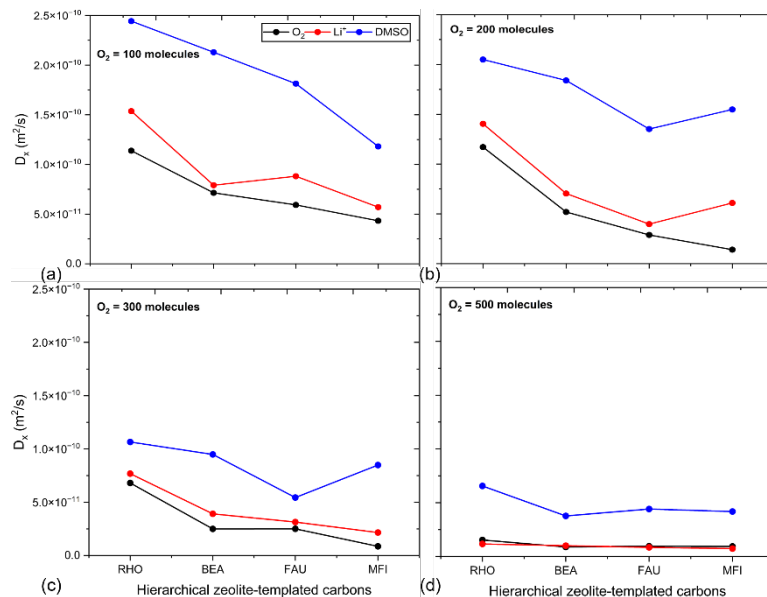


Figure 2. Self-diffusivity (x-direction) trends for oxygen, Lithium-ion, and DMSO electrolyte through hierarchical zeolite-templated carbons at various discharge states (a) 100, (b) 200, (c) 300, and (d) 500.

### 3. Results and discussions

Figure 2 (a-d) is plotting the self-diffusivities curves of oxygen, Li<sup>+</sup>, and DMSO along the preferential direction through hierarchical zeolite-templated carbon electrodes at various oxygen molecules (100, 200, 300, and 500). The results demonstrate that the self-diffusivities of all the species are decreasing as we move from RHO-ZTC to MFI-ZTC at all oxygen consumption cases. This is because of the reducing micropore size from RHO to MFI. The superior species diffusion through hierarchical RHO-ZTC porous framework

(among all the ZTCs) could be induced by the large pore volume ( $\sim 1.85 \text{ cm}^3\text{g}^{-1}$ ) due to the presence of high density of micro- ( $\sim 14.5 \text{ \AA}$ ), and meso-pores ( $\sim 25 \text{ \AA}$ ). Consequently, RHO-ZTC not only accommodate large fraction of solid discharge product but also it provides enough space for molecules (particularly O<sub>2</sub>) to transport without pore clogging. Another reason for the facilitated diffusion might be the fast molecular exchange between slit-mesopore and microporous walls of the hierarchical RHO-ZTC. This finding corresponds to the claim reported by one of our group member (Elabyouki *et. al.*, 2019) who stated that the enhanced discharge characteristics of hierarchical electrodes are mainly due to the continuous molecular exchange occurring at meso-micro pores interphase thereby preventing the micropores clogging by solid product. Further, for RHO-ZTC, the attained diffusivity ranges for Li<sup>+</sup> and O<sub>2</sub> are ( $1.54 \times 10^{-10} - 2.69 \times 10^{-11} \text{ m}^2\text{s}^{-1}$ ) and ( $1.14 \times 10^{-10} - 1.53 \times 10^{-11} \text{ m}^2\text{s}^{-1}$ ), respectively. Further, it is evident, that as oxygen consumption increases (from 100 to 500 molecules) the species self-diffusivity coefficients decrease. This is because of the generation of solid discharge product (Li<sub>2</sub>O<sub>2</sub>) which not only reduces the free Li<sup>+</sup> and reactive O<sub>2</sub> but also imposes extra resistance towards their transport. At the lower oxygen consumption (i.e., 100 oxygen molecules), small size discharge intermediates (e.g., LiO, Li<sub>2</sub>O, LiO<sub>2</sub>) are the dominant products, however, higher oxygen consumption (i.e., 500 O<sub>2</sub> molecules) permits the formation of large size discharge product aggregates (specifically Li<sub>2</sub>O<sub>2</sub>) which are strongly adsorbed to the electrode pore surface and fill the entire pore volume resulting in lower diffusivities of Li<sup>+</sup>, oxygen, and DMSO (electrolyte solvent).

Figure 3 illustrates the comparison of discharge profiles for hierarchical zeolite-templated carbon electrodes and experimental data from Read's work (Read *et. al.*, 2002) at current density value of 0.1 mA/cm<sup>2</sup>. The discharge capacity trends obtained from 2-D modelling of hierarchical ZTCs are as follows; RHO (2848 mAh/g) > FAU (960 mAh/g) > MFI (740 mAh/g) > BEA (690 mAh/g). These discharge capacities correspond to the self-diffusivity trends obtained from reaxFF-MD simulation for individual hierarchical ZTCs. It is noticed that specifically hierarchical RHO-ZTC have delivered 3 times much higher discharge capacity than that of literature data generated from conventional SP carbon electrode ( $\sim 941 \text{ mAh/g}$ ). For BEA-ZTC, unlike superior diffusivities, the capacity is lower which is mainly due to the lower porosity, surface area, and cell density of the

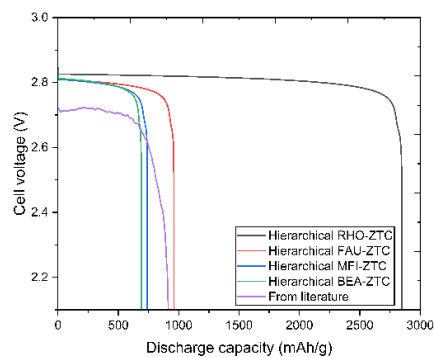


Figure 3. Comparative analysis of discharge profiles of hierarchical-ZTCs and liter. (Read *et. al.*, 2002).

structure. Further, hierarchical electrodes also improved the average discharge voltage which may lead to improved specific energy density of LOB. This finding explains the fact that facilitated mass transport (besides electrochemical kinetics) through hierarchical

structures not only improve the storage capacity but also helps to mitigate the effect of voltage loss during galvanostatic discharge resulting in improved energy density. Another reason of improved storage capacities is attributed to the more utilization of pore volume available in the hierarchical structure (particularly RHO-ZTC). Overall hierarchical ZTC based air-electrodes could be promising for the development of future generation Li-O<sub>2</sub> power system.

#### 4. Conclusion

Dealing with the challenge of lower capacity due to limited mass transport, here we have proposed novel hierarchical zeolite-templated carbon as potential air-electrodes for high-capacity Li-O<sub>2</sub> batteries. An integrated multiscale model is utilized to screen the developed electrodes based on their cell performance, particularly discharge capacity. The reaxFF-MD study revealed that hierarchical structures have facilitated the mass transport characteristics. Specifically, RHO-ZTC showed dominant transport of Li<sup>+</sup> and O<sub>2</sub> compared to rest of the structures. Subsequently, hierarchal RHO-ZTC generated much higher discharge capacity (~ 2848 mAh/g) than the rest of the electrode structures. The great improvement in capacity is mainly attributed to the inclusion of mesopores which not only facilitated species diffusion but also improved the accessible free volume to store the solid discharge product i.e., Li<sub>2</sub>O<sub>2</sub>. This multiscale study could guide us to design state-of-the-art Li-O<sub>2</sub> batteries with ultrahigh capacities.

#### References

- K. Hayat, L. F. Vega, A. AlHajaj, 2022, What have we learned by multiscale models on improving the cathode storage capacity of Li-air batteries? Recent advances and remaining challenges, *Renew Sustain Energy Rev*, 154, 111849.
- K. Hayat, L. F. Vega, A. AlHajaj, 2021, Modeling of hierarchical cathodes for Li-air batteries with improved discharge capacity, *J Electrochem Soc.* 168, 120534.
- D. Schneider, D. Mehlhorn, P. Zeigermann, J. Kärger, R. Valiullin, 2016, Transport properties of hierarchical micro-mesoporous materials, *Chem Soc Rev.* 45, 3439–67.
- M. Elabyouki, D. Bahamon, M. Khaleel, L. F. Vega, 2019, Insights into the transport properties of electrolyte solutions in a hierarchical carbon electrode by molecular dynamics simulations. *J Phys Chem C.* 123, 27273–85.
- E. Braun, Y. Lee, S. M. Moosavi, S. Barthel, R. Mercado, I. A. Baburin, 2018, Generating carbon schwarzites via zeolite-templating, *Mater Cloud Arch.* 115, E8116–24.
- R. G. LeBel, D. A. I. Goring, 1962, Density, viscosity, refractive index, and hygroscopicity of mixtures of water and dimethyl sulfoxide, *J Chem Eng Data*, 7, 100–1.
- M. M. Islam, V. S. Bryantsev, A. C. T. Van Duin, 2014, ReaxFF reactive force field simulations on the influence of teflon on electrolyte decomposition during Li/SWCNT anode discharge in lithium-sulfur batteries, *J Electrochem Soc.* 161, E3009.
- K. A. O’Hearn, M. W. Swift, J. Liu, I. Magoulas, P. Piecuch, A. C. T. Van Duin, 2020, Optimization of the Reax force field for the lithium–oxygen system using a high fidelity charge model, *J Chem Phys.* 153, 84107.
- M. Raju, P. Ganesh, P. R. C. Kent, A. C. T. Van Duin, 2015, Reactive force field study of Li/C systems for electrical energy storage, *J Chem Theory Comput.* 11, 2156–66.
- M. M. Islam, A. Ostadhossein, O. Borodin, A.T. Yeates, W.W. Tipton, R.G. Hennig, N. Kumar, A.C.T. Van Duin, 2015, ReaxFF molecular dynamics simulations on lithiated sulfur cathode materials, *Phys. Chem. Chem. Phys.* 17, 3383–3393.
- J. Read, 2002, Characterization of the lithium/oxygen organic electrolyte battery, *Journal of the Electrochemical Society*, 149, A1190.

# Numerical Evaluation of Fuel Utilisation Enhancement Strategies in a Solid Oxide Fuel Cell System Fuelled by Hydrocarbons Mixtures

Clyde-Theodore Nguimbous Batista,<sup>a</sup> Amirpiran Amiri,<sup>a</sup> Robert Steinberger-Wilckens,<sup>b</sup>

<sup>a</sup>*European Bioenergy Research Institute (EBRI), School of Engineering and Applied Science, Aston University, Birmingham B4 7ET, UK*

<sup>b</sup>*School of Chemical Engineering, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK*

## Abstract

In this paper an internal reforming SOFC is simulated, including the Balance-of-Plant (BoP) it is embedded in. Various approaches are evaluated to provide data for the trade-off between system efficiency and fuel cell health. The results reveal that controlled single-pass fuel utilisation (~50%) with a maximised system fuel utilisation of 85% is feasible. The operating voltage, current trajectories and detailed patterns of voltage losses and efficiencies are captured. The anode activation loss is observed to be the key contributor to the operating voltage variation when anode-off gas recycling is applied to improve the system's  $U_f$ . The BoP modifications are presented to identify the optimal recycling point in the process flowsheet. The simulation results showed that in a comprehensive BoP, optimisation of system fuel utilisation and efficiency through tuning of the anode-off gas recycle needs exploring of both flowsheet and operating parameters enhancement.

**Keywords:** SOFC, fuel utilisation, efficiency, health.

## 1. Introduction

SOFCs are promising devices towards highly efficient electrical power generation from chemical energy. The technology can be used in a wide range of applications for both domestic and industrial sectors, from distributed power/CHP generation to light traction vehicles (O. Sharaf and M. Orhan, 2014). The high operating temperatures make them ideal for combined heat and power systems, utilising high-grade waste heat, wherein, home applications can replace boilers for heating. The SOFCs boast high fuel flexibility, low emission and relatively low operating costs. However, the main drawbacks of this technology are the high costs of manufacture and the rapid deterioration of the fuel cell performance. The performance losses can be associated with the complex degradation process of the SOFC, impacted by both intrinsic and extrinsic factors. Despite the abundant research into the explicit impact of factors such as the presence of impurities, and mechanical deterioration on performance and losses during operation (from an electrochemical standpoint), long-term durability remains a major concern. This, in turn, prevents the development of accurate lifetime predictive models, and so current research focalises on the maximising of the SOFC health, i.e. the performance of the fuel cell over

time, until failure. In this consideration, fuel utilisation ( $U_f$ ), describing the quantity of fuel utilised towards the production of power, is an important aspect to consider particularly at the system level. In this work, two definitions of fuel utilisation will be employed: i. single-pass fuel utilisation (SP  $U_f$ ) that refers to the stack and ii. Overall system fuel utilisation (System  $U_f$ ) refers to the system or the BoP. Fuel utilisation maximisation can have both a positive impact (producing more energy with lower fresh fuel input requirements and increased efficiency), and a negative impact correlated to voltage losses and localised degradation by various causes. This impedance on SOFC health, with respect to  $U_f$ , can typically be associated with high SP  $U_f$ , as such, strategies to maximise the system  $U_f$  whilst maintaining moderate SP  $U_f$  are important (J. Yi and T. Kim, 2017). Employable strategies include multistaging (MS) and the introduction of anode off-gas recycling (AOGR). In this work, the various methods of implementing the AOGR will be investigated within a single internal reforming SOFC system and the impact on SOFC health explored, with the aim to provide insight into the optimum strategy for improving long-term fuel cell performance. The aim is to investigate the impact of flowsheeting on AOGR effectiveness.

## 2. Simulation

The SOFC and BoP are simulated on Aspen HYSYS® utilising the built-in modules and functions. Three flowsheets are used for the investigation of the effectiveness of the implementation of AOGR on the SOFC performance and health. The base case system depicted in Figure 1, captures the pre-AOGR implementation performance data. Two modifications were implemented in Platform A to create Platform B and C as shown in Figure 2 and Figure 3, respectively.

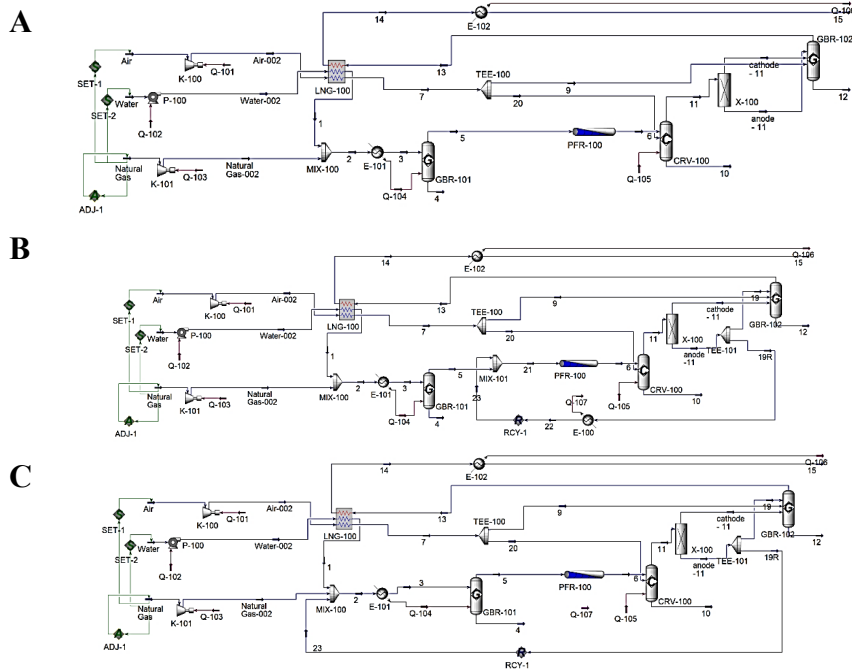


Figure 1: Platform A - base case model of the BoP with no AOGR; Platform B - AOGR connected to the stack inlet point; Platform C - AOGR connected to the external pre-reformer inlet point.

The SOFC electrochemical calculations are carried out within a series of inbuilt SPREADSHEET using the governing equations used by W. Zhang et al., 2005.

### 2.1. Simulation Validation

The model validation is presented in Table 1. To this end, the results of Platform C were compared against the results of similar systems existing in literature where the models' inputs are identical.

Table 1: Simulation results compared to the literature data for power =120kW DC.

Variable/Parameter	W. Zhang et al., 2005	W. Doherty et al., 2010	T. R. Tanim, 2012	This work
Voltage, (V)	0.70	0.68	0.69	0.69
Current density, (mA/cm <sup>2</sup> )	178.0	182.9	182.0	180.7
Pre-reforming temperature, (°C)	536.0	535.1	535.0	535.1
Cathode inlet temperature, (°C)	821.3	823.7	826.0	826.0
Stack exhaust temperature, (°C)	834.0	833.7	836.0	793.9
Anode outlet composition				
H <sub>2</sub> O	50.9%	50.9%	50.9%	50.9%
CO <sub>2</sub>	24.9%	24.9%	24.9%	24.9%
H <sub>2</sub>	11.6%	11.6%	11.6%	11.6%
CO	7.4%	7.4%	7.4%	7.4%
N <sub>2</sub>	5.1%	5.1%	5.1%	5.1%
Stack exhaust composition				
N <sub>2</sub>	77.3%	77.3%	77.2%	77.3%
O <sub>2</sub>	15.9%	15.9%	15.7%	15.9%
H <sub>2</sub> O	4.5%	4.5%	4.7%	4.5%
CO <sub>2</sub>	2.3%	2.3%	2.4%	2.3%
Gross electrical efficiency (LHV), (%)	52.0%	51.3%	51.7%	52.0%

As shown in Table 1, under the same operating conditions, the simulation is capable of simulating an internal reforming SOFC operating at 85% fuel utilisation and 120kW output power. The data shows the results obtained to be within 5% of that from the literature. Key assumptions made in the work are the inlet fuel composition as CH<sub>4</sub> 81.3%, C<sub>2</sub>H<sub>6</sub> 2.9%, C<sub>3</sub>H<sub>8</sub> 0.4%, C<sub>4</sub>H<sub>10</sub> 0.2%, N<sub>2</sub> 14.3% and CO 0.9%, an inlet air temperature of 630°C and fuel feed of 200 °C. A total active area of 96.1m<sup>2</sup> over 1152 cells was assumed (W. Zhang et al., 2005).

### 3. Results and discussions

The base case system shown in Platform A is used as a point of reference for the system performance. Figure 2a presents the impact of recycle ratio (RR) on the system U<sub>f</sub> and operating voltage captured in Platform B and C for 50kW power output. In Figure 2b variations of voltage and system efficiency, against RR are presented.



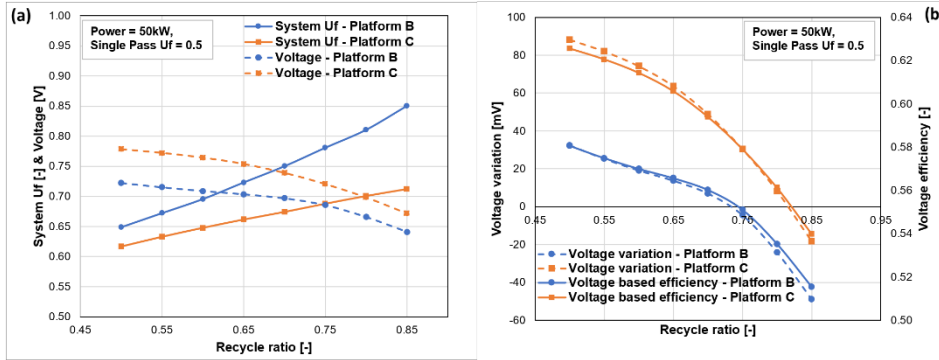


Figure 2: Platform B and C performance comparison at constant power and fueling rate and varying RR; (a) impact on system Uf and SOFC operating voltage, (b) impact on voltage variation  $\Delta V = V - V_{ref}$  and voltage efficiency.

From Figure 2a, at RR = 0.85 Platform B and C can reach a system Uf of 0.85 and 0.71 respectively, at a constant fuel rate that is obtained from the base case model of Platform A for 50kW DC power output and SP Uf = 0.5. Despite this, Platform B displays more cause for concern in terms of the SOFC voltage level. As shown in Figure 2a, the SOFC voltage in Platform B is much lower than that of Platform C, ranging from 0.72 to 0.64 compared to 0.78 to 0.67, with a larger difference at lower RRs. The voltage variation against RR (shown in Fig 2b) does not differ significantly comparing these two platforms. The voltage drop can be a health risk for the SOFC by reaching a thermodynamic voltage below which the anode, in this case Ni, will change its oxidation state i.e., conversion from Ni to NiO. This voltage threshold is a function of temperature. So, if higher voltage loss results in dropping the operating voltage below the thermodynamic Nernst voltage limit at that temperature, then anode re-oxidation gets momentum. Subsequently, depending on the voltage drop extent and time, there can be a significant change in the anode microstructure. So even if the voltage drop is temporary and the voltage recovers, the damage can be significant and lead to an increased rate of degradation. To safeguard against this, there is a need to have a control measure for RR to not allow the operating voltage to drop below its temperature-dependent thermodynamic voltage for re-oxidation. Further, the impact of optimising the fueling of the system on the performance is assessed for both Platform B and C, at a constant SP Uf of 0.5. This allows for the determination of the optimum method for the maximisation of system fuel utilisation whilst averting system performance degradation. With this operation mode, both Platforms B and C are seen to be capable of attaining a system Uf of 0.87 at a recycle ratio of 0.85. However, the impact of this on the system varies by platform and with increasing power demand. Figure 3 below shows the fuel input requirements and the associated efficiency for both platforms B and C at a recycle ratio of 0.5, achieving a system Uf of 0.67 each.

As can be seen in Fig 3, at all power values the flowsheet variations presented in Platform B and C show increases in the electrical efficiency of the system when compared to that of the base case (Platform A). For DC power output of 50kW, the base case simulation provided an efficiency of 44% with a SP Uf of 0.5 and consequently a system Uf of 0.5. However, Platforms B and C results showed the efficiency to increase to 47% and 50%, respectively, both showing an increased system Uf of 0.67 achieved through 50% recycling of the anode off-gas. At all DC power values investigated and 50% AOG, R,

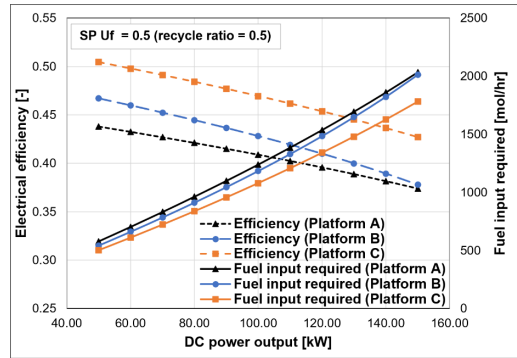


Figure 3: Systems efficiency analysis at different powers, constant SP Uf and RR.

Platform C achieved the greatest overall efficiency, with 47% and 43% efficiencies observed at 100kW and 150kW respectively. This indicated an average of 6% efficiency increase from the base case model. Conversely Platform B achieved 3-5% lower efficiencies when compared to Platform C, while an increase of 1-2% from the base case was observed. As can be seen in Fig 4b, for all platforms the maximum achievable efficiencies occur at higher RR when the drawn power reduces.

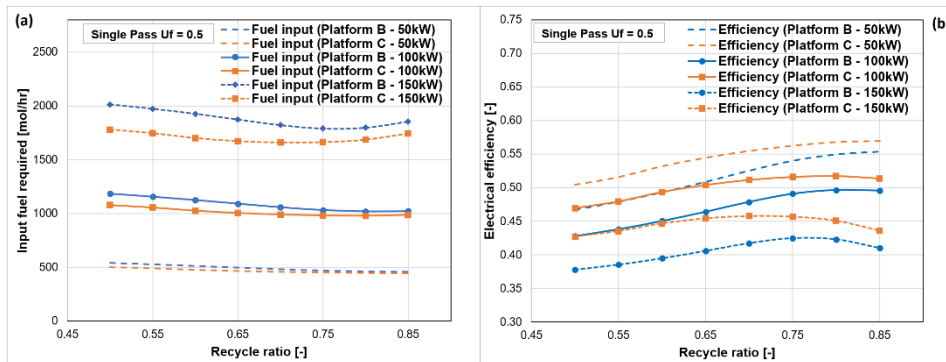


Figure 4: Platforms B and C, efficiency, and fuel input requirement versus RR at different powers.

The incremental increase in DC power output, for all cases, showed to increase in the fuel input required, as shown in Figure 3. As shown in Figure 4a, with increasing RR both Platforms B and C display consistently decreasing fuel input requirements at 50kW, 100kW and 150kW reaches a plateau at a RR about 0.75. for the 150kW case models reaching an optimum recycling ratio at 0.75 followed by an increase in fuel requirement that is because of promoted voltage loss at high RR values. This is consistent with the observations depicted in Figure 4b for efficiency decrease after the optimum RR point.

This is due to greater AOGR ratios reducing the partial pressures of hydrogen and hydrogen carriers within the fuel cell stack inlet, leading to the observed reduction in the electrical efficiency (above 0.75 AOGR). The increasing recycle ratio shows both Platforms B and C to observe changes in efficiency proportional to the changes in the input fuel requirements. For both platforms, operating at 50kW and 100kW shows efficiency to increase to a plateau and 150kW to reach a peak efficiency of 42% and 46% for Platforms B and C respectively, at a recycling ratio of 0.7.

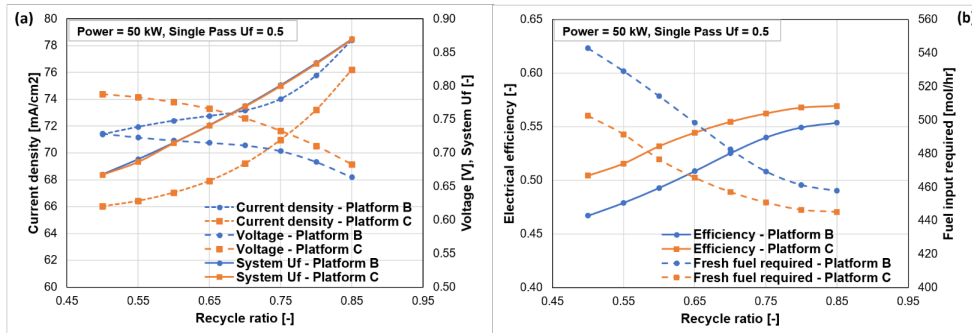


Figure 5: Comparison at 50kW DC power output.

From the results, it can be deduced that AOGR in Platform C is a more desirable method at all simulated system Ufs due to the greater overall increase in efficiency and reduction in fuel input required (Figure 4b and 5b). At 50kW DC power output, the efficiency is shown to increase from 50% to 57%, with Platform B increasing from 47% to 55% for a recycle ratio of 0.5 to 0.85. The maximum system Uf and efficiency reached by both are 0.87 with 55% and 57% efficiency respectively. Furthermore, in the interest of SOFC health, Platform B shows to provide the system with a more stable voltage (decreasing from 0.73V to 0.66V between recycle ratio of 0.5 and 0.85), with a steady increase in current density until a recycle ratio of 0.7, followed by the steep increase (Figure 5a). On the other hand, Platform C shows a more consistent change in both voltage and current density, with the minimum voltage being 0.68V.

#### 4. Conclusions

In literature, moderated SP Uf and high system Uf have been recommended for SOFC system operation for sake of SOFC health and overall efficiency, respectively. This result of this work revealed that the system Uf needs to be optimised along with BoP operating conditions due to the concerns associated with high system Uf that should be taken into account in any process optimisation. In this paper, the system level Uf was evaluated for two BoPs with different AOGR points. The results have shown that efficiencies, voltage losses and system Uf can significantly be different influencing the SOFC lifetime and system overall efficiency.

#### References

- O. Sharaf and M. Orhan, 2014, An overview of fuel cell technology: Fundamentals and application, *Renewable and Sustainable Energy Reviews*, 32, 810-853
- J. Yi and T. Kim, 2017, Effects of fuel utilisation on the performance of SOFC/gas turbine combined power generation systems, *Journal of Mechanical Science and Technology*, 31, 6, 3091-3100
- W. Zhang et al., 2005, Simulation of a tubular solid oxide fuel cell stack using AspenPlus™ unit operation models, *Energy Conversion and Management*, 46, 2, 181-196
- W. Doherty et al., 2010, Computer simulation of a biomass gasification-solid oxide fuel cell power system using Aspen Plus, *Energy*, 35, 12, 4545-4555
- T. R. Tanim, 2012, Modelling of a 5kWe Solid Oxide Fuel Cell Based Auxiliary Power Unit Operating on JP-8 Fuel

# Heat pumping and renewable energy integration for decarbonizing brewery industry and district heating

Daniel Flórez-Orrego<sup>a,b</sup>, Meire Ribeiro Domingos<sup>a</sup>, François Maréchal<sup>a</sup>

<sup>a</sup>*IPÉSE, Federal Polytechnic School of Lausanne, Switzerland*, <sup>b</sup>*School of Mines, National University of Colombia, Colombia*, [daniel.florezorrego@epfl.ch](mailto:daniel.florezorrego@epfl.ch).

## Abstract

The energy integration of brewery plants and urban agglomerations has many benefits in terms of rational energy use and reduced environmental impact, especially during scenarios of uncertain supply chain and volatile market prices. In fact, the diversification of the energy inputs along with the waste heat upgrading through the implementation of optimal energy integration systems become crucial for defining the new energy pathways towards a sustainable and robust energy security. The enhanced waste heat recovery and the residues upgrading also increase the revenues of the industrial energy systems, as they can be used to supply the energy demands of urban agglomerations. Yet, the combination of a large number of energy technologies and the time-varying energy demands calls for the application of a systematic approach capable of identifying the operating conditions and arrangements that minimize the energy resources consumption without impairing the financial feasibility. Thus, in this work, a mixed integer linear problem is solved using OSMOSE platform in order to determine the most profitable arrangement that meets supply and demand profiles of the industrial and urban systems when different energy conversion technologies are considered. More precisely, this analysis sheds light on the relevance of relying on a combination of renewable (e.g. solar, anaerobic digestion, heat pumps) and nonrenewable (e.g. natural gas-based) solutions to supply the time-varying electricity and heating demands of a beer factory that settles nearby a city, which also has variable energy demands. In this way, the optimal load distribution between components such as internal combustion engines and high temperature heat pumps can be achieved. The incremental capital cost between the typical and integrated brewery scenarios (5.3 M€/year) is offset by an incremental operating income, leading to an incremental total revenue of about 66.4 M€/year in the integrated scenario. As a result, more stringent environmental and economic scenarios, triggered by a shortage of fossil resources, could be endured thanks to the higher flexibility achieved using a set of renewable, and heat pump and power-to-gas technologies.

**Keywords:** heat pumping, renewable energy, brewery, district heating

## 1. Introduction

Brewing is responsible for a large amount of carbon dioxide emissions, derived not only from the fermentation process, but also from burning natural gas and from the indirect emissions associated to the supply chains of the utilities consumed (e.g. electricity from grid). Fermentation yields approximately three times more CO<sub>2</sub> (about 30 g per liter of beer), than what is actually required per each batch of beer, including brewing, canning, and bottling. Some breweries have already implemented CO<sub>2</sub> management strategies, although it may become prohibitively expensive, unless efforts to find additional and more profitable uses for CO<sub>2</sub> capture are realized. In this regard, brewery, as well as other food and biomass-based industries (e.g. sugar and ethanol mills, pulp and paper plants, etc.) (Domingos 2022, Telini 2022), may benefit from a promotion of net negative CO<sub>2</sub>

emissions in future scenarios of energy transition. Indeed, biomass-based industries that capture more CO<sub>2</sub> than they can use may choose either to sell it, use it in greenhouses to boost plant growth, or trade and inject the biogenic emissions, which directly contributes to a depletion of the CO<sub>2</sub> in the atmosphere and increases the plant revenues (Nakashima 2019, Domingos 2021). On the other hand, the increasing integration of renewable, yet intermittent energy technologies, is an option to decarbonize the industrial applications and calls for better carbon management strategies, such as its storage and conversion into synthetic natural (SNG) gas using power-to-gas systems subject to seasonal availability of inexpensive electricity (Flórez-Orrego 2022). Moreover, the surplus power and low-grade waste heat derived from additional energy conversion technologies in the beer industry can be used to supply the energy requirements of an urban agglomeration that settles nearby the brewery process. As a result, synergistic waste heat and mass upgrading processes, along with suitable fuel storage and utilization approaches may help tackling inherent drawbacks of the renewable energy systems and decarbonize the process heating demand, while increasing the process revenues and reducing the cooling duty. Thus, in this work, the integration of more advanced energy technologies, such as supercritical heat pumping and energy storage and conversion systems is compared to the conventional operating scenario of a brewery plant and a city district heating based on natural gas. The cross-cutting challenges to develop and adopt those low carbon industrial heat and power technologies are weighed using thermo-economic and environmental indicators. As a result, despite more equipment and electrification are required, the potential reduction of CO<sub>2</sub> emissions and the calculated incremental revenues point towards a promising path for beer industry decarbonization while maintaining high profits.

## 2. Process description and methods

Figure 1 shows the flowsheet of the industrial brewery along with the cogeneration and storage technologies able to be integrated. The brewery process starts when milled barley and corn are mixed with water and heated at 65°C, so that a sugar-rich wort is produced due to saccharification of the mash. In the lautering process, the wort is filtered and the husk can be exported or sent to a biodigestion unit. The filtered mixture is boiled slightly above 100°C (10% wt. evaporation) and then the hops are added. Next, it is stirred and the trub is removed in a centrifuge. The mixture is cooled down to 10°C and goes through the fermentation process, wherein the sugars are converted to alcohols and carbon dioxide using yeasts under anaerobic conditions by releasing heat. Fermented wort is matured at low temperature and then chilled before the yeast is eliminated by filtration. Cold water is mixed up to obtaining the desired alcohol concentration. The pasteurization process occurs hereafter by heating the mixture up to 70°C and cooled to 5°C. The beer is stored and bottled. The bottle cleaning process is also energy consuming and produces waste water. Meanwhile, the utility system options include gas fired boiler, combustion engines with waste heat recovery, sub and supercritical heat pumps, carbon abatement unit, water cooling and refrigerator, biodigestion, electrolysis, methanation, liquefaction, storage and injection compression units. SNG and CO<sub>2</sub> are stored in liquid form at -162°C and -50°C, respectively. The district heating relies on a CO<sub>2</sub> network system that supplies the domestic hot water, space heating, air conditioning and refrigeration demands (Fig. 2) (Suciu 2019). The waste heat released by the industrial process and the utility systems can be upgraded using surplus electricity and heat pumping systems. In this work, two scenarios are analyzed, i.e. a typical brewery operating using natural gas and electricity from the grid, and other scenario integrating the energy conversion and storage systems shown in Fig. 1, but still allowing for electricity import. Optimal flows have been determined using Osmose Lua platform (Yoo, Lessard et al. 2015), which handles a MILP that identifies

the best energy technologies and operating conditions that minimize the operating costs subject to feasibility and minimum energy requirements constraints, shown in Eqs. (1-6).

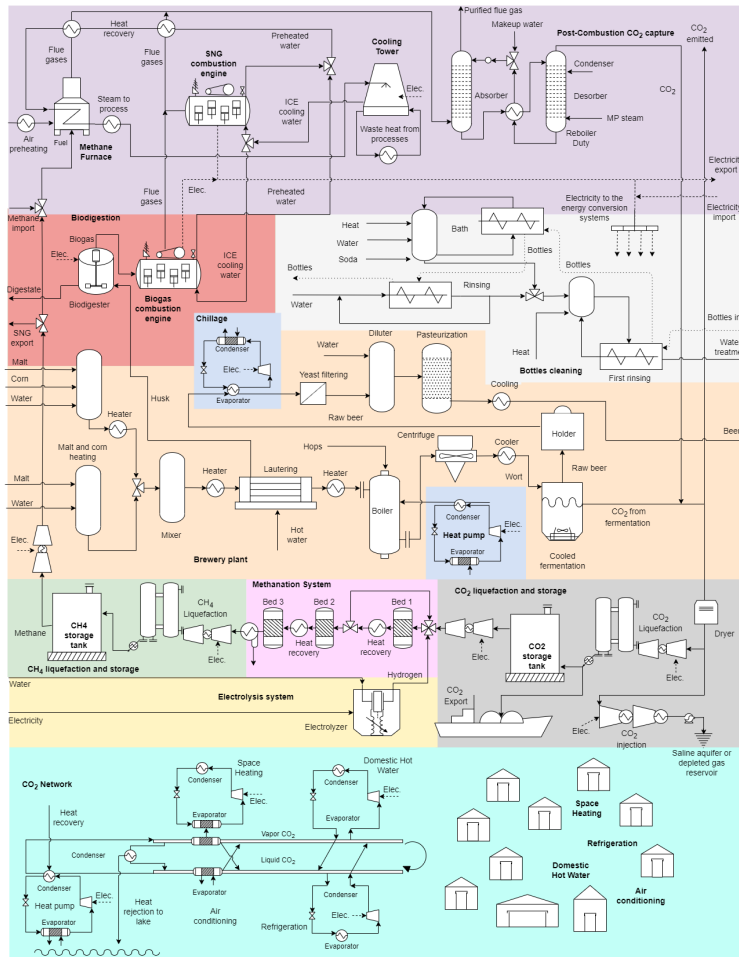


Figure 1. Flowsheet of the brewery and potential cogeneration and storage technologies, and the district heating CO<sub>2</sub> network.

The binary  $y_w$  and load factor  $f_w$  optimization variables are used to define the best size and arrangement based on assumed prices of material and energy inputs and CO<sub>2</sub> tax.

$c_{NG}$  (€/kWh): 1;  $c_{EE}$  (€/kWh): 0.35 (Jan/Feb/Nov/Dec), 0.001 (other months);  $tax_{CO_2}$  (€/t): 100;

$$\text{Min} \left[ f_{\text{material}} \times (c \cdot m)_{\text{material}} + f_{\text{EnvEm}} \times (m_{CO_2} \cdot tax) - f_{\text{material}} \times (c \cdot m)_{\text{material}} \right] \quad (\text{Eq.1})$$

$$\pm f_{EE} \times (c \cdot W)_{EE} + \frac{Z_{\text{equip}} \times \text{Ann\_factor}}{N_{\text{hours per year}}} \quad (\text{Eq.2})$$

$$\text{s.t.} \sum_{\omega=1}^{N_{\omega}} f_{\omega} q_{\omega,r} + \sum_{i=1}^N Q_{i,r} + R_{r+1} - R_r = 0 \quad \forall r = 1 \dots N \quad (\text{Eq.2})$$

$$\sum_{\omega=1}^{N_{\omega}} f_{\omega} W_{\omega} + \sum_{\text{chemical units}} W_{\text{net}} + W_{\text{imp}} - W_{\text{exp}} = 0 \quad (\text{Eq.3})$$

$$f_{\min,\omega} Y_{\omega} \leq f_{\omega} \leq f_{\max,\omega} Y_{\omega} \quad \forall \omega = 1 \dots N_{\omega} \quad (\text{Eq.4})$$

$$W_{mp} \geq 0, W_{exp} \geq 0 \quad (\text{Eq.5})$$

$$R_1 = 0, R_{N+1} = 0, R_r \geq 0 \quad (\text{Eq.6})$$

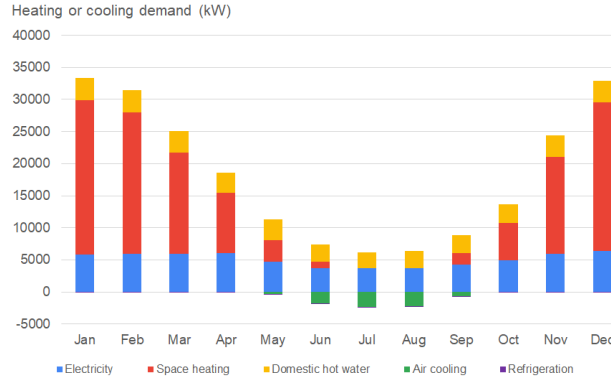


Figure 2. Seasonal variation of the heating and cooling demands of the city.

### 3. Results and discussion

For both typical and integrated beer production scenarios, the total malt, corn, soda and water consumption are 129, 11 and 19 kt/year and 973,263 m<sup>3</sup>/year, respectively, for a total production of 892,534 m<sup>3</sup>/year of beer and 38 kt/year of husk with a potential biogas yield of 75 Nm<sup>3</sup>/t<sub>husk</sub>. Annual direct CO<sub>2</sub> emission from natural gas use (64.95 GWh/year) in the typical scenario is 13 kt/year, whereas 24 kt/year are vented from fermentation process and other 6 kt/year are related to indirect emissions of NG/EE supply chains (1.15 kt/year for NG, and 5.14 kt/year for EE), entailing a yearly tax of 4.2 M€. In the typical brewery scenario, the chiller system consumes up to 15 GWh/year of EE, being the second largest EE consumer, after the city (44.6 GWh/year of EE, 12.3 GWh/year for space heating heat pumps). Meanwhile, in the advanced scenario, all husk is biodigested to cut down fossil fuel consumption. Husk-derived biogas falls short to supply the energy demands of the plant, thus requiring to import 133.1 GWh/year of electricity to drive the electrolysis unit (76.6 GWh/year, 12,532 m<sup>3</sup><sub>water</sub>/year, 1,39 kt<sub>H<sub>2</sub></sub>/year). Electricity import is about 41.25% of the total power consumed in the integrated brewery plant, leading to indirect emissions of 8.34 kt/year. In the integrated scenario, 16 kt/year of CO<sub>2</sub> from the fermentation unit are also injected and only 8 kt/year are emitted. Apart from biogas, the SNG produced by methanation (39 GWh/year) is also used in internal combustion engines to cogenerate the heat and power required by the whole system. Other major power consumers are sCO<sub>2</sub> heat pump (11.9 GWh/year), while gas liquefaction, recompression and injection compression together only demand about 3 GWh/year of EE. Carbon credits (1.6 M€/year) associated to the injection of CO<sub>2</sub> definitely contribute to the increased revenues of the integrated brewery scenario. It is worth noticing that the carbon capture unit becomes prohibitively expensive and is not activated by the optimization routine. Figure 3 shows the integrated curves of the typical (Fig. 3 left) and the integrated (right) brewery scenarios. The blue curve represents the heat cascaded plot of the utility systems, whereas the red curve corresponds to process units (brewery plant). The typical scenario, based on the natural gas combustion, not only has higher associated CO<sub>2</sub> emissions, but also entails larger exergy destruction due to the use of high-grade energy resources, such as natural gas, to only generate steam and hot water (Fig. 3 left). In fact, self-sufficient zones indicate a potential to install a temperature-gliding supercritical heat pump to partly

deliver the heating demand at temperatures between 50 - 100°C. The heat pumping of the waste heat available at much lower temperatures reduces the need of water cooling, while provides the refrigeration duty. But it also implies a compromise between the utilization of the combustion gases of the engines and the heat pump devices to deliver waste heat to the brewery and the district heating. Anyhow, surplus electricity generated in the integrated brewery scenario allows to simultaneously export electricity to drive the heat pumps of the CO<sub>2</sub> network and capitalize on the low-grade waste heat of all the energy technologies to reduce the environmental footprint of the heating processes.

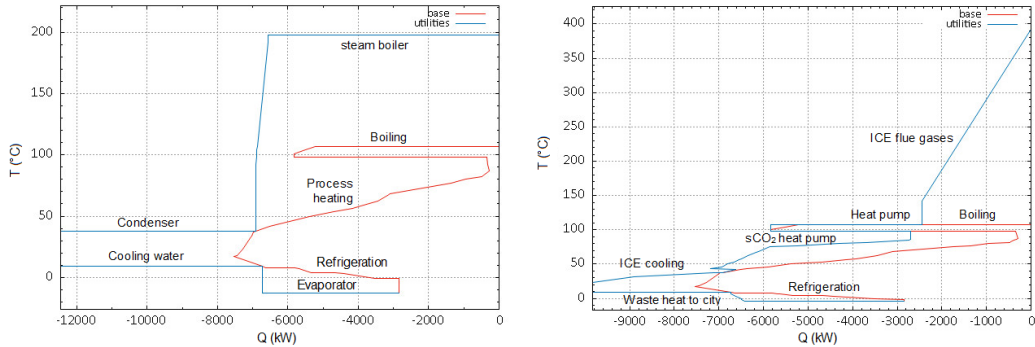


Figure 3. Integrated curves of the typical (left) and the integrated (right) brewery scenarios.

Finally, Figure 4 shows the yearly variation of the storage tanks levels. It is observed that the CO<sub>2</sub> tank level increases gradually starting from October till end of February, which coincides with the emptying of the SNG tanks during the period of more expensive electricity price (Nov.-Feb). Thus, optimal CO<sub>2</sub> management implies a higher monthly power consumption during the months with surplus cheap electricity generation (e.g. excess solar, nuclear or hydro power), which in turn can be stored as fuels.

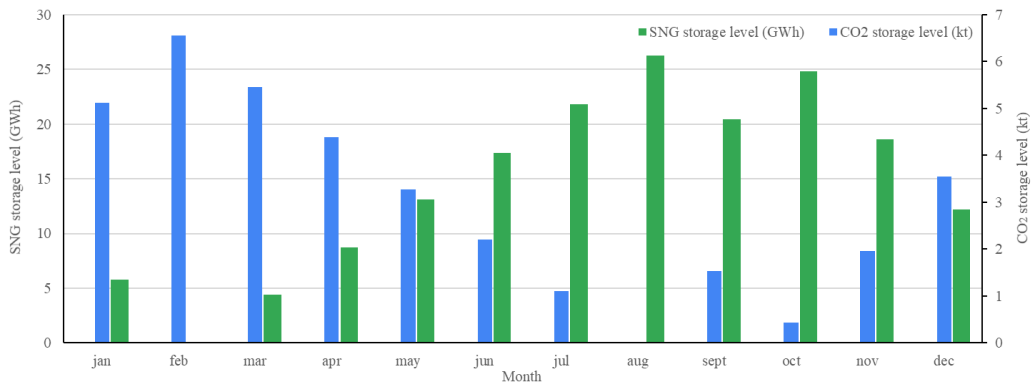


Figure 4. Seasonal variation of the storage tanks levels.

As a result, the operation cost is optimized, because the system uses the most affordable energy resources only when they are available and convert them into other valuable resources that can be used to supply the variable energy demands. In this way, the upgrading of the industrial residues and their ingenious reutilization increases the plant energy security, and contributes to increased revenues of the whole system. Indeed, the incremental capital expenditure between the typical and the integrated brewery scenarios (5.3 M€/year) is completely offset by an incremental operating income (71.7 M€/year),



thus leading to an incremental total revenue of about 66.4 M€/year in the second scenario. Naturally, the proposed approach represents a riskier investment, since it depends on the commodities prices, the carbon tax, the emissions trading system, the reliability of the newly integrated energy systems and the space budget of the brewery unit. Nonetheless, the proposed setup sheds light on the thermoeconomic and environmental feasibility of integrating advanced energy technologies for tackling the problem of carbon management and intermittency of renewable energy.

#### 4. Conclusions

In this work, the integration of various energy conversion and storage technologies to decarbonize a beer production industry is discussed in the light of thermodynamic and environmental indicators and economic revenues. Future economic and environmental costs of natural gas calls for more electrification to supply heating in a more efficient way (HPs and HTHPs). Meanwhile, the integration of husk biodigestion process, along with the reuse of CO<sub>2</sub> from fermentation as a source to store intermittent renewable energy in the form of synthetic natural gas may help to avoid the import of fossil natural gas and still provide half of the power required by the brewery and the heat pumps of the district heating network. The import of electricity is still largely required to balance the capacity of the cogeneration systems, thus, more environmentally friendly and cheaper electricity generation technologies are crucial for the process heat decarbonization in the industry. The carbon trading is an important factor that may boost the deployment of some energy conversion approaches, such as power-to-gas systems, and also promote better CO<sub>2</sub> management approaches. As a result, chemical energy storage systems allowed for a cost-beneficial operating condition, in which the incremental capital expenditure seems to be offset by a substantial gain in plant revenues.

#### Acknowledgments

The authors thank the Swiss Federal Office of Energy grant no. SI/502336-01. First author also thanks the Colombian Ministry of Science, Technology and Innovation.

#### References

- Domingos, M. E. G., Flórez-Orrego, D., Teles dos Santos, M., Oliveira Junior, S., Maréchal, F. (2022). "Techno-economic and environmental analysis of methanol and dimethyl ether production from syngas in a kraft pulp process." *Computers & Chemical Engineering*: 107810.
- Domingos, M. E. G., Florez-Orrego, D., Teles, M., Velasquez, H., de Oliveira Jr, S. (2021). "Exergy and environmental analysis of black liquor upgrading gasification in an integrated kraft pulp and ammonia production plant." *International Journal of Exergy* **35**(1): 35-65.
- Flórez-Orrego, D., Domingos, M. E. G., Maréchal, F. (2022). A systematic framework for the multi-time integration of industrial complexes and urban systems. *7th International Conference on Contemporary Problems of Thermal Engineering - CPOTE, 20th-23th September, 2022*. Warsaw, Poland.
- Nakashima, R. N., Flórez-Orrego, D., de Oliveira Junior, S. (2019). "Integrated anaerobic digestion and gasification processes for upgrade of ethanol biorefinery residues." *Journal of Power Technologies* **99**(2): 104-114.
- Suciu, R., Stadler, P., Kantor, I., Girardin, L., Maréchal, F. (2019). "Systematic Integration of Energy-Optimal Buildings With District Networks." *Energies* **12**(15): 1-38.
- Telini, R., Florez-Orrego, D., Oliveira Jr., S. (2022). "Techno-economic and environmental assessment of ammonia production from residual bagasse gasification: a decarbonization pathway for nitrogen fertilizers." *Frontiers in Energy Research* **10**(881263).
- Yoo, M., L. Lessard, M. Kermani and F. Maréchal (2015). *OSMOSE Lua: A Unified Approach to Energy Systems Integration with Life Cycle Assessment*. 12th International conference PSE 2015 and 25th International conference ESCAPE 2015., Copenhagen, Denmark.

# Dynamic Simulation to Verify Operability of LNG Plants with Post Combustion Carbon Capture

Jaleel Valappil<sup>a</sup>

<sup>a</sup>*Bechtel Energy, 3000 Post Oak Boulevard, Houston, TX - 77407, USA*

## Abstract

Global warming and climate change caused by greenhouse gases have received widespread and growing concern in recent years. The natural gas/LNG industry has seen tremendous over the years. Most of the facilities use gas turbine drivers which emit significant amount of flue gases with CO<sub>2</sub>. The carbon capture and storage technology (CCS) is one opportunity to capture CO<sub>2</sub> emissions from flue gases of combustions processes. The retrofitting of existing natural gas facilities with a carbon dioxide capture offers a promising opportunity to achieve the global target reduction in CO<sub>2</sub> emissions. Previous studies have examined the integration of power plants with carbon capture units and the resulting operability challenges and optimization opportunities. Integrating power plant with carbon capture simulation models for analysis has also been performed in the past [1]. To get a full understanding of the transient characteristics of a hydrocarbon facility with post combustion carbon capture (PCC), both the units can be combined in an integrated dynamic model for analysis. A case study is presented where the exhaust from LNG facility is send to carbon capture unit. Several valuable recommendations were derived from the study to ensure the operability of the integrated facility.

**Keywords:** Post combustion carbon capture, Natural gas, Dynamic simulation, Operability, Amine system.

## 1. Introduction

Global warming and climate change caused by greenhouse gases have received widespread and growing concern in recent years. At the United Nations Climate Change Conference in 2015, 196 countries accepted the Paris agreement to reduce their carbon emissions with the goal to keep the global warming below 2 °C. Further, The COP 27 conference held in Sharm-El-Sheikh recognized that limiting global warming to 1.5 °C requires rapid, deep, and sustained reductions in global greenhouse gas emissions of 43% by 2030 relative to the 2019 level. The natural gas/Liquefied Natural Gas (LNG) industry has seen rapid growth over the years with several grassroots plant built around the world to meet the global energy needs. Most of the facilities use gas turbine drivers which emit significant amount of flue gases with CO<sub>2</sub>. In addition, fired heater used in these facilities also emit CO<sub>2</sub>. CO<sub>2</sub> is also emitted by acid gas removal processes in the pre-treatment of feed gas prior to liquefaction. The carbon capture and storage technology (CCS) is one opportunity to capture CO<sub>2</sub> emissions from flue gases of combustions processes. The captured CO<sub>2</sub> can be stored in underground formations afterwards. LNG facilities are expected to be designed or retrofitted with CCS in future [4]. However, most of the LNG

plants in operation currently are not designed for carbon capture. For this reason, post-combustion CO<sub>2</sub> capture (PCC) has an advantage over other alternatives because the technology can simply be implemented as an ‘end-of-the-pipe’ retrofit without the need for significant changes to existing facilities. The most widely used technology for post combustion CO<sub>2</sub> capture is amine solvent-based process.

With the retrofitting of existing natural gas facilities with a carbon dioxide capture, any operational disturbances in the hydrocarbon facility can impact the operation of the carbon capture in an adverse way and vice versa. Accounting for these interactions in the design of carbon capture facility is important in ensuring operation of the integrated facility. This paper discusses the integration of an amine-based post combustion carbon capture unit with an LNG facility and the use of dynamic models to account for resulting effects of the integration on the LNG facility operation.

In order to get a full understanding of the transient characteristics of a hydrocarbon facility with PCC, both the units can be combined in an integrated dynamic model for analysis. Dynamic simulation can play a pivotal role in identifying any operational bottlenecks at transient conditions for the integrated hydrocarbon and CO<sub>2</sub> capture plants. The integrated model can also be valuable to establish operational strategies on both sides to account for the various upstream scenarios. As an alternative, individual unit models can be used to establish boundary conditions for each other to get a representative behavior of the entire facility. A case study is presented where the exhaust from LNG facility is sent to carbon capture unit. The dynamic model of the Waste Heat Recovery Unit (WHRU), duct work and the associated equipment is used to ensure that the various transient scenarios in LNG facility or in the PCC unit does not adversely impact the LNG plant operation.

## **2. Integration of Post Combustion Carbon Capture Systems with Existing Plants**

Post-combustion attractive for new LNG plants or as a retrofit to existing plants is an attractive option as it can be implemented without major modifications to the existing facility. There is also a good amount of experience with post combustion capture in power plants which makes it a less risky option. Also, downtime in carbon capture plant does not have to affect LNG production as the flue gas can be diverted to atmosphere with appropriate design [1]. The use of amine-based solvent for post-combustion CO<sub>2</sub> capture (PCC) has attracted considerable interest in CO<sub>2</sub> capture. This is a proven technology for CO<sub>2</sub> capture to date. Existing PCC pilot plants already demonstrate the feasibility of CO<sub>2</sub> removal with amine solvent from flue gas albeit at modest scale. Mono ethanolamine (MEA) is the most commonly used solvent in chemical absorption processes for CO<sub>2</sub>. A typical PCC process consists of two main columns, an absorber and a regenerator. These may be packed or tray columns, but packed columns are typically used due to their larger contact area. In the absorber, lean amine solvent is counter-currently contacted with flue gas produced from the LNG plant and CO<sub>2</sub> is chemically absorbed from the flue gas. The CO<sub>2</sub>

rich solvent is regenerated in the stripper using heat from either the hot oil, steam or other source as the heating medium.

The integration of post combustion carbon capture with LNG plant is done by directing the flue gas from gas turbines to amine-based carbon capture unit. This will include additional ducting, dampers and possibly a blower to increase flue gas pressure. A Waste Heat Recovery Unit (WHRU) may be used to further recover the heat from flue gas. The heat requirements for the reboiler can be provided by the existing hot oil or steam system in the facility.

The following aspects are important when integrating post combustion carbon capture into an existing facility considered to ensure reliable and safe operation [2].

**Flexibility of operation:** Ability to operate in an acceptable manner over a range of steady-state conditions. For example, this can include operation at turndown rates (Half rate/other) due to changing ambient conditions or other operational reasons. Also, operation with no LNG production (with refrigeration compressors running in recycle with gas turbines partially loaded).

**Controllability:** Ability to recover from process disturbances and move to new set-points in a measured and timely fashion. This is dependent on both the design of process and the underlying process control schemes.

**Start-up and shutdown:** The startup and shutdown of LNG facility/PCC unit is a key operating scenario for the amine system that can induce major changes.

**Reliability/Safety of operation:** Ability to accommodate equipment failures and trips in a safe manner. This includes trips of various equipment, fail closing of valves etc.

The design and integration post combustion carbon capture needs to account for the various operational scenarios and disturbances in the LNG facility and the carbon capture unit. One of the main concerns with integration is the impact of upsets in PCC facility affecting the LNG production. Large variations in gas turbine exhaust pressure can cause the turbine driver to trip, stopping the LNG production with a very significant economic cost. The key scenarios in LNG plant that impact the integrated system include LNG train trip/facility wide shut down, trip of individual gas turbines and turndown to half rate/standby mode with refrigeration compressors in recycle. Startup of refrigeration compressors/PCC unit is also an important case to consider. The scenarios to consider on the PCC side include trip of flue gas blower, fail closing of flue gas damper and other operational upsets in PCC unit. The operability of the integrated LNG – PCC plant can be studied with high fidelity dynamic simulation tools to ensure the design can handle these scenarios.

### *2.1. Application of Dynamic Models in Design and Analysis*

Dynamic simulation utilizing rigorous mathematical models has become an influential tool in the process design, design validation, control system verification, startup support

and troubleshooting. Many recent developments in the simulation software and technology have led to the development of large-scale dynamic models and their use in the entire project life. Dynamic simulation can be used to evaluate the operational response and controllability of a plant for a wide range of transient conditions, which may include disturbances due to equipment malfunction or human error, fluctuations in ambient temperature, or changes in feed conditions. Dynamic models can also be used for development of online process optimization strategies is essential to ensure the process operates under optimal conditions.

The scope of the dynamic modeling for the integrated LNG-PCC facility can be divided into three areas as shown in Figure 1. There are different options (model scope/methodology) to study the overall behavior of the integrated LNG plant – PCC facility. This can be selected based on the scope and objectives of the analysis.

For a rigorous understanding of the transient characteristics of an LNG plant with PCC, a combined dynamic model is preferred. The transient responses of the LNG plant are well understood and has been studied over the years. The behavior of integrated LNG-PCC facility is a relatively recent development and can benefit from integrated model-based analysis. The dynamic behavior of amine system is more complicated and less understood compared to that of the LNG plant.

Other option is to have individual models of the three sections of the integrated facility. This way, individual models can be run and used as the boundary condition for the next model. The case study given below is intended to look at upsets in the interconnecting units affecting the LNG facility and vice versa.

The modeling of LNG facilities is normally carried out during the detailed design. This could be a plantwide model developed for engineering studies or an Operator Training simulator (OTS). The use of these models for design and operation of LNG plants have become more widespread in the recent years [4]. The dynamic modeling of interconnecting ductwork, WHRU and flue gas blower can also be performed with available commercial dynamic modeling tools.

Many of the commercially available process simulation software can be used to develop a dynamic model of amine-based CO<sub>2</sub> capture processes [3]. The first option to model this is to use equilibrium-based modeling approach. The equilibrium-based modelling assumes that liquid and vapor phases reach equilibrium at theoretical stages and perfect mixing occurs at each stage. The equilibrium-based approach adjusts the performance at each stage with an efficiency correction factor (tray efficiency for tray columns or height equivalent of a theoretical plate (HETP) for packed). The equilibrium approach simplifies the mass transfer process to reduce the complexity of the model, neglecting the limitations of mass and heat transfer.

Equilibrium is rarely acquired in a reactive CO<sub>2</sub> absorption process especially during dynamic operation. The rate-based approach is more rigorous and considers the actual rates of mass and heat transfer including chemical reactions. The reaction depends on the kinetics regime and can either be considered instantaneous or kinetically controlled. Rate based approach is preferred for modeling reactive columns as in the PCC process. The

disadvantage of this approach is the computational requirements as rigorous rate-based approach results in extended systems of equations to be solved at each integration step.

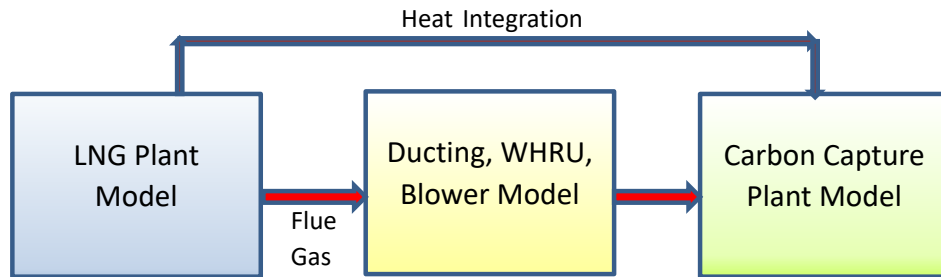


Figure 1: Different section of an Integrated LNG-PCC unit dynamic model

### 3. Case Study – Analysis of Operability of LNG Plant with PCC Unit

This case study evaluates the integration of amine-based PCC unit with an LNG facility utilizing gas turbines to drive refrigeration compressors. The addition of the PCC unit involves introduction of additional ductwork, bypass dampers (to atmosphere and to PCC unit) and a flue gas blower to increase pressure. The flue gas is further cooled a heat exchanger and then routed to the Absorber where it is contacted with amine solvent in the Absorber to remove the CO<sub>2</sub> from flue gas. The CO<sub>2</sub> depleted flue gas is then released to atmosphere. The main concern with the addition of PCC unit is the possibility of impacting LNG production during various transient upsets in the PCC unit. The gas turbine exhaust outlet must be maintained within a pressure range to avoid a reduction in power output or a total turbine shutdown. Several scenarios were studied to evaluate the transient impact on the gas turbine outlet conditions. A dynamic simulation model of the interconnecting unit (WHRU, flue gas blower, other equipment, and associated piping) was developed to perform this study.

The flue gas blower trip is simulated to verify that the pressure at GT outlet can be maintained within limits so that LNG plant operation is not affected. As soon as the flue gas blower is tripped, the damper to stack is signaled to open to avoid the excessive pressure increase in the system. The blower will coast-down with speed response based on the rotational inertia of the string. The pressure at the GT outlet is shown in Figure 2. As can be seen, the pressure drops rapidly as the damper is opened along with the blower trip. This preemptive action prevented the pressure at gas turbine outlet from increasing due to trip of blower.

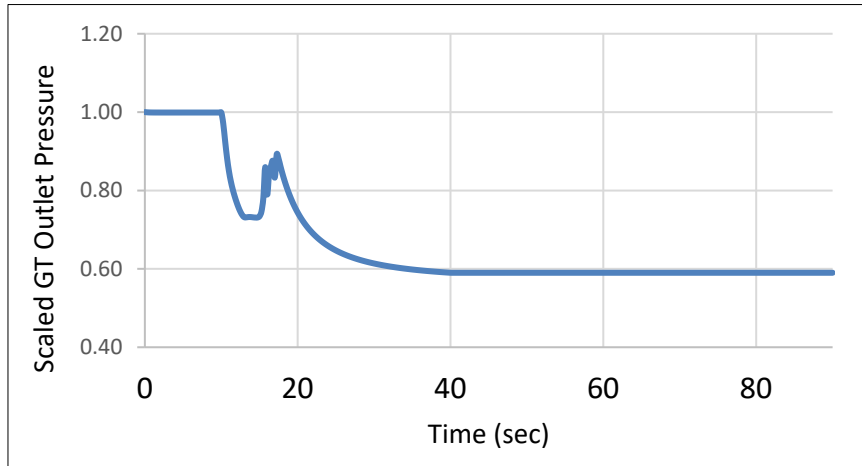


Figure 2: Variation in gas turbine outlet pressure during flue gas blower trip.

The trip of a Gas Turbine (GT) will result in rapid change in temperature and flow rate of flue gas at the inlet of WHRU. The variation of flue gas flowrate and temperature from gas turbine vendor is used. Once the GT is tripped, the flue gas flow reduces, decreasing the pressure at the inlet of flue gas blower. The flue gas blower must be shutdown to prevent significant vacuum formation at suction. Shutting the flue gas blower too early can result in over pressuring the duct as the flue gas flow continues for some time. An alternate option is to shut down the flue gas blower and open the damper to stack as soon as the GT trips. This will prevent over pressuring the duct as the flue gas flow gradually comes down. This operational procedure and timing was tested in the dynamic simulation and the pressure profile at the GT outlet for this scenario is shown in Figure 3.

Several other scenarios were also evaluated with this model to identify any operability concerns. These included dampers fail closing, shutdown of PCC unit, startup of PCC unit and others. The transition from operation without PCC unit to PCC unit was simulated to verify the operating procedure for this transition.

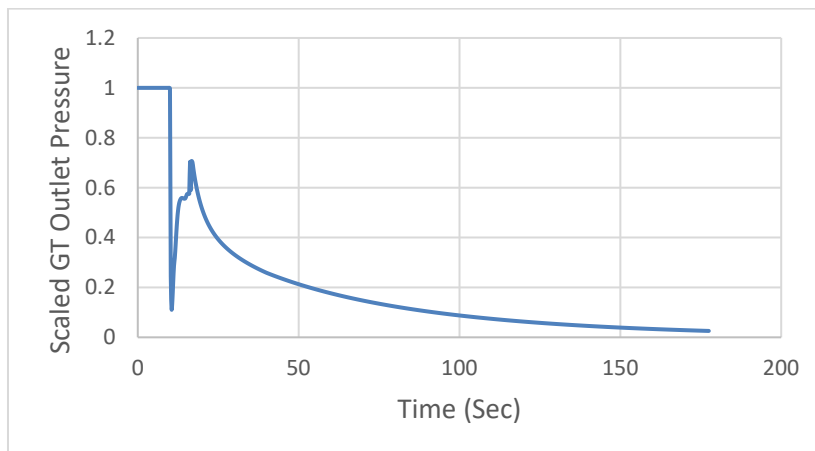


Figure 2: Variation in gas turbine outlet ducting pressure during a gas turbine trip.

#### **4. Conclusions and Recommendations**

The concerns about global warming have spurred interest in application of carbon capture and storage in the hydrocarbon industry. This paper examined the benefits of using first principle based commercially available dynamic simulation tools to optimize and enhance the design of integrated of PCC - LNG facility. The key operating scenarios that impact the facility operation were selected and simulated.

The insights gained from dynamic simulation was valuable in verifying the operability and identifying design parameters. The key operational recommendations included 1. Shutting down the flue gas blower when damper fail closes and during the GT trip, 2. Opening bypass damper when a damper fail closes or during flue gas blower trip and 3. Sequencing of the blower startup and damper opening during startup. The opening time for the dampers and the sequence of operation of various components were verified using simulation.

Further optimization of the integrated facility is possible using rigorous integrated dynamic models for LNG plant and the PCC unit. This will enable optimizing the entire facility dynamically by selecting the right extent of carbon capture while considering the energy cost (Mainly in the amine regenerator reboilers). This has been studied for power plants in the past and can be extended to LNG facilities.

#### **References**

1. Bui, M., Gunawan, I., Verheyen, V., Feron, P., Meuleman, E. and Adeloju, S., 2014, Dynamic modelling and optimisation of flexible operation in post-combustion CO<sub>2</sub> capture plants—A review, *Computers and Chemical Engineering*, 61 (2014) 245-265.
2. Chukukwa, A., Enaasen, N., Kvamsdal, H.M. and Hillestad, M., 2012, Dynamic modeling of post-combustion CO<sub>2</sub> capture using amines – A review, *Energy Procedia*, 23(2012) 82-91.
3. Mokhatab, S., Mak, J., Valappil, J. and Wood, D., 2014, *Handbook of Liquefied Natural Gas*, Mokhatab, S., Mak, J., Valappil, J. and Wood, D., Elsevier Publications.
4. Techno-economic evaluation of CO<sub>2</sub> capture in LNG, IEAGHG Technical Report, October 2019.





# Optimization of wind farm layout and cable network topology under generalized wake effects

Demian J. Presser,<sup>a,b</sup> M. Josefina Fiorini,<sup>a</sup> Diego C. Cafaro<sup>a,b</sup>

<sup>a</sup>*Fac. de Ing. Química, UNL, Santiago del Estero 2829, Santa Fe (3000), Argentina*

<sup>b</sup>*INTEC, UNL-CONICET, Guemes 3450, Santa Fe (3000), Argentina*

## Abstract

The world is undergoing a significant change of the energy matrix towards a sustainable future. Wind energy has proved to be one of the key players in this transition. A new historical record of installed capacity was achieved in 2021, currently supplying 7% of the world energy demand. The optimal design of wind farms is a still challenging problem for industry, aiming at maximizing power generation while minimizing costs and energy loss, also accounting for uncertain wind speed and direction. We propose a mixed integer linear programming formulation coupled to a wake effect computation algorithm to optimally locate and connect turbines in a wind farm under uncertain conditions. In contrast to previous contributions, the proposed framework accounts for combined wake effects that can significantly compromise turbine efficiency and drastically affect optimal wind farm designs and cable network topologies. An illustrative case shows improvements of 6.8% relative to state-of-the-art approaches.

**Keywords:** Wind Farm, Layout, Design, Cable Network, Wake Effect, MILP

## 1. Introduction

The total power produced by a wind farm is directly related to the number of turbines installed. However, the location of these turbines has a considerable impact on their performance due to wind interference interactions known as wake effects. These phenomena occur when upstream turbines reduce the wind speed for downstream turbines. Upstream turbines produce a wake on its back side where the wind speed is reduced, yielding poor performance and eventual damage in downstream turbines. Naturally, interference effects across the farm depend on the wind speed and direction, which vary frequently. Even more, combined wake effects coming from multiple turbine interactions may severely impact the performance and the need for maintenance. Therefore, the optimally design of wind farms over a given region usually seeks to maximize the number of turbines installed while minimizing the impact of interferences. Sparse turbine locations reduce wake effects, but the total energy produced over the same region will be rather poor due to the limited number of turbines installed.

On the other hand, turbines must be connected to collection stations to further deliver power to the grid. Cable networks and stations can represent up to 30% of the cost of a wind farm, since more than one collection hub may be required depending on the spatial dispersion of turbines. In this sense, sparser designs to avoid wake effects, regardless of the wind direction, can significantly impact cable interconnection and stations costs.

One of the earliest contributions to the optimal design of wind farms can be found in Mosetti et al. (1994). The authors use the Jensen (1983) model to find the best arrangement of wind turbines over a region based on genetic algorithms. Herbet-Acero

et al. (2009) compare different heuristic approaches based on the Jensen’s model. Fischetti et al. (2015) propose a mixed integer linear programming (MILP) model to optimize wind farm layout accounting for wake effects between pairs of turbines. A thorough review of wind farm layout optimization using Jensen’s model is later presented by Shakoor et al. (2016). More recently, Mohki et al. (2020) and Bellat et al. (2022) use meta-heuristics to address the cable interconnection problem for a wind farm with a given layout. Finally, Fischetti et al. (2022) extend their approach to address one of the first problems integrating wind farm layout and cable routing. They develop Benders-like cuts and a tailored branch-and-cut strategy to deal with combinatorial complexity. Nevertheless, wake effects are still assessed between every pair of turbines. In this work, we simultaneously address wind farm layout and cable network problems under generalized wake effects. A novel algorithm accurately assesses combined effects prior to solve an MILP model dealing with interferences of higher order.

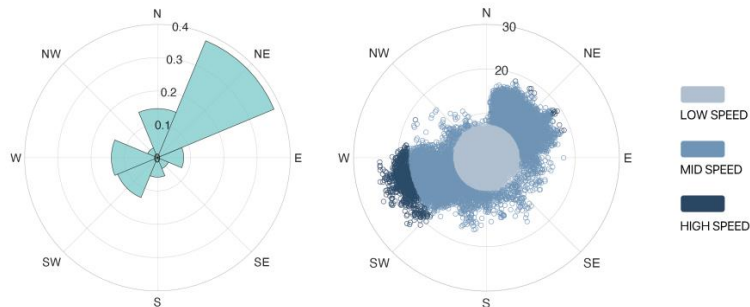
## 2. The wind farm layout and cable network problem

### 2.1. Problem statement

The wind farm layout and cable network problem aims to optimally place turbines, collection stations and cable interconnections between them. The problem can be stated as follows. Given: (a) a set of potential sites  $i \in I$  for the location of wind turbines and  $s \in S$  for collection stations; (b) power curves of the turbines; (c) possible wind scenarios ( $\pi(u)$  is the probability of scenario  $u \in U = D \times V$ , with direction  $d \in D$  and speed  $v \in V$ ); (d) the loss due to wake effects between potential turbine locations for each wind scenario; and (e) different types of cable and collection stations with capacities  $pl \in P$  and  $c \in C$ , respectively, for conveying and collecting power. We seek to determine: (i) the optimal number and location of wind turbines in the field, (ii) the optimal placement of stations to collect power, and (iii) the optimal cable interconnection between selected turbines and stations, in order to maximize the profit of the project over a long-term time horizon. Since wind behavior is uncertain, the optimal solution does not necessarily correspond to the best arrangement for a particular direction. One must weight all possible wind directions and corresponding speeds.

### 2.2. Assumptions

- Wind profiles are predicted from statistical data, discretized into directions ( $D$ ), speeds ( $V$ ) and possible scenarios ( $U = D \times V$ ). Figure 1 illustrates an example with  $D = \{N, NE, E, SE, S, SW, W, NW\}$  and  $V = \{\text{Low, Medium, High}\}$ .
- Jensen (1983) model is used to estimate the impact of wake effects between turbines. However, any other model can be also used in the proposed algorithm.
- The terrain is assumed to be flat and all turbines have the same height.



**Figure 1** – Histogram for wind direction probabilities (left) and speed discretization (right)

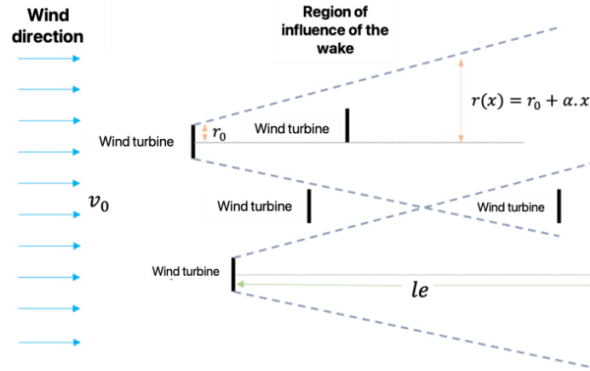


Figure 2 - Illustration of the area of interference due to wake effects.

### 3. Interferences and wake effect computation

Prior to optimization, an efficient algorithm is required to estimate interferences (i.e., productivity loss due to wake effects) in wind farms. Previous contributions underestimate the impact of interferences by only considering pairs of turbines. Instead, we evaluate impacts of higher order, when multiple turbines combine their wake effects. We address two phenomena: chained and overlapping effects. Chained effects refer to the successive wind energy loss when three or more turbines are placed one behind the other following the wind direction. On the other hand, two wind turbines may affect a third turbine in combination when their wakes overlap. The wake can be visualized as a conical region where the initial air velocity ( $v_0$ ) is reduced (see Figure 2). There is also a distance  $l_e$  beyond which the wake effect vanishes. If  $r_0$  is the radius of the wind turbine including the blades,  $r(x)$  defines the radius of the cone as a function of the distance ( $x$ ).

#### 3.1. Algorithm for computing interferences

Let  $K$  be the set of wind directions and  $WS(k)$  the set of velocities for each direction. Let  $INT_{i,j,k}$  be a 0-1 parameter indicating if a turbine placed in  $i$  interferes another turbine in  $j$  under direction  $k$ . Let  $vr_{i,j,k}$  be the proportion of the original wind speed that can be harvested by the turbine in  $j$  due to interference with  $i$  in direction  $k$ . Finally,  $C_T$  is the thrust coefficient of the turbines. The algorithm can be outlined as follows.

#### Algorithm

For all  $k$ , set  $i' = i_0$  as the first node to be reached by wind in such direction.

1. Build a line  $R_k$  orthogonal to direction  $k$ , passing through  $i'$ .
2. Compute the minimum distance  $d_{ik}$  from each potential location  $i$  to  $R_k$ .
3. Sort the elements in  $I$  by increasing  $d_{ik}$ .
4. For all  $i' \in I$  and  $i > i'$ :
  - 4.1. Let  $\mu$  be the displacement vector between  $i'$  and  $i$ .
  - 4.2. Compute the module of the projection of  $\mu$  over  $k$ ,  $\lambda = |\text{proj}_k(\mu_{i',i})|$
  - 4.3. If  $\lambda > l_e$ , set  $INT_{i',i,k} = 0$ ;  $vr_{i',i,k} = 1$ . Go 4.5. Else, go 4.4.
  - 4.4. Compute the module of the projection of  $\mu$  over  $R_k$ ,  $\gamma = |\text{proj}_{\perp k}(\mu_{i',i})|$
  - 4.5. If  $\gamma > r(|\mu|)$ , set  $INT_{i',i,k} = 0$ ;  $vr_{i',i,k} = 1$ . Go 4.7. Else, go 4.6.
  - 4.6. Node  $i$  is interfered by  $i'$ ,  $INT_{i',i,k} = 1$ ,  $vr_{i',i,k} = 1 + (\sqrt{1 - c_T} - 1) \left(\frac{r_0}{r}\right)^2$  (Jensen, 1983).
5. **Direct wake effects** (by-pair). Let  $Vf_{i,i',k}$  be the proportion of wind speed in direction  $k$  received by  $i$ , exclusively due to  $i'$ . Set  $Vf_{i,i',k} = vr_{i,i',k}$ .

6. **Indirect wake effects** (refinement). For all  $i'$  in the ordered set  $I$ :

- 6.1. If  $\sum_{i=1}^I \text{INT}_{i,i',k} = 0$ ,  $Vf_{i',k}=1$ ,  $i'$  is not interfered.
- 6.2. If  $\sum_{i=1}^I \text{INT}_{i,i',k} = 1$ , traverse the elements  $i$  in  $I$  until  $\text{INT}_{i,i',k}=1$  and update received velocity as  $Vf_{i',k} = vr_{i,i',k} Vf_{i,k}$  (chain effect).
- 6.3. If  $\sum_{i=1}^I \text{INT}_{i,i',k} > 1$ , add all  $i$  such that  $\text{INT}_{i,i',k}=1$  to set  $J$ . Then,

$$Vf_{i',k} = 1 - \sqrt{\sum_{j=1}^J (1 - Vf_{j,k} vr_{j,i',k})^2} \quad (\text{overlap effect})$$

The output of this algorithm is used to set input parameters  $ip_{i',i}$  (for direct effects) and  $is_{i',j,i}$  (for indirect effects) to the optimization model presented in the next section.

#### 4. MILP formulation

A mixed integer linear programming (MILP) formulation is proposed to simultaneously address the optimal wind turbines location and cable network interconnection for a wind farm site. The 0-1 variable  $x_i$  accounts for the location of a wind turbine in node  $i$ . Eq. (1) determines when two turbines  $i$  and  $j$  are installed ( $y_{i,j} = 1$ ) to account for direct wake effects. Notice that minimum distances between turbines are also imposed.

$$x_i + x_j - 1 \leq y_{i,j} \quad \forall (i,j), \quad \text{with } y_{i,j} = 0 \text{ if } d(i,j) < D_{min} \quad (1)$$

Additional constraints are set to address indirect effects. Whenever three or more nearby turbines are selected for placement, indirect effects may have to be assessed. Eq. (2) determines when three interacting turbines are installed, by means of the 0-1 variable  $ys_{i,i',j}$ . Similar constraints might be added to account for more turbine interactions.

$$x_i + x_{i'} + x_j - 2 \leq ys_{i,i',j} \quad \forall i, i', j \quad (2)$$

In Eq. (3), direct and indirect effects are combined into a total wake effect whenever the generator  $i$  is installed. Parameters  $ip_{i',i}$  and  $is_{i',j,i}$  refer to direct and indirect (by-three chained and overlap effects). They are obtained by computing the expected value of inferences coming from calculations provided by the algorithm in Section 3, after reducing the nodes involved to  $i$ ,  $i'$  and  $j$ , instead of the full set. Parameters  $wk_{k,ws,i',i}$  and  $wk_{k,ws,i',j,i}$  indicate power reduction in turbine  $i$  due to  $i'$ , and due to  $i'$  plus  $j$  in the latter case. Both of them are derived from the power curve by comparing generation under unrestricted and interfered airstream conditions. Eqs. (4) and (5) show how expected power reduction is obtained by weighing direction and speed probabilities.

$$\sum_{i'} ip_{i',i} y_{i,i'} + \sum_{j,i'} is_{i',j,i} ys_{i,i',j} \leq W_i + Wmax(1 - x_i) \quad \forall i \in I \quad (3)$$

$$ip_{i',i} = \sum_k \sum_{ws(k)} \pi_k \gamma_{k,ws} wk_{k,ws,i',i} \quad \forall i, i' \in I, i \neq i' \quad (4)$$

$$is_{i',j,i} = \sum_k \sum_{ws(k)} \pi_k \gamma_{k,ws} wk_{k,ws,i',j,i} - ip_{i',i} \quad \forall i, i', j \in I, i \neq i' \neq j \quad (5)$$

On the other hand, binary variable  $xs_{i,s,pl}$  indicates if turbine located at position  $i$  is connected to a collection station located in  $s$  with a cable of type  $pl$ . Eq. (6) states that a

turbine might only be connected to one station while Eq. (7) imposes that the capacity of station  $s$  must be unique, if installed ( $y_{s,c} = 1$ ).

$$\sum_{pl} \sum_s x_{s,i,pl} \leq 1 \quad \forall i \in I \quad (6)$$

$$\sum_c y_{s,c} \leq 1 \quad \forall s \in S \quad (7)$$

Furthermore, variable  $FP_{i,j}$  assesses the power conveyed from turbine  $i$  to station  $s$ . This flow must not be greater than the power that  $s$  is capable of managing (see Eq. 8).

$$FP_{i,s} \leq \sum_c pmax_c y_{s,c} \quad \forall i \in I, j \in J \quad (8)$$

Since each location  $i$  may serve as a power gathering node, a power balance is required to account for the incoming and outgoing currents. Continuous variable  $PE_i$  is the effective power generated by turbine located at  $i$  and is computed by discounting wake losses over theoretical power generation capacity (see Eq. 9).

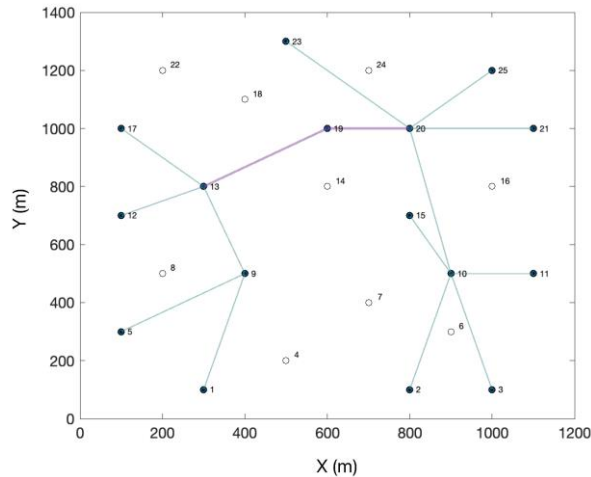
$$\sum_{o \neq i} FP_{o,i} + PE_i = \sum_{j \neq i} FP_{i,j} + \sum_s FP_{i,s} \quad \forall i \in I \quad (9)$$

Finally, the objective function accounts for investment costs from wind turbines, stations and cable connections, as well as operating costs of turbines and stations. Benefits from energy production over a time period  $\Delta$  are included to determine profit.

$$\begin{aligned} \max z = & \sum_i ((pw_i x_i - W_i) pv \Delta - inv x_i - opex x_i \Delta) \\ & - \sum_c \sum_s sub_{pl} y_{s,c} + \sum_i \sum_j \sum_{pl} cc_{pl} dist_{i,j} x_{s,i,j,pl} \end{aligned} \quad (10)$$

## 5. Results

An illustrative case study is solved to assess the potential of the optimization tool. The available land size is 1.68 km<sup>2</sup> with 25 potential sites for turbine placement. Generators involved are Vestas V80 of 2 MW and the wind profile over the region is the same as in Figure 1. The project is evaluated for a 7-year horizon and electricity price is assumed to be constant at \$0.06/kWh. Economies of scale are stated for collecting stations and cables. Three capacities are allowed for the installation of stations and 4 types of cable are available. Indirect wake effect calculations are extended up to 4 turbine interactions. The optimization model is implemented in software GAMS 40.2 and solved with GUROBI 9.5.2 in 75 seconds using an Intel Core i7 2.6 Ghz (6 cores) system with 16 GB of RAM. The model involves 21,993 equations and 3,967 variables (with 3,290 of them being 0-1). Solution suggests the installation of 16 turbines (2.5 MUSD each) with a total profit of 3.58 MUSD after 7 years. This implies 6.8% improvement with regards to separate optimization of layout and cable network. Investment costs for stations and cable interconnection amount to 4.02 MUSD. As shown in Figure 3, several turbines are connected with the cheapest cable of 70 mm<sup>2</sup> section. After collecting power of several turbines, the solution suggests to convey energy to the final station with 180 mm<sup>2</sup> cable.



**Figure 3** – Solution suggested by the model for the 25 potential sites illustrative case.

## 6. Conclusions

In this work, we have developed a novel preprocessing algorithm and an MILP model to integrally solve wind turbines layout and cable network design problems for a given wind farm. In contrast to previous contributions, the model involves the evaluation of generalized wake effects comprising direct and indirect interferences. That is, the evaluation of multiple interactions by means of chained and overlapping effects. The model is coupled with an efficient algorithm that allows accurately estimating wind energy loss from interactions of higher order using any surrogate model, even those of maximum complexity. Although the model is effective in solving moderate dimensional cases, combinatorics grows fast when more accuracy is required, leading to refined grids with more potential sites. Future work will be focused on developing solution strategies to address problems with denser grids and continuous modeling approaches, as well as the inclusion of losses in the cabling and decisions over transmission lines.

## References

- A. Bellat, I. Tyass, K. Mansouri, A. Raihani, 2022, New approach to optimize the cost and intercon. of wind turbines using the PSO alg. <https://doi.org/10.1051/e3sconf/202233600013>.
- M. Fischetti, J. Leth, A. B. Borchersen, 2015, A MILP approach to wind farm layout and inter-array cable routing. Proc. of Am. Control Conf. <https://doi.org/10.1109/ACC.2015.7172266>.
- M. Fischetti, M. Fischetti, 2022, Integrated Layout and Cable Routing in Wind Farm Optimal Design. Management Science, <https://doi.org/10.1287/mnsc.2022.4470>.
- J. F. Herbert-Acero, J.R. Franco-Acevedo et al., 2009, Linear wind farm layout opt. through comp. inter. Adv. in Artificial Int., 692–703. [https://doi.org/10.1007/978-3-642-05258-3\\_61](https://doi.org/10.1007/978-3-642-05258-3_61).
- N.O. Jensen, 1983, A Note on Wind Generator Interaction. Risø National Laboratory, No. 2411.
- C.E. Mokhi, A. Addaim, 2020, Optimization of Wind Turbine Interconnections in an Offshore Wind Farm Using Metaheuristic Alg., Sustainability, 12. <https://doi.org/10.3390/su12145761>.
- G. Mosetti, C. Poloni, B. Diviacco, 1994, Opt. of wind turbine positioning in large windfarms by means of a genetic algorithm. Journal of Wind Eng. and Ind. Aerodynamics, 51, 105-116.
- R. Shakoor, M. Hassan, A. Raheem, Y. Wu, 2016, Wake effect modeling: A review of wind farm layout optimization using Jensen's model. Ren. and Sust. Energy Reviews, 58, 1048-1059.

# Mixed-integer recourse in industrial demand response scheduling with interruptible load

Jnana Sai Jagana<sup>a</sup>, Satyajith Amaran<sup>b</sup>, Qi Zhang<sup>\*a</sup>

<sup>a</sup>*Department of Chemical Engineering and Material Science, University of Minnesota, Minneapolis, MN 55414, USA*

<sup>b</sup>*The Dow Chemical Company, Lake Jackson, TX 77566, USA*

<sup>\*</sup>*qizh@umn.edu*

## Abstract

While industrial demand response has become a prominent strategy for power-intensive chemical plants to remain cost-competitive, the financially incentivized provision of load reduction capacities, also called interruptible load, to the power grid is still a less explored topic. Here, a major challenge lies in dealing with the uncertainty that one does not know when load reduction will be requested. In this work, a scheduling model for a continuous industrial process providing interruptible load is developed, where we apply an adjustable robust optimization approach to address the uncertainty. The main difference to previous works is that we incorporate both continuous and binary recourse variables. When applied in our computational case study, the proposed model achieves significant cost savings when compared with a model that only considers continuous recourse.

**Keywords:** Interruptible load, multistage robust optimization, mixed-integer recourse

## 1. Introduction

In recent years, there have been significant advances in industrial demand response (DR) (Zhang and Grossmann, 2016), which has enabled large industrial consumers of electricity, such as power-intensive chemical plants, to reduce their operating costs by altering their power usage in response to varying electricity prices. DR is also beneficial to the power grid as it helps maintain grid stability, and a particularly effective way to do so is through the provision of interruptible load, which constitutes a form of ancillary services (Dowling et al., 2017). Here, economic incentives are offered to electricity consumers for committing load reduction capacities (that is, interruptible load) up to the agreed amount when requested by the grid operator.

The major challenge in providing interruptible load is that load reduction demand is not known in advance, but one must still guarantee dispatch upon request. Disregarding this uncertainty may jeopardize plant safety or lead to situations in which it is no longer possible to satisfy all product demand. Zhang et al. (2015) capture the uncertainty using a tailored uncertainty set and apply robust optimization (Ben-Tal et al., 2009) to this scheduling problem. However, they only model the static case where no recourse is considered, which leads to very conservative solutions. Zhang et al. (2016) address this shortcoming by incorporating continuous recourse decisions using affine decision rules to show significant increases in cost savings enabled by flexible recourse. Yet the proposed approach still cannot realize the full potential of interruptible load since it does not consider discrete recourse and hence does not allow, for example, full plant shutdowns when load reduction is required, which is what is often done in practice. In this work, we extend the previous framework to also include binary recourse decisions.



## 2. Problem statement

Consider a power-intensive continuously operated plant that manufactures a set of products. To satisfy demand, the products can be produced in the plant or purchased at a higher expense. The plant consumes electricity whose varying price is assumed to be known over the scheduling horizon. In addition, the plant can cut costs by providing interruptible load to the grid. To this end, the plant makes a commitment to reduce its electricity consumption at the grid operator's request. Load reduction may not always be requested but the plant earns revenue regardless. We assume that no additional payment is made to the plant when load reduction is requested; this assumption can be tweaked depending on the electricity market with no major changes to the model.

Given the plant model parameters (as discussed in Section 3) and a scheduling horizon, the goal is to determine the production schedule that is feasible for all possible realizations of load reduction while minimizing the total operating cost. Note that recourse variables, e.g., production rates and purchase amounts, depend on the realization of the uncertainty.

## 3. Deterministic model formulation

The model presented here is based on formulations developed in previous works (Mitra et al., 2012; Zhang et al., 2015) and is a direct extension of the model used in Zhang et al. 2016. Hence, only a brief description of the various constraints is provided.

### 3.1. High-level formulation

The structure of the deterministic model can be broadly expressed as follows:

$$\text{minimize } \textit{Net operating cost} \quad (1a)$$

$$\text{subject to } \textit{Feasible regions of operating modes} \quad (1b)$$

$$\textit{Transition constraints} \quad (1c)$$

$$\textit{Mass balance constraints} \quad (1d)$$

$$\textit{Initial conditions} \quad (1e)$$

The objective function is the cost of electricity plus the cost of additional products purchased minus the revenue from providing interruptible load. We assume that the plant can operate in different operating modes whose *feasible regions* are given in the form of polytopes. One such example of a plant with two products, P1 and P2, is shown in Fig. 1. Here, Region 1 could denote the off mode where no products are produced. *Transition constraints* are used to enforce feasible mode switching. Fig. 2 shows an example of requirements from the transition constraints. For instance, the process shown needs to stay in the off mode for at least 8 h before it can switch to the startup mode. *Mass balance constraints* ensure demand is satisfied through production or purchase; the remaining products are stored. *Initial conditions* are required for the problem to be well-defined.

### 3.2. Interruptible load constraints

As mentioned previously, the grid operator can request load reduction (no greater than the agreed amount) from the IL provider and the plant needs to alter its production schedule to cater to that request. These alterations in production schedule are modeled by introducing variables of the form:

$$\overline{PD}_{mit} = \widehat{PD}_{mit} + \widetilde{PD}_{mit} \quad \forall m \in M, i \in I, t \in T, \quad (2a)$$

$$\bar{y}_{mt} = \hat{y}_{mt} + \tilde{y}_{mt} \quad \forall m \in M, t \in T, \quad (2b)$$

where  $T$  denotes the set of time periods and  $\overline{PD}_{mit}$  is the production rate of product  $i$  in mode  $m$  in time period  $t$ . The nominal production rate is denoted by  $\widehat{PD}_{mit}$ , and  $\widetilde{PD}_{mit}$  is the deviation from the nominal value when load reduction is requested. The binary variable  $\bar{y}_{mt}$  equals 1 if the plant operates in mode  $m$  in time period  $t$ ;  $\hat{y}_{mt}$  is its nominal

value while  $\tilde{y}_{mt}$  is a discrete recourse variable that can take the values -1, 0, or 1, depending on the amount of load reduction requested. The decrease in power consumption associated with the decrease in production or, in some cases, a process shutdown must be at least as much as the amount of load reduction requested ( $LR_t$ ). This is modeled as the following constraint:

$$\sum_m \delta_m \tilde{y}_{mt} + \sum_m \sum_i \gamma_{mi} \tilde{P}D_{mit} \leq -LR_t \quad \forall t \in T, \quad (3)$$

where we assume that the electricity consumption is a linear function of production rates with a constant  $\delta_m$  and coefficient  $\gamma_{mi}$  for the selected operating mode.

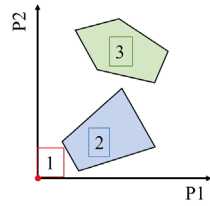


Fig. 1: Example for feasible operating regions of a process with three operating modes.

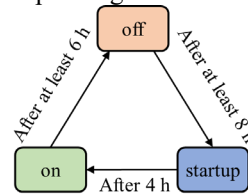


Fig. 2: Example for possible transitions between operating modes and corresponding operational constraints.

## 4. Multistage robust formulation with mixed-integer recourse

### 4.1. Uncertainty set

We adopt a “budget of uncertainty” approach (Bertsimas and Sim, 2004; Zhang et al., 2015) to formulate the following uncertainty set  $W$  for the load reduction demand:

$$W = \left\{ w \in \mathbb{R}^{|T|} : \left( 0 \leq w_k \leq 1 \quad \forall k = 1, \dots, t, \sum_{k=1}^t w_k = \Gamma_t \right) \quad \forall t \in T \right\}, \quad (4)$$

where  $w_t$  is the normalized requested load reduction, i.e.,  $LR_t = IL_t w_t$ ,  $IL_t$  is the amount of interruptible load provided, and  $\Gamma_t$  is a budget parameter limiting the cumulative load reduction required up to time  $t$ . The choice of  $\Gamma_t$  can be based on historical data or the electricity market rules. Note that  $\Gamma_t$  must be a monotonically increasing parameter. We now write Eq. (3) in terms of the normalized load reduction:

$$\sum_m \delta_m \tilde{y}_{mt}(w) + \sum_m \sum_i \gamma_{mi} \tilde{P}D_{mit}(w) \leq -w_t IL_t \quad \forall t \in T. \quad (5)$$

### 4.2. Adjustable robust formulation with mixed-integer decision rules

The overall adjustable robust optimization problem can be formulated as follows:

$$\begin{aligned} & \text{minimize} && \text{Net operating cost at } w = 0 \text{ or worst-case net operating cost} \\ & \text{subject to} && \text{Eqs. (1b)–(1e), (2), (5)} \quad \forall w \in W, \end{aligned} \quad (6)$$

where we either minimize the net operating cost in the nominal or in the worst case. Here, the deviation variables  $\tilde{P}D_{mit}$ ,  $\tilde{y}_{mt}$ , etc., serve as the recourse variables and are hence functions of the uncertain parameters  $w$ . Following the concept of lifted uncertainty (Bertsimas and Georghiou, 2018) and the derivation for multistage problems in Feng et al. 2021, we implement (potentially discontinuous) piecewise linear decisions rules for the continuous variables and piecewise constant decision rules for the binary variables. As an example, the decision rules for  $\tilde{P}D_{mit}$  and  $\tilde{y}_{mt}$  have the following form:

$$\widetilde{PD}_{mit} = \sum_{t'=t-\bar{\zeta}}^t \sum_{k=1}^{K_{t'}} (\bar{X}_{mit'k}^t \bar{w}_{t'k} + \hat{X}_{mit'k}^t \hat{w}_{t'k}) \quad \forall m \in M, i \in I, t \in T, \quad (7a)$$

$$\tilde{Y}_{mt} = \sum_{t'=t-\bar{\zeta}}^t \sum_{k=1}^{K_{t'}} \hat{Y}_{mt'k}^t \hat{w}_{t'k} \quad \forall m \in M, t \in T, \quad (7b)$$

where  $\bar{w}$  and  $\hat{w}$  are the auxiliary lifted uncertain parameters,  $K_t$  is the number of breakpoints associated with  $w_t$  that define the piecewise structures of the decision rules, and  $\bar{\zeta}$  determines how many uncertain parameters from previous time periods are considered in the decision rules. The coefficients  $\bar{X}$ ,  $\hat{X}$ , and  $\hat{Y}$  define the decision rules and are to be optimized. Problem (6) is a semi-infinite program, which we solve using the reformulation approach that leverages linear programming duality (Yanikoğlu et al., 2019). For the sake of brevity, we refer the reader to Feng et al., 2021 for more details including the full reformulation, which results in a mixed-integer linear program (MILP).

## 5. Case study

In this section, the proposed robust model with mixed-integer recourse is applied to an illustrative example. All models were implemented in Julia v1.7 using the modeling environment JuMP v0.22.3 and solved to 1% optimality gap using Gurobi v9.5.1 on an Intel Core i7-8700 machine at 3.20 GHz with 8 GB RAM.

In this case study, a single-product plant is considered, and the scheduling problem is solved over a time horizon of 48 hours with hourly time discretization. The plant can operate in three different modes: off, startup, and on. Table 1 shows the details of the polytopes (which here are simple ranges) that characterize the operating modes and the electricity consumption in each operating mode. Table 2 shows the possible mode transitions and the respective minimum stay times. Fig. 3 shows the electricity prices and revenue from providing interruptible load over the time horizon. The plant is operating in the on mode at the start and it is assumed that no mode switching has occurred in the eight time periods prior to the beginning of the scheduling horizon.

The initial inventory is 1,000 kg. The minimum and maximum inventory levels are 0 and 5,000 kg, respectively, for all time points. At the end of the time horizon, the minimum inventory level is set to 1,000 kg. The cost of purchasing additional products is \$3/kg. If interruptible load is provided in a time period, then the provided amount must be between 200 and 5,000 kWh. We assume that the budget parameter  $\Gamma_t$  increases every 8 time periods by 1, i.e., maximum load reduction can only be requested once during the first 8 h, twice during the first 16 h, etc., and at most six times during all 48 h.

**Table 1:** Polytope equations, fixed ( $\delta_m$ ) and unit ( $\gamma_m$ ) electricity consumption for each operating mode (product indices have been omitted).

Operating mode	Polytope	$\delta_m$ [kWh]	$\gamma_m$ [kWh/kg]
Off	$0 \leq PD_{mt} \leq 0$	0	0
Startup	$5 \leq PD_{mt} \leq 5$	0	60
On	$100 \leq PD_{mt} \leq 160$	1,200	20

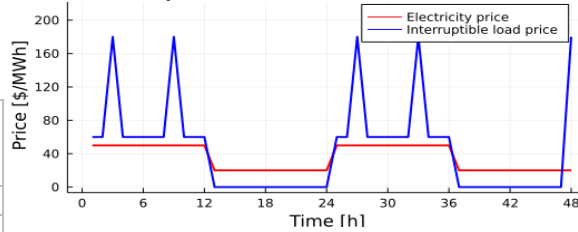
### 5.1. Advantages of discrete recourse

We solve the problem for different  $\bar{\zeta}$ , which controls the amount of past information used in the decision rules. The flexibility in the solution increases with  $\bar{\zeta}$ , but it also increases the model size and hence the computational intensity. Table 3 compares the costs between the continuous recourse only case and the mixed-integer recourse case. The results show

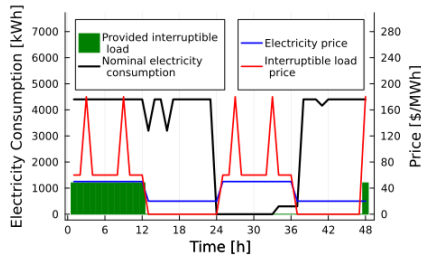
significant cost reductions, of up to 14%, for mixed-integer recourse case over continuous recourse case. In Figs. 4 and 5, we see that the nominal electricity consumption profiles in both cases are very similar. This implies that a major factor in the cost reduction is the additional interruptible load provided in the mixed-integer recourse case. Figs. 5 and 7 show that this additional interruptible load is provided in time periods 27 and 48, where process shutdown is a feasible recourse action. This is also reflected in the time period 48 of Fig. 7, where the nominal and recourse production amounts add up to zero, indicating a process shutdown in the worst-case uncertainty realization.

**Table 2:** Possible transitions between operating modes and minimum stay times.

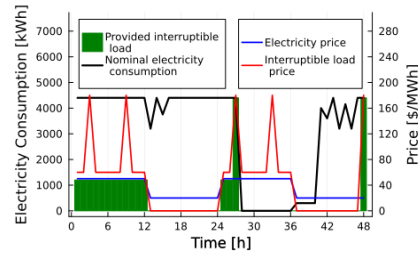
Transition from mode $m$ to mode $m'$	Minimum stay time in $m'$ [h]
Off $\rightarrow$ startup	4
Startup $\rightarrow$ on	6
On $\rightarrow$ off	8



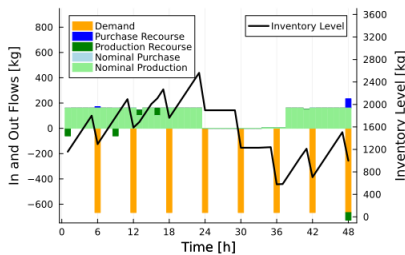
**Fig. 3:** Electricity and interruptible load prices for case study



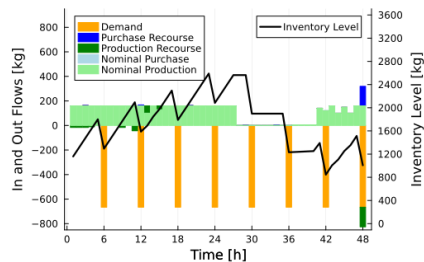
**Fig. 4:** Nominal electricity consumption and interruptible load provided with continuous recourse only and  $\bar{\zeta} = 47$ .



**Fig. 5:** Nominal electricity consumption and interruptible load provided with mixed-integer recourse and  $\bar{\zeta} = 47$ .



**Fig. 6:** Nominal and worst-case recourse product flows and nominal inventory profile for the case with continuous recourse only and  $\bar{\zeta} = 47$ .



**Fig. 7:** Nominal and worst-case recourse product flows and nominal inventory profile for the case with mixed-integer recourse and  $\bar{\zeta} = 47$ .

**Table 3:** Net operating cost ( $C$ ) in \$ for cases with different  $\bar{\zeta}$ . Here,  $C^{\text{nominal}}$  is the cost when no load reduction is requested and  $C^{\text{wc}}$  is the worst-case cost when load reduction is requested.

$\bar{\zeta}$	Continuous recourse only ( $C^{\text{nominal}}$ , \$)	Mixed-integer recourse ( $C^{\text{nominal}}$ , \$)	Continuous recourse only ( $C^{\text{wc}}$ , \$)	Mixed-integer recourse ( $C^{\text{wc}}$ , \$)
47	3,214.1	2,755.3	3,266.3	2,917.8

23	3,214.1	2,755.3	3,266.3	2,917.8
11	3,214.1	2,810.4	3,282.7	2,999.8

### 5.2. Computation time

The computation times required to solve the models are shown in Table 4. As expected, considering mixed-integer recourse is computationally significantly more expensive. Adjusting the parameter  $\bar{\zeta}$  can help reduce the solution time but may give a solution with higher objective value. Clearly, a trade-off exists between computational performance and solution quality. In the example considered, setting  $\bar{\zeta} = 23$  provides a solution of the same quality as  $\bar{\zeta} = 47$  in much shorter time.

**Table 4:** Computation times for cases with different  $\bar{\zeta}$ .

$\bar{\zeta}$	Continuous recourse only	Mixed-integer recourse
47	90 s	3,626 s
23	32 s	1,228 s
11	10 s	253 s

## 6. Conclusion

In this work, we developed a multistage robust optimization model for the scheduling of power-intensive plants that also participate in the interruptible load market. Piecewise linear/constant decision rules are used to determine the recourse actions necessary in both continuous and discrete variables. When applied to an illustrative example, the proposed model achieves significant cost savings compared to the formulation that only considers continuous recourse. The flexibility that mixed-integer recourse provides comes at a substantial computational cost. However, our model provides a way of exploring this trade-off by setting the amount of past information allowed to be considered in the decision rules. By doing so, we find that the problem can often be solved in much less time with little sacrifice on the solution quality.

## 7. Acknowledgements

The authors gratefully acknowledge the financial support from the National Science Foundation under Grant No. 2215526.

## References

- Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. (2009). Robust optimization (Vol. 28). *Princeton University Press*.
- Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations Research*, 52(1), 35-53.
- Dowling, A.W., Kumar, R., and Zavala, V. M. 2017. A multi-scale optimization framework for electricity market participation. *Applied Energy*, 190, 147-164.
- Feng, W., Feng, Y., and Zhang, Q., 2021. Multistage robust mixed-integer optimization under endogenous uncertainty. *European Journal of Operational Research*, 294(2), 460-475.
- Mitra, S., Grossmann, I.E., Pinto, J.M., & Arora, N. (2012). Optimal production planning under time-sensitive electricity prices for continuous power-intensive processes. *Computers & Chemical Engineering*, 38, 171-184.
- Yanikoğlu, İ., Gorissen, B. L., & den Hertog, D. (2019). A survey of adjustable robust optimization. *European Journal of Operational Research*, 277(3), 799-813.
- Zhang, Q. and Grossmann, I.E., 2016. Enterprise-wide optimization for industrial demand side management: Fundamentals, advances, and perspectives. *Chemical Engineering Research & Design*, 116, 114-131.

- Zhang, Q., Grossmann, I.E., Heuberger, C.F., Sundaramoorthy, A. and Pinto, J.M., 2015. Air separation with cryogenic energy storage: optimal scheduling considering electric energy and reserve markets. *AIChE Journal*, *61*(5), 1547-1558.
- Zhang, Q., Morari, M.F., Grossmann, I.E., Sundaramoorthy, A., and Pinto, J.M., 2016. An adjustable robust optimization approach to scheduling of continuous industrial processes providing interruptible load. *Computers & Chemical Engineering*, *86*, 106-119.



# Technoeconomic Assessment of Digestate Valorization for Biogas Plant

Shivom Sharma<sup>a</sup>, Rafael Graf<sup>b</sup>, Jaroslav Hemrle<sup>b</sup>, Adrian Schneider<sup>b</sup>,  
Francois Maréchal<sup>a</sup>

<sup>a</sup>*IPESE, EPFL Valais Wallis, 1950 Sion, Switzerland*

<sup>b</sup>*Hitachi Zosen Inova AG, Zürich, 8005, Switzerland*

*shivom.sharma@epfl.ch*

## Abstract

Anaerobic digestion (i.e., biogas plant) converts organic waste into biogas that can directly be used or upgraded to bio-methane for natural gas grid injection. The digestate (residues) from biogas plant has nutrients and unconverted organic matters. The dry digestate can be used to produce hydrogen and methane *via* thermochemical route. This study has modelled a complete process or value chain for converting digestate into methane and hydrogen. The developed process model includes gasification, water gas shift and methanation as the main conversion units, whereas pressure swing adsorption, membrane separation and amine absorption were used to separate different gaseous mixtures. This study focuses on systematic generation and comparison of digestate valorization options. The formulated optimization problem is a mixed integer linear programming problem, and total annual cost is considered as the performance criterion. Several economic scenarios have been generated based on selling prices of hydrogen and methane, and the optimization problem was solved repeatedly for each scenario. The obtained solutions are divided into three categories: only producing methane, producing methane and hydrogen as main products, and producing hydrogen as main product and methane as side product. As expected, the selection of valorization route and relative production of hydrogen and methane depend upon their selling prices. Finally, use of biogenic carbon dioxide for storing excess renewable electricity has been explored for selected scenario.

**Keywords:** Biogas Plant; Digestate Valorization; Hydrogen Production; Methane Production; Carbon Capture and Utilization.

## 1. Introduction

Anaerobic digestion is a process to convert organic matters (e.g., food waste, sewage sludge, agriculture waste, etc.) into biogas. The composition of biogas varies for different feed-stocks. Typically, 50-80 v/v% methane, 20-50 v/v% carbon dioxide, and small quantities of hydrogen, nitrogen, oxygen and carbon dioxide are present in biogas (Chen et al., 2015). The biogas can be upgraded into bio-methane by separating carbon dioxide and other impurities. The biogas and/or bio-methane can be used as fuel to produce heat and electricity. The residues from biogas plants are called digestate that contain valuable nutrients and unconverted organic matters. The water content of digestate can be as high as 90% (Ehmann et al., 2018), and digestate is dried to achieve 10-20% water content before further usage (Dziedzic et al., 2021). The lower heating value of dry digestate varies between 15 and 20 MJ/kg. The digestate can be used as organic fertilizer, soil conditioner or heat source. In European Union (EU28) about 180 million tonnes of



digestate is produced annually (Corden et al., 2019). In literature, several studies have focused on the conventional use of digestate as fertilizer and soil conditioner. However, there are limited number of studies that have focused on the valorization of digestate for producing liquid and gaseous fuels/products (e.g., methane, oil, etc.). Wisniewski et al. (2015) studied pyrolysis and gasification of digestate from agricultural biogas plant. Chen et al. (2017) studied gasification of digestate in a downdraft fixed bed gasifier. Castro-Amoedo et al. (2021) studied biowaste valorization along with CO<sub>2</sub> removal from biogas, so that produced bio-methane can be injected into natural gas grid for transportation purpose. Renewable energy requires short- and long-term storage due to its variable or intermittent nature. In case of biogas upgradation and digestate valorization into liquid and gaseous fuels, CO<sub>2</sub> is produced as a side product that can be used for storing excess renewable energy. Recently, Wang and Lee (2021) reviewed different valorization options for digestate of different origins (e.g., agriculture waste, food waste, municipal solid waste etc.).

This study focuses on detailed evaluation of high value fuel (CH<sub>4</sub>, H<sub>2</sub>) production from digestate. Figure 1 presents a simplified flowchart of digestate conversion into methane and hydrogen. The digestate contains 62 wt.% water and 38 wt.% solid materials. The solid material has 79 wt.% organic matters, and remaining inert. The organic materials are broken down into small gaseous molecules (CH<sub>4</sub>, H<sub>2</sub>, CO and CO<sub>2</sub>) and char inside the gasifier. The product gases from gasifier are processed in different conversion and separation units to produce high quality methane and hydrogen. This study has modelled a complete process or value chain for converting digestate into methane, hydrogen and carbon dioxide. The developed process model includes digestate pre-processing, conversion and gas separation units. The digestate value chain has the possibility to convert the captured CO<sub>2</sub> into additional methane, using renewable electricity. The formulated optimization problem is a mixed integer linear programming problem (MILP) that is implemented in AMPL (A Mathematical Programming Language). This study systematically generates and compares different scenarios for digestate valorization, based on the product prices. The obtained solutions are compared based on total annual cost (TAC), and they provide more insights on digestate valorization options.

## 2. Digestate Valorization Model

### 2.1 Process Description

The digestate from biogas plant contains valuable nutrients, unconverted organic matters, and significant amount of water, and so it is dried to about 20% moisture content before gasification. After the gasification, the product gases are cleaned and cooled to remove solid particles, sulphur and other impurities. The clean gases are handled differently for methane (dotted box X, Figure 1) and hydrogen productions (dotted box Y, Figure 1). In case of methane production, water gas shift (WGS) reactor is used to convert part of the carbon mono-oxide into hydrogen. This is required to ensure adequate carbon to hydrogen ratio for the subsequent methane synthesis. The carbon dioxide is removed using pressure swing adsorption (PSA) after WGS reactor. After methanation reactor, product gases are dried and upgraded (*via* membrane unit) for natural gas grid injection. The methane production route has some off-gases that are either recycled to upstream units or burnt in a burner to supply the process heat. In case of hydrogen production, WGS reactor is used to convert most of the carbon mono-oxide into hydrogen. The product gases include mainly methane, hydrogen and carbon dioxide. The carbon dioxide is separated from the gaseous mixture using amine absorption process, followed by methane and hydrogen separation using a membrane unit.



$$\min TAC (\$/y) = \sum_{t \in T} \sum_{i \in L} (c_{i,t}^{op1} \cdot \psi_{i,t} + c_{i,t}^{op2} \cdot f_{i,t}) \cdot t_t^{op} + \frac{i(1+i)^n}{(1+i)^n - 1} \cdot \sum_{i \in Lu} (c_i^{inv1} \cdot \psi_i + c_i^{inv2} \cdot f_i) \quad (1)$$

$$f_i^{min} \cdot \psi_{i,t} \leq f_{i,t} \leq f_i^{max} \cdot \psi_{i,t}, \quad \forall i \in L, \quad \forall t \in T \quad (2)$$

$$M_{r,i,t}^+ = f_{i,t} \cdot \dot{m}_{r,i,t}^+, \quad \forall r \in R, \quad \forall i \in L, \quad \forall t \in T \quad (3)$$

$$M_{r,i,t}^- = f_{i,t} \cdot \dot{m}_{r,i,t}^-, \quad \forall r \in R, \quad \forall i \in L, \quad \forall t \in T \quad (4)$$

$$\sum_{i \in L} M_{r,i,t}^- = \sum_{i \in L} M_{r,i,t}^+, \quad \forall r \in R, \quad \forall t \in T \quad (5)$$

$$\sum_{i \in L} \dot{Q}_{i,t,k} \cdot f_{i,t} + \dot{R}_{t,k+1} - \dot{R}_{t,k} = 0, \quad \forall t \in T, \quad \forall k \in K, \text{ with } T_{k+1} \geq T_k, \quad (6)$$

$$\dot{R}_{t,k} \geq 0, \quad \dot{R}_{t,k_{max}+1} = \dot{R}_{t,1} = 0, \quad \forall t \in T \quad (7)$$

### 3. Digestate Valorization: A Case Study

For this case study, we considered that 5 t/h digestate is available for valorization. The purchase prices of digestate, water and electricity are 0.01 \$/kg, 0.0005 \$/kg and 0.15 \$/kWh, respectively. The wastewater treatment cost is 0.001 \$/kg. The carbon dioxide storage cost is 20 \$/t. The relative production of methane and hydrogen depends on their selling prices. The optimization method optimizes distribution of digestate carbon between methane and carbon dioxide. The selling prices of methane (synthetic natural gas, SNG) and hydrogen are 1-1.5 \$/kg and 3-5 \$/kg, respectively. In order to explore digestate valorization options, we have generated 9 scenarios, by using three different selling prices for methane (1, 1.25 and 1.5 \$/kg) and hydrogen (3, 4, and 5 \$/kg).

**Table 1:** Digestate Valorization Scenarios based on Selling Prices of CH<sub>4</sub> and H<sub>2</sub> (Elec: Electricity; WW: Wastewater; S: Solution)

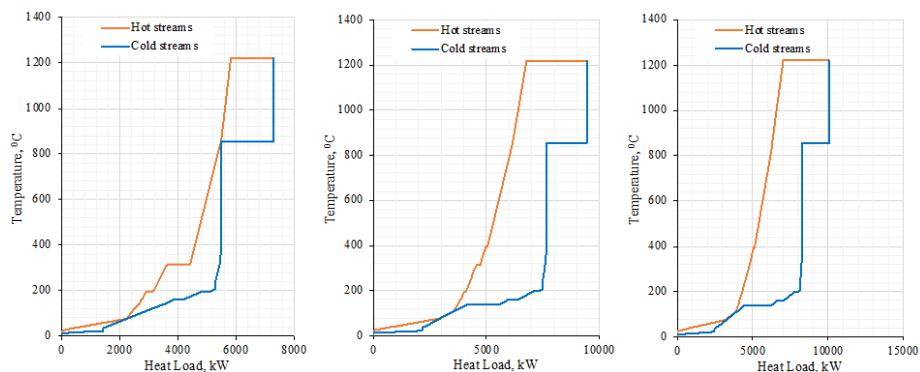
H <sub>2</sub> \$/kg	CH <sub>4</sub> \$/kg	H <sub>2</sub> O kg/h	Elec kWh	H <sub>2</sub> kg/h	CH <sub>4</sub> kg/h	CO <sub>2</sub> kg/h	WW kg/h	CAPEX 10 <sup>3</sup> .\$/y	OPEX 10 <sup>3</sup> .\$/y	S. No
3	1	876.7	602.7		398	1040.6	912.8	2258	-1998	A:1
3	1.25	876.7	602.7		398	1040.6	912.8	2258	-2870	A:2
3	1.5	876.7	602.7		398	1040.6	912.8	2258	-3741	A:3
<b>4</b>	<b>1</b>	<b>1160.4</b>	<b>522.2</b>	<b>62.4</b>	<b>104.4</b>	<b>950.3</b>	<b>764.7</b>	<b>1666</b>	<b>-1687</b>	<b>C:1</b>
4	1.25	876.7	602.7		398	1040.6	912.8	2258	-2870	A:4
4	1.5	876.7	602.7		398	1040.6	912.8	2258	-3741	A:5
<b>5</b>	<b>1</b>	<b>1160.4</b>	<b>522.2</b>	<b>62.4</b>	<b>104.4</b>	<b>950.3</b>	<b>764.7</b>	<b>1666</b>	<b>-2233</b>	<b>C:2</b>
<b>5</b>	<b>1.15</b>	<b>1111.3</b>	<b>543.7</b>	<b>49.8</b>	<b>170.5</b>	<b>986.5</b>	<b>810.3</b>	<b>1807</b>	<b>-2613</b>	<b>B:1</b>
5	1.5	876.7	602.7		398	1040.6	912.8	2258	-3741	A:6

In Table 1, the obtained solutions are divided into three categories based on the selection of valorization route and relative production of products. Solution A only produces methane, without using technologies in “dotted box Y” (Figure 1). Solution B uses all technologies in Figure 1, and produces both methane and hydrogen. Finally, Solution C does not use technologies in “dotted box X” (Figure 1), and produces hydrogen as the main product and methane as the side product. For Solutions A, B and C, annual capital (CAPEX) and operating (OPEX) costs are shown in Table 1, and all major flows are presented in Table 2. Finally, Figure 2 shows hot and cold composite curves for different solutions. The solid carbon and part of the product gases from gasifier were used in burner

to supply the process heat. Solution A produces maximum amount of CO<sub>2</sub> (1040.6 kg/h), as it does not produce any hydrogen. Solutions B and C have 986.5 kg/h and 950.3 kg/h CO<sub>2</sub>, respectively.

**Table 2:** Flow Details (kg/h) for Solutions A, B and C (Stream Numbers are Shown in Figure 1; \*solid carbon flow and +product gases flow)

Stream	Stream Details	Solution A	Solution B	Solution C
1	Digestate to Grinder	5000	5000	5000
2	Grinder to Air dryer	5000	5000	5000
3	Air dryer to Gasification	2037.5	2037.5	2037.5
4	Gasification to Cleaning	1882.5	1317.8	1134.2
5	Cleaning to Compressor	1562.3	1093.4	941.1
6	Compressor to WGS (1)	1562.3	342.2	
7	WGS (1) to PSA	1673.2	366.5	
8	PSA to Methanation	632.3	138.5	
9	Methanation to Dryer	1027	225	
10	Dryer to Membrane (1)	434.7	95.2	
11	Water to Boiler	876.7	1111.3	1160.4
12	Boiler to Gasification	407.5	407.5	407.5
13	Boiler to WGS (1)	110.9	24.3	
14	Boiler to Methanation	358.3	78.5	
15	Gasification to Burner	77.3*, 489.7 <sup>+</sup>	77.3*, 1054.5 <sup>+</sup>	77.3*, 1238.1 <sup>+</sup>
16	Cleaning to Wastewater	320.5	224.3	193.1
17	PSA to CO <sub>2</sub>	1040.6	228	
18	Dryer to Wastewater	592.3	129.7	
19	Membrane (1) to CH <sub>4</sub>	398	87.2	
21	WGS (2) to Wastewater		456.3	571.6
22	Compressor to WGS (2)		751.2	941.1
23	MEA to CO <sub>2</sub>		758.6	950.3
24	Membrane (2) to H <sub>2</sub>		49.8	62.4
25	Membrane (2) to CH <sub>4</sub>		83.3	104.4
26	Boiler to WGS (2)		601	752.9
27	WGS (2) to MEA		896	1122.4
28	MEA to Membrane (2)		142.2	178.1



**Figure 2:** Hot and Cold Composite Curves for Solutions A, B and C (Left to Right)

### 3.1 Integration of Power to Methane System

To explore the potential of power to methane concept, three scenarios were selected for further analysis: (1)  $H_2 = 3$  \$/kg,  $CH_4 = 1.25$  \$/kg (Solution A:2), (2)  $H_2 = 5$  \$/kg,  $CH_4 = 1.25$  \$/kg (Solution B:1), and (3)  $H_2 = 5$  \$/kg,  $CH_4 = 1$  \$/kg (Solution C:2). For this analysis, electricity price was reduced to 0.1 \$/kWh, carbon credit (96 \$/t- $CO_2$ ) was used for converting  $CO_2$  into green  $CH_4$ , and no investment was considered for power to methane system. Table 3 presents important results for integrated system. The integration of power to methane system with digestate valorization provides better heat integration possibilities. For scenarios 2 and 3, flow of gases from gasification to burner has reduced (489.7 and 986.7 kg/h in Table 3 compared to 1054.5 and 1238.1 kg/h in Table 2). This increases  $CH_4$  production for scenarios 2 and 3. Further,  $O_2$  is also produced from electrolyser that may be used for oxy-combustion (burner) to produce  $CO_2$  and  $H_2O$ .

**Table 3:** Results for Integration of Power to Methane System with Digestate Valorization (EC: Electrolysis, G: Gasifier, B: Burner)

Scenario No.	CAPEX 10 <sup>3</sup> .\$/y	OPEX 10 <sup>3</sup> .\$/y	CH <sub>4</sub> kg/h	H <sub>2</sub> kg/h	Green CH <sub>4</sub> kg/h	O <sub>2</sub> kg/h	EC Elec, kWh	G to B kg/h
1	2258	-2560	398		378.4	1513.6	7881.5	489.7
2	2177	-3314	<b>297.1</b>	46.5	466	1864.1	9706.2	<b>489.7</b>
3	1805	-1661	<b>127.5</b>	76.2	422.2	1688.6	8792.8	<b>986.7</b>

## 4. Conclusions

This study explores digestate valorization that comprises gasification, water gas shift and methanation as the main conversion units. Pressure swing adsorption, amine absorption and membrane separation were considered for separating gaseous mixtures of carbon dioxide, methane and hydrogen. A mixed integer linear programming problem was formulated for digestate value chain, and total annual cost was used as objective function. In total, eight scenarios were generated based on the selling prices of hydrogen and methane. The obtained solutions were categorized into three types, based on the selection of valorization route/technologies and relative production of final products. The power to methane system has good synergies with digestate valorization into methane/hydrogen.

## References

- X. Y. Chen, H. Vinh-Thang, A. A. Ramirez, D. Rodrigue, S. Kaliaguine, 2015, Membrane gas separation technologies for biogas upgrading, *RSC Adv*, 5, pp. 24399-24448.
- A. Ehmann, U. Thumm, I. Lewandowski, 2018, Fertilizing potential of separated biogas digestates in annual and perennial biomass production systems, *Frontiers in Sustainable Food System*, 2 (12), pp. 1-14.
- K. Dziedzic, et al., 2021, Solid digestate - Physicochemical and thermal study, *Energies*, 14, 7224.
- C. Corden, et al., Digestate and compost as fertilisers: Risk assessment and risk management options, European Commission, 2019.
- D. Wisniewski, J. Gołaszewski, A. Białowiec, 2015, 3 The pyrolysis and gasification of digestate from agricultural biogas plant, *Archives of Environmental Protection*, 41 (3), pp. 70-75.
- G. Chen, et al., 2017, Air gasification of biogas-derived digestate in a downdraft fixed bed gasifier, *Waste Management*, 69, pp. 162-169.
- R. Castro-Amoedo, J. Morisod, J. Granacher, F. Maréchal F., 2021, The role of biowaste: a multi-objective optimization platform for combined heat, power and fuel, *Front. Energy Res.* 9:718310.
- W. Wang and D. J. Lee, 2021, Valorization of anaerobic digestion digestate: A prospect review, *Bioresource Technology* 323:124626.
- F. Maréchal and Kalitventzeff, B., 1998, Energy integration of industrial sites: tools, methodology and application, *Appl. Therm. Eng.*, 18, pp 921-933.

# Optimization of Hydrogen systems for prospective Life Cycle Assessment: Well-to-Tank approach

Rocky Mashi<sup>a</sup>, Yohan Vincotte<sup>a</sup>, Sofía De-León Almaraz<sup>b</sup>, Catherine Azzaro-Pantel<sup>a,\*</sup>

<sup>a</sup>*Laboratoire de Génie Chimique, Université Toulouse, CNRS, INPT, UPS, Toulouse France*

<sup>b</sup>*Corvinus University of Budapest, Department of Supply Chain Management, 8 Fővám tér. 1093 Budapest, Hungary*

\* *catherine.azzaropantel@toulouse-inp.fr*

## Abstract

The optimization of hydrogen energy systems generally focuses on minimizing total cost and greenhouse gas (GHG) emissions. While climate neutrality by 2050 is the objective set by the European Green Deal, with a target on GHG emissions, other environmental impacts such as resource demand and land use are also of great importance and are mostly not taken into account in these models. This study uses the Life Cycle Assessment (LCA) method to analyze 18 Key Environmental Indicators (KEI). The framework combines a multi-objective design strategy for hydrogen networks with LCA evaluation. The optimization strategy uses a Mixed Integer Linear Programming (MILP) approach with an augmented epsilon-constraint, implemented in the GAMS environment. The Pareto solutions for hydrogen supply chains obtained from optimization were analyzed using the SimaPro Power user package, the Ecoinvent database, and the ReCiPe2016 method with a Well-to-Tank approach. The results showed that considering only GHG emissions leads to poor results for traditional solutions like SMR, and low performance in other KEIs such as water consumption and mineral resource scarcity for the electrolysis process using PV. These findings emphasize the importance of multi-objective optimization that considers multiple environmental criteria, not just GHG emissions.

**Keywords:** Hydrogen Systems, Modelling, Optimization, Prospective Life Cycle Assessment.

## 1. Introduction

Clean hydrogen (H<sub>2</sub>) from renewable or low-carbon sources is seen as a key solution in achieving net-zero emissions by 2050 and limiting global warming to 1.5 degrees Celsius. This offers a long-term, scalable, and cost-effective option in decarbonizing hard-to-abate sectors (IRENA, 2022). Optimizing hydrogen energy systems thus provides effective support for policy decision-making by prioritizing the minimization of both cost and environmental impact. Although the European Green Deal aims for climate neutrality by 2050, the current optimization of these systems (Li *et al.*, 2019) primarily addresses carbon emissions and does not fully consider other environmental impacts such as resource use and land utilization. This raises various methodological questions, such as expanding the assessment of carbon to a comprehensive, multi-criteria evaluation of environmental impact.

Decision support tools, particularly Life Cycle Assessment (LCA), are essential for guiding stakeholders in their strategic choices and identifying potential ways to reduce environmental impacts. The hydrogen supply chain is here the studied system, covering all stages from resource extraction to production, storage, transportation, and use. LCA, as a widely accepted systemic environmental assessment method, is the best approach for these challenges. Although traditional attributional LCA describes the system and its impacts by assigning a portion of the technosphere's impacts, it is inadequate for capturing the changing emission factors, such as technology advancements over time.

With growing interest in hydrogen, many review papers on hydrogen LCAs have been published, such as (Rinawati *et al.*, 2022). The scientific objective of this paper is to explore the synergies between LCA, energy system optimization, and multi-criteria decision support approaches in selecting a sustainable "hydrogen" system. However, various hydrogen pathways to achieve similar GHG emission reductions may differ in terms of both costs and environmental impacts, which raises several issues such as: which ecological co-benefits and negative side effects can be expected from a climate-friendly hydrogen energy system? How do the various hydrogen supply chains differ in terms of other ecological impacts? The novelty of this work is thus to propose a methodological framework that can aid in selecting a hydrogen supply chain that incorporates a set of environmental criteria.

The paper is organized as follows: the solution approach is developed in Section 2. The main assumptions used for the case study, i.e. the deployment of the hydrogen in Occitanie, are presented in Section 3. Section 4 discusses and highlights some significant results. Finally, some conclusions and perspectives for future work are proposed in Section 5.

## 2. Methods and tools

The hydrogen supply chain for mobility purposes encompasses production, liquefaction, transportation, storage, and delivery at hydrogen refueling stations (HRS). The production methods include two types of electrolysis processes - Alkaline Electrolyser Cell (AEC) and Proton Exchange Membrane (PEM) - utilizing renewable energy sources such as wind, hydro, and solar photovoltaic, as well as grid electricity. Other hydrogen production methods considered are steam methane reforming (SMR) and steam methane reforming with carbon capture, utilization, and storage (CCUS). The possible hydrogen pathways are depicted in Figure 1.

The methodological framework adopts a 3-tier approach.

- Firstly, the design and operation of a future Hydrogen Supply Chain (HSC) is considered with a geographical and multi-period (6 periods of 5 years each, from 2020 to 2050) approach to meet the hydrogen demand profile of a specific geographic region. The optimization problem is modeled as a Mixed Integer Linear Programming (MILP) using the GAMS® modeling system and CPLEX 12 solver. The comprehensive formulation of the model includes constraints (such as demand satisfaction, mass balance, production, transportation, and storage constraints, primary energy source availability for electricity production) and objective functions (presented in De-León Almaraz *et al.*, 2015). Three optimization objectives are considered: an economic objective based on minimizing the total network cost, an environmental objective related to the global greenhouse gas

(GHG) emissions of the supply chain, and a risk index. The multi-objective optimization is formulated using an augmented epsilon-constraint approach and generates the Pareto front of optimal HSC configurations as outputs.

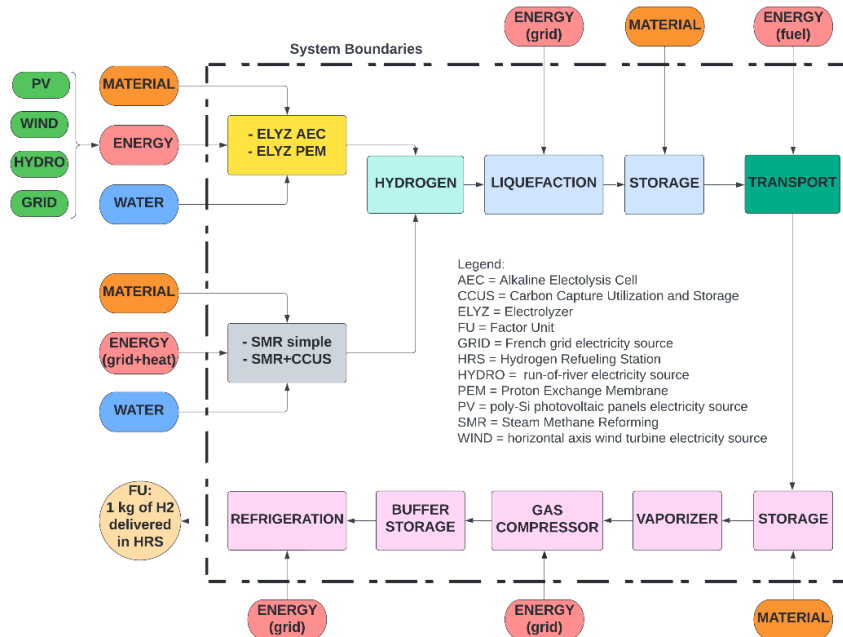


Figure 1 – System study and associated LCA System Boundaries

- Second, the Pareto solutions were evaluated from an environmental perspective using the Life Cycle Assessment (LCA) method. This method was selected for its systemic approach, international standardization, and use of inventory databases such as EcoInvent. The LCA method requires the definition of a functional unit (FU), the implementation of life cycle inventory (LCI), and the life cycle impact assessment (LCIA). The LCA system boundaries for the hydrogen supply chain are shown in Figure 1. The functional unit (FU) was set as 1 kg of hydrogen delivered at the Hydrogen Refueling Station (HRS) using a Well-to-Tank approach. The Life Cycle Inventory (LCI) was supported by the Ecoinvent database in SimaPro Power. Data collection for materials, energy and water use was also been carried out for specific components of the hydrogen supply chain (i.e., electrolyzers). The design of specific modules dedicated to inventory production for the different hydrogen production methods are based on the data provided in (Spath *et al.*, 2000); (Koj *et al.*, 2017); (Bareiß *et al.*, 2019). The ReCiPe2016 midpoint (H) method was used as the LCIA, and 18 Key Environmental Indicators (KEI) were calculated for each configuration of the hydrogen supply chain. A parametric approach was used to automatically generate the LCA of the Pareto solutions using GAMS and SimaPro Power.
- Third, the modified-TOPSIS (m-TOPSIS), a variant of the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) is applied to choose the best compromise configuration from the Pareto front.



### 3. Case study

The methodology was applied to study hydrogen deployment in the Occitanie region of France with the aim of decarbonizing the mobility sector (passenger cars, light commercial vehicles, buses, and heavy goods vehicles). The data used was updated from previous studies (REPOS, 2017) and the demand estimation was based on the pluriannual plan of the minister of ecological transition in France. Two scenarios were analyzed: the "Base Scenario" which allows for hydrogen production using all technologies (electrolysis and SMR with and without CCS), and the "e-H<sub>2</sub> Scenario" which only allows for hydrogen production through electrolysis technologies (green hydrogen and grid make-up if necessary). The methodology was applied in two variants for each scenario. Variant #1 considered Life Cycle Assessment (LCA) after the multi-criteria decision analysis, while variant #2 considered LCA after the multi-objective optimization, as can be seen in Figure 2 and Figure 3 respectively. The Pareto front consisted of 21 optimal compromise configurations.



Figure 2 – Methodological framework (Variant #1)



Figure 3– Methodological framework (Variant #2)

For the application of the multi-criteria decision method (TOPSIS), an equal weight was assigned to each group of criteria: 1/3 for the cost, risk, and environmental parameters. In variant #1, 1/3 weight was given to the greenhouse gas (GHG) criteria. In variant #2, a weight of 1,852e-2 was given to each Key Environmental Indicator (KEI).

### 4. Results

Table 1 reports the quantitative results of all criteria obtained after conducting a Life Cycle Assessment (LCA) for all variants. The study found that it is feasible to design a hydrogen ecosystem using only renewable energy-powered electrolyzers, with grid support, in line with the region's decarbonization goals. However, the average Levelized Cost of Hydrogen (LCOH) over the study period was 10.22/10.28 €/kgH<sub>2</sub>, with an average emission of 4.51/4.82 kgCO<sub>2</sub>eq/kgH<sub>2</sub>. The optimization of all technologies resulted in the selection of the SMR+CCS version, with a lower LCOH of 5.15/5.34 €/kgH<sub>2</sub> and an average GHG emission of 6.92 kgCO<sub>2</sub>eq/kgH<sub>2</sub>. The risk index was higher for the SMR version, which is primarily due to its larger plant size. A radar graph (Figure 4) is also provided to analyze the differences between the solutions obtained.

The order of application of the Life Cycle Assessment (LCA) method and the choice of weights in the TOPSIS method did not significantly affect the selected solution. However, these results cannot be generalized at this time.

Index	Unit	Base Scenario: Variant #1	Base Scenario: Variant #2	e-H <sub>2</sub> Scenario: Variant #1	e-H <sub>2</sub> Scenario: Variant #2
<b>Cost</b>	billion €	8,85E+00	7,63E+00	1,60E+01	1,65E+01
<b>LCOH</b>	€/ kg H <sub>2</sub>	5,15	5,34	10,22	10,28
<b>Risk</b>	-	1,03E+03	1,26E+03	9,57E+02	9,18E+02
<b>Global warming</b>	kg CO <sub>2</sub> eq / kg H <sub>2</sub>	6,92E+00	6,91E+00	4,51E+00	4,82E+00
<b>Stratospheric ozone depletion</b>	kg CFC11 eq / kg H <sub>2</sub>	5,26E-06	5,26E-06	3,76E-06	4,18E-06
<b>Ionizing radiation</b>	kBq Co-60 eq / kg H <sub>2</sub>	7,53E+00	7,53E+00	1,61E+01	1,87E+01
<b>Ozone formation, Human health</b>	kg NOx eq / kg H <sub>2</sub>	1,36E-02	1,36E-02	1,19E-02	1,25E-02
<b>Fine particulate matter formation</b>	Kg PM2.5 eq / kg H <sub>2</sub>	9,27E-03	9,24E-03	1,36E-02	1,42E-02
<b>Ozone formation, Terrestrial ecosystems</b>	kg NOx eq / kg H <sub>2</sub>	1,42E-02	1,42E-02	1,22E-02	1,29E-02
<b>Terrestrial acidification</b>	kg SO <sub>2</sub> eq / kg H <sub>2</sub>	2,62E-02	2,62E-02	3,69E-02	3,88E-02
<b>Freshwater eutrophication</b>	kg P eq / kg H <sub>2</sub>	2,28E-03	2,28E-03	3,02E-03	3,11E-03
<b>Marine eutrophication</b>	kg N eq / kg H <sub>2</sub>	1,84E-04	1,83E-04	4,89E-04	5,39E-04
<b>Terrestrial ecotoxicity</b>	kg 1,4-DCB / kg H <sub>2</sub>	2,15E+01	2,12E+01	1,15E+02	1,18E+02
<b>Freshwater ecotoxicity</b>	kg 1,4-DCB / kg H <sub>2</sub>	1,50E-01	1,50E-01	1,52E+00	1,56E+00
<b>Marine ecotoxicity</b>	kg 1,4-DCB / kg H <sub>2</sub>	2,11E-01	2,10E-01	1,94E+00	1,99E+00
<b>Human carcinogenic toxicity</b>	kg 1,4-DCB / kg H <sub>2</sub>	3,74E-01	3,58E-01	8,68E-01	8,47E-01
<b>Human non-carcinogenic toxicity</b>	kg 1,4-DCB / kg H <sub>2</sub>	3,05E+00	3,04E+00	1,89E+01	1,98E+01
<b>Land use</b>	m <sup>2</sup> crop eq / kg H <sub>2</sub>	2,39E-01	2,39E-01	6,55E-01	6,61E-01
<b>Mineral resource scarcity</b>	kg Cu eq / kg H <sub>2</sub>	2,56E-02	2,47E-02	1,27E-01	1,31E-01
<b>Fossil resource scarcity</b>	kg oil eq / kg H <sub>2</sub>	5,47E+00	5,47E+00	1,20E+00	1,28E+00
<b>Water consumption</b>	m <sup>3</sup> / kg H <sub>2</sub>	5,17E-02	5,17E-02	1,74E-01	1,89E-01

Table 1: Values of cost, risk and environmental indicators for the studied scenarios

The analysis highlights that considering only GHG emissions as an environmental impact criterion, traditional solutions like SMR have the poorest results, while other Key Environmental Indicators (KEIs) such as water consumption or mineral resource scarcity show low performance for the electrolysis process using photovoltaic (PV) technology.

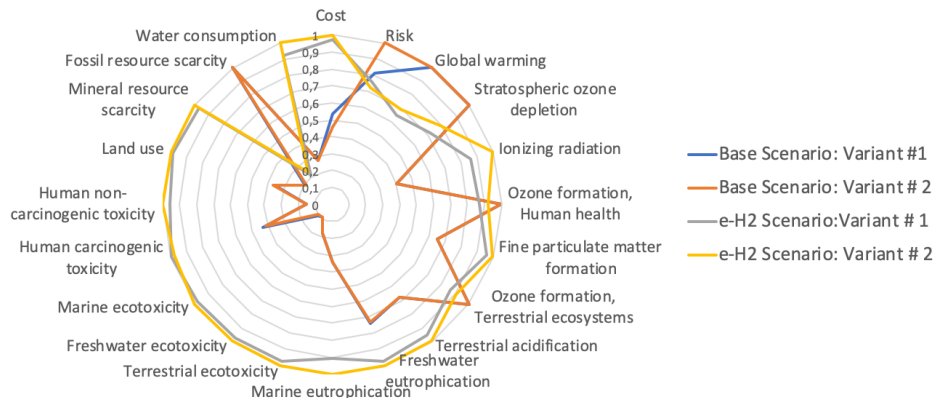


Figure 4 –Radar chart for comparison of all scenarios

## 5. Conclusions

The present paper emphasizes the significance of incorporating LCA into multi-objective optimization analysis as a crucial step in the decision-making process. The proposed framework automates the integration of LCA results with those of the prospective hydrogen energy system model through scenario parameterization. Moreover, this study conducts a thorough well-to-tank analysis, whereas most studies only consider the operational phase impacts. The case study results obtained reinforce the importance of multi-objective optimization approaches that take into account various environmental criteria, not just GHG emissions.

## References

- Acar, C., Bicer, Y., Demir, M.E., Dincer, I., 2019. Transition to a new era with light-based hydrogen production for a carbon-free society: An overview. *Int. J. Hydrogen Energy* 44, 25347–25364. <https://doi.org/10.1016/j.ijhydene.2019.08.010>
- Bareiß, K., de la Rua, C., Möckl, M., & Hamacher, T., 2019. Life cycle assessment of hydrogen from proton exchange membrane water electrolysis in future energy systems. *Applied Energy*, 237, 862-872.
- De-León Almaraz, S., Azzaro-Pantel, C., Montastruc, L., Boix, M., 2015. Deployment of a hydrogen supply chain by multi-objective / multi-period optimisation at regional and national scales. *Chem. Eng. Res. Des.* 104, 11–31. <https://doi.org/10.1016/j.cherd.2015.07.005>
- Koj, J. C., Wulf, C., Schreiber, A., & Zapp, P. (2017). Site-dependent environmental impacts of industrial hydrogen production by alkaline water electrolysis. *Energies*, 10(7), 860.
- IRENA, 2022, Global hydrogen trade to meet the 1.5°C climate goal: Part III – Green hydrogen cost and potential. International Renewable Energy Agency, Abu Dhabi
- Li, L., Manier, H., Manier, M.A., Hydrogen supply chain network design: An optimization-oriented review, *Renewable and Sustainable Energy Reviews*, Volume 103, 2019, Pages 342-360, ISSN 1364-0321, <https://doi.org/10.1016/j.rser.2018.12.060>.
- REPOS, 2017, Scénario région à énergie positive de la Région Occitanie / Pyrénées-Méditerranée.
- Rinawati, D.I., Ryota Keeley, A., Takeda, S., Managi, S., Life-cycle assessment of hydrogen utilization in power generation: A systematic review of technological and methodological

choices, *Front. Sustain.*, 28 July 2022, Sec. Quantitative Sustainability Assessment, Volume 3 - 2022 | <https://doi.org/10.3389/frsus.2022.920876>

Spath, P. L., & Mann, M. K. (2000). *Life cycle assessment of hydrogen production via natural gas steam reforming* (No. NREL/TP-570-27637). National Renewable Energy Lab.(NREL), Golden, CO (United States).



# A techno-economic assessment of biomass combustion with CO<sub>2</sub> capture technology

Nela Slavu<sup>ab\*</sup>, Maytham Alabid<sup>a</sup>, Marius Sandru<sup>c</sup>, Cristian Dinca<sup>ab</sup>

<sup>a</sup>*University POLITEHNICA of Bucharest, Splaiul Independentei, 313, Bucharest, 060042, Romania*

<sup>b</sup>*Academy of Romanian Scientists, Ilfov 3, 050044 Bucharest, Romania*

<sup>c</sup>*SINTEF Materials and Chemistry, Sem Saelands vei 2A, 7465, Trondheim, Norway*  
*slavunela@yahoo.com*

## Abstract

In this study the technical, economic, and environmental performance of a 50 MW biomass power plant with a post-combustion CO<sub>2</sub> capture process by chemical absorption was analysed. The type of biomass used was poplar, with a lower heating value of 18.36 MJ/kg. In the case of CO<sub>2</sub> capture process by chemical absorption, three types of amines were analysed in a mass concentration of 30% (monoethanolamine, diethanolamine, methyl-diethanolamine). For a capture efficiency of 90%, the specific heat consumption for solvent regeneration is 2.76 for MEA, 2.65 for DEA and 2.48 GJ/tCO<sub>2</sub> for MDEA. Thus, after integrating the CO<sub>2</sub> capture process, the overall efficiency of the biomass power plant decreases by up to 8.14 percentage points. After integration of the CO<sub>2</sub> capture process a negative CO<sub>2</sub> emission factor of -746 kg/MWh is obtained. In the case of economic indicators, the levelised cost of electricity increased by up to 129% after the integration of the CO<sub>2</sub> capture process.

**Keywords:** biomass, power plant, CO<sub>2</sub> capture, negative CO<sub>2</sub> emissions.

## 1. Introduction

Biomass use for power generation is a solution to reduce CO<sub>2</sub> emissions from the energy sector. Biomass is considered the renewable energy source with the greatest potential, covering about 10% of global primary energy demand (M.M. Tun, et al., 2019). Biomass accounts for 65% of renewable energy potential in Romania. The energy potential from biomass, estimated at about 7.6 million tonnes/year or 318,000 TJ/year, accounts about 19% of the total primary energy consumption in Romania (R.L. Berevoianu, 2020). There are several methods for using biomass to produce energy, in heat or electricity: thermochemical processes – direct combustion, gasification and pyrolysis, and biochemical processes – anaerobic digestion. The most well-known, verified, and available commercially is direct combustion. Biomass combustion releases about the same amount of CO<sub>2</sub> as coal combustion. However, it is considered neutral because it absorbs the CO<sub>2</sub> released by combustion during growth (A. Sertolli, et al., 2022). CO<sub>2</sub> capture technologies have been developed over the last decades to be integrated into thermal power plants (N.S. Sifat and Y. Haseli, 2019). The CO<sub>2</sub> capture technologies can be integrated pre-/post- or oxy-combustion. The most developed method of capturing CO<sub>2</sub> from a gas stream is by chemical absorption based on aqueous amines. As biomass is considered CO<sub>2</sub> neutral, when a CO<sub>2</sub> capture technology is integrated, negative CO<sub>2</sub> emissions can be obtained. Therefore, a biomass combustion technology coupled with a

CO<sub>2</sub> capture technology can contribute to the decarbonisation of the environment. It is also included in CDR technologies (Carbon Dioxide Removal technologies). This study investigated the integration of the CO<sub>2</sub> capture process in a biomass power plant to produce electricity with negative CO<sub>2</sub> emissions.

## 2. Biomass power plant with post-combustion CO<sub>2</sub> capture

The power plant analysed in this study is based on woody biomass (poplar) with the chemical composition and higher and lower heating value shown in Table 1. The pulverised combustion of biomass was considered for electricity generation. The SST-400 steam turbine was used, with a maximum power of up to 60 MW, and a live steam pressure and temperature of up to 140 bar and 540 °C respectively. The wet and dry flue gas composition resulting from biomass combustion for an excess air coefficient value of 1.4 is shown in Table 2.

Table 1. Biomass characterization

Content	Value
Carbon, wt.%	50.02
Hydrogen, wt.%	6.28
Oxygen, wt.%	42.17
Nitrogen, wt.%	0.19
Sulphur, wt.%	0.02
Ash, wt.%	1.32
Higher heating value, HHV, MJ/kg	18.95
Lower heating value, LHV, MJ/kg	18.36

Table 2. Flue gas composition

Component	Wet flue gas	Dry flue gas
CO <sub>2</sub> , vol.%	10.141	10.973
SO <sub>2</sub> , vol.%	0.002	0.002
N <sub>2</sub> , vol.%	76.465	82.741
O <sub>2</sub> , vol.%	5.807	6.284
H <sub>2</sub> O, vol.%	7.585	0

The main data of the biomass power plant are presented in Table 3. For an 85% load factor of the biomass power plant, corresponding to an annual operating time of 7500 h/year, the biomass flow rate required for 50 MW power is 6.28 kg/s. The electricity produced annually is 375 GWh. The CO<sub>2</sub> emission factor is 829.01 g/kWh. If biomass is considered CO<sub>2</sub> neutral, the emission factor is 0 g/kWh.

Table 3. Power plant main data

Data	Value
Net power, MW	50
Overall efficiency, %	42
Annual operation, h/year	7500
Load factor, %	85
Electric energy, GWh/an	375
Biomass flow, kg/s	6.28
Flue gas flow, kg/s	71.84
CO <sub>2</sub> flow, kg/s	11.51
CO <sub>2</sub> emission factor, gCO <sub>2</sub> /kWh	829.01

The cases analysed are:

- Case 1: Biomass power plant without CO<sub>2</sub> capture;
- Case 2: Biomass power plant with post-combustion chemical absorption process based on monoethanolamine (MEA 30 wt.%);
- Case 3: Biomass power plant with post-combustion chemical absorption process based on diethanolamine (DEA 30 wt.%);
- Case 4: Biomass power plant with post-combustion chemical absorption process based on metil-diethanolamine (MDEA 30 wt.%).

The conventional amine-based CO<sub>2</sub> capture process is shown in Figure 1. The absorption process is characterised by the L/G ratio (ratio of chemical solvent flow rate to flue gas flow rate entering the absorption column). The CO<sub>2</sub> capture efficiency depends on the L/G ratio and the type of amine used in the process. In the case of the regeneration process, the lean loading in CO<sub>2</sub> of 0.15 was constantly maintained for all types of amines studied. The steam required for solvent regeneration is taken from the biomass power plant. Thus, after post-combustion integration of the CO<sub>2</sub> capture process the overall efficiency of the power plant decreases. The CO<sub>2</sub> capture process by chemical absorption was modelled in Chemcad. The thermodynamic package Amine was used.

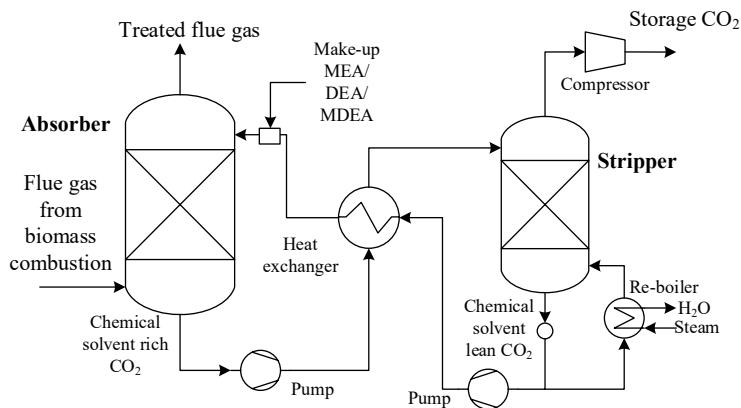


Figure 1. Post-combustion chemical absorption process

Figures 2-4 show the results obtained in terms of CO<sub>2</sub> capture efficiency and thermal energy consumption for chemical solvent regeneration. In the case of using MEA in 30 wt.% concentration, the L/G ratio varied between 0.5-2.1 kmol<sub>solvent</sub>/kmol<sub>flue\_gas</sub>. The CO<sub>2</sub> capture efficiency increases with increasing L/G ratio due to the higher amount of chemical solvent introduced into the absorption column. The CO<sub>2</sub>-rich loading chemical solvent varied in 0.46-0.55 kmol<sub>CO<sub>2</sub></sub>/kmol<sub>MEA</sub>. Heat duty for solvent regeneration increases with increasing L/G ratio, ranging from 31960-135800 MJ/h. For a CO<sub>2</sub> capture efficiency of 90%, the L/G ratio is 1.60 kmol<sub>solvent</sub>/kmol<sub>flue\_gas</sub>, with a specific thermal energy for solvent regeneration of 2.76 GJ/tCO<sub>2</sub>. Regarding the DEA and MDEA, they have a lower CO<sub>2</sub> absorption capacity, resulting in higher amounts of solvent that must be used to achieve CO<sub>2</sub> capture efficiencies up to 100%. Thus, in the case of DEA 30 wt.%, for a CO<sub>2</sub> capture efficiency of 90%, the L/G ratio is 2.85, respectively 6.54 kmol<sub>solvent</sub>/kmol<sub>flue\_gas</sub> for MDEA 30 wt.%. In terms of specific thermal energy consumption, it is 2.65 GJ/tCO<sub>2</sub> for DEA and 2.48 GJ/tCO<sub>2</sub> for MDEA, lower than for MEA due to lower temperatures in the regeneration process (125 °C for MEA, 118 °C for DEA and 104 °C for MDEA). The specific solvent loss is 0.1 kg/tCO<sub>2</sub> for MEA, 0.05 kg/tCO<sub>2</sub> for DEA, respectively 0.31 kg/tCO<sub>2</sub> for MDEA.



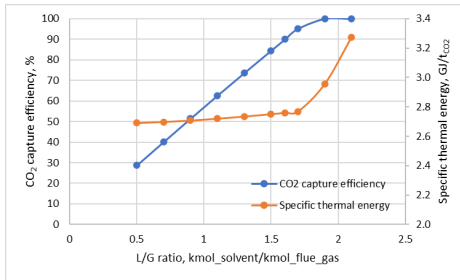


Figure 2. CO<sub>2</sub> capture efficiency and specific thermal energy according to L/G ratio for MEA 30%

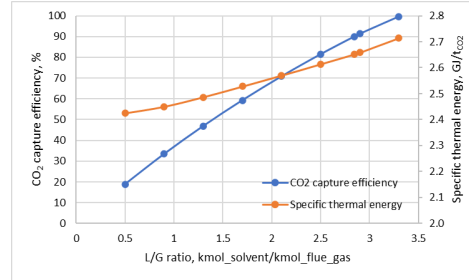


Figure 3. CO<sub>2</sub> capture efficiency and specific thermal energy according to L/G ratio for DEA 30%

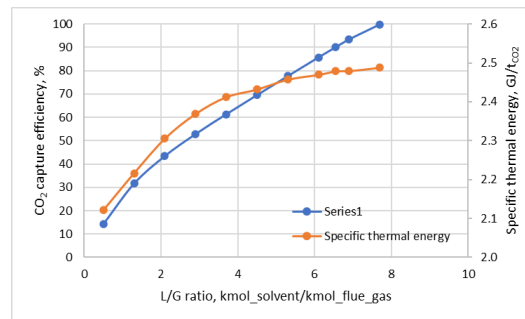


Figure 4. CO<sub>2</sub> capture efficiency and specific thermal energy according to L/G ratio for MDEA 30%

### 3. Techno-economic assessment

The results obtained for the techno-economic indicators are centralised in Table 4. Overall biomass power plant efficiency decreases by up to 20% after integration of CO<sub>2</sub> capture technology (for a CO<sub>2</sub> capture efficiency of 90%). Due to the neutrality of biomass and the integration of CO<sub>2</sub> capture technology, a negative emission factor of 746 g/kWh was achieved. The methodology used to determine technical and economic indicators is presented in another article by the authors (N. Slavu, et al., 2022). The update rate assumed was 8%. The study period is 30 years, during the first 3 years the investment is made. The specific investment cost increases by 55% with the integration of the chemical absorption capture process. Due to the lower purchase price and the smaller amount of solvent used to achieve 90% capture efficiency, the lowest chemical solvent costs were obtained when using MEA (Case 2). The levelised cost of electricity is 80.69 €/MWh in Case 1, for Cases 2, 3 and 4 it increases by 109.27%, 113.74% and 128.27% respectively. The CO<sub>2</sub> capture cost ranges from 95–115 €/tCO<sub>2</sub>, the lowest being in Case 2 (MEA 30 wt.%), and the CO<sub>2</sub> avoided cost ranges from 118–139 €/tCO<sub>2</sub>. In the study by B. Yang et al. (2021), where pulverised biomass combustion with CO<sub>2</sub> capture was analysed, the value obtained for levelised cost of electricity was 168.60 €/MWh, the CO<sub>2</sub> capture cost was 89.21 €/tCO<sub>2</sub>, and the CO<sub>2</sub> avoided cost was 158.85 €/tCO<sub>2</sub>.

Figure 5 shows the influence of the biomass price on the levelised cost of electricity. At a decrease/increase of the biomass price by 10% compared to the one initially considered in the analysis, the levelised cost of electricity decreases/increases by 2.8% in Case 1, by 1.66% in Case 2, by 1.61% in Case 3, and by 1.49% in Case 4.

Table 4. Technical and economic indicators

Technical indicator	Case 1	Case 2	Case 3	Case 4
Specific thermal energy for solvent regeneration, GJ/tCO <sub>2</sub>	-	2.76	2.65	2.48
CO <sub>2</sub> capture efficiency, %	-	90	90	90
CO <sub>2</sub> amount captured, t/year	-	279868.2	279835.2	279823.1
Chemical solvent amount (with losses), t/year	-	317.6	602.6	1448.1
CO <sub>2</sub> amount (neutral biomass-negative emissions), t/year	0.00	-279868.2	-279835.2	-279823.1
Electric energy, MWh/year	375000	302343.7	304662.0	308448.3
Overall efficiency, %	42	33.86	34.12	34.55
Overall efficiency penalty, percentage points	-	8.14	7.88	7.45
CO <sub>2</sub> emission factor, g/kWh	829.01	102.57	101.90	100.69
CO <sub>2</sub> emission factor (neutral biomass - negative emissions), g/kWh	0	-746.32	-746.23	-746.20
Economic indicator	Case 1	Case 2	Case 3	Case 4
Update rate, %	0.08	0.08	0.08	0.08
Operating time, year	30	30	30	30
Investment completion time, year	3	3	3	3
Specific investment cost, €/kW	3917.68	6111.04	6111.04	6111.04
Specific fixed costs O&M, €/kW	42.20	61.92	61.92	61.92
Specific variable costs O&M, €/MWh	4.68	11.42	11.42	11.42
Fuel price, €/kg	0.05	0.05	0.05	0.05
Chemical solvent price, €/kg	-	2.6	3.8	4.5
Investment cost, M€	195.8	305.5	3055.5	305.5
Fixed costs O&M, M€/year	2.11	3.09	3.09	3.09
Variable costs O&M, M€/year	1.75	3.45	3.47	3.52
Fuel cost, M€/year	8.48	8.48	8.48	8.48
Chemical solvent cost, M€/year	-	0.82	2.29	6.51
Levelised cost of electricity, €/MWh	80.69	168.86	172.47	184.19
CO <sub>2</sub> capture cost, €/tCO <sub>2</sub>	-	95.25	99.92	114.09
CO <sub>2</sub> avoidance cost, €/tCO <sub>2</sub>	-	118.14	122.09	138.71

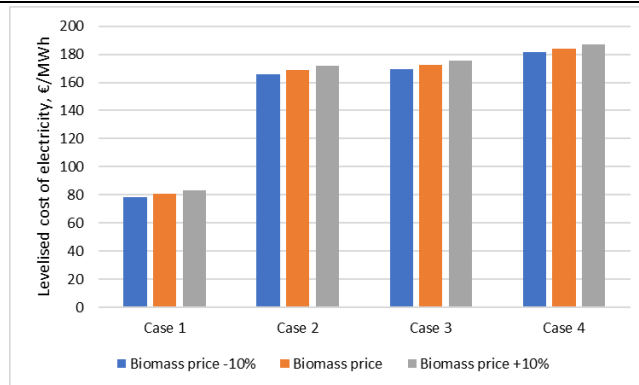


Figure 5. Levelised cost of electricity variation with  $\pm 10\%$  of biomass price

The financial-economic indicators are presented in Table 5. The price of electricity on sale was 250 €/MWh. In all the cases studied, the values of the indicators indicate that the proposed solution is economically feasible (NPV is greater than 0, IRR is greater than the update rate considered, and UPP is less than the operating life). Figures 6 and 7 show the variation of the NPV and the UPP as a function of the update rate. An update rate of 8,

10 and 12% was considered. The best values are obtained at the lowest update rate of 8%. As the update rate increases the NPV decreases and the UPP period increases. In Case 3 and Case 4, for an update rate of 12% the project is no longer feasible.

Table 5. Financial-economic indicators

Indicator	Case 1	Case 2	Case 3	Case 4
Net present value, NPV, M€	668	182	174	146
Internal rate of return, IRR, %	30.96	16.39	16.16	15.31
Updated payback period, UPP, year	7.04	14.61	14.87	15.92

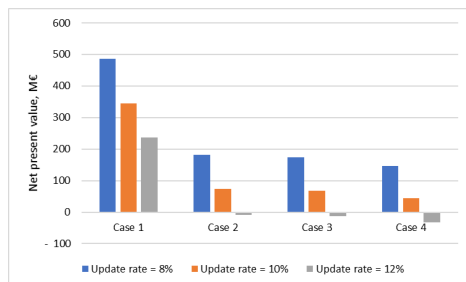


Figure 6. NPV according to update rate

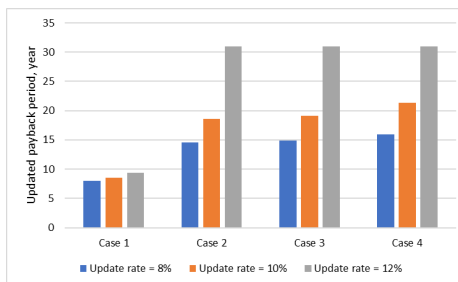


Figure 7. UPP according to update rate

#### 4. Conclusions

Biomass combustion coupled with CO<sub>2</sub> capture technology can significantly contribute to CO<sub>2</sub> emission reductions, as both the electricity generation process and the post-combustion capture by chemical absorption process are mature and can be implemented on industrial scale. In this study, for a 50 MW biomass power plant, a negative CO<sub>2</sub> emission factor of 746 g/kWh was obtained after integration of CO<sub>2</sub> capture technology. Therefore, over a year, this technology can absorb about 279.8 ktCO<sub>2</sub> from the environment. The levelized cost of electricity for the biomass power plant without CO<sub>2</sub> capture process is 80.69 €/MWh. In the case of integrating the CO<sub>2</sub> capture process by chemical absorption, the cost obtained is between 168–185 €/MWh, depending on the type of amine used.

#### Acknowledgments

The study was funded by the UEFISCDI within the National Project number 106PTE/2022. The research leading to these results received funding from the NO Grants 2014–2021, under project contract No. 13/2020.

#### References

- M.M. Tun, et al., 2019, Biomass energy: An overview of biomass sources, energy potential, and management in Southeast Asian countries, *Resources*, 8(2), 81.
- R.L. Berevoianu, 2020, Main directions for energy use of biomass in Romania, *International Symposium*, 11th Edition (pp. 333–339).
- A. Sertolli, et al., 2022, Biomass Potential and Utilization in Worldwide Research Trends—A Bibliometric Analysis. *Sustainability*, 14(9), 5515.
- N.S. Sifat and Y. Haseli, 2019, A critical review of CO<sub>2</sub> capture technologies and prospects for clean power generation. *Energies*, 12(21), 4143.
- N. Slavu, et al., 2022, Technical and Economical Assessment of CO<sub>2</sub> Capture-Based Ammonia Aqueous. *Processes*, 10(5), 859.
- B. Yang, et al., 2021, Life cycle cost assessment of biomass co-firing power plants with CO<sub>2</sub> capture and storage considering multiple incentives. *Energy Economics*, 96, 105173.

### **Reviewers' answers**

First, we would like to thank you for your recommendations and remarks.

1. The novelty is not clear. Amine based CO<sub>2</sub> capture has been studied for long. So the specific objective of this new study is not clear.

As you said, the CO<sub>2</sub> capture process by chemical absorption-based amine has been studied over the years and is a mature process. The aim of this study was to integrate this process in a biomass power plant, and to analyze economically the use of different amines. Considering that we have a biomass power plant with CO<sub>2</sub> capture, this technology is included in the CDR technologies (carbon dioxide removal technologies).

2. How are data reported in table 3 obtained, particularly numbers for CO<sub>2</sub> emission factor?

Table 3 reports the data obtained from the thermodynamic cycle calculation. The CO<sub>2</sub> emission factor was determined based on the flue gas composition, the flue gas flow rate and the amount of electricity produced.

3. What do you mean by lean loading of CO<sub>2</sub> in the regeneration system?

The lean loading solvent is the molar ratio of CO<sub>2</sub> to the used amine (MEA, DEA, MDEA) in the stream leaving the regeneration column.

4. Please add some information regarding Chemcad simulation (thermodynamic models, source of kinetic data etc.)

The thermodynamic package Amine was used to simulate the CO<sub>2</sub> capture process by chemical absorption in Chemcad. We have added this information in the text.

5. The negative emission factor seems to be calculated based on 90% capture efficiency. I do not think it is from life cycle perspective where the emissions related to amine production as well as recycling will be considered. Please clarify.

Yes, you are right. In this study no life cycle analysis was done for either the biomass or the amines used.

6. The difference between CO<sub>2</sub> capture cost and CO<sub>2</sub> avoidance cost is not clear.

The difference between the two calculated costs can be seen from the relationships applied. I did not add them in this article because a reference (another article by the authors) has been placed where the methodology for calculating all economic indicators is described. In the case of CO<sub>2</sub> capture cost, the difference between levelized cost of electricity with capture and without capture is reported to the amount of CO<sub>2</sub> captured. In the case of the CO<sub>2</sub> avoidance cost, the same difference is reported to the difference between the CO<sub>2</sub> emission factor with capture and without capture.

7. Comparison of results with prior literature will be helpful.

We compared with the literature on economic indicators. Thus, in the study by B. Yang et al. (2021), where pulverized biomass combustion with CO<sub>2</sub> capture was analyzed, the value obtained for levelized cost of electricity was 168.60 €/MWh, the CO<sub>2</sub> capture cost was 89.21 €/tCO<sub>2</sub>, and the CO<sub>2</sub> avoided cost was 158.85 €/tCO<sub>2</sub>. In our study the levelized cost of electricity with CO<sub>2</sub> capture is between 168-185 €/MWh, the CO<sub>2</sub> capture cost is between 95-115 €/tCO<sub>2</sub>, and the CO<sub>2</sub> avoided cost is between 118-139 €/tCO<sub>2</sub>.



# Process Modelling and Pinch Analysis for an Integrated System of Anaerobic Digestion with Digestate Recycling via Hydrothermal Gasification

Fadilla Noor Rahma<sup>a</sup>, Khanh-Quang Tran<sup>a</sup>, Roger Khalil<sup>b</sup>

<sup>a</sup>*Norwegian University of Science and Technology, Høgskoleringen 1, Trondheim 7034, Norway*

<sup>b</sup>*SINTEF Energy Research, Sem Sælands vei 11, Trondheim 7034, Norway*

## Abstract

Anaerobic digestion (AD) of low-grade biomass materials is one of the strategical technologies to achieve the climate-neutral target. Its implementation is however limited due to problems related to digestate management and feedstock scarcity. Digestate recycling to the AD process via hydrothermal gasification (HTG) may be an answer to the problems. This work develops and analyzes thermodynamically a novel integrated process including among others AD, HTG, and power generation unit for energy recovery from the HTG, using Aspen Plus software, with a focus on the HTG block. Effects of HTG process parameters on biogas production and energy recovery are investigated. In addition to that, a Pinch analysis was performed. The result shows that higher energy efficiency is obtained at higher temperature and higher D/W ratio. The syngas recycling to the AD increases biomethane production by 26.10%, power production by 21.42%, and total energy efficiency by 20.05%, compared to the process without syngas recycle.

**Keywords:** Anaerobic digestion; hydrothermal gasification, Aspen Plus; process integration; Pinch analysis.

## 1. Background

Biogas production from low-grade biomass via anaerobic digestion (AD) is essential in the transition towards sustainable and renewable energy. However, AD implementation has been hindered by challenges related to the digestate management [1] and the limited sustainable feedstocks [2]. Integration of AD and hydrothermal gasification (HTG) is a potential answer to solve both the digestate management and feedstock scarcity issues.

Indeed, the AD-HTG integration allows valorization of AD digestate through HTG, facilitating the digestate management [3]. Another key benefit of the AD-HTG integration is the potential to recycle HTG-generated syngas to the AD system [4].

Previous studies have confirmed that addition of syngas produced from the digestate via HTG into the AD accelerates the digestion process and enhances CH<sub>4</sub> production [5, 6]. Prior to syngas recycle, the energy associated with HTG product fluid streams at high pressures and temperatures can be recovered for electricity generation through a series of turbines and heat exchangers. A number of process configurations integrating HTG with power production has been proven to be more efficient than the stand-alone processes [7].

In this paper, a conceptual process design of AD-HTG integration with syngas recycling and power production is proposed. To the best of the authors' knowledge, such system has not been previously studied in the literature. This research aims to predict the performance of the integrated system under various operating conditions, focusing on the HTG block. The process performance is indicated by biomethane and power production capacity, as well as the process energy efficiency. In addition, the heat recovery potentials of the integrated process are evaluated using Pinch analysis.

## 2. Methods

### 2.1. Description of the Process Simulation

The conceptual design of the integrated system under study and its model in Aspen Plus are presented in Figure 1 and Figure 2, respectively.

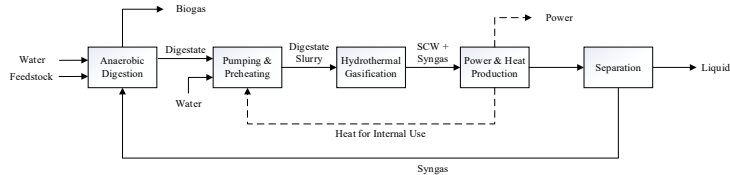


Figure 1. Process Design of the Integrated System under Investigation

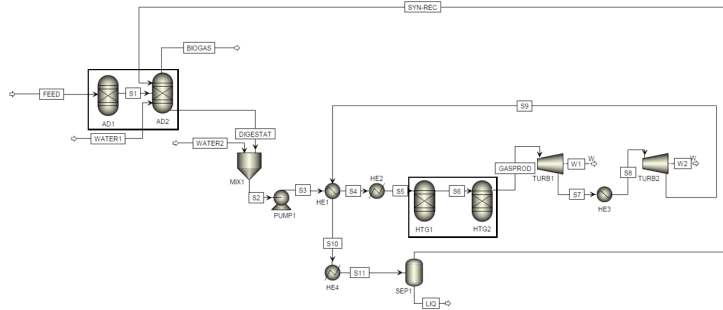
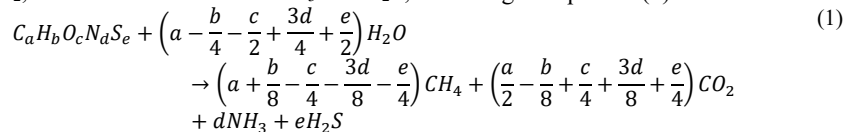


Figure 2. Process Flow Diagram of the Integrated System in Aspen Plus

#### 2.1.1. Anaerobic Digestion (AD)

In Aspen Plus, AD process is modelled using a kinetic-free equilibrium model based on Buswell equation [8]. This equation uses theoretical stoichiometric method to estimate the products of the AD process, including biogas and digestate. The method was widely applied in earlier studies [9-11] and adopted for the present work to reduce the model complexity. According to this method, the volatile matter in the biomass feedstock is converted into  $\text{CH}_4$ ,  $\text{CO}_2$ , and smaller amounts of  $\text{NH}_3$  and  $\text{H}_2\text{S}$ , according to Equation (1) below:



For the Aspen Plus modelling, two blocks of RYIELD reactor, each incorporated with a calculator block, are employed to represent AD. Biomass feedstock (*FEED*) is defined in Aspen Plus as a non-conventional compound consisting of volatile matter, fixed carbon, moisture, and ash content. The first RYIELD reactor (*AD1*) is used to breakdown the volatile matter in the feedstock into its components, consisting of carbon, hydrogen, oxygen, nitrogen, and sulfur. The output of *AD1* is the fictitious stream *SI*, which does not represent a physical stream in the real condition. The *SI* stream flows into another RYIELD reactor (*AD2*) which calculates the biogas and digestate production based on Equation (1). In the simulation, the recycle stream (*SYN-REC*) and water stream (*WATER1*) also enter the AD system through this reactor. During the AD process, the fixed carbon and ash contents are assumed to remain undigested [10]. The AD system produces two product streams, i.e., the biogas (*BIOGAS*) and digestate (*DIGESTAT*), which is further processed in the HTG section.

#### 2.1.2. Hydrothermal Gasification (HTG)

The HTG reactor model is based on the Gibbs free energy minimization principle. This approach is considered suitable to evaluate the performance and predict the equilibrium composition of HTG products [7]. Similar approach has been widely adopted in numerous

HTG modelling works utilizing Aspen Plus [12, 13], where the modelling results have been verified to accurately represent the experimental data from previous HTG studies.

Prior to entering the reactor, the digestate from the AD reactor (*DIGESTAT*) is mixed with water (*WATER2*) in a mixer (*MX1*) and pumped to the reactor pressure in *PUMPI*. Subsequently, the stream *S3* is preheated in heat exchanger *HE1*, which utilizes heat from the turbine's exhaust stream, and sent to heat exchanger *HE2* for further heating to its supercritical state. Since the *DIGESTAT* stream still contains the non-conventional fixed carbon compound from the undigested AD product, an RYIELD reactor block (*HTG1*) is employed to breakdown the compound into its elements. The output of *HTG1* is a fictitious stream *S6*, which flows to the RGIBBS reactor (*HTG2*), where the Gibbs free energy minimization takes place. The reaction produces gaseous product (*GASPROD*) consisting of supercritical water and syngas, which is mainly composed of H<sub>2</sub>, CO, CO<sub>2</sub>, and CH<sub>4</sub>.

### 2.1.3. Power generation for energy recovery from HTG

The hot-pressurized gaseous product from the HTG (*GASPROD*) contains a high amount of energy which can be recovered via power generation. For this purpose, it is important to consider a key requirement of the system that the syngas can be recycled to the AD reactor. Therefore, the power generation in this integrated system is designed to utilize the sensible heat in the *GASPROD* stream without chemical conversion. This is achieved by employing two turbines (*TURB1* and *TURB2*) for power generation with a reheater (*HE3*) between the turbine stages. Two-stage turbines with reheater are employed to obtain a higher energy recovery. A number of previous studies have also reported using turbine for direct expansion and energy recovery of HTG gaseous product [14, 15]. The heat remaining in the turbine exhaust stream (*S9*) undergoes cooling in *HE1* and *HE4* before the stream is sent to *SEP1* for the separation of syngas and water phase.

## 2.2. Performance Evaluation of the Process

The performance of the proposed integrated system is evaluated by investigating the effect of important process parameters, i.e., HTG temperature and digestate-to-water (D/W) ratio, on biomethane and power production. As the proposed system requires additional heat supply from the utility, the total heat requirement is also assessed. In addition, the overall energy efficiency,  $\eta$ , of the process is calculated using Equation (2) [16], where the *total energy recovered* is the sum of recovered energy in the form of electricity and lower heating value (LHV) of the produced biogas; and the *total energy input* is the sum of input energy from available sources, including heating value (LHV) of the feedstock.

$$\eta = \frac{\text{Total energy recovered}}{\text{Total energy input}} \quad (2)$$

### 2.3. Pinch Analysis

Pinch analysis is performed in this study to identify the possibility for reducing the energy consumption (heating and cooling duties) of the system by heat integration. The first step is analyzing the available heat sinks and sources in the processes. A grand composite curve is developed based on the heat sinks and sources data obtained from the simulation [17].

## 3. Results

To evaluate the effect of syngas recycling, a comparison of the proposed process with and without syngas recycle was carried out. The comparison was performed under the base-case condition (Table 1).

Table 1. Process Condition for the Aspen Plus Simulation

Process parameter	Value
HTG Temperature	600 °C
HTG Pressure	250 bar
Feedstock mass flow	10000 kg/hr



Figure 3 (a) and (b) show a clear correlation between the amount of recycled syngas from the HTG and total biomethane production from the AD, indicating that the amount of recycled syngas directly affects biomethane production. According to the figures, both syngas and biomethane production increase at higher HTG temperature and lower D/W ratio. The HTG temperature positively affects syngas production by enhancing the highly endothermic stream reforming reaction within the HTG [12, 18]. On the other hand, a lower D/W ratio means providing more water into the system, making the HTG feed more dilute. Water played a significant role in HTG as a reactant for the steam-reforming reaction and water-gas shift reaction, both are enhanced with an increase in water concentration [19]. Previous studies have confirmed that a lower feedstock concentration leads to higher gasification efficiency and increasing gas yield [20]. Furthermore, both figures indicate that the influence of temperature is more prominent at lower D/W ratio. A possible reason is that the lower feedstock concentration minimizes competing reaction pathways [21]. Therefore, the positive effect of temperature towards the main reactions in HTG are more significant at lower D/W ratio.

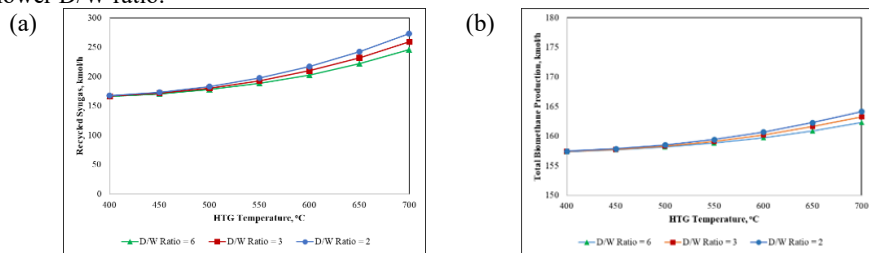


Figure 3. Effect of HTG Temperature and D/W Ratio on (a) Syngas Production; and (b) Biomethane Production ( $P = 250$  bar)

The influence of temperature and D/W ratio on the energy performance of the process is shown on Figure 4. Operating the HTG reactor at a higher temperature requires higher heat input prior to the HTG reactor (HE 1 and HE 2). Accordingly, the total heat requirement increases as the temperature rises. However, higher HTG temperature also leads to an increase in the total power recovery. This is due to the higher syngas production, as previously indicated in Figure 3 (a), which leads to pressure increase in the steam-syngas mixture entering the steam turbines. Consequently, higher HTG temperature results in an increased overall energy efficiency of the process (Figure 4 (b)). On the other hand, lowering D/W ratio, which means providing more water into the system, also increases both the power production and heat requirement. Operating at a lower D/W ratio increases the heat requirement substantially, whereas the increase in power recovery is less prominent, as suggested by Figure 4 (a). Consequently, as shown in Figure 4 (b), the overall energy efficiency of the process decreases at lower D/W ratios.

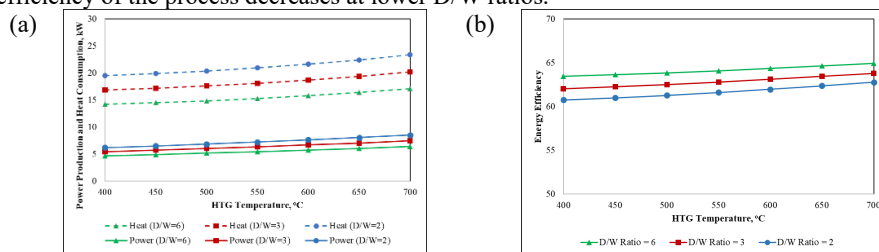


Figure 4. Effect of HTG Temperature and D/W Ratio on (a) Total Produced Power and Total Required Heat; and (b) Energy Efficiency ( $P = 250$  bar)

Figure 5 displays the grand composite curve obtained from pinch analysis of the system. The curve was obtained using  $\Delta T_{\min}$  of 10 °C. The analysis found that the system does not

have further potential for heat integration, as the availability of heat sinks are significantly larger than those of heat sources. The utilization of the turbine exhaust stream (S9) for the preheating of the stream entering HTG reactor (S3) reduces 1.1 MW of both the hot and cold utility requirement. However, 14.08 MW of heating duty and 0.22 MW of cooling duty should be provided by the utility.

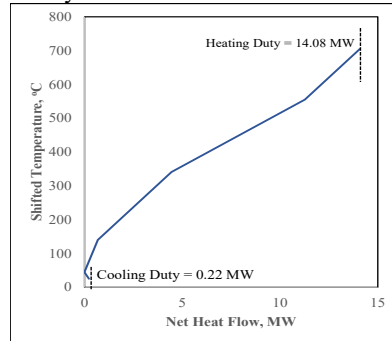


Figure 5. Grand Composite Curve from Pinch Analysis

#### 4. Discussion

The present study confirms the potential to improve biomethane production in AD through an integrated AD-HTG system with syngas recycle. The developed process model has been able to investigate the effect of key process parameters with a focus on the HTG block, without modifying process parameters in the other parts of the process. Overall, the result shows a substantial increase of 26.10% in total biomethane production when syngas is recycled to the AD. In addition, power production increases by 21.42% and total energy efficiency is 20.05% higher compared to the process without syngas recycle. However, evaluation of other process parameters as such of the AD block would be of interest. Future works might benefit from utilizing a more complex model that can better represent the complex biochemical reactions in the AD system.

On the other hand, the Pinch analysis was performed in this study with a  $\Delta T_{\min}$  of 10 °C. The consequence of using a low  $\Delta T_{\min}$  is that a large heat transfer area might be required, leading to higher heat exchanger capital cost. A more comprehensive evaluation with techno-economic analysis should be performed in the future work. It also suggests that the process design requires further modification to obtain a higher heat integration. A possible option for example is integrating a gas turbine combusting a fraction of the syngas.

#### 5. Conclusion

A process integration of AD, HTG, and power production is proposed in this paper to improve the efficiency of the AD. A process model is developed using Aspen Plus to assess the performance of the integrated process. The effect of temperature (400-700 °C) and digestate-to-water ratio (2, 3, and 6) is investigated. The result shows that higher energy efficiency is obtained at higher temperature and higher D/W ratio. The syngas recycling to the AD increases biomethane production by 26.10%, power production by 21.42%, and total energy efficiency by 20.05%, compared to the process without syngas recycle. The proposed process design provides an innovative solution for digestate management and feedstock scarcity, as well as enhancing biomethane production via AD, with an additional benefit of producing power from the integrated process.

#### Acknowledgements

This work was supported by the BioSynGas project partly funded by the Research Council of Norway (Project Number: 319723).

## References

1. R. Nkoa, 2014, Agricultural benefits and environmental risks of soil fertilization with anaerobic digestates: a review, *Agronomy for Sustainable Development*, 34(2), 473-492.
2. D. Divya, L. Gopinath, and P.M. Christy, 2015, A review on current aspects and diverse prospects for enhancing biogas production in sustainable means, *Renewable and sustainable energy reviews*, 42, 690-699.
3. R. Kumar, S. Singh, and O.V. Singh, 2008, Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives, *Journal of industrial microbiology and biotechnology*, 35(5), 377-391.
4. A. Molino, et al., 2019, Experimental and theoretical investigation on the recovery of green chemicals and energy from mixed agricultural wastes by coupling anaerobic digestion and supercritical water gasification, *Chemical Engineering Journal*, 370, 1101-1110.
5. Z. Yang, et al., 2020, Improvement of biofuel recovery from food waste by integration of anaerobic digestion, digestate pyrolysis and syngas biomethanation under mesophilic and thermophilic conditions, *Journal of Cleaner Production*, 256, 120594.
6. Y. Li, Y. Chen, and J. Wu, 2019, Enhancement of methane production in anaerobic digestion process: A review, *Applied energy*, 240, 120-137.
7. D. Hantoko, et al., 2019, Supercritical water gasification of sewage sludge and combined cycle for H<sub>2</sub> and power production—a thermodynamic study, *International Journal of Hydrogen Energy*, 44(45), 24459-24470.
8. G. Symons and A. Buswell, 1933, The methane fermentation of carbohydrates, *Journal of the american chemical society*, 55(5), 2028-2036.
9. H. Nguyen, S. Heaven, and C. Banks, 2014, Energy potential from the anaerobic digestion of food waste in municipal solid waste stream of urban areas in Vietnam, *International Journal of Energy and Environmental Engineering*, 5(4), 365-374.
10. A. Naqi, J.N. Kuhn, and B. Joseph, 2019, Techno-economic analysis of producing liquid fuels from biomass via anaerobic digestion and thermochemical conversion, *Biomass and Bioenergy*, 130, 105395.
11. A. Skorek-Osikowska, et al., 2020, Thermodynamic, economic and environmental assessment of energy systems including the use of gas from manure fermentation in the context of the Spanish potential, *Energy*, 200, 117452.
12. D. Hantoko, et al., 2018, Thermodynamic study on the integrated supercritical water gasification with reforming process for hydrogen production: Effects of operating parameters, *International Journal of Hydrogen Energy*, 43(37), 17620-17632.
13. J.A. Okolie, et al., 2020, Hydrothermal gasification of soybean straw and flax straw for hydrogen-rich syngas production: Experimental and thermodynamic modelling, *Energy Conversion and Management*, 208, 112545.
14. A. Molino, et al., 2016, Process innovation via supercritical water gasification to improve the conventional plants performance in treating highly humid biomass, *Waste and biomass valorization*, 7(5), 1289-1295.
15. C. Cao, 2017, et al., System analysis of pulping process coupled with supercritical water gasification of black liquor for combined hydrogen, heat and power production, *Energy*, 132, 238-247.
16. M. Luqman and T. Al-Ansari, 2021, A novel solution towards zero waste in dairy farms: A thermodynamic study of an integrated polygeneration approach, *Energy Conversion and Management*, 230, 113753.
17. I.C. Kemp, 2011, *Pinch analysis and process integration: a user guide on process integration for the efficient use of energy*, Elsevier.
18. D. Macri, et al., 2020, Supercritical water gasification of biomass and agro-food residues: Energy assessment from modelling approach, *Renewable Energy*, 150, 624-636.
19. M. Yan, et al., 2019, Experimental study on the energy conversion of food waste via supercritical water gasification: improvement of hydrogen production, *International Journal of Hydrogen Energy*, 44(10), 4664-4673.
20. Y. Matsumura, et al., 2005, Biomass gasification in near-and super-critical water: status and prospects, *Biomass and Bioenergy*, 29(4), 269-292.
21. J.A. Onwudili and P.T. Williams, 2014, Production of hydrogen from biomass via supercritical water gasification, in *Near-critical and supercritical water and their applications for biorefineries*, Springer, 299-322.

# Optimization of a continuous multi-stage fluidized bed system for CO<sub>2</sub> capture utilizing temperature swing adsorption

Yuri Souza Beleli,<sup>a</sup> José Luis de Paiva,<sup>a</sup> Marcelo Martins Seckler,<sup>a</sup> Galo Antonio Carrillo Le Roux<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, Polytechnic School, University of São Paulo, Av. Prof. Luciano Gualberto, trav. 3, 380, São Paulo 05508-010, Brazil.*

## Abstract

In this work we study a system to capture the CO<sub>2</sub> resulting from processes of power generation combustion by capturing it with a solid adsorbent in a fluidized bed, thus obtaining a process with low carbon emission. The flue gas firstly goes through the adsorption column in an upward flow, contacting an adsorbent fluidized at each stage, which captures the CO<sub>2</sub>. The adsorbent flows downward, from a stage to the next right below. The resulting bottom solid flow containing the loaded adsorbent is transported to the desorption column so that a stripping gas removes the CO<sub>2</sub> from the adsorbent. The simulation is carried out in GAMS so that the conditions necessary to achieve the adsorption and the desorption processes could be analyzed and optimized to best synthesize the full system.

**Keywords:** CO<sub>2</sub> adsorption; fluidized bed; CO<sub>2</sub> capture; Temperature Swing Adsorption.

## 1. Introduction

For the mitigation of the effects of greenhouse gases (GHG) concentration in the atmosphere, CO<sub>2</sub> capture is a possible solution because the captured CO<sub>2</sub> can be concentrated and stored or used as a chemical product.

Recently, the use of solid adsorbents has been studied due to their economic and energetic advantages<sup>1</sup>. In trying to progress this technology further, it is possible to use solid fluidized bed adsorbents, which have even better thermal and material exchange conditions<sup>2</sup>. Adsorption is an exothermic process, while desorption is an endothermic process. To capture CO<sub>2</sub>, both adsorption and desorption steps are necessary, which makes the temperature swing adsorption process studied in this paper a possibility. The temperature swing adsorption defines that the continuous system composed of the two parts (adsorption and desorption) should be operated at temperatures that favor the respective mechanics, thus modulating the temperatures between them.

The reduction of CO<sub>2</sub> emissions has become of great interest worldwide in order to reduce the effects of this gas in the atmosphere. An issue of great interest is controlling the rate of global warming, discussed in the treaty of Paris, which proposes to keep it between 1.5 and 2.0°C on average<sup>3</sup>. Among the alternatives for reducing the emission of greenhouse gases, carbon capture technologies stand out. Post-combustion capture technologies have been considered the most viable because they can be implemented directly in the effluent gases of the combustion energy production process<sup>4</sup>. The CO<sub>2</sub> resulting from capture technologies can be transported for storage or used as feed for industrial processes.

For the adsorption process to be used continuously, it is necessary that there be recovery of the adsorbent (desorption of the adsorbate), which can be done through a cyclic chain where the situations occur one after the other.

The adsorption process is favored at high pressures and low temperatures while the adsorbent recovery process (desorption) is favored at low pressures and high temperatures. In cyclic adsorption processes there are the options of pressure modulation (PSA), temperature modulation (TSA) and a combination of these processes (PTSA).

In pressure modulation, the adsorbed gases in the adsorption step are desorbed by reducing the pressure in the bed, while in temperature modulation the desorption occurs by heating the adsorbent<sup>5</sup>.

The system studied is composed of two columns, in which the first one is the adsorption column and the second one is the desorption column. In both columns the solid flows downwards while the gas flows upwards. The columns are connected by the solid flow, which after coming out of the adsorption column at its base, is transported to the desorption column at its top so the CO<sub>2</sub> adsorbed can be stripped and captured for stocking or posterior usage.

## 2. Methodology

For multi-stage adsorption and desorption columns, stage modeling is used, where the balances of each stage are described in a general way so that the set of stages can be used in the optimization step. The stage has two countercurrent flows, the gas flows upwards and the solid flows downwards. The stages are designed so the gas flow is able to fluidize the solid flow, resulting in a downward fluidized bed. A representation of a generic adsorption and desorption column stage can be found in Figure 1. The mass balance is represented in Equation (1).

$$F_{j-1}y_{i,j-1} - ms(q_{i,j} - q_{i,j+1}) - F_jy_{i,j} = 0 \quad (1)$$

The adsorbed amount is simulated using Toth's equilibrium, represented in Equation (2), which relates the partial pressures of the components to the equilibrium variables.

$$q_i^* [\text{mol/kg}] = \frac{q_{mi}(b_iP_i)^{n_i}}{(1 + \sum_i (b_iP_i)^{n_i})^{\frac{1}{n_i}}} \quad (2)$$

In the energy balance, the total enthalpies of the inlet and outlet currents are evaluated, as well as the energy added to or removed from the system. It is represented by Equation (3).

$$F_{j-1}h_{j-1} + ms(H_{j+1} - H_j) - F_jh_j + Q_j = 0 \quad (3)$$

The enthalpies of the gas and solid streams are evaluated, with the enthalpies of the components present, the enthalpy of the solid and the enthalpy of the adsorption process, they are presents in Equation (4) and Equation (5).

$$h_j [\text{J/mol}] = \sum_i y_{i,j} \int_{T_r}^{T_j} cp_i dT \quad (4)$$

$$H_j [\text{J/kg}] = \int_{T_r}^{T_j} cp_s dT + \sum_i q_{i,j} \int_{T_r}^{T_j} cp_i dT - 1000 \sum_i \Lambda(q_{i,j}) \quad (5)$$

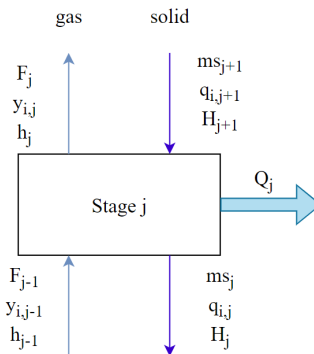


Figure 1: Stage representation with its gas and solid flows.

The parameters of the equilibrium model were previously validated<sup>6</sup> and used to simulate the stages. This model was then simulated in GAMS so the conditions necessary to achieve the adsorption and the desorption processes could be studied.

### 3. Results and discussion

#### 3.1. Adiabatic Simulation

##### 3.1.1. Adsorption column

For the adsorption simulations, a gas stream with 15% CO<sub>2</sub> and 85% N<sub>2</sub> was considered. The solid inlet stream is composed of pure adsorbent (zeolite 13x BF). The column inlet streams are at 323 K<sup>7</sup>. The simulation was done by varying the ratio between the gas-solid mass flow rates, fixing the gas flow rate and varying the pure adsorbent flow rate at the column inlet (S/G ratio). The results of the solid stream outlet temperature as a function of the S/G ratio of the column feed (with a different number of stages) are shown in Figure 2.

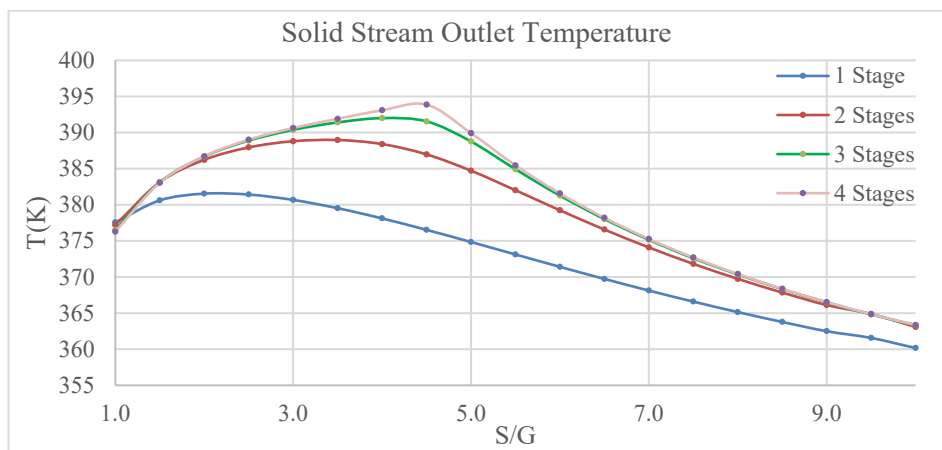


Figure 2: Bottom solid stream temperature related to column feed S/G ratio.

By increasing the amount adsorbed (increasing S/G), more adsorption heat is released and then the temperature increases. However, by increasing S/G, more sensible heat will be required for the temperature of the currents to increase. In each simulation there is a

maximum temperature point, and from this point on, the column temperature decreases as S/G increases, because the amount of solid becomes more prominent than the amount adsorbed, in terms of energy distribution.

The results of the total percentage adsorbed (of the amount of CO<sub>2</sub> available in the gas) function of the S/G ratio of the column feed are shown in Figure 3.

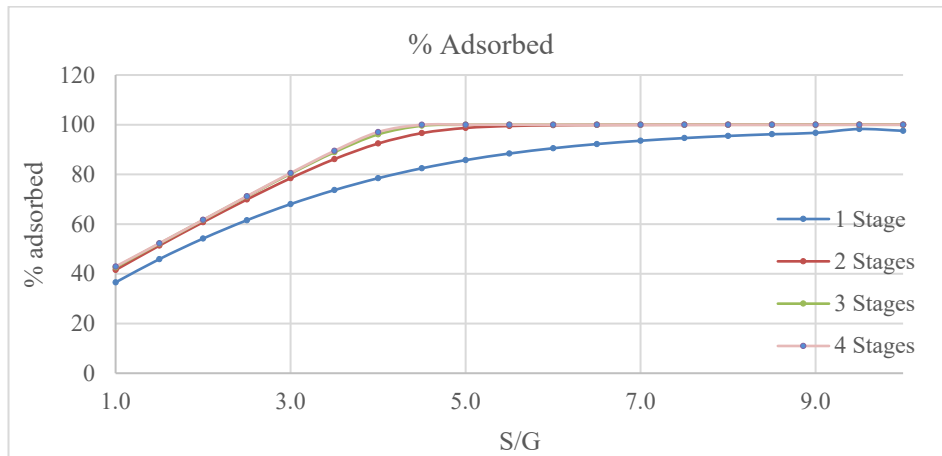


Figure 3: Percentage adsorbed related to column feed S/G ratio.

By increasing the number of stages, a lower S/G ratio is required for the adsorbed percentage to reach significant numbers, as well as the solid flow rate. As stated before, the relationship between number of stages and solid flow rate will be an essential part of the optimization, since their costs will be of extreme importance when calculating the total cost of the project.

To confirm how adsorption is influenced by column temperature, a one (1) stage simulation was performed, with S/G ratio equal to five (5), varying the temperature between 313 K and 413 K. The results are shown in Figure 4.

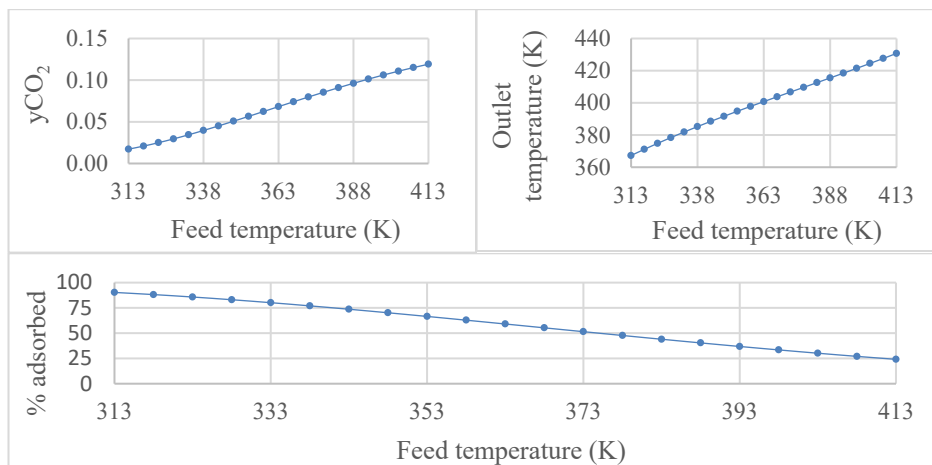


Figure 4: Variables of interest related to the feed temperature in 1 stage.

Since the adsorption reaction is exothermic, increasing the temperature will disfavor it. By increasing the temperature in the column, the percentage of CO<sub>2</sub> adsorbed will be lower, causing its mole fraction in the outlet gas to be higher. As less adsorption occurs, causing less heat to be released into the column, the difference between the column inlet and column outlet temperatures decreases.

### 3.1.2. Desorption Column

For the recovery simulations a gaseous stream of pure CO<sub>2</sub> was considered so that the driving forces relating to the gas composition and temperatures can be studied. The solid inlet stream is composed of adsorbent (zeolite 13x BF) and adsorbed CO<sub>2</sub>. The column inlet streams are at 513 K<sup>7</sup>. Simulations of the desorption column with different inlet temperatures were made to evaluate the temperature influence in the process. The results of the percentage recovered and outlet temperatures of the solid outlet stream are shown in Figure 5.

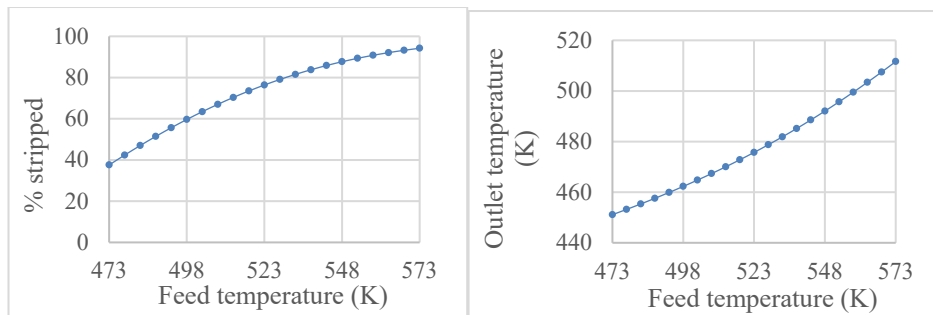


Figure 5: Percentage stripped and outlet temperature related to the feed temperature.

As expected, when increasing the inlet temperature of the currents, the recovery is favored and the percentage recovered is higher. Like adsorption, when the desorption mechanism is favored the difference between the temperatures of the inlet and outlet column currents increases.

### 3.2. Isothermal simulation

For the isothermal simulation of the adsorption column, a gas with 15% CO<sub>2</sub> and 85% N<sub>2</sub> was again considered. The solid inlet stream is composed of pure adsorbent (zeolite 13x BF). The column inlet and outlet streams are at 323 K<sup>7</sup>.

Simulations were done by varying the ratio between the S/G ratio. One (1) stage results of the gas mole fraction at the outlet, percentage adsorbed are shown in Figure 6.

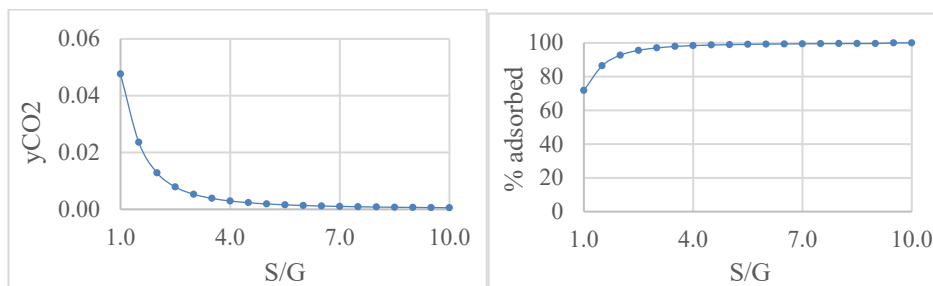


Figure 6: Outlet gas molar fraction and percentage related to feed column S/G ratio.



Later simulations with more than 1 stage showed that there is no significant increase in the adsorption process in this specific temperature when compared to 1 stage, due to the saturation of the adsorbent (lower S/G) or adsorption near 100% of equilibrium (higher S/G).

#### 4. Conclusion

A mathematical model for stage adsorption has been developed so that it can be extended to multi-stage configurations and to the desorption process. The model consists of mass and energy balances coupled to the thermodynamics of adsorption of the mixture of carbon dioxide and nitrogen on solid adsorbents. The Toth equilibrium isotherm was considered, employing experimentally determined parameters. Initial studies indicate that less than four stages are sufficient for effective CO<sub>2</sub> removal. The mass ratio of solid to gas proved to be important for the adsorption process because the amount of pure solid present in the inlet stream favors the adsorption mechanics. In the future, a cost analysis will be performed, and it is expected that there should be a trade-off between the energy cost of heat exchanger and the cost of the adsorbent that will be used in the column. This happens because these factors will directly influence the percentage adsorbed, since in adsorption both the temperature of the stages and the S/G ratio are key factors in obtaining a high percentage adsorbed, while in desorption the energy cost will be the main factor to consider. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. We gratefully acknowledge support of the RCGI – Research Centre for Greenhouse Gas Innovation, hosted by the University of São Paulo (USP) and sponsored by FAPESP – São Paulo Research Foundation (2014/50279-4 and 2020/15230-5) and Shell Brasil, and the strategic importance of the support given by ANP (Brazil's National Oil, Natural Gas and Biofuels Agency) through the R&D levy regulation.

#### References

1. R. Morales-Ospino, et al., 2018, Assessment of temperature swing adsorption configurations for CO<sub>2</sub> capture on zeolite 13X, Dissertation (Masters), Universidade Federal do Ceará.
2. S.E. Zanco et al., 2018, Modeling of circulating fluidized beds systems for post-combustion CO<sub>2</sub> capture via temperature swing adsorption, *AIChE Journal*, v. 64, n. 5, p. 1744–175.
3. S. Neto et al., 2021, Calcium looping post-combustion CO<sub>2</sub> capture in sugarcane bagasse fuelled power plants. *International Journal of Greenhouse Gas Control*, v. 110, p. 103401.
4. Y. Wang et al., 2017, A Review of Post-combustion CO<sub>2</sub> Capture Technologies from Coal-fired Power Plants. *Energy Procedia*, v. 114, p. 650–665.
5. F. Dietrich et al., 2018, Experimental study of the adsorber performance in a multi-stage fluidized bed system for continuous CO<sub>2</sub> capture by means of temperature swing adsorption. *Fuel Processing Technology*, v. 173, p. 103–111.
6. R. Morales-Ospino et al., 2020, Assessment of CO<sub>2</sub> desorption from 13X zeolite for a prospective TSA process. *Adsorption* 26, 813–824.
7. R. Morales-Ospino et al., 2021, Parametric Analysis of a Moving Bed Temperature Swing Adsorption (MBTSA) Process for Postcombustion CO<sub>2</sub> Capture. *Industrial & Engineering Chemistry Research*, v. 60, n. 29, p. 10736–10752.

# Regional sustainability of food-energy-water nexus considering water stress using multi-objective modeling and optimization

Anupam Satyakam,<sup>a</sup> Rashi Dhanraj,<sup>a</sup> Yogendra Shastri,<sup>a</sup>

<sup>a</sup>*Indian Institute of Technology Bombay, Mumbai 400076, India*  
[yshastri@iitb.ac.in](mailto:yshastri@iitb.ac.in)

## Abstract

In this work, we developed a regional optimization model to achieve food-energy-water nexus sustainability in the context of water stress. An ethanol-based bioenergy system is developed. The model considers district-level water and land availability. The critical decision variables are the allocation of land area to chosen crops and the residue selection for ethanol production. The primary goals are to maximize farmers' profit and minimize water withdrawals while meeting the ethanol production targets. The resulting multi-objective mixed-integer linear programming problem was applied to a case study of 33 districts in Maharashtra, India. Results are compared for two cases, the land practices suggested by the optimization model and the existing practices. Upon reducing water consumption by 40.7% compared to the base case, only a 2.2% reduction in profits was observed from 229 \$/ha to 224 \$/ha. For ethanol water footprint minimization, 60% of available water consumption is found to be optimal where profits stand at 117.4 \$/ha.

**Keywords:** Water stress, food-energy-water nexus, regional sustainability, optimization.

## 1. Introduction

The security of food, energy, and water has been a subject of serious concern for the last few decades. These resources are being depleted at an alarming rate. The food, energy, and water nexus provides a strategy for achieving their security. According to United Nations projections, India's population will reach 1.5 billion people in 2030. This demand growth causes more than 60% of energy consumption (UN,2014). Water stress is present in 45% of the regions globally, with 18% of them experiencing high water stress and 9% experiencing severe water stress. In India, 600 million people face high to extreme water stress (UN-Water,2021). Additionally, the area under the crops will grow by 6% between 2015 and 2050, while the water withdrawn for the same will grow by 42% (IPCC, 2018). Particularly, in the semi-arid and arid regions like India, the competition for water resources between the production of food and energy can result from the unsustainable use and management of water resources. Therefore, water resources should be handled sustainably. In this study, we introduced a water stress constraint that would limit the consumption of water for irrigational purposes to a much lower value compared to the existing practices and simultaneously maintain the economic benefits. The bio-refinery model considered in this work was developed earlier in the literature (Dhanraj et al., 2021). Their work didn't take surface water as one of the sources for irrigation. Also, the rainfall recharge was neglected. This is not desirable as some crops rely on surface water for irrigation and rainfall recharge availability relaxes the stress on groundwater. To

overcome these shortcomings, changes have been made to this model to obtain optimal land patterns in the context of water stress.

## 2. Model description

A multi-objective optimization approach is used to build an ethanol-based bioenergy system. The objective of this study is to develop a bioenergy system that is limited by the availability of district-level land, regional water, and conflicting water demands. The choice of residues for ethanol production and the allocation of land areas to the chosen crops are the crucial choices. There are four objectives: minimization of ethanol cost, ethanol's water footprint, irrigation water, and farmers' profit maximization. The resulting problem was formulated as a mixed integer linear problem (MILP) and solved using the epsilon constraint method (EPCM). The optimization model, constraints, parameters, key variables, and objectives are represented in Figure 1.

The model considers the amount of water used by agricultural crops as well as the economic rewards to those involved in agriculture and ethanol production. For water balance modeling, the model uses the district as the base unit. The consumption and groundwater recharge for each district are then estimated based on the agricultural land use pattern. The water stress constraint is applied in each district for each season separately. The time frame of the model is one year, and is divided into three seasons: Kharif, Rabi, and Summer.

Important assumptions of this model:

- Inter-district groundwater recharge flow is ignored, which implies that the recharge occurring in a district is used in the same district.
- Ethanol production was based on the total transport fuel consumption and a particular petrol blending rate.
- Groundwater recharge during Kharif season was available for utilization during Rabi season and groundwater recharge during Rabi season was available for utilization during the summer.

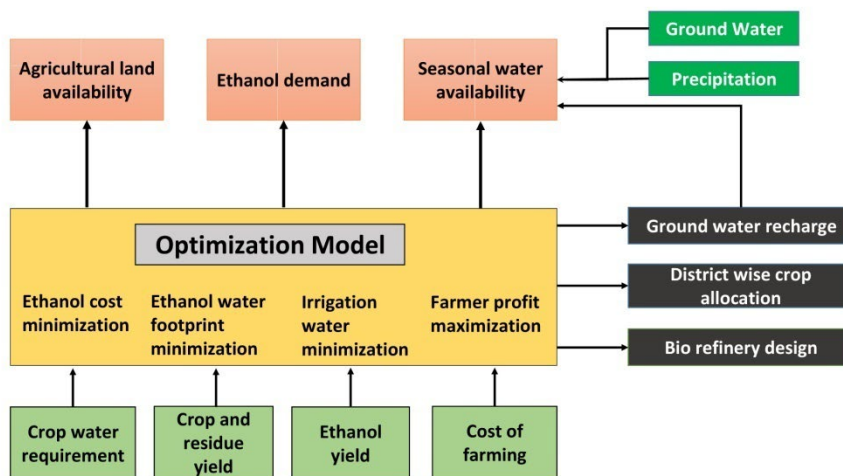


Figure 1: Schematics representing the scope of the model: orange, green, yellow, and grey boxes represent constraints, parameters, optimization model and the key variables, including final decisions, respectively (Adapted from Dhanraj et al. (2021)).

### 2.1. Optimization model formulation

The mixed-integer linear programming problem was formulated to solve the multi-objective optimization problem in General Algebraic Modeling System (GAMS) version 24.9.2 and solved using the CPLEX® MILP solver (Robichaud, 2010). This section discusses the constraints, decision variable and the objective functions of the model.

#### 2.1.1. Constraints

The amount of land allotted  $x(i, j, t)$  to various crops is the key decision variable in this model. This cannot exceed the amount of agricultural land available  $AL(i, t)$  in each district. Here index  $i, j$ , and  $t$  denotes the location of farm, crops and season, respectively. Agricultural land is often used for other purposes, such as cultivating fruits and vegetables. Therefore, a fraction of the land  $F_{AL}(i, t)$  is used for specific crops.

$$\sum_j x(i, j, t) \leq AL(i, t) \times F_{AL}(i, t) \quad \forall i, t \quad (1)$$

Rainwater or irrigation can be used to meet the crop's water needs. The rainfall absorbed by plants to account for crop water requirement is termed green water requirement or effective rainfall (ER), and the irrigation water (IR) used to meet the crop water requirement (CWR) is termed blue water. Irrigation water requirement is calculated as:

$$IR(i, j, t) = CWR(i, j, t) - ER(i, j, t) \quad \forall i, j, t \quad (2)$$

In the constraints related to water, the water balance for Kharif season is discussed first. Equation 3 represents the constraint on effective rainfall (green water)

$$\sum_j ER(i, j, t_k) \times x(i, j, t_k) \leq R(i, t_k) \times (1 - \zeta) \times AL(i, t_k) \times F_{AL}(i, t_k) - \sum_j GWR(i, j, t_k) \times x(i, j, t_k) \quad \forall i \quad (3)$$

Here, the LHS term represents the rainwater consumed by crop  $j$ , in the district  $i$  during the Kharif season  $t_k$ . On the righthand side, the first term is the total rainfall available after run-off consideration. The land fraction in this term ensures the rainfall is considered only within the area covered by the crops of interest.  $\zeta$  is the run-off fraction, assumed to be the same for all the districts. The second term indicates the groundwater recharge that is not utilized as rainwater source for crop utilization.

Irrigation water requirements for the crops can be met in two ways, the crops that use groundwater for irrigation and the others that use surface water. The resulting constraints are then modeled using equations 4 and 5.

$$\sum_j IR(i, j, t_k) \times x(i, j, t_k) \leq GW(i, t_k) \times F_{AL}(i, t_k) + \sum_j GWR(i, j, t_k) \times x(i, j, t_k) + GRO(i, t_k) \quad \forall i \quad (4)$$

$$\sum_j IRS(i, j, t_k) \times x(i, j, t_k) \leq SW(i, t_k) \times F_{AL}(i, t_k) \quad \forall i \quad (5)$$

Here,  $IR(i, j, t_k)$  denotes the groundwater needed by crop  $j$  in the district  $i$  during the Kharif season and  $IRS(i, j, t_k)$  denotes the surface water needed by crop  $j$  in the district  $i$  during the Kharif season. The RHS term in equation 4 is the sum of three terms, a fraction of total groundwater that is available for utilization of selected crops, groundwater getting recharged within the same season, and recharge through the land other than the selected crops. Similarly, in equation 5, the RHS term indicates the fraction of surface water available for the consumption of selected crops. After green water and blue water requirements are met, the net groundwater left in the Kharif season is represented through equation 6.

$$N_{GW}^K(i) = [R(i, t_k) \times (1 - \zeta) \times AL(i, t_k) \times F_{AL}(i, t_k) - \sum_j ER(i, j, t_k) \times x(i, j, t_k)] \\ + [(GW(i, t_k) \times F_{AL}(i, t_k) + GRO(i, t_k) \times F_{AL}(i, t_k)) \\ - \sum_j IR(i, j, t_k) \times x(i, j, t_k)] \quad \forall i \quad (6)$$

Here, the RHS denotes the difference between the sum of water available in the form of rainwater, groundwater, and surface water minus the total crop water requirement. Water balances for Rabi and summer seasons are modeled in a similar manner as that of Kharif. In the equations 4 and 6, for the Rabi season, the net groundwater left after the Kharif season is available for irrigation usage. Similarly, for the Summer season, the net groundwater left after the Rabi season is available for irrigation usage.

To reduce this over-dependence on groundwater, a groundwater stress constraint that puts a limitation on the total available water for agricultural purposes is included in the model. This constraint is imposed seasonally and for each individual district. Groundwater stress is defined as the ratio of demand for water to the available water, and it is an indicator of competition for water resources (Luck M. et al., 2015). Its value ranges from ‘0’ to ‘1’ where 0 indicates no availability of water for irrigation and 1 indicates entire water availability. The groundwater stress factor accounts for the groundwater stress level across each individual district. This can be mathematically written as:

$$g_k = \frac{\sum_j IR(i, j, t_k) \times x(i, j, t_k) + DIW(i, t_k)}{[GW(i, t_k) \times F_{AL}(i, t_k) + \sum_j GWR(i, j, t_k) \times x(i, j, t_k) + GRO(i, t_k) + SW(i, t_k) \times F_{AL}(i, t_k)]} \quad \forall i \quad (7)$$

The total available water for irrigational purposes is modeled from equation 10 as:

$$\sum_j IR(i, j, t_k) \times x(i, j, t_k) \leq g_k \times [GW(i, t_k) \times F_{AL}(i, t_k) + \sum_j GWR(i, j, t_k) \times \\ x(i, j, t_k) + GRO(i, t_k) + SW(i, t_k) \times \\ F_{AL}(i, t_k)] - DIW(i, t_k) \quad \forall i \quad (8)$$

This puts a constraint on the sum of blue water for the Kharif season. In the equation 7, numerator represents the sum of blue water requirement and water consumed by domestic and industrial sectors  $DIW(i, t_k)$  and the denominator represents the total available water. In the stress equations for the Rabi and Summer, water left after Kharif and Rabi will be added to the total available water.

### 2.1.2. Objective functions

There are four objectives in this multi-objective optimization model. The first objective is minimization of ethanol cost, and is formulated as follows:

$$\text{Min } J_c = \sum_{i,j,t} C_E(i, j, t) \quad (9)$$

Here,  $C_E(i, j, t)$  is the cost of ethanol production. Cost of ethanol includes the cost of residue, cost of transport, cost of ethanol production process and storage cost. The objective of maximization of profit earned by farmers is formulated as follows.

$$\text{Max } J_P = \sum_{i,j,t} P_F(i, j, t) \quad (10)$$

The profit earned by farmers  $P_F(i, j, t)$  is the difference between the price of agricultural product and residue to the cost of cultivation. The third objective, minimization of water footprint of ethanol, is formulated as follows:

$$\text{Min } J_{WF} = \sum_{i,j,t} W_{FR}(i, j, t) \tag{11}$$

Here,  $W_{FR}(i, j, t)$  is the water footprint of ethanol produced from different agricultural residue. The fourth objective, minimization of irrigation water consumption is formulated as follows:

$$\text{Min } J_{IR} = \sum_{i,j,t} IR(i, j, t) \times x(i, j, t) \tag{12}$$

Here,  $IR(i, j, t)$  is the irrigation water required by the crop  $j$  in  $\text{m}^3/\text{ha}$ , which is multiplied by land allocated to the crop. This single objective optimization problem is solved for various combinations of groundwater stress to obtain the trade-offs between different objective functions and find out the optimal water practices.

### 3. Case study of Maharashtra, India

The model has been applied to a case study of 33 districts covered under 12 river basins. Cotton, sorghum (both Kharif and Rabi varieties), wheat, rice, and sugarcane are the five crops considered in the case study. The area allocated to crops other than the five crops of interest is assumed to be constant. The groundwater recharge data for the Solapur district was calibrated for different crops. Based on the data for the Solapur district, the correlation between rainfall and groundwater recharge was obtained using the linear regression method. This correlation was used for all districts.

### 4. Results and discussion

Based on historical water use from the years 2004 to 2020, the degree of groundwater stress in various districts was computed. It was observed that the number of districts moving towards high water stress has increased from 17 districts (58.6%) in 2004 to 23 districts (67.6%) in 2020.

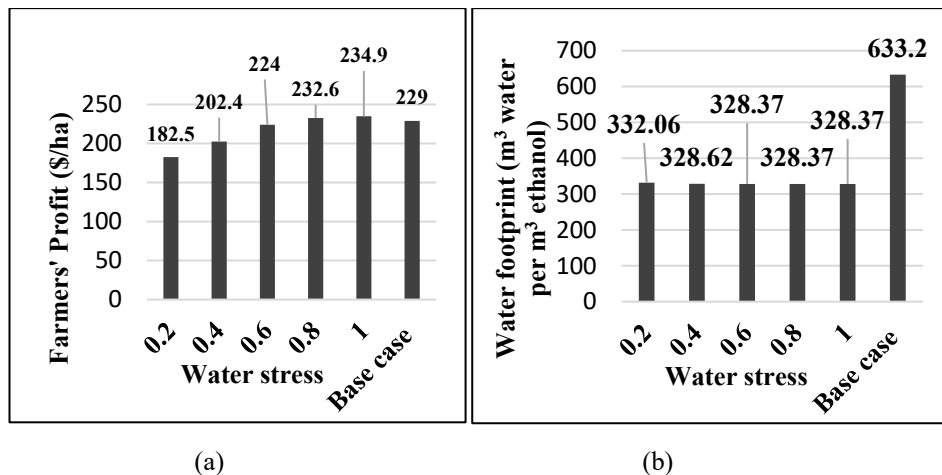


Figure 2: Effect of groundwater stress factor on two objectives (a) Maximum farmers' profit and (b) Minimum ethanol water footprint.

Profit increased with the rising groundwater stress factor, or higher water availability, as shown in figure 2(a). This can be explained by the rise in the amount of land allocated for rice and wheat, as well as the overall acreage. The stress factor was also calculated for the current land use practices and its value was found to be 1.07. As compared with the base case, for a stress factor of 0.6, a reduction in 40.7% of water consumption, profit only reduced by 2.2% and ethanol cost had a least value of 0.71 \$/lit. ethanol. From figure 2(b), water footprint value for 0.6 stress factor is almost half (328.3) that of the base case (633.2). Also, farmers' profit has a value of 117.4 \$/ha and ethanol cost has a value of 0.69 \$/lit. ethanol.

## 5. Conclusion

In this work, we tried to reduce the water consumption for irrigation and simultaneously maintain economic benefits. We quantified the trade-offs between profits and water withdrawals. In the future, the model can be used to capture the food production aspects and also consider the impacts of climate change on decision-making. Important observations from this study are as follows:

- In case of reduction in water consumption, more land should be allocated to sorghum and less land should be allocated to rice, cotton, and sugarcane.
- On comparing the current land use practices and the land allocation suggested by the model for reduced water consumption, it is evident that there is a need to change the cropping pattern.
- As the change in cropping pattern is farmers choice, government should provide incentives for farmers to shift to less profitable crops.
- Decision makers should make a choice on the water stress factor as increase in profit comes at an expense of increased water consumption.

## References

- The United Nations World Water Development Report. 2014. Water and Energy, Paris, UNESCO, Volume 1, ISBN: 978-92-3-104259-1.
- The United Nations World Water Development Report, 2021 SDG 6 Water and sanitation for all, Geneva, Switzerland.
- IPCC, Global Warming of 1.5 °C. Special Report, Intergovernmental Panel on Climate Change, Geneva, 2018
- Dhanraj, R., Punnathanam, V., Shastri, Y. 2021. "Multi objective optimization of ethanol production based on regional resource availability". *Sustainable Production and Consumption*, 27, 1124-1137.
- Robichaud, V., 2010. An Introduction to GAMS 1. Matrix 1–14.
- Luck, M., M. Landis, F. Gassert. 2015. "Aqueduct Water Stress Projections: Decadal Projections of Water Supply and Demand Using CMIP5 GCMs." Technical Note. Washington, D.C.: World Resources Institute. Available online at: [wri.org/publication/aqueduct-water-stress-projections](http://wri.org/publication/aqueduct-water-stress-projections)

# Toluene as effective LOHC: detailed techno-economic assessment to identify challenges and opportunities

Elvira Spatolisano<sup>a</sup>, Angelo Matichecchia<sup>a</sup>, Laura A. Pellegrini<sup>a</sup>  
Alberto R. de Angelis<sup>b</sup>, Simone Cattaneo<sup>b</sup>, Ernesto Roccaro<sup>b</sup>

<sup>a</sup>*GASP - Group on Advanced Separation Processes & GAS Processing, Dipartimento di Chimica, Materiali e Ingegneria Chimica “G. Natta”, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy*

<sup>b</sup>*Eni S.p.A. Research and Technological Innovation Department, via F. Maritano 26, 20097 San Donato Milanese, Italy*

[elvira.spatolisano@polimi.it](mailto:elvira.spatolisano@polimi.it)

## Abstract

Liquid Organic Hydrogen Carriers (LOHC) are organic molecules that can be reversibly hydrogenated and dehydrogenated to release H<sub>2</sub>. Due to their easy transport and adaptability to the existing infrastructures, they are gaining attention in research panorama. However, despite their flexibility, LOHC are not yet exploited as large-scale energy storage/transport medium due to the too high costs. In this respect, this work is aimed at performing a detailed techno-economic assessment of H<sub>2</sub> distribution through toluene as LOHC. Several scenarios are discussed, at variable distance covered via ship transport together with utilities and raw materials costs, to identify the weaknesses of the overall value-chain and know where to focus research efforts to improve the existing technologies.

**Keywords:** LOHC, hydrogen distribution, hydrogen storage, techno-economic assessment.

## 1. Introduction

To address climate change and achieve the decarbonization target, power generation from fossil fuels has to be limited in favor of an increase in the shares of renewable energies (Pellegrini et al., 2018). The intermittent and seasonal nature of these sources has raised issues on their effective application, also considering that they are often produced in remote sites. In this panorama, where energy storage and transport solutions are needed, hydrogen is a key player supporting the indirect electrification (Hurskainen, 2019). For large-scale distribution and utilization of hydrogen as an energy vector, its safe, efficient and economical storage and transportation is needed (Aakko-Saksa et al., 2018). Among the various alternatives for H<sub>2</sub> transportation and storage, a promising one, which has gained increased attention in the last years, is the one based on the so-called liquid organic hydrogen carriers (LOHC), which can be reversibly hydrogenated and dehydrogenated. The hydrogenation involves the chemical bonding of hydrogen to the organic liquid. In this way, the hydrogenated compound can be transported at atmospheric pressure like many other oil-like substances. At the destination, the hydrogen is released via an endothermic dehydrogenation reaction (Roland Berger, 2021). In most of the proposed supply chains, the dehydrogenated LOHC is intended to be transported back to the



hydrogen source for reuse. Examples of LOHCs include toluene, dibenzyltoluene or other N-substitute heterocycles. Among them, toluene is the most mature option, used in a demonstration project by Chiyoda Corporation (Naruke, 2017).

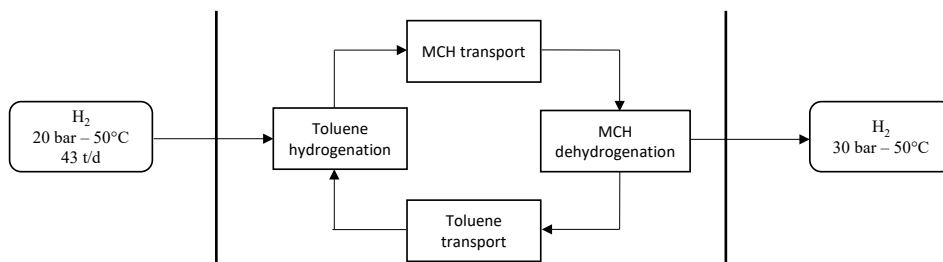
Despite the high technology readiness level, no detailed analysis of the overall value-chain of hydrogen transport through toluene is available in literature. To fill this gap and to understand the opportunities and challenges of the LOHCs, this work focuses on an in-depth techno-economic assessment of hydrogen transport considering toluene as LOHC medium. When possible, a detailed Aspen Plus® simulation is carried out for evaluating process performances. For each section of the value-chain (i.e., storage at the loading terminal, hydrogenation, transportation, dehydrogenation, storage at the unloading terminal, distribution to the end-users), drawbacks and weaknesses are pointed out for the identification of future research developments. Several scenarios are discussed, depending on utilities, raw materials and carbon tax costs. A sensitivity analysis as a function of the harbor-to-harbor distance is performed, to cover different situations. Due to the encouraging outcomes of the technology assessment and the easy adaptability to the existing infrastructures, the potential of LOHC is demonstrated, proving that they can become a feasible hydrogen distribution choice.

## 2. Toluene value-chain

H<sub>2</sub> distribution through LOHC is typically based on a reversible hydrogenation and dehydrogenation of carbon double bonds. The system boundaries for toluene value-chain are reported in Figure 1. In our study, 43 t/d of H<sub>2</sub> from alkaline electrolyzers at 20 bar and 50 °C enter the process battery limits at the departure site. Hydrogen is reacted together with toluene into the hydrogenation stage, to produce methylcyclohexane (MCH). The process is exothermic and takes place at elevated temperatures and pressures. After being produced, MCH is stored in tanks and shipped from the loading terminal to its final destination (unloading terminal), through cargo powered with traditional fossil fuels. A sensitivity analysis as a function of the distance covered via ship in the range 2500 – 10000 km has been performed.

When arrived at the unloading terminal, MCH is routed via truck to the H<sub>2</sub> valley, far 100 km from the unloading terminal. Here, MCH undergoes a catalytic endothermic dehydrogenation reaction, operated at high temperatures and low pressures, and H<sub>2</sub> is released, with a purity of 99.97% (ISO 14687:2019) and a pressure of 30 bar, to be collected into the H<sub>2</sub> valley. For the dehydrogenation process, heat is provided through the combustion of part of the H<sub>2</sub> produced.

The hydrogenation and dehydrogenation stages have been simulated through Aspen Plus V11® simulation software, to retrieve heat and material balances and, thus, evaluate the associated fixed and operating costs.



**Figure 1.** Toluene value-chain.

### 2.1 Toluene hydrogenation

The developed process scheme for toluene hydrogenation, together with the specifications of inlet and outlet streams from the process battery limits, is reported in Figure 2. Toluene from the unloading terminal, *TOL* in Figure 2, which contains benzene impurities from the dehydrogenation section, is mixed with the make-up toluene, *MAKEUP* in Figure 2, to recover unavoidable losses in the overall value-chain. The mixed stream *TOL-MIX* is pumped up to 20 bar and, after preheating, is routed to *REACTOR*, together with H<sub>2</sub> and N<sub>2</sub>. The inlet nitrogen acts as thermal diluent, being the hydrogenation reaction exothermic. The *REACTOR* operates at  $P = 20$  bar and  $T = 210$  °C in the presence of a Pt-based catalyst. Due to the lack of experimental data and a well-established kinetic expression, the reactor has been simulated through the *RGibbs* module of Aspen Plus. Thus, the toluene conversion is the thermodynamic equilibrium one at the fixed  $T$  and  $P$ . Steam is generated to benefit from the reaction exothermicity. Downstream the reactor, methylcyclohexane is purified from the heavy byproducts eventually produced from side reactions. Unconverted hydrogen, fed in excess of the stoichiometric, is recycled back to the reactor, after compression up to 20 bar, while the MCH produced is collected for storage and transport to the unloading terminal.

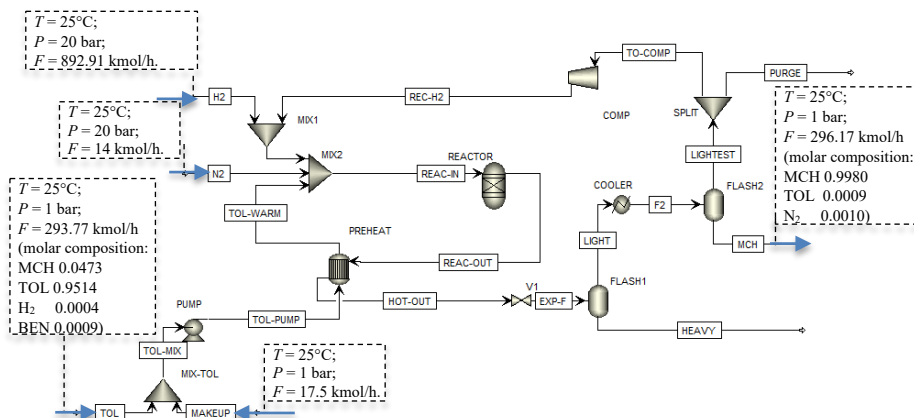


Figure 2. Toluene hydrogenation. Simulation in Aspen Plus® V11.

### 2.2 MCH dehydrogenation

Figure 3 reports the flowsheet of the toluene dehydrogenation. Methylcyclohexane from the hydrogenation section is pumped, preheated and routed to the dehydrogenation reactor (*REACTOR* in Figure 3). The dehydrogenation reaction, strongly endothermic, is performed at  $P = 3$  bar and  $T = 350$  °C in the presence of a Pt-based catalyst. The dehydrogenation reactor is simulated as a black box with fixed conversion, whose value is in line with the literature one. The reaction products, toluene and hydrogen, are separated in a series of flash units downstream the reaction section. H<sub>2</sub> is compressed up to 30 bar, the pressure needed for its destination to H<sub>2</sub> valley, and is recovered with a purity suitable for its commercialization, while toluene can be sent back to the hydrogenation section. Part of the produced hydrogen (297.60 kmol/h of the 843.39 kmol/h shown in Figure 3) is burned to cope the reaction endothermicity

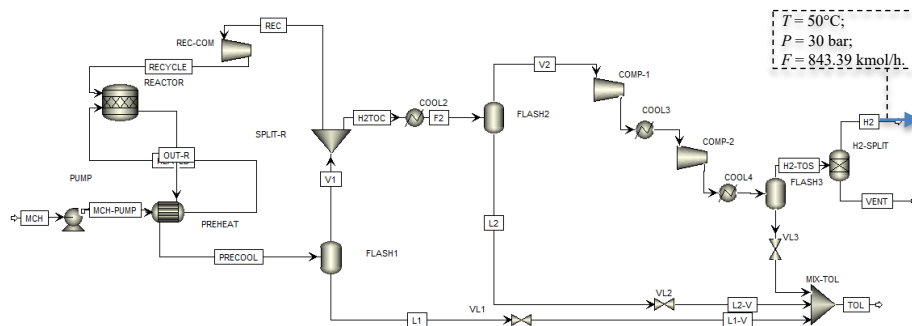


Figure 3. Methylcyclohexane dehydrogenation. Simulation in Aspen Plus V11<sup>®</sup>.

### 3. Methodology

The toluene pathway described in section 2 has been analysed through a techno-economic assessment, to identify the cost-drivers of the whole value-chain.

Fixed and operating costs for the hydrogenation and dehydrogenation sections have been evaluated according to the Turton methodology (Turton et al., 2008).

For fixed costs, the purchased base cost ( $C_{p,i}^0$ ) is calculated for each piece of equipment of Figure 2 and Figure 3, through equation (1), where  $K_{1,i}$ ,  $K_{2,i}$  and  $K_{3,i}$  are specific constants available for each equipment type and  $A_i$  represents the equipment capacity.

$$\log_{10} (C_{p,i}^0 (2001)) = K_{1,i} + K_{2,i} \log_{10} (A_i) + K_{3,i} [\log_{10} (A_i)]^2 \quad (1)$$

The purchased base cost, referred to 2001 year, is actualized through the 2022 Chemical Engineering Plant Cost Index (CEPCI), which is CEPCI (2022) = 825 (https://chemengonline.com/). Equation (2) allows the calculation of the bare module cost ( $C_{BM,i}$ ), to account for the equipment material of construction and operating pressure.  $B_{1,i}$ ,  $B_{2,i}$ ,  $F_{M,i}$  and  $F_{P,i}$  are pressure and material dependent constants.

$$C_{BM,i} = C_{p,i}^0 F_{BM,i} = C_{p,i}^0 (B_{1,i} + B_{2,i} F_{M,i} F_{P,i}) \quad (2)$$

The total module cost ( $C_{TM}$ ) is evaluated from the bare module cost according to equation (3) and the grassroots cost according to equation (4), where  $C_{BM,i}^0$  is the bare module cost of each equipment evaluated at atmospheric pressure and with carbon steel as the material of construction. The Fixed Capital Investment ( $FCI$ ) is retrieved as equation (5).

$$C_{TM} = 1.18 \sum_{i=1}^n C_{BM,i} \quad (3)$$

$$C_{GR} = 0.5 \sum_{i=1}^n C_{BM,i}^0 \quad (4)$$

$$FCI = C_{TM} + C_{GR} \quad (5)$$

As regards the operating costs, they include the Direct Manufacturing Costs ( $DMC$ ), the Fixed Manufacturing Costs ( $FMC$ ) and the General Expenses ( $GE$ ), all evaluated as in Turton (2008). The cost of waste treatment ( $C_{WT}$ ), as well as the depreciation, has been neglected in the present analysis, while all the assumptions introduced in the economic evaluations are listed in Table 1.

Two different scenarios (actual and future) have been considered, at variable cost of utilities, raw materials and CO<sub>2</sub> emissions. In the future scenario (in 5 years) a cost reduction is foreseen, after the huge 2022 inflation. When possible, the hourly utility

consumption has been calculated through the energy balance retrieved from process simulations.

**Table 1.** Assumptions introduced in the economic assessment.

Assumption	value	
payback time	10 years	
operator yearly wage	40000 €/y	
plant operability	8000 h/y	
exchange rate (2022)	0.915 €/€	
<b>Utilities cost</b>		
	<i>actual</i>	<i>future</i>
cooling water (30 to 40°C)	0.015 €/t	0.015 €/t
MP steam (200°C, 15 bara)	130 €/t	50 €/t
boiler feed water	1.15 €/t	1.20 €/t
electric energy	500 €/MWh	220 €/MWh
<b>Raw materials cost</b>		
	<i>actual</i>	<i>future</i>
toluene	1300 €/t	850 €/t
nitrogen	0.2 €/Nm <sup>3</sup>	0.15 €/Nm <sup>3</sup>

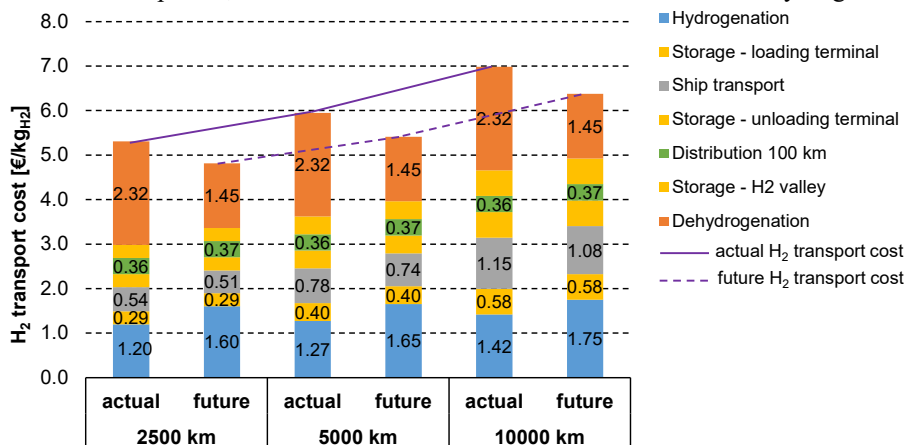
Costs of harbor infrastructures at both the loading and unloading terminals have been neglected, as well as catalysts costs, spare units and indirect CO<sub>2</sub> emissions.

The *FCI* has been divided for the payback time and added to the operating expenses. Their sum was then divided for the H<sub>2</sub> production to retrieve the H<sub>2</sub> transport cost [€/kg<sub>H2</sub>].

As regards ship transport (<https://compassmar.com>), storage (IEA, 2019) and distribution (Teichmann et al., 2012), costs are retrieved from typical oil products applications.

#### 4. Results and discussion

Results for the economic assessment performed are reported in Figure 4 considering the different distances covered through ship transport (2500-5000-10000 km) and the actual and future scenarios. H<sub>2</sub> transport costs increase as the distance covered via ship increases: in this case, a higher fuel consumption and a higher fixed cost of the tanker is observed. As expected, the cost-driver of the overall value-chain is the dehydrogenation.



**Figure 4.** Results of the economic assessment at variable ship distances. Actual and future scenarios.

Figure 5 reports the CAPEX and OPEX for the value-chain cost driver (MCH dehydrogenation). Costs are almost entirely due to operating expenses, whose most-significant cost-item is utilities consumption. Future research has to be focused for sure on the reaction section, to identify active catalysts such that the reaction can be efficiently performed at lower temperatures.

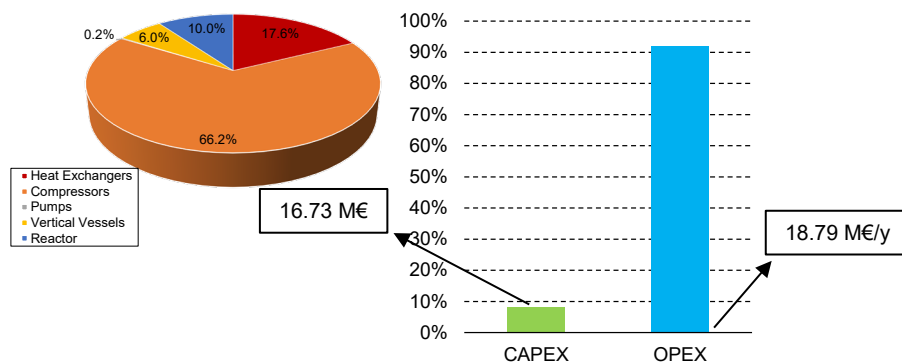


Figure 5. CAPEX and OPEX for MCH dehydrogenation stage.

## 5. Conclusions

Nowadays, exploiting the existing installations, only a small amount of the H<sub>2</sub> produced in Europe is transported: it is rather produced onsite from natural gas via steam reforming. The techno-economic assessment presented in this work shows that H<sub>2</sub> transport via LOHC can be economically more attractive in the near future, considering the increasing cost associated to CO<sub>2</sub> emissions together with the expected higher H<sub>2</sub> demand.

## References

- A. Ajanovic, M. Sayer, R. Haas, 2022, The economics and the environmental benignity of different colors of hydrogen. *International Journal of Hydrogen Energy* 47, 24136-24154. <https://chemengonline.com/>. Accessed: October 2022.
- <https://compassmar.com>. Accessed: October 2022.
- IEA G20 Hydrogen report: Assumptions. 2019.
- ISO 14687:2019. Hydrogen fuel quality — Product specification.
- Naruke, 2017, The world's first global hydrogen supply chain demonstration project. [https://www.chiyodacorp.com/meida/170727\\_e.pdf](https://www.chiyodacorp.com/meida/170727_e.pdf).
- P. T. Aakko-Saksa, C. Cook, J. Kiviaho, T. Repo, 2018, Liquid organic hydrogen carriers for transportation and storing of renewable energy – Review and discussion. *Journal of Power Sources* 396, 803-823.
- M. Hurskainen, 2019, Liquid organic hydrogen carriers (LOHC): Concept evaluation and techno-economics. VTT Technical Research Centre of Finland. Research Report No. VTT-R-00057-19Y.
- Roland Berger, 2021, Hydrogen transportation | The key to unlocking the clean hydrogen economy. ROLAND BERGER GMBH, Sederanger 1, 80538 Munich, Germany.
- R. Turton, R.C.Bailie, W.B.Whiting, J.A. Shaeiwitz, 2008, Analysis, Synthesis and Design of Chemical Processes. Pearson Education.
- L. A. Pellegrini, G. De Guido, S. Langé, 2018, Biogas to liquefied biomethane via cryogenic upgrading technologies. *Renewable Energy*, 124, 75-83.
- D.Teichmann, W.Arlt, P. Wasserscheid, 2012, Liquid Organic Hydrogen Carriers as an efficient vector for the transport and storage of renewable energy. *International Journal of Hydrogen Energy*, 37, 18118-18132.

# Evaluation of Methane Mitigation Technologies in the Upstream Oil and Gas Sector using the TEAM Digital Platform

Marcelo Mathias, Ahad Sarraf Shirazi, Jairo Duran, Alberto Alva-Argaez<sup>a</sup>

<sup>a</sup> *Process Ecology Inc., 401 – 301 14 St. NW, Calgary, AB T2N 2A1, Canada*  
[alberto@processecology.com](mailto:alberto@processecology.com)

## Abstract

Greenhouse gas emissions from upstream operations account for most of the methane emitted in the oil and gas industry in Canada. To achieve its national emissions reduction target, Canada intends to make significant investments, however there is a need to design specific mitigation strategies at a regional level to implement cost-effective mitigation technologies. In this work, a detailed mitigation techno-economic analysis for a subset of oil and gas facilities in Alberta was performed using the TEAM digital platform. The inputs of the model include production levels and vented and flared volumes for each facility obtained from publicly available data. The analysis shows that 14.8 Mtpa CO<sub>2e</sub> of carbon emissions can be removed from the atmosphere per year at marginal abatement costs of \$50/tCO<sub>2e</sub> or below. Regional differences in production and facility types suggest that gas conservation technologies such as tie-in to gathering systems or onsite power generation are well suited for facilities with large reported vented emissions, whereas sites in eastern Alberta that predominantly report fugitive emissions can benefit more from facility upgrades and leak detection and repair programs (LDAR). About 70% of methane emissions in the selected facilities can be mitigated at nearly zero abatement cost.

**Keywords:** methane mitigation, carbon abatement, marginal abatement cost curve, upstream emission mitigation, carbon credits.

## 1. Introduction

Methane is responsible for approximately 30% of the global rise in temperatures to date and accounts for approximately 13% total greenhouse gas emissions in Canada (UNEP, 2021; ECCC, 2021). The Canadian government is committed to ambitious climate action, including a new target to reduce methane emissions in the oil and gas sector by at least 75% by 2030 and capping and reducing emissions from the oil and gas sector (O&G). Cutting methane emissions is one of the fastest and lowest-cost ways to fight climate change and is one of the top recommendations made by climate scientists (IPCC, 2022).

Several publications have assessed the technical and economic feasibility of implementing mitigation technologies in the sector, especially in the upstream segment which contributes most of the emissions. Proven technologies to tackle this problem tend to concentrate on gas conservation or destruction via combustion. For example, some studies have shown the impact and costs of flaring stranded gas, leak detection and repair

(LDAR) programs, replacing pneumatic pumps, and increased frequency for replacing rod packing seals in reciprocating compressors. The cost of these technologies is usually assessed over a certain time duration and is expressed as a marginal abatement cost (MAC), typically in dollars per tCO<sub>2</sub>e abated. The viability of these mitigation projects usually depends on site-specific factors, technologies, commodities, labor prices, among others. The most viable projects have MAC close to zero or even return profit over its duration (negative MAC). For the thousands of facilities that comprise the upstream O&G in Alberta (oil batteries, wells, etc.), the evaluation of baseline GHG emissions, the potential mitigated emissions achieved with a set of technologies, and the economic evaluation of each of these is time consuming. In this work, we used the TEAM (Techno-Economic Analysis Model) digital platform developed by Process Ecology Inc. to automate all calculations and perform sensitivity analysis of key technical and economic parameters for a collection of oil and gas facilities in Alberta.

In this study, a total of 7,177 installations were selected, including oil batteries, gas gathering systems, and gas plants. This selection of sites generated nearly 60% of reported provincial GHG emissions in the sector in 2020 (the year of the data selected). Volumetric emissions data from publicly available sources and engineering assumptions were used to generate a detailed inventory of emission sources, estimate the emissions profile of each facility, and then perform a detailed feasibility and economic assessment of selected methane mitigation technologies.

## 2. Methodology

### 2.1. Oil and Gas Facility Data

Volumetric production and flaring data for 2020 were retrieved from Petrinex (Petrinex, 2022) database. Of the total 7,177 facilities, 5,955 are oil batteries (BT), 861 gas gathering stations (GS), and 361 gas plants (GP). To avoid including non-routine large flaring events, median values for the reported vented and flared volumes were considered. The facilities were organized into subregions, as geographical location provides indications of fluid properties, facility type and number of equipment or emissions control methods used in each site. The PSAC (PSAC, 2022) and heavy oil migration classifications (AER, 2020) were applied here and the Montney region was added separately due to the large number of hydraulic fracturing projects in the area. Figure 1 shows the regional distribution considered. It is worth noting that the AB4 (Lloydminster area) is characterized by heavy oil production, whereas the AB6 region has the largest portion of oil sands production in the province. Based on location, a fluid characterization (heavy or light/medium oil) was derived for each facility. The PetroNinja database (Stack Technologies, 2022) was used to sample “typical” oil and gas compositions in each region. The fluid type impacts how the facility is designed and operated and ultimately determines the type of equipment and emissions sources that are likely to be found.

### 2.2. Distance to Critical Infrastructure

In this study, critical infrastructure refers to gas gathering pipelines, electricity grid, and qualified labor pools. This information is required to determine the feasibility and costs of implementation of some key mitigation technologies. Geospatial pipeline data from The Alberta Energy Regulator was loaded into TEAM to calculate distances for all facilities. A total of 120,226 pipelines were considered as part of the existing gas gathering network (AER, 2020). Each pipeline was discretized into segments which represent potential tie-in options for each facility. For each site, the shortest distance to a tie-in point was determined applying the KD-Tree search algorithm (Skrodzki, 2019). The

cost of tie-in depends on distance and compression needed to bring the recovered gas to the necessary tie-in pressure. The search algorithm was expanded to allow the identification of the closest pipeline with low pressure. Information for the electricity grid distribution in Alberta is sparse, therefore it was assumed that distances from facilities to a grid connection follow a Poisson distribution. For the distance to labor, the previously mentioned KD-Tree search algorithm was applied considering only population centers of more than 12,000 inhabitants in Alberta.

### *2.3. Facility Configuration*

TEAM relies on a detailed bottom-up model of the emissions sources present at each facility. Since this information is usually not publicly available, TEAM was equipped with “autoconfigure” rules that help to fill the gaps in information and determine the number and type of equipment present at each facility. The resulting process configuration depends on the type of fluid and production rates of each facility and was validated through conversations with SMEs, operators, and after inspection of process flow diagrams of representative facilities. For example, heavy oil batteries were assumed to contain a line heater, a flare stack (if the facility reported flare emissions), production tanks for 7 days of production, and a glycol dehydrator if gas throughput exceeded 0.5 MMSCFD. A set of process design rules were derived for each facility subtype and the selected equipment counts depend on the production rates. Pneumatic devices and fugitive component counts depend on the types of equipment assigned to each facility and were obtained from industry literature (Clearstone, 2018). Additional details and assumptions for other facility types can be found elsewhere (Process Ecology, 2020).

### *2.4. Other Considerations*

Fugitive emissions from each main equipment type (*i.e.*, reciprocating compressors, centrifugal compressors, glycol dehydrators, etc.) are calculated in TEAM using the equipment inventory determined in the autoconfiguration step, and fugitive components and associated emission factors as reported in the literature (IPCC, 2006; Clearstone, 2018). To reflect recent efforts in replacing high-bleed pneumatic controllers, it was assumed that 75% of the devices were low-bleed. The selected sites capture approximately 60% and 70% of total reported volumes for flaring and venting, respectively. From the collection of facilities studied here, the region identified as AB4 contains most of the heavy oil batteries (1,566) and the Montney region has the largest number of gas plants (120). This is expected to impact the most economical mitigation strategies by region.

### *2.5. The TEAM Digital Platform*

The digital platform allows for easy upload of the input data and automates the calculation of GHG emissions and mitigation strategies. TEAM calculates emissions profiles for each facility (vented, flared, fugitive, and combustion), as well as a marginal abatement cost curves (MACC) for single sites and for any subset of facilities. For each technology assessed, the program estimates the anticipated emission reductions and the marginal abatement cost (*i.e.*, \$/tCO<sub>2</sub>e). For each mitigation project (*i.e.*, the implementation of a given technology at a given site), the tool produces an estimated cash flow for the lifetime of the project (assumed to be 10 years). Net present value (NPV), return of investment (ROI), and payback period (PB) are reported as quantitative indicators of the performance of each project.



### 3. Results and Discussion

A techno-economic analysis model was configured with over 7,000 facilities that include oil batteries, gas plants, and gas gathering systems. The model was then solved to calculate emission profiles and mitigation strategies for each facility. It is worth noting that the reported volumetric venting and flaring data were used to configure the facilities but are different to the calculated values. The model estimates 14.8 Mtpa CO<sub>2</sub>e including carbon dioxide and methane, and 233,906 ton/year of methane emissions (GWP of 25 was applied for methane). Total venting emissions were estimated at 1.04 Mtpa CO<sub>2</sub>e and 1.3 Mtpa of flaring emissions. Pneumatic devices represent 1.8 Mtpa CO<sub>2</sub>e while fugitive emissions contribute an additional 1.82 Mtpa CO<sub>2</sub>e. The remaining 8.79 Mtpa CO<sub>2</sub>e corresponds to fuel gas combustion.

Oil batteries represented 40% of the total GHG emissions, gas plants and gas gathering stations contributed 34% and 26%, respectively. About 60% of total GHG emissions originate from fuel gas combustion. Pneumatic devices and fugitives contribute about 30% each to methane emissions.

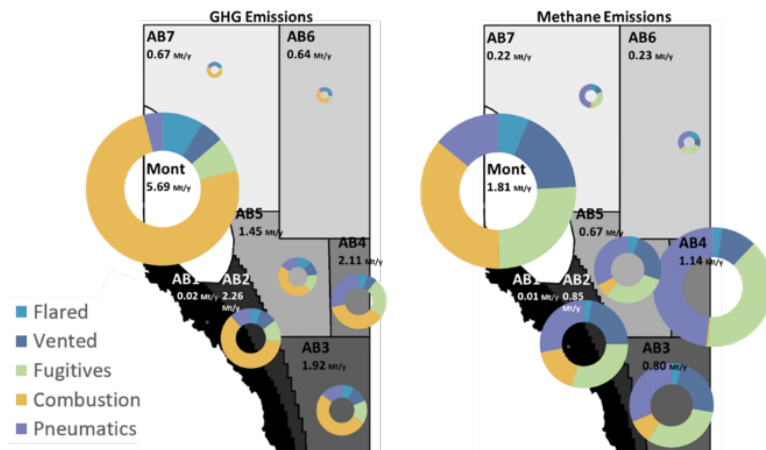


Figure 1. GHG and methane emissions per region calculated in TEAM for a subset of facilities in Alberta.

The Montney region represented the largest emitting area having almost double the emissions of the next largest region AB2 (5.7 vs 2.3 Mtpa CO<sub>2</sub>e). Most of the emissions from Montney originate from combustion sources, which is related to the large number of gas plants present in the region. In contrast, pneumatic and fugitive emissions in region AB4 have the largest contribution. A reasonable agreement between the model results and Petrinex data was found for flaring emissions. The greater deviations occurred for venting emissions, further investigation would be required to explain these differences, however it is likely that vented volumes have not been consistently reported by all operators.

TEAM provides economic indicators for each potential project and for all facilities. Conventional MAC curves display all possible technologies and their potential for emissions reductions, in general several mitigation technologies would act to conserve the same emissions source, *i.e.*, mitigation projects that are mutually exclusive. A function was implemented to eliminate these competing projects by selecting the lowest MAC project at each facility. After this selection, a total of 19,559 mitigation projects

were assessed by TEAM, resulting in potential for total emissions reduction of 4,6 MtCO<sub>2</sub>e/y for an estimated capital investment of \$222 million (USD). This represents a 75.4% emissions reduction compared to the business-as-usual scenario (BAU). Figure 2 presents the MACC for the maximum reduction scenario (after screening out mutually exclusive technologies). About 72.5% of potential mitigation can be achieved at zero or negative marginal cost making most of these projects profitable. Around 66% of projects were projected to cost \$0 USD/tCO<sub>2</sub>e or cheaper. It is worth noting that in our analysis the abatement cost considers the NPV of the projects and the carbon tax \$15 CAD/tCO<sub>2</sub>e is assumed to be a revenue stream for the project (tax credit or trading of carbon credits in the market). Subject matter experts advised that projects under \$50 USD/tCO<sub>2</sub>e are worth pursuing given current conditions. For this MAC threshold, 14% of the assessed mitigation projects are left out and the remaining projects provide an emissions reduction potential of 4.53 MtCO<sub>2</sub>e with a capital investment of \$129.1 MMUSD. The largest opportunities are the replacement of chemical pumps with solar or electric pumps (27% of total reductions) and installing vapor recovery units (23%).

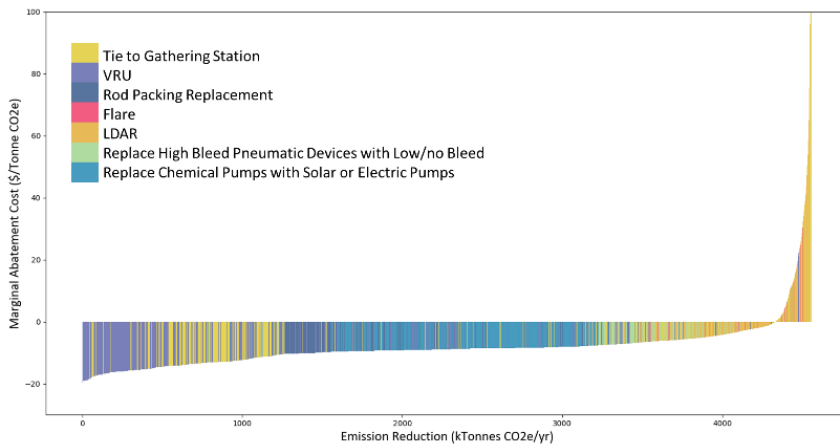


Figure 2. Maximum reduction MACC analysis for individual mitigation projects.

Figure 2 shows the distribution of ranked emission reduction projects. Replacing chemical pumps with solar or electric pumps accounts for 27% of total emissions reduction below \$50/tCO<sub>2</sub>e. This technology displays a negative average marginal abatement cost, meaning the installation of these is profitable. Chemical pump replacement and VRU installations represent the majority of the selected projects, several of which are economical to implement (MAC below \$0 USD/tCO<sub>2</sub>e). The option of tying in facilities to nearby gas gathering systems also appears promising with fewer projects delivering significant emissions reductions. The average tie-in to gathering station project reduces 1.09 ktCO<sub>2</sub>e while chemical pump replacement and VRUs have an average emission reduction of 0.27 and 0.67 ktCO<sub>2</sub>e per facility, respectively.

The Montney region contains the largest reduction potential (1.13 MtCO<sub>2</sub>e). Region AB4 has a similar reduction potential of 0.94 MtCO<sub>2</sub>e. However, these two regions present different mitigation strategies. While the Montney region would derive most reductions from installing VRUs, the AB4 region obtains its reductions from replacing chemical pumps. This aligns with the type of operations found in each region. Interestingly, only about a third of the emissions reduction potential are attained with technologies currently

addressed by regulations (LDAR programs, replacement of high-bleed devices, and rod packing replacement). Initially, the analysis did not consider onsite power generation as literature usually reports that this technology is not economical. With the limited data set and high uncertainty, it should be noted that power generation projects displace both VRU and tie-in to gathering projects in multiple sites, as this is a more economic option in some circumstances given the additional revenue from selling power to the grid or to neighbouring facilities. Power generation is most promising for larger facilities with large volumes of vented gas and in proximity to connection points to the grid.

#### 4. Conclusions

The results of the techno-economic assessment performed in the TEAM digital platform indicate that there are significant benefits in applying differentiated mitigation strategies on a regional basis as opposed to a province-wide approach. Fugitive emissions and pneumatic devices remain a significant source of emissions from the sector as well as uncombusted methane from natural gas engines. There are regulations in place for these sources, however there may be still a significant potential for further emissions reduction by increasing the adoption of gas conservation technologies (VRU, tie-in to gathering, onsite power generation, etc.). Similarly, the application of mitigation technologies to small and remote facilities with low gas venting volumes resulted in the most expensive abatement costs. Most emissions reductions can be achieved with projects below \$50 CAD/tCO<sub>2</sub>e, with about two thirds of this potential achieved at a profit. It must be noted that the mitigation technology costing information used in this study does not consider sub regional differences and would need further refinements. As such these results should be construed as preliminary indications of the total cost of mitigation from the sector. Even though a rigorous uncertainty analysis is still necessary to further validate these results, the regional analysis still suggest that a single mitigation strategy province-wide may not be optimal, but rather certain regions would have better mitigation results per dollar invested using a particular set of technologies.

#### Acknowledgments

The authors would like to acknowledge the support provided by Natural Resources Canada and Clearstone Engineering in providing financial support and costing information.

#### References

- United Nations Environment Programme and Climate and Clean Air Coalition, 2021, Global Methane Assessment: Benefits and Costs of Mitigating Methane Emissions, Nairobi, UNEP.
- Environment and Climate Change Canada (ECCC), 2021, Review of Canada's Methane Regulations for the Upstream Oil and Gas Sector, Gatineau, QC.
- Intergovernmental Panel on Climate Change (IPCC), 2022, Summary for Policy Makers, In: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of IPCC, Cambridge University Press, UK.
- Petrinex, Canada Petroleum Information Networks, 2022. Available at <https://www.petrinex.ca/>.
- Petroleum Services Association of Canada (PSAC), 2022, PSAC Resources. Retrieved April 25, 2022, from <https://www.pzac.ca/resources/>
- Alberta Energy Regulator (AER), 2020, Atlas of the Western Canada Sedimentary Basin, Chapter 31: Petroleum Generation and Migration in the Western Canada Sedimentary Basin, Edmonton, AB.
- Stack Technologies Ltd., PetroNinja, 2022, Retrieved May 20, 2022, from <https://petroninja.com/>

- Skrodzki, M., 2019, The K-D Tree Data Structure and a Proof for Neighborhood Computation in Expected Logarithmic Time, Cornell University, <https://doi.org/10.48550/arXiv.1903.04936>
- Clearstone Engineering, 2018. Update of Equipment, Component, and Fugitive Emission Factors for Alberta Upstream Oil and Gas, Calgary, AB.
- Process Ecology Inc., 2020, Techno-Economic Assessment Model: Product Documentation, Calgary, AB.
- Delphi Program, 2017, Methane Abatement Costs: Alberta. Report for the Alberta Energy Regulator, Ottawa, ON.



# Thermodynamic description of the CO<sub>2</sub>-AMP-H<sub>2</sub>O system by ENRTL-RK model

Stefania Moioli<sup>a,\*</sup>, Laura A. Pellegrini<sup>a</sup>, Ricardo R. Wanderley<sup>b</sup>, Hanna K. Knuutila<sup>b</sup>

<sup>a</sup>*GASP, Group on Advanced Separation Processes & GAS Processing, Dipartimento di Chimica, Materiali e Ingegneria Chimica "Giulio Natta", Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133 Milano, Italy*

<sup>b</sup>*Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), NO-7491 Trondheim, Norway*  
*stefania.moioli@polimi.it*

## Abstract

Aqueous MonoEthanolAmine (MEA) solution is a well-established benchmark solvent for capturing CO<sub>2</sub>, however it is characterized by some relevant drawbacks that make its substitution with other solvents an important target to be achieved for Carbon Capture, Storage and Utilization.

2-Amino-2-Methyl-1-Propanol (AMP) is a sterically hindered primary amine which is being under consideration because of some advantages as those related to its high loading capacity and high thermal stability.

This work focuses on studying of the CO<sub>2</sub> chemical absorption in an aqueous solvent containing AMP. ASPEN Plus<sup>®</sup> process simulator has been used. After determining the thermodynamic method and the reaction set to be employed, the obtained model has been used for the simulation of a pilot plant for validation and has been employed for the study of possible application of the AMP solvent to the upstream CO<sub>2</sub> removal of an Integrated Biomass Gasification Combined Cycle (IBGCC).

**Keywords:** CCUS, AMP, IBGCC, chemical absorption.

## 1. Introduction

Aqueous MonoEthanolAmine (MEA) solution is a well-established benchmark solvent for capturing CO<sub>2</sub>. However it has some drawbacks that make research focusing on alternative solvents for Carbon Capture, Utilization and Storage (CCUS) necessary.

2-Amino-2-Methyl-1-Propanol (AMP) is being considered one possible substitute to the already industrially employed MEA, despite its lower kinetics, because of some advantages such as those related to its high loading capacity and high thermal stability (Barzagli et al., 2019) and low heat of absorption of CO<sub>2</sub>. AMP is a sterically hindered amine, a primary amine with the amino group attached to a tertiary carbon atom, and is not expected to form a stable carbamate, as primary and secondary amines do. Because of the different stoichiometry of the reaction, the maximum loading of CO<sub>2</sub> in AMP solvent is higher when compared to MEA (Gabrielsen et al., 2006). It has been studied in aqueous solution (Barzagli et al., 2013; da Silva and Svendsen, 2006), in mixtures with other amines in aqueous solutions, such as AMP+PZ (Artanto et al., 2014; Nwaoha et al., 2018) and also as a component of water-lean or water-free solvents. Water-lean solvents could reduce the energy demand due to lower heat capacity, thus reducing the heat demand and potentially lowering the cost of carbon capture (Sanku and Svensson, 2019).

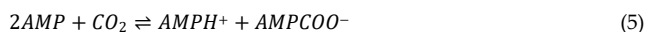
This work focuses on the thermodynamics of the system of CO<sub>2</sub>+AMP+water intending to provide a vapor-liquid equilibrium model suitable for process simulations.

The proposed model, after validation on experimental data of a pilot plant (Gabrielsen, 2007), is employed for the study of possible application of the AMP solvent to the upstream CO<sub>2</sub> removal of an Integrated Biomass Gasification Combined Cycle (IBGCC) (Carpentieri et al., 2005), with optimization of the main process characteristics.

## 2. Thermodynamic modeling with chemical reaction

A rigorous thermodynamic model must take into account the Vapor-Liquid-Equilibrium (VLE), physical solubility of gases in the solvent, and the chemical reactions occurring in the liquid phase.

In this work, the set of chemical reactions often employed to describe AMP reacting with CO<sub>2</sub> in aqueous systems (AspenTech, 2021) has been modified to take into account the reactions required to describe water-lean or water-free AMP-based solvents. In particular, in order to consider only linear independent reactions on the basis of the rank of the matrix species-elements, the considered set is:



The set includes the formation of carbamate by a reaction between AMP and CO<sub>2</sub> as considered in Sanku and Svensson (2019), not including water.

The ENRTL-RK method in ASPEN Plus<sup>®</sup> was used. This method is based on the Electrolyte-NRTL theory developed by Chen et al. (Chen et al., 1979; Chen et al., 1982; Chen and Evans, 1986; Mock et al., 1986), though with some modifications. In particular, for mixed electrolyte systems, the ENRTL-RK method uses the mixing rules only to calculate interaction parameters, instead of calculating both interaction parameters and Gibbs free energy from mixing rules. Moreover, the ENRTL-RK uses a single thermodynamics framework to calculate the activity coefficients, Gibbs free energy and enthalpy, instead of using separate models. The two methods (ENRTL-RK and Chen et al.) are identical for systems containing a single electrolyte.

Binary interaction parameters, without default values for ENRTL-RK in the software, have been obtained by using Data Regression System. Both in-house experimental data and literature data (Montalbano, 2022) of solubility of CO<sub>2</sub> in the aqueous AMP solution, and heat of absorption data (Arcis et al., 2007; Kim, 2009) were used. A non-randomness factor of 0.2 resulted to best represent the experimental values together with the determined values for  $a_{ij}$  and  $a_{ji}$ ,  $b_{ij}$  and  $b_{ji}$  determining the  $\tau$  parameter of H<sub>2</sub>O-(AMPH<sup>+</sup>, HCO<sub>3</sub><sup>-</sup>), of AMP-(AMPH<sup>+</sup>, HCO<sub>3</sub><sup>-</sup>) and of CO<sub>2</sub>-(AMPH<sup>+</sup>, HCO<sub>3</sub><sup>-</sup>). The % Absolute Average Deviations (AAD%) on the partial pressure of CO<sub>2</sub> of all the considered sources of experimental data is 13.86%, with a minimum of 1.63% for one source and a maximum of 43.97% for another source. Results are presented in Figure 1 for one set of experimental data as example, for which the AAD% resulted to be equal to 35.9% for data at 313 K,

equal to 3.79% for data at 353 K and 12.15% for data at 393 K, considering all the available data. The AAD% for 313 K resulted high because of five points not well represented at low loadings, with a % error > 30%. Without these points the resulting AAD% would be equal to 7.83%. The obtained AAD% for the heat of absorption data is 0.41% for all the considered data.

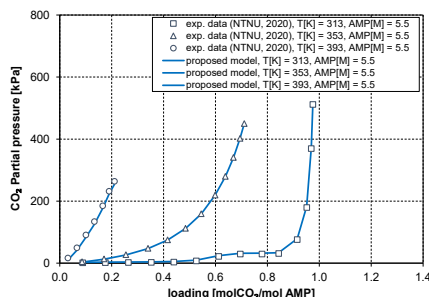


Figure 1. Comparison between the experimental points from NTNU and the calculated values of CO<sub>2</sub> partial pressure for varying loading and temperatures obtained in this work.

### 3. Simulation of the process

#### 3.1. Validation on pilot plant data

The experimental points obtained in runs of a pilot plant providing data for both the absorption section (with diameter equal to 0.15 m and packing height of 4.36 m) and the regeneration section (with diameter equal to 0.10 m and packing height of 3.89 m) by Gabrielsen (2007) have been used for model validation. Rate-based simulations have been performed, using reaction kinetics as reported in ASPEN Plus® (AspenTech, 2021). On the basis of the experimental data, the absorber and the regeneration column have been simulated separately. The input streams for absorber simulations were the lean solvent flow rate, temperature and loading, column specifications (packing type, height and diameter) and flowrate, temperature and pressure of the inlet gas stream. In the stripper simulation, the rich solvent composition, temperature and pressure, column specifications, condenser temperature and reboiler duty were used as input.

The temperature profile of the absorber for one run (RUN4) is reported in Figure 2a) and the obtained lean loadings for a fixed duty required at the reboiler are shown in Figure 2b), showing a good representation of the main characteristics of the process, including the position and the extension of the temperature bulge in the absorber, due to the liquid to gas ratio and to the exothermicity of the reactions (Kohl and Nielsen, 1997).

#### 3.2. Application to a IBGCC plant

The developed model has been used to study the application of the AMP aqueous solvent for removing CO<sub>2</sub> from a gaseous stream at atmospheric pressure in an IBGCC plant previously reported in the literature (Carpentieri et al., 2005). In the plant biomass is used to reduce greenhouse gas emissions and natural reserve depletion. CO<sub>2</sub> removal is applied to obtain a gas with high hydrogen and low carbon dioxide contents to have a fuel with high lower heating value suitable for utilization in a gas turbine for power production. The considered scheme is the base one, composed of one absorber and one regeneration column, with a lean-rich heat exchanger for heat recovery between the hot lean solvent and the cold rich solvent entering the regeneration section. The internals for both columns is Mellapak 250Y, which has an increased contact area and higher mass transfer than Mellapak 250X (Wang et al., 2013).



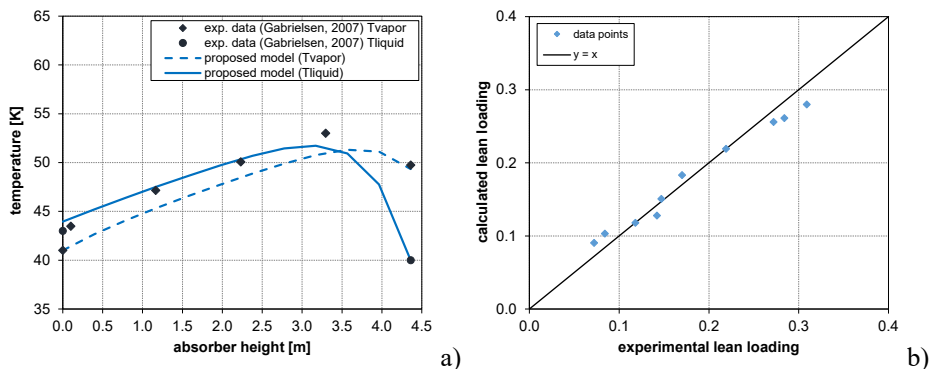


Figure 2. a) Profiles of temperature for the vapor and the liquid phase in the absorption column for RUN 4 and b) parity plot of CO<sub>2</sub> lean loading in the lean solvent exiting the regeneration section for the different runs of the pilot plant obtained with the model proposed in this work.

A parametric study aimed at optimizing the capture process for lowest thermal energy requirement, by analyzing the influence of lean CO<sub>2</sub> loading, absorber height and AMP concentration in the lean solvent has been carried out.

The obtained results are reported in Figure 3 and Figure 4. The lean solvent flow rate increases as the lean loading increases, and the energy requirement is characterized by a minimum (detailed in Figure 3b)), that decreases for higher heights of the absorption column. An absorber height of 40 m has been selected in order to reduce the energy consumptions and the operating costs in line with the recent trends in process design and, on the basis of the obtained results at different AMP concentrations reported in Table 1, the 46.3% wt. AMP solution gives the lowest energy requirement. However, the impact of the concentration on the energy consumption is minor. The reboiler duty is quite high and similar to that of 30 wt% MEA. This is due to low kinetics (the  $k$  value for the direct reaction of AMP with CO<sub>2</sub> is more than one order of magnitude lower than the reaction of MEA), reducing the rich loading that can be reached and also increasing the required liquid flow rate. The results here illustrate why AMP should always be combined with an amine, like Piperazine in CESAR1 blend, able to promote the absorption kinetics.

Table 1. Optimal lean loading and minimum energy requirement for different AMP concentrations in the lean solvent with an absorber height of 40 m.

AMP mass fraction	Optimal lean loading [mol CO <sub>2</sub> /mol AMP]	Minimum energy requirement [MJ/kgCO <sub>2</sub> ]
26.3 wt%	0.125	4.274
30.0 wt%	0.136	4.269
36.7 wt%	0.138	4.241
46.3 wt%	0.130	4.193

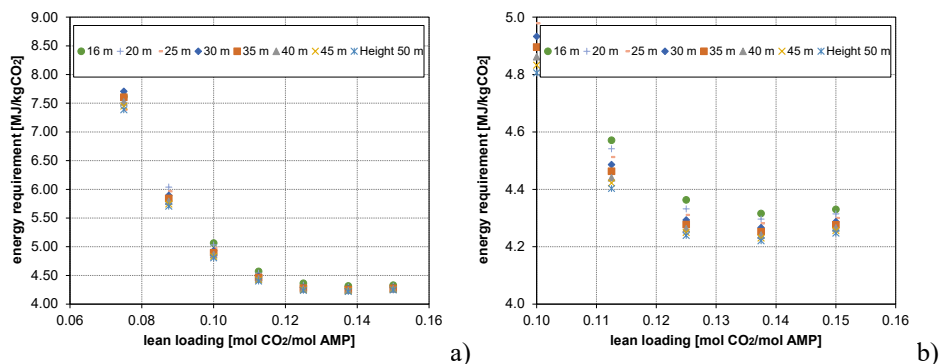


Figure 3. a) Full profile and b) detail of the profiles of energy requirement for different lean loadings and absorber heights for 90% CO<sub>2</sub> removal with an aqueous solution of AMP 36.3% wt.

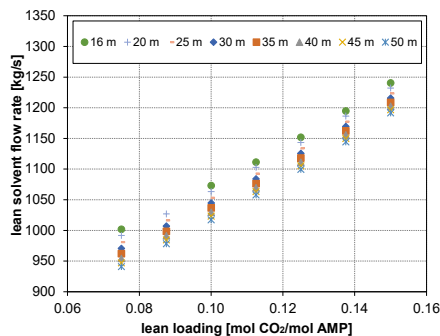


Figure 4. Profile of the needed solvent flowrate of an aqueous solution of AMP 36.3% wt. for different lean loadings and absorber heights to remove 90% of CO<sub>2</sub> present in the raw gas to be treated.

In this work, the  $\Delta T$  of approach in the lean-rich heat exchanger has been fixed to be 10 K and the regeneration pressure has been fixed to 2 bar. An analysis of the influence of the regeneration pressure is of interest in future development, considering that, as reported by Osagie (Osagie et al., 2018), the AMP solvent can be regenerated at higher temperatures (140°C) and higher pressures (3.5 bar) than MEA, potentially reducing the reboiler duty.

#### 4. Conclusions

This work has focused on studying the CO<sub>2</sub> chemical absorption with aqueous AMP, by employing ASPEN Plus<sup>®</sup> process simulator. The binary interaction parameters for the description of the thermodynamics of the system AMP-H<sub>2</sub>O-CO<sub>2</sub> with the ENRTL-RK method have been obtained by regression using experimental data for the solubility of CO<sub>2</sub> in the solvent at different conditions and using data for the heat of absorption of CO<sub>2</sub>. In addition, a new set of chemical reactions has been proposed to consider the contribution of a chemical reaction between AMP and CO<sub>2</sub> without water with the aim of also describing water-lean systems in the future.

The obtained model has been used for the simulation of a pilot plant for validation and has been employed for the study of possible application of the AMP solvent to the upstream CO<sub>2</sub> removal of an Integrated Biomass Gasification Combined Cycle to evaluate its substitution to the already considered MEA.

## References

- H. Arcis, L. Rodier, J.-Y. Coxam, 2007, Enthalpy of solution of CO<sub>2</sub> in aqueous solutions of 2-amino-2-methyl-1-propanol, *J. Chem. Thermodyn.*, 39, 878-887.
- Y. Artanto, J. Jansen, P. Pearson, G. Puxty, A. Cottrell, E. Meuleman, P. Feron, 2014, Pilot-scale evaluation of AMP/PZ to capture CO<sub>2</sub> from flue gas of an Australian brown coal-fired power station, *Int. J. Greenh. Gas Control*, 20, 189-195.
- AspenTech, 2021, Rate-Based Model of the CO<sub>2</sub> Capture Process by AMP using Aspen Plus. AspenTech, Burlington, MA.
- F. Barzagli, C. Giorgi, F. Mani, M. Peruzzini, 2019, Comparative Study of CO<sub>2</sub> Capture by Aqueous and Nonaqueous 2-Amino-2-methyl-1-propanol Based Absorbents Carried Out by <sup>13</sup>C NMR and Enthalpy Analysis, *Ind. Eng. Chem. Res.*, 58, 4364-4373.
- F. Barzagli, F. Mani, M. Peruzzini, 2013, Efficient CO<sub>2</sub> absorption and low temperature desorption with non-aqueous solvents based on 2-amino-2-methyl-1-propanol (AMP), *Int. J. Greenh. Gas Control*, 16, 217-223.
- M. Carpentieri, A. Corti, L. Lombardi, 2005, Life cycle assessment (LCA) of an integrated biomass gasification combined cycle (IBGCC) with CO<sub>2</sub> removal, *Energy Conversion and Management*, 46, 1790-1808.
- C.C. Chen, H.I. Britt, J.F. Boston, L.B. Evans, 1979, Extension and application of the Pitzer equation for vapor-liquid equilibrium of aqueous electrolyte systems with molecular solutes, *AIChE J.*, 25, 820-831.
- C.C. Chen, H.I. Britt, J.F. Boston, L.B. Evans, 1982, Local composition model for excess Gibbs energy of electrolyte systems. Part I: single solvent, single completely dissociated electrolyte systems, *AIChE J.*, 28, 588-596.
- C.C. Chen, L.B. Evans, 1986, A local composition model for the excess Gibbs energy of aqueous electrolyte systems, *AIChE J.*, 32, 444-454.
- E.F. da Silva, H.F. Svendsen, 2006, Study of the Carbamate Stability of Amines Using ab Initio Methods and Free-Energy Perturbations, *Ind. Eng. Chem. Res.*, 45, 2497-2504.
- J. Gabrielsen, 2007, CO<sub>2</sub> Capture from Coal Fired Power Plants. Technical University of Denmark.
- J. Gabrielsen, M.L. Michelsen, E.H. Stenby, G.M. Kontogeorgis, 2006, Modeling of CO<sub>2</sub> absorber using an AMP solution, *AIChE J.*, 52, 3443-3451.
- I. Kim, 2009, Heat of reaction and VLE of post combustion CO<sub>2</sub> absorbents. Norwegian University of Science and Technology (NTNU).
- A.L. Kohl, R. Nielsen, 1997, *Gas Purification*, 5th ed. Gulf Publishing Company, Book Division, Houston, Texas, USA.
- B. Mock, L.B. Evans, C.C. Chen, 1986, Thermodynamic Representation of Phase Equilibria of Mixed-Solvent Electrolyte Systems, *AIChE J.*, 32, 1655-1664.
- D. Montalbano, 2022, Improved Thermodynamic Model and Simulation of the Absorption Process by Aqueous AMP Solvent for CO<sub>2</sub> Removal. Politecnico di Milano, Milano, Italy.
- C. Nwaoha, M. Beaulieu, P. Tontiwachwuthikul, M.D. Gibson, 2018, Techno-economic analysis of CO<sub>2</sub> capture from a 1.2 million MTPA cement plant using AMP-PZ-MEA blend, *Int. J. Greenh. Gas Control*, 78, 400-412.
- E. Osagic, C. Biliyok, G. Di Lorenzo, D.P. Hanak, V. Manovic, 2018, Techno-economic evaluation of the 2-amino-2-methyl-1-propanol (AMP) process for CO<sub>2</sub> capture from natural gas combined cycle power plant, *Int. J. Greenh. Gas Control*, 70, 45-56.
- M.G. Sanku, H. Svensson, 2019, Modelling the precipitating non-aqueous CO<sub>2</sub> capture system AMP-NMP, using the unsymmetric electrolyte NRTL, *Int. J. Greenh. Gas Control*, 89, 20-32.
- C. Wang, M. Perry, F. Seibert, G.T. Rochelle, 2013, Characterization of Novel Structured Packings for CO<sub>2</sub> Capture, *Energy Procedia*, 37, 2145-2153.

# An Automated Approach for Emission Reduction Cost Calculation

Maria Victoria Migo-Sumagang,<sup>a,b</sup> Raymond R. Tan,<sup>a</sup> Kathleen B. Aviso,<sup>b</sup>  
Dominic C. Y. Foo,<sup>c\*</sup>

<sup>a</sup> *Department of Chemical Engineering, De La Salle University, 2401 Taft Avenue, 0922 Manila, Philippines*

<sup>b</sup> *Department of Chemical Engineering, College of Engineering and Agro-Industrial Technology, University of the Philippines Los Baños, College, Los Baños, Laguna, 4031, Philippines*

<sup>c</sup> *Department Chemical and Environmental Engineering/Centre of Excellence for Green Technologies, University of Nottingham Malaysia, Broga Road, 43500 Semenyih, Selangor, Malaysia*

## Abstract

The marginal abatement cost (MAC) curve is a popular visual tool in emissions reduction planning. The curve is generated by plotting the emission reduction measures by increasing MAC or specific cost on the vertical axis vs. their cumulative emissions reduction potential on the horizontal axis. The resulting graph shows the different measures arranged by cost-effectiveness, which can be easily communicated to stakeholders. The practicality of the tool makes it useful in evaluating the tradeoffs between emissions reduction and cost. However, the main limitation of the tool is the indirect determination of the total cost, which requires tedious steps in calculating the areas under the curve. This work addresses this limitation by an automated-MAC (AMAC) approach. The latter performs the calculations on the total emission reduction cost. A case study on emissions reduction planning with negative emission technologies illustrates the proposed method. This new approach removes the tediousness of calculating the total cost for emission reduction.

**Keywords:** negative emission technologies, process integration, targeting, net-zero emissions planning

## 1. Introduction

The latest IPCC reports show that achieving net-zero emissions by 2050 is critical to limiting global warming to 2 °C and even 1.5 °C (IPCC, 2022). Net-zero emissions are being targeted by firms (Chrobak, 2021) and countries worldwide (Iyer et al., 2021). Aside from the conventional approaches like turning to renewable energy and increasing energy efficiency, negative emission technologies (NETs) are now part of the solution portfolio to address the residual emissions from the hard-to-abate sectors (IPCC, 2022). These technologies remove carbon dioxide from the atmosphere resulting in a net decrease in its atmospheric concentration (The Royal Society, 2018). Examples of NETs include the ones utilizing the biological pathway: afforestation/reforestation, bioenergy with carbon capture and storage (BECCS), biochar (BC), and soil carbon sequestration (The Royal Society, 2018). The advantages of BECCS and BC over the other NETs are

the generation of energy on top of their carbon sequestration function and the increased permanence of their carbon storage (The Royal Society, 2018).

Several studies have investigated the integration of NETs into energy systems to achieve net-zero emissions, particularly energy systems models involving BECCS (Köberle, 2019). A linear program was used to transition the global energy mix to net-zero emissions using BECCS (Selosse and Ricci, 2014). The same approach was done in a country-scale energy mix (Kato and Kurosawa, 2019). BECCS supply chain optimization using process systems engineering has also been investigated (Fajardy and Mac Dowell, 2017). Optimization techniques using process integration also play a part in implementing cost-effective decarbonization strategies while simultaneously considering energy and resource use efficiencies (Klemeš, 2022). For instance, a NET portfolio was optimized under multi-footprint resource constraints (Migo-Sumagang et al., 2021).

The mentioned studies utilize mathematical programming, which can handle complex problems in planning/integrating NETs. However, mathematical programming has less interactivity compared to graphical techniques. On the other hand, graphical approaches provide immediate visualization of the solutions, which can support the first steps of decision-making (Klemeš et al., 2013). Graphical approaches in carbon footprint planning have also been developed to include algebraic targeting techniques such as cascade analysis to improve their targeting features (Foo et al., 2008). Graphical approach studies that integrate NETs in energy mixes and decarbonization strategies include pinch analysis (Nair et al., 2020) and marginal abatement cost (MAC) curves (Lameh et al., 2022).

MAC curves were derived from the energy conservation supply curves in the 1980s (Meier et al., 1982). They involve plotting the specific cost (or MAC) vs. the cumulative abatement potential of mitigation options in rectangular coordinates. The mitigation options are arranged in increasing specific costs to enable the easy determination of the most cost-effective options visually. MAC curves gained attention when they were used to rank global decarbonization options (Enkvist et al., 2007). The main limitation of the tool is the indirect determination of the total cost, which requires tedious steps in calculating the areas under the curve. So far, only two studies have been found utilizing MAC curves for decarbonization planning with NETs (Lameh et al., 2022) and net-zero planning (Migo-sumagang et al., 2022).

There is an overall lack of studies in graphical approaches to net-zero planning with NETs. Although MAC curves are a practical tool in decarbonization planning, the calculation of the total cost using this method is tedious. To address the research gap and to remove the tediousness in the cost calculation of MAC curves, this work develops an automated-MAC (AMAC) approach to calculate the total cost in MAC curves. Additionally, the developed technique enables the targeting of net-zero emission. The rest of the paper is organized as follows. The problem statement is given in the next section, followed by the proposed methodology. Next, an illustrative case study on net-zero planning with NETs is presented. Lastly, the conclusions and future work are given.

## 2. Problem Statement

The net-zero planning problem is stated as follows. Given a set of NETs ( $N_1, N_2, \dots, N_i$ ), with their specific costs and their negative emission capacities ( $A_i$ ). Given a set of CO<sub>2</sub>-emitting sources ( $P_1, P_2, \dots, P_j$ ), with their specific costs and emissions ( $G_j$ ). The objective of the current work is to determine the total cost in the MAC curve using an automated approach. The work also enables the determination of the negative emissions deficit to achieve net-zero emission, as well as the total cost to reach such target.

### 3. Proposed methodology

This section describes the newly proposed automated-MAC (AMAC) technique, with its basic structure shown in Table 1. Procedure for the AMAC is given as follows.

- i. Arrange the specific costs ( $C_k$ ) of the NETs and the CO<sub>2</sub>-emitting sources in descending order in column 1. The last value in column 1 should be zero.
- ii. Get the difference between the adjacent rows of the specific cost ( $\Delta C_k$ ) using Eq. (1), placing the results in the intervals as shown in column 2.

$$\Delta C_k = C_{k+1} - C_k \tag{1}$$

- iii. Place the total negative emission capacity of each NET ( $\sum_i A_i$ ) in column 3 and the total emission of each CO<sub>2</sub>-emitting source ( $\sum_j G_j$ ) in column 4 at their respective specific cost level  $k$ .

Table 1 Basic structure of the AMAC technique.

$C_k$	$\Delta C_k$	$\sum_i A_i$	$\sum_j G_j$	$E_k$	Cum. $E_k$	Cum. $\sum_i A_i$	Cum. $\sum_j G_j$	$TC_k$	Cum. $TC_k$
$C_1$		$\sum_i A_i$	$\sum_j G_j$	$E_1$					
	$\Delta C_1$				Cum. $E_1$	Cum. $\sum_i A_i$	Cum. $\sum_j G_j$	$TC_1$	
$C_2$				$E_2$					Cum. $TC_1$
	$\Delta C_2$				Cum. $E_2$			$TC_2$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$					Cum. $TC_2$
	$\vdots$				$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$C_{n-1}$				$E_{n-1}$					$\vdots$
	$\Delta C_{n-1}$				Cum. $E_{n-1}$	$\vdots$	$\vdots$	$TC_{n-1}$	
$C_n$		$\sum_i A_i$	$\sum_j G_j$	$E_n$					Cum. $TC_{n-1}$
	$\Delta C_n$				Cum. $E_n$	Cum. $\sum_i A_i$	Cum. $\sum_j G_j$	$TC_n$	
0									Cum. $TC_n$

- iv. Subtract the negative emission capacity ( $\sum_i A_i$ ) from the total emission ( $\sum_j G_j$ ) at each specific cost level  $k$  and assign it as  $E_k$  in column 5 and is calculated using Eq. (2). A positive value indicates a positive emission or CO<sub>2</sub> released while a negative value indicates a negative emission or CO<sub>2</sub> removed from the atmosphere. The last value in column 6 (Cum.  $E_n$ ) represents the negative emissions deficit; the latter refers to the amount of negative emissions required to reach net zero.

$$E_k = \sum_j G_j - \sum_i A_i \tag{2}$$

- v. Get the cumulative values of  $E_k$  (column 6),  $\sum_i A_i$  (column 7), and  $\sum_j G_j$  (column 8) using Eqs. (3) to (5).

$$\text{Cum. } E_k = \text{Cum. } E_{k-1} + E_k \tag{3}$$

$$\text{Cum. } \sum_i A_i = \text{Cum. } \sum_{i-1} A_{i-1} + \sum_i A_i \tag{4}$$

$$\text{Cum. } \sum_j G_j = \text{Cum. } \sum_{j-1} G_{j-1} + \sum_j G_j \tag{5}$$

- vi. Calculate the total cost ( $TC_k$ ) in column 9 using Eq. (6). This value represents the area of a rectangle formed by the cumulative specific cost and emissions in the MAC curve.

$$TC_k = (\text{Cum. } \sum_i A_i + \text{Cum. } \sum_j G_j) \Delta C_k \tag{6}$$

- vii. Get the cumulative total cost (Cum.  $TC_k$ ) using Eq. (7) in the last column. The last value in column 10 (Cum.  $TC_n$ ) represents the total cost of the system.

$$\text{Cum. } TC_k = \text{Cum. } TC_{k-1} + TC_k \tag{7}$$

- viii. Formulate the objective function and constraints. The former is set to minimize the total cost of the system as found in Eq. (8), subject to all constraints in Eqs. (1) – (7), as well as a target net-zero emissions in Eq. (9). The optimization model is a linear programming (LP), which can be solved using any LP optimization tool such as Matlab, Lingo, MS Excel, etc.

$$\min \text{Cum. } TC_n \quad (8)$$

$$\text{Cum. } E_n \leq 0 \quad (9)$$

#### 4. Illustrative Case Study

The following case study is adapted from the work of Migo-sumagang et al. (2022). The net-zero emissions targeting of a power industry using coal and natural gas as energy sources and BC and BECCS as NETs is investigated. Table 2 shows the specific costs, emissions, and NET capacities of the said system.

Table 2 Data of the NETs and CO<sub>2</sub>-emitting sources

Entity	Power Rating (MW)	Specific Cost (\$/t CO <sub>2</sub> )	Emissions (Mt CO <sub>2</sub> /y)	Reference
Coal power plant	500	9.91	4.14	(Lameh et al., 2022)
Natural Gas power plant	500	12.21	1.69	(Lameh et al., 2022)
BC	0	30	To be determined	(Fuss et al., 2018)
BECCS	500	100*		(Donnison et al., 2020)

\*(Fuss et al., 2018)

Implementing the AMAC procedure in MS Excel by solving Eq. (8), subject to the constraint in Eqs. (1) – (7) and (9), with results shown in Table 3 (scenario 1). Based on the results, the system now achieves net-zero emissions (see column 6, last row) with a total cost of 236.56 M \$/y (see column 10, last row).

Table 3 Results of Scenario 1.

$C_k$	$\Delta C_k$	$\sum A_i$	$\sum G_j$	$E_k$	Cum. $E_k$	Cum. $\sum A$	Cum. $\sum G$	$TC_k$	Cum. $TC_k$
\$/t CO <sub>2</sub>	\$/t CO <sub>2</sub>	Mt CO <sub>2</sub> /y	Mt CO <sub>2</sub> /y	Mt CO <sub>2</sub> /y	Mt CO <sub>2</sub> /y	Mt CO <sub>2</sub> /y	Mt CO <sub>2</sub> /y	M \$/y	M \$/y
100		0		0					
	70				0	0	0	0	
30		5.83		-5.83					0
	17.79				-5.83	5.83	0	103.72	
12.21			1.69	1.69					103.72
	2.3				-4.14	5.83	1.69	17.30	
9.91			4.14	4.14					121.01
	9.91				0	5.83	5.83	115.55	
0									236.56

In scenario 2, the model is set to minimize cost, subject to the maximum BC capacity of 1.25 Mt CO<sub>2</sub>/y (Migo-sumagang et al. (2022), with a new constraint in Eq (10).

$$\sum A_{BC} \leq 1.25 \quad (10)$$

Solving the AMAC with objective in Eq. (8), subject to the constraints in Eqs. (1) – (7) and (9) – (10), with results given in Table 4. As shown, the system has an increased BECCS capacity of 4.58 Mt CO<sub>2</sub>/y, while BC is kept to the maximum of 1.25 Mt CO<sub>2</sub>/y. The new system has a higher total cost of 557.16 M \$/y (column 10, last row). Note that these results were identical to those reported by Migo-sumagang et al. (2022) using graphical method.

Comparison of the two scenarios is shown in the MAC curves in Figure 1. Note that in the original graphical procedure, the total cost has to be calculated by manually summing the areas below the MAC curves (see details in Figure 1; Migo-sumagang et al. (2022)). With the AMAC, the total cost values can be easily determined. Besides, changes due to process system improvement can also be updated easily with the AMAC model.

Table 4 Results of Scenario 2.

$C_k$ \$/t CO <sub>2</sub>	$\Delta C_k$ \$/t CO <sub>2</sub>	$\sum A_i$ Mt CO <sub>2</sub> /y	$\sum G_j$ Mt CO <sub>2</sub> /y	$E_k$ Mt CO <sub>2</sub> /y	Cum. $E_k$ Mt CO <sub>2</sub> /y	Cum. $\sum A$ Mt CO <sub>2</sub> /y	Cum. $\sum G$ Mt CO <sub>2</sub> /y	$TC_k$ M \$/y	Cum. $TC_k$ M \$/y
100		4.58		-4.58					
	70				-4.58	4.58	0	320.6	
30		1.25		-1.25					320.6
	17.79				-5.83	5.83	0	103.72	
12.21			1.69	1.69					424.32
	2.3				-4.14	5.83	1.69	17.30	
9.91			4.14	4.14					441.61
	9.91				0	5.83	5.83	115.55	
0									557.16

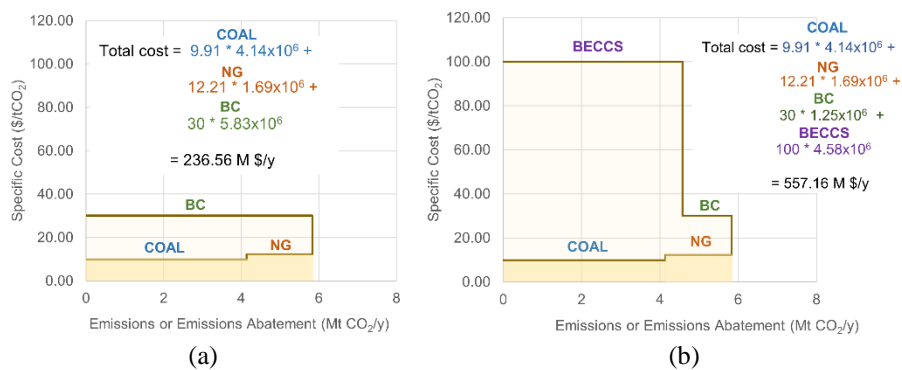


Figure 1 MAC curves for (a) scenario 1 and (b) scenario 2 (with a capacity constraint on BC)

### 5. Conclusions

An automated MAC (AMAC) technique for targeting net-zero emissions was developed. The technique enables the automated calculation of the total cost for net-zero scenarios. A case study on emissions reduction planning with NETs was used to illustrate the method. The results of the automated method match those from graphical approach, and removes the tediousness of calculating the total cost in emission reduction planning. Future work can include the calculation of the cost of the alternative energy, as well as the consideration of other environmental footprints of NETs aside from cost.



## References

- Chrobak, U., 2021. Corporate Climate Pledges Pile Up—Will It Matter? *Engineering* 7, 1044–1046.
- Donnison, C., Holland, R.A., Hastings, A., Armstrong, L.M., Eigenbrod, F., Taylor, G., 2020. Bioenergy with Carbon Capture and Storage (BECCS): Finding the win–wins for energy, negative emissions and ecosystem services—size matters. *GCB Bioenergy* 12, 586–604.
- Enkvist, P.-A., Nauclér, T., Rosander, J., 2007. A cost curve for greenhouse gas reduction [WWW Document]. McKinsey. URL <https://www.mckinsey.com/business-functions/sustainability/our-insights/a-cost-curve-for-greenhouse-gas-reduction> (accessed 3.21.22).
- Fajardy, M., Mac Dowell, N., 2017. Can BECCS deliver sustainable and resource efficient negative emissions? *Energy and Environmental Science* 10, 1389–1426.
- Foo, D.C.Y., Tan, R.R., Ng, D.K.S., 2008. Carbon and footprint-constrained energy planning using cascade analysis technique. *Energy* 33, 1480–1488.
- Fuss, S., Lamb, W.F., Callaghan, M.W., Hilaire, J., Creutzig, F., Amann, T., Beringer, T., De Oliveira Garcia, W., Hartmann, J., Khanna, T., Luderer, G., Nemet, G.F., Rogelj, J., Smith, P., Vicente, J.V., Wilcox, J., Del Mar Zamora Dominguez, M., Minx, J.C., 2018. Negative emissions - Part 2: Costs, potentials and side effects. *Environmental Research Letters* 13, 063002.
- IPCC, 2022. Summary for Policymakers. *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK and New York, NY, USA.
- Iyer, G., Clarke, L., Edmonds, J., Fawcett, A., Fuhrman, J., McJeon, H., Waldhoff, S., 2021. The role of carbon dioxide removal in net-zero emissions pledges. *Energy and Climate Change* 2, 100043.
- Kato, E., Kurosawa, A., 2019. Evaluation of Japanese energy system toward 2050 with TIMES-Japan – deep decarbonization pathways. *Energy Procedia* 158, 4141–4146.
- Klemeš, J.J., 2022. Process Integration (PI): An Introduction. In: *Handbook of Process Integration (PI): Minimisation of Energy and Water Use, Waste and Emissions*. Woodhead Publishing, Cambridge, UK, pp. 3–27.
- Klemeš, J.J., Varbanov, P.S., Kravanja, Z., 2013. Recent developments in Process Integration. *Chemical Engineering Research and Design* 91, 2037–2053.
- Köberle, A.C., 2019. The Value of BECCS in IAMs: a Review. *Current Sustainable/Renewable Energy Reports* 6, 107–115.
- Lameh, M., Al-Mohannadi, D.M., Linke, P., 2022. Minimum marginal abatement cost curves (Mini-MAC) for CO<sub>2</sub> emissions reduction planning. *Clean Technologies and Environmental Policy* 24, 143–159.
- Meier, A., Rosenfeld, A.H., Wright, J., 1982. Supply curves of conserved energy for California's residential sector. *Energy* 7, 347–358.
- Migo-Sumagang, M.V., Aviso, K., Tapia, J.F., Tan, R.R., 2021. A Superstructure Model for Integrated Deployment of Negative Emissions Technologies under Resource Constraints. *Chemical Engineering Transactions* 88, 31–36.
- Migo-sumagang, M.V., Aviso, K.B., Bhasker, P.N.S., Short, M., Tan, R.R., Foo, D.C.Y., 2022. A Graphical Technique for Net-Zero Emissions Planning Based on Marginal Abatement Cost (MAC) Curves. *Chemical Engineering Transactions* 96.
- Nair, P.N.S.B., Tan, R.R., Foo, D.C.Y., 2020. Extended Graphical Approach for the Deployment of Negative Emission Technologies. *Industrial and Engineering Chemistry Research* 59, 18977–18990.
- Selosse, S., Ricci, O., 2014. Achieving negative emissions with BECCS (bioenergy with carbon capture and storage) in the power sector: New insights from the TIAM-FR (TIMES Integrated Assessment Model France) model. *Energy* 76, 967–975.
- The Royal Society, 2018. *Greenhouse Gas Removal*, The Royal Society. London, UK.

# Lumped-Parameter Heat Exchanger Models for the Robust Dynamic Modelling of Power Generation Cycles

Oliver Ward<sup>a</sup>, Federico Galvanin<sup>a</sup>, Nelia Jurado<sup>b</sup>, Chris Clements<sup>c</sup>, Mohamad Abdallah<sup>d</sup>, Daniel Blackburn<sup>c</sup>, Eric Fraga<sup>a</sup>

<sup>a</sup>*Department of Chemical Engineering, University College London, Torrington Place, London, WC1E 7JE, UK*

<sup>b</sup>*Department of Mechanical Engineering, University College London, Torrington Place, London, WC1E 7JE, UK*

<sup>c</sup>*United Kingdom Atomic Energy Authority, Culham Science Centre, Abingdon, OX14 3DB, UK*

<sup>d</sup>*Formerly at United Kingdom Atomic Energy Authority, Culham Science Centre, Abingdon, OX14 3DB, UK*

## Abstract

In this work, novel lumped-parameter dynamic models are presented for one-phase and two-phase heat exchangers, based on the logarithmic mean temperature difference. These models allow for the more robust and computationally efficient simulation of thermal power plants than traditional discretized finite volume models. Comparison of the one-phase lumped model and the finite volume model shows good agreement in the computed outlet temperatures under transient inlet conditions, with steady state differences of <0.15°C. The dynamic responses of outlet temperatures from the models show good agreement, besides a response delay that the lumped model cannot replicate. The lumped models are then implemented in a model of a steam Rankine cycle driven by a source of hot molten salt. The behavior of the system is investigated under a 10% drop in the molten salt mass flow rate, with comparable drops in the main steam pressure, mass flow rate and net power output but a negligible impact on the main steam temperature. The future use of the models is in the design and control optimization of the power conversion system for the STEP fusion power plant being designed by UKAEA.

**Keywords:** Heat exchanger modelling; Dynamic simulation; Steam Rankine cycle; Fusion power; Modelica.

## 1. Introduction

The Spherical Tokamak for Energy Production (STEP) is a project run by the United Kingdom Atomic Energy Authority (UKAEA) to demonstrate the commercial viability of fusion power generation [UKAEA (2022)]. Nuclear fusion is an attractive prospect to meet growing energy demands, using a fusion reactor as the thermal energy source for a thermal power plant. Fusion energy has the potential for an abundant fuel supply that can be extracted from water or bred within the reactor, low environmental impact (emissions and radioactive by-products) and safe operation. It also poses unique design challenges relative to current thermal power plants, such as transient reactor energy output and multiple heat sources. These need to be accounted for in the design of the power

conversion system, which converts the thermal energy into electrical energy. Hence, dynamic models are required for the designs to evaluate their transient performance; these evaluations will be performed within an optimization algorithm that aims to search for promising designs.

Robust and computationally inexpensive models that still capture the main system dynamics are a key requirement; effectively exploring the design space during optimization will need a sizeable number of simulations. Novel lumped-parameter dynamic models of one-phase and two-phase heat exchangers are developed to meet these requirements, with their use demonstrated in the simulation of a steam Rankine cycle.

## 2. Heat Exchanger Modeling

The modelling of heat exchangers varies in complexity depending on the specific component and the intended application of the model. For example, an evaporator model will need to be able to handle phase-change in the fluid whereas a preheater model can ignore this for normal operation. The finite volume method is a common approach [Desideri (2016)], which discretizes a one-dimensional thermofluid flow into several lumped storage volumes, each with dynamic balance equations.

Lumped-parameter models may be preferable for computer aided design and optimization. Their simplicity means that a minimal amount of geometric parameterization is required, useful at the conceptual design stage when no fixed details are known. Lumped models also improve initialization and reduce the required simulation time. Both aspects will be advantageous when optimizing the power plant design as each design will need to be simulated without user intervention to evaluate its effectiveness.

The development of distinct one-phase and two-phase heat exchanger models avoids an issue known as chattering that occurs in finite volume models. Chattering occurs when there are discontinuities in the model variables occur for the density on the saturated liquid curve. Chattering can lead to simulation failure unless numerical methods are employed to improve the robustness [Quoilin (2014)].

The models presented below have been implemented in the open-source OpenModelica [Open Source Modelica Consortium (2022)]. Modelica has been chosen as it is an acausal and object-oriented language [Modelica Association (2021)].

### 2.1. One-Phase Heat Exchangers

A counter-current one-phase heat exchanger model may be used to model preheaters, superheaters and reheaters in a thermal power plant. Each fluid side is modelled as a lumped thermal mass [Kannaiyan (2019)]. With the assertion that only one phase is present on each side of the heat exchanger, the assumption is made that density time derivatives are negligibly small in both fluids (see equation 1). This simplifies the resulting mass and energy balance equations of the fluids. The mass balance equation is static, with no mass storage of the fluids and hence equal inlet and outlet mass flow rates for both fluids. Equal inlet and outlet pressures are assumed here.

The energy balances are dynamic with one state variable, the outlet specific enthalpy, which is taken to be the specific enthalpy in the lumped fluid mass:

$$\frac{d}{dt} h_{f,out} = \frac{\dot{m}_f \cdot (h_{f,in} - h_{f,out}) + \dot{Q}_f}{\rho_{f,out} \cdot V_f} \quad (1)$$

Subscript  $f$  denotes either the hot or cold fluid.  $h_{f,in}$ ,  $h_{f,out}$  are the fluid's inlet and outlet specific enthalpy [ $\text{J} \cdot \text{kg}^{-1}$ ],  $\dot{m}_f$  is the fluid's mass flow rate [ $\text{kg} \cdot \text{s}^{-1}$ ],  $\dot{Q}_f$  is the heat flow rate into the fluid [W],  $\rho_{f,out}$  is the fluid's outlet density [ $\text{kg} \cdot \text{m}^{-3}$ ] and  $V_f$  is the volume of the fluid's heat exchanger side [ $\text{m}^3$ ].

Thermal storage in the heat exchanger walls is modelled assuming a linear temperature profile via two temperature state variables [Altés-Buch (2015)]. Both fluids exchange heat with the wall only. The heat transfer on each fluid side is modelled using a robust logarithmic mean temperature difference proposed for dynamic simulations [Altés-Buch (2015)]:

$$\frac{\Delta T_1 - \Delta T_2}{\ln(\Delta T_1) - \ln(\Delta T_2)} \quad \begin{array}{l} \text{if } \Delta T_1 \neq \Delta T_2, \\ \Delta T_1 \geq \varepsilon, \Delta T_2 \geq \varepsilon \end{array} \quad (2)$$

$$\frac{\Delta T_1 + \Delta T_2}{2} \quad \begin{array}{l} \text{if } \Delta T_1 = \Delta T_2, \\ \Delta T_1 \geq \varepsilon, \Delta T_2 \geq \varepsilon \end{array} \quad (3)$$

$$\Delta T_{RLMTD} = \frac{\Delta T_1 - \varepsilon}{\ln\left(\frac{\Delta T_1}{\varepsilon}\right) \cdot (1 - \xi \cdot (\Delta T_2 - \varepsilon))} \quad \begin{array}{l} \text{if } \Delta T_1 \geq \varepsilon, \\ \Delta T_2 < \varepsilon \end{array} \quad (4)$$

$$\frac{\Delta T_2 - \varepsilon}{\ln\left(\frac{\Delta T_2}{\varepsilon}\right) \cdot (1 - \xi \cdot (\Delta T_1 - \varepsilon))} \quad \begin{array}{l} \text{if } \Delta T_2 \geq \varepsilon, \\ \Delta T_1 < \varepsilon \end{array} \quad (5)$$

$$\frac{\varepsilon}{(1 - \xi \cdot (\Delta T_1 - \varepsilon)) \cdot (1 - \xi \cdot (\Delta T_2 - \varepsilon))} \quad \begin{array}{l} \text{if } \Delta T_1 < \varepsilon, \\ \Delta T_2 < \varepsilon \end{array} \quad (6)$$

$\Delta T_1$ ,  $\Delta T_2$  are the temperature differences at each heat exchanger end [K],  $\varepsilon$  is the threshold temperature difference for the alternative equations [K] and  $\xi$  is a parameter that controls how fast the function goes to zero [ $\text{K}^{-1}$ ]. More detailed discussion of the parameters can be found in [Altés-Buch (2015)]. The formulation avoids simulation failures by allowing for all possible temperature differences as inputs, which is useful for initialization when large transients are common. The function is also continuous to help with solver convergence. Note that equations 4-6 are non-physical and normal model operation should not be within these regions.

The lumped one-phase heat exchanger model can be compared with the finite volume model, using the upwind discretization scheme and 100 cells [Desideri (2016)], considering transient inlet conditions. Parameters are taken from [Altés-Buch (2015)] with steel walls, and are subjected to the same transient inlet conditions, with both fluids being water at 1bar. The initial inlet conditions are  $3\text{kg} \cdot \text{s}^{-1}$  at  $20^\circ\text{C}$  and  $2\text{kg} \cdot \text{s}^{-1}$  at  $80^\circ\text{C}$ . A sigmoidal ramp in the cold fluid inlet temperature from  $20^\circ\text{C}$  to  $10^\circ\text{C}$  between  $t=25\text{s}$  and  $t=50\text{s}$  is considered. The results are shown in figure 2.

Model agreement in the outlet temperatures can be seen to be within  $0.15^\circ\text{C}$  for both steady states, with the lumped model over-estimating the total heat transferred relative to

the finite volume model. The dynamic behavior of the lumped model closely follows the finite volume model, with a maximum difference of  $1.4^{\circ}\text{C}$  in the outlet temperatures. This is caused by a delayed response in the finite volume model due to the temperature front propagation. Only the spatially discretized finite volume model can capture this effect. Otherwise, the lumped model reflects the key behavior of the heat exchanger. The finite volume simulation required  $\sim 1000\times$  the time of the lumped-parameter simulation.

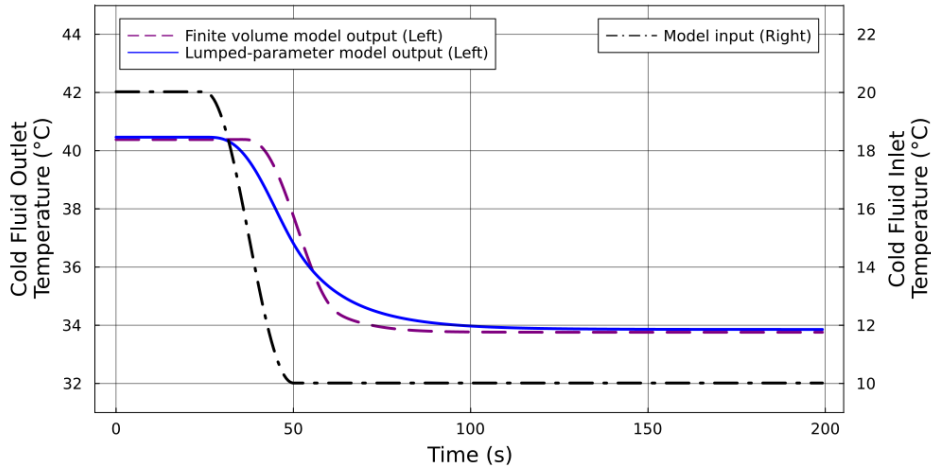


Figure 1: Comparison of results from the one-phase lumped-parameter (LP) heat exchanger model (solid blue line) with a discretized finite volume (FV) model (dashed purple line) under a  $10^{\circ}\text{C}$  decrease in the cold fluid's inlet temperature over 25s (dash-dot black line). Both models are initially at steady state. A  $\sim 10\text{s}$  delay is seen in the response of the FV model that is not seen in the LP model.

## 2.2. Two-Phase Heat Exchangers

The model introduced in the previous section is extended for two-phase heat exchangers such as evaporators and condensers. The model structure of the side undergoing phase-change, the working fluid, is modified: the working fluid is modelled as a storage volume at the saturation condition; all fluid is either saturated liquid or saturated vapor. The secondary fluid, wall and heat transfer are modelled as in the one-phase model.

The working fluid side is governed by dynamic mass and energy balance equations [Zhang (2021)]. For an evaporator, the outlet working fluid is saturated vapor, while in a condenser, it is saturated liquid. As the phases share the same physical space of volume in the heat exchanger, there is an additional constraint: the volumes occupied by each phase must add up to the total working fluid side's volume.

Both phases are in pressure equilibrium and, hence, also thermal equilibrium. The model accounts for phase-change by allowing for the phase volumes to change dynamically, dependent on the heat flow from the wall, without encountering any numerical issues from discontinuities in the density model at the saturated liquid curve.

## 3. Example: Steam Rankine Cycle Modeling

The case study is the modelling of a 45MW steam Rankine cycle, shown in figure 3. The cycle fed by a  $350\text{kg}\cdot\text{s}^{-1}$  molten salt stream at  $550^{\circ}\text{C}$ , as in a power plant with direct thermal energy storage [Zhang (2021)].

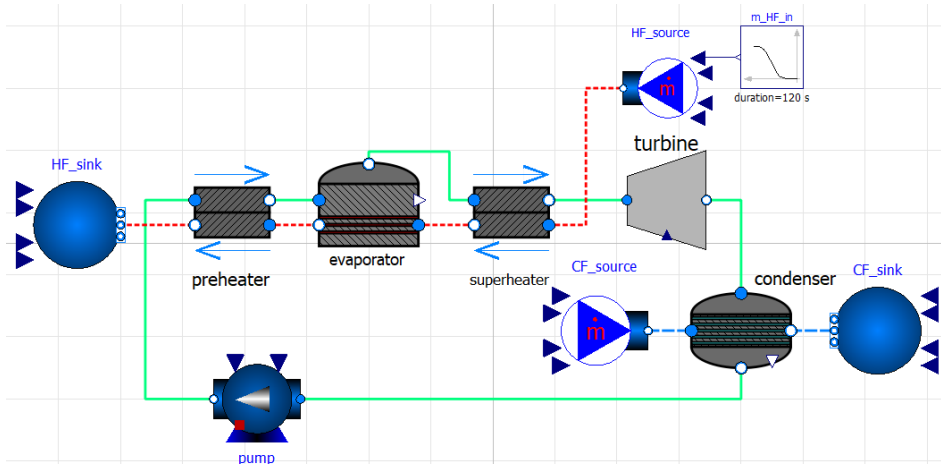


Figure 2: Diagram of the modelled steam Rankine cycle in OpenModelica. Water, the working fluid (solid green line), is heated by a source of hot molten salt at constant temperature (dotted red line) and cooled by a constant source of cooling water (dashed blue line).

The preheater and superheater are one-phase exchangers, and the evaporator and condenser are two-phase. The turbine and pump are both modelled as static components, each assuming a constant isentropic efficiency. The mass flow rate and pressure drop through the turbine is modelled using the Stodola equation [Hefni (2019)]. The pump is modelled using the Modelica Fluid library, producing an ideally constant mass flow rate [Modelica Association (2021)]. The molten salt is modelled assuming constant specific heat capacity and density, while water is modelled using the IF97 standard implemented in the Modelica Media library. Parameter values are extrapolated from [Zhang (2021)].

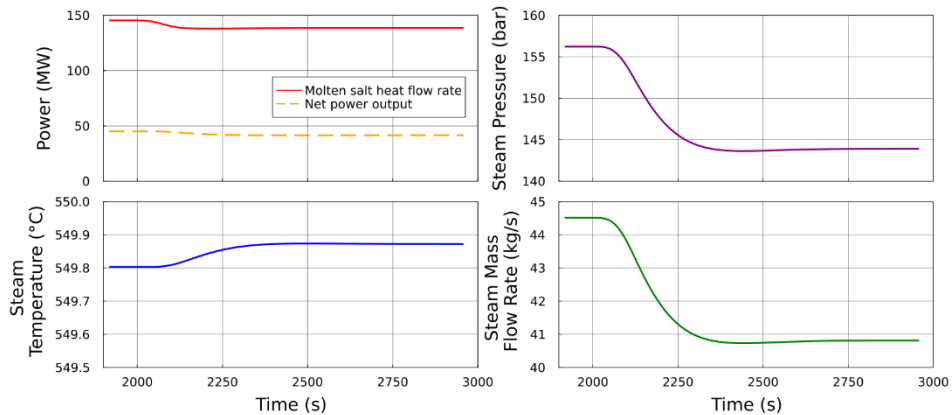


Figure 3: Dynamic response of the steam Rankine cycle to a 10% drop in the molten salt mass flow rate, between  $t=2000s$  and  $t=2120s$ , from  $350kg \cdot s^{-1}$  to  $315kg \cdot s^{-1}$ . Comparable drops in steady state values can be seen in the main steam mass flow rate, pressure, and the net power output, while the main steam temperature remains relatively unchanged. The heat flow rate delivered by the molten salt to the cycle undergoes a smaller relative drop.

The cycle is simulated undergoing a 10% sigmoidal drop over 120s in the molten salt mass flow rate, from  $350kg \cdot s^{-1}$  to  $315kg \cdot s^{-1}$ . The results are shown in figure 4. The 10% drop in molten salt mass flow rate results in comparable drops in the steady state values of the net power output, main steam (at the turbine inlet) pressure and main steam mass

flow rate, with percentage changes of 7.9%, 7.9% and 8.3% respectively. The heat flow rate from the molten salt to the cycle drops by 4.5%, which differs from the net power output drop due to a simultaneous drop in cycle efficiency. The main steam temperature increases, but to a much smaller degree than the other variables. This is because the superheater, as specified, is sufficiently sized to heat the steam to near the constant molten salt inlet temperature. The 0.1°C temperature increase of the steam temperature is due to the reduced steam mass flow rate through the superheater, reducing the heat capacity of the steam flow and allowing for a greater temperature change.

Each of the four decreasing variables shows a similar dynamic profile, with a slight overshoot of the new steady state value due to the lower inlet temperature of the molten salt to the evaporator while the superheater's steam outlet temperature is still increasing.

#### 4. Conclusion

In this work, robust lumped-parameter models were developed for one-phase and two-phase heat exchangers that make use of the logarithmic mean temperature difference. This robustness is desirable for the simulation of thermal power plants in the context of the early design and control optimization of the STEP fusion power plant being designed by UKAEA. Comparison of the one-phase heat exchanger model with a finite volume model showed good agreement in both the dynamic and steady-state outlet temperatures for changes in the inlet temperature, except for the temperature front propagation through the heat exchanger which the lumped model could not capture.

The new lumped heat exchanger models were then used in a steam Rankine cycle model, driven by a source of molten salt, as an example of their use. The system dynamics under a 10% drop in the molten salt mass flow rate showed comparable drops in the net power output, main steam pressure and main steam mass flow rate. The main steam temperature increased by only 0.1°C due to the sufficient sizing of the superheater. These results demonstrate the potential of using lumped heat exchanger dynamic models in a computer aided design environment aiming to identify not only the design of the power conversion system for STEP but also the control structure that may be necessary. The next steps include extending the Rankine cycle model with a thermal energy storage system, implementing, and comparing alternative control strategies, including energy input from the reactor, and modelling reheat and regeneration cycles.

#### References

- Q. Altés-Buch et al., 2015, *Dynamic modeling of thermal systems using a semi-empirical approach and the ThermoCycle Modelica library*.
- A. Desideri et al., 2016, *Comparison of moving boundary and finite-volume heat exchanger models in the Modelica language*, Energies, 9, 5
- B. E. Hefni and D. Bouskela, 2019, *Modeling and Simulation of Thermal Power Plants with ThermoSysPro*, Springer Cham.
- S. Kannaiyan et al., 2019, *Dynamic modeling and simulation of a hybrid solar thermal power plant*, Industrial Engineering Chemistry Research, 58, 18
- Modelica Association, 2021, *Modelica Language*.
- Open Source Modelica Consortium. 2022, *OpenModelica*.
- S. Quoilin et al., 2014, *Methods to increase the robustness of finite-volume flow models in thermodynamic systems*, Energies, 7, 3
- UKAEA, 2022, *Spherical Tokamak for Energy Production*.
- S. Zhang et al., 2021, *Dynamic simulation and performance analysis of a parabolic trough concentrated solar power plant using molten salt during the start-up process*

# Energy and economic analysis of integration of water electrolysis with an oxygen-enriched combustion power plant

Yuxing Peng,<sup>a</sup> Pei Liu,<sup>\*</sup> Zheng Li

*<sup>a</sup>State Key Lab of Power Systems, Department of Energy and Power Engineering, Tsinghua-BP Clean Energy Center, Tsinghua University, Beijing 100084, China  
[liu\\_pei@tsinghua.edu.cn](mailto:liu_pei@tsinghua.edu.cn)*

## Abstract

Hydrogen is an ideal energy carrier to promote clean utilization of fossil energy and support large-scale development of renewable energy. Green hydrogen production has no carbon emissions therefore is an ideal hydrogen source, albeit its relatively high cost. One way to reduce the cost of green hydrogen is to effectively utilize the oxygen by-product in a water electrolysis process, for instance, in a transformation of oxygen-enriched combustion coal-fired power plant. This paper discusses potential of integrating a coal-fired power plant with a water electrolysis process via oxygen production and utilization. This provides opportunities for more cost-effective utilization of oxygen produced in a water electrolysis process, and opportunities for reduction in cost of supplying oxygen to an oxygen-enriched combustion coal-fired power plant. Analysis results indicate that the oxygen supply using green hydrogen system can save 8.08 million euros compared to the oxygen supply using air separation device, which lowers the price per kilogram of green hydrogen generation by 0.07 euros, for a 2×300MW oxygen enhanced burning coal-fired power plant with 4000 hours of operation per year.

**Keywords:** green hydrogen, water electrolysis by-product oxygen, coal-fired power plants, oxygen-enriched combustion

## 1. Introduction

Hydrogen is regarded as a promising alternative energy carrier in a low/zero carbon emission energy system, especially for sectors that have difficulties with carbon emissions reduction using existing technologies, for instance, transportation and steel industry (IRENA, 2022). Water electrolysis using renewable energy, typically solar and wind, appears to be a technically viable way to produce hydrogen with low carbon emissions, namely green hydrogen (Ge, 2022). Production of green hydrogen is now significantly more expensive than that of gray hydrogen.

As a result, the most pressing issue is lowering the cost of green hydrogen generation (Wang, 2022). Utilizing the by-product oxygen from the water electrolysis process is an excellent cost-cutting method. When hydrogen is produced in a water electrolysis process, half a mole of oxygen is produced as a byproduct, and this oxygen is currently vented straight into the atmosphere. If the scale of green hydrogen production is considerably enlarged in the future, the byproduct oxygen would also be mass produced. If oxygen is



completely used, the resulting economic benefits could lower the cost of hydrogen generation.

An effective and energy-saving combustion method called oxygen-fuel combustion uses pure oxygen or oxygen-enriched air with a higher oxygen concentration than regular air as the combustion gas. Horn and Steinburg first put up the idea of oxy-fuel combustion in 1981 (Horn, 1982), and Argonne National Laboratory confirmed it later (Abraham, 1982). Oxygen-enriched combustion can effectively minimize the emissions of various pollutants after combustion, which has a significant application potential for carbon neutralization. It can also increase the efficiency of fuel consumption.

However, when the coal-fired power plant is transformed by oxygen-enriched combustion, which will significantly increase the power plant's energy consumption, the additional air separation oxygen system is required. If the oxygen-enriched combustion is changed, the entire plant's net efficiency will be decreased by at least 10% (Pikkarainen, 2014). Therefore, one significant drawback of this technology is its enormous energy usage. The air separation unit (ASU) is completely removed if the oxygen generated by electrolyzed water is used for oxy-fuel combustion in coal-fired power plants. This not only reduces the initial cost of the air separation unit but also reduces the significant energy consumption of the air separation system.

Although some coal-fired power plants will be replaced by renewable energy in the future, due to the intermittent nature of renewable energy, coal-fired power plants will still be required to perform peak shaving. As a result, it is important to investigate the oxygen-enriched combustion of coal-fired power plants for low carbon emissions. Simultaneously, an energy system comprised of renewable energy units and coal-fired power plants not only realizes renewable power absorption, but also ensures the safety and stability of power supply. As a result, using the oxygen produced by the electrolytic cell for oxygen-enriched combustion power stations is a good choice.

This paper proposed an energy system consisting of a wind turbine, an electrolyzer, and an oxy-fuel combustion coal-fired power station in order to effectively use the oxygen produced by water electrolysis from renewable energy sources. It also compared the costs of the new energy system's oxygen supply transformation to the costs of the traditional air separation oxygen supply transformation.

## 2. Methodology

### 2.1. System description

This new energy system is depicted in Figure.1. The electrolytic cell uses power from the wind farm to operate normally and generate hydrogen and oxygen. High-purity hydrogen can be used in many fields such as energy and chemical industry, while oxygen is transported to the compressor. The alkaline electrolytic cell used in the electrolytic cell is now the most developed. The compressor increases the by-product oxygen's pressure to an appropriate level before transferring it via the oxygen pipeline to the liquefaction equipment. The oxygen storage tank holds the liquid oxygen. Designers usually select the liquid form for long-term storage of huge amounts of oxygen due to the volatility of wind turbines, which may cross seasons. ASU also employs a liquid oxygen storage tank since it often has a low-capacity standby oxygen storage tank that needs to be kept for a long time. Before delivering appropriate oxygen for the power plant, the high-pressure pump and vaporizer are the necessary treatment devices. It will be present in the conventional air separation system as well, so it won't be taken into account in the following cost analysis.

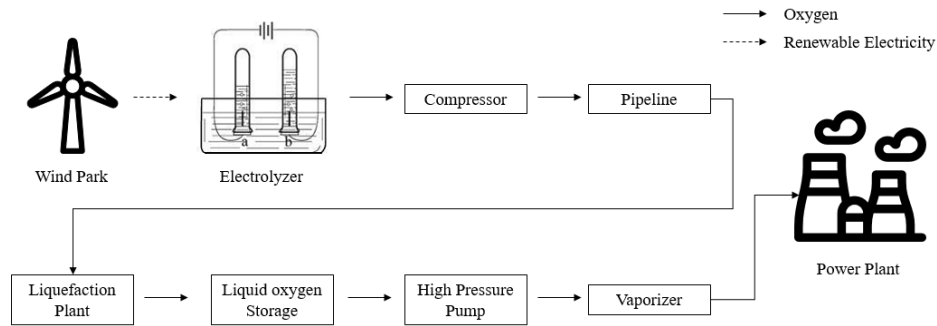


Figure 1. Composition of Energy System integrated by Electrolyzed Water and Oxygen - enriched Combustion Power Plant

Before determining the system's overall cost, make the following assumptions: assume that every equipment has a 25-year service life and that the capital investment cost depreciates linearly. Annual calculations are made for operational costs (OPEX). Some operational costs are computed as a percentage of CAPEX, while others are computed based on a specific energy use. 72 €/MWh is the projected electricity cost for operating compressors, the ASU, the liquefaction plant, etc. (Gonzá, 2021). This is the transaction price for major electricity users which is greater than 150,000 MWh in 2019, including energy taxes but excluding VAT.

2.2. Components description

The energy system requires a number of components that are important to our study. The conventional 2x300 MW unit of oxy-fuel coal-fired power plants is used. ASU is a widely used technology and the industry standard for producing huge volumes of oxygen; in this study, the ASU parameters are chosen based on the oxygen requirements of coal-fired power plants. When oxygen is transferred from an electrolytic tank to a coal-fired power plant, compressors are needed. This complies with applicable industry standards. Carbon steel pipelines can be utilized to transmit oxygen because no particular restrictions exist. There are some study models on gas pipeline costs that relate to the pipeline's diameter, length, gas pressure, and other elements. Tanks for liquid oxygen (LOX) are typically used in the ASU design. The LOX tanks of an ASU can be built to have a minimal capacity, while the LOX tanks of a green hydrogen module must be greatly expanded due to the erratic nature of renewable electricity. Assuming a particular energy consumption of 17% of the ASU, the liquefaction plant can be regarded as a component of the ASU. Table 1 displays the parameters for these components (Berenschot. 2019)(Cai, 2014).

Table 1. Components' parameters of the new energy system and ASU

Component	Unit	Parameter
Electrolyzer	Nominal power	1680MW
	Oxygen production rate	480t/h
	Yearly working hours	2000h
Power plants	Nominal power	2x300 MW
	Coal consumption per unit power	300 t/MWh

	Coal consumption per unit time	180 t/h
	Yearly working hours	4000 h
<b>ASU</b>	CAPEX	200 M€
	Oxygen feeding rate	240 t/h
	Specific energy consumption	210 kWh /t O <sub>2</sub>
	LOX capacity	1200 t
<b>Compressor</b>	CAPEX	11.6 M€
	Specific energy consumption	33 kWh /t O <sub>2</sub>
<b>Pipeline</b>	Investment cost per unit length	1.08 M€/km
	Pipeline length	5 km
<b>LOX storage tank</b>	Unit volume investment cost	1500€/m <sup>3</sup>
	Liquid oxygen density	1.413 t/m <sup>3</sup>
	LOX capacity	144000 t
<b>Liquefaction plant</b>	CAPEX	100 M€
	Specific energy consumption	40.8 kWh /t O <sub>2</sub>

### 3. Results and discussion

#### 3.1. Cost comparison

The cost comparison results are shown in Figure 2: the yearly total savings is 8.08 million euros. The main reason of this saving is the complete removal of the ASU, which saves 18.59 million euros in OPEX and 8.00 million euros in CAPEX.

In the new energy system, since the wind park is intermittent, the gaseous oxygen from the electrolyzer needs to be stored in the oxygen tank. If oxygen is overproduced, gaseous oxygen needs to be liquefied. The liquefaction process increases OPEX by about 3.92 million euros and increases CAPEX by about 4.00 million euros. The capacity of oxygen storage tank is obviously increased, the CAPEX of oxygen tank reaches 7.56 million euros, while the CAPEX of oxygen tank in BAU is 0.50 million euros. The remaining costs include compressors and oxygen pipelines, these costs are 3.13 million euros.

The annual saving of 8.08 million euros is used to produce green hydrogen. Divided by the amount of green hydrogen produced (120 million kilogram), 0.07 euros can be saved per kilogram of H<sub>2</sub>.

#### 3.2. Sensitivity analysis

The results of this analysis will be somewhat impacted by the actual adjustments made to the LOX tank capacity and pipeline length based on the particular project. Therefore, the

length of the pipeline and the oxygen storage tank's capacity received a sensitivity analysis in this study.

It is obvious that a significant share of the costs of new energy systems are accounted for by LOX tank CAPEX. The cost can be further decreased if we can reduce the capacity of the LOX storage tanks. The following expression is used to define the storage factor: LOX tank capacity/total LOX volume passing through the tank. The conclusion we previously described has a storage factor of 15%, and the tank can hold 144000 tonnes of liquid oxygen. Table 2 shows the influence of tank capacity on savings: savings decrease when tank capacity increases. When the storage factor is more than 0.31, the savings become negative. In order to save more costs for new energy systems, we need to minimize tank capacity while ensuring the normal operation of the system, This can be done by adjusting the rated power of the wind park and the rated power of the electrolyzer.

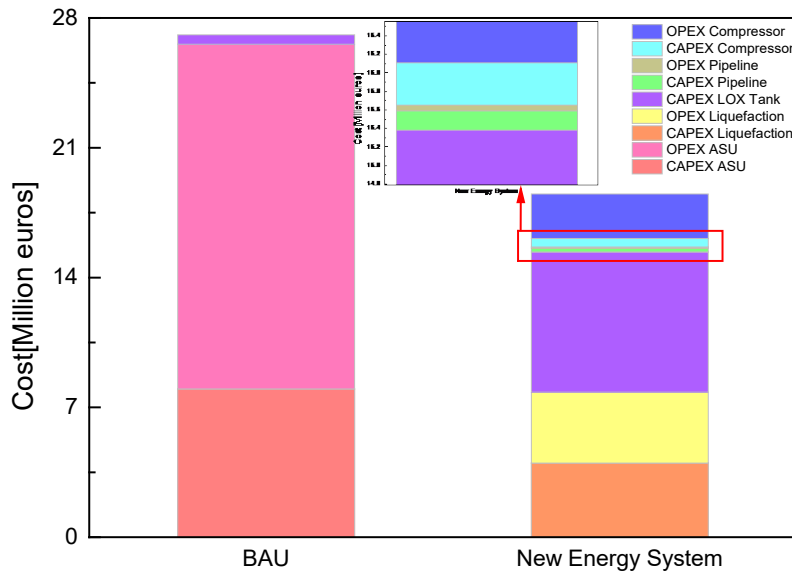


Figure 2. Yearly CAPEX and OPEX for BAU and New Energy System (Million Euros)

Table 2. Sensitivity Analysis of LOX Tank Capacity

The storage factor	Tank capacity/t	Savings/M€
0.05	48000	13.12
0.10	96000	10.06
0.15	144000	8.08
0.20	192000	5.56
0.25	240000	3.04

In this analysis, the pipeline's length is fixed at 5 km, which is the optimal circumstance. The length of the pipeline will actually be significantly extended in several nations with greater landmasses. Savings are impacted by pipeline length in Table 3 and decline as pipeline length rises. If the pipeline is longer than 155 kilometers, the savings are zero. As a result, we must minimize the distance between the green hydrogen system and the power plant in order to save expenditures. Therefore, it is more cost-effective to convert

coal-fired power plants to oxy-fuel combustion in regions with a wealth of renewable energy supplies.

Table 3. Sensitivity Analysis of pipeline length

The pipeline length/km	Pipeline cost/M€	Savings/M€
5	0.27	8.08
25	1.35	7.00
50	2.7	5.65
100	5.4	2.95

#### 4. Conclusion

The integration of wind energy, a water electrolysis system, and an oxygen-enriched combustion power plant creates the suggested new energy system as a whole. In order to achieve a low-carbon, economic goal, the electrolytic water system by-product of oxygen for oxygen-enriched combustion power plants thoroughly remove the air separation plant. This increases coal-fired power plants' efficiency, lowers carbon dioxide emissions, and lowers the cost of retrofitting coal-fired power plants for oxygen-enriched combustion. From the results of cost comparison, when the tank capacity is 144,000 tonnes, the new energy system can save 8.08 million euros annually compared to BAU. Green hydrogen generation can be produced for 0.07 euros less per kilogram if cost savings are computed to do so. Larger cost savings are possible with smaller tank capacities and shorter pipeline length. The goal can then be achieved by optimizing the rated power of the wind park and the rated power of the electrolyzer in accordance with the 1-year wind turbine output curve and converting coal-fired power plants to oxy-fuel combustion in regions with a wealth of renewable energy supplies.

#### References

- IRENA. 2022. GLOBAL HYDROGEN TRADE TO MEET THE 1.5°C CLIMATE GOAL.
- L. Ge, Q. Cui, M. Li, et al. 2022. Review on water electrolysis for hydrogen production powered by fluctuating wind power and PV. *Integrated Intelligent Energy*, 2022,44(05):1-14.
- H. Wang, W. Xu, Z. Zhang. Development status and suggestions of green hydrogen energy produced by water electrolysis from renewable energy. *Chemical Industry and Engineering Progress*, 2022,41(S1):118-131.
- J. Deng. Study on carbon content per unit calorific value of coals in China. *Coal Processing & Comprehensive Utilization*, 2022,(01):58-62.
- Horn F L, Steinberg M. Control of carbon dioxide emissions from a power plant (and use in enhanced oil recovery). *Fuel*, 1982, 61(5)415-422.
- B. M. Abraham, J. G. Asbury, E. P. Lynch, et al. Coal-oxygen process provides CO<sub>2</sub> for enhanced recovery. *Oil Gas J*, 1982, 80(11).
- Pikkarainen T, Saastamoinen J, Saastamoinen H, et al. Development of 2nd generation oxyfuel CFB technology-Small scale combustion experiments and model development under high oxygen concentrations. *Energy Procedia*, 2014, 63: 372-385.
- Berenschot. 2019. Oxygen synergy for hydrogen production.
- G. Cai. 2022. Analysis on the influence of changing operation conditions on energy consumption of air separation plant. *Copper Engineering* 2022(02):81-83.
- Javier Serrano González, César Álvarez Alonso. Industrial electricity prices in Spain: A discussion in the context of the European internal energy market. *Energy Policy*, 2021, 148.

# CO<sub>2</sub> water-lean capture: Mathematical modelling and analysis of the industrial process

Ulderico Di Caprio<sup>a</sup>, Emine Kayahan<sup>a</sup>, Min Wu<sup>a</sup>, Peter Hellinckx<sup>b</sup>, Tom Van Gerven<sup>c</sup>, Steffen Waldherr<sup>d,e</sup> and M. Enis Leblebici<sup>a,\*</sup>

<sup>a</sup>Center for Industrial Process Technology, Department of Chemical Engineering, KU Leuven, Agoralaan Building B, 3590 Diepenbeek, Belgium

<sup>b</sup>Faculty of Applied Engineering, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerpen, Belgium

<sup>c</sup>Process Engineering for Sustainable Systems, Department of Chemical Engineering, KU Leuven, Celestijnenlaan 200F, B-3001 Heverlee, Belgium

<sup>d</sup>KU Leuven, Department of Chemical Engineering, Celestijnenlaan 200F-bus 2424, 3001 Leuven, Belgium

<sup>e</sup>University of Vienna, Molecular Systems Biology (MOSYS), Department of Functional and Evolutionary Ecology, Faculty of Life Sciences, 1030 Vienna, Austria

\*[muminenis.leblebici@kuleuven.be](mailto:muminenis.leblebici@kuleuven.be)

## Abstract

The most mature post-combustion capture technology is absorption of CO<sub>2</sub> into monoethanolamine (MEA). Typically, this process utilizes a solution of 30wt.% aqueous MEA as absorbent liquid; however, recent studies showed how increasing MEA concentration enhances mass-transfer coefficient in absorption columns. This work investigates effect of MEA concentration on the overall process. The main focus is given to regeneration heat. Heat duty at the reboiler per unit of recovered CO<sub>2</sub> (HDUC) is utilized as the primary performance indicator. The analysis employs an Aspen HYSYS<sup>®</sup> model. Increasing MEA concentration reduces HDUC by 74% when MEA concentration goes from 30wt.% to 90wt.%. Another parameter affecting HDUC is stoichiometric ratio between moles of MEA and CO<sub>2</sub>. Halving molar ratio between MEA and CO<sub>2</sub>, at 30 wt.% MEA concentration returns a reduction in HDUC of 48%, while at 90wt.% MEA concentration, the observed reduction is 9%. Overall, increasing MEA concentration improves process efficiency.

**Keywords:** CO<sub>2</sub> capture, monoethanolamine, water-lean solvent, HYSYS, heat duty

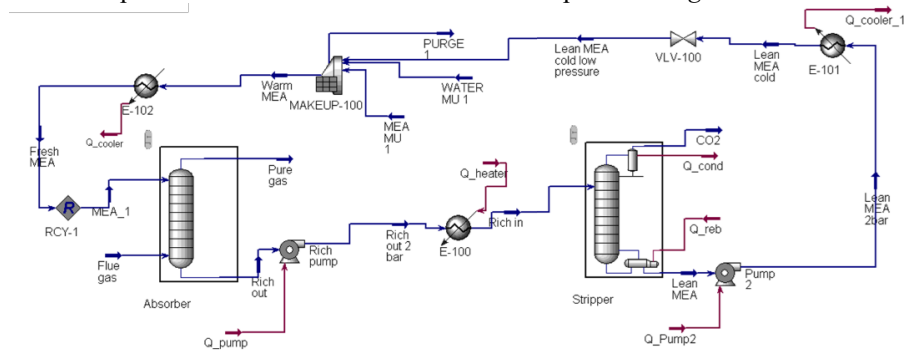
## 1. Introduction

The anthropogenic CO<sub>2</sub> emission reinforces the global warming phenomena. The International Panel on Climate Change recently highlighted how humanity is urged to reduce CO<sub>2</sub> emission to prevent the global average temperature from rising and hitting 1.5°C. Besides alternative energy sources, several technologies are available to reduce anthropogenic CO<sub>2</sub> emissions, such as oxyfuel technologies, precombustion capture and post-combustion capture. Among these technologies, post-combustion capture is the most mature solution and requires a minor modification to the currently available production processes (Luis, 2016). This process is executed in a cycling system. CO<sub>2</sub> in a flue gas is captured within absorption column using monoethanolamine (MEA) aqueous solutions. CO<sub>2</sub>-rich amine solution is then processed in stripping column. Here, CO<sub>2</sub> is removed from the amine solution at a high temperature and collected from the head of the column.

The CO<sub>2</sub>-lean amine solution is then conveyed to absorption column to serve as absorbent liquid. Typically, this process utilizes a solution of 30wt.% aqueous MEA. Despite this being the most mature technology, its widespread usage is limited by high regeneration costs, as highlighted by Sreedhar et al. (2017). To reduce the required regeneration heat and increase mass-transfer coefficients within absorption columns, Kayahan et al. (2022) proved the benefits of the CO<sub>2</sub> post-combustion process when employing an MEA solution with low water concentration. The increased transfer coefficients allow lower dimensions of the absorption columns or a higher flue gas flow rate processable by the plant. It has been shown that reducing water amount in the MEA solution can lead to 3-10 fold decrease in the reboiler heat duty. However, this calculation did not consider the kinetics of the reaction. In the paper, it has been shown by NMR analysis that using 90wt.% MEA leads to a significant increase in the carbamate formation compared to other aqueous MEA concentrations. This suggests change in the reaction mechanism of CO<sub>2</sub> absorption into MEA at and above 90wt.% MEA. In this work, we are still not accounting for this change. However, by using a commonly accepted rate-based model of CO<sub>2</sub> absorption into MEA, we aim to calculate the heat of desorption more accurately with increasing MEA concentrations. The analysis employs a mathematical model of the post-combustion CO<sub>2</sub> capture process. More specifically, the analysis will be executed in Aspen HYSYS<sup>®</sup>. Despite the advanced optimisation techniques deployed in the simulation software, this study will not employ them since it does not aim to obtain the optimal process condition. On the contrary, this study focuses on the effect of the MEA concentration on process efficiency. In addition, such behaviour will be correlated to, and explained with, the physical and chemical phenomena underlying the capture process.

## 2. Methodology

This work investigates effect of MEA concentration using a mathematical model of the process. The model is executed using Aspen HYSYS<sup>®</sup> as simulation software, and its property package “Acid Gas – Chemical Solvents” has been employed to predict physicochemical behaviour of the mixture. It considers the rate-based model proposed by Zhang et al. (2009). In this work we assume that the CO<sub>2</sub> absorption into MEA follows the same kinetic within the entire MEA concentration range. The plant scheme implemented in the simulation was reported in the literature by Øi (2007). The absorber column comprises 10 equilibrium stages, and the desorption column comprises 6 equilibrium steps. The stripper worked at 120°C with an overall recovery efficiency of 85%. The representation of the simulation model is reported in *Figure 1*.



*Figure 1. Overall representation of the HYSYS model employed in this work. The system comprises an absorber, reported in the left-hand of the picture, and a stripper, reported in the right-hand of the picture.*

The gas composition was obtained from the paper of Arachchige et al. (2012). It considers the typical flue gas obtained by a gas-fired plant. Namely, the gas composition is the following: 4% CO<sub>2</sub>, 8% H<sub>2</sub>O, 76% N<sub>2</sub>, and 12% O<sub>2</sub>. The inlet temperature of the gas is 40°C.

The absorbent liquid was composed of MEA and water. The model also includes make-up for absorbent solution, which tackles absorbent liquid losses because of evaporation in the reboiler. Fresh absorbent mixture is fed at this point and is mixed with the CO<sub>2</sub>-lean absorbent solution leaving the stripper. The latter contains a load of not-stripped CO<sub>2</sub>; therefore, the absorber is fed with an absorbent solution containing a certain amount of CO<sub>2</sub>.

MEA concentration was varied in the study to investigate their effects on the energy efficiency of the reboiler. More specifically, the evaluation was executed on stripper column reboiler heat duty. Together with the reboiler duty, the metric utilized to evaluate the system performance is heat duty at reboiler per unit of recovered CO<sub>2</sub> (HDUC). In this work we assume that the pumping cost of the liquid does not change with the MEA concentration. Therefore, lower values of HDUC indicate more energy-efficient process. The following MEA concentrations were utilized in this study: 20wt.%, 30wt.%, 50wt.%, 70wt.%, and 90wt.%. Effect of absorbent liquid inlet temperature on the absorber column on the HDUC was also investigated in this study. The absorber inlet liquid temperature was set to 30°C, 40°C and 50°C. All the experiments mentioned above were carried out with a molar ratio between MEA in absorbent liquid and CO<sub>2</sub> in flue gas ( $n_{\text{MEA}}/n_{\text{CO}_2}$ ) of 10. The effect of  $n_{\text{MEA}}/n_{\text{CO}_2}$  was investigated. In addition,  $n_{\text{MEA}}/n_{\text{CO}_2}$  of 6 was tested to investigate the effect of this parameter. The molar ratio was tuned by changing liquid flow rate going to absorption column.

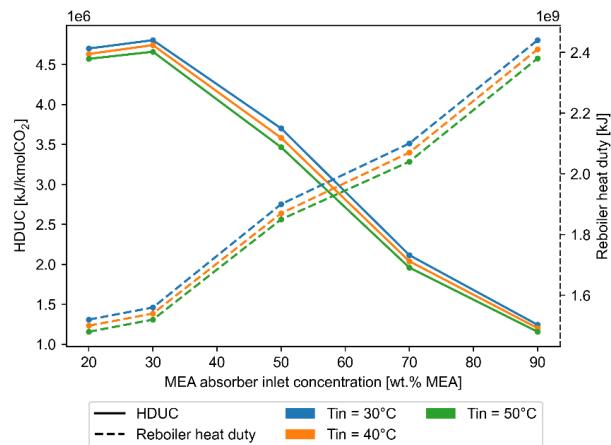
### **3. Results and discussion**

#### *3.1. Effect of the MEA concentration*

In *Figure 2*, profile of HDUC and reboiler duty are reported with changing MEA mass fraction at the make-up outlet. The curves within the plot are reported at different liquid inlet temperature to the absorber. Increasing MEA mass fraction, HDUC slightly increases from 20wt.% to 30wt.% and steadily decreases when moving MEA fraction from 30wt.% to 90wt.%. On the contrary, overall heat duty at reboiler steadily increases with mass fraction of MEA. The presence of the peak is investigated in Section 3.2. Changing MEA mass fraction from 30wt.% to 90wt.%, decreased HDUC by 74% while it increased reboiler heat duty by 60%. We hypothesize it is related to the absorption heat and evaporation capacity of the absorbent liquid. Energy given by the reboiler of the stripping column is employed to 1) evaporate water and MEA, and 2) strip the CO<sub>2</sub> molecules from the liquid. At 120°C, the latent heat of the MEA is 54.7kJ/mol MEA (McDonald et al. (1959)), while for water, it is 36.71kJ/mol H<sub>2</sub>O. The absorption heat of CO<sub>2</sub> into aqueous MEA solution at 120°C is around 105 kJ/mol CO<sub>2</sub>; this value is almost independent of MEA concentration in the range 10wt.% to 70wt% MEA (Kim et al. (2014)). Therefore, moving from 30wt.% to 70wt.% MEA concentration, absorption heat does change its value. Consequently, in this range, the increasing of the heat duty is only related to the change in the absorption liquid composition. Using the value reported above and assuming the mixture enthalpy to be negligible, latent heat of an aqueous MEA solution at 30wt.% is 30.72kJ/mol while a solution with 70wt.% MEA has a latent heat of 44.04kJ/mol. Thus, the latent heat increases by 43% when the MEA mass fraction increases, which explains the reboiler heat duty increase. However, the value increases when energy is normalized with the amount of released CO<sub>2</sub> (i.e., HDUC). Increasing



MEA concentration increases absorbent amount faster than the latent heat. The result is in line with our previous findings. Kayahan et al. (2022) proved that increasing MEA concentration increased overall mass transfer coefficients and overall capture efficiencies. In addition, in *Figure 2*, the curves are parametric with the inlet temperature of the absorbent liquid within the absorption column ( $T_{in,abs}$ ). Even if the effect is negligible, when increasing  $T_{in,abs}$  both reboiler duty and HDUC decrease. In other words, increasing the temperature at the absorber decreases the amount of energy required in the stripping step. We hypothesize this is related to the temperature profiles established within the loop at the steady state. Increasing the inlet temperature increases the temperatures of the various streams in the system. Therefore, increasing  $T_{in,abs}$  also increases the liquid inlet temperature at the stripper. Therefore, the reboiler must supply less sensible heat to let the stream arrive at boiling.



*Figure 2. Profile of reboiler heat duty per unit of recovered CO<sub>2</sub> (HDUC) and reboiler duty. The HDUC is reported on the left-hand y-axis. The reboiler heat duty is reported on the right-hand y-axis. The HDUC decreases by 74% when the concentration of MEA is increased from 30wt.% to 90wt.% within the absorbent liquid. On the contrary, the reboiler duty increases by 60% in the same MEA mass fraction range.*

### 3.2. Effect of the molar ratio between MEA and CO<sub>2</sub>

*Figure 3* shows the behavior of the HDUC parametric in nMEA/nCO<sub>2</sub>. Namely, two cases are analyzed in this study with an nMEA/nCO<sub>2</sub> of 6 and 10. For both the molar ratios, the HDUC decreases when increasing the MEA mass fraction from 20wt.% to 90wt.%. However, as also previously mentioned, the curves present a local maximum. The local maximum is located around 30wt.% MEA when the molar ratio is 10, while it is around 50wt.% MEA when the molar ratio is 6. In other words, the peak shifts to higher values of MEA mass fraction when decreasing the nMEA/nCO<sub>2</sub>. We hypothesize that the existence of the peak and its shifting toward higher MEA mass fractions with decreasing nMEA/nCO<sub>2</sub> is related to the amount of the absorbed CO<sub>2</sub> in the absorption column. In the left-hand of the peak, increasing the MEA mass fraction increases the HDUC; therefore, the energy required for the stripping increases faster than the released CO<sub>2</sub> at the stripper. On the right side of the peak, the released CO<sub>2</sub> increases faster than the energy required for the stripping; this is related to the absorbed amount of CO<sub>2</sub> within the absorbent liquid. Here, increasing the MEA mass fraction increases the amount of absorbed CO<sub>2</sub>; however, as previously highlighted, the desorption heat increase is negligible with the MEA mass fraction in the liquid. Therefore, since the required heat is

mainly given by the latent heat of the absorbent liquids, the reboiler duty is spread over a higher amount of CO<sub>2</sub> at high MEA mass fractions. Thus, the energy spent to strip one mole of CO<sub>2</sub> decreases by increasing the MEA mass fraction within the absorbent liquid. The peak shift to higher MEA mass fractions when decreasing the molar ratio means that the trade-off between the energies is reached for higher MEA mass fractions when the molar ratio is lower. This effect is related to the relative presence of MEA. The MEA is the main component driving the CO<sub>2</sub> capture into the absorbent liquid; thus, increasing the concentration of MEA increases the number of molecules absorbed by the liquid. Decreasing nMEA/nCO<sub>2</sub> decreases the amount of MEA molecules available to run the capture. Therefore, a higher MEA mass fraction in the absorbent liquid is needed to reach the absorption energy trade-off. We hypothesize this is the main reason driving the peak shift toward higher MEA mass fraction when decreasing the nMEA/nCO<sub>2</sub> ratio.

Figure 3 shows a different decrease in the HDUC with the nMEA/nCO<sub>2</sub> at 20wt.% and 90wt.%. At 20wt.%, the HDUC decreases by 48% when moving the nMEA/nCO<sub>2</sub> ratio from 10 to 6. On the contrary, the decrease is only 9% when the MEA mass fraction is 90wt.%. This effect is connected to the amount of water utilized in the system. Increasing the nMEA/nCO<sub>2</sub> ratio causes an increase in the ratio between the water in the absorbent liquid and the CO<sub>2</sub> in the flue gas (nH<sub>2</sub>O/nCO<sub>2</sub>). For low MEA mass fractions, nH<sub>2</sub>O/nCO<sub>2</sub> increases faster than nMEA/nCO<sub>2</sub> because of the higher presence of water within the liquid. However, the water weakens the absorption of the CO<sub>2</sub> molecules, but it still requires energy from the reboiler duty to arrive at the stripping temperature. Therefore, it behaves as an inert specie that still requires energy. Consequently, by reducing the ratio nMEA/nCO<sub>2</sub>, the HDUC decreases more than having a solution with a higher MEA mass fraction. The effects of the water as inert are mitigated at high MEA mass fractions. The substantial increase in the HDUC at higher nMEA/nCO<sub>2</sub> ratios does not occur since nMEA/nCO<sub>2</sub> changes faster than nH<sub>2</sub>O/nCO<sub>2</sub>. Therefore, the increase in the reboiler duty when increasing the ratio nMEA/nCO<sub>2</sub> is compensated by more CO<sub>2</sub> molecules absorbed within the liquid.

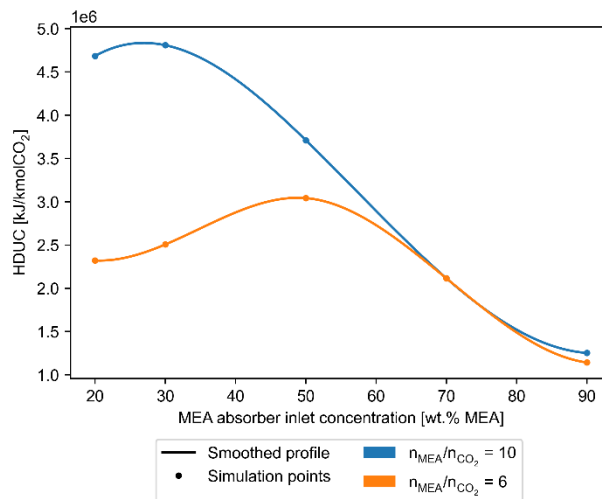


Figure 3. Heat reboiler duty per unit of recovered CO<sub>2</sub> as a function of the MEA wt.% in the absorbent liquid. The curves are parametric with the molar ratio between CO<sub>2</sub> and MEA. By decreasing the molar ratio from 10 to 6, the HDUC decreases. At 20wt.% MEA, the HDUC decreases 48%, while at 90wt.% MEA the HDUC decreases 9%.

#### 4. Conclusion

In this work, we analysed the effect of the monoethanolamine (MEA) concentration on the reboiler energy efficiency of the CO<sub>2</sub> capture process. The heat duty per mole of recovered CO<sub>2</sub> (HDUC) has been utilized as the main key performance indicator. Increasing the MEA mass fraction in the absorbent liquid increases the heat duty at the stripper by 60%, but it decreases the HDUC by 72%. This effect is mainly related to the latent heat of the absorbent liquid and the independency of the CO<sub>2</sub> desorption heat with the MEA mass fraction in the absorbent liquid. In addition, we showed how the HDUC is affected by the molar ratio between MEA in the absorbent liquid and CO<sub>2</sub> in the flue gas. Increasing the ratio at a fixed MEA mass fraction generally corresponds to an increase in the HDUC. However, this effect is not constant toward the entire MEA mass fraction range. At 20wt.%, the HDUC decreases by 48% when moving the ratio nMEA/nCO<sub>2</sub> from 10 to 6, while at 90wt.%, the HDUC decreases by only 9%. Overall, this work showed how increasing the MEA mass fraction within the absorbent liquid increase the energy efficiency of the reboiler. However, for this study, only the energy of the reboiler has been considered. Further investigations are required to investigate the impact of the higher MEA concentration on the pumping cost and evaporation losses.

#### Acknowledgements

The authors acknowledge funding from VLAIO Moonshot HBC.2019.0109 and HBC.2021.0255 “Intensification of CO<sub>2</sub> capture processes (CAPTIN)”.

#### References

- U.S.P.R. Arachchige, M.C. Melaaen, 2012, Aspen Plus Simulation of CO<sub>2</sub> Removal from Coal and Gas Fired Power Plants, *Energy Procedia*, 23, 391-399
- E. Kayahan, U. Di Caprio, A. Van den Bogaert, M.N. Khan, M. Bulut, L. Braeken, T. Van Gerven, M.E. Leblebici, 2023, A new look to the old solvent: Mass transfer performance and mechanism of CO<sub>2</sub> absorption into pure monoethanolamine in a spray column, *Chemical Engineering and Processing - Process Intensification*, 184, 109285
- I. Kim, K.A. Hoff, T. Mejdell, 2014, Heat of absorption of CO<sub>2</sub> with aqueous solutions of MEA: new experimental data, *Energy procedia*, 63, 1446-1455
- P. Luis, 2016, Use of monoethanolamine (MEA) for CO<sub>2</sub> capture in a global scenario: Consequences and alternatives, *Desalination*, 380, 93–99
- R.A. McDonald, S.A. Shrader, D.R. Stull, 1959, Vapor Pressures and Freezing Points of Thirty Pure Organic Compounds, *Journal of Chemical and Engineering Data*, 4, 4, 311-313
- L. E. Øi, 2007, Aspen HYSYS simulation of CO<sub>2</sub> removal by amine absorption from a gas based power plant, *SIMS2007*, 73-81
- I. Sreedhar, T. Nahar, A. Venugopal, B.Srinivas, 2017, Carbon capture by absorption–Path covered and ahead, *Renewable and Sustainable Energy Reviews*, 76, 1080-1107
- Y. Zhang, H. Chen, C. Chen, J.M. Plaza, R. Dugas, G.T. Rochelle, 2009, Rate-Based Process Modeling study of CO<sub>2</sub> Capture with Aqueous Monoethanolamine Solution, *Industrial & Engineering Chemistry Research*, 48, 20, 9233-9246

# Stochastic pinch analysis to target for resources in batch processes

Md Alquma Haider, Nitin Dutt Chaturvedi\*

*Process Systems Engineering Laboratory, Department of Chemical and Biochemical Engineering, Indian Institute of Technology Patna, Bihta, Patna, 801106, Bihar, India*

*nitind@iitp.ac.in*

## Abstract

Resource conservation in batch processes is critical for process industries. For minimizing resource consumption in batch and continuous processes, Pinch Analysis is one of the best insight-based optimization techniques. Variation in operational conditions and/or environmental changes may cause changes in quality and flow. Uncertain quality and flow can be addressed using distinct probability distributions with known mean and standard deviations. The objective of this work is to minimize the resource in a batch process with uncertain flow. In order to transform the probabilistic flow constraints associated with the sources into deterministic ones, a chance-constraint programming approach is used. An illustrative example is given to show how the suggested method works, with analysis based on three distinct distributions (the normal distribution, the lognormal distribution, and Chebyshev's one-sided inequality).

**Keywords:** Stochastic Pinch Analysis, Batch process, Resource

## 1. Introduction

. Process Systems Engineering's contributions in the area of optimization under uncertainty can enhance numerous fields, including synthesis, design, planning, scheduling, control, and supply chain management. Studying optimization under uncertainty has been prompted by the prevalence of uncertain parameters in optimization models for synthesis, design, planning, scheduling, and supply chains. Stochastic programming is an approach to the problem of uncertainty in which a subset of decisions is made in the belief that corrective actions will be available over a specified scenario tree with discrete probabilities of the uncertainties[1].

Batch processes make it possible to make many different kinds of products and share resources [2]. The pharmaceutical, polymer, food, and specialty chemical industries have all made extensive use of batch processing.[3] The most important area for enhancement in batch processes is resource conservation or the efficient use of resources. Haider and Chaturvedi[4] have created a Pinch Analysis(PA)-based methodology for specialising in the conservation of resources through the use of dedicated sources in segregated targeting problem for batch processes.The addition of time makes the design and synthesis of resource allocation networks for a batch process more difficult. However, as product-oriented industrialization has become increasingly emphasized, more systematic integration and design methodologies have been developed for batch processes considering uncertain parameters.

This paper focuses on the water allocation problem (WAN) and how to handle it when there are unknown parameters. Parametric uncertainty associated with source flows is incorporated into resource allocation networks using the chance-constrained approach[5]

and the PA framework. Modeling these parameters' unknown values requires the use of probability distribution functions whose mean and standard deviation are already known. Through the use of a chance-constraint programming strategy, the probabilistic flow constraints related to the sources can be converted into deterministic ones. The proposed method is explained with the help of an example, which employs analysis based on three different distributions (the normal distribution, the lognormal distribution, and Chebyshev's one-sided inequality).

## 2. Problem statement and mathematical formulation

The following mathematical expression describes the general stochastic Pinch problem, which seeks to optimise the resource in a batch network subject to uncertainties in source parameters.

There is a set of  $N_s$  internal sources available. Each internal source generates a uncertain flow  $F_{is}$  (with a mean and standard deviation) at a certain quality  $q_{is}$  at a definite time interval. It is also known that there is a set of  $N_d$  internal demands that must be met, and that each of these demands is willing to accept a flow of some quality that is below some maximum quality limit  $F_{dj}$ . Each demand is also bound with definite time interval. To meet the unmet needs, there is also access to an external source called resource, which has an known quality  $q_r$  and no flow restriction. Resource is available all the time. Unused flows from internal sources are routed to an external demand known as waste, with no caps on quality or quantity.

Let  $f_{ij,t}$  denote the flow from  $i^{th}$  source to  $j^{th}$  demand at a given time interval (t). For a given time interval (t),  $f_{iw,t}$  denote flow from  $i^{th}$  source to the waste (w) while  $f_{oj,t}$  represent flow from fresh water (resource) to the  $j^{th}$  demand. In accordance with the flow conservation principle, the following expressions can be used to express the flow balance for all internal sources and sinks for batch processes.  $\sum_{j=1}^{N_d} f_{ij,t} + f_{iw,t} = F_{si,t} \quad \forall i$

(1)

$$\sum_{i=1}^{N_s} f_{ij,t} + f_{oj,t} = F_{dj,t} \quad \forall j \quad (2)$$

Where

$F_{si,t}$ : total flow from  $i^{th}$  source at a particular time interval (t).

$F_{dj,t}$ : total flow in  $j^{th}$  demand at a particular time interval (t).

$C_{si,t}$ : contaminant concentration of  $i^{th}$  source at a particular time interval (t).

$C_{s0,t}$ : contaminant concentration of fresh source at a particular time interval (t).

$C_{dj,t}$ : contaminant concentration of  $j^{th}$  demand at a particular time interval (t).

The contaminant constraint can be mathematically expressed for any internal demand as

$$\sum_{i=1}^{N_s} f_{ij,t} C_{si,t} + f_{oj,t} C_{s0,t} \leq F_{dj,t} C_{dj,t} \quad \forall j \quad (3)$$

A WAN's primary goal is to decrease the demand for fresh water, and this can be expressed mathematically as follows:

$$\text{Minimize } R(t) = \sum_{j=1}^{N_d} f_{oj,t} \quad (4)$$

Since it is not possible to guarantee that the constraints in eqs. 1 and 3 will be met in the presence of uncertainty in the source flow ,these constraints are treated as probabilistic in order to account for the parametric uncertainty.The source flow constraints in eq. 1 has reliability parameter, and denote them by  $\beta(t)$ .The probabilistic constraints in Equations 2 is rewritten as follows.

$$Prob\left(\sum_{j=1}^{N_d} f_{ij,t} \leq F_{si,t}\right) \geq \beta(t) \quad \forall i \tag{5}$$

A stochastic technique called chance-constrained programming uses the probability distribution of the random variables to transform a probabilistic constraint into a deterministic constraint. It can therefore be utilised to resolve an optimization issue with parametric uncertainty. The probabilistic restrictions in equations 5 is transformed into deterministic equivalents in this study using the chance-constrained programming methodology.It should be noted that uncertainty in source quality is not considered here.

Because  $f_{iw,t}$  in equation (1) is a slack variable, the probabilistic constraint (eq. (5)) can be rewritten as

$$Prob\left(\sum_{j=1}^{N_d} f_{ij,t} \leq F_{si,t}\right) \leq (1 - \beta(t)) \quad \forall i \tag{6}$$

It is possible to write a deterministic form of eq. (7) using chance-constrained programming as follow:

$$\sum_{j=1}^{N_d} f_{ij,t} \leq P_{F_{si,t}}^{-1} (1 - \beta(t)) \quad \forall i \tag{7}$$

Where  $P_{F_{si,t}}$  represent source flow probability distribution time function

For the purpose of determining effective flow, the following equations[6] are used:

$$Effective\ flow = \mu_{i,t} - z_{\beta,t} \sigma_{i,t} \quad \forall i \tag{8}$$

$$Effective\ flow = e^{m_{i,t} - z_{\beta,t} s_{i,t}} \tag{9}$$

Where

$$s_{i,t} = \sqrt{\ln\left(1 + \left(\frac{\sigma_{i,t}}{\mu_{i,t}}\right)^2\right)} \quad \text{and} \quad m_{i,t} = \ln(\mu_{i,t}) - \frac{s_{i,t}^2}{2}$$

- $\mu_{i,t}$ : mean of  $i^{th}$  source for a given time interval (t)
- $z_{\beta,t}$ : inverse cumulative probability distribution function of a standard normal distribution for a given time interval (t)
- $\sigma_{i,t}$ : standard deviation of  $i^{th}$  source for a given time interval (t)
- $s_{i,t}$  and  $m_{i,t}$ : two parameters to characterize the log-normal distribution for a given time interval (t)

$$Effective\ flow = \mu_{i,t} - \sqrt{\frac{\beta(t)}{1-\beta(t)}} \sigma_{i,t} \quad \forall i \tag{10}$$

Eq. 8 is used for calculation of effective flow for normal distribution . Eq. 9 is used for calculation of effective flow for lognormal distribution. Eq. 10 is used for calculation of effective flow for Chebyshev’s one-sided inequality.

The overall goal of stochastic optimization problem is to reduce the amount of resources needed (eq. 4) in a batch resource allocation network while maintaining the required reliability of all sources and demands.

### 3. Proposed Algorithm

To reduce the amount of resources needed for a batch process, the following algorithm has been proposed:

1. First, the entire time horizon of the batch process is divided into several time intervals (say  $(I_1, I_2, I_3, \dots, I_m)$ ) such that all sources and/or demands must start or end at the end points of these intervals.
2. Determine the bare minimum of resources needed for each time interval without taking uncertainty into account using any PA technique like source composite curve [7] (i.e., first solve for the deterministic case).
3. The waste profile is transferred in a sequential manner for all upcoming time periods [8].
4. Using any PA technique, calculate the minimum resource requirements for each interval while taking into account uncertainty in one parameter, say source flow. Make an effective flow calculation for each of the three possible distributions. For the purpose of determining effective flow eqs. (8),(9),and (10) are used.
5. Create a comparison plot for each interval for all three distributions.
6. Develop a network for allocating resources, making sure to account for efficient flows and specific time periods.

### 4. Illustrative example

In order to demonstrate how the proposed algorithm operates, an example will be used. There are six different sources of water, but only one freshwater resource, so the water conservation problem has six different need demands. The fact that the source flow is uncertain by 10% is taken into consideration here. The data for the source and the demand are shown in Table 1, with the only uncertain variable being the source flow. Table 1. Source and demand data

Source	Concentration(ppm)	Flow(t)		Duration(h)
		Mean	S.D	
S1	330	40	4	1-3
S2	300	45	4.5	1-3
S3	255	42	4.2	0-3
S4	200	32	3.2	0-2
S5	150	15	1.5	0-1
S6	100	50	5	2-3
FW	10			0-3
Demand		Flow(t)		
D1	110	40		0-3
D2	100	45		0-2
D3	125	42		2-3
D4	50	32		2-3
D5	20	15		0-3
D6	15	50		0-2

In accordance with Step 1, the entire batch process is broken up into three distinct intervals, which are referred to as 0-1, 1-2, and 2-3 h. The next step, which is called Step

2, is to calculate the minimum resource requirements. For the first interval, the minimum resource requirements are calculated to be 43.288 t/h without taking into consideration any uncertainty. In accordance with step 3, the waste that was produced during the initial interval will be transported to the subsequent interval in order for it to be incorporated between the intervals. In the source flow with interval integration, the minimal resource requirements for upcoming intervals are 47.235 and 30.692 t/h, respectively. Equations 8, 9, and 10 are used to figure out the effective flow for the normal distribution, the lognormal distribution, and Chebyshev's one-sided inequality, as demonstrated in Step 4. These effective flows are computed for a range of source reliabilities, including 90%, 80%, 70%, 60%, and 50% respectively. It is highly likely that the freshwater requirement ought to be increased in order to make source flows more dependable. Figure 1 illustrates how the freshwater requirements of various distributions shift in response to increasing levels of source reliability. Chebyshev's distribution requires more fresh water than a normal distribution does, and the log-normal case requires a significantly greater quantity of fresh water at each interval.

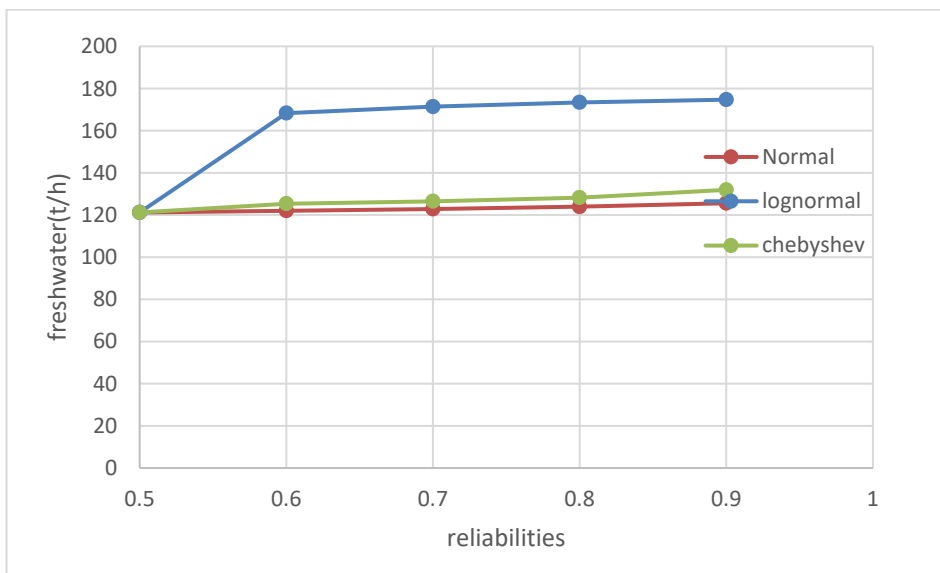


Figure 1: Variation of freshwater with reliability of sources with uncertain flow for different distributions

## 5. Conclusion

This study presents a conceptual methodology for resource targeting in batch processes utilising the idea of stochastic optimization. The methodology is based on pinch analysis, and it uses the idea. The constraints are probabilistic because of the presence of source flow uncertainty; however, they can be transformed into a deterministic form by utilising chance-constrained programming with the normal distribution or Chebyshev's one-sided inequality. The waste profile is transferred sequentially for all future time periods in order to reduce the amount of waste produced. It should be noted that the normal distribution is a natural choice for such investigations; Chebyshev's one-sided inequality, which



produces very conservative results, can be used for any unknown distribution. Pinch analysis techniques that have been tried and tested have developed the concepts of effective flow in order to solve the resulting problem in a manner that can be anticipated. The mean and standard deviation of flow metrics, in addition to the desired level of constraint reliability, are used to calculate these effective values. The concept of effective flow was first presented as part of the standard pinch analysis methods in order to deterministically address this issue.

## References

1. I.E.Grossmann,R.M.Apap,B.A.Calfa,P.García-Herreros,Q. Zhang,2016. Recent advances in mathematical programming techniques for the optimization of process systems under uncertainty. *Computers & Chemical Engineering*,91,3-14.
2. N.D.Chaturvedi,Z.A.Manan,2021. Batch process integration for resource conservation toward cleaner production–A state-of-the-art review. *Journal of Cleaner Production*,318, .128609.
3. G.Shukla,N.D.Chaturvedi,2021.Targeting compression energy in batch process. *Cleaner Engineering and Technology*, 5, 100315.
4. M.A.Haider,N.D.Chaturvedi,2021.Segregated targeting for resource conservation with dedicated sources for batch process. *Chem. Eng. Trans.* 88, 175–180.
5. A.Charnes,W.W. Cooper,1959.Chance-constrained programming. *Management Science*. 6, 73–79.
6. D.Arya,K. Shah,A.Gupta,S.Bandyopadhyay, 2018. Stochastic pinch analysis to optimize resource allocation networks. *Industrial & Engineering Chemistry Research*,57(48), 16423-16432.
7. S.Bandyopadhyay,2006. Source composite curve for waste reduction. *Chemical Engineering Journal*, 125(2), 99-110.
8. N.D.Chaturvedi,S.Bandyopadhyay,2013 Targeting for multiple resources in batch processes, *Chem. Eng. Sci.* 104 ,1081–1089.

# MINLP modelling and optimization of the supply chain for the renewable production of methanol in Mexico

Nereyda Vanessa Hernández-Camacho,<sup>a</sup> Fernando Israel Gómez-Castro,<sup>a,\*</sup> José María Ponce-Ortega,<sup>b</sup> Mariano Martín<sup>c</sup>

<sup>a</sup> *Universidad de Guanajuato, Campus Guanajuato, División de Ciencias Naturales y Exactas, Departamento de Ingeniería Química, Noria Alta S/N, Guanajuato, Guanajuato 36050, Mexico*

<sup>b</sup> *Universidad Michoacana de San Nicolás de Hidalgo, Departamento de Ingeniería Química, Francisco J. Múgica S/N, Morelia, Michoacán 58060, Mexico*

<sup>c</sup> *Universidad de Salamanca, Departamento de Ingeniería Química, Pza. Caídos 1-5, Salamanca 37008, Spain*  
*fgomez@ugto.mx*

## Abstract

Methanol production in Mexico is limited since it is produced from natural gas and only in a few facilities, causing a lack of supply and the need for import from other countries. Due to this, finding production alternatives such as biomass is necessary to satisfy the national demand. In this work, a supply chain to produce methanol from renewable and waste materials available in Mexico is proposed. The supply chain has been modelled following a generalized disjunctive programming approach, including as elements of the supply chain the election of the type and source of raw materials, the selection of the location of the facilities, the type of technology to be used, and the election of the market to be satisfied. The optimal route involves using 8 raw materials, 19 processing plants and the use of three different process technologies to help satisfying national demand with a profit of 7.34 MUSD/y.

**Keywords:** methanol, renewable sources, industrial waste, supply chain, optimization.

## 1. Introduction

Methanol is a highly relevant product. It is used as raw material to produce a diversity of derivatives, such as dimethyl ether, formaldehyde, among others. It is commonly obtained from sources such as natural gas and coal. Global methanol demand has grown over the last years, from 80 Mt in 2016 to 91.55 Mt by 2018 (Alvarado, 2016; Maddren, 2019). The only methanol production facility in Mexico is managed by the national enterprise PEMEX. Unfortunately, difficulties have been identified to satisfy the methanol demand, due to the uncertainties regarding the availability of natural gas. This opens an area of opportunity to search for alternative raw materials for methanol production. Such alternative routes must have the potential to reduce the environmental impact associated with the production of methanol, while efficiently satisfying the demand. However, it is important to evaluate the economic potential of such production schemes from a macroscopic point of view. The use of renewable sources is an interesting approach,

where biomass is one of the renewable sources that has presented the greatest potential for use as carbon source. It is used from 10 to 14% in the world (Shen y Yoshikawa, 2013), and methanol can be obtained from biomass or industrial residues. The use of biomass in the world is lower than fossil fuels. However, within clean energy sources, biomass presents a greater advance because it includes any biological material such as wood, energy crops, agricultural residues, food waste, etc. (Valdez-Vazquez et al., 2010). Biomass can be used as a renewable source to produce various fuels and chemical products. Products as biomethanol can be produced from non-food biomass (inedible lignocellulosic biomass residues) (Adelabu et al., 2017). Currently, the development of products from second-generation biomasses is receiving more attention to avoid competition between food and fuels. In countries like Mexico, the use of food-grade biomass is not allowed since it could generate shortages. Moreover, due to population expansion, it is probable that in the future farmland will be more limited. This shows that second-generation raw materials have an advantage as they do not require additional farmland that is already available for other purposes.

The conversion of biomass into products can take place by various routes, such as thermochemical or biological processes. Biochemical processes follow microbial/enzymatic pathways, such as anaerobic digestion. In the case of thermochemical processes, these include three basic subcategories: combustion, pyrolysis, and gasification (Demirbas, 2009). It is important to know what type of biomass will enter the system, since residues with high water content are more favorable to be processed by anaerobic digestion, and biomass residues with low water content are more favored by thermochemical processes.

Decision-making surrounding the supply chain for methanol production from biomass involves different aspects to consider. All these potential decisions and their interactions can be represented in terms of superstructures. This allows decision-making based on technical judgment and established rules. Generalized Disjunctive Programming (GDP) offers an intuitive way of representing the relationship between the different process alternatives represented in the superstructure, while capturing the connection between logical clauses and algebraic logic (Mercarelli et al., 2020). In this work, the supply chain to produce methanol from renewable and waste materials available in Mexico is modelled and optimized. The model follows a GDP approach, including as elements of the supply chain the election of the type and source of raw materials, the election of the location of the facilities and the kind of technology to be used, and the election of the market to be satisfied.

## **2. Methodology**

### *2.1. Raw material selection*

For the selection of the types of biomasses that have been considered for this study, the National Atlas of Biomass (ANBIO, 2012) is used. Within this Atlas, the availability of biomass residues in Mexico for the entire country is reported. The selected residues with the highest annual availability are shown in Table 1. These values are limiting for each type of residue.

### *2.2. Processing technologies*

Within the raw materials, there are three subdivisions: lignocellulosic waste with low water content, waste with high water levels, and materials with specific treatments, such as CO<sub>2</sub>. Due to this, 4 different technologies are defined as potential conversion routes: gasification, pyrolysis, anaerobic digestion, and CO<sub>2</sub>-based synthesis. Depending on the

type of waste, one technology or another is selected according to the properties of this waste. In addition, there is the condition of selecting only one type or thermochemical technology, either gasification or pyrolysis, but never both together. Table 2 shows the yields considered for each of the processing technologies.

Table 1. Availability of biomass residues (ANBIO, 2012).

Biomass	Availability [t/y]
<i>Solid urban waste</i>	17,586,209.83
<i>Cane bagasse</i>	17,689,266.38
<i>Agave bagasse</i>	2,026,575.00
<i>Vegetable oil production waste</i>	1,494,257.00
<i>Beer production waste</i>	1,430,959.00
<i>Coffee bagasse</i>	977,261.00
<i>Sawmill waste</i>	728,846.00
<i>Pig manure</i>	55,328.32
<i>Corn bagasse</i>	442,335.00
<i>Chicken manure</i>	217,407.00
<i>Sorghum bagasse</i>	139,689.00
<i>Bovine manure</i>	125,241.40
<i>Wheat bagasse</i>	65,527.00
<i>CO<sub>2</sub></i>	755,000,000.00

Table 2. Process yields.

Process	Yield [t MeOH/ t raw material]	Source
Gasification	50.35	Yang et al. (2018)
Pyrolysis	51.20	Kasmuri et al. (2017)
Anaerobic Digestion	46.8	Patel et al. (2021)
Synthesis from CO <sub>2</sub>	50.00	Martín and Grossmann (2017)

### 2.3. Selected markets

The studies carried out by the Ministry of Economy through Data Mexico regarding the purchase and sale of methanol in Mexico were taken into consideration. It was detected that the states that require purchasing methanol from abroad due to lack of supply are Baja California, Nuevo León, Chihuahua, the state of Mexico and Mexico City, with methanol requirements in 2021 equivalent to 162,442.4 tons (Harris, et al., 2021; Secretaría de Economía, 2021). The superstructure that includes the aspects raised is shown in Figure 1. 12 types of waste, 29 states for obtaining raw materials, 32 states for

the analysis of the location of processing plants and 5 sales markets are included in the superstructure.

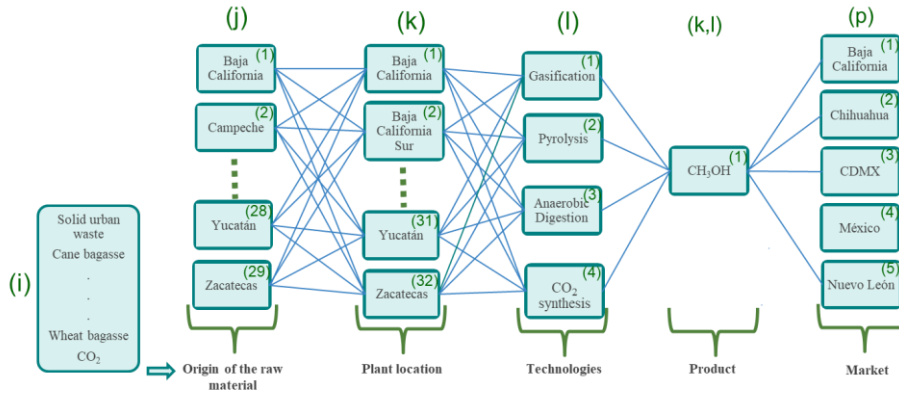


Figure 1. Superstructure for the wastes-to-methanol supply chain.

#### 2.4. Mathematical model

Nested disjunctions are used for the location of the plants and the technology used in each of these facilities, recalling that for each location there may be one or more of the four available technologies. The nested disjunction system is shown in equation 1, while equation 2 represents the objective function for a maximum profit given in USD/y.

$$\left[ \begin{array}{l}
 CBT_k = \sum_i \sum_l Y_{k,l} \cdot CB_i \cdot F_{i,j,k} \\
 CTOP_{j,k} = \sum_i \sum_l F_{i,j,k} \cdot D_{j,k} \cdot C_{comb} \\
 CTPM_{k,p} = \sum_k \sum_p F_{MeOH,k,p} \cdot D_{k,p} \cdot C_{comb} \\
 CV_{MeOH,k} = \sum_k \sum_p F_{MeOH,k} \cdot PV_{MeOH}
 \end{array} \right] \vee \left[ \begin{array}{l}
 -Y_k \\
 CBT_k = 0 \\
 CTOP_{j,k} = 0 \\
 CTPM_{k,p} = 0 \\
 CV_{MeOH,k} = 0
 \end{array} \right] \forall k \quad (1)$$

$$\left[ \begin{array}{l}
 \sum_i \sum_l F_{i,j,k} \cdot n_G = FME_{1,k} \\
 CPr_{1,k} = FME_{1,k} \cdot C_{prod1,k} \\
 COpe_{1,k} = FME_{1,k} \cdot C_{oper1,k} \\
 \sum_i \sum_l F_{i,j,k} \cdot n_p = FME_{2,k} \\
 CPr_{2,k} = FME_{2,k} \cdot C_{prod2,k} \\
 COpe_{2,k} = FME_{2,k} \cdot C_{oper2,k} \\
 \sum_i \sum_l F_{i,j,k} \cdot n_{DA} = FME_{3,k} \\
 CPr_{3,k} = FME_{3,k} \cdot C_{prod3,k} \\
 COpe_{3,k} = FME_{3,k} \cdot C_{oper3,k} \\
 \sum_i \sum_l F_{i,j,k} \cdot n_C = FME_{4,k} \\
 CPr_k = FME_{4,k} \cdot C_{prod4,k} \\
 COpe_{4,k} = FME_{4,k} \cdot C_{oper4,k}
 \end{array} \right]$$

$$Profit = \sum_k (CV_{MeOH})_k - \sum_k (C_{Pr} * F_{MeOH})_k - \sum_k (C_{Ope} * F_{MeOH})_k - \sum_k (C_{BT})_k - \sum_k (C_{TOP,j,k}) - \sum_k (CTPM_{k,m}) \quad (2)$$

Where  $i$  represents the raw material,  $j$  is the origin of raw material,  $k$  is location of plants and  $l$  is the kind of technology.  $Y_k$ ,  $Y_{k,l}$  are logical variables for plant location and type of technology.  $CBT_k$  is the total cost of raw material and  $CB_i$  is the cost for each raw material [USD/t],  $F_{i,j,k}$  is the biomass rate sent to  $k$  [t/y],  $CTOP_{j,k}$  and  $CTPM_{k,p}$  represent the cost of transportation from origin to facility and facility to market, respectively [USD/y].  $D_{j,k}$ ,  $D_{k,p}$  are transport distances [km],  $C_{comb}$  is transport cost [USD/km · t]  $CV_{MeOH,k}$  are the total sales of methanol [USD/t],  $F_{MeOH,k}$  is produced methanol [t/y],  $PV_{MeOH}$  is unitary methanol sales price [USD/t].  $n_G$ ,  $n_p$ ,  $n_{DA}$ ,  $n_C$  are technologies yields [t MeOH/ t raw material].  $FME_{1,k}$  is the methanol produced by the technology  $l$  [t MeOH/ t raw material],  $CPr_{1,k}$  represent the cost of production for each technology and  $COpe_{1,k}$  represents the operating cost for each technology.  $C_{prod1,k}$ ,  $C_{oper1,k}$  are the unitary cost

of production and cost of operation [USD/y]. Model is reformulated into an MINLP using the convex hull method and solved in the software GAMS to maximize the profit. The purpose of the model is to satisfy a production of 162,442.4 t/y.

### 3. Results

The relaxed model using the convex hull reformulation was coded in the GAMS software, obtaining a nonlinear mixed integer programming (MINLP) model, which was solved using the BARON solver. The model has 3,509 equations and 9,641 variables. The solution results in a profit of 7,336,700.00 USD/y, this profit is obtained with the route shown in Figure 2.

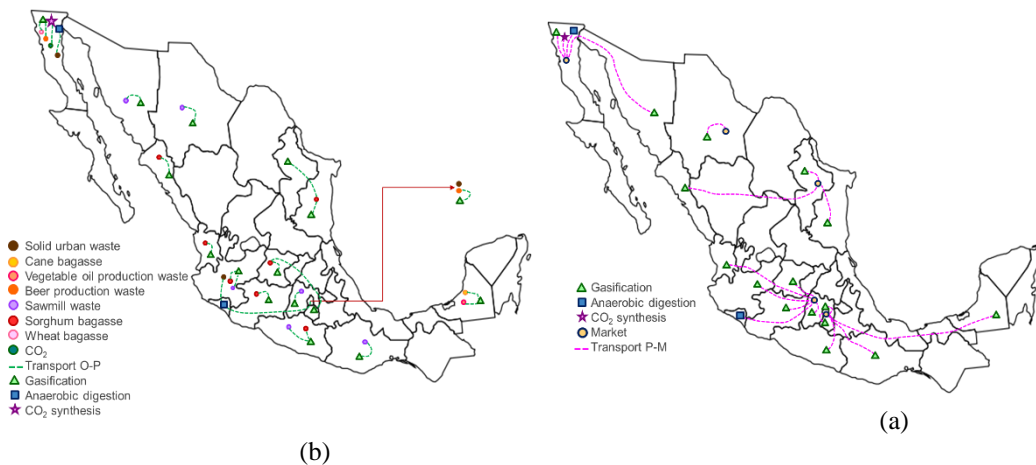


Figure 2. Transportation route, (a) from origin to facility, (b) from facility to market.

According to Figure 2.a, the solution implies the use of eight raw materials: urban solid wastes, cane bagasse, vegetable oil production waste, beer production waste, sawmill waste, sorghum waste, wheat bagasse and CO<sub>2</sub>. These raw materials are obtained from 14 states: Baja California, Sonora, Chihuahua, Sinaloa, Tamaulipas, Nayarit, Jalisco, Guanajuato, Michoacan, Guerrero, México, CDMX, Oaxaca and Campeche. The predominant biomass is sorghum bagasse and sawmill waste because its sale price is low. The facilities are placed in 17 states, which are the same from where waste is obtained, plus Nuevo León, Colima, and Morelos. It is important to mention that in Baja California there will be 3 different processing routes: gasification, anaerobic digestion, and CO<sub>2</sub>-based synthesis. This is the only state that has all 3 types of technology, because it handles 3 different types of waste. Colima is the other state that has an anaerobic digestion plant installed. The rest of the installed plants use gasification. The pyrolysis process was not selected for any state, because of the high cost of this technology, which is not compensated by its slightly higher yield. Figure 2.b shows the methanol transportation route from the production plant to the markets. It is observed that the production of the Baja California plants, and the production of Sonora are sent to the Baja California market. Chihuahua is supplied with methanol from its own facility. Nuevo León market is supplied by its own plant, receiving additional product from the plants in Tamaulipas and Sinaloa. The State of Mexico market is supplied by six different plants located in Colima, Nayarit, Jalisco, Michoacan, Guanajuato and State of Mexico. Finally, the

Mexico City market is supplied by four plants located in Guerrero, Oaxaca, Morelos, Campeche and Mexico City. This distribution shows that methanol plants in the north supply the northern markets and the plants in the center and south of the country satisfy the markets in that area. The installed plants vary according to the location in terms of capacity to produce methanol, the minimum is 5.88 kg/h and the maximum is 1,910.85 kg/h.

#### 4. Conclusions

In this work, a mathematical model to represent the methanol supply chain in Mexico has been developed. The supply chain has been optimized through the software GAMS, obtaining the maximum-profit route from the origin of the raw materials to their delivery to the market. The solution allows a profit of 7.34 MUSD/y, making use of 8 raw materials from 14 different states of obtaining, with 19 facilities installed in 17 different states, supplying 5 different markets. It has been possible to observe the versatility of the route, which transforms several raw materials using 3 different processing technologies, demonstrating that, regardless of the type of waste to be treated, there is a processing option available for its use. Finally, it is important to mention that this model was constrained to a maximum methanol production, which is satisfied, but there are still available wastes in the country. Thus, the potential for exporting renewable methanol could be explored in future works.

#### References

- Atlas Nacional de Biomasa (ANBIO), 2012, <https://www.gob.mx/sener/articulos/atlas-nacional-de-biomasa>
- B. A. Adelabu, S. O. Kareem, F. Oluwafemi, I.A. Adeogun, 2017, Bioconversion of corn straw to ethanol by celluotyc yeasts immobilized in *Mucuna urens* matrix, *Journal of King Saud University-Science*, 1-6.
- M. Alvarado, 2016, The changing face of the global methanol industry, *HIS Chemical*, 10-11.
- A. Demirbas, 2009, Biofuels securing the planet's future energy needs, *Energy Conversion and Management*, 50, 2239-2249.
- K. Harris, R. G. Grim, Z. Huang, L. Tao, 2021, A comparative techno-economic analysis of renewable methanol synthesis from biomass and CO<sub>2</sub>: Opportunities and barriers to commercialization, *Applied Energy*, 303, 117637.
- S. Maddren, 2019, Asia methanol – China and the rest, 22<sup>nd</sup> IMPCA Asian Methanol Conference, Singapore, November 5-7.
- M. Martín, I.E. Grossmann, 2017, Towards zero CO<sub>2</sub> emissions in the production of methanol from switchgrass. CO<sub>2</sub> to methanol, *Computers and Chemical Engineering*, 105, 308-316.
- L. Mercarelli, Q. Chen, A. Pagot, I. Grossmann, 2020, A review on superstructure optimization approaches in process system engineering, *Computers and Chemical Engineering*, 1-15.
- S.K.S. Patel, R.K. Gupta, V.C. Kalia, J.K. Lee, 2021, Integrating anaerobic digestion of potato peels to methanol production by methanotrophs immobilized on banana leaves, *Bioresources Technology*, 323, 124550.
- Secretaría de Economía, 2021, Metanol: alcohol metílico, Data Mexico, <https://datamexico.org>
- Y. Shen, K. Yoshikawa, 2013. Recent progress in catalytic tar elimination during biomass gasification or pyrolysis- A review, *Renewable and Sustainable Energy Reviews*, 21, 371-392.
- I. Valdez- Vazquez, I., J. A. Acevedo-Benitez, & C. Hernández-Santiago, 2010, Distribution and potential of bioenergy resources from agricultural activities in México, *Renewable and Sustainable Energy Reviews*, 14, 2147-2153.
- S. Yang, B. Li, J. Zheng, R.K. Kankala, 2018, Biomass-to-Methanol by dual-stage entrained flow gasification: Design and techno-economic analysis based on system modeling, *Journal of Cleaner Production*, 205, 364-374.

# Supply chain optimal planning for revalorization of empty plastic pesticide and fertilizer containers

Sergio Iván Martínez-Guido<sup>a</sup>, Rubén Bernabé-Martínez<sup>b</sup>, Claudia Gutiérrez-Antonio<sup>a\*</sup>, Salvador Hernández<sup>c</sup>

<sup>a</sup>*Universidad Autónoma de Querétaro, El Marqués, Querétaro, 76265, Mexico*

<sup>b</sup>*PRANA Proyectos de Reciclaje y Acciones por la Naturaleza, Tuxtla Gutiérrez-Chiapas, 29000, Mexico*

<sup>c</sup>*Universidad de Guanajuato, Guanajuato, 36050, Mexico*  
*claudia.gutierrez@uaq.mx*

## Abstract

Particularly, in agro-industry most of the plastic products are single-use and can persist in the environment long after their intended use, threatening food security and safety (FAO, 2021). Then, in this work is proposed the optimal planning of the supply chain for the revalorization of empty plastic pesticide and fertilizer containers (bottles) generated during the production period. In this way, a mathematical model was proposed to represent all the supply chain steps, and codified in GAMS<sup>®</sup>. As a case of study, three states in Mexico were selected, taking into consideration some of the higher agricultural sites. Main results show that is possible to obtain 715 tonnes of pellets as products from pesticide and fertilizer containers. Moreover, it is possible to forecast an economic benefit of USD\$164,480 per year, and an annual reduction of 1,136 tonnes of CO<sub>2</sub>, in comparison with the current situation.

**Keywords:** waste management, circular economy, supply chain, optimization, mathematical model.

## 1. Introduction

According to Pohjakallio (2020), if we continue producing plastics as we do now, the plastic industry would consume 20% of the total oil production by 2050. Additionally, whether the behavior in the plastics demand continues, the global plastic waste volumes would grow 77% by 2030 in comparison with the 260 million tonnes per year in 2016. Moreover, the constant use of plastic is increasing the number of their pollutants in the environment, and today it's possible to find these in our food chain and even in our blood, placing them as a threat to human health (Moshood *et al.*, 2022).

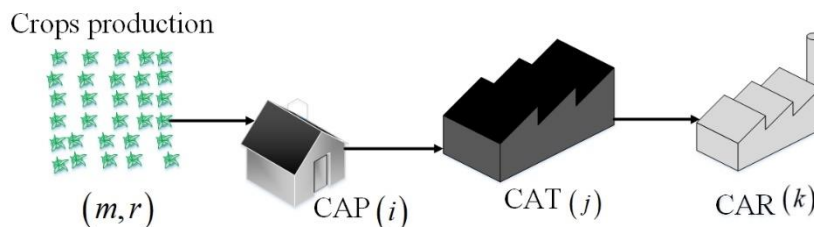
On the other hand, in agriculture, plastic products are used in all aspects as the livestock production (feed and animal care), plant production, fisheries, and aquaculture. Plastic also are used in distribution and retail activities, with the goal to keep the agricultural products quality. Furthermore, some of the benefits from using these products in crops production are the reduction water demand, optimizing seed germination, increasing crop yields, crop protection, among others. However, most of the agricultural plastic products are single-use, and these can persist in the environment long after their intended use, threatening food security, food safety and potentially human health (FAO, 2021). In the face of this bleak outlook, bio-based plastics are already in the market as carbon neutral solutions. Notwithstanding, these renewable solutions have been mostly applied into packaging and agriculture activities in less of 2% of the total production plastics; this due



the high demand of biomass to other uses, and the lower cost of fossil sources (Patrício-Silva, 2021). Hence, in this work is proposed the optimal planning of the supply chain for the revalorization of empty plastic pesticide and fertilizer containers (bottles) generated during the crops production. In this way, a mathematic model was proposed to represent all the steps involved in the supply chain, and codified in GAMS®. As a case of study, three states in Mexico were selected: Jalisco, Guanajuato and Colima, considering some of its municipalities with higher agricultural production. To obtain the supply chain optimal planning warehouses currently installed in each state were included, as well as the seasonality of crops, distances between the production sites, warehouses, and revalorization sites. Main results show that is possible to obtain 715 tonnes of pellets and flakes as products from pesticide and fertilizer containers. Moreover, it is possible to forecast an economic benefit of USD\$164,480 per year, and an annual reduction of 1,136 tonnes of CO<sub>2</sub>, in comparison with the current situation in which these containers are left in the fields and burned. In this way, all the strategies with emissions reduction goals can be improved with the integration of optimization tools, even when part of the chain links already exist, or are installed. Additionally, all the information obtained from these optimal solutions represent a starting point for the promotion of public policies as well as a watershed for decision makers at all levels.

## 2. Problem statement

As a first approach to a regional analysis, 10 municipalities from Jalisco and Guanajuato states and eight from Colima state were taken into consideration as case study, due their highest respective amount of agricultural land. As strategy, it is proposed that in each selected municipality a primary storage center (CAP) will be installed, providing an easily identified by local farmers and does not need to be moved to another producing municipality. Subsequently, there are temporary storage centers (CAT) where more detailed handling of the packaging can be given (See figure 1). As installation sites for this kind of centers are considered in first stay existing, located in Tamazula de Gordiano (Jalisco), Juventino Rosas (Guanajuato), and Tecomán (Colima). In addition, it is possible, if it isn't enough with these three current CAT, to install some others in any of the municipalities identified as producers. For this, disjunctive equations are proposed that allow the existence decisions, which as a result will activate fixed and variable costs of each installed center. Finally, a specialized handling center (CAR) is proposed for all packaging with the capacity to generate value-added products, which has the possibility of being installed in any of the 28 selected municipalities.



**Figure 1.** Proposed supply chain

### 3. Methodology

In the present work is proposed the optimal supply chain planning determination of empty agricultural containers, through the use of mathematical tools developed by the Operations Research area (Goldbeck et al., 2020). In this sense, as a case study, the current generation of residual containers of agricultural production of different selected crops in the states of Jalisco, Guanajuato and Colima was considered. As a starting point, a review was made of the current real situation, useful information for the choice of the 10 most representative crops, in each of the 28 selected municipalities.

**Table 1.** Selected crops.

<b>Jalisco</b>	Corn, fodder corn, sugarcane, agave, sorghum, avocado, wheat, oats, bean and fodder sorghum
<b>Guanajuato</b>	Corn, sorghum, wheat, bean, barley, alfalfa, broccoli, oats, agave, onion.
<b>Colima</b>	Lemon, copra, sugarcane, corn, banana, mango, papaya, rice, coffee, onion.

Subsequently, the formulation of a mathematical model was carried out, representing the matter and energy balances of the real problem. In this model were included all the activities involved in the supply chain, and described in detail in the section "Mathematical model". To feed the model it was necessary to collect real parameters of each selected municipality including crop kind, agricultural needs, land used for cultivation, distances, costs, CO<sub>2</sub> emissions (using the GREET software), among others. The model and other parameters were codified in the GAMS® platform, where it was resolved by integrating economic performance maximization and environmental emissions minimization. The solutions generated were classified into four main points, solution A (optimal economic performance), solution D (optimal emission reduction), and solutions B and C which are compromise solutions between both objectives.

#### 3.1. Mathematical model

As was described before, the proposed approach takes into consideration all the steps involved to achieve the optimal planning of the supply chain for the revalorization of empty plastic pesticide and fertilizer containers (bottles), which are generated during the crops production. Equation 1 represents the balance to estimate the number of produced containers in each municipality per week ( $FB_{j,i,t}$ ), which depends from the amount of area used to produce each kind of crop ( $AUSE_{j,r,t}$ ) as well as the pack type factor linked to each crop ( $FGEN_{r,i,t}$ ). Afterwards, the number of empty containers is transformed into mass flow ( $FMB_{j,t}$ ), multiplying the particular weight of each container ( $DB_i$ ) by the generated quantity from these ( $FBC_{j,i,t}$ ), as equation 2 describes. Total containers mass flow ( $FMB_{j,t}$ ) have two options the revalorization into pellets pathway ( $FBE_{j,m,t}^{CAP}$ ) or to be sent it to a final disposition site, where normally are burned it ( $FBE_{j,t}^{QUE}$ ). Hence, in equation 3 is shown this mass balance. If the plastic flow take the revalorization pathway, hence, equation 4 represents the balance in each CAP, giving the total current flow ( $Val_{m,t}^{RealCAP}$ ), from the subtraction of the flow that it sends to the CAT ( $ED_{m,l,t}^{CAT}$ ) from the total received ( $FBE_{j,m,t}^{CAP}$ ).

$$FBC_{j,i,t} = \sum_r AUSE_{j,r,t} \cdot FGEN_{r,i,t}, \quad \forall j \in J, i \in I, t \in T \quad (1)$$

$$FMB_{j,t} = \sum_i FBC_{j,i,t} \cdot DB_i, \forall j \in J, t \in T \quad (2)$$

$$FMB_{j,t} = \sum_m FBE_{j,m,t}^{CAP} + FBEQUE_{j,t}, \forall j \in J, t \in T \quad (3)$$

$$Val_{m,t}^{RealCAP} = \sum_j FBE_{j,m,t}^{CAP} - \sum_l ED_{m,l,t}^{CAT}, \forall m \in M, t \in T \quad (4)$$

Afterwards, in equation 5 is shown the disjunction for the CAP installation, using the binary variable decision ( $x_{m,c}^{CAP}$ ), constrained by the processing limits ( $Limmi_{m,c}^{CAP}$  and  $Limma_{m,c}^{CAP}$ ). Additionally, if the CAP is installed, then the fixed ( $CF_{m,c}^{CAP}$ ) and variables cost ( $CV_{m,c}^{CAP}$ ) are activated, giving the capital CAP installation cost ( $Cost_{m,c}^{CAP}$ ). Similarly, the installation decisions of CAT's and CAR's were obtained in the same way that CAP. Profit (*Profit*) was obtained from the earnings by the pellets sale ( $Gan_l^{Real}$ ), minus all the involved cost in the supply chain as equation 6 describes. Similarly, all the CO<sub>2</sub> emissions generated by each supply chain activity were calculated ( $Emiss^{Re}$ ), resulting in emissions (*BE*) that can be avoided if the plastic container's takes the revalorization pathway instead of the final disposition site (see equation 7). Finally, the multi-objective function (*F.O*) was the profit maximization (*MaxProfit*) and the emissions reduction (*MinBE*) as equation 8 shown.

$$\forall_{m,c} \left[ \begin{array}{l} x_{m,c}^{CAP} \\ Limmi_{m,c}^{CAP} \leq EN_{m,c}^{CAP} \leq Limma_{m,c}^{CAP} \\ Cost_{m,c}^{CAP} = CF_{m,c}^{CAP} + CV_{m,c}^{CAP} \cdot EN_{m,c}^{CAP} \end{array} \right] \quad (5)$$

$$Profit = \sum_l Gan_l^{Real} - \left( \begin{array}{l} Cost^{Quema} + \sum_{m,c} Cost_{m,c}^{CAP} + Cost^{Tr-CAP} + \sum_l Cost_l^{CAr} \\ + Cost^{Tr-CAr} + \sum_p Cost_p^{CAr} + Cost^{Tr-CAr} \end{array} \right) \quad (6)$$

$$BE = Emiss^{Que} - Emiss^{Re} \quad (7)$$

$$F.O = Max Profit; Min BE \quad (8)$$

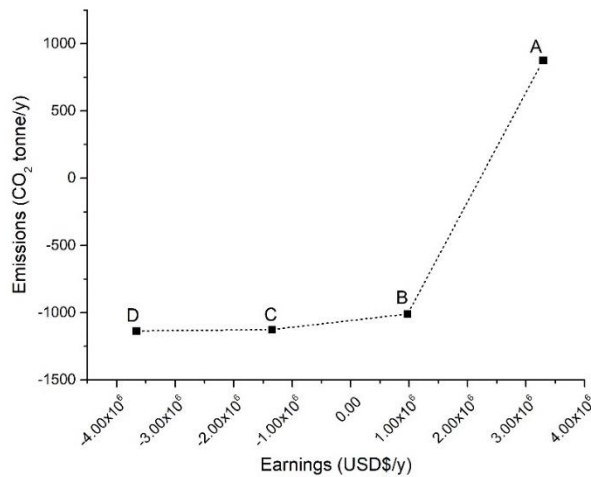
#### 4. Results

As was described in methodology section, the model and other parameters were codified in the GAMS<sup>®</sup> platform, where it was resolved by integrating economic performance maximization and environmental emissions minimization using the  $\epsilon$  constrain method (Diwekar, 2010). The model is a Mixed Linear Integer model, since there are no equations with division of variable by variable, multiplication of these, or some other algebraic functions. The solutions were obtained using the CPLEX solver, with an analysis time of 0.844 seconds per test, calculating 141,231 variables, 14,680 equations, 142 binary variables and 451,389 elements. The CPU execution time per assay was 2.360 seconds. In order to obtain the solution for each scenario the CPLEX solver was used in an AMD<sup>®</sup> A10 processor running at 2.60 GHz with 8GB of RAM memory. Obtained solutions were classified in four different scenarios. The scenario A which represent the optimal economic performance, while the scenario D represents the optimal emission reduction. Scenarios B and C which show compromise solutions between both objectives, as is shown by Figure 2. It is possible to notice in Table 1 that solutions A and B are the only where the supply chain optimal planning is perceived to be profitable, due to the existence

of earnings, while solution C and D show negative numbers. However, the results are given taken into consideration one year of operation; hence, all the installation cost are paid in the same year. Therefore, solutions C and D doesn't have any earnings, due that in these solutions all the CAT's installed are new; this means that the existent installation are not considered as strategic points for emission reduction. This could be avoided by considering the acquisition of a loan or sub-loan over a period of more than one year, perhaps considering a life period between 10 and 15 years.

**Table 1.** Main results obtained in each scenario.

Scenario	A	B	C	D
Earnings (USD\$/y)	164,480	48,620	-67,240	-183,100
Emissions (Ton CO <sub>2</sub> /y)	877	-1,010	-1,126	-1,136



**Figure 2.** Pareto solution diagram

Results show for all the scenarios analyzed, the best solutions consider the existence of at least one CAP (Primary collection center) in each of the municipalities (28 overall). These have functions such as material storage as well as basic sorting, size reduction by crushing, as well as sending material to CAT (Revalorization center). The CAR's (Packaging and recycling center) installation are not taken into consideration, due the respective capacities; hence, in the installed CAT are carry out the rest of the revalorization activities. A final size reduction and pellets extrusion. For the CAT where considered 3 different capacities 0.76, 1.52, 2.3 tonnes/d; for all the solutions was selected by the model the lowest capacity. In scenario A, exchanges are allowed only between municipalities in the same state for savings in transport costs. While in scenario D only one municipality exchanges to three others, the rest to only two; this is due that transportation activities are constrained by emissions. In solution A, Romita and Dolores do not use CAT, their material is burned in filling, even when it was considered an economic penalty for carrying out this activity. Moreover, in scenario D burning is not

allowed, due that in this scenario the emissions reduction tamed the model. Furthermore, in solution A all the existent installed CAT are taken into consideration as part of the supply chain, additionally to the installation of two new CAT's, having the capacity to process 665 tonnes of plastic containers per year. While in solution B are installed three new CAT's plus to the existing with a process capacity of 715 tonnes of plastic containers per year. In solution C only two of the existing CAT's are part of the supply chain planning, and four new are installed, with a total processing capacity of 715.5 tonnes/y. Lastly, as was described before in solution D any of the existing CAT's are including in the planning the six are new, with a total capacity to process 715.7 tonnes of empty plastic containers.

## 5. Conclusion

This kind of analysis presents an overview of the optimal configuration based on the data provided; however, considering any relevant change does not represent complexity for its study. Results shown that this kind of revalorization contributes with the emissions reduction and with an economic profit. Particularly, in scenarios with higher emissions reduction it is possible to obtain an economic benefit; however, this earning is reflected up to year two of operation, due that the investment required for new links in the supply chain is paid in first year. In this way, it was possible to notice that installation costs are the most relevant points in the economic performance; therefore, for the implementation of these kinds of strategies a network of decision makers is required, as well as the governmental and private financing. Additionally, government incentives or subsidies in loan scheme, which can cushion the higher installation costs this in order to observe the behavior of the configuration. Current links installed (mainly storage sites) turn out to be unattractive optimal configurations since the environmental point of view, due that these are located so far from the main distribution ways. Hence, it is possible to conclude that each supply chain design and planning has to be analyzed using optimal configuration methods such as the proposed tool.

## 6. Acknowledgments.

Financial support provided by CONACYT, through grant 320583, is acknowledged.

## References

- M. Pohjakallio, 2020, Secondary plastic products-examples and market trends, *Plastic Waste and Recycling*, 1, 467-479. 10.1016/B978-0-12-817880-5.00018-9
- Patrício-Silva, A.L. Future proofing plastic waste management for a circular bioeconomy, *Current Opinion in Environmental Science & Health*, 22,100263, 10.1016/j.coesh.2021.100263
- T.D. Moshood, 2022, Sustainability of biodegradable plastics: New problem or solution to solve the global plastic pollution?, *Current Research in Green and Sustainable Chemistry*, 5, 100273, 10.1016/j.crgsc.2022.100273
- “FAO” Food and Agriculture Organization of the United Nations, 2021, Assessment of agricultural plastics and their sustainability: a call for actions, available on: <https://www.fao.org/policy-support/tools-and-publications/resources-details/es/c/1460015/>
- U. Diwekar, 2010, *Introduction to Applied Optimization*, Second ed, Springer, 10.1007/978-0-387-76635-5
- GREET® “The Greenhouse gases, Regulated Emissions, and Energy use in Technologies Model”, 2020, Argonne National Laboratory, Energy Systems, available: <https://greet.es.anl.gov/>
- N. Goldbeck, P. Angeloudis, W. Ochieng, 2020, Optimal supply chain resilience with consideration of failure propagation and repair logistics, *Transport Res E-Log*, 133, 101830. 10.1016/j.tre.2019.101830

# Developing a Comprehensive Decision Support Optimization Model for Biofuel Supply Chain

Brook Tesfamichael,<sup>a</sup> Ludovic Montastruc,<sup>b,c</sup> Stéphane Negny,<sup>b,c</sup>

<sup>a</sup>*Addis Ababa University, Addis Ababa Institute of Technology, School of Chemical and Bio Engineering, Addis Ababa, Ethiopia*

<sup>b</sup>*Université de Toulouse, INP-ENSIACET, LGC (Laboratoire de Génie Chimique), 4, allée Emile Monso, F-31432 Toulouse Cedex 04, France*

<sup>c</sup>*CNRS, LGC (Laboratoire de Génie Chimique), F-31432 Toulouse Cedex 04, France*

## Abstract

Although several optimization-based decision support mathematical models are developed so far, they have limitations to be directly adapted and applied to real country-level cases. Therefore, the objective of this study is to develop a comprehensive and realistic mathematical optimization-based decision support model for the strategic design and tactical planning of a country's biofuel supply chain. First, a bi-objective economic objective optimization model, with objectives of profit maximization and investment cost minimization, is formulated. The model is devised to provide both the strategic and tactical decisions simultaneously. Then, the previous model is extended to a multi-objective optimization model by adding an environmental objective that aims to minimize the ecocost of the biofuel supply chain after conducting a comprehensive Life Cycle Assessment (LCA). The developed model in this study is applied to a case study in Ethiopia to support the policymakers and other stakeholders engaged in the country's biomass-to-biofuel projects.

**Keywords:** biofuel, supply chain, multi-objective optimization, economic-environment.

## 1. Introduction

Biofuel derived from biomass, on account of being renewable energy source and having high potential to substitute fossil fuels, have attracted considerable interest in both developed and developing countries. Despite there is a strong interest in biofuels, there are several challenges along the supply chain that hinder biofuel development and utilization progresses. The many challenges and uncertainties related to dynamic criteria need more quantitative research that supports policymakers to make evidence-based decisions along the biofuel supply chain.

Several mathematical models were developed to support decision-making in different supply chains. Among the mathematical models, optimization techniques have been employed in most of the studies to manage the biofuel supply chain from strategic, tactical and operational perspectives (Barbosa-Póvoa et al., 2018). The models coupled with an optimization method are effective when different parts of the supply chain (procurement, preprocessing, production, distribution, and transportation), different sustainability dimensions (economic, environmental, social, technological, and political) and different decision levels (strategic, tactical and operational) are integrated.

However, the previous studies have limitations to directly adapt and apply to real country-level cases.

Therefore, this research aims to develop a comprehensive and realistic mathematical optimization-based decision support model, for the strategic design and tactical planning of country's biofuel supply chain over a 20-year horizon. The economic-environment optimization model developed in this study broadens the preceding research endeavors by incorporating the following issues in the model simultaneously:

- A comprehensive biofuel supply chain, which includes all the principal supply chain components - upstream and downstream of the biorefinery along with a decentralized preprocessing unit is considered.
- A multi-product biofuel supply chain, with the intention of producing bioethanol and biodiesel along with the associated co-products, is considered in the model.
- The development of an integrated strategic-tactical level model is considered to prevent non-optimal and infeasible results, which may arise from hierarchical models.
- Several impact categories are considered to assess the impact of the biofuel supply chain on the environment. Moreover, a monetary-based environmental impact indicator is considered to compare with the economic objective clearly.
- A new approach that analyzes the compromise between objectives of large size model within a justifiable computational time and efficient optimization resolution is considered.
- Decisions associated with location, technology, and capacity is considered unlike to many researches, which consider either of these decisions exclusively or combination of two of them only, especially location and capacity.
- The investment decisions associated to location, technology, and capacity of the facilities are appeared on yearly basis contrarily to many studies that consider this decision only at the beginning of the planning period.

The model is applied to the real case of Ethiopia to support the policymakers and other stakeholders engaged in the country's biofuel projects.

## 2. Methodologies

### 2.1. General structure of the biofuel supply chain

This study considers all the principal biofuel supply chain components - upstream and downstream of the biorefinery, which includes biomass purchasing center, biomass preprocessing unit, biorefinery, biofuel distribution center, and demand zone. The superstructure representing the consecutive activities of the supply chain is illustrated in Figure 1 and the problem is described as follows. Various types of biomass feedstocks  $b$ , which are purchased from different supply sources  $h$ , are considered. For the sake of pretreatment, these biomasses are transported to a preprocessing unit at location  $i$  with capacity  $c$  and technology  $f$ . Then, the preprocessed biomasses  $d$  are conveyed to a biorefinery located at  $j$  with capacity  $e$  and technology  $g$  for the production of the respective biofuel  $u$  and co-product  $v$ . Then, the biofuel  $u$  goes to the final demand center  $m$  directly or it is shipped to a distribution center located at  $k$  with capacity  $a$  and then to a demand zone  $n$ . Meanwhile, the co-product  $v$  directly goes from the biorefinery  $j$  to the respective demand zone  $n$ .

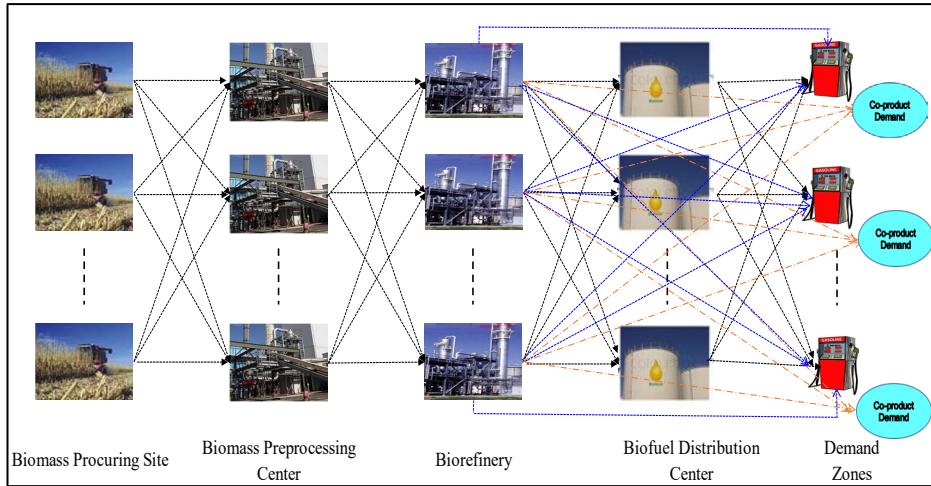


Figure 1: Biofuel supply chain structure

### 2.2. Model formulation

A deterministic model is selected since the solution of this model gives decision makers excellent insights for making excellent choices. To address the whole behavior of the supply chain, the problem is designed as a spatially explicit, multi-product, multi-feedstock, multi-period, and multi-echelon MILP modeling framework.

First, a strategic-level mono-objective optimization model, which intends NPV maximization of the biofuel supply, is developed. Here, the model considers a yearly time period. The objective function, as illustrated in Equation (1), comprises of revenue, total cost and initial investment.

$$NPV = \sum_{y \in Y} \left[ \begin{array}{l} \left[ \frac{1}{(1+ir)^y} (Revenue_y - TotalCost_y) \right] \\ - \left[ \frac{1}{(1+ir)^{y-1}} (InvestmentCost_y) \right] \end{array} \right] \quad (1)$$

Where  $y$  is the year and  $ir$  is the discount rate

However, the supply chain design recommended by the strategic optimization model has resulted in infeasible solutions at tactical level, which makes this hierarchical approach not reliable to apply for long-term planning. Henceforth, second, an integrated strategic-tactical model is developed in order to prevent non-optimal and infeasible results, which arise from the conventional hierarchical approach. Nevertheless, no solution is acquired as the numerical solver-CPLEX was run out of memory. This is attributed to the sharp increment of the model size due to (a) requirement of vast data as it is a real and comprehensive national-level case study for a long-term planning of a sector and (b) existence of many binary variables due to numerous annual strategic choices related to the location, technology, and capacity of the supply chain components. Therefore, a new methodology is proposed to convert the mono-objective (NPV maximization) to bi-objective optimization model, which aims to maximize the profit and minimize the investment cost of the biofuel supply chain as illustrated in Equation (2).



$$\begin{aligned}
\text{Profit} &= \sum_{y \in Y} \frac{1}{(1+ir)^y} \left[ \text{Revenue}_y - \left( \text{BiomassPurchase}_y + \text{ProductionCost}_y \right. \right. \\
&\quad \left. \left. + \text{InventoryCost}_y + \text{TransportCost}_y \right) \right] \\
\text{InvestmentCost} &= \sum_{y \in Y} \frac{1}{(1+ir)^{y-1}} \left( \sum_{i \in I} \sum_{f \in F} \sum_{c \in C} (\text{Inv1}_{f,c} * (S_{i,f,c,y} - S_{i,f,c,y-1})) \right) + \\
&\quad \left( \sum_{j \in J} \sum_{g \in G} \sum_{e \in E} (\text{Inv2}_{g,e} * (R_{j,g,e,y} - R_{j,g,e,y-1})) \right) + \\
&\quad \left( \sum_{k \in K} \sum_{a \in A} (\text{Inv3}_a * (Q_{k,a,y} - Q_{k,a,y-1})) \right)
\end{aligned} \tag{2}$$

Where  $\text{Inv1}_{f,c}$ ,  $\text{Inv2}_{g,e}$  and  $\text{Inv3}_a$  are capital investments to install the preprocessing facility with technology  $f$  and capacity  $c$ , the biorefinery with technology  $g$  and capacity  $e$ , and the distribution center with capacity  $a$  respectively.  $S_{i,f,c,y}$ ,  $R_{j,g,e,y}$  and  $Q_{k,a,y}$  are binary variables with a value of 1 if preprocessing facility is installed in location  $i$  with technology  $f$  and capacity  $c$  at year  $y$ , if biorefinery is installed in location  $j$  with technology  $g$  and capacity  $e$  at year  $y$  and if distribution center is installed in location  $k$  with capacity  $a$  at year  $y$  respectively; otherwise 0.

Third, taking the previous bi-objective economic optimization model as the base model, an environmental objective function is added to it, which results a multi-objective MILP model. The environmental objective is to minimize total ecocost, which comprises the biomass processing, biofuel production and material transportation Ecocosts as indicated in Equation 3.

$$\text{TotalEcocost} = \sum_{y \in Y} \left( \begin{aligned} &\text{ProcessingEcocost}_y \\ &+ \text{ProductionEcocost}_y \\ &+ \text{TransportationEcocost}_y \end{aligned} \right) \tag{3}$$

Therefore, the mathematical model is comprised of two economic (profit maximization and investment cost minimization) and one environmental (ecocost minimization) objective functions. The design process is subjected to logical constraints and mass balances that must be satisfied at each supply chain nodes.

### 2.3. Case study economic, technical and environmental data

The developed model in this study will be applied to a case study in Ethiopia. The 58 zones in the country are considered as the candidate locations for the biomass preprocessing, biorefinery, and biofuel distribution centers. The prior feedstocks for biofuel development in Ethiopia are considered in this study; sugar cane molasses, bagasse and corn stover for bioethanol and jatropha and castor bean for biodiesel production. The annual production of the biomasses are estimated based on the current data obtained from different Ethiopian Ministry of Agriculture, Central Statistical Agency of Ethiopia, Ethiopian Sugar corporation, and biomass developers. The preprocessing mechanism considered for the two bioethanol feedstocks, i.e. bagasse and corn stover, is drying and size reduction. However, no preprocessing is considered for the molasses. On the other hand, technologies considered for biodiesel preprocessing are mechanical and solvent extraction. Biochemical and thermochemical conversion processes for bioethanol

production and homogenous and heterogeneous catalysis transesterification for biodiesel production are considered as biorefinery technologies. Six potential capacity levels are selected for the preprocessing, biorefinery and distribution facilities. Due to the lack of real preprocessing and biorefinery investment and production costs data from commercial plant within Ethiopia, these costs are estimated with bibliographic data, assuming that the components follow the economy of scale, and then corrected with the available information and recent data. The biofuels' (bioethanol and biodiesel) demand are forecasted by considering the government plan. Based on the case study area features, only road transportation mode is assumed for transporting different materials within the supply chain (Tesfamichael et al., 2021).

LCA method is applied for evaluating the environmental impact. As indicated in Figure 1, the scope of this work encompasses the upstream stage (biomass feedstocks procurement, preprocessing, storage, and transportation), midstream stage (preprocessed biomasses conversion to biofuels and co-product), and downstream stage (biofuels distribution and transportation to final demand zones). The life cycle analysis of biomass feedstocks production and biofuels use are excluded because these activities are outside of the system boundary of this work. Furthermore, the life cycle analysis of material storage in all stages of the supply chain is excluded since there is insignificant emission and energy requirement associated with storage of materials. The LCA is carried out using primary data, information and data from prior studies, and LCA software - GREET LCA 2016 (a freely available tool from Argonne National Laboratory). The GREET software package is used to generate emission factors for the various biomass preprocessing and biofuel production (biorefinery) technologies by considering each type of biomass and preprocessed biomass. Moreover, the emission associated to transportation of materials is determined using prior vehicular emission study in Ethiopia. Finally, the Ecocost method is used to quantify the impacts of all the biofuel supply chain activities on the environment.

### **3. Result and Discussion**

Since this planning problem relies on three objective functions, it requires a solution of multi-objective MILP. The three mono-objective optimization problems were considered independently before dealing with the multi-objective problem. The mono-objective optimization models are solved using CPLEX 12.8 algorithm embedded in IBM ILOG. Then, the intention was to make a payoff table and acquire balanced solution that helps to look for the best compromise among the three antagonistic issues. Nevertheless, the attempt to solve one of the objective functions (investment cost minimization) could not be successful as the numerical solver run out of memory due to the many binary variables associated with several strategic choices as well as vast data of national level. Therefore, a semi-heuristic approach was gradually developed through the following procedure. First, the two objective functions, namely profit maximization and ecocost minimization, were simulated with respect to different investment cost ranges. The results showed that there is a significant variation on the profit by varying the investment cost. Contrary, the ecocost variation is almost insignificant. Knowing that the profit is more sensitive parameter, second, the profit maximization was simulated with respect to different investment cost and ecocost ranges. Here, the profit maximization was run by taking four intervals for the ecocost and twenty-two intervals for the investment cost that generates a total of 88 feasible solutions that offers a good compromise between the three objectives. The model for this case had 1,348,772

constraints and 2,578,430 decision variables, of which 174,000 and 2,404,430 were binary and continuous variables respectively. The results were obtained in the ranges of 157.21 to 668.58 minutes with 0.01% optimality gap on an Intel 2.60-GHz processor. The simulation result is depicted in Figure 2.

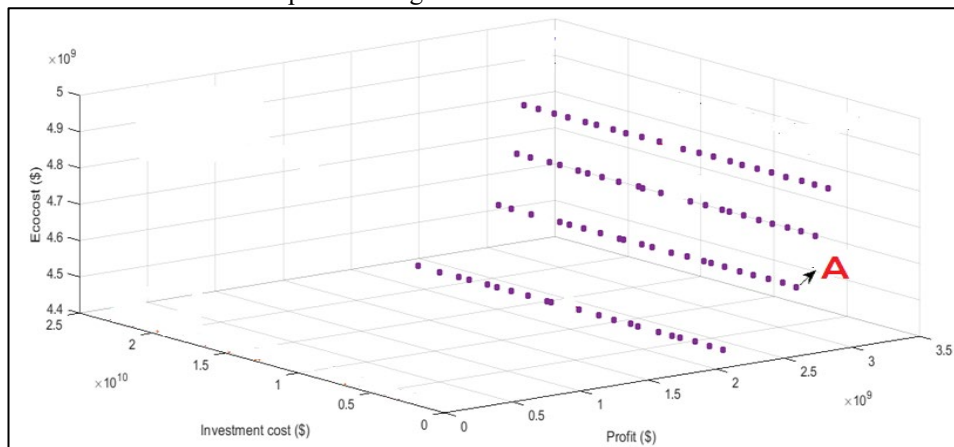


Figure 2: Feasible solutions to the multi-objective problems

The most feasible solution that compromises the three objectives is indicated as point A in Figure 2. This solution has exclusively proposed the installation of biochemical and heterogeneous transesterification refineries for bioethanol and biodiesel production respectively throughout the planning period. Most of the refineries are located in areas proximate to the biomass supply locations rather than biofuel demand locations. Moreover, this solution favors the utilization of molasses in the first 10 planning years, which is then joined by both corn stover and bagasse to satisfy the bioethanol production in the next 11-20 years. Besides, it recommends the utilization of both jatropha and castor for biodiesel production in the planning period.

#### 4. Conclusion

In this work, a comprehensive and realistic mathematical optimization-based decision/planning support model is developed by considering various issues simultaneously. The problem relies on three objective functions, namely, profit, investment cost and ecocost. A semi-heuristic strategy, which offers a feasible solution to the problem, is explored to analyze the compromise between the objectives. A first potential future expansion of this work is to broaden the supply chain boundary by incorporating biomass cultivation activity. A second extension could be to incorporate the social and political dimensions. A third expansion is to consolidate operational level planning in the integrated strategic-tactical level model developed here.

#### References

- Barbosa-Póvoa, A.P., da Silva, C., Carvalho, A., 2018. Opportunities and challenges in sustainable supply chain: An operations research perspective. *Eur. J. Oper. Res.* 268, 399–431. <https://doi.org/10.1016/j.ejor.2017.10.036>
- Tesfamichael, B., Montastruc, L., Negny, S., Yimam, A., 2021. Designing and Planning of Ethiopia's Biomass-to-Biofuel Supply Chain through Integrated Strategic-Tactical Optimization Model considering Economic Dimension. *Comput. Chem. Eng.* <https://doi.org/10.1016/j.compchemeng.2021.107425>

# Investigation of economic and environmental tradeoffs of spatially explicit biomass supply chains towards the production of net-negative biofuels

Eric G. O'Neill<sup>a,b</sup>, Caleb H. Geissler<sup>a,b</sup>, Christos T. Maravelias<sup>a,b,c</sup>

<sup>a</sup>*Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08544, USA*

<sup>b</sup>*DOE Great Lakes Bioenergy Research Center, USA*

<sup>c</sup>*Andlinger Center for Energy and the Environment, Princeton University, Princeton, NJ, 08544, USA*

*maravelias@princeton.edu*

## Abstract

Cellulosic biomass is an attractive source of renewable fuel because of its high greenhouse gas (GHG) mitigation potential. However, the optimal fuel conversion technology, amount of carbon capture and storage (CCS), and supply chain (SC) design depend on spatial features of the system which are essential to the systemwide GHG mitigation potential. We analyze the cost and GHG mitigation of a cellulosic biofuel SC with CCS using mixed-integer linear programming. While previous studies examine small high-resolution regions or large coarsely-represented regions, we consider a high-resolution SC for an 8-state region in the USA Midwest using, importantly, realistic biomass data. We show how the amount of biofuel produced and the level of carbon incentive contribute to substantial changes in the optimal SC configuration, biofuel conversion, and CCS technologies installed. While significant GHG mitigation is possible, sequestration credits may neglect to incentivize the further mitigation attainable by considering all sources of emissions.

**Keywords:** Supply chain, carbon capture and storage, biofuels

## 1. Introduction

For cellulosic biofuels to become economically and environmentally feasible, an efficient field-to-product supply chain (SC) is necessary. The field-to-product system consists of biomass growth and harvesting, the SC logistics network, and the biorefinery design and operation. Potential for biomass growth is highly distributed across the landscape, and careful selection of which fields to plant and the level of fertilization to apply to each field can significantly impact crop yields and the greenhouse gas (GHG) balance of the crop harvesting (Field et al. 2018). Within the SC, the facility location, transportation, and inventory management are highly dependent on the landscape design decisions, and spatial tradeoffs exist between the two problems. The final part of the value chain is producing biofuels at the biorefinery. Biomass can be converted into a variety of liquid fuels using different conversion technologies. The conversion technologies have different fuel yields, capital costs, and operating costs, which makes technology selection nontrivial. Additionally, carbon capture and storage (CCS) technologies can be installed at the biorefinery to improve the GHG mitigation potential of cellulosic biofuel, and, depending on the conversion technology, there is the potential to install CCS on a variety

of process streams with varying CO<sub>2</sub> concentration and flow rates. Landscape, SC, and biorefinery design have largely been studied independently, which overlooks the interdependencies between these systems. By considering landscape, SC, and biorefinery design simultaneously at a high spatial resolution, optimization models can be used to study the tradeoffs between these systems to minimize cost and GHG emissions.

The fuel conversion and CCS technologies at the biorefinery are especially important for the design of the integrated system. The most technologically mature conversion routes for lignocellulosic biomass are fermentation, pyrolysis, and gasification. For a biorefinery using fermentation, CO<sub>2</sub> is emitted in three streams: from the fermentation process at >95 wt.%, in biogas from anaerobic digestion of wastewater at ~73 wt.%, and in flue gas from solid residue combustion at ~20 wt.% (Humbird et al., 2011). For pyrolysis, the biomass is thermochemically converted to a bio-oil that requires hydrogen to be upgraded into a mixture of gasoline and diesel. This hydrogen could be purchased from off-site (assumed to be from a natural gas source) or be produced on-site by converting some of the bio-oil into hydrogen to upgrade the remaining bio-oil fraction. For either source of hydrogen, the pyrolysis process emits CO<sub>2</sub> only in a flue gas stream at ~32 wt.% (Wright et al., 2010). Lastly, in gasification, the biomass is thermochemically converted to syngas. As a part of generating and cleaning the syngas, a >95 wt.% CO<sub>2</sub> stream is generated. The syngas can be converted into liquid fuels by Fischer-Tropsch synthesis, which results in an additional flue gas stream with ~15 wt.% CO<sub>2</sub> (Swanson et al., 2010). From a cost perspective, fermentation and pyrolysis with hydrogen purchase have relatively low capital investment, but high operating costs, while gasification and pyrolysis with hydrogen production have higher capital costs and lower operating costs.

The different conversion technologies have been compared at a biorefinery with CCS, but without the additional consideration of spatially explicit landscape and SC design (Geissler and Maravelias, 2022). Conversely, other studies that have considered SC and CCS simultaneously have only one conversion and CCS technology at the biorefinery (Lainez-Aguirre et al., 2017). Modeling the landscape, SC, and biorefinery design simultaneously at a high spatial resolution can identify tradeoffs between decisions at each stage of the value chain. This modeling improves the understanding of the complex system interactions but is not meant to be a full decision support tool. In section 2 we present a model for simultaneous optimization of landscape, SC, and biorefinery design. In section 3 we discuss the results of applying this model to an 8-state region in the USA Midwest. Finally, in section 4 we present conclusions and discuss further research avenues.

## 2. Supply chain and biorefinery model

The integrated modeling approach for simultaneous landscape, SC, and biorefinery optimization largely follows the work of O'Neill et al. (2022) with several key modifications and extensions to account for the additional spatially explicit aspects of the large-scale system and the inclusion of CCS (O'Neill et al., 2022). The data sources and processing strategies are described in section 2.1 and important model constraints are presented in section 2.2. Many of the model constraints are omitted in the interest of brevity.

### 2.1 Data

We consider an 8-state region in the USA Midwest, an area of interest to the US Department of Energy. The study area consists of roughly 400,000 fields identified as historically abandoned lands with the potential to establish biomass (Lark et al., 2020). The growth of native bioenergy grass was simulated over a 40-year time horizon at 0

kgN/ha and 50 kgN/ha of fertilization (Martinez-Feria and Basso, 2020). The average value for biomass yield (Mg/ha) and the annualized soil organic carbon sequestration (Mg/ha/y) at each field are taken as inputs to the landscape design model. The 400,000 fields are aggregated according to the unique soil-type/weather combinations and landscape design decisions are made on the aggregated fields to reduce the model size while maintaining landscape design accuracy.

The large study area motivates the consideration of additional spatially explicit features. The price and CO<sub>2</sub> impact of electricity available at the biorefinery depend on the location of the installed biorefinery. Furthermore, the transportation and injection cost of captured CO<sub>2</sub> depends on the location of both the injection sites and the installed biorefineries. Figure 1(a) shows the study area including 30 potential biorefineries, 800 potential depots and the high-level distribution of biomass yield. Figure 1(b) shows the different fuel conversion and CCS technologies considered at the biorefinery.

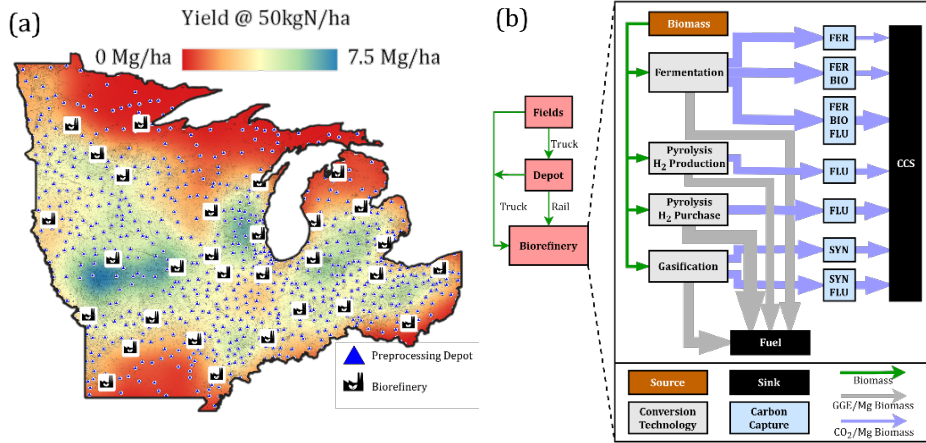


Figure 1. (a) Map of the study area and (b) summary of technologies considered at the biorefinery. The thickness of the gallons of gasoline equivalent (GGE) and CO<sub>2</sub> arrows is proportional to the relative mass/energy flows.

## 2.2 Selected model constraints

The objective of the integrated optimization model (equation 1) is to minimize the total annualized cost of the full field-to-product system consisting of capital cost  $C^{CAP}$ , feedstock cost including the annualized cost of crop establishment, harvesting, and handling  $C^{LND}$ , operational costs  $C^{PRD}$ , inventory costs  $C^{INV}$ , transportation costs  $C^{TRA}$ , and the cost (or credit) of electricity purchased (sold) at the biorefinery  $C^{ELC}$ .

$$OBJ = C^{CAP} + C^{LND} + C^{PRD} + C^{INV} + C^{TRA} + C^{CCS} + C^{ENV} + C^{ELC} \quad (1)$$

There are two terms in the objective function that account for the GHG emissions of the system. The credit for CO<sub>2</sub> captured at the biorefinery,  $C^{CCS}$ , is defined in equations 2-4.

$$C^{CCS} = \pi^{CCS} \sum_{i \in \{CO_2\}, l, t} L_{i, l, t} \quad (2)$$

$$P_{i', l, m, t}^{C-CB} = \sum_{i \in \{F\}} \eta_{i, i', m} G_{i, l, m, t}^{F-CB} + \sum_{i \in \{D\}} \eta_{i, i', m} G_{i, l, m, t}^{D-CB} \quad \forall i \in \{CO_2\}, l, m, t \quad (3)$$

$$L_{i, l, t} \leq P_{i, l, m, t}^{C-CB} \quad \forall i \in \{CO_2\}, l, m, t \quad (4)$$

where  $\pi^{CCS}$  is the carbon sequestration credit,  $L_{i, l, t}$  is the amount of species  $i$  (CO<sub>2</sub>) sequestered at refinery  $l$  during time  $t$ , and  $P_{i', l, m, t}^{C-CB}$  is the amount of CO<sub>2</sub> that is produced by technology  $m$  that is available for capture. We assume the balance of the CO<sub>2</sub> in

equation 4 is vented as biogenic CO<sub>2</sub> and does not contribute to the net emissions of the system.  $\eta_{i,i',m}$  is the conversion coefficient of biomass to CO<sub>2</sub> able to be captured.  $G_{i,l,m,t}^{F-CB}$  and  $G_{i,l,m,t}^{ID-CB}$  are the consumption of raw and densified feedstock. The set  $m \in \mathbf{M}$  includes the 11 possible combinations of conversion and CCS technology installations (Figure 1).

The GHG emissions from SC activities are described by equation 5 and include direct and indirect emissions from all sources other than the CO<sub>2</sub> captured at the biorefinery.

$$C^{ENV} = \pi^{SCC} (GHG^{TRA} + GHG^{LND} + GHG^{PRD} + GHG^{ELC}) \quad (5)$$

where  $\pi^{SCC}$  is the social cost of carbon, the monetary value placed on SC emissions (Nordhaus 2017). Emission sources include transportation, landscape (including annualized soil carbon sequestration), production, and the indirect contribution from grid electricity. Due to the large study area, the location of biorefineries and preprocessing depots determines the price and CO<sub>2</sub> impact of the available electricity and the parameters are considered in a spatially explicit manner in equations 6-8.

$$P_{i',l,m,t}^{E-CB} = \sum_{i \in I^F} \eta_{i,i',m} G_{i,l,m,t}^{F-CB} + \sum_{i \in I^{ID}} \eta_{i,i',m} G_{i,l,m,t}^{ID-CB} \quad \forall i \in \{ELC\}, l, m, t \quad (6)$$

$$P_{i',l,m,t}^{E-PD} = \sum_{i \in I^F} \eta_{i,i',m} G_{i,k,m,t}^{F-PD} \quad \forall i \in \{ELC\}, k, m, t \quad (7)$$

$$GHG^{ELC} = \sum_{i,l,m,t} \Gamma_l P_{i,l,m,t}^{E-CB} + \sum_{i,k,m,t} \Gamma_k P_{i,k,m,t}^{E-PD} \quad (8)$$

where  $\Gamma_l$  and  $\Gamma_k$  are the CO<sub>2</sub> impact at each biorefinery  $l$  and depot  $k$  from the local grid. Note that the production of electricity at the biorefinery  $P_{i',l,m,t}^{E-CB}$  is allowed to be negative which represents a technology configuration that produces excess electricity that is sold back to the grid and offsets grid emissions.

### 3. Results

Model instances are implemented in GAMS and solved using Gurobi 9.5.1. We consider two design cases to analyze the spatial interactions involved in the system. First, we consider the 'status-quo' case which corresponds to the current incentive structure in the USA where a credit is only applied to CO<sub>2</sub> captured using CCS (case A:  $\pi^{SCC} = 0, \pi^{CCS} > 0$ ). Case A neglects to account for GHG emissions from SC activities and indirect emissions from other sources. The second case (case B:  $\pi^{CCS} = \pi^{SCC}$ ) considers all CO<sub>2</sub> emissions equally. Case B treats avoided emissions in the SC the same as CO<sub>2</sub> captured with CCS. That is, the net emissions in CO<sub>2</sub> equivalents, including both CO<sub>2</sub> from CCS and from the SC receive the same economic credit or penalty in the objective function. By considering the net emissions of the entire system, the emissions at the landscape and SC level contribute to spatial tradeoffs that affect the optimal design, operation, and biorefinery technology portfolio.

Figure 2 shows a breakdown of the cost and GHG emissions for a representative sample of biofuel demands and sequestration credits for the two cases discussed above. In case A, the cost contributions remain relatively constant from each source. Transportation costs increase slightly at higher demands as biomass must be transported farther. The sequestration credits of 85 \$/Mg and 125 \$/Mg incentivize CCS and, as a tradeoff, there are higher contributions from the electricity and capital cost needed to perform CCS and receive the credit. In case A, pyrolysis with hydrogen purchase with (P<sup>H+</sup>) and without (P<sup>H</sup>) CCS is incentivized due to its high fuel yield. When CCS is incentivized, the biorefinery requires electricity, which is purchased from the grid.

In case B, P<sup>H</sup> and P<sup>H+</sup> are incentivized at lower credits, and gasification with syngas CCS (G<sup>+</sup>) and gasification with both syngas and flue gas CCS (G<sup>++</sup>) are incentivized at high credits. Fermentation results in two relatively concentrated CO<sub>2</sub> streams that are easy

to capture, and pyrolysis with H<sub>2</sub> production does not have the indirect emissions from the purchase of H<sub>2</sub>; however, because of the relatively high costs and low yields of fermentation and pyrolysis with H<sub>2</sub> production, they are never the optimal technology for any combination of demand and sequestration credit.

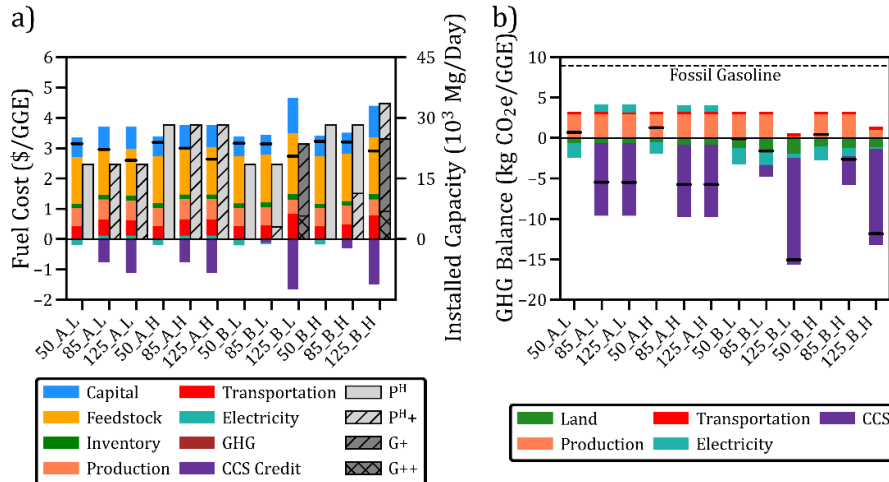


Figure 2. (a) Breakdown of costs and installed capacity of each technology and (b) GHG emissions for representative scenarios. The three indicators in the scenario names are the sequestration credit, case designation (A or B), and demand (L or H for a demand of 500MM of 850MM GGE/y, respectively). The horizontal black lines on the stacked bar show the overall cost and GHG balance.

When all GHG emissions are valued equally, CCS is no longer incentivized at every biorefinery at 85 \$/MgCO<sub>2</sub>. Instead, the spatial features of the system dictate where CCS is incentivized. Interestingly, at 85 \$/MgCO<sub>2</sub> the net emissions for the cost-optimal system are higher when all emissions are considered equally. This is because in most locations, the sequestration credit is not sufficient to incentive CCS due to the economic penalty of GHG emissions from electricity. The sequestration credit does not offset the capital and operating cost of the additional CCS technology when also considering the indirect emissions of electricity except in areas where electricity is very cheap and clean.

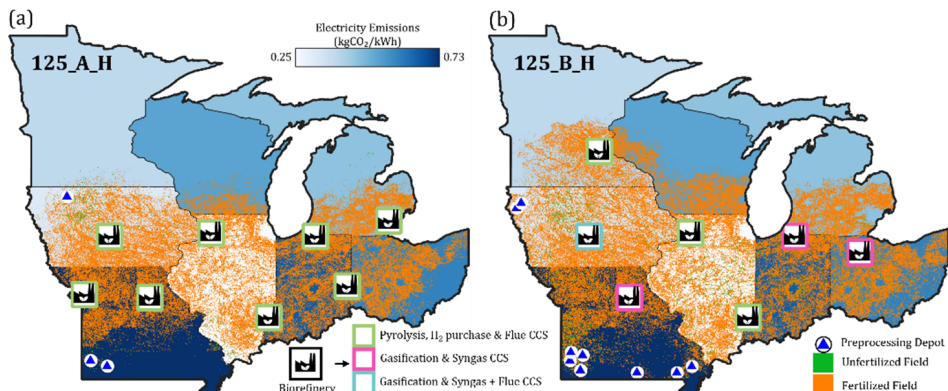


Figure 3. The supply chain configuration for a sequestration credit of 125 \$/MgCO<sub>2</sub> and a demand of 850MM GGE/y for case A and B.



Figure 3. shows the spatial changes at a high sequestration credit (125 \$/MgCO<sub>2</sub>). In case B, there is a dramatic shift in both the net emissions and the optimal technology portfolio. Gasification is preferred because it avoids indirect emissions from H<sub>2</sub> needed for pyrolysis with H<sub>2</sub> purchase and is essentially energetically self-sufficient. The high demand incentivizes constructing biorefineries with high yielding P<sup>H+</sup> to meet demand without having to plant poor fields. Refineries are strategically positioned so that P<sup>H+</sup> is installed in states with low electricity emissions and G+/G++ are installed in states with higher electricity emissions. Gasification has a lower fuel yield than pyrolysis which increases the amount of biomass that is planted leading to higher transportation and handling costs, but results in an additional opportunity for soil carbon sequestration. This increase in the amount of required biomass is somewhat offset by careful landscape design in which fields with high soil carbon sequestration are prioritized.

#### 4. Conclusions

In this paper we used mixed-integer linear programming methods to analyze the spatially explicit tradeoffs for a full field-to-product biofuel production system under two incentive structures. We showed that for the USA Midwest, the current sequestration credit incentive may neglect to account for emissions related to the system's spatial interactions such as soil carbon sequestration, the local price and CO<sub>2</sub> emission rate of electricity, and biomass availability. When valuing all CO<sub>2</sub> equivalents equally from an economic perspective, the cost-optimal SC and technology portfolio requires higher credits to fully incentivize CCS. Additionally, the optimal technology shifts from pyrolysis to gasification because of the lower indirect emissions and the resulting landscape design plants additional biomass to compensate for the lower fuel yield. We also demonstrated the importance of large-scale modeling that considers the interactions of the full field-to-product system, particularly at a high resolution. The high-resolution modeling captures the system wide effects of CO<sub>2</sub> mitigation incentive structures and highlights the sensitivity of cellulosic biofuel landscapes and SCs to those incentives.

#### References

- J.L. Field, S.G. Evans, E. Marx, M. Easter, P.R. Adler, T. Dinh, B. Wilson, K. Paustian, 2018, High-Resolution Techno-Ecological Modelling of a Bioenergy Landscape to Identify Climate Mitigation Opportunities in Cellulosic Ethanol Production, *Nat. Energy*, 3, 211-19
- C.H. Geissler, C.T. Maravelias, 2022, Analysis of alternative bioenergy with carbon capture strategies: present and future, *Energy Environ. Sci*, 15, 2679-2689.
- D. Humbird, R. Davis, L. Tao, C. Kinchin, D. Hsu, A. Aden, 2011, Process design and economics for biochemical conversion of lignocellulosic biomass to ethanol: dilute acid pretreatment and enzymatic hydrolysis of corn stover, NREL/TP-5100-47764.
- J.M. Lainez-Aguirre, M. Pérez-Fortes, L. Puigjaner, 2017, Economic evaluation of bio-based supply chains with CO<sub>2</sub> capture and utilisation, *Comput. Chem. Eng.*, 102, 213-225.
- T.J. Lark, S.A. Spawn, M. Bougie, H.K. Gibbs, 2020, Cropland expansion in the United States produces marginal yields at high costs to wildlife, *Nat. Commun.*, 11, 1, 1-11
- R.A. Martinez-Feria, B. Basso, 2020, Predicting soil carbon changes in switchgrass grown on marginal lands under climate change and adaptation strategies, *GCB Bioen.*, 12, 9, 742-755
- W.D. Nordhaus, 2017, Revisiting the social cost of carbon, *PNAS*, 114, 7, 1518-1523
- E.G. O'Neill, R.A. Martinez-Feria, B. Basso, C.T. Maravelias, 2022, Integrated Spatially Explicit Landscape and Cellulosic Biofuel Supply Chain Optimization Under Biomass Yield Uncertainty, *Comput. Chem. Eng.*, 160, 107724.
- R.M. Swanson, J.A. Satrio, R.C. Brown, A. Platon, D.D. Hsu, 2010, Techno-economic analysis of biofuels production based on gasification, NREL/TP-6A20-46587.
- M.M. Wright, J.A. Satrio, R.C. Brown, D.E. Daugaard, D.D. Hsu, 2010, Techno-economic analysis of biomass fast pyrolysis to transportation fuels, NREL/TP-6A20-46586.

# A MILP-based approach to manage logistics in large industrial gas supply chains

Sergio G. Bonino,<sup>a</sup> Luis J. Zeballos,<sup>a</sup> Akash Moolya,<sup>c</sup> Jose Lainez,<sup>c</sup> Jose M. Pinto,<sup>c</sup> Ignacio E. Grossmann,<sup>b</sup> Carlos A. Méndez<sup>a</sup>

<sup>a</sup>*Intec (UNL-CONICET), Güemes 3430, Santa Fe 3000, Argentina*

<sup>b</sup>*Carnegie Mellon University, 5000 Forbes, Pittsburgh 15213, United States*

<sup>c</sup>*Praxair, Inc., a Linde company, Connecticut 06810, United States*

*cmendez@intec.unl.edu.ar*

## Abstract

The high competitiveness between companies dedicated to the production and distribution of industrial gases make necessary to optimize their activities in order to improve their profitability. This paper addresses the production routing problem (PRP) of an industrial gas supply chain (SC). The work proposes a MILP-based approach to deal with this problem. The framework seeks to improve the company's profit by optimizing production and distribution activities. For the distribution part, a set of feasible routes are previously generated and selected. Then, considering only the selected routes, the model chooses the best combination of them to satisfy all the constraints. It is important to note that the model takes into account the effect of lead time on production and distribution decisions. Finally, a real case study is presented and solved. The results obtained demonstrate the effectiveness of the proposed approach in reasonable CPU times.

**Keywords:** Supply chain management, Production routing problem, Mathematical modelling.

## 1. Introduction

Recently, Vendor-managed inventory (VMI) policy has begun to be widely used in various industrial sectors. The gas industry has been one of them, and inventory is now controlled by the companies for most of its customers. This policy consists of the company constantly monitoring its customers' inventory to make decisions on when and how much product to deliver to its customers to ensure the proper performance of its daily tasks. Under this policy, customers should no longer issue replenishment orders to the companies (Dong et al., 2017). Due to this paradigm shift, companies are focusing their efforts on improving the SC performance in order to remain competitive in an increasingly aggressive market. One of the most important changes so far to try to optimize the logistics of companies is to simultaneously solve the problems of production and distribution of the products (Zhang et al., 2017). In a conventional approach, the production problem would be solved first, followed by distribution decisions. This sequential method may generate suboptimal solutions or in the worst case infeasible. By solving the integrated problem, lower-cost solutions can be found than those obtained with the sequential approach (Lee et al., 2022).

This paper addresses the problem in an integrated manner, considering explicitly the lead times in a VMI policy.

## 2. Problem Statement

This work addresses an industrial gas SC's production routing problem (PRP). The objective is to optimize the operation of the SC, seeking to minimize the total cost of production and distribution activities. The particularity of this SC is that the company that supplies industrial gases is in charge of controlling and managing the customers' inventory. This practice is known as vendor-managed inventory (VMI). For this reason, knowing the monthly demands of customers and their maximum and minimum inventory capacities, the company must generate optimal production and distribution strategies that allow it to visit customers on time to meet their demands. The products considered in this case are liquid oxygen and nitrogen. To meet daily demands, the company has production plants and inventory storage tanks. Each plant has a maximum production rate of product per day. Regarding distribution, each plant has a fleet of trucks for each type of product. A relevant characteristic of the problem considered here is that trucks can make several short trips per day or a single trip of several days. The trip duration depends on the distance to the clients considered in the trip, the order in which customers are visited, and the loading and unloading times of the products. Since the trips made can last several days, this problem explicitly considers the effect that the lead time has on the following:

- The time at which the product is released from the plant.
- The time at which the product arrives at the customer.
- The amount of time the trucks are not available for other trips.

## 3. Solution strategy and main assumptions

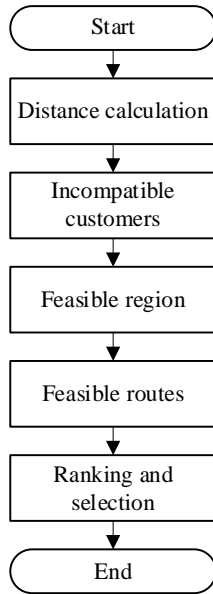
For this problem, an efficient discrete-time MILP approach to simultaneously minimize the production and distribution costs of the PRP is proposed. To address the mentioned problem characteristics, the model is based on the idea of the explicit generation/representation of the trips, thus when a trip is selected all the information about it is available and considered in the approach. The central decision variables of the model are a) the amount of each product manufactured in each plant per day, b) the customers visited per day, c) the customers visited on the same trip, d) the order in which customers are visited, e) the trips that trucks make each day, f) the amount of product loaded in the trucks per trip, g) the amount of product delivered to each customer included on the trips.

The main assumptions under which the model was developed are as follows:

- Trucks can only be used 12 hours per day. More time implies a multiday trip.
- Routes of up to 48 hours (4 days) are considered.
- Routes of up to 3 customers are considered. The features of this problem makes this assumption valid.
- Lead time is taken into account when scheduling a customer visit.
- Trucks making trips longer than twelve hours are not available for use until their return to plant.
- Each customer's ending inventory is allowed to be below a certain percentage of the beginning inventory.
- The total ending inventory must be greater than or equal to the total beginning inventory.
- A truck can make 2 trips in one day only if the total time of both trips is less than or equal to 12 hours.
- A customer can be visited at most twice in the same day.

### 3.1. Route generation procedure

This section introduces a summary of the route generation procedure. The structure of this procedure is shown in Figure 1.



- The algorithm starts with the calculation of the distance and travel times between nodes, (i.e., customer-customer and plant-customer distance).
- Once the travel times between nodes are known, customer combinations that may lead to infeasible routes are discarded. This is because the total travel time of the route exceeds the maximum allowed or the customer demands are not compatible to be realized in the same trip.
- Step 3 defines a feasible region of visit for each customer starting from a given plant. This idea was initially presented by Dong et al. (2017) and consists of determining a region through a cone. Subsequently, each customer will be able to share routes with customers located within its region.
- In step 4, all possible feasible routes are generated, taking into account the regions determined in the previous step and the constraints between customers that cannot be on the same route.
- Finally, for each customer, a ranking of the routes is made taking into account the total time of the route and the arrival time to the customer. The ranking is done for each plant and each type of route (i.e., direct shipment, 2-customer route, 3-customer route). Subsequently, for each customer, the best 'x' routes per plant and per type are selected. This selection is made to reduce the number of possible alternatives to visit a client and reduce the combinatorial complexity of the problem to be solved.

**Figure 1.** Structure of Route Generation Procedure

The feasible routes generated with the procedure described above, provide the following information a) travel time of the route, b) clients visited, c) arrival time at each client, d) plant from where the route departs.

### 3.2. Model formulation

The MILP model developed aims to minimize the costs of production and distribution operations. In the subsections, due to the lack of space only the main equations of the model are presented. The definition of sets, parameters and variables of the models are defined as follows:

#### Sets

- $T$  time periods
- $I$  products
- $C$  customers
- $P$  plants
- $S$  routes
- $SI$  routes  $s$  transporting product  $i$
- $SP$  routes  $s$  departing from plant  $p$
- $SC$  routes  $s$  visiting the customer  $c$
- $IC$  customer  $c$  consuming product  $i$
- $IP$  plants  $p$  producing product  $i$

#### Parameters

- $day_{c,s}$  day of arrival at customer  $c$  with route  $s$
- $dayRoute_s$  duration in days of route  $s$

#### Continuous Variables

- $LD_{i,p,t}$  Product  $i$  transported from plant  $p$  during period  $t$
- $\overline{DL}_{i,p,s,t}$  Product  $i$  transported from plant  $p$  in period  $t$  using route  $s$
- $\widehat{DL}_{i,p,s,c,t}$  Product  $i$  transported to customer  $c$  from plant  $p$  in period  $t$  using route  $s$
- $DL_{i,c,t}$  Product  $i$  received by customer  $c$  in period  $t$
- $TU_{i,p,t}$  Trucks not available at plant  $p$  for period  $t$
- $CPP_{i,p,t}$  Production cost of plant  $p$  in period  $t$
- $CTP_{i,p,t}$  Transportation cost of plant  $p$  in period  $t$
- $PCC_t$  Purchase cost for period  $t$
- $TC$  Objective function (total cost)

#### Binary variable

- $A_{i,p,s,t}$  1 if route  $s$  is used in period  $t$

### 3.2.1. Inventory update

This set of constraints is intended to update customer and plant inventories. In a given period, the inventory in the plants will be modified by the production made and the amount of product used to satisfy customer demand. The inventory of customers is updated taking into account their daily demand and the amount of product received through shipments or purchases.

Generally, in the literature, updating customer inventory takes into account the current period (t) and the previous period (t-1). In the proposed formulation, this idea is extended so that the quantity received by shipments ( $DL_{i,c,t}$ ) takes into account products that left the plant up to t-3 periods ago.

### 3.2.2. Production and Distribution

All decisions related to the daily quantity to be produced at each plant are considered here, as well as determining which customers will be visited, and defining the assigned route and the quantity to be transported.

In this section, new restrictions appear for the treatment of lead time in deliveries. Equations 1 to 3 show how lead time is treated concerning product distribution.

$$LD_{i,p,t} = \sum_{s \in (SP \wedge SI)} \overline{DL}_{i,p,s,t}, \quad \forall p, i \in (IP_{p,i}) \quad (1)$$

$$\overline{DL}_{i,p,s,t} = \sum_{c \in (IC_{c,i} \wedge ASC)} \widehat{DL}_{i,p,s,c,t}, \quad \forall p, i \in (IP_{p,i} \wedge SI \wedge SP) \quad (2)$$

$$DL_{i,c,t} = \sum_{(p,s) \in (IP \wedge ASC \wedge SI \wedge SP \wedge (tt=t-day_{c,s}+1))} \widehat{DL}_{i,p,s,c,tt}, \quad \forall c, i, t \in (IC_{c,i}) \quad (3)$$

Equation 1 calculates the quantity of product that is transported from a given plant in a given period. Equation 2 indicates how much product is delivered to each customer when they are visited with route 's'. Finally, in equation 3, it is calculated when the customer receives the delivery taking into account the lead time, this is considered with the condition  $tt = t - day_{c,s} + 1$ .

### 3.2.3. Truck availability

The set of equations belonging to this subsection limits the number of available trucks per period and avoids the use of trucks that are traveling for more than 12 hours. These trucks become available again once the trip has been completed.

Constraint 4 shows how to calculate the number of unavailable trucks in period t for a given plant ( $TU_{i,p,t}$ ). This is done only for trucks performing trips longer than 12 hours.

$$TU_{i,p,t} = \sum_{(s,tt) \in (SP \wedge SI \wedge (t-dayRoute_s+1 \leq tt \leq t-1) \wedge (dayRoute_s > 1))} A_{i,p,s,tt} \quad (4)$$

$$\forall p, i, t \in (IP_{p,i} \wedge (t > 1))$$

### 3.2.4. Cost estimation and Objective Function

The objective function considered seeks to minimize the total operating cost, which is made up of the cost of production, transportation and purchasing. It is shown in equation 5.

$$\text{minimize } TC = \sum_{(p,i,t) \in (IP_{p,i})} CPP_{i,p,t} + CTP_{i,p,t} + \sum_t PCC_t \quad (5)$$

#### 4. Case study and computational results

The real case study involves an industrial gases SC with 5 production plants, 2 products (liquid oxygen and nitrogen), approximately 500 monthly deliveries to 114 customers, 13 vehicles and a 30-day time horizon. Due to the large size of the addressed instance, several heuristics strategies and fit constraints were introduced to reduce the combinatorial explosion issue. Heuristics strategies are mainly dedicated to efficiently maintaining limited the number of trips considered. On the other hand, fit constraints are aimed at limiting the search for solutions to promote better truck-loading capacity utilization.

Due to space limitations, only the results obtained for the case of liquid oxygen (LO2) will be presented. It should be noted that the model can be used for both products. The LO2 problem have 5 production plants, 47 customers and 7 trucks for shipments. The location of customers and plants is shown in Figure 2. Plants 4 and 5 (P4 and P5) have the highest production rate and lowest costs, but shipping from there involves several days.

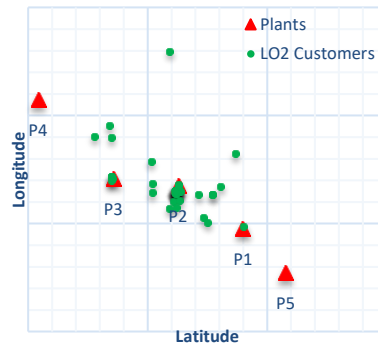


Figure 2. Plant and customer location

Three different scenarios were run for this case. In the first scenario, only direct shipment routes were considered. In scenario 2, all feasible routes generated with the algorithm were used. Finally, in the third scenario, the best route per plant and per type was selected for each customer. All scenarios were implemented in GAMS 32.0 and solved with GUROBI 10.0.0 on a desktop computer with 2.80GHz Intel(R) Core(TM) i9-10900F processor and 16 GB RAM on Windows 10. The maximum run time for each scenario was set to 3600 seconds.

Table 1 shows the results obtained for the different scenarios. In all 3 solutions, the model was able to satisfy all demand through shipments, thus avoiding purchases from external suppliers. As can be seen, when the model is given the possibility of combining customers on the same route, trips are better utilized and, consequently, the total cost is reduced (Scenario 1 vs. Scenarios 2 and 3).

Table 1. Results of the proposed scenarios

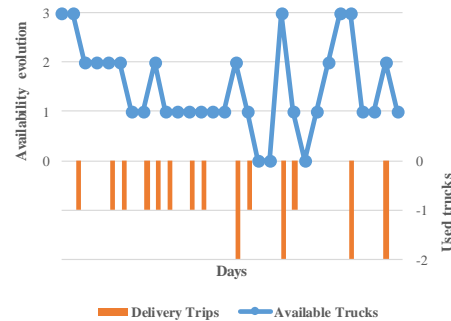
Scenario	Model Size		Number of Routes	Total Cost [\$]	GAP [%]
	Continuous Variables	Discrete Variables			
1	12,421	3,780	121	203,490.99	2.54
2	756,181	194,340	6,473	190,216.49	1.83
3	30,481	9,000	295	190,155.24	0.63

By comparison of scenarios 2 and 3, it can be observed that the route selection criteria considered for solving the model has an important impact on the complexity of the problem and the quality of the solutions that can be obtained in the same processing time. Another positive aspect to highlight of scenario 3 is that it was able to reach a solution with a 1.8% GAP in 100 seconds, while scenario 2 reached it in 3600 seconds of CPU time.

Figures 3 and 4 show examples of how lead time influences decisions. Figure 3 presents the inventory profile of customer 110, who receives shipments of product on days 3, 7, 13, 17, 22 and 27. These shipments leave plant 5 on days 2, 5, 12, 16, 21 and 26. The graph clearly shows the difference between the day the truck leaves the plant and the day it arrives at the customer. Figure 4 shows available trucks per day in plant 5, taking into account when trucks leave and return to the plant. Plant 5 has a fleet of 3 trucks.



**Figure 3.** Inventory profile for customer 110



**Figure 4.** Truck availability at plant 5

## 5. Conclusion

The proposed MILP-based approach was found helpful to optimize decisions concerning daily production and customer deliveries. Lead time considerations were implemented, resulting in more accurate solutions. The computational results demonstrate the effectiveness of the proposed approach, obtaining efficient solutions in reasonable CPU times.

## References

- Dong, Y., Maravelias, C. T., Pinto, J. M., & Sundaramoorthy, A. (2017). Solution methods for vehicle-based inventory routing problems. *Computers and Chemical Engineering*, *101*, 259–278.
- Lee, Y., Charitopoulos, V. M., Thyagarajan, K., Morris, I., Pinto, J. M., & Papageorgiou, L. G. (2022). Integrated production and inventory routing planning of oxygen supply chains. *Chemical Engineering Research and Design*, *186*, 97–111.
- Zhang, Q., Sundaramoorthy, A., Grossmann, I. E., & Pinto, J. M. (2017). Multiscale production routing in multicommodity supply chains with complex production facilities. *Computers and Operations Research*, *79*(April 2016), 207–222

# Modelling of reverse supply chains in the context of circular economy

Christiana M. Papapostolou, Emilia M. Kondili

*Optimisation of Production Systems Laboratory*

*Dept. of Mechanical Engineering, University of West Attica, Ancient Olive Grove*

*Campus, P. Ralli & Thivon Street, 12244, Egaleo, Greece*

## Abstract

The circular economy model has created a shift paradigm in materials and products utilisation throughout and at the end of their life cycle, by improving the economic, social and environmental performance of the respective supply chains both in terms of operation as well as in terms of interaction with the entire ecosystem. On that note, reverse supply chains offer an opportunity for manufacturing companies and practitioners to improve the supply chain performance too. By gaining the customers insights, the entire supply chain may be redesigned to fit the repurposing of improved services, eco-friendliness (also in line with Green Supply Chain Management) whilst at the same time meeting as well the cost minimisation goal. Reverse supply chains coupled with circular economy model principles can provide an optimised, modern model of supply chains operation, as waste-to-products resources and flows can be redesigned so as to maximise their use in the improved SC network.

The aim of this paper is to introduce via indicative case-studies, from the field of manufacturing, a modelling approach of reverse supply chains in the context of circular economy and provide valuable knowledge, alternatives and possibly organisation insights to make supply chains more sustainable. The developed modeling framework is based on the Resource-Task-Network leading to a novel approach for reverse supply chain optimisation.

**Keywords:** sustainability, optimisation, reverse supply chains, RTN

## 1. Introduction

Over the last decades the scope and definition of Supply Chains (SCs) and the relative Management (SCM) was widen to include aspects dealing with the sustainability at all SC stages. SCM, which is as an integrated network of business units, producers, traders, retailers and consumers, with the relative information and materials flow from the supplier, to the producer of the final product up to the final consumer, requires as minimum operational conditions: “A systems approach to viewing the supply chain as a whole, and to managing the total flow of goods inventory from the supplier to the ultimate customer; a strategic orientation toward cooperative efforts to synchronize and converge intrafirm and interfirm operational and strategic capabilities into a unified whole; a customer focus to create unique and individualized sources of customer value, leading to customer satisfaction” (Mentzer et al., 2001).



Owing that to a nexus of reasons (increased consumer awareness, legally imposed environmental constraints), the final SC stage which delivers the end product to customer, has grown in importance, necessitating the need for safe product returns as well as more environmentally friendly products and by-products, such as recyclable packaging materials. Reverse Supply Chain (RSC) which is the SC that starts from the end consumer and goes up to the step of materials acquisition, also entails the collaborative responsibility of suppliers, producers and consumers to reduce the waste by recycling.

Transforming the current, linear economic model towards a resource-efficient Circular Economy (CE) model requires a uniform understanding of the interactions amongst the various decision factors. Integrating the CE perspective implies that emphasis will be given on product transformation, reverse loops of materials and energy, circularity of resources in general, also enabling waste elimination and environmental pollution decrease. Hence, it is very critical to study the incorporation of the CE aspect in RSCs and which is the appropriate evaluation framework.

## **2. Reverse Supply Chains and Circular Economy**

At the moment the classic, linear SC now called forward SC is at a transformative phase aiming at including responsible management of end-of-life (EoL) products and also sustainable manufacturing. Under that pressure, firms are considering designing their SC network to take control of the reverse flow of products as well. Reverse supply chain management (RSCM) focuses on materials recovery and inventory management issues with typical operational decisions including design, planning, coordination of inventory management, outsourcing/ 3rd party logistics selection, vehicle routing problems, recovery, recycling and return of materials. Transforming forward into reverse SCs (RSCs) aims at ensuring waste minimisation while reducing the excessive usage of raw materials and at the same time protecting the environment.

CE approach “focuses on the circularity of energy and material resources which provides economic, environmental and social benefits for manufacturing organisations” (Yadav et al., 2020). CE adoption concept supports the vision for transforming the linear Take-Make-Dispose to Take-Make-Use of material and energy flows to a model that entails reusability and recyclability along with environmental, economic and social benefits.

Energy and materials reuse and recycle are common industrial practices. However, the CE approach recently embraced, has put emphasis on: a) processes dealing with repairing and remanufacturing and b) the potential utilisation of renewable energy sources (RES) throughout the product value chain. As Korhonen et al., (2018) states, “CE promotes high value material cycles alongside more traditional recycling and develops systems approaches to the cooperation of producers, consumers and other societal actors in sustainable development work”. As an evidence to that, lately, the coupling of industry 4.0 (I4.0) and CE is also widely examined by researchers (Dev et al., 2020, Stylianopoulou et al., 2022), recognising the synergies at SC level.

The CE business model reorients efficient, waste / secondary material utilisation and also creates multidimensional values; economic value (profit maximisation, cost reduction), customer value (satisfaction, awareness), environmental (sustainability, etc) (Pal et al. 2019). Effective management of reverse logistics provides the possibility of value maximisation across the entire SC by choosing the most appropriate recovery alternative

that can generate additional profitability and reduce the utilisation of resources. Communication of information across different levels, concerning the volume and type of returned products, the customers' attitudes, the customers' preferences regarding the environmental impact of products, the changes in the consumer model (i.e. e-commerce), are pushing the firms for improved customer services, new business strategies, changes in products' materials and their respective life cycle.

The CE model employed in each RSC it's a nexus of decisions based on the type of business and inventory model and to the required network and logistics design. Therefore, in the following section three representative case studies are listed from the field of manufacturing, highlighting the benefits of coupling CE model principles in RSC remodelling.

### 3. Case studies from the field of Manufacturing

Selected cases studies from different field of manufacturing will be presented in this section; cases of reverse supply chain modelling coupled with CE approach and related decision, highlighting the importance of CE and RSCs "engagement".

*Field: pharmaceutical waste CE Decisions in RSCs: coordination of inventory management, third party logistics operation, recycling and return of materials*

Nami and Hosseini-Motlagh (2022) developed a non-linear programming model that also includes a flexible transfer price agreement, and the Nash-bargaining game model to coordinate all parties of pharmaceutical RSC (i.e., pharma-manufacturer (DP Company), 3PL company, and customers), considering the initial number of unwanted medications as a source of uncertainty. Two reverse SC types are examined: a) the medications with an expiry date of more than six months from the collection date can be legally resold by DP in secondary markets and b) the medications with an expiry date of less than six months from the collection date, are safely discarded in disposal sites by DP under a fixed disposing cost. **Objective function:** profit maximisation of 3PL. **Model added value:** Improved responsiveness of the SC to revenue fluctuations, improved profitability of the entire production system, in the sustainability pillar / social helpful when the customers are reluctant to return their items.

*Field: car sharing waste components CE Decisions in RSCs: remanufacturing, recycling*

Rentizelas and Trivyza (2022) propose an optimised design of a RSC for reusable parts in sharing vehicles (type of vehicles of high mileage and short service life). They focus in the reusability and re-manufacturability of carbon fibre reinforced polymer car frame, which has a long-life span and light weight properties. The network design was optimised using a MILP model for three scenarios of car reusability at the end of the car life, for the case study of the UK. **Objective function:** minimisation of RSC annual cost. **Model added value:** Improved percentage of remanufactured frames / impact of economies of scale in cost reduction demonstrated.

*Field: membrane fuel cell used in hydrogen vehicles CE Decisions: network design, remanufacturing, recycling*

Alkahtani and Ziout, (2019) developed a multi-objective mixed integer linear model in order to design the reverse logistic, recovery supply chain network for remanufacturing the proton-exchange membrane fuel cell used in hydrogen vehicles of General Motors. **Objective function:** minimisation of the RSC network costs, including environmental

factors (CO<sub>2</sub> emission) and social aspects (number of job openings), and to maximisation of social benefits. **Model added value:** Improved collection network design (economically, socially and environmentally) for collecting used products and locating to the remanufacturing plant.

*Field: food SC – case study: coffee SC CE Decisions in RSCs: alternative process pathways, minimisation of waste losses*

(Baratsas *et al.*, 2021) developed a novel CE system engineering framework and decision-making tool for the modeling and optimisation of food supply chains. The MILP developed model is tested for a coffee SC under different goals and objectives. Five scenarios for the supply of coffee cherries and the demand of final coffee products are investigated for each of the single or multi-objective optimisation problem. **Objective functions:** Minimisation of generated waste, Maximisation of recycling, Minimisation of consumed raw materials and natural resources, Maximisation of total energy and/or renewable energy output, Minimisation of GHG emissions, Maximisation of product, equipment and packaging durability. **Model added value:** More efficient process pathways, improved design and operation of the SC.

#### 4. The Reverse Supply Chain framework features

As evidenced from the cases listed above, typically in RSCs the decisions deal with: a) the design of the SC (i.e. improvement additional manufacturing-oriented methods for product design) b) the logistics/ supply networks, c) the inventory/ capacity of the warehouses d) the inventory management methods are of critical decisions to be made within the scope of the redesigned SC, e) the product recovery, while maximising the benefit along the entire SC.

While the CE aspect inclusion in the design and operation of the optimised RSCs can provide insight to the following:

- For which selling price of secondary raw materials is the RSC-CE economically viable?
- To which kind of remanufacturing and repairing operation should we focus?
- To which operation should we invest? i.e. raw material production or secondary raw materials?
- Which is the market demand for secondary raw materials? What is the required purity for the secondary raw materials?
- Which is the environmental footprint of the “downstream” operations such as the recycling and waste treatment?

The proposed RSC-CE framework seeks to appropriately model the interdependencies among the supply chain “parts” by providing key focus both on the forward and the reverse operations according to the scope of the optimisation problem. It entails the construction of Resource-Task-Network (RTN) representation and a generic model formulation linking interactions between tasks and resources. Resources, typically may include any type of material and/or energy in different states represented as a circle and tasks, represented as a rectangle, include processes (i.e. production, recycling, repairing, sorting, energy recovery etc) (Figure 2). Resources modelling also enables inventory

management decisions (i.e. flows of materials supplier to consumers, selection of the supply network etc).

The benefit of the present modelling framework, which is meant to be formulated as a MILP model to include binary variables (0,1) for all type of decisions points, is that allows based on the focus aspect either of the CE or of the RSC to formulate the relevant optimisation criterion. Some indicatives examples are illustrated in Table1, whilst the key components of the model are listed in Table 2.

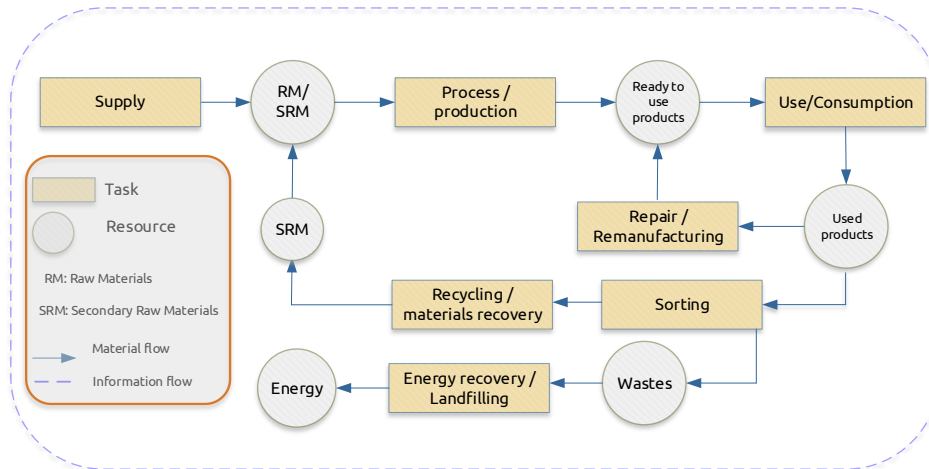


Figure 2: The remodelled RSC under the CE and the RTN concept

Table 1: Optimisation criterion and RSC and CE focus points.

RSC focus	Optimisation Criterion	CE - Focus
On RMs supply in the production stage (decrease)	Minimisation of RMs supply	Increase of SRMs production
On wastes generated (decrease)	Maximisation of SRMs production	Invest/ select RMs with longer life span
On wastes generated (decrease)	Maximisation of recycling rate	Recovered materials to meet the market needs
SC environmental footprint (decrease)	Minimise CO <sub>2</sub> eq through the entire SC	More efficient processes, maximisation of RES utilisation

Table 2: Modelling

Sets
Resources (raw materials, secondary raw materials, ready-to-use products, used products, wastes, energy)
Time interval for which the problem is considered (hour, day, week, month, year)
Processes (production, consumption, sorting, repairing, recycling, landfilling)
Variables
Decision variables (Based on the selected optimisation criterion): Quantities of resources /year (produced, recycled, returned, reused)
Binary variables - For selecting the appropriate process pathway

Economic parameters
Prices of resources (€/item, €/kg...)
Process - based costs (€/processed item)
Transportation costs (€/processed item)
Process-conversion coefficients
Demand – Balance Constraints
Demand for raw materials (kg/year)
Demand for new products (items/year)
Demand for secondary raw materials (kg/year)
Variables (cont.)
Demand for ready-to-use products (new and refurbished (items/year)

Capacities of the processes (items or kg or m <sup>3</sup> /year)
<b>Environmental – CE</b>
Environmental footprint of the processes (tn CO <sub>2</sub> /per item recovered)
Emissions constraints over the planning horizon (tn CO <sub>2</sub> /year)
<b>Demand – Balance Constraints (cont.)</b>

Wastes constraints over the planning horizon (kg, m <sup>3</sup> /year)
Energy recovered over the planning horizon (% of total input)
Circular material use rate
End-of-life recycling input rates

## 5. Conclusions

In this work the basic structure and principles of an optimisation framework aiming at modelling RCSs in the context of CE is presented. The existing experience evidences that CE can strongly relate to RSCs remodeling and optimisation as it highlights the practical pathways towards that transition. Moreover, the literature review revealed that companies are interested in becoming more efficient and sustainable along the different SC manufacturing stages and are making real efforts towards that direction. A very interesting aspect, that will be investigated in a later work is the coupling of closed – loops SCs with the CE and RSCs in a holistic way.

## Acknowledgments

*The publication /registration fees were totally covered by the University of West Attica.*

## References

- M. Alkahtani, A. Ziout, 2019, Design of a sustainable reverse supply chain in a remanufacturing environment: A case study of proton-exchange membrane fuel cell battery in Riyadh, *Advances in Mechanical Engineering*, 11, 4, 1–14
- S.G. Baratsas, E.N. Pistikopoulos, S. Avraamidou, 2021, A systems engineering framework for the optimization of food supply chains under circular economy considerations, *Science of the Total Environment*, Elsevier B.V., 794, -, 148726
- N.K. Dev, R. Shankar, F.H. Qaiser, 2020, Industry 4.0 and circular economy: Operational excellence for sustainable reverse supply chain performance, *Resources, Conservation and Recycling*, Elsevier, 153, -, 104583
- J. Korhonen, C. Nuur, A. Feldmann, S. Eshetu, 2018, Circular economy as an essentially contested concept, 175, -, 544-552
- J.T. Mentzer, J.S. Keebler, N.W. Nix, C.D. Smith, and Z.G. Zacharia, 2001, *Journal Of Business Logistics*, 22, 2, 1–25
- N. Nami, S.M. Hosseini-Motlagh, 2022, Central robust decision-making structure for reverse supply chain: a real pharmaceutical case, *Comp. and Ind. Eng.*, Elsevier Ltd, 173, -, 108726
- R. Pal, E. Sandberg, M.K. Paras, 2019, Multidimensional value creation through different reverse supply chain relationships in used clothing sector, *Supply chain management*, 24,6, 729-747.
- A. Rentizelas, N.L. Trivyza, 2022, Enhancing circularity in the car sharing industry: Reverse supply chain network design optimisation for reusable car frames, *Sustainable Production and Consumption*, 32, -, 863–879
- K.G. Stylianopoulou, E.M. Kondili, J.K. Kaldellis, 2022, Process Systems Engineering prospects in Circular Economy implementation in industry, in Montastruc, L. and Negny, S. (Eds.), 32nd European Symposium on Computer Aided Process Engineering, 51, -, 1309–1314
- G. Yadav, S.K. Mangla, A. Bhattacharya, S. Luthra, 2020, Exploring indicators of circular economy adoption framework through a hybrid decision support approach, *Journal of Cleaner Production*, 277, -, 24186
- R. Pal, E. Sandberg, M.K. Paras, 2019, Multidimensional value creation through different reverse supply chain relationships in used clothing sector, *Supply chain management*, 24,6, 729-747.

# Modular vs centralized manufacturing supply chain: identifying the best solution under uncertainty

Alessandro Di Pretoro<sup>a</sup>, Stéphane Negny<sup>a</sup>, Ludovic Montastruc<sup>a\*</sup>

*<sup>a</sup>Laboratoire de Génie Chimique, Université de Toulouse, CNRS/INP, Toulouse, France*  
[alessandro.dipretoro@ensiacet.fr](mailto:alessandro.dipretoro@ensiacet.fr)

## Abstract

Although the economy of scale has been the production systems golden standard for investment costs minimization during the last century, the modular manufacturing supply has recently seen a renewed interest during the last decades. This approach allows to considerably reduce transportation and operating expenses based on small capacity units that can be installed on site and that, after their usage, can be transferred to a new customer. Moreover, when uncertain variables are involved, the optimal choice is not that trivial anymore and a quantitative approach is required to have a clear and conscious view of the best supply chain design solution for the specific application. For this purpose, a dedicated case study concerning production and supply of a multicomponent resin was set up. In the centralized scenario the entire production process is carried out in the factory and the final product is delivered to the customers while, in the distributed one, the raw materials are provided to the final user and the production process takes place in the customers' site. Transportation costs and optimal centralized factory location are then calculated by means of an adaptation of the methodology proposed by Di Pretoro et al. (2022). Customers position and demand are randomly generated within the expected range of variable fluctuation. The uncertainty is addressed by means of a stochastic methodology and the obtained results show that it is possible to generate mono- and multi-variate probability distributions to describe the optimal centralized factory location as well as capital and operating expenses. This study proves that proposed methodology achieves its goal to identify the best solution between modular and centralized supply layout. Furthermore, it allows to quantify the expected costs for given sets of known and aleatory parameters. The study is then surely worth a deeper investigation for more complex and detailed case studies to further improve the quality of the outlined procedure.

**Keywords:** modular manufacturing, centralized manufacturing, supply chain, uncertainty, stochastic optimization.

## 1. Introduction

The transition process towards sustainable raw materials and renewable energy sources as well as the fast market dynamics and the volatility of goods price imply the spreading of uncertainty over all the sectors of the industrial domain. In particular, from a logistic perspective, customers are more and more frequently changing suppliers according to their particular needs. This trend results in non-negligible difficulties during the supply chain design in terms of uncertain demand and customers location. Moreover, the optimization related to the less expensive delivery layout and the optimal travel paths to meet the European standards (EEA, 2022) in terms of minimal emissions is compromised

in terms of optimal solution accuracy whether the uncertainty is not taken into account since the early design phases. Therefore, the need for more flexible solution resulted in a renewed interest towards the modular unit approach (Baldea et al., 2017). With respect to the conventional centralized approach, that exploits the economy of scale to minimize investments, the modular approach is based on units (skid) with a limited capacity but that can be used to perform a part or the entire process on the customer site. Moreover, these units can be transferred to another customer when the contract with the previous one expires.

In this research work we address the comparison between centralized and modular approaches under uncertain customers location and demand. The selected stochastic method for this study is the Sample Average Approximation. This method allows to decouple the uncertainty from the system model at the expenses of the computational costs, that is a negligible drawback given the size of the system under analysis. Further details about the case study are provided in the following section.

## 2. Case study

The case study presented in this work is based on a supply chain case study for a chemical product that is made up of two components, namely A and B, and water. The geographic domain of interest is given by a rectangle whose sizes are 1500 X 2000 km respectively. A set of 20 customers with a demand of either 1 or 2 kton/y with random location are generated for each simulation. They are modeled as “new” customers, each of those appears every year and stays until the end of the simulation duration.

Two scenarios for the logistics approach are then analyzed. The first one is based on a centralized distribution strategy for which the optimal location of the centralized factory is obtained by minimizing the Total Annualized Cost function (TAC) described in detail in the next section. An illustrative example is shown in Figure 1.

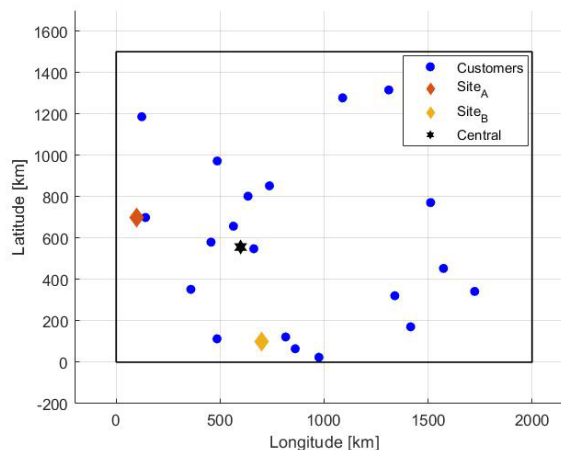


Figure 1 – Raw materials sites, centralized factory location and aleatory customers distribution

The second scenario is based on a modular approach. In this case water does not need to be delivered to the customers since water available on-site is used. As a consequence the delivery truck has a higher content of raw materials and the amount of required travels is way lower. Investment costs are given by the sum of the 1 or 2 kton skid units required to fulfill each customer demand according to the correlations of section 3.2.

Product composition, raw material supply centers location and other operating variables are listed in Table 1.

System parameters	Value	Unit	Abbreviation
Site A location	100 x 700	km x km	$x_a \times y_a$
Site B location	700 x 100	km x km	$x_a \times y_a$
Truck capacity	20	ton	$Cap_T$
<b>Composition</b>			
A fraction	0.3	kg/kg	$w_A$
B fraction	0.4	kg/kg	$w_B$
Water fraction	0.3	kg/kg	$w_W$

Table 1 – System operating parameters

Next section discusses in detail the methodology concerning optimization criteria and costs calculation previously introduced.

### 3. Methodology

The approach developed in this research work is a bi-level optimization based on a combination between Sample Average Approximation for the upper level and deterministic optimization for the lower one. The first method one is used to model aleatory customers, appearing and leaving during time in different geographical sites with different demand. The second one is used to find the optimal solution for a given scenario in terms of delivery paths and centralized factory location.

#### 3.1. Objective function

The lower level optimization loop has the purpose to define the shortest travel path and the corresponding transport costs. is based on the cost function minimization. The objective function to minimize is defined as:

$$F_{obj}[\$/y] = d[km/y] \cdot \hat{C}_{tr}[\€/km] \quad (1)$$

where the considered specific transportation costs  $\hat{C}_{tr}$  are equal to 1.2 €/km price and the travel distance is given, for the centralized scenario, by the sum of the terms:

$$d[km/y] = n_{A,C} \cdot dist_{A,C} + n_{B,C} \cdot dist_{B,C} + \sum_{i=1}^{n_{cust}} n_{C,i} \cdot dist_{C,i} \quad (2)$$

The number of amount of travels  $n_{ji}$  from the A, B and centralized site respectively are based on the ratio between demand composition and truck capacity as follows:

$$n_{A,C}[1/y] = \frac{w_A[kg/kg] \cdot dem_{tot}[kg/y]}{cap_{truck}[kg]} \quad (3)$$

where  $dem_{tot}$  is the total yearly demand. Differently from the first scenario, for the distributed approach the two terms related to raw material sites should be expressed as the sum weighted with respect to their distance from each customers and the third term is not taken into account since there is no centralized distribution center.



### 3.2. Costs calculation

Once the lower level optimization is completed, operating expenses (OPEX) can be calculated and combined to capital expenses (CAPEX) in order to quantify the Total Annualized Costs according to the equation:

$$TAC[\$/y] = OPEX[\$/y] + CRF[1/y] \cdot CAPEX[\$] \quad (4)$$

The Capital Recovery Factor (CRF) is defined as:

$$CRF[1/y] = \frac{i \cdot (i + 1)^n}{(i + 1)^n - 1} \quad (5)$$

where the discount rate  $i$  and the plant lifetime  $n$  for this case study are respectively equal to and 20. For CAPEX evaluation the Cobb-Douglas correlation is used as follows to resize the process unit cost according to its capacity:

$$Cost_{Centr} = C_{skid2} \cdot \left( \frac{Cap_{Centr} + 8000}{2000} \right)^{0.8} \quad (6)$$

$$Cost_{Centr} = C_{skid2} \cdot \left( \frac{1000}{2000} \right)^{0.8} \quad (7)$$

where the cost of a 2 kton skid is 300 k€ and the centralized factory is oversized by 8 kton. Finally, the production costs, i.e. raw materials and duties, are neglected during the comparison since they are the same for both scenarios.

## 4. Results

The case study simulation was run 5000 times. This value allows to have a results that can be represented by sufficiently significative and smooth Probability Distribution Functions (hereafter PDF).

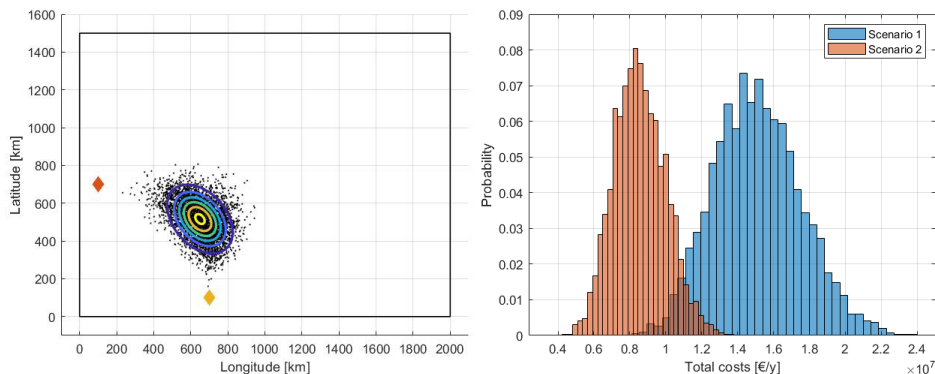


Figure 2 – left) Optimal centralized factory location and TAC distributions

As a first result, the 5000 positions for the optimal centralized factory location are outlined as shown in Figure 2. Based on the outcome, it is possible to regress the obtained data by means of a bi-variate probability distribution. This function could be then exploited to have a quantitative idea of the best places to locate the centralized distribution factory according to geographical (e.g. mountains, rivers etc.) and operational (e.g. electricity grid, duties or water sources availability etc.) constraints.

The second relevant result given by the TAC probability distribution. At a first glance, we can notice that, although the two distributions overlap over a small range, the modular scenario exhibits lower TAC with respect to the centralized one. Moreover, we can identify an average TAC value (without accounting for operating and raw materials cost for this example) equal about to 8.2 and 15.3 M€ per year respectively. The same analysis can be carried out concerning the standard deviation. Furthermore, if needed, the cumulative function could be derived in order to associate a given value of costs to the percentage of cases that fall outside the affordability threshold.

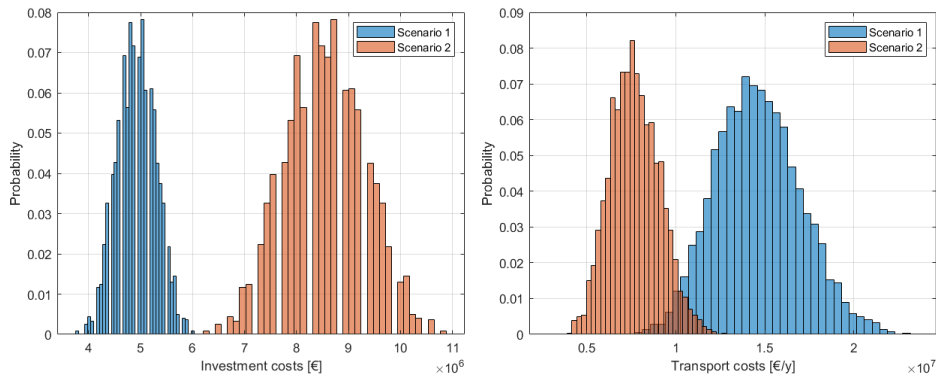


Figure 3 – left) Investment and right) operating costs distribution comparison

In order to analyze in deep the economic aspects, the TAC function could be decomposed into its parts. Figure 3 shows the outcome of this analysis for CAPEX (left) and transport costs (right).

As concern investment costs we can notice that they do not show a continuous trend. This is due to the fact that the customers demand could be either 1 or 2 kton per year resulting in a discretized distribution. In fact, it can be observed that, in terms of distribution shape, scenario 1 and 2 have the same trend. However, the modular approach results slightly less than twice as expensive as the centralized one as expected. On the other hand, in the transport costs plot the two distributions exhibit a different shape due to the fact that the travel paths are different in case of raw materials instead of final product distribution. Moreover, as it can be suddenly noticed as well, the magnitude of transport costs per year is comparable with the investments costs over the entire 20 years life of the plant with an average value of about 8.05 and 14.9 M€/y respectively.

For the sake of completeness, given that the truck travels are always performed at its full charge, the transport costs are proportional to the travel distance, i.e. with a scale factor represented by the ratio between specific fuel costs and specific fuel emissions, the equivalent CO<sub>2</sub> value (i.e. the GWP indicator) has the same trend as the plot in Figure 3 – right.

The obtained results and graphs allow then to conclude that, for this specific case study, due to the considerable impact of transport costs on the TAC, the modular solution results to be much more convenient than the conventional centralized one.

A more complete summary of the outcome of this research work is nevertheless provided in the next section.

## 5. Conclusions

The presented research work discusses a stochastic approach to optimize a supply chain system under uncertainty. The outcome of this study can be resumed in three main relevant points.

First, the proposed stochastic methodology based on Sample Average Approximation, although rather simple, allows to define a probability domain where the optimal centralized factory location should lie. Second, costs and emissions can be quantified both in terms of average value and standard deviation so that they can be compared for different scenarios in term of residual values associated to a given cost-emissions threshold and compromise. Finally, besides the results that are specifically related to the proposed case study, the comparison between a centralized and a modular industrial system configurations is viable. This aspect allows the decision maker to draw conscious conclusions based on thorough analysis and make the most suitable choice according to its specific needs.

As concerns further perspectives, the study could be extended to more complex systems in order to face the computational issue related to the fast growth of the amount of calculations with respect to the increase of the system parameters (e.g. number of customers, number of raw materials sites etc.). This aspect makes the research worth a deeper analysis in the future.

## References

- M. Baldea, T.F. Edgar, B.L. Stanley, A. Kiss, 2017, Modular Manufacturing Processes: Status, Challenges, and Opportunities. *AICHE Journal*, 63 (10), 4262–4272. <https://doi.org/10.1002/aic.15872>.
- A. Bhosekar, M. Ierapetritou, 2021, A Framework for Supply Chain Optimization for Modular Manufacturing with Production Feasibility Analysis. *Computers & Chemical Engineering*, 145, 107175. <https://doi.org/10.1016/j.compchemeng.2020.107175>.
- C.W. Cobb, P.H. Douglas, 1928, A Theory of Production. *American Economic Review*. 18 (Supplement): 139–165. JSTOR 1811556
- A. Di Pretoro, S. Negny, L. Montastruc, 2022, The Traveling Deliveryman Problem under Uncertainty: Fundamentals for Flexible Supply Chains. *Computers & Chemical Engineering*, 160, 107730. <https://doi.org/10.1016/j.compchemeng.2022.107730>.
- European Environment Agency, Greenhouse gas emissions from transport. Retrieved at <https://www.eea.europa.eu/ims/greenhouse-gas-emissions-from-transport>. Accessed on 27/11/2022.
- E. Alarcon-Gerbier, U. Buscher, 2022, Modular and mobile facility location problems: A systematic review. *Computers & Industrial Engineering* 173, 108734. <https://doi.org/10.1016/j.cie.2022.108734>

# Stochastic Optimization of Agrochemical Supply Chains with Risk Management

Saba Ghasemi Naraghi,<sup>a</sup> Zheyu Jiang<sup>a\*</sup>

<sup>a</sup>*Oklahoma State University, 420 Engineering North, Stillwater, Oklahoma, USA 74078*

*\*Corresponding author: zheyu.jiang@okstate.edu*

## Abstract

The global agrochemical market is highly consolidated, with large multinational companies accounting for a major share of the market. Thus, even for a single agrochemical product, its supply chain typically involves many possible paths connecting the raw material sources of active ingredients to final customers. In addition to structural complexity, agrochemical supply chains are also subject to seasonality and various unique uncertainties, thereby demanding high system resilience and implementation of risk management strategies in the face of these uncertainties and disruptions. In this study, we formulate and optimize the supply chain of an agrochemical active ingredient by formulating a stochastic mixed-integer nonlinear programming (MINLP) model. This MINLP formulation is scenario-based with demand uncertainty addressed by Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR). For the first time, we propose to reformulate these nonlinear CVaR constraints using perspective reformulation techniques. We show that these perspective cuts give a tight approximation of the original MINLP model. Through an illustrative case study, we compare the results and performance of the original MINLP and the reformulated MILP.

**Keywords:** Agrochemical supply chain, Value-at-Risk (VaR), Conditional Value-at-Risk (CVaR), perspective cuts, mixed-integer nonlinear programming.

## 1. Introduction

In 2050, the global population is expected to increase by 2 billion people to 9.7 billion, which puts unprecedented stress on food, energy, and water resources as the global food production must increase by at least 70% (Searchinger et al., 2018). Therefore, the manufacturing and supply chain of agrochemicals, including pesticides, herbicides, fungicides, and insecticides, are critical to ensuring food production and security. The global agrochemical market is highly consolidated, with 60-70% of the global market share dominated by four agrochemical companies alone (IEPS-Food, 2017). Each of these leading companies has a diversified product line, and its supply chains are multistage networks involving many possible paths connecting the raw material sources to active ingredients to final products. In addition, agrochemical supply chains are further complicated by seasonality and various uncertainties due to climate change, more frequent black swan events (e.g., COVID-19 pandemic), and increasingly complex geopolitical landscape (e.g., the Russia-Ukraine war). In particular, seasonal demand is a unique characteristic for agrochemical supply chains. To design cost-effective, resilient, and well-managed agrochemical supply chains, in this work, we develop an optimization

framework to effectively model the risks associated with these seasonality and uncertainties and propose a reformulation strategy to solve the optimization problem.

Among numerous recent works on supply chain optimization (Garcia and You, 2015), Bassett and Gardner (2009) presented one of the first mixed-integer linear programming (MILP) formulations for global agrochemical supply chain optimization considering seasonality and uncertainties in customer demand. Liu and Papageorgiou (2012) extended the agrochemical supply chain optimization framework by modeling and comparing different plant expansion strategies. To further ensure continuous use and inactivity of warehouses for continuous periods of time, Brunaud et al. (2017) developed dynamic contract policy constraints for warehouses and incorporated them to the agrochemical supply chain model. In terms of quantifying uncertainties and risks, You et al. (2009) proposed a scenario-based two-stage stochastic linear programming framework and decomposition strategies for mid-term planning of multi-product supply chain under demand and freight rate uncertainties. Later, Carneiro et al. (2010) focused on the oil supply chain optimization problem, in which they incorporated Conditional Value-at-Risk (CVaR) as a risk assessment measure that quantifies the tail risk in their investment portfolio. In this work, we develop a scenario-based two-stage mixed-integer nonlinear programming (MINLP) model for global agrochemical supply chain optimization and adopt the concepts of Value-at-Risk (VaR) and CVaR to quantify and control the risks associated with demand unfulfillment. Note that these risk measures are highly nonlinear. Therefore, we introduce perspective cuts to linearize the CVaR constraint and reformulate the MINLP model. Perspective cuts were first introduced by Frangioni and Gentile (2006), who showed that the convex envelope of the objective function containing semicontinuous variables in a general mixed-integer program (MIP) is the perspective function of MIP's continuous part. More recently, Bestuzheva et al. (2021) extended perspective cuts to nonlinear constraints in MINLPs. Consider a MINLP with a linear objective function  $f: \min f(x, y, z)$  subject to nonlinear constraints  $g(x) \leq 0$ , in which  $y \in \Omega$ ,  $z \in \{0,1\}^n$ , and  $x \in \mathbb{R}^n$  are semi-continuous variables (for every  $i = 1, \dots, n$ ,  $x_i = 0$  when  $z_i = 0$  and  $x_i \in [l, u]$  when  $z_i = 1$ ). The perspective cuts for linearizing nonlinear constraints  $g(x) \leq 0$  are given by:

$$\langle \nabla g(\bar{x}), x \rangle + [g(\bar{x}) - \langle \nabla g(\bar{x}), \bar{x} \rangle]z \leq 0 \quad (1)$$

where  $\bar{x} \in [l, u]$  is an arbitrary parameter. After replacing  $g(x) \leq 0$  with these perspective cuts, the MINLP is reformulated to a MILP, which can be solved iteratively. Specifically, starting from the second iteration,  $\bar{x}$  is chosen to be the solution of the MILP from the previous iteration. Bestuzheva et al. (2021) also conducted a detailed computational study of perspective reformulation for MINLPs with convex and nonconvex nonlinear constraints. They showed that the perspective reformulation of convex MINLPs provides much tighter approximation of the original problems compared to conventional branch-and-cut approaches, thereby leading to significant computational time reduction. Nevertheless, they also reported that adding perspective cuts for nonconvex MINLPs had less impact on computational time, although it reduces the size of branch-and-cut trees and strengthens the root node relaxation.

## 2. Problem Statement and Model Formulation

In this illustrative case study, we consider the three-echelon supply chain of one active ingredient (AI) produced in an agrochemical company involving five AI production

plants, four warehouses/distribution centers, and three market regions. We are interested in midterm planning (1 year) divided into 52 periods (i.e., 1 week per period). AI production plants and warehouses are connected by one or more transportation links, and so are warehouses and market regions. We allow different types of transportation modes for each transportation link. AI production plants can either be active or inactive during each time period. The manufacturing capacity of an AI production plant can be expanded at most once in a year. During the expansion period, AI production must be inactive. The AI produced can either be transported to warehouses or stored as inventory. The inventory level must be larger than or equal to the safety stock. As shown in Figure 1, the yearly demand of an agrochemical product typically follows a bimodal distribution (Bassett, 2018). For this case study, we consider three scenarios in customer demand. Also, we consider the risk of demand loss or unsatisfaction due to uncertainties related to production planning and warehouse capacity limitations. We are given the safety stock of an inventory, initial inventory, as well as unit costs associated with inventory holding, AI manufacturing and expansion, material transportation, warehouse storage and expansion, and demand loss or unsatisfaction (see Table 1). The objective function is to minimize the total cost of the supply chain. Due to space limitations, we only highlight some of the key points in our MINLP model:

1. Following You et al. (2009), decision variables of the first time period (Week 1) are first-stage variables and are independent of scenarios. Second-stage variables and scenario-based stochasticity begins at the second time period (Week 2).
2. We adopt the dynamic contract policy formulation from Brunaud et al. (2018) and extend it to AI production plants – Each AI production plant must remain in production for at least  $U$  time periods, after which it might undergo cleanup for  $F$  time periods, during which no production activities would take place:

$$\begin{aligned}
 & -\alpha_i + \alpha_i^{start} \geq 0, \quad \forall i \in I \\
 & \alpha_{i,t-1}^s - \alpha_{i,t}^s + \alpha_{i,t}^{s,start} \geq 0, \quad \forall i \in I, s \in S, t \in T(t > 1) \\
 & \alpha_i + \sum_{\tau=2}^R \alpha_{i,\tau}^s \geq U \alpha_i^{start}, \quad \forall i \in I, s \in S \\
 & \sum_{\tau=t}^{t+R-1} \alpha_{i,\tau}^s \geq U \alpha_{i,t}^{s,start}, \quad \forall i \in I, s \in S, t \in T(t > 1, t + R - 1 \leq |T|) \\
 & \alpha_{i,t+1}^s - \alpha_{i,t}^s + \alpha_{i,t}^{s,final} \geq 0, \quad \forall i \in I, s \in S, t \in T \\
 & \sum_{\tau=t+1}^{t+F} \alpha_{i,\tau}^s + F \alpha_{i,t}^{s,final} \leq F, \quad \forall i \in I, s \in S, t \in T(t + F \leq |T|)
 \end{aligned} \tag{2}$$

where  $\alpha_i$ ,  $\alpha_i^{start}$ ,  $\alpha_{i,t}^s$ ,  $\alpha_{i,t}^{s,start}$ ,  $\alpha_{i,t}^{s,final}$  are binary variables indicating whether the AI production plant  $i \in I$  is in production (1) or not (0) at time period 1, starting at time period 1, at time period  $t$  under scenario  $s$ , starting at time  $t$  under scenario  $s$ , and ending at time  $t$  under scenario  $s$  respectively.  $|T|$  is the total number of time periods.

3. The total demand loss  $\chi_t^s$  for time period  $t$  under scenario  $s$  is the sum of demand loss from all market regions. To quantify the risks due to demand loss, we introduce the CVaR constraint in the model. We assume that the demand loss follows a normal distribution, which enables us to express VaR in terms of  $\chi_t^s: \text{VaR}^s = \frac{\sigma^s z_\beta}{\sqrt{|T|}}$ , where  $z_\beta$

is z-score at confidence interval  $\beta$ ,  $|T|$  is the total number of time periods (i.e., 52), and  $\sigma^s = \sqrt{\frac{1}{|T|-1} (\sum_{t \in T} \chi_t^{s2} - \frac{(\sum_{t \in T} \chi_t^s)^2}{|T|})}$  is the standard deviation of the total demand loss as for scenario  $s$ . From VaR, we calculate CVaR as  $\frac{1}{1-\beta} \sum_{s \in S} p_s \text{VaR}^s$  following Carneiro et al. (2010), where  $p_s$  is the probability of scenario  $s$ . By specifying a lower bound  $P$  on CVaR (e.g., 30 mass units), we obtain the following CVaR constraint and add it into the formulation:

$$\text{CVaR} = \frac{z_\beta}{(1-\beta)\sqrt{|T|(|T|-1)}} \sum_{s \in S} p_s \sqrt{\left( \sum_{t \in T} \chi_t^{s2} - \frac{(\sum_{t \in T} \chi_t^s)^2}{|T|} \right)} \geq P \quad \forall s \in S, \quad (2)$$

- Since Equation (2) is nonconvex and nonlinear, we introduce perspective cuts and reformulate the original MINLP into a MILP by substituting Equation (2) to Equation (1):

$$\left( P - R \sum_{s \in S} p_s Q_s - R \sum_{s \in S} p_s \frac{\bar{\chi}_t^s (\bar{\chi}_t^s - \frac{\sum_{t \in T} \bar{\chi}_t^s}{|T|})}{Q_s} \right) z_t - R \sum_{s \in S} p_s \frac{\chi_t^s (\bar{\chi}_t^s - \frac{\sum_{t \in T} \bar{\chi}_t^s}{|T|})}{Q_s} \leq 0 \quad \forall t \in T, \quad (3)$$

where  $R := \frac{z_\beta}{(1-\beta)\sqrt{|T|(|T|-1)}}$ ,  $Q_s := \sqrt{\left( \sum_{t \in T} (\bar{\chi}_t^s)^2 - \frac{(\sum_{t \in T} \bar{\chi}_t^s)^2}{|T|} \right)}$ ,  $z_t \in \{0,1\}$ , and  $\bar{\chi}_t^s$  is the optimal demand loss from the previous iteration of introducing perspective cuts.

- We point out that this is the first linearization of CVaR constraint reported in the literature. Both problems are solved using SCIP v8.0 in GAMS 40.2.0 in a Dell Precision 7920 workstation with Intel Xeon Gold 6226R CPU @ 2.90 GHz and 96 GB of RAM, and the results are compared.

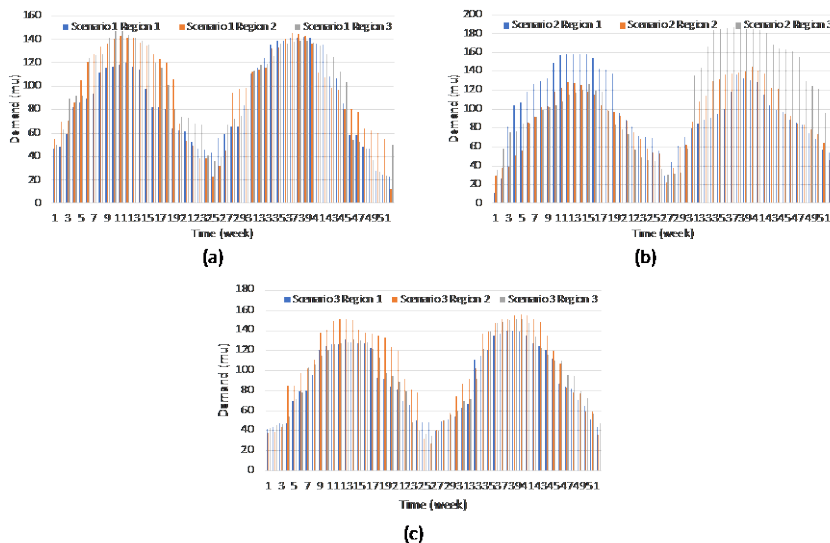


Figure 1. Bimodal demand curve for each market region under different scenarios.

Table 1. List of model parameters used in the illustrative case study.

Model parameters	Values
Initial capacity in each of the 5 AI production plants (mu)	100, 115, 100, 120, 110
Safety stock in each AI production plant (mu)	5, 5, 10, 10, 5
Initial inventory level in each AI production plant (mu)	5, 5, 15, 0, 15
Transportation capacity from AI production plant to warehouse (mu)	400 for mode 1 350 for mode 2
Transportation capacity from warehouse to market region (mu)	100 for mode 3 200 for mode 4
Storage capacity in each of the 4 warehouses (mu)	550, 650, 500, 450
Initial inventory level in each warehouse (mu)	25, 20, 30, 10
Probability of the three possible scenarios	0.25, 0.45, 0.30
Fixed cost in each of the 5 AI production plants ( $\times 10^3$ \$)	75, 100, 150, 100, 75
Variable cost in each of the 5 AI production plants (\$)	10, 15, 10, 5, 20
Expansion cost in each of the 5 AI production plants ( $\times 10^3$ \$)	10, 50, 25, 35, 15
Fixed transportation cost from AI production plant to warehouse (\$)	25 for mode 1 30 for mode 2
Fixed transportation cost from warehouse to market region (\$)	15 for mode 3 30 for mode 4
Variable transportation cost from AI production plant to warehouse (\$/mu)	2 for mode 1 2.5 for mode 2
Variable transportation cost from warehouse to market region (\$/mu)	1.5 for mode 3 2.5 for mode 4
Maximum total demand loss in each time period (mu)	15
Variable cost on demand loss (\$/mu) for each of the 3 market regions	50, 55, 70

### 3. Discussion

The original MINLP model, which contains 9255 continuous variables and 1874 binary variables, has 14499 linear constraints and 2 nonlinear constraints of Equation (2) (one equality constraint and one inequality constraint). After the first iteration of adding perspective cuts, the reformulated MILP contains 14551 linear constraints (14499 linear constraints and 52 perspective cuts, each corresponding to one time period), 1874 binary variables, and 9254 continuous variables. We specify a solving time of 150 seconds, at which the original MINLP model has an objective function value of  $\$1.509 \times 10^6$  and a gap of 2.39%, whereas the reformulated model shows an objective function value of  $\$1.498 \times 10^6$  and a closer gap of 1.64%. We emphasize that the reformulated model always yields a feasible solution in the original formulation, suggesting that it provides a better optimal solution. This is due to the fact that the reformulated model is able to identify a solution with less overall demand loss  $\sum_{s \in S} \sum_{t \in T} \chi_t^s$  compared to the solution obtained from the original MINLP. This suggests that the reformulated model produces a more efficient supply chain compared to the original model.

In a separate numerical study, we analyze the impact of problem size on the computational benefits of introducing perspective cuts. When we consider only 13 time periods (monthly) in the model, no significant computational speed improvement is observed in the reformulated model compared to the original MINLP. However, when considering 26 time periods (biweekly) in the model, the reformulated model solves to 1% gap in 29 seconds, whereas the original MINLP solves to the same objective function value as well



as the same 1% gap in 41 seconds. This observation is consistent with that made by Bestuzheva et al. (2021), in the sense that the computational benefits of perspective cuts depend on the size of the problem. Perspective cuts should be accompanied by other reformulation strategies and bound-tightening constraints to maximize their benefits.

#### 4. Conclusion

In this work, we optimize the supply chain of an agrochemical active ingredient by formulating and solving a scenario-based stochastic MINLP problem. The nonlinearity of the model comes from the CVaR constraint used to quantify risks associated with unforeseen demand loss. For the first time, we propose to reformulate the CVaR constraints using perspective reformulation techniques. The reformulated model, which is a MILP, always gives a feasible solution to the original MINLP model. Using a simple case study, we demonstrate the effectiveness of perspective cuts in fostering convergence and reducing computation time. In particular, we show that the reformulated model typically produces optimal solutions with less overall demand loss compared to the solution obtained from the original MINLP. Thus, adopting perspective reformulation could help identify a more efficient supply chain network and production/distribution plans with lower costs and carbon emissions. On the other hand, we point out that, depending on the problem size, adding perspective cuts for nonconvex MINLPs may not lead to significant computational time improvements. In this case, perspective cuts should be accompanied by other reformulation strategies and bound-tightening constraints to synergistically facilitate the solution of nonconvex MINLPs.

#### References

- M. Bassett, L. Gardner, 2009, Optimizing the design of global supply chains at Dow AgroSciences, *Computers & Chemical Engineering*, 34, 10, 254-265.
- M. Bassett, 2018, Optimizing the design of new and existing supply chains at Dow AgroSciences, *Computers & Chemical Engineering*, 114, 18, 191-200.
- K. Bestuzhev, A. Gleixner, S. Vigerske, 2021, A Computational Study of Perspective Cuts, preprint available at <https://arxiv.org/abs/2103.09573>
- B. Brunaud, M.H. Bassett, A. Agarwal, J.M. Wassik I.E Grossmann, 2017, Efficient formulation for dynamic warehouse location under discrete transportation costs, *Computers & Chemical Engineering*, 111, 18, 311-323.
- M.C. Carneiro, G.P Ribas, S. Hamacher, 2010, Risk management in supply chain: CVaR approach, *Industrial & Engineering Chemistry Research*, 49, 3286-3294.
- A. Frangioni, C. Gentile, 2005, Perspective cuts for a class of convex 0-1 mixed integer programs, *Mathematical Programming*, 106, 225-236.
- D.J. Garcia, F. You, 2015, Supply chain design and optimization: Challenges and opportunities, *Computers & Chemical Engineering*, 81, 15, 153-170.
- IEPS-Food, 2017, Too big to feed: Exploring the impacts of mega-mergers, concentration, concentration of power in the agri-food sector, accessed via [http://www.ipcs-food.org/images/Reports/Concentration\\_FullReport.pdf](http://www.ipcs-food.org/images/Reports/Concentration_FullReport.pdf)

- S. Lui, L.G. Papageorgiou, 2012, Multiobjective optimization of production, distribution and capacity planning of global supply chains in the process industry, *Omega*, 41,13, 369-382.
- T. Searchinger, R. Waite, J. Ranganathan, P. Dumas, E. Matthews, 2018, World Resources Report: Creating a Sustainable Food Future, World Resources Institute.
- F. You, J.M Wassik, I.E Grossmann, 2009, Risk management for a global supply chainplanning under uncertainty: Models and algorithms, *AIChE Journal*, 55, 4, 931-946.



# Supply Chain Optimization Considering Disruption Demand Uncertainty

Oluwadare Badejo <sup>a</sup>, Marianthi Ierapetritou <sup>a</sup>

<sup>a</sup> *Department of Chemical and Biomolecular Engineering,  
University of Delaware, Newark, 19711, USA,*

## Abstract

In this paper, a mathematical framework was developed to optimize supply chains under disruption and demand uncertainty. A four-tier supply chain network is used as a directed graph with nodes representing supply chain entities and arcs representing the transportation system. The model was a two-stage stochastic optimization, aimed at minimizing costs while considering alternative sourcing, inventory levels, capacity flexibility, and demand uncertainty. The stochastic model was compared with a deterministic model using metrics such as total cost, service level, and cost of production. The results showed that the stochastic model made better decisions, achieved higher service levels, and reduced total costs in the long run. The model was also used in a rolling horizon framework, and the recommended actions for certain demand periods reduced costs and increased service levels. Future directions include exploring ways to model policies that incorporate risk averseness.

**Keywords:** Supply chain disruption, stochastic programming, rolling horizon

## 1. Introduction and Literature Review

The global supply chain network has experienced unprecedented challenges in recent times, such as the COVID-19 pandemic, geopolitical tensions, and economic uncertainties, leading to substantial disruptions. The ramifications of these disruptions on supply chain operations are widespread, affecting various industries and economies, with a majority of Fortune 1000 companies reporting significant supply chain disruptions (Supply Chain Disruption & How to Respond | Accenture, 2022). These disruptions not only result in revenue loss, but also a tarnished reputation, potentially causing customer loss to competitors (Ivanov et al., 2017; Shekarian and Mellat Parast, 2021). Enterprises are struggling with the management of these disruptions while also preparing for future recovery.

Supply chain disruptions and operational uncertainty present significant obstacles in enterprises' operations. These disruptions stem from unplanned, unexpected events such as man-made or natural disasters, disrupting the existing supply chain network topology (Badejo and Ierapetritou, 2022; Dolgui and Ivanov, 2021). Operational uncertainty is characterized by unpredictability in demand, supply, and other operational variables, hindering organizations' ability to make accurate forecasts and leading to inefficiencies. The operational uncertainty is well addressed in the literature (Bhosekar et al., 2021; Charitopoulos et al., 2018; Guillén-Gosálbez and Grossmann, 2009). The coexistence of supply chain disruptions and operational uncertainty presents a challenging situation for enterprises in managing their supply chains effectively. While previous studies have

explored these uncertainties independently, the interaction between them has not been well studied. To address these challenges, this paper proposes a stochastic optimization framework for supply chain management, aiming at simultaneously mitigating the impact of disruptions and reducing the effects of operational uncertainty.

## 2. Methodology

### 2.1. Problem Description and Assumptions

The following assumptions were made regarding the disruptions and operational uncertainties. Known distributions were assumed for operational parameters, with a normal distribution for demand uncertainty and a uniform distribution for the cost parameters. Supply chain entities have two states: normal and disrupted, with normal entities functional and disrupted ones not.

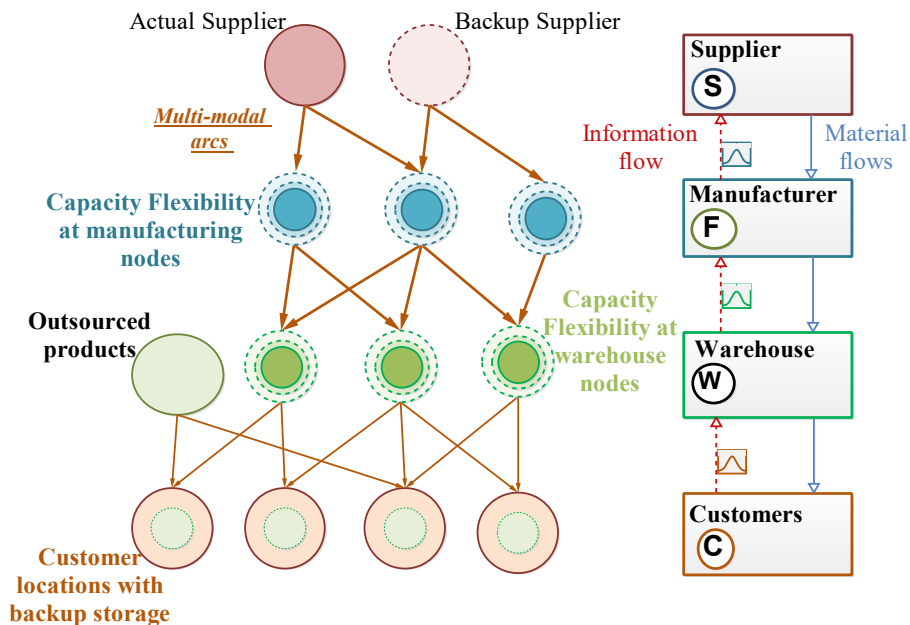


Figure 1: Supply chain topology

The problem considers a multi-products customer-driven supply chain network. This network produces a variety of products to meet the needs of different customer zones. The supply chain network consists of four echelons: supplier nodes  $s$ , manufacturing facilities  $f$ , warehouses  $w$ , and customer zones  $c$  as shown in Figure 1. A discrete time representation of the time horizon is considered. Raw material ordering, product manufacturing, storage, and shipping are executed to fulfill customer demand with unsatisfied demand resulting in lost sales and backorder penalty cost. The supply chain entities are vulnerable to disruptions, but buffer mechanisms are employed to mitigate their impact. The optimization problem aims to balance customer service level and minimize operating cost in a tactical supply chain problem under demand uncertainty and disruptions.

### 2.2. Mathematical Formulation

The objective of the problem is to minimize the total cost associated with the supply chain operations through optimization of raw materials and product flows across the network,

subject to constraints on the capacities of the nodes and arcs, continuity balance and demand satisfactions. This is shown in equation (1), where the cost associated with the operations at each node includes the cost of normal operations, as well as recovery costs in case of disruptions. The cost of flow across the arcs is captured by the multi-modal transportation options, each incurring a different cost.

$$\min\{ \mathbb{E}(\text{nodeCost}(\theta) + \text{arcCost}(\theta)) \} \quad (1)$$

The constraints are augmented by incorporating integer variables to account for supplier selection, facility capacity utilization during disruptions, as expressed in equations (2a) to (2e).

$$y_{n,t}^l - y_{n',t}^l \geq 0 \quad \forall n \in \{S_a^n, W_a^n\}; n' \in S_b \quad (2a)$$

$$y_{n,t}^l - y_{n',t}^{l'} \geq 0 \quad \forall l < l', n \in \{w, f\}, t \in T \quad (2b)$$

$$y_{nt}^{l=1} := \begin{cases} 1 & \forall n \in (w, f) | n \in N_a^n, t \in T \\ 0 & \forall n \in (w, f) | n \in N_a^d, t \in T, t < t_R \end{cases} \quad (2c)$$

$$\sum_{\substack{i \in \{r,p\} \\ n',m}} Q_{i,n,n',m,t}(\theta) \leq \sum_l y_{n,t}^l \times \text{cap}_{n^l} \quad \forall n \in (s, f); t \in T; (n, n') \in \{(s, f), (f, w)\} \quad (2d)$$

$$\sum_{i \in \{r,p\}} Q_{i,n,n',m,t}(\theta) \leq x_{m,t}^{n,n'} \times t\text{Cap}_m^{n,n'} \quad \forall (n, n') \in \{(s, f), (f, w), (w, c)\} \quad (2e)$$

The precedence constraints are established by Equations (2a) to (2c) where the binary variables  $y_{n,t}^l$  indicate the operation of capacity level  $l$  at node  $n$  in time period  $t$ . Equation (2a) determines the priority of regular nodes  $n$  over backup nodes  $n'$ , and equation (2b) and (2c) outlines the ordered selection of capacity levels. Once the capacity level is determined, equation (2d) assigns the corresponding activity level. Additionally,  $Q_{(i,n,n',m,t)}(\theta)$  which represents the flow of commodity  $i$  from through arc connecting nodes  $n$  and  $n'$  is bounded to the capacity constraints of the transportation modes, as expressed in Equation (2e).

At the warehouse nodes, constraints (3a) – (3b) for inventory management are used. Equation (3a) shows that the minimum inventory level is proportional to the standard deviation of the products and replenishment lead time, while equation (3b) limits the lower bound of the inventory to the safety stock. This equation assumes independent and identically distributed product demand (Brunaud et al., 2019). The cumulative normal distribution coefficient,  $z$ , is used to determine the required service level. In this study, a service level of 95% was assumed with a value of 1.65 for  $z$ .

$$I_{i,n,t}^{SS} = z \times \sqrt{L} \times \sigma_p \quad \forall i \in p; n \in w \in W_{\{a\}}^n; t \in T \quad (3a)$$

$$I_{i,n,t}(\theta) \geq I_{i,n,t}^{SS} \quad \forall i \in p; n \in w \in W_{\{a\}}^n; t \in T \quad (3b)$$

The proposed model adopts a multi-period mixed integer linear programming (MILP) formulation. The product demands are characterized by a temporal discretization, with a segment of certain demands and a segment forecasted from a distribution described by parameter  $\theta$ . A two-stage stochastic model is adopted, as shown in equation (4): the first

stage decisions  $x$ , represents the decisions at the time where information is certain and include the node configuration, the flow between the nodes to satisfy the product demands, and the inventory level at the warehouse to hedge against future demand variability, while in the second stage decisions  $y(\theta)$  are adjusted with respect to the uncertainty realized. These decisions include the flow across adjacent nodes for all possible scenarios of the uncertain period, and the inventory policies to be adopted for all scenarios.

$$\min \left\{ c^T x + \mathbb{E}[d^T y(\theta)] \mid \begin{array}{l} \mathbb{C}_i(x, y(\theta), p(\theta)) \leq 0 \\ x, y(\theta) \in \mathbb{X} \\ x \in \{\mathbb{R}_+ \times \mathbb{Z}\}, y(\theta) \in \mathbb{R} \end{array} \right\} \quad (4)$$

To compare with a deterministic model, the expected value of the uncertain demands is used as value for the uncertain period. Thus, the deterministic formulation is shown in Equation (5). In this case we have decisions across time periods but implement only the decisions in the certain time segment.

$$\min \left\{ c^T x \mid \begin{array}{l} \mathbb{C}_i(x, \mathbb{E}[p(\theta)]) \leq 0 \\ x \in \mathbb{X} \\ x \in \{\mathbb{R}_+ \times \mathbb{Z}\} \end{array} \right\} \quad (5)$$

To address the dynamic and uncertain nature of supply chain operations, it employs a rolling horizon optimization approach. This approach involves solving the problem iteratively over multiple periods, with only the first stage being implemented. The model leverages the current state of the supply chain, including inventory levels and capacity utilization, as an initial condition for updating the solution in response to new information.

### 3. Results and Discussion

In this section a four-echelon supply chain with three products manufactured using two raw materials and six available suppliers (four actual and two backup) is studied. The enterprise operates four manufacturing facilities, two warehouses and supplies to five customer zones. Additionally, there is capacity for product storage at the customer locations. A monthly demand forecast with known demand for one month and forecasting for the next four is made to hedge against uncertainty. The goal is to minimize the total supply chain operation cost for a given period, including optimal tactical decisions during disruptions. In this case the disruptions considered are the breakdown of some arcs and the disrupted manufacturing facility. Decisions made include flow of materials between nodes, production amount, inventory, product delivery, outsourcing, and unmet demand.

Table 1: Metrics to compare the deterministic and stochastic solution.

Metrics	Deterministic	Stochastic
Total Cost [\$]	200,498	235,659
Implemented Cost [\$]	34,696.1	31,215.4
Service Level	80%	98.7%
Cost Per Period [\$/unit]	65.60	47.84
SC Efficiency	71%	89.65%

The proposed model was evaluated on a PC with Intel Core i7-10510U, 2.30 GHz, and 16GB of RAM. The results of the two-stage model, which considers the disruptions and operational uncertainty, were compared to the deterministic model. 20 demand scenarios were sampled for each product over 5 time periods, with the first period being certain. The deterministic model utilized the expected values of the scenarios, while the two-stage model considered all scenarios. Both models were solved in GAMS/CPLEX (v 38.2.1).

The deterministic model had 6851 constraints, 8809 continuous variables, and 2262 binary variables, while the two-stage model had 63,781 constraints, 67,139 continuous variables, and 2262 discrete variables. The deterministic model was solved in 25s, and the two-stage model took 260s.

Comparative analysis between the two-stage stochastic model and deterministic model is presented in Table 1, where the total cost and implemented cost for the two-stage model are higher. This results in a higher service level and supply chain efficiency due to integer decisions for node and arc selection and utilization as shown in Figure 2 and Figure 3, respectively.

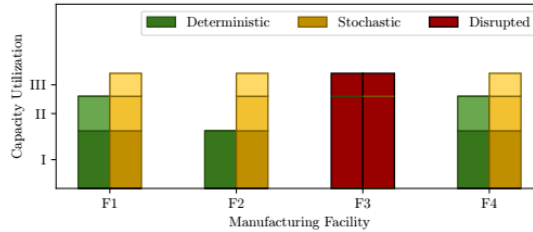


Figure 2: Manufacturing site capacity selections.

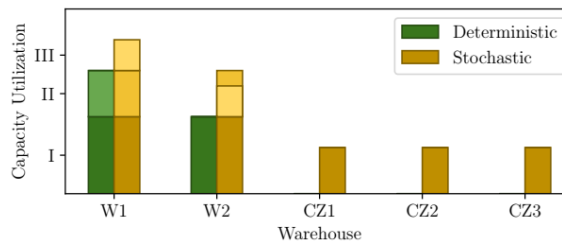


Figure 3: Warehouse capacity selections.

Table 2 provides the breakdown of implemented cost. The results from the stochastic model indicate a preference for higher utilization of manufacturing sites and warehouses, compared to the deterministic model, in order to minimize costs associated with node operations and uncertainty mitigation. This is evident from the higher values for service level and supply chain efficiency in the stochastic model. Conversely, the deterministic model optimizes only for certain demands and the average of all scenarios, resulting in increased backorder costs.

Table 2: Breakdown of implemented cost for the deterministic and two-stage stochastic model.

	Implemented Cost [\$]		
	Deterministic	Stochastic	Difference
<b>Supplier Cost</b>	1,763.49	2,664.03	900.54
<b>Facility</b>	10,989.70	16,287.30	5,297.60
<b>Outsourcing</b>	3,222.08	3,222.08	0
<b>Inventory</b>	228.44	1,044.47	816.03
<b>Transportation</b>	4,612.30	7,079.21	2466.91
<b>Backorder cost</b>	13,880.10	918.37	-12,961.73

#### 4. Conclusions

In this study, a resilient supply chain network model is presented that considers both supply chain disruptions and operational uncertainty. The objective is to minimize the



total cost of the supply chain while ensuring demand fulfillment. A deterministic multi-period model and a two-stage stochastic model are compared, and the results indicate that the stochastic model performs better in terms of service level and decision-making for uncertainty. The rolling horizon approach is used to evaluate the spatial-temporal decisions made by both models and the results show that the stochastic model has better performance. However, incorporating risk measures into the stochastic model can provide risk-averse decisions and this will be explored in future work.

## 5. Acknowledgements

The authors are grateful for the financial support provided by the National Science Foundation through awards 2134471, OIA-2119754, and 2217472. These grants have greatly contributed to the successful completion of the research project.

## References

- Badejo O, Ierapetritou M. Mathematical Programming Approach to Optimize Tactical and Operational Supply Chain Decisions under Disruptions. *Ind Eng Chem Res* 2022. <https://doi.org/10.1021/acs.iecr.2c01641>.
- Bhosekar A, Badejo O, Ierapetritou M. Modular supply chain optimization considering demand uncertainty to manage risk. *AIChE Journal* 2021;n/a:e17367. <https://doi.org/10.1002/aic.17367>.
- Brunaud B, Láinez-Aguirre JM, Pinto JM, Grossmann IE. Inventory policies and safety stock optimization for supply chain planning. *AIChE Journal* 2019;65:99–112. <https://doi.org/10.1002/aic.16421>.
- Charitopoulos VM, Aguirre AM, Papageorgiou LG, Dua V. Uncertainty aware integration of planning, scheduling and multi-parametric control. In: Eden MR, Ierapetritou MG, Towler GP, editors. *Computer Aided Chemical Engineering*, vol. 44, Elsevier; 2018, p. 1171–6. <https://doi.org/10.1016/B978-0-444-64241-7.50190-7>.
- Dolgui A, Ivanov D. Ripple effect and supply chain disruption management: new trends and research directions. *International Journal of Production Research* 2021;59:102–9. <https://doi.org/10.1080/00207543.2021.1840148>.
- Guillén-Gosálbez G, Grossmann IE. Optimal design and planning of sustainable chemical supply chains under uncertainty. *AIChE Journal* 2009;55:99–121. <https://doi.org/10.1002/aic.11662>.
- Ivanov D, Dolgui A, Sokolov B, Ivanova M. Literature review on disruption recovery in the supply chain. *International Journal of Production Research* 2017;55:6158–74. <https://doi.org/10.1080/00207543.2017.1330572>.
- Shekarian M, Mellat Parast M. An Integrative approach to supply chain disruption risk and resilience management: a literature review. *International Journal of Logistics Research and Applications* 2021;24:427–55. <https://doi.org/10.1080/13675567.2020.1763935>.
- Supply Chain Disruption & How to Respond | Accenture. 2022. <https://www.accenture.com/us-en/insights/consulting/coronavirus-supply-chain-disruption> (accessed February 6, 2023).

# Integrating economic, environmental, and social concerns into the design and planning of supply chains using monetization strategies

Cátia da Silva<sup>a\*</sup>, Ana Paula Barbosa-Póvoa<sup>a</sup>, Ana Carvalho<sup>a</sup>

<sup>a</sup>*CEG-IST, Instituto Superior Técnico, University of Lisbon, Av. Rovisco Pais, 1049-001 Lisboa, Portugal*

*catia.silva@tecnico.ulisboa.pt*

## Abstract

The growing concerns with sustainable aspects have influenced supply chain decision-makers to incorporate them into their efficient management. Thus, the decision-making process, which is traditionally based on economic profit, today also aims to take into account the environmental and social impacts of the supply chain so that decision-making is completer and more efficient, although complex. This work contributes to help decision-makers to take an informed decision and develops a decision support tool based on a mixed integer linear programming (MILP) model for the design and planning of supply chain that accounts for economic, environmental, and social performances in the same monetary unit. Environmental performance is considered in monetary units using the life cycle assessment methodology by applying the Environmental Priorities Strategies (EPS) as the life cycle assessment method. Social performance is measured through a social monetization approach based on job creation indicator (translated into a monetary units). In addition, this paper accounts for uncertainty in products' demand. Conclusions are drawn on how monetization can be a useful tool to support decision-making process and support supply chain's sustainable performances. A real-based case study is solved showing the applicability of the tool.

**Keywords:** sustainable supply chain, sustainability, monetization, uncertainty

## 1. Introduction

Sustainability is a relevant topic in today's world, having gradually grown in recent years. For this reason, supply chain management has also started to consider sustainability as a relevant aspect and a reason that allows supply chains to be competitive in the market. In this way, the supply chain previous focus on economic performance was expanded to include environmental and social performances. In fact, the importance of the overall supply chain performance reflects all three pillars of sustainability, namely the economic, environmental, and social pillars. Moreover, scientific papers and academic research increasingly highlight the need to answer to both environmental and social challenges combined with an improvement in the economic performance (Barbosa-Póvoa, da Silva, and Carvalho 2018). However, efficient supply chain management is not an easy task as supply chains are complex systems, and its complexity increases when the three

performances are taken into account. Additionally, the evaluation of these performances can be challenging, particularly when considering both environmental and social pillars. Economic performance is known for years as several authors intend to quantify it using optimization models to reduce costs and improve financial profits. However, the quantification of environmental and social performances is yet difficult, as there are different valuation methods to assess them, and they are usually expressed in abstract units that are hard to understand by decision-makers. Thus, it can be important to realize how environmental and social supply chains' performances can be assessed and they can be translated in monetary units easily understood by decision-makers (Beske-Janssen, Johnson, and Schaltegger 2015).

## **2. Environmental and social monetization methodologies**

Regarding environmental performance, life cycle assessment (LCA) has been considered the best methodology to assess environmental impacts related to a specific product or process (European Commission, 2013). This methodology includes four main steps: goal and scope definition, inventory analysis, life cycle impact assessment (LCIA), and life cycle interpretation. The most critical phase is LCIA, where decision-makers can assign weighting factors for objectives, which leads to subjectivity in results. In addition, there are several different LCIA methods that can be chosen to assess environmental impacts, particularly monetary and non-monetary methods. Monetary methods are the ones that can translate environmental impacts into monetary units, while others convert impacts into abstract or dimensionless units. In this work, Environmental Priorities Strategies (EPS) is the LCIA method chosen to convert impacts into monetary units.

There are several social indicators within the literature (Popovic et al., 2018), which are usually represented in non-monetary units (dimensionless units). Considering the need to better understand social impacts, this work intends to convert job creation indicator in monetary units so as decision-makers can really understand the real value of supply chains' social performance. To convert social indicators into monetary units, a correlation between economic indicators and supply chain variables is developed.

## **3. Problem description and model characterization**

This study considers the design and planning of a sustainable supply chain, which is generically represented in Figure 1. The main flows are as follows: i) raw materials flow from suppliers to factories; ii) final products go from factories to warehouses or directly to markets; iii) end-of-life products can be recovered from the markets and sent to warehouses or directly to factories to be reprocessed. All these flows are guaranteed by road, air, or sea transport. Given: i) a possible set of entity locations; ii) production and remanufacturing technologies; iii) possible transport modes between entities; and iv) SC's products. The objective is to obtain: i) the structure of the SC network; ii) supply and purchase levels; iii) the capacities of the entities; iv) the transport network; v) production, remanufacturing, and storage levels; and vi) supply flows and the recovery levels of the materials. In order to: maximize profit, maximize social impact and minimize environmental impacts while guaranteeing the establish minimum market requirements.

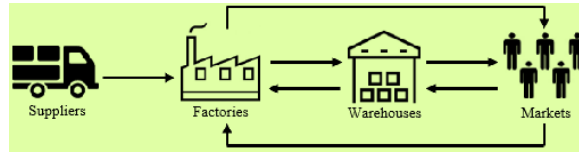


Figure 1. Scheme of SC network

To solve the problem on hands it is used a MILP model based on da Silva et al. (2020). The model was extended to consider the impact of supply chain social performance. The economic, social, and environmental performances are considered independently and simultaneously in the objective function. The economic objective is obtained through Eq. (1) that maximizes the Net Present Value (NPV), which considers the discounted cash flows (CF) in each period at a given interest rate. For the last time period, it is also considered the salvage values of the SC (FCI<sub>γ</sub>). The environmental impact is obtained using the EPS method. LCA is performed on the transportation activities and on entities installed in the SC boundaries. The Life Cycle Inventory is retrieved from the Ecoinvent database (assessed through SimaPro 9.4.0 software). From these results, it is determined the environmental impact of transportation and entities installation. The result values for transportation are used as input data in the first term of Eq. (2), while the values for entities installation are used in the second term. Moreover, considering that the focus of European Commission is to bet on promoting job creation and regional development, the social performance based on that is represented by Eq. (3). The monetarized social impact of job creation is obtained through the relationship between a macroeconomic indicator (i.e., GDP per capita) and the microeconomic variable (i.e., number of employees). The promotion of jobs in regions with higher unemployment rate,  $\lambda_n$  is also taken into account. Finally, Eq. (4) reflects the objective function that is obtained through maximization of the equilibrium between economic (NPV), environmental (EnvImpact) and social (SocImpact) impacts.

$$\max NPV = \sum_{t \in T} \frac{CF_t}{(1+ir)^t} - \sum_{\gamma} FCI_{\gamma} \quad (1)$$

$$\min EnvImpact = \sum_{(a,m,i,j) \in NetP} \sum_{t \in T} e_{iac} p w_m d_{ij} X_{mai} j t + \sum_{i \in I_f \cup I_w} e_{ic} Y C_i \quad (2)$$

$$\max SocImpact = \sum_{\substack{t \in T \\ n \in N \\ m \in M}} \frac{GDP_n^{pc} \cdot NEmployees_{nmct} \cdot \lambda_n}{(1+intr)^t} \quad (3)$$

$$\max NPV - EnvImpact + SocImpact \quad (4)$$

#### 4. Case study

The model is applied to a chemical components' company located in Lyon, France (Silva et al., 2020). The main suppliers are also placed in Lyon. The company supplies three markets that are located in different countries: Portugal, France, and Germany. France is the market that owns the highest percentage of company's sales, followed by Germany, and finally Portugal. Considering the willingness of the company's decision makers that want to expand its supply chain to three new markets (Ireland, Spain, and Canada), the company wants to know what changes could result from this expansion in financial terms. Currently, transportation is only performed by road mode. However, with the expansion there is the need of combining road transportation with air and sea transportation. Regarding company's portfolio, it sells three different types of chemical products (fp1, fp2, fp3) that can be sold within the chemical industry and to other industries. In the current production, three technologies (pr1, pr2, pr3) are used that produce respectively products fp1, fp2, and fp3. Furthermore, end-of-life products can be recovered and remanufactured into final products. This work accounts for product's demand uncertainty through a scenario tree approach.

#### 5. Results

In this work first the objective functions were studied individually, that is, the different performances of the supply chain are studied separately. Then, the three performances are studied in the same objective function in order to perceive the different trade-offs between the three pillars of sustainability. Thus, the cases analysed are: i) Case A - represents the optimal economic performance; ii) Case B - represents the optimal social performance; iii) Case C - represents the optimal environmental performance; iv) Case D - represents the optimal equity between social, environmental, and economic performances. From the analysis of these cases, the obtained values are stated in Table 1.

Table 1 Outcomes for the economic, social, and environmental impacts.

	Case A	Case B	Case C	Case D
Obj. Function	Max eNPV	Max eSocImpact	Min eEnvImpact	Max (eNPV+eSocImpact-eEnvImpact)
eNPV	1.563x10 <sup>9</sup> €	1.486x10 <sup>9</sup> €	1.199x10 <sup>9</sup> €	1.512x10 <sup>9</sup> €
eSocImpact	0.571x10 <sup>8</sup> €	1.415x10 <sup>8</sup> €	0.710x10 <sup>8</sup> €	0.856x10 <sup>8</sup> €
eEnvImpact	1.024x10 <sup>9</sup> €	1.132x10 <sup>9</sup> €	9.446x10 <sup>8</sup> €	9.798x10 <sup>8</sup> €

Analyzing the obtained results, it is possible to verify that case A denotes the highest value of eNPV (1.56x10<sup>9</sup> €) since its objective function is maximized. On the other hand, case B returns the highest value of eSocImpact (1.42x10<sup>8</sup> €) because in this case the social component is maximized. Case C that considers the minimization of the environmental impact has the lowest value of eEnvImpact which corresponds to 9.45x10<sup>8</sup> €. Finally, in

case D, where the three performances are considered in the same objective function a trade-off between the three goals was obtained. The network obtained in this case results from the trade-off between i) transport, production, and installation costs; ii) environmental impact of the transport and installation of entities; and iii) social impact of job creation. On the economic approach the global profit decrease of 5 million euros is observed when comparing to case A. On the environmental impact this has also decreased by around 45 million euros. For the social performance the value achieved between the values of cases A and B. Finally, it is possible to verify that the overall profit improved by more than 310 million euros between cases C and D, while the impact of the environment is worse at nearly 35 million. The resulting supply chain network structure for case D is described in Table 2. It can be seen that a new factory is needed in addition to the existing factory. In addition, there is a need to open new warehouses in order to increase the overall existing capacity. It is also possible to observe that all factories are supplied by all suppliers, which results from the balance between economic and environmental costs. In terms of intermodal transport, air transport is not used, only sea transport is used.

Table 2 Supply chain network structure for case D

Allocation of suppliers	All suppliers supply all factories	
Factories	Existing one plus Galway	
Warehouses	Existing one plus Galway and Bremen	
Production	Most production of fp1 is in the existing factory	
	Most production of fp2 is shared between factories	
	Most production of fp3 is in the new factories	
Remanufacture	Most of rp1 is in the existing facility	
	rp2 is only performed in the existing facility	
	Most of rp3 is in the existing facility	
Recovered products	Minimum possible	
Technologies	Alternative technologies are preferred in the new factories	
Inventory	fp1	Divided between all warehouses
	fp2	60% in Lyon, 25% in Galway, and 15% in Bremen
	fp3	25% in Lyon, 55% in Galway, and 20% in Bremen
Transportation	Road	12 trucks
	Air	Not adopted
	Sea	Adopted

## 6. Conclusions

This paper allows us to understand the importance of monetization in helping to understand the overall performances of the supply chain. Considering that this work includes the three pillars of sustainability, the process of monetization allows to convert both environmental and social performances into monetary units, which can be particularly useful for decision-makers that are used to deal with money when managing their supply chain. Such option permits to obtain in an easier form solution where a trade-off between the three sustainability goals is obtained.

Further research should be done on this topic to develop monetization methods to be a trustworthy alternative for multi-objective optimization approaches. Also, an extension of this work should account for other uncertainty sources.

## 7. Acknowledgements

This work is financed by national funds through the FCT-Foundation for Science and Technology, I.P., under the project UIDB/00097/2020.

## References

- A. Barbosa-Póvoa, C. da Silva, and A. Carvalho. 2018. “Opportunities and Challenges in Sustainable Supply Chain: An Operations Research Perspective.” *European Journal of Operational Research* 268 (2): 399–431. <https://doi.org/10.1016/J.EJOR.2017.10.036>.
- C. da Silva, Ana Barbosa-Póvoa, and Ana Carvalho. 2020. “Environmental Monetization and Risk Assessment in Supply Chain Design and Planning.” *Journal of Cleaner Production* 270 (October): 121552.
- P. Beske-Janssen, Matthew Phillip Johnson, and Stefan Schaltegger. 2015. “20 Years of Performance Measurement in Sustainable Supply Chain Management – What Has Been Achieved?” *Supply Chain Management: An International Journal* 20 (6): 664–80. <https://doi.org/10.1108/SCM-06-2015-0216>.
- European Commission. 2013. “Commission Recommendation of 9 April 2013 on the Use of Common Methods to Measure and Communicate the Life Cycle Environmental Performance of Products and Organisations.”
- T. Popovic, Ana P. Barbosa-Póvoa, Andrzej Kraslawski, and Ana Carvalho. 2018. “Quantitative Indicators for Social Sustainability Assessment of Supply Chains.” *Journal of Cleaner Production* 180 (April): 748–68. <https://doi.org/10.1016/j.jclepro.2018.01.142>.

# Strategic low-carbon hydrogen supply chain planning under market price uncertainty

Tushar Rathi<sup>a</sup>, Jose M. Pinto<sup>b</sup>, Qi Zhang<sup>\*a</sup>

<sup>a</sup>*Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455, USA*

<sup>b</sup>*Linde Digital Americas, Linde plc, Danbury, CT 06810, USA*

*\*qizh@umn.edu*

## Abstract

With the increased emphasis on decarbonization in recent years, there is a growing interest in the development of sustainable industrial supply chains. With its widespread applications, clean hydrogen holds immense potential to push forward the goal of a net-zero economy. For this reason, we develop mathematical models to optimize the design and planning of low-carbon hydrogen supply chains, specifically when the potential markets are willing to pay certain premiums for hydrogen of different levels of carbon intensity. We introduce the concept of a hydrogen supply chain with virtual distribution of carbon intensities, which allows more cost-effective supply chain operations compared with the conventional model in which different carbon intensities associated with a product can only be achieved through physical blending. Additionally, owing to the difficulty in predicting future hydrogen prices, we propose a multistage stochastic programming framework to account for uncertainty in market prices. The proposed model is applied to a network of hydrogen production plants and potential markets, highlighting differences in decisions across different scenarios.

**Keywords:** hydrogen-based economy, virtual supply chain, mixed-integer nonlinear programming, nonconvex optimization, multistage stochastic programming

## 1. Introduction

Clean hydrogen is expected to play an important role in the decarbonization of multiple sectors, including energy, transportation, chemicals, and agriculture. This presents a major opportunity for hydrogen producers; however, hydrogen supply chain planning (Li et al., 2019; Zhang and Pinto, 2022) is a challenging problem as future demands and prices are difficult to predict. Also, with growing awareness of global environmental issues, we see an increasing number of organizations pledging to invest in decarbonization initiatives and meet sustainability goals. Over the next few years, we expect hydrogen markets to follow suit with a greater appetite for low-carbon hydrogen. However, the prices offered in different markets may vary substantially depending on the consumers' budgets, demands, and preferences. Additionally, we assume the price offered in each market to be a function of the "grade" of the low-carbon hydrogen requested, which corresponds to the carbon intensity of the hydrogen in the supplied batch. For example, 60%-grade hydrogen is equivalent to hydrogen that is composed of 60% hydrogen produced with zero carbon emissions and 40% conventionally produced hydrogen. To achieve a certain hydrogen grade, one may have to blend the hydrogen from different production facilities that exhibit different carbon intensities. Alternatively, since the chemical properties of the final product, namely hydrogen, do not depend on its grade (per our definition), one could also assign the grades virtually based on a global



accounting of the carbon emissions, which may significantly reduce the need of physically transporting and blending the hydrogen. In this work, we explore the potential benefits of the latter approach over the former in the planning of a future low-carbon hydrogen supply chain over multiple years.

The proposed notion of a “virtual” hydrogen supply chain is inspired by the concept of renewable energy credits (RECs), which are certificates that represent proof that a given amount of electricity was generated from a renewable energy source (Berry, 2002). RECs can be traded and used by the buyers to offset their carbon emissions. As such, RECs allow renewable energy generated at one location to be counted toward the reduction in carbon intensity at a different location. This is analogous to the clean hydrogen produced at one location but virtually assigned to another location in our approach.

The prices offered in different markets will depend on, among other factors, the willingness of the consumers to pay a premium for clean hydrogen, which is not only highly uncertain but will likely change over time. Hence, we also extend the model to a multistage stochastic programming formulation (Ruszczyński and Shapiro, 2003) that effectively accounts for uncertainty in future market prices. We analyze the resulting investment strategies and draw initial insights via a case study involving a small network of hydrogen production plants and potential markets.

## 2. Mathematical formulation

We consider a network consisting of a set of hydrogen production plants  $I$  and a set of potential buyers (markets)  $J$ . A hydrogen plant  $i$  has initial production capacity  $\bar{C}_{i0}$  and can produce hydrogen of a fixed grade,  $\lambda_i$ , which can vary from 0% (gray hydrogen) to 100% (green hydrogen). Additionally, a plant  $i$  has  $|\mathcal{K}_i|$  permissible expansion points, with the difference between two consecutive expansion points  $k$  and  $k - 1$  defined by the parameter  $\bar{\Delta}_{ik}$ . We consider a planning horizon  $\mathcal{T}$ , and a potential market  $j$  is assumed to have a maximum demand  $d_{jt}^{\max}$  in time period  $t \in \mathcal{T}$ . Importantly, a market  $j$  offers different prices for different grades of hydrogen, which is captured by function  $f_{jt}(c_{jt})$ , where  $c_{jt}$  is the grade of hydrogen promised to (requested by) market  $j$  in time period  $t$ . The goal is then to determine the optimal capacity expansion and supply chain decisions to maximize the overall profit during the planning horizon. In the subsequent subsections, we discuss hydrogen supply chain optimization models with different characteristics.

### 2.1. Basic supply chain model

We first present a basic supply chain (BSC) model formulated as the following nonconvex mixed-integer nonlinear program (MINLP):

$$\begin{aligned} \underset{x,d,c,F,C,\Delta}{\text{maximize}} \quad & \sum_{t \in \mathcal{T}} \alpha_t \left[ \sum_{i \in I} \sum_{j \in J} f_{jt}(c_{jt}) F_{ijt} - \sum_{i \in I} \sum_{k \in \mathcal{K}_i} q_{ikt} \bar{\Delta}_{ik} x_{ikt} - \right. \\ & \left. \sum_{i \in I} \sum_{j \in J} r_{ij} F_{ijt} - \sum_{i \in I} b_{it} \sum_{j \in J} F_{ijt} \right] \end{aligned}$$

$$\text{subject to } C_{i0} = \bar{C}_{i0} \quad \forall i \in I \quad (1)$$

$$C_{it} = C_{i,t-1} + \Delta_{it} \quad \forall i \in I, t \in \mathcal{T} \quad (2)$$

$$\Delta_{it} = \sum_{k \in \mathcal{K}_i} x_{ikt} \bar{\Delta}_{ik} \quad \forall i \in I, t \in \mathcal{T} \quad (3)$$

$$x_{ikt} \leq \sum_{\tau=1}^t x_{i,k-1,\tau} \quad \forall i \in I, k \in \mathcal{K}_i \setminus \{1\}, t \in \mathcal{T} \quad (4)$$

$$\sum_{\tau=1}^t x_{ik\tau} \leq 1 \quad \forall i \in I, k \in \mathcal{K}_i, t \in \mathcal{T} \quad (5)$$

$$\sum_{j \in J} F_{ijt} \leq C_{it} \quad \forall i \in I, t \in \mathcal{T} \quad (6)$$

$$\sum_{i \in I} F_{ijt} = d_{jt} \quad \forall j \in J, t \in \mathcal{T} \quad (7)$$

$$\sum_{i \in I} \lambda_i F_{ijt} = c_{jt} d_{jt} \quad \forall j \in J, t \in \mathcal{T} \quad (8)$$

$$0 \leq d_{jt} \leq d_{jt}^{\max} \quad \forall j \in J, t \in \mathcal{T} \quad (9)$$

$$0 \leq c_{jt} \leq 1 \quad \forall j \in J, t \in \mathcal{T} \quad (10)$$

$$F_{ijt} \geq 0 \quad \forall i \in I, j \in J, t \in \mathcal{T} \quad (11)$$

$$C_{it}, \Delta_{it} \geq 0 \quad \forall i \in I, t \in \mathcal{T} \quad (12)$$

$$x_{ikt} \in \{0,1\} \quad \forall i \in J, k \in \mathcal{K}_i, t \in \mathcal{T}, \quad (13)$$

where the binary variable  $x_{ikt}$  equals 1 if the capacity of plant  $i$  is increased to at least point  $k$  in time period  $t$ . The capacity increase and the total capacity of plant  $i$  in time period  $t$  are denoted by  $\Delta_{it}$  and  $C_{it}$ , respectively. The amount of hydrogen transported from plant  $i$  to market  $j$  in time period  $t$  is denoted by  $F_{ijt}$ , whereas the net demand satisfied at market  $j$  in time period  $t$  is captured by variable  $d_{jt}$ . The grade of hydrogen for which a market  $j$  eventually makes the payment in time period  $t$  is denoted by  $c_{jt}$ . The cost coefficient  $q_{ikt}$  is associated with the unit expansion for plant  $i$  from point  $k - 1$  to  $k$  in time period  $t$ ,  $r_{ij}$  is the unit transportation cost between plant  $i$  and market  $j$ , and the unit production cost for plant  $i$  in time period  $t$  is represented by coefficient  $b_{it}$ . The capacity constraints are adopted from our previous work (Rathi and Zhang, 2022), where constraints (1)-(3) represent the capacity balance. Constraints (4) ensure that to increase the capacity to point  $k$ , we must have already increased the capacity to  $k - 1$  in the current or one of the preceding time periods. Constraints (5) state that the capacity cannot be increased to point  $k$  more than once in any time period  $t$ . Constraints (6) ensure that production must not exceed the capacity, and constraints (7) enforce the usual demand satisfaction condition for each market. Equations (8) are the blending constraints. Lastly, constraints (9)-(13) enforce bounds on the decision variables. With  $\alpha_t$  as the discount factor in time period  $t$ , the objective is to maximize the overall profit resulting from a difference of revenue generated from supplying hydrogen and costs incurred from capacity expansion, transportation, and production.

### 2.2. Virtual supply chain model

We now present the model that allows virtual distribution of carbon intensities and hence indirect satisfaction of hydrogen grade, which we refer to as the virtual supply chain (VSC) model. Compared to the BSC model, it requires the following modifications:

$$\bar{F}_{it} = \sum_{j \in J} F_{ijt} \quad \forall i \in I, t \in \mathcal{T} \quad (14)$$

$$\sum_{i \in I} \lambda_i \bar{F}_{it} = \sum_{j \in J} c_{jt} d_{jt} \quad \forall t \in \mathcal{T} \quad (15)$$

$$\bar{F}_{it} \geq 0 \quad \forall i \in I, t \in \mathcal{T}, \quad (16)$$

where we introduce  $\bar{F}_{it}$  to denote the net hydrogen production at plant  $i$  in time period  $t$ . Constraints (14) represent flow distribution from plants to markets. Unlike in the BSC model, where we have a blending constraint for each market in each time period, here we only consider an overall blending constraint per time period represented by constraints (15). Hence, the VSC model is a variant of the BSC model with the same objective function, new decision variable  $\bar{F}_{it}$ , additional constraints (14) and (16), and constraints (8) replaced by (15).

### 2.3. Stochastic virtual supply chain model

To account for market price uncertainty in the VSC model, we first define a set of possible realizations for the price profile of market  $j$  in time period  $t$ ,  $\Xi_{jt} := \{f_{jt}^m(c_{jt}) : m =$

$1, 2, \dots, M_{jt}$ , where  $m$  and  $M_{jt}$  denote the index of a specific realization and the number of possible realizations, respectively. Then the set of all possible scenarios is  $S = \mathbf{x}_{t \in \mathcal{T}} (\mathbf{x}_{j \in \mathcal{J}} \Xi_{jt})$ . The following is a compact representation of the corresponding multistage stochastic virtual supply chain (SVSC) model:

$$\begin{aligned} & \text{maximize } \mathbb{E}[Z^{\text{VSC}}] \\ & \text{subject to } w_s \in \mathcal{W}_s^{\text{VSC}} \quad \forall s \in S \\ & \quad \text{NACs,} \end{aligned}$$

where  $Z^{\text{VSC}}$ ,  $w_s$ , and  $\mathcal{W}_s^{\text{VSC}}$  denote the objective function, the set of variables for scenario  $s$ , and the feasible set for scenario  $s$ , respectively. NACs represent the nonanticipativity constraints on binary variables  $x$ , ensuring that the same capacity expansion decisions are made for indistinguishable scenarios at different time points during the planning horizon.

### 3. Case study

We consider a network with three hydrogen plants with gray (0%-grade), blue (50%-grade), and green (100%-grade) hydrogen production capacities, respectively. Each plant has an initial installed capacity of 10 units. Further, the network has five hydrogen markets, with linear price profiles positively correlated with the hydrogen grade, i.e.,  $f_{jt}(c_{jt})$  is linear in  $c_{jt}$  with a nonnegative slope. The assumption of linear price profiles transforms the model into a nonconvex mixed-integer quadratically-constrained quadratic program (MIQCQP). We consider a planning horizon with three time periods and a discount factor of 3%. All models were implemented using JuMP v1.4.0 in Julia v1.8.2 and were solved using Gurobi v9.5.2.

For the deterministic cases (BSC and VSC), the variation in premium offered for 100%-grade hydrogen by different markets across the planning horizon is shown in Figure 1. Note that we assume the price offered by all markets for 0%-grade hydrogen is the same across all time periods. Figure 1 further summarizes the solutions for the deterministic cases. The plants and markets are represented using square and circular nodes, respectively. The color coding used is as follows: (i) For plants, gray, blue, and green nodes indicate 0%-, 50%-, and 100%-grade hydrogen production, respectively. The size of a plant node corresponds to its current installed capacity. For clarity, we also mention the additional capacity installed for each plant in each time period next to each plant node, e.g., “+5” implies a 5-unit capacity increase. (ii) For markets, the node color indicates the grade of hydrogen requested, and the color shade indicates the relative proportion of clean hydrogen in the requested batch. For example, 20-45%-grade hydrogen has a lighter shade of gray compared to a darker shade for <20%-grade hydrogen. The regular blue and green shades represent 45-60%- and >85%-grade hydrogen, respectively.

One can observe the following key differences in the expansion and supply chain decisions between the BSC and VSC models: (i) In the BSC case, the grade of hydrogen requested and the grade of hydrogen supplied (which may require blending) match exactly, whereas in the VSC case, this is often not the case. Since the chemical properties of hydrogen are the same regardless of its grade, we do not need to exactly match the requested hydrogen grade at each market and only need to ensure that the net production of a particular grade matches its overall demand within the network. For example, because Market 1 pays the highest premium for 100%-grade hydrogen, we simply supply it the same from Plant 3 in the BSC case, whereas in the VSC case, we supply it gray hydrogen from Plant 1 instead, which is more cost-effective due to lower transportation cost, yet it pays for clean hydrogen. (ii) In the BSC case, the price offered by Market 5 for any grade of hydrogen is not sufficient to offset the production and transportation cost, resulting in

no hydrogen supply at all times, whereas in the VSC case, the liberty to supply a different grade of hydrogen than the one requested (due to having the same chemical properties) allow us to supply hydrogen to Market 5. (iii) Unlike in the BSC case, we see more localized transportation in the VSC case, owing to the VSC model's emphasis on cost savings through lower transportation costs while maximizing profits through higher premiums offered by markets for low-carbon hydrogen. In particular, the net profit earned increased from \$504,926 in the BSC case to \$725,903 in the VSC case, a significant 43.8% increase. (iv) Finally, in the VSC case, we see greater investment in the capacity expansion of the gray hydrogen Plant 1 because Market 1, which pays the highest premium, is located in its proximity. Since the objective is to maximize profit, it may or may not support the expansion of clean hydrogen plants. Instead, determining whether VSC promotes or stifles the expansion of clean hydrogen plants necessitates a careful examination of relative plant and market locations, as well as the various logistics costs.

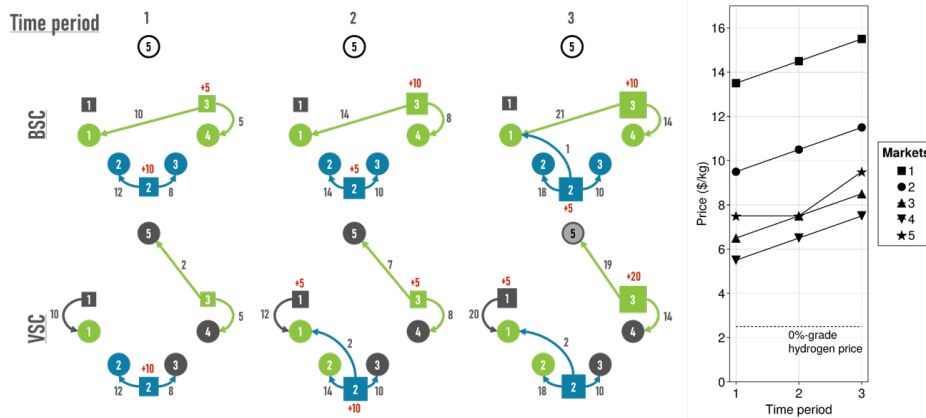


Figure 1: Illustrating differences in hydrogen supply chain decisions obtained via the BSC and VSC models. Plants and markets are represented by square and circular nodes, respectively. The color of the nodes represents the grade of hydrogen produced/requested at a plant/customer location. The numbers on the edges represent the flows; the size of a plant node represents its current production capacity. The number in red font next to a plant node represents the additional capacity installed. The plot on the right depicts the premiums offered by each market for 100%-grade hydrogen.

In the SVSC case, we consider eight equally probable scenarios with varying market prices across the planning horizon. To emphasize the importance of accounting for uncertainty in market price, we discuss differences in the decisions for two particular scenarios. The premium offered by different markets for 100%-grade hydrogen under the two scenarios and the corresponding solutions are shown in Figure 2. Scenarios 1 and 2 have very different price profiles, which significantly impact the grade of hydrogen requested by each market in each scenario. For example, in the first time period, Market 1 offers to pay the highest and lowest premium for 100%-grade hydrogen in Scenarios 1 and 2, respectively, which is reflected in the eventual decision of the grade requested in these scenarios. Similarly, in Scenario 2, Market 2 offers a significantly higher price for 100%-grade hydrogen than Scenario 1, making it more profitable to make it pay for green hydrogen. In addition to the transportation and production costs, different price profiles across scenarios affect the amount of hydrogen supplied by each plant, resulting in different capacity expansion decisions that satisfy the nonanticipativity conditions. For example, at the start of the planning horizon, both scenarios are indistinguishable; thus, the capacity expansion decisions made in each are identical, as illustrated in Figure 2.

Finally, we discover that a market's proximity to a hydrogen plant does not guarantee that its hydrogen demand will be met. For example, in the second time period of Scenario 1, Plant 3 has enough capacity to meet Market 4's maximum demand; however, the low price offered by Market 4 is insufficient to offset the transportation and high production costs. In the first time period, on the other hand, we see that Plant 3 does supply some hydrogen to Market 4, which was made possible by the high price offered for 100%-grade hydrogen by Markets 1 and 5, which helps offset the production and transportation costs.

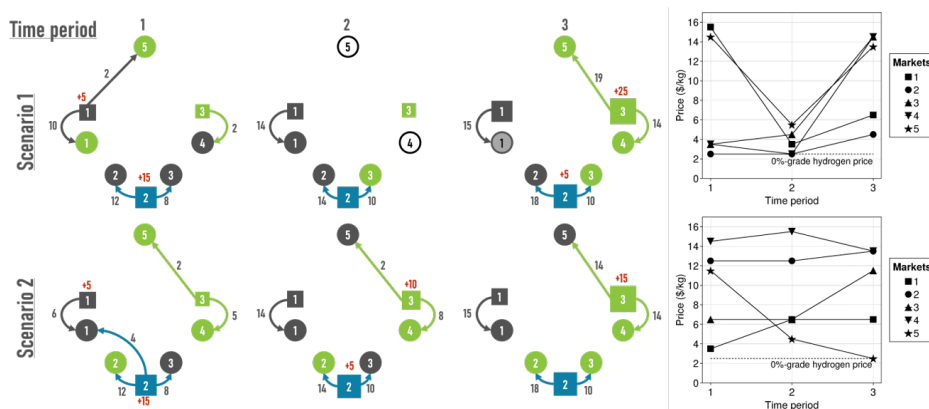


Figure 2: Illustrating differences in decisions between two particular scenarios, obtained by solving the SVSC model. The plots on the right depict the premiums offered by each market for 100%-grade hydrogen in the two scenarios.

#### 4. Conclusions

In this work, we developed a multiperiod supply chain planning model for low-carbon hydrogen distribution, where each potential market offers a different price for different grades of hydrogen; here, a higher grade corresponds to hydrogen with a lower carbon intensity. The invariability in chemical properties of hydrogen with the mode of production relaxes the need to satisfy the exact hydrogen grade requests as long as the hydrogen production satisfies certain global production constraints within the network. As shown in our case study, this can lead to significant cost savings primarily from avoiding unnecessary transportation costs and relying more on the local distribution of hydrogen. Furthermore, accounting for market price uncertainty using the proposed stochastic programming framework allows for the generation of investment strategies tailored to each specific scenario while minimizing the overall expected cost.

#### References

- D. Berry, 2002. The market for tradable renewable energy credits. *Ecological Economics*, 42(3), 369-379.
- L. Li, H. Manier, and M.A. Manier, 2019. Hydrogen supply chain network design: An optimization-oriented review. *Renewable and Sustainable Energy Reviews*, 103, pp. 342-360.
- T. Rathi and Q. Zhang, 2022. Capacity planning with uncertain endogenous technology learning. *Computers & Chemical Engineering*, 164, 107868.
- A. Ruszczyński and A. Shapiro, 2003. Stochastic programming models. *Handbooks in Operations Research and Management Science*, 10, pp. 1-64.
- Q. Zhang and J.M. Pinto, 2022. Energy-aware enterprise-wide optimization and clean energy in the industrial gas industry. *Computers & Chemical Engineering*, 165, 107927.

# Ontology Modelling for Valorisation of Sugarcane Bagasse

Maureen Chiebonam Okibe,<sup>a</sup> Michael Short,<sup>a</sup> Franjo Cecelja,<sup>a</sup> Madeleine Bussemaker,<sup>a\*</sup>

<sup>a</sup>*Department of Chemical and Process Engineering, University of Surrey, Guildford, Surrey GU2 7XH, United Kingdom.*

[m.bussemaker@surrey.ac.uk](mailto:m.bussemaker@surrey.ac.uk)

## Abstract

Sugarcane bagasse (SCB) is an agro-industrial residue extracted during sugar processing. Sugar mills use basic process technologies for bagasse valorisation. Residues are often disposed improperly or burned inefficiently by thermal boilers causing environmental pollution. Biorefining bagasse as substrate for bio-based chemical production is superior to biomass disposal by incineration, burning and landfill. The conversion of bagasse for value-added applications may be economical with less environmental impact on humans and the ecosystem. Computer aided tools such as ontologies for knowledge modelling represent available information for bagasse feedstocks. This paper presents a reference model referred to as the SCB Ontology which is limited to a framework for the modelling of knowledge on the current utilisation of SCB feedstocks from principal sugarcane-cultivating countries. The SCB Ontology identifies opportunities for efficient bagasse valorisation by principal sugarcane producers. Potential application of the SCB Ontology for bagasse valorisation which is a circular bioeconomy initiative was also discussed.

**Keywords:** Sugarcane Bagasse, Biorefining, Knowledge Modelling, Ontology

## 1. Introduction

Sugarcane is cultivated in over 100 countries around the world, mainly from tropical and sub-tropical regions. The global annual production of sugarcane is ~2 billion tonnes. For every tonne of sugarcane processed in sugar mills, ~30% of sugarcane bagasse (SCB) is generated (Narisetty et al., 2022). SCB are fibrous residues derived during sugar processing which are dense in polymeric substances. Consequently, SCB feedstocks are considered the world's biggest second-generation lignocellulose biomass (LCB) (Konde et al., 2021). These complex polymers are the energy carriers found in the SCB biomass. SCB polymers require processing or valorisation using a range of process technologies prior its utilization (Konde et al., 2020). Commercial biorefining of SCB feedstocks has opened new horizons towards achieving circular bioeconomy initiatives by reusing second generation biomass for value-added products (Narisetty et al., 2022). However, the economic and environmental benefits of SCB valorisation are currently unmaximized.

Ontologies are tools for knowledge representation and modelling (Liyana et al., 2019). Scientific and engineering researchers use ontologies for modelling knowledge and data representation in an organised and structured manner. The level of organisation involved in Web Ontology Language (OWL) makes it possible for new knowledge to be inferred (Liyana et al., 2019). Computer aided tools can accelerate strategic decision-

making for simpler and sustainable feedstock supply chains using knowledge modelling ontologies (Trokanas et al., 2018). OWL semantic infrastructure was promising for knowledge representation in lignocellulose biorefinery domain, where biorefining of various agricultural residues was modelled (Bussemaker et al., 2018; Trokanas et al., 2018). Ontologies were demonstrated to improve search and retrieval of biomass information, facilitate data integration, and enhance decision-making for biomass component utilization (Trokanas et al., 2018). Classifications of LCB with respect to biomass composition and potential use have also been discussed (Bussemaker et al., 2018). However, agro-industrial residues such as SCB were not considered in previous OWL studies.

The objective here, is to demonstrate a systematic approach to knowledge modelling for SCB availability and current use toward a proposed ontology framework. The framework was developed using semantic knowledge from SCB feedstocks and is interpretable by both computers and humans. The SCB Ontology aims to formalise knowledge representation toward decision support for future valorisation of SCB.

## 2. The SCB Ontology

Previous studies have shown that knowledge represented in OWL can be superior to data stored in traditional relational database management systems (RDBMS). To model complex highly connected data, an underlying graph model used by OWL is often both more intuitive and more efficient than RDBMS model, which can require several complex joins to link data (Ferrer et al., 2021). OWL framework provides directly accessible graph model i.e., knowledge graphs. In addition, OWL provides a semantic layer that is more abstract and closer to business and chemical process engineering models than relational models. OWL satisfies a World Wide Web Consortium (W3C) standard so it can facilitate interoperability (Dutta & DeBellis, 2020).

The SCB Ontology (Figure 1) is designed using the OWL with the Protégé 5.5.0 ontology editor from Stanford University (McGuinness, 2001). Designing the SCB Ontology follows an iterative process using a semantic framework. This Ontology is made up of classes, subclasses, properties, individuals, and rules. The axioms and data were gathered from literature sources. The semantics of OWL are defined by Description Logic which is a decidable subset of First Order Logic (FOL). Each subclass represents one aspect of SCB as gathered from literature sources on sugarcane bagasse.

Agwa et al. (2022) reported that major sugar-producing countries (Brazil, India, China, and Thailand) produce over 100 MT (Million Tonne) of sugarcane per year. These SCB nations were represented in the classes: SCB\_Producer which is EquivalentTo Sugarcane\_Producer (Figure 2a). Principal\_Producer is a subclass of: SCB\_Producer and is an example of a defined class. A defined class is represented in Protégé using a yellow circle with three small white lines (Figure 1). A defined class has both the necessary and sufficient conditions required for an Individual to be classified as an instance of that class completely defined. Thus, the OWL reasoner can automatically infer when an individual is an instance of a defined class based on the property values for that individual and the axioms defining the class (Figure 2a).

Classifications of producing nations were defined following classifications in literature (Gikonyo, 2021; Karp et al., 2022) into Top\_Producer, Medium\_Producer and

Low\_Producer. An additional class (Principal\_Producer) was created to encompass the four largest SCB producers, responsible for 70% of global SCB production. These defined subclasses possess data properties modelled on Sugarcane\_Tonnage (Figure 2a). Each Principal\_Producer produces >100 MT sugarcane (Figure 2a). The Top\_Producer generates 8-100 MT of sugarcane. The Medium\_Producer produces 1-8 MT. The Low\_Producer nations produce <1 MT. (Figure 2a).

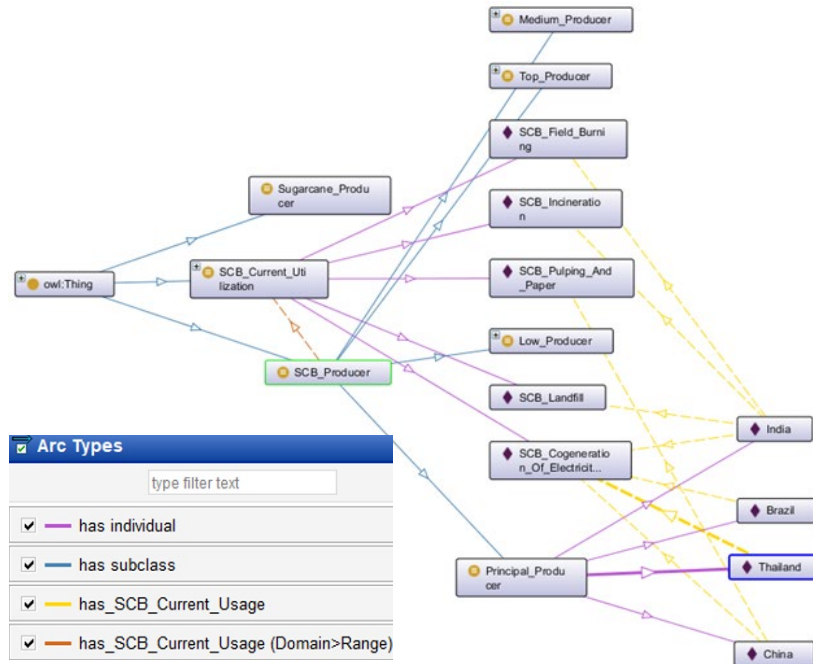


Figure 1: SCB Ontology with SCB Producers and SCB Current Utilization Domains

The SCB\_Producer class has a text descriptor known as the ‘rdfs comment’ (Figure 2a), providing descriptions which is characterised by the human readable property. This holds additional information that can be read by humans. Brazil, India, China, and Thailand were modelled in the SCB Ontology under Principal\_Producer as individual instances illustrated using the purple diamonds. In the SCB Ontology, the object property holds property for SCB usage as ‘has\_SCB\_Current\_Usage’. Principal\_Producer was modelled set against SCB\_Current\_Usage (Figure 1) which is an object property that has SCB\_Current\_Utilization as its domain (Figure 1). The class, SCB\_Current\_Utilization (Figure 1) was used to represent uses of SCB. Gikonyo (2021) reported on the uses of SCB by sugar processing mills in different countries. Principal\_Producer models the producing nations against utilization of SCB (Figure 1). For example, Vandenberghe et al. (2022) reported that Brazil utilizes SCB for cogeneration of electricity and that usage was represented linking Brazil, a Principal\_Producer (subclass of SCB\_Producer EquivalentTo Sugarcane\_Producer) to the individual: Cogeneration\_Of\_Electricity which is an instance of the class SCB\_Current\_Utilization (Figure 1).





Figure 2a: Schematic representation of User Interface showing Search on SCB\_Producer

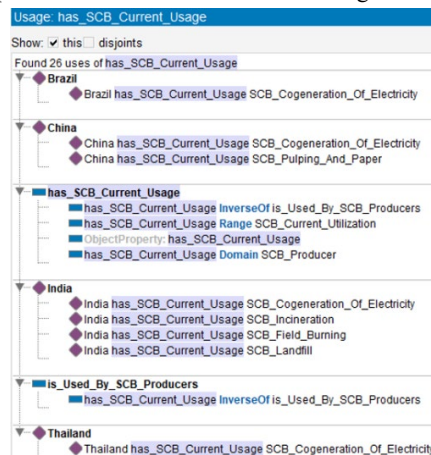


Figure 2b: Information Retrieval on Current Usage of Principal\_Producer Countries

### 3. Potential Applications of the SCB Ontology

The SCB Ontology (Figures 2a and 2b) demonstrates some of the reasoning capabilities of OWL. Principal\_Producer (i.e., China, Thailand, Brazil, and India) is a defined subclass of SCB\_Producer. This means that the axioms used to define it provide both necessary and sufficient conditions for an individual to be an instance of that class. In the case of Principal\_Producer the important axiom is that the producer creates more than 100 MT of sugarcane per year. The Ontology also records how the SCB is utilized, e.g., for fuels by burning in thermal boilers for cogeneration of electricity in sugar producing mills (Vandenberghe et al., 2022).

Utilizing bagasse for cogeneration of electricity is astronomical in Brazil as a current practice to curb the overdependency on fossil fuel energy and associated emissions preventing the climatic impact of global warming (Vandenberghe et al., 2022). However, Gil et al. (2013) reported that bagasse fuel generates emissions with significant impact on humans and the ecosystem. SCB feedstocks may be considered for more

meaningful use with lower impacts on the environment and humans. For example, numerous organic acids and polyols which are platform chemicals may be explored by understanding the lifecycle stages during their production (Gil et al., 2013).

India, one of the Principal\_Producer of SCB still disposes bagasse (~1500 metric tonne per day) in landfills and burns bagasse in incinerators and open fields (Konde et al., 2021). Data modelled from the SCB Ontology suggest that SCB\_Producer may have newer opportunities from the diversification of SCB residues derived from sugar-processing mills, especially in India since SCB is still sent to landfills, burned in fields, and incinerated as disposal measures. This supports the circular bioeconomy initiative for the upcycle of SCB feedstocks (Narisetty et al., 2022).

The presented Ontology shows a framework for development of an ontology around SCB feedstocks and use. In this study, other subclasses of the SCB\_Producer class were defined. For example, the SCB\_Current\_Utilization class defines the various ways (e.g., Cogeneration) that SCB can be utilized. The OWL reasoner automatically tracks the connections between the two domains: 21 and 26 uses of SCB\_Producer and has\_SCB\_Current\_Usage respectively (Figures: 2a and 2b).

Once all available SCB production knowledge is modelled and incorporated into the SCB Ontology, searches could infer new knowledge on domains, classifications and relationships. Through modelling environmental and economic knowledge on existing valorisation processes, new knowledge for cost-benefit analysis and life cycle assessment (LCA) can be inferred (Liang, 2021). Consequently, inferred data generated from new searches may inform decision-making and strategic initiatives for development of new, sustainable products with lower carbon footprints and environmental impact. Ontologies facilitate improved use of knowledge and will deliver more strategic decision which can mitigate scientific guesses by researchers. This type of semantic data representation meets the 'FAIR' requirements (Findable, Accessible, Interoperable and Reusable) of information during searches and information retrieval (Dutta & DeBellis, 2020). This can benefit LCB commercial players and researchers when making and justifying decisions.

#### **4. Conclusions and Future Work**

An overview of the SCB Ontology framework, development of its structure and potential applications have been presented. This study discusses the background and usefulness of using the OWL framework to share data and represent SCB knowledge. Ontologies are useful for enhancing decision-making processes and supporting dynamic opportunities which can create rapid industrial change within the LCB biorefineries (Liyana et al., 2019). SCB Ontology benefits from the semantic interoperability of multiple subclasses and domains which would assist in driving optimal solutions for more strategic decisions. In future work, modelling the environmental and economic sustainability of SCB valorisation will be explored through semantics around carbon or greenhouse gas emissions and cost data (Liang, 2021). OWL will therefore be used to create a knowledge repository to support LCA and technoeconomic assessment (TEA) studies.

#### **Acknowledgements**

The team would like to offer their gratitude to EPSRC Doctoral Training Programme for the PhD funding through the University of Surrey Doctoral College. Special thanks to Michael DeBellis for his continuous support on OWL Tutorial throughout this research.

## References

- Agwa, I., Zeyad, A. M., Tayeh, B. A., Adesina, A., de Azevedo, A. R. G., Amin, M., & Hadzima-Nyarko, M. (2022). A comprehensive review on the use of sugarcane bagasse ash as a supplementary cementitious material to produce eco-friendly concretes. *Materials Today: Proceedings*
- Bussemaker, M., Trokanas, N., Koo, L., & Cecelja, F. (2018). Ontology Modelling for Lignocellulosic Biomass: Composition and Conversion. In *Computer Aided Chemical Engineering* (Vol. 43, pp. 1565–1570). Elsevier B.V.
- Dutta, B., & DeBellis, M. (2020). *CODO: An Ontology for Collection and Analysis of Covid-19 Data*.
- Ferrer, B., Mohammed, W. M., Ahmad, M., Iarovyi, S., Zhang, J., Harrison, R., & Martinez Lastra, J. L. (2021). Comparing ontologies and databases: a critical review of lifecycle engineering models in manufacturing. *Knowledge and Information Systems*, 63(6), 1271–1304.
- Gikonyo, B. (2021). *Sugarcane as biofuel feedstock: advances toward a sustainable energy solution*.
- Karp, S. G., Schmitt, C. C., Moreira, R., de Oliveira Penha, R., de Mello, A. F. M., Herrmann, L. W., & Soccol, C. R. (2022). Sugarcane Biorefineries: Status and Perspectives in Bioeconomy. In *Bioenergy Research*. Springer.
- Konde, K. S., Nagarajan, S., Kumar, V., Patil, S. v., & Ranade, V. v. (2021). Sugarcane bagasse based biorefineries in India: Potential and challenges. *Sustainable Energy and Fuels*, 5(1), 52–78.
- Liang, J. S. (2021). A knowledge with ontology representation for product life cycle to support eco-design activities. *Journal of Engineering, Design and Technology*.
- Liyana, N., Mohd, L., Law, F., Mahmoud, M. A., Tang, A. Y. C., Lim, F.-C., Kasim, H., Othman, M., & Yong, C. (2019a). A Review of Ontology Development Aspects. In *IJACSA International Journal of Advanced Computer Science and Applications* (Vol. 10, Issue 7).
- McGuinness, D. L. (2001). *Ontology Development 101: A Guide to Creating Your First Ontology*.
- Narisetty, V., Okibe, M. C., Amulya, K., Jokodola, E. O., Coulon, F., Tyagi, V. K., Lens, P. N. L., Parameswaran, B., & Kumar, V. (2022). Technological advancements in valorization of second generation (2G) feedstocks for bio-based succinic acid production. *Bioresource Technology*, 360, 127513.
- Gil, M., Contreras Moya, A. M., & Rosa Domínguez, E. (2013). Life cycle assessment of the cogeneration processes in the Cuban sugar industry. *Journal of Cleaner Production*, 41, 222–231.
- Trokanas, N., Koo, L., & Cecelja, F. (2018). Towards a Methodology for Reusable Ontology Engineering: Application to the Process Engineering Domain. In *Computer Aided Chemical Engineering* (Vol. 43, pp. 471–476). Elsevier B.V.
- Vandenbergh, L. P. S., Valladares-Diestra, K. K., Bittencourt, G. A., Zevallos Torres, L. A., Vieira, S., Karp, S. G., Sydney, E. B., de Carvalho, J. C., Thomaz Soccol, V., & Soccol, C. R. (2022). Beyond sugar and ethanol: The future of sugarcane biorefineries in Brazil. *Renewable and Sustainable Energy Reviews*,

# Modelling availability and affordability concerns in the design and planning of pharmaceutical supply chains

Inês Duarte<sup>a</sup>; Bruna Mota<sup>a</sup>; Tânia Pinto-Varela<sup>a</sup>; Ana Amaro<sup>b</sup>; Ana Paula Barbosa-Povoa<sup>a</sup>

<sup>a</sup> *Centre for Management Studies of IST (CEG-IST), University of Lisbon.  
Av Rovisco Pais, 1, 1049-001 Lisbon, Portugal*

<sup>b</sup> *Coimbra Business School-ISCAC/ IPC, Quinta Agrícola, 3045-601 Coimbra, Portugal  
[ines.r.duarte@tecnico.ulisboa.pt](mailto:ines.r.duarte@tecnico.ulisboa.pt)*

## Abstract

Although sustainability principles have been gaining attention when managing supply chains, integrating social objectives remains an essential concern to be explored. In particular, social impact concerns related to equity in access to pharmaceutical products are crucial to be considered. This work addresses this goal and proposes a decision-support tool based on a multi-objective MILP model that integrates strategic-tactical decisions while considering sustainability principles. The social pillar is modelled using two approaches: the location of facilities to maximize access to medicines where the incidence of disease is higher; and a second approach where a metric based on the economic concept of demand-to-price elasticity is used so as to measure how sensitive markets are to medicines' prices allowing producers to manage the trade-off between prices' portfolio and markets, enhancing social equity criteria, without profit losing. The economic pillar is addressed through the Net Present Value, and the environmental pillar follows a Life Cycle Analysis methodology. The model is applied to a representative case study and different scenarios are analysed where it is shown how strategic and tactical decisions impact these performance indicators.

**Keywords:** Pharmaceutical Supply Chains; Sustainability; Equity; Demand-to-price elasticity; Optimization

## 1. Introduction

Nowadays, pharmaceutical supply chain operations and their management face numerous challenges triggered by the fast-changing and competitive market, scientific and technological breakthroughs, as well as increased pressures from society. Sustainability has been recognized as one of the driving forces impacting supply chains. However, integrating sustainability aspects considerably increases the complexity of the supply chain design and planning decisions. The United Nation's 2030 agenda for sustainable development set 17 developing goals dedicated to helping build a path to a more sustainable and resilient world, that benefits current and future generations, as well as the planet. In particular, the third and tenth United Nations' sustainable developing goals underline the need to guarantee healthy lives and enhance universal coverage of well-being, reduce inequalities within and among countries, and ensure responsible consumption and production patterns, respectively (United Nations, 2015). Besides the undoubtedly economic and environmental impact, the pharmaceutical sector significantly impacts society, making it crucial to improve the supply chain network and promote global access to medicines. However, assessing the social sustainability dimension still remains a research (Duarte et al., 2022). Some authors evaluated social performance at

an operational level while others assessed social concerns at the strategical level accounting for the number of jobs created.

In the context of the pharmaceutical sector, other social objectives reveal to be crucial. According to the Access to Medicine Index (AtMI), pharmaceutical products must be equitably accessible in all countries, and their accessibility should lean on two main pillars: availability and affordability (Menou et al., 2021). More precisely, according to this index, medicines are to be made, not only, available, but also affordable in countries with a higher burden of disease and less capacity to acquire these products, to improve equity in access among countries (Duarte et al., 2022). Based on this need, this work aims to contribute towards this end and so its main contributions are listed as follows:

- Proposing a pharmaceutical supply chain network design model that includes strategical and tactical decisions;
- Introducing a novel social approach regarding equity in access to pharmaceuticals, where availability and affordability objectives are explored;
- Considering economic, environmental, and social concerns as objective functions and analysing trade-offs between these three sustainability pillars;
- Applying the proposed model to a representative case study of a vaccine supply chain.

## 2. Problem Definition and Mathematical Formulation

The mathematical model, thoroughly described in Duarte et al. (2022), includes **decisions** at the strategical level, such as the number, location, and capacities of factories and warehouses, the selection of technologies for production and storage, as well as transportation modes for the distribution between entities; at the tactical level, decisions on defining supply, production, inventory and distribution planning are included. The problem was subjected to **material balance constraints** at factories, warehouses and hub terminals (Cross-docking at airports and seaports), **entity location and capacity constraints**, **transportation constraints** such as transshipment constraints at the airports and seaports, contracted capacity constraints with road, air and sea carriers, and quantification of the necessary number of trips between entities; Moreover, **technology constraints** regarding production and storage operations such as capacity, production and storage levels, as well as allocation of technologies are included. Finally, **demand constraints** are included to guarantee demand satisfaction of each market.

The sustainability objectives are integrated in the model through three objective functions: the economic performance of the supply chain; the environmental impact assessment; and the social performance. Moreover, an additional social approach is added and integrated to, together with the social objective function, fulfil two strong pillars of equity in access to pharmaceutical products: availability and affordability.

### 2.1. Social Assessment

The first social approach is introduced and explored in Duarte et al., (2022), through the objective function defined in Eq.1). The aim of this objective function is to improve the availability of pharmaceutical products by prioritizing the location of supply chain entities in geographical areas with higher burden of disease. The metric Disability-Adjusted Life Years (DALYs) reflects the burden of disease and is used as a rate per 100,000 population. In Eq. 1) the parameter  $p_i^{DALY}$  is used together with the decision variable  $Y_i$  (equals 1 if entity  $i$  is installed) to prioritize the location of entity  $i$  in geographical areas with higher DALY value.

$$\max \text{PharmaAccess} = \left( \sum_{i \in (I_f \cup I_w)} p_i^{DALY} \cdot Y_i \right) \quad (1)$$

The second social aspect reflects affordability and is based on the economic concept of demand-to-price elasticity. This relation was firstly introduced in the context of pharmaceutical supply chain design in the work of Amaro & Barbosa-póvoa (2009). The elasticity coefficient ( $\epsilon_{mi}$ ) translates the relation between demand and price and is defined by the absolute value of the logarithmic derivative of demand to price. Hence, demand can be considered elastic if its relative change is greater than or equal to the corresponding relative price variation ( $|\epsilon_{mi}| \geq 1$ ), and inelastic otherwise ( $|\epsilon_{mi}| < 1$ ). The elasticity parameter ( $\epsilon_{mi}$ ) is dependent from the product type ( $m$ ) and the market ( $i \in I_C$ ) where the products are distributed and sold. This is because different geographical areas may have different affordability levels, leading to different sensibilities to price variations. Therefore, the main goal of this approach is to explore how the demand-to-price elasticity can be related to the quantity and the price that a company can apply, and to what extent the population is able to pay that price, defined in Eq. 2). The demand variable ( $D_{mit}$ ) is obtained through polynomial relation with the corresponding price ( $ps_{mit}$ ), and  $B_{mit}$  defines the lower bound of demand for the minimum feasible price (Eq.3).

$$Dm_{mit} = \beta_{mit} (pm_{mit})^{\epsilon_{mi}} \quad (2) \leftrightarrow \ln(Dm_{mit}) = \ln(\beta_{mit}) + \epsilon_{mi} \ln(pm_{mit})$$

And a linear relation is achieved  $\begin{matrix} \longleftrightarrow \\ D_{mit} = \ln(Dm_{mit}); \\ B_{mit} = \ln(\beta_{mit}) \\ ps_{mit} = \ln(pm_{mit}) \end{matrix} D_{mit} = B_{mit} + \epsilon_{mi} ps_{mit} \quad (3)$

### 2.2. Economic and Environmental assessment

The economic assessment is made through the maximization of the Net Present Value (NPV), which corresponds to the sum of the discounted cash flows of each time-period, at an interest rate. In turn, these cash flows are obtained through the net earnings ( $NE_t$ ), which are given by the difference between incomes and overall costs, where the former corresponds to the amount of products sold ( $X_{mai jt}$ ) at a certain price ( $ps_{mit}$ ), and the latter by the following costs. As seen in Eq. 4), these supply chain costs include raw material costs (first term), production operating costs (second term), storage costs (third term), transportation costs (fourth term), hub handling costs (fifth term), airline/freighter contracted costs (sixth term), inventory costs (seventh term), fixed and variable labour costs (eighth and ninth terms), and finally the labour costs for technologies' use (tenth term).

$$NE_t = (1 - tr) \left[ \sum_{\substack{(m,i,j) \in F_{INCFP} \\ (a,m,i,j) \in NetP}} ps_{mi} X_{mai jt} - \left( \sum_{\substack{(m,i,j) \in F_{OUTSUPRM} \\ (a,m,i,j) \in NetP}} rmc_{mi} X_{mai jt} + \sum_{\substack{(m,g) \in H_{prod} \\ i \in I_f}} opc_g P_{mgit} \right. \right. \\ + \sum_{\substack{(m,g) \in H_{stor} \\ i \in (I_f \cup I_w)}} opc_g S_{mgit} + \sum_{\substack{(a,m,i,j) \in NetP \\ \alpha \in (A_{plane} \cup A_{boat} \cup A_{trucks})}} tc_a \cdot pw_m \cdot d_{ij} \cdot X_{mai jt} + \sum_{\substack{(a,m,i,j) \in NetP \\ (j \in I_{plane/boat} \wedge i \in I_{plane/boat})}} hhc_j \cdot X_{mai jt} \\ + \sum_{i \in (I_{plane} \cup I_{boat})} cfp_i \cdot Y_i + \sum_{\substack{(m,g) \in H_{stor} \\ (m,i) \in V}} sc_m S_{mgit} + \sum_{i \in (I_f \cup I_w)} w_i \cdot lc_i \cdot wwh \cdot wpt \cdot Y_i \\ \left. \left. + \sum_{i \in (I_f \cup I_w)} wpsq \cdot lc_i \cdot wwh \cdot wpt \cdot YC_i + \sum_{\substack{(m,g) \in H \\ i \in I_f}} w_g \cdot lc_i \cdot wwh \cdot wpt \cdot Z_{gmi} \right) + tr \cdot DP_t \right] \quad (4)$$

The environmental objective function is defined using the Life Cycle Analysis (LCA) methodology, where the minimization of the environmental impact of production, storage, transportation, and entity installation is calculated for each midpoint category  $c$ , summed, and normalized as represented in Eq. 5).

$$\begin{aligned} \min EnvImpact = \sum_c \eta_c \left( \sum_{\substack{t \in T \\ (m,g) \in H}} e i_{mgc} p w_m P_{mgt} + \sum_{\substack{t \in T \\ (m,g) \in H}} e i_{mgc} v p u_m S_{mgt} \right. \\ \left. + \sum_{\substack{t \in T \\ (a,m,i,j) \in H}} e i_{ac} p w_m d_{ij} X_{mai jt} + \sum_{\substack{t \in T \\ (a,m,i,j) \in H}} e i_{ic} Y C_i \right) \end{aligned} \quad (5)$$

### 3. Case-Study

The analysed case-study refers to the production and distribution of a meningitis conjugate vaccine by a specific pharmaceutical company. The vaccine in study is characterized by significant use outside of routine immunization. Low- and middle-income countries do not purchase this vaccine due to its price. Moreover, countries in Africa and Eastern Mediterranean Region indicated an interest in introducing this vaccine into routine immunization programs but have struggled to access these vaccines due to the high price and limited availability (WHO, 2019). This pharmaceutical company's supply chain is currently composed of suppliers and factories, located in Europe and U.S., warehouses located in Europe, U.S and Australia, airports, seaports, and five markets, which are Europe, U.S., Australia, Africa and Middle East. A possible location for a new factory in Africa is included in the study to evaluate the possibility of increasing accessibility for this vaccine in this market. The Global Burden of Disease Results Tool is used to obtain the DALY Rate per 100.000 population for meningitis disease in each geographical area of the corresponding markets. Moreover, these markets have different demand and prices for this vaccine. Through the database MI4A/VP3 (Market Information for Access to Vaccines/Vaccine Product, Price and Procurement) published by WHO in 2022, information on demand and prices for the vaccine in study can be obtained. Based on the available data, demand to price elasticity coefficients for each market were estimated. As previously explained, demand is a decision of the model and its obtained value can depend on price variations, introduced by the demand to price elasticity coefficient. In the developed case-study the elasticity coefficient for the vaccine being distributed and sold are -5.81, -4.94, -3.45, -0.69, -0.05 for Europe, U.S., Australia, Middle East, and Africa, respectively. The first three markets are characterized by a more elastic demand-price elasticity, while Middle East and Africa are inelastic.

In the developed model, apart from demand, other decisions are made on facility location (suppliers, factories, warehouses and hub terminals), production and storage levels as well as transportation network established. Additionally, demand values for each market are also a decision of the model, whose values are influenced by price variations.

### 4. Results and Discussion

The results focus on three cases studies (A, B and C) over two scenarios (1 and 2) where a lexicographic optimization was used. Case A corresponds to the optimum economic performance. Case B corresponds to the optimum social performance and Case C to the optimum environmental performance. In both scenarios, Cases A, B and C were studied to explore the trade-off between each sustainability pillar. In the first scenario 1) no price variations are introduced, while the second scenario 2) considers an increase of 2% in the vaccine's price in U.S., Australia, and Europe markets, whereas a 15% decrease is applied to markets in Middle East and Africa. These price variations are introduced in the third year of a planning period of 10 years.

The obtained supply chain decision for each case and scenario are shown in Table 1, from where it is possible to observe that the same supply chain structure was obtained for all cases and scenarios. However, depending on the case and scenario, different decisions are taken, resulting in different performances for each sustainability objective. Table 1 summarizes the major results regarding the decisions taken on facility location and transportation network.

Table 1. Supply chain topology and performance indicators' results comparison

		Case A	Case B	Case C
Scenarios 1 and 2	Suppliers and factories location	U.S., Africa	U.S., Africa, Europe	U.S., Africa, Europe
	Transport Network	Trucks of big capacity, Intercontinental connections by boat (4 seaports)	Trucks of big capacity, Intercontinental connections by boat (3 seaports)	Trucks of big capacity, Intercontinental connections by boat (4 seaports)
Scenario 1	NPV (€)	1.832E+09	1.810E+09	1.807E+09
	PharmAccess	3.740E+03	3.798E+03	3.798E+03
	EnvImpact	3.274E+06	3.274E+06	3.271E+06
Scenario 2	NPV (€)	1.723E+09	1.711E+09	1.707E+09
	PharmAccess	3.740E+03	3.798E+03	3.798E+03
	EnvImpact	3.268E+06	3.274E+06	3.264E+06

Case A corresponds to the most profitable solution but worst social and environmental performance. In Cases B and C, a small profit decrease of 1% and 2% is observed, but the best social and environmental performance, respectively are obtained. The small variation between cases is justified by the high amount of demand being allocated to only two markets: 54% of the total demand corresponds to U.S market, 45% to Middle East with the remaining percentage of demand corresponding to the other three markets. Moreover, African market has less than 1% of the total demand, possibly because the price of the meningitis vaccine considered is not affordable for the market, leading to very low demand. The performance of the social indicator is directly related to the decision of entities' locations. From Case A (best economic performance) to Cases B and C, the three factories located in Europe, Africa, and U.S are used for production, whereas in Case A only the factory in U.S. and Africa are used. This is explained by the more social beneficial solution being considered in Case B where the best performance of this indicator is achieved by deciding to allow all the three factories to open and contribute to the total value of the social factor  $p_i^{DALY}$ . In case C, since the environmental indicator has the highest priority a 12% decrease in the environmental impact is achieved by opening all facilities in order to reduce transportation between entities.

Table 2. Comparison between Scenario 1 and Scenario 2

Market	Elasticity coefficient $\epsilon_{mi}$	Price variations	Scenario 1 Contracted Demand (SKU*)	Scenario 2 Price variations Demand (SKU*)
Europe	-5.81	↑ 2%	2000	$t \leq 2 : 2000$ ; $t > 2 : 3740$
U.S.	-4.94	↑ 2%	992965	$t \leq 2 : 992965$ ; $t > 2 : 899771$
Australia	-3.45	↑ 2%	3616	$t \leq 2 : 3616$ ; $t > 2 : 3356$
Middle East	-0.69	↓ 15%	831400	$t \leq 2 : 831400$ ; $t > 2 : 918697$
Africa	-0.05	↓ 15%	1200	$t \leq 2 : 1200$ ; $t > 2 : 1210$

\*Each product stock keeping unit (SKU) contains 5 single-doses of meningitis vaccines



In Scenario 2, where price variations are introduced in the third year of the planning period, different demand responses depending on the market are obtained, as depicted in Table 2. In the markets where a price increase of 2% was introduced, a demand reduction of 11.8% in Europe, 9.4% U.S., 7.2% in Australia is obtained. On the other hand, a demand increase of 10.5 % in Middle East and 0.75% in Africa is obtained caused by a price reduction of 15%. Even though the effect of the same variation results in a consistent behaviour among markets, the magnitude of these responses is different. This is due to the different demand-price elasticities considered among markets. Europe, U.S., and Australia have elastic demand-price elasticities, resulting in great sensibilities to price variations. This can be explained by the existence of substitute vaccines produced by other companies in these three markets. Middle East and Africa do not experience the same, meaning that in these two markets there is no other company producing and distributing a substitute product for this vaccine, which results in a more inelastic relation between demand and price. The markets' response to price variations results in a 6% decrease in the NPV for the company. Therefore, it is important to explore trade-offs, not only between sustainability pillars, but also between the amount of products and price that a company can practice, and the amount of products that each market needs, at a price that is affordable for them.

## 5. Conclusions and Future work

This work presents an optimal approach for the design and planning of sustainable pharmaceutical supply chains, where the social component is the focus. Through the application of the model, it is possible to understand the existence of trade-offs between each sustainability pillar, and opportunities can be explored envisioning the improvement of availability and affordability of medicines. The concepts of demand-price elasticity are modelled to analyse the affordability of pharmaceutical products as in countries where the need is the highest this can be improved by analysing each market's response to price variation. Despite the promising results achieved in the present work, further research is required. In particular, it is crucial to study the possibility of including lower- and middle-income countries with significant lower prices when compared to the prices practiced in the HIC and UMIC studied in the present work.

## Acknowledgements

The authors acknowledge the support provided by FCT under the project PTDC/EME-SIS/6019/2020 and UIDB/00097/2020, and PhD Grant 2022.11365.BD.

## References

- Amaro, A. C. S., & Barbosa-póvoa, A. P. F. D. (2009). *The effect of uncertainty on the optimal closed-loop supply chain planning under different partnerships structure*. 33, 2144–2158. <https://doi.org/10.1016/j.compchemeng.2009.06.003>
- Duarte, I., Mota, B., Pinto-Varela, T., & Barbosa-Póvoa, A. P. (2022). Pharmaceutical industry supply chains: How to sustainably improve access to vaccines? *Chemical Engineering Research and Design*, 182, 324–341. <https://doi.org/10.1016/j.cherd.2022.04.001>
- Menou, V., Hornstein, A., & Lipton-McCombie, E. (2021). *Access to Medicine Index 2021*. 1–72. [http://www.accesstomedicineindex.org/sites/www.accesstomedicineindex.org/files/publication/2008\\_Index\\_Report.pdf](http://www.accesstomedicineindex.org/sites/www.accesstomedicineindex.org/files/publication/2008_Index_Report.pdf)
- United Nations. (2015). *UN General Assembly, Transforming our world: the 2030 Agenda for Sustainable Development, 21 October 2015, A/RES/70/1, available at: https://www.refworld.org/docid/57b6e3e4.html [accessed 28 November 2019]. 16301(October), 1–35.*

# Optimization of Decentralized Pine Needle-to-Electricity Supply Chain for Almora District of Uttarakhand, India

Gayatri Gawande,<sup>a</sup> Ankush Halba,<sup>b</sup> Pratham Arora,<sup>b,\*</sup>

<sup>a</sup>*Department of Mechanical Engineering, National Institute of Technology Tiruchirappalli, Tiruchirappalli-620015, India*

<sup>b</sup>*Hydro and Renewable Energy Department, Indian Institute of Technology Roorkee, Roorkee-247667, India*

*\*pratham.arora@hre.iitr.ac.in@Indian Institute of Technology Roorkee, Roorkee, India.  
+91-1332284921*

## Abstract

Approximately  $3.4 \times 10^4$  km<sup>2</sup> of forest land in Uttarakhand is susceptible to wildfires, with an economic loss of around 1200 USD per hectare due to dry and flammable Pine needles. The Chir Pine occupies 26% of the state's forest, producing  $1.67 \times 10^6$  tonnes of Pine needles annually. These Pine needles could be utilized in a gasifier system to provide electricity and solid cooking fuel to people residing in remote and rural areas of Uttarakhand. However, the short window for collection (100 days), inefficient supply chain, and high transportation costs pose significant challenges for implementing any gasifier-based project. The current study involves finding locations for these gasifier-based plants and optimizing the supply chain to reduce cost, distance, and CO<sub>2</sub> emissions. Economic analysis for ten different plant capacities ranging from 25 to 250 kW showed that the estimated cost ranged from  $33 \times 10^3$  to  $146 \times 10^3$  USD.

**Keywords:** Supply chain, Optimization, Gasification, Economic analysis, Electricity.

## 1. Introduction

The Uttarakhand state of India has an abundant Pine Forest, which is responsible for forest fires during the summer season because the leaves of Pine (Pine needles) are susceptible to catching fire due to high temperatures and low humidity levels in summer. Bisht and Thakur (2022) have suggested the utilization of Pine needles in a decentralized gasifier to produce electricity, which could provide electrification in the rural and remote regions of Uttarakhand state. They also looked at the roadblocks to the Pine needle-based gasifier project, discovering that an inefficient supply chain is the most significant obstacle. An efficient biomass supply chain is required to improve the utilization and conversion of biomass into bioenergy. The three most common components of the biomass supply chain are harvesting, transportation, and storage. Razak et al. (2021) optimized the supply chain for bioethanol production from Palm oil via integrated GIS-AHP for Malaysia. However, they have only considered the minimum travel distance parameter to minimize the logistic cost and GHG emissions. Fan and Klemeš (2020) used the Pinch Analysis based on the months to optimize the biomass supply chain for different types and amounts of biomass. They observed that each biomass has a different harvesting period, making the biomass supply chain complex. Durmaz and Bilgen (2020) have used multi-objective optimization for sustainable biomass supply chain design to produce biogas from poultry farms. They

proposed a three-stage sequential solution to maximize profit and minimize the distance between the biogas facility and poultry farms. However, the long payback period and the high operational cost is the main hindrance to the biogas-based project.

To the best of the authors' knowledge, no significant research on Pine needle-based supply chain optimization for a gasification plant in India has been carried out, including the biomass pretreatment scenario such as shredding, pelletizing, and briquetting. In light of this, the present study aims to optimize the supply chain to reduce cost, distance, and CO<sub>2</sub> emissions for various densification and non-densification scenarios in the Indian context.

## 2. Methodology

The Pine needle to electricity supply chain optimization was performed to reduce the transportation distance, costs, and CO<sub>2</sub> emissions by considering various densification scenarios. Shredding, pelletizing, and briquetting are considered to densify Pine needles. The flow chart of densification routes is presented in Figure 1. Baling, bundling, module construction, boxing, silo-bagging, packing, pelletizing, and briquetting or cubing are all examples of biomass densification (Ibitoye et al., 2021). Local merchants provided the prices for raw Pine needles and densified Pine needles. The average amount of Rupees 1 equals 0.0134 USD based on the market currency exchange rate in 2021.

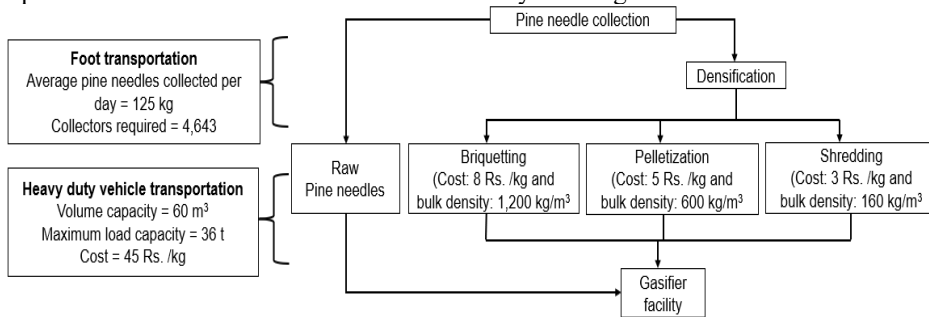


Figure 1: Densification routes

Data regarding the Pine needle availability for the forest blocks in the 13 districts of Uttarakhand is taken from the Forest department of the Uttarakhand Government (Forest Department in Uttarakhand, 2021). The data gives information about a total of 8,922 points yielding  $1.67 \times 10^6$  t of Pine needles per annum from the forest compartments having a geographical area of  $1.03 \times 10^4$  km<sup>2</sup>. The flow chart of the Pine needle-based gasifier project for the Almora district of Uttarakhand is presented in Figure 2.

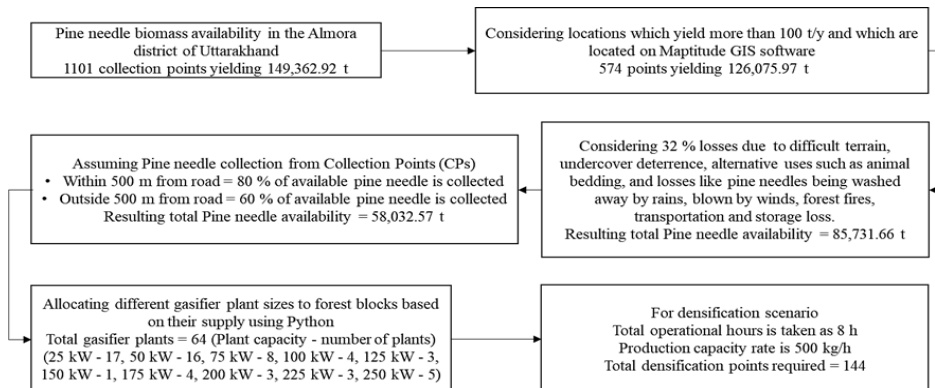


Figure 2: Flow chart of Pine needle-based gasifier project

### 3. Results and Discussion

Maptitude Geographical Interface System software was used to determine locations for densification centers and gasifier facilities based on the collection point candidates set. The facility allocation system employs Greenfield Analysis (GFA or "center of gravity analysis") to identify the locations for densification centers and the gasifier facilities. The results of Maptitude GIS software are presented in Figure 3.

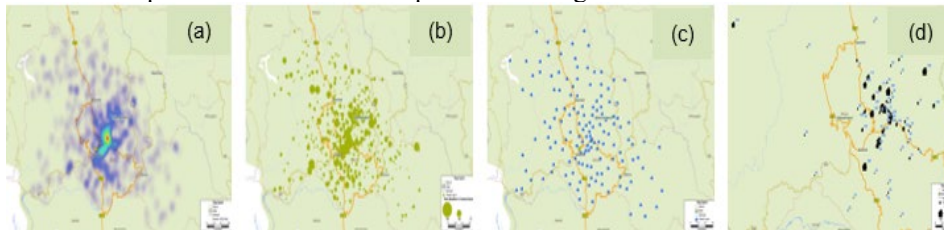


Figure 3: Results of Maptitude software for (a) Pine needle density map; (b) Pine needle collection points; (c) densification points; (d) gasifier facility points for Almora district

Using the Routing and Direction Tool of Maptitude GIS Software, straight-line distances are calculated considering Collection points as the origin/supply nodes and densification points as the destination/demand nodes, as shown in Figure 4.

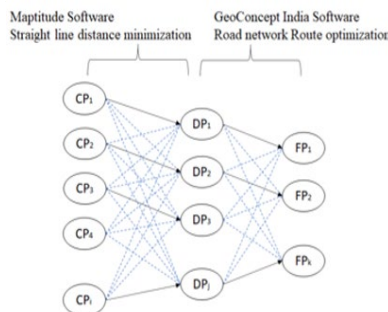


Figure 4: Transportation network flow

Foot transportation is done by Pine needle collectors from CPs to DPs in the hilly forest area of Uttarakhand, where there is a lack of a proper road network, and straight-line distances are considered. A matrix  $[X_{ij}]$  is generated for straight-line distances that the Pine needle collector has to travel from each collection point to each densification point. The average distances a Pine needle collector travels for the first 5, 10, 15, 20, 25, and 30 closest points from the supply nodes, CP<sub>i</sub> to the demand nodes DP<sub>j</sub> are calculated to be 2.9 km, 4.2 km, 5.3 km, 6.2 km, 7 km, 7.7 km. The relationship between the nth closest point and the average distance is derived using linear regression fit as  $y$  (km) = (0.207166  $x$  + 1.91276) (km) with the R squared value of 0.9733, where  $y$  denotes the average distance and  $x$  denotes the  $n^{\text{th}}$  closest distance. The route optimization results from GeoConcept India for the average distance a truck travels in the case of a Pine needle gasifier plant of 10 different capacities are depicted in Figure 5. On average, 125 kg of Pine needles per day is collected by a collector, so approximately 4,643 collectors would be required. Heavy-duty vehicle/ truck with a volume capacity of 60 m<sup>3</sup>, maximum load capacity of 36 tonnes, and cost of 45 Rs./km is considered for transportation from the DP to FP. Figure 6 depicts the transportation costs for the ten different gasifier facilities in the case of different densification scenarios based on the number of trips the truck has to make to transport the required Pine needle biomass in the 100-day window.

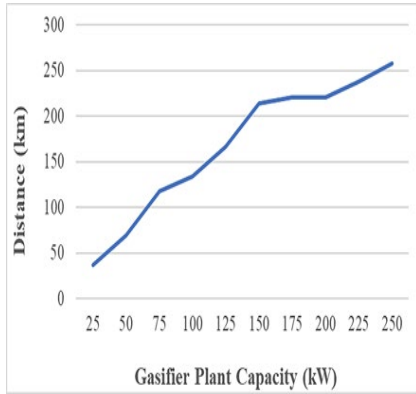


Figure 5: Average distance for heavy-duty vehicle/truck transportation

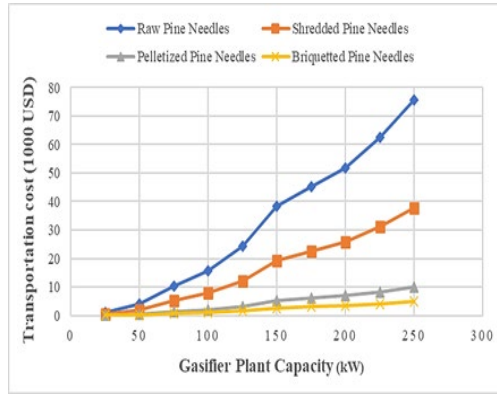


Figure 6: Transportation costs for densification scenarios

The Pine needle consumption is 1.5 kg/kWh, and the auxiliary power consumption is taken as 5%. The electricity sent out after auxiliary consumption is assumed to be sold at 8 Rs./kWh, and the charcoal produced (10% of Pine needle consumed) as a by-product from gasification is assumed to be sold at 20 Rs./kg as a cooking fuel (Renewable Energy Government of Uttarakhand, 2018). The procurement cost for raw Pine needles is considered as 1.5 Rs./kg, shredded Pine needles as 3 Rs./kg, pelletized Pine needles as 5 Rs./kg, and briquette Pine needles as 8 Rs./kg. Local sellers provided the prices for raw Pine needles and densified Pine needles. The cost model, including revenue, procurement, capital, and operational and maintenance costs, is considered for the gasifier plant, as presented in Figure 7.

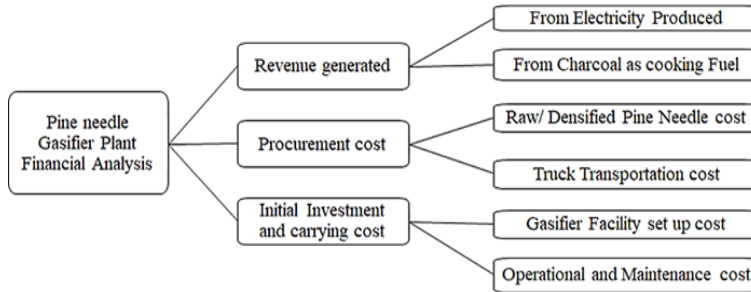


Figure 7: Cost model of Pine needle-based Gasifier plant

The cost of 25 to 250 kW was calculated using Eq. (1), ranging from 33 x 10<sup>3</sup> to 146 x 10<sup>3</sup> USD. The annual capacity factor was calculated to be 0.64, considering actual annual energy generation of 140,160 kWh and for a 25 kW Pine needle gasifier plant, which is operational for 16 hours. The annual profit/loss and return on investment (ROI) for raw, shredded, pelletized, and briquette Pine needles are presented in Figure 8.

$$\frac{C_2}{C_1} = \left(\frac{Q_2}{Q_1}\right)^x \tag{1}$$

Where, C = Facility cost, Q = Plant capacity, and x = Annual capacity factor.

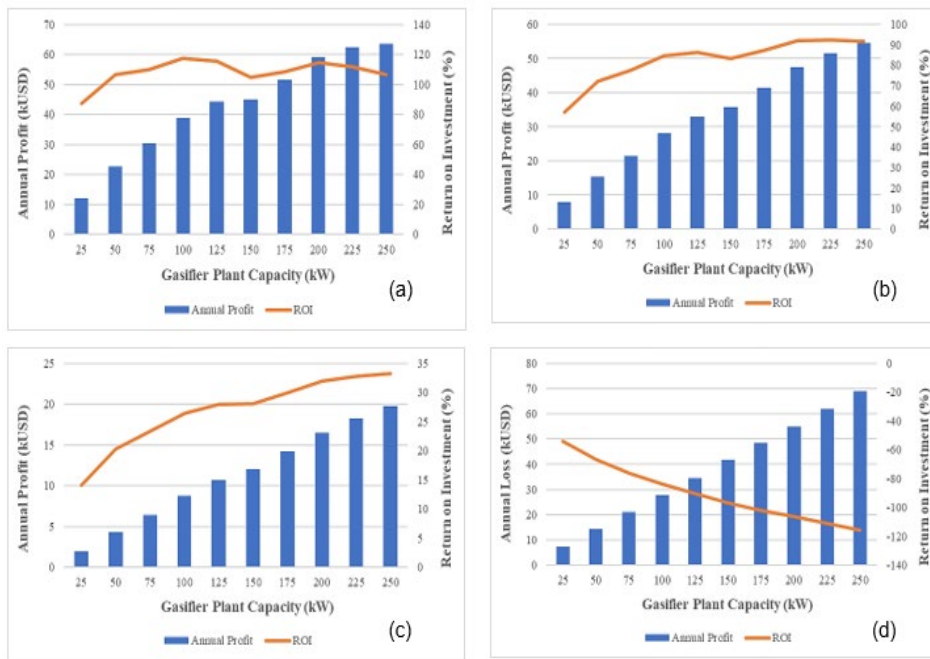


Figure 8: Annual profit/ loss and return on Investment (ROI) for (a) raw Pine needles scenario; (b) shredded Pine needles scenario; (c) pelletized Pine needles scenario; (d) briquette Pine needles scenario

The Almora district comes under the Northern Indian electricity grid, where the GHG emissions are estimated as 1.24 kg CO<sub>2</sub> eq/kWh. In contrast, GHG emissions from electricity generation through Pine needle gasification are 0.132 kg CO<sub>2</sub> eq/kWh (Briones-Hidrovo et al., 2021). The total GHG emissions from the conventional Northern Indian grid and Pine needles gasification are estimated as 41,537.82 tonnes of CO<sub>2</sub> eq and 4,421.77 tonnes of CO<sub>2</sub> eq for the 33.5 × 10<sup>6</sup> kWh of electricity generation in the Almora district. The CO<sub>2</sub> emissions for the raw, shredded, pelletized, and briquette Pine needles scenario are calculated to be 419.9, 209.9, 55.9, and 27.9 tonnes based on the emission factor of 0.7375 kg CO<sub>2</sub>/km for HDV in India as per the recommendation of India GHG Program 2015.

#### 4. Conclusions

The CO<sub>2</sub> emissions from truck transportation are based on the number of vehicles and trips needed to supply the biomass to the gasifier facility in 100 days time window. A total of 64 Gasifier facilities comprising different capacities can be located in Almora based on the Pine needle biomass supply from different forest blocks. The total electricity produced by the 64 gasifier plants would be 33.5 × 10<sup>6</sup> kWh, and the total capital cost would be 44 × 10<sup>6</sup> USD. Among all the costs, procurement of raw/densified Pine needles accounts for 80% of the total costs, transportation accounts for 15%, and O&M accounts for 5% of the total costs. All scenarios except briquetting are profitable due to the high procurement cost of Pine needle briquettes at 0.104 USD/kg. The payback period for the raw Pine needles scenario ranged from 0.85-1.14 years, 1.08-1.76 years for the shredding scenario, and 3-7 years for the pelletizing scenario. The potential annual reduction in CO<sub>2</sub>

emissions in Almora by generating electricity from a Pine needle biomass-based gasifier compared to the conventional grid is 37,116 tonnes of CO<sub>2</sub> eq.

## References

- A. S. Bisht, N. S. Thakur, 2021, Identification & Prioritisation of Barriers in the Growth of Pine Needle Biomass Gasification Plants (<250 kW) for Electricity Generation in the Western Himalayan Region: Uttarakhand, India, Process Integration and Optimization for Sustainability, 6, 37-60. <https://doi.org/10.1007/s41660-021-00199-y>.
- A. Briones-Hidrovo, J. Copa, L. A. Tarelho, C. Gonçalves, T. P. da Costa, A. C. Dias, 2021, Environmental and energy performance of residual forest biomass for electricity generation: Gasification vs. combustion, Journal of Cleaner Production, 289, 125680. <https://doi.org/10.1016/j.jclepro.2020.125680>.
- Y. G. Durmaz, B. Bilgen, 2020, Multi-objective optimization of sustainable biomass supply chain network design, Applied Energy, 272, 115259. <https://doi.org/10.1016/j.apenergy.2020.115259>.
- V. Y. Fan, J. J. Klemeš, 2020, Biomass supply and inventory management for energy conversion, Chemical Engineering Transactions, 78, 421-426. <https://doi.org/10.3303/CET2078071>.
- Forest Department in Uttarakhand, 2021, Pine Needle Abundance Area in Uttarakhand <[www.forest.uk.gov.in/pirul-abundance-areas](http://www.forest.uk.gov.in/pirul-abundance-areas)>, accessed on 30.10.2022.
- S. E. Ibitoye, T. C. Jen, R. M. Mahamood, E. T. Akinlabi, 2021, Densification of agro-residues for sustainable energy generation: an overview, Bioresources and Bioprocessing, 8, 75. <https://doi.org/10.1186/s40643-021-00427-w>.
- ISFR, 2021, State of Forest Report 2021, Forest Survey of India <[www.fsi.nic.in/forest-report-2021](http://www.fsi.nic.in/forest-report-2021)>, accessed on 1.04.2022.
- M. Kumar, A. Kumar, R. Kumar, B. Konsam, N. A. Pala, J. A. Bhat, 2021, Carbon stock potential in Pinus roxburghii forests of Indian Himalayan regions, Environment, Development and Sustainability, 23, 12463-12478. <https://doi.org/10.1007/s10668-020-01178-y>.
- MoEFCC, 2018, National Action Plan on Forest Fire <[www.forests.tn.gov.in/app/webroot/img/NAPFF-Final%20Draft%202023.04.2018.pdf](http://www.forests.tn.gov.in/app/webroot/img/NAPFF-Final%20Draft%202023.04.2018.pdf)>, accessed on 19.10.2022.
- N. H. Razak, H. Hashim, N. A. Yunus, J. J. Klemeš, 2021, Integrated GIS-AHP optimization for bioethanol from oil palm biomass supply chain network design. Chemical Engineering Transactions, 83, 571-576. <https://doi.org/10.3303/CET2183096>.
- Renewable Energy Government of Uttarakhand, 2018, Pine Needle Power Generation, Uttarakhand <[https://ureda.uk.gov.in/departament2/library\\_file/file-17-07-2021-10-54-40.pdf](https://ureda.uk.gov.in/departament2/library_file/file-17-07-2021-10-54-40.pdf)>, accessed on 08.10.2022.
- C. Venkataraman, G. Habib, D. Kadamba, M. Shrivastava, J. F. Leon, B. Crouzille, O. Boucher, D. G. Streets, 2006, Emissions from open biomass burning in India: Integrating the inventory approach with high-resolution Moderate Resolution Imaging Spectroradiometer (MODIS) active-fire and land cover data, Global biogeochemical cycles, 20(2). DOI: 10.1029/2005GB002547.

# Life Cycle Assessment of Two Liquid Organic Hydrogen Carriers

Camille Bontron,<sup>a</sup> Diogo Rodrigues,<sup>b</sup> Catarina G. Braz,<sup>b</sup> Henrique A. Matos<sup>b</sup>

<sup>a</sup>*INP-ENSIACET, 4 All. Emile Monso, CS 44362, 31030 Toulouse Cedex 4, France*

<sup>b</sup>*Centro de Recursos Naturais e Ambiente, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal*  
*dfmr@tecnico.ulisboa.pt*

## Abstract

Liquid organic hydrogen carriers (LOHCs) can store H<sub>2</sub> through hydrogenation and dehydrogenation processes. Two systems are based on Toluene-methylcyclohexane (MCH) and dibenzyltoluene-perhydro-dibenzyltoluene (DBT-PHDBT). In this study, the life cycle assessment (LCA) methodology is applied to the two systems. The goal is to determine which molecule is the most suitable one. The general system comprises hydrogenation in Sines (Portugal), transport, and dehydrogenation in Rotterdam (Netherlands). The H<sub>2</sub> released during dehydrogenation is the functional unit. The amounts of waste, makeup, cooling water, electricity, heat, and transported quantity are calculated. SimaPro is used as the LCA tool, considering three impact assessment methods, IMPACT 2002+, ILCD MidPoint 2011, and ReCiPe MidPoint (H). The system based on DBT-PHDBT is found to have a smaller impact for all methods. In particular, the process based on Toluene-MCH emits more than twice as much CO<sub>2</sub> eq/kg of H<sub>2</sub>. Energy efficiency improvements and renewable energy sources should be considered.

**Keywords:** Liquid organic hydrogen carriers, Hydrogenation, Dehydrogenation, Life cycle assessment.

## 1. Introduction

Liquid organic hydrogen carriers (LOHCs) can be used for long-term storage and long-distance transport of hydrogen. Hydrogen is stored in the organic molecule through a hydrogenation process and then unloaded via the reverse process of dehydrogenation. These organic molecules allow hydrogen to be handled at ambient conditions and transported using the existing crude oil-based global supply chain. In a previous study by Carvalho (2020), suitable organic molecules were reviewed, and two carrier systems based on Toluene-methylcyclohexane (MCH) and dibenzyltoluene-perhydro-dibenzyltoluene (DBT-PHDBT) were chosen. Toluene-MCH was selected as it is well studied and DBT-PHDBT is convenient for handling and safe. For both molecules, a techno-economic analysis was performed, which is the basis for this study. In this study, the life cycle assessment (LCA) methodology is applied to the two LOHC systems to evaluate and compare the corresponding processes. SimaPro is used as the LCA tool to determine the processes' impact on the resources, environment, and human health factors.

## 2. Context

The goal of this study is to compare the processes based on Toluene-MCH and on DBT-PHDBT and to determine which molecule is most suitable to be used as a LOHC. The



Ecoinvent 3.6 database is used for the background data assuming the Portuguese context for the hydrogenation process, and the Netherlands for the dehydrogenation process. First, the boundaries of the system must be determined. The boundaries are gate to gate, meaning only hydrogen, LOHC makeup and purge will be used as material flows in this study. The general system used for both molecules is represented in Figure 1. The functional unit in this system is the release of 1 kg of hydrogen during dehydrogenation.

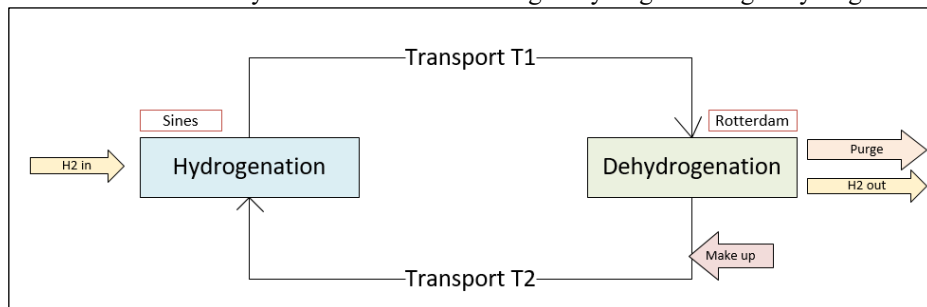


Figure 1 – General system for LCA study.

The general system comprises the hydrogenation process, located in Sines (Portugal), transport, and dehydrogenation in Rotterdam (Netherlands). The system used to transport the hydrogen is a two-way carrier. The dehydrogenated molecule is transported back to the hydrogenation process with the added makeup of toluene or DBT.

The first LOHC molecule studied is Toluene-MCH. It can transport 3 molecules of hydrogen. The other LOHC molecule studied and compared to Toluene-MCH is DBT-PHDBT. It can transport 9 hydrogen molecules.

The dehydrogenation processes release waste via a purge flow. The “Cut off” approach is used, meaning that the creator of pollution is the one responsible for it. Indeed, the pollutant molecules in the purge flow will be accounted for in the LCA.

The reactions need energy to be carried out. Electricity and heat exchangers are used for both the hydrogenation and dehydrogenation processes, and the amount needed comes from Carvalho (2020). For the Toluene-MCH and DBT-PHDBT systems, both the hydrogenation process and the dehydrogenation process may need fired heat, low-pressure steam (LPS), and high-pressure steam (HPS). Cooling water is also needed for both hydrogenation and dehydrogenation processes.

The LOHCs need to be transported from Sines (Portugal) where the hydrogenation takes place to Rotterdam (Netherlands) for dehydrogenation. Transport by ship is used as transport by rail or by road consumes more energy and produces more pollution. The maritime route is extracted from Godinho (2021).

### 3. LCA Inventory

To be used in SimaPro, DBT synthesis data had to be entered, as it was not initially in the database. To create this molecule, the RREM method is used (Huber et al., 2022), which consists of four steps: general research on the chemical process, setting up the reaction equation, researching the required thermal energy, and modeling of the dataset.

First, DBT has a molar mass of 272.38 g/mol. The reaction between toluene and benzyl chloride creates DBT and hydrogen chloride. We suppose the only side products are reactants that did not react. These reactants are recycled.

The production of DBT is 660 g, which represents 2.423 mol. The catalyst used is iron chloride. For catalyst and solvent, a 50% recycling rate is assumed. For 660 g of products,

50 g of catalyst is needed. With that information, and knowing there is a 70% yield, the toluene makeup and benzyl chloride makeup streams can be calculated. The hydrogen chloride waste is also calculated. As the energy for this reaction was not found in the literature, it is assumed that 1 kg of product needs 1.8 MJ to be created (Huber et al., 2022). Therefore, for 660 g of DBT produced, 1.188 MJ is needed.

The transport used is by ship, selected as “Transport, freight, sea, tanker for liquid goods other than petroleum and liquefied natural gas {GLO} | Cut-off, S”. The impact of transport considers the quantity transported in both ways and the distance between Sines and Rotterdam by boat, which is 2537.24 km according to Godinho (2021).

As mentioned, both processes need electricity. Toluene hydrogenation also consumes HPS and MCH dehydrogenation consumes fired heat and LPS. PHDBT dehydrogenation also consumes fired heat. In SimaPro, those parameters are represented with heat from steam (called “heat, from steam, in chemical industry (RER)”) or from natural gas (called “heat, district or industrial, natural gas (RER)). The data are taken from Carvalho (2020). On the other hand, DBT hydrogenation consumes fired heat and generates high-pressure steam (HPS). HPS is an avoided product as it is not directly used by the DBT/PHDBT process but can be used for other applications.

Cooling water is used during both hydrogenation and dehydrogenation processes. The used water can be recycled, but a makeup flow is needed to compensate the losses. In Carvalho (2020), the energy needed for cooling water is calculated. With that information, the makeup mass flow rate, which represents 2% of the cooling water mass flow rate, can be calculated (Vengateson, 2017). The water selected for the SimaPro simulation is “Water, cooling, unspecified natural origin, PT/NL” in the natural resources’ entry.

Regarding the different out flows, waste, purge, and gases out, several compounds are released from the process. In the Toluene-MCH process, xylene, methane, benzene, carbon dioxide, and nitrogen dioxide are released. The DBT synthesis creates a hydrogen chloride release and the DBT-PHDBT process releases benzene, benzyltoluene, toluene, cyclohexane and MCH. By adding the different output flows, it is possible to calculate the emissions for each molecule.

In the inputs, the makeup needs to be entered. For the Toluene-MCH process, the makeup represents the quantity of toluene that must be added so that the recycled toluene plus makeup is equal to the quantity of toluene input needed. The toluene makeup needed is 915.4 kg. For DBT/PHDBT, the makeup needed is DBT. In the same way, DBT makeup is calculated from the DBT output and DBT input and equals 134.1 kg. The quantity of hydrogen needed in the input comes from the mass balance and equals 6312.5 kg for the Toluene-MCH process and 6509.1 kg for the DBT-PHDBT process.

#### **4. Methodologies**

LCA relies on different methods using several impact categories. The main issues treated are human health impact, environmental impact, and resources impact. The goal of the study is to determine which LOHCs create fewer effects for the three main impact categories, based on three methodologies. The three methods are ReCiPe MidPoint (H), IMPACT 2002+ and ILCD MidPoint 2011, which are international reference life cycle data systems in the European Union. Those methods have different minor categories to classify and focus on different pollution sources. They were chosen as they are the most universal and complete methodologies. They are all based on long-term data.

## 5. Results and analysis

The data are calculated for 1 kg of products created. The goal of this part is to analyze and compare the results for the DBT/PHDBT and Toluene/MCH processes, so one can choose which one impacts the least the environment, human health, and resources. To this purpose, the total for each category of impact is calculated.

### 5.1. ReCiPe MidPoint (H)

The ReCiPe methodology considers environmental, human health and resource issues with the same importance. The main impacts created by both processes, shown in Figure 2, are on global warming, terrestrial ecotoxicity, human non carcinogens effects, and fossil resource scarcity. Regarding impact on resources, the Toluene/MCH process has a clearly higher impact than DBT/PHDBT. Indeed, the total value of fossil resource scarcity is 5.63 kg oil eq for the Toluene/MCH process and only 3.98 kg oil eq for the DBT/PHDBT system. Regarding human health impact, more precisely human non carcinogens category, the DBT/PHDBT process has a smaller impact than the Toluene/MCH process. Indeed, the impact for Toluene/MCH is approximately twice as much as the one for DBT/PHDBT. Finally, regarding environmental impacts, terrestrial ecotoxicity and global warming are both at least two times smaller for the DBT/PHDBT process compared to the Toluene/MCH process. Moreover, Toluene/MCH releases 8.39 kg CO<sub>2</sub> eq and DBT/PHDBT releases 3.42 kg CO<sub>2</sub> eq.

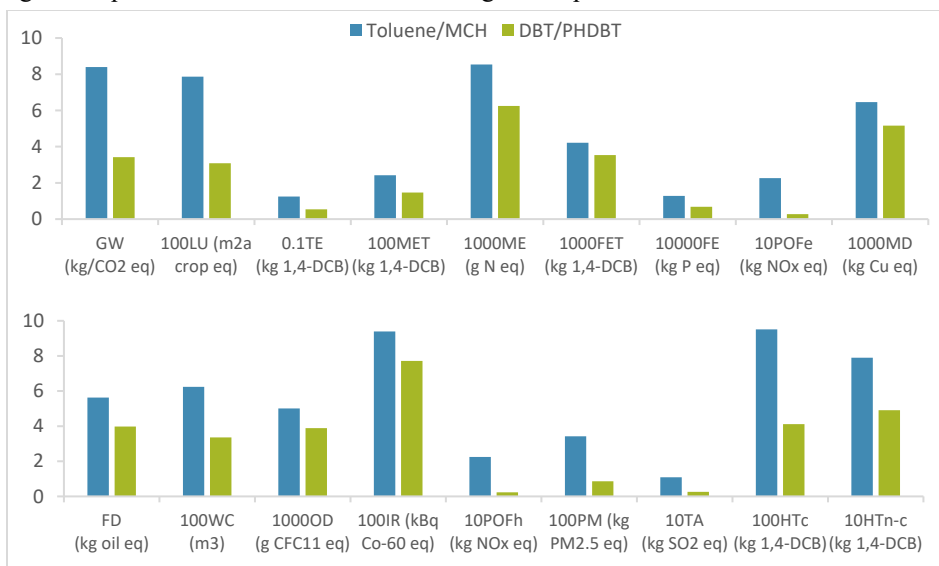


Figure 2 – Comparison of the impacts of the DBT/PHDBT and Toluene/MCH systems using the ReCiPe method. GW: global warming; LU: land use; TE: terrestrial ecotoxicity; MET: marine ecotoxicity; ME: marine eutrophication; FET: freshwater ecotoxicity; FE: freshwater eutrophication; POFe: photochemical ozone formation ecosystems; MD: mineral depletion; FD: fossil depletion; WC: water consumption; OD: ozone depletion; IR: ionizing radiation; POFh: photochemical ozone formation human health; PM: fine particulate matter formation; TA: terrestrial acidification; HTc: human carcinogenic toxicity; HTn-c: human non-carcinogenic toxicity.

### 5.2. Impact 2002+

With the IMPACT 2002+ method, one can observe in Figure 3 that the impact on human health is not significantly generated by the DBT/PHDBT process, in contrast to the Toluene/MCH process. Both processes have a significant impact on the environment and resources. One can compare the impact of each process on the different impact categories. DBT/PHDBT process mainly affects nonrenewable energy, respiratory inorganics, global warming, non-carcinogens effects and terrestrial ecotoxicity. Also, Toluene/MCH affects the same categories as well as terrestrial acid/nutrients. Regarding human health factors, the Toluene/MCH process has a significantly bigger impact on non-carcinogens effects and respiratory inorganics. For resource impacts, it is also the Toluene/MCH process which makes the largest impact as the total for nonrenewable energy is higher than with the DBT/PHDBT process. Finally, for environmental impact, one can see that the values for global warming, terrestrial ecotoxicity and terrestrial acid/nutrients are lower if the DBT/PHDBT process is used. Moreover, if one looks at the value of CO<sub>2</sub> equivalent for a kg of hydrogen, one can observe that, with the IMPACT 2002+ method, the Toluene/MCH process releases 7.92 kg CO<sub>2</sub>/H<sub>2</sub>, and the DBT/PHDBT process releases 3.21 kg CO<sub>2</sub>/H<sub>2</sub>.

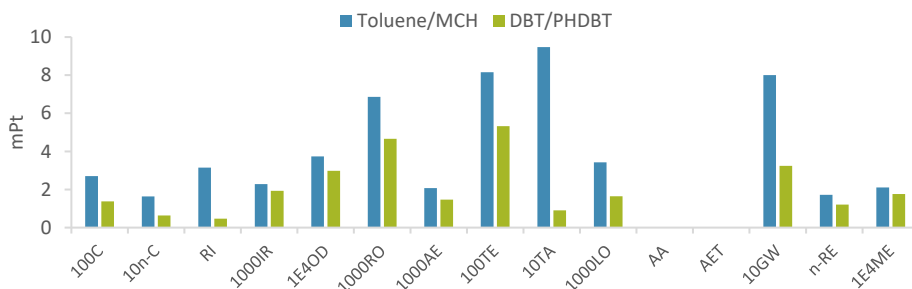


Figure 3 – Comparison of the impacts of the DBT/PHDBT and Toluene/MCH systems using the Impact 2002+ method. C: Carcinogens; n-C: Non-carcinogens; RI: Respiratory Inorganics; IR: Ionizing radiation; OD: Ozone layer depletion; RO: Respiratory organics; AE: Aquatic ecotoxicity; TE: Terrestrial ecotoxicity; TA: Terrestrial acidification/nutritification; LO: Land occupation; AA: Aquatic acidification; AET: Aquatic eutrophication; GW: Global warming; n-RE: Non-renewable energy; ME: Mineral Extraction.

### 5.3. ILCD MidPoint 2011

The results in Figure 4 show that this method does not focus on resources impact, for both processes. One can compare the impact of each process on the different categories, bearing in mind that the scale of Figure 4 is from 0 to 1 mPt, while the scale of Figures 2 and 3 is from 0 to 10 mPt. The resource impact is lower than the others. However, the impact on resources made by the Toluene/MCH process is always slightly higher than for DBT/PHDBT. Regarding human health factors, human toxic carcinogens impact is lower for DBT/PHDBT process than it is for Toluene/MCH, as well as human non carcinogens effects, particulate matter, and ionizing radiation HH. Hence, with the ILCD method, it appears that the DBT/PHDBT process is less harmful for human health than the one for Toluene/MCH. Finally, for the environmental impact, one can see that, once again, the Toluene/MCH process indicators are all higher than the ones for the other process. The sum of all impact categories is equal to 2.4 mPt for the Toluene/MCH process, and 0.59 mPt for the DBT/PHDBT process. Consequently, one can say that the DBT/PHDBT process has a lower impact than the one based on Toluene/MCH. Moreover, one can observe that the Toluene/MCH process releases 8.21 kg of CO<sub>2</sub>/kg of hydrogen while the one for DBT/PHDBT only releases 3.33 kg of CO<sub>2</sub>.

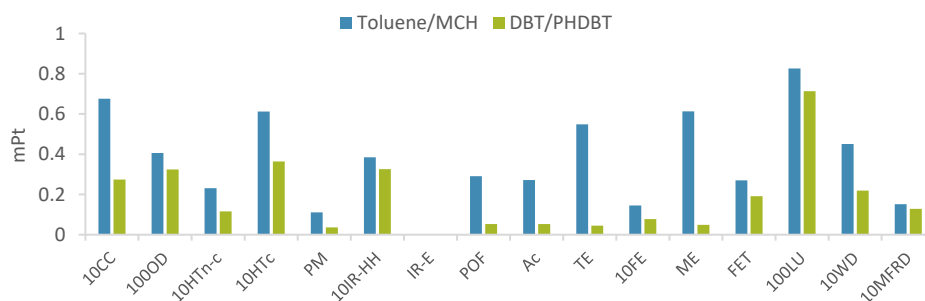


Figure 4 – Comparison of the impacts of the DBT/PHDBT and Toluene/MCH systems using the ILCD MidPoint 2011 method. CC: Climate change; OD: Ozone depletion; HTn-c: Human toxicity, non-cancer effects; HTc: Human toxicity, cancer effects; PM: Particulate matter; IR-HH: Ionizing radiation HH; IR-E: Ionizing radiation E (interim); POF: Photochemical ozone formation; Ac: Acidification; TE: Terrestrial eutrophication; FE: Freshwater eutrophication; ME: Marine eutrophication; FET: Freshwater ecotoxicity; LU: Land use; WD: Water resource depletion; MFRD: Mineral, fossil & ren resource depletion.

## 6. Discussion and conclusions

This study evaluates the impact of two systems for hydrogen storage and transport using two LOHCs, Toluene-MCH and DBT-PHDBT, on the resources, environment, and human health thanks to a LCA with the software SimaPro. For this study, the methods ReCiPe MidPoint (H), IMPACT 2002+ and ILCD MidPoint 2011 were considered. They were chosen as they are the most universal and complete methodologies.

Once the impacts are calculated by SimaPro for Impact 2002+ and ILCD MidPoint 2011 methods, a normalization is applied. After comparing the data, one can conclude that, for all methods, the DBT/PHDBT process appears to have a smaller impact on resources, environment, and human health factors than the Toluene/MCH process. In particular, the process based on Toluene-MCH emits more than twice as much CO<sub>2</sub> eq/kg of H<sub>2</sub> as the process based on DBT-PHDBT.

To conclude, some improvements could be made regarding data selection. Indeed, SimaPro does not allow the use of green hydrogen in the process. However, it is an important aspect since green hydrogen avoids using fossil energy or other non-renewable energies, which constitute a significant pollution source. The energy used for DBT synthesis could also be improved, as the value used is an approximation but could be more exact. Finally, to improve the impact of both processes, energy efficiency improvements and renewable energy sources should be considered, as both processes are energetically intensive. One could work on the heat exchanger and electricity values, as those units are the most consuming ones.

## References

- M. Carvalho, 2020, Liquid Organic Hydrogen Carriers (LOHCs): A review and techno-economic analysis, MSc Thesis
- J. Godinho, 2021, A greenhouse gas footprint and economic assessment of importing renewable H<sub>2</sub> from Portugal to the Netherlands, using liquid organic hydrogen carriers, MSc Thesis
- E. Huber, V. Bach, P. Holzapfel, D. Blizniukova, M. Finkbeiner, 2022, An Approach to Determine Missing Life Cycle Inventory Data for Chemicals (RREM), Sustainability, 14, 3161
- U. Vengateson, 2017, Cooling Towers: Evaporation Loss and Makeup Water, <https://www.chemengonline.com/cooling-towers-estimate-evaporation-loss-and-make-up-water-requirements/> (accessed Dec. 09, 2022)

# Distribution planning of medical oxygen supply chains under uncertainty

Georgios L. Bounitsis,<sup>a</sup> Yena Lee,<sup>a</sup> Karthik Thyagarajan,<sup>b</sup> Jose M. Pinto,<sup>b</sup>  
Lazaros G. Papageorgiou,<sup>a</sup> Vassilis M. Charitopoulos<sup>a</sup>

<sup>a</sup> *The Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, University College London, Torrington Place, London WC1E 7JE, UK*

<sup>b</sup> *Linde Digital Americas, Linde PLC, 10 Riverview Drive, Danbury CT, 06810, US*  
[v.charitopoulos@ucl.ac.uk](mailto:v.charitopoulos@ucl.ac.uk)

## Abstract

Recently, supply chain operations of medical products were put to test over the extreme uncertainties that COVID-19 pandemic induced. For instance, medical oxygen distribution to hospitals involves complex decision making due to volatility, which concern both industrial gas manufacturers and healthcare managers. In this work, we address the production and inventory routing problem (PIRP) of medical oxygen under demand uncertainty. A two-stage stochastic programming (TSSP) formulation is proposed where inventory decisions are set as the here-and-now variables. Uncertainty is captured via a novel framework, which exploits demand's forecasts and is coupled with data-driven scenario generation approaches. The methodology is examined on a case-study of medical oxygen distribution to hospitals in the UK during COVID-19 pandemic.

**Keywords:** Healthcare supply chain; Distribution Planning; Inventory Routing Problem; Optimisation under uncertainty; Stochastic Programming.

## 1. Motivation

During COVID-19 pandemic, medical oxygen demand at hospitals sporadically rocketed. Hence, efficient oxygen supply chain's optimal planning emerged as a very challenging problem as it requires simultaneous decisions on production, inventory, and distribution under extreme volatile circumstances. Oxygen supplier's objective is to minimise the cost while guaranteeing quality of service to hospitals and honouring related safety guidelines. In this work, we aim to extend the proposed Mixed-Integer Linear Programming (MILP) model of Lee et al. (2022) to account for uncertainty on oxygen supply chains. Thus, we propose a reformulation using scenario-based two-stage stochastic programming (TSSP), which accounts for uncertainties on COVID-19 cases and consequently oxygen demand at hospitals. Eventually, demand's scenario paths are generated from a novel framework which exploits both data from COVID-19 cases forecast models (COVID-19 Scenario Analysis Tool, 2022) and data-driven scenario generation methods (Bounitsis et al., 2022). As far as the modelling of the PIRP problem, the proposed TSSP remains MILP while decisions regarding inventory are the here-and-now decisions of the problem. Moreover, modelling includes the additional capability of emergency deliveries to the hospitals, in order to keep intact the security inventory levels of the hospitals for extreme scenarios with high demands of oxygen. Finally, a regional case study for hospitals in the UK is investigated to demonstrate the applicability of the proposed approach.

The remainder of the article is organised as follows: in Section 2 a brief literature review is conducted while the scenario generation framework and modelling advances are

presented Section 3. In Section 4, we study a regional case study of medical oxygen distribution to hospitals in the UK. Finally, conclusions are summarised in Section 5.

## 2. Literature review

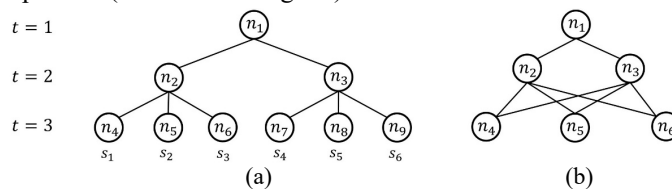
### 2.1. Production and Inventory Routing Problem

Integrated supply chain optimisation techniques have been studied in the literature aiming to make optimal production, inventory and distribution decisions simultaneously (Papageorgiou, 2009; Fahimnia et al., 2013). Production and Inventory Routing problem (PIRP) constitutes such an integrated problem, whose solution on chemical process industries can lead to significant benefits (Barbosa-Povoa & Pinto, 2020). In this context, You et al. (2011a) proposed an MILP model and solution techniques for the Inventory Routing problem (IRP) on industrial gas supply chains. Recently, Lee et al. (2022) developed an optimisation model for the PIRP of oxygen supply chains, using a two-level hybrid solution technique, which has been tested on a real-world supply chain planning problem in the UK during the COVID-19 pandemic. Although COVID-19 has brought uncertainties and risk on the planning problem, the latter work proposed a deterministic approach to confront the PIRP. IRP of industrial gases under uncertainty was examined by You et al. (2011b), where MINLP stochastic formulations were investigated. Focusing on stochastic IRP for medical applications, Nikzad et al. (2019) proposed TSSP model, chance constraint models and solution methods for a medical drug distribution problem under demand uncertainty.

### 2.2. Stochastic Programming & Scenario generation

Stochastic Programming (SP) is a risk-neutral mathematical approach for optimisation problems under uncertainty, in which the uncertainty is modelled via a known probability distribution. In cases of known discrete realisations, also called *scenarios*, optimisation problems are easily reformulated as scenario-based stochastic programs (so-called *deterministic equivalent problems*), which seek to optimise the expected value of the problem over the uncertain set. Stochastic programming problems are characterised as two-stage (TSSP) or multi-stage (MSSP) if uncertainty is revealed at a single or multiple steps respectively (Li & Grossmann, 2021). Scenario Generation (SG) methods are useful techniques to create a small set of scenarios, which is representative of the original uncertain set or the continuous stochastic process. Ultimately, scenario generation or reduction methods assist to efficiently solve computationally demanding SP problems.

Focusing on stochastic process,  $F_1, \dots, F_T$ , discretisation through a scenario tree is the most popular approach. A scenario tree (visualized in Fig. 1a) comprises of stages  $t \in T$ , which establish the time representation, and nodes  $n \in N$ , which include information regarding the possible realisations for the uncertain parameters and their probabilities of occurrence. So, in a multi-stage setting, a scenario path  $s \in S$  captures the full history of realisations from the root node till the leaf node (Li & Grossmann, 2021). However, when a Markovian data process is considered (i.e., the value of a parameter depends only on its precedent value), the less complex scenario lattice approach can be exploited to discretise the stochastic process (visualised in Fig. 1b).



**Figure 1:** Three-stage scenario paths using (a) scenario tree & (b) scenario lattice.

A scenario lattice is also organised by finite stages  $t \in T$ , but, in contrast to a scenario tree, a node of a lattice at stage  $t$  succeeds all the nodes of the previous stage  $t - 1$ . Hence, if  $\bar{N}_t$  denotes the number of nodes at stage  $t$ , the total number of scenario paths is equal to  $\prod_{t \in T} \bar{N}_t$ . Finally, probability weights correspond to the outgoing arcs of a node, and their sum is equal to one (Löhndorf & Wozabal, 2021).

### 3. Methodology and mathematical developments

#### 3.1. Proposed scenario generation methodology

Our ultimate goal is to generate scenario paths for oxygen demand at hospitals. This directly depends on the COVID-19 cases for each of the 3 countries of Great Britain (England, Scotland, Wales). Thus, towards the generation of future uncertain profiles of oxygen demand the input data include: (i) historical data for COVID-19 hospitalisations reported by the UK government (UK Health Security Agency, 2022), and (ii) predicted profiles of COVID-19 cases for the countries provided by advanced forecasting tools (COVID-19 Scenario Analysis Tool, 2022). These data are converted to medical oxygen demand and consequently the generated scenario paths refer directly to oxygen demand. The SG framework extends the work by Bounitsis et al. (2022) in a multi-stage setting, by assuming the data process as Markovian and employing the scenario lattice approach. Considering discrete time stages  $t \in T$ , the SG framework is outlined in Table 1.

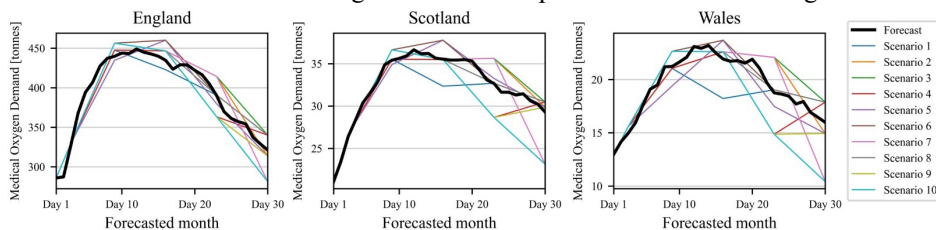
**Table 1:** Steps of the proposed multi-stage scenario generation framework.

---

<b>Step 1a</b>	Historical data are used as input to identify a copula function with goal to capture the correlations between oxygen demands of the 3 countries
<b>Step 1b</b>	Normal distributions are simulated for each country and each $t \in T$ using forecasting values as mean and an increasing variance as time progresses
<b>Step 2</b>	Copula-based sampling of scenarios is performed as in Bounitsis et al. (2022)
<b>Step 3</b>	MILP Distribution Matching Problem (DMP) integrating clustering by Bounitsis et al. (2022) is executed for each $t \in T$ separately, generating $\bar{N}_t$ nodes per stage $t$ with corresponding values and probabilities
<b>Step 4</b>	Scenario lattice is constructed using the nodes resulting to $\prod_t \bar{N}_t$ scenario paths
<b>Step 5</b>	OSCAR (Li & Floudas, 2014) MILP probabilistic distance minimisation model for scenario reduction is extended to estimate the multi-stage distance by Pflug & Pichler (2012) towards the further reduction of scenario paths

---

For instance, implementing the proposed SG framework, a final set of 10 scenario paths for the uncertain demand can be generated. These paths are visualised in Fig. 2.



**Figure 2:** Scenario paths of total hospital oxygen demand for a future month in the UK.

The small number of scenario paths for this example is selected because of the computational limitations of the case study problem. An exemplary SG process to generate 10 scenario paths is described as follows. In particular, the SG framework provides scenario paths with certain probabilities and values of oxygen demand for the 3



countries at each day. However, the time horizon of 30 days can be further discretised in stages  $t \in T$ . For instance, considering 5 time stages  $t$  (which correspond to future days 2, 9, 16, 23, 30, i.e., weekly steps to cover the whole horizon) and a branching vector (4,4,4,4,4), demand scenarios for each  $t$  are obtained by solving DMP. Then, a scenario lattice with 1024 scenario paths can be constructed. The demands on the in-between days can be determined by linear interpolation between the values of the nodes of the scenario paths. Finally, employing Step 5 of the SG framework, the set can sequentially be reduced to the 10 final scenario paths of Fig. 2, whose probabilities are also determined.

### 3.2. PIRP of medical oxygen under uncertainty

The proposed MILP is an extension of the upper-level problem for the PIRP of oxygen by Lee et al. (2022). This is reformulated as a TSSP to account for demand's uncertainty, as it is captured by scenario paths. For the TSSP the decisions are classified into the "here-and-now" and the "wait-and-see" decisions. The proposed model sets the security inventory levels at the hospitals as here-and-now variables, which are determined before the realisation of the uncertainty and are common for every scenario in the imminent step. The rest wait-and-see variables are dependent on each scenario and assist to alleviate infeasibilities for extreme scenario paths while inducing corrective impact. To this end, additional assumptions to the problem statement are beneficial such as the allowable violations of the prespecified minimum inventory levels at hospitals (but obeying security guidelines), which should be satisfied via emergency deliveries. Thus, modifications on the upper-level model by Lee et al. (2022) are presented and we refer the interested readers to the original work for notation and complete model formulation. The scenario-based formulation is developed using index  $s \in S$  for demand's scenario paths,  $D_{kts}$ , with corresponding probabilities  $p_s$ . Indices  $i \in I, k \in K, t \in T$  refer to plants, customers, and days, respectively. Objective function aims at minimisation of expected cost:

$$\text{minimise } \sum_{s \in S} p_s \cdot (TP_s + TT_s + TI_s + TED_s + TOP_s) \quad (1)$$

where  $TP_s, TT_s$  denote the production and transportation costs while  $TI_s, TED_s$  and  $TOP_s$  are new terms regarding plants' inventory cost, emergency delivery cost, and hospitals' inventory opportunity cost respectively. All emergency deliveries are conducted exclusively by vehicles from the plants  $i \in I$ . Terms regarding the costs are given below:

$$TI_s = \sum_{i \in I} \sum_{t \in T} f^{inv} C^{elec} USP I_{its} \quad \forall s \in S \quad (2)$$

$$TED_s = f^{tr} \sum_{i \in I} \sum_{k \in K} \sum_{v \in V} \sum_{t \in T} C_v^{veh} L_{ik} A_{ikvts}^{ed} \quad \forall s \in S \quad (3)$$

$$TOP_s = f^{oc} \sum_{k \in K} \sum_{t \in T} (I_{kts} - ID_k) \quad \forall s \in S \quad (4)$$

Parameters  $f^{inv}, f^{out}, f^{tr}$  are factors for the estimation of costs. Then,  $TI_s$  and  $TOP_s$  are estimated using corresponding unit costs, i.e.,  $f^{inv} C^{elec} USP$ , and  $f^{oc}$ , which is the constant of opportunity cost [ $\text{£/kg} - \text{day}$ ]. These are multiplied by the stored amounts  $I_{its}$  and  $(I_{kts} - ID_k)$  respectively. In this work,  $ID_k$  denotes the here-and-now decisions on the optimal hospitals inventory levels, while prespecified (agreed-upon) lower bounds by the hospitals are denoted as  $IH_k$ . When such a lower bound is violated, hospital is on emergency and emergency deliveries are conducted, denoted as  $A_{ikvts}^{ed}$ . In particular, variables  $Q_{kts}^{d,minus}$ , which capture negative deviation from  $IH_k$ , are combined with binary variables,  $E_{kts}$ , in Eq. (5) to indicate if a hospital is on an emergency:

$$Q_{kts}^{d,minus} = (IH_k - I_k^{min}) E_{kts} \quad \forall k \in K, t \in T, s \in S \quad (5)$$

Parameter  $I_k^{min}$  denotes the hard security bound for the normal operation of the hospitals, which must not be violated. Hence, Eq. (5) provides an upper bound for the allowable violated levels which varies for hospitals. For the sake of brevity, revised balances for hospitals inventory,  $I_{kts}$ , are provided in Eq. (6), including emergency loads ( $Q_{ikvts}^{ed}$ ). Rest of the constraints are reformulated accordingly based on the model by Lee et al. (2022).

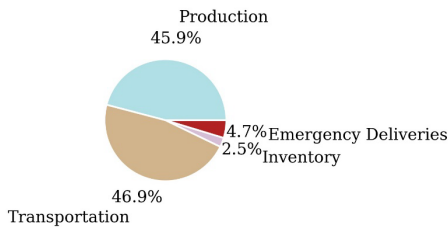
$$I_{kts} = I_k^{ini} |_{t=1} + I_{k,t-1,s} |_{t>1} + \sum_{i:k \in K_{iv}} \sum_{v \in V_i} (Q_{ikvts} + Q_{ikvts}^{ed}) - D_{kts} \quad \forall k \in K, t \in T, s \in S \quad (6)$$

#### 4. Case study

This synthetic case study examines the PIRP problem of medical oxygen distribution in a certain geographical region in the UK and was developed in collaboration with our industrial partners. The regional case comprises of one plant, which supplies 56 hospitals using 6 vehicles over a time horizon of 30 days. As the presented modification of the model in Section 3.2 can easily be modified for various assumptions, we assume that emergency deliveries can be conducted by all available vehicles of the plant. Moreover, emergency deliveries are forced to be conducted to a customer  $k$  the next day of its  $IH_k$  level’s violation. The latter is set by appropriate manipulation of binary variables  $E_{kts}$ . Model statistics are presented in Table 2 and costs are on monetary units [RMU] for confidentiality reasons. The deterministic model (DET) considers only one scenario path, i.e., the forecasted values of Fig. 2. Stochastic model (STOC) employs the 10 generated scenario paths, which are also visualised in Fig. 2. The executions are performed in a Dell workstation with Intel® Core™ i9-10900K CPU @ 3.70 GHz and 32.00 GB RAM using solver GUROBI 9.5 in GAMS 41.3 modelling and optimisation system. An optimality gap of 5% or 3,600s of execution time are set as stopping criteria.

**Table 2:** Model statistics on the proposed case study.

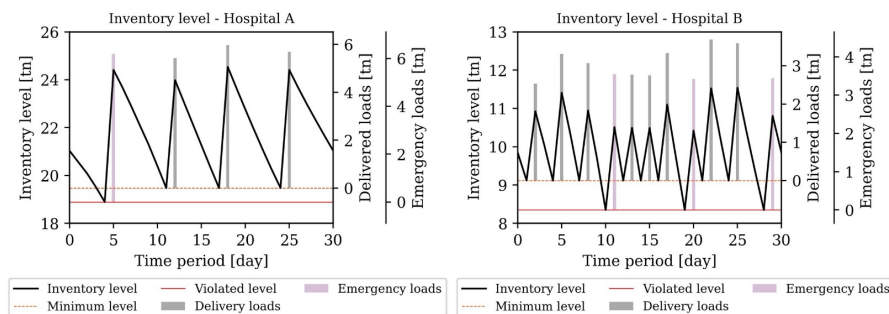
Model	Solution [RMU]	Opt. gap [%]	CPU [s]	Total Equations	Total Variables	Discrete Variables
DET	100.00	4.88	41	38,378	41,942	18,360
STOC	101.21	5.72	3,600	383,771	418,907	183,600



**Figure 3:** TSSP expected cost breakdown.

Statistics from Table 2 demonstrate that the problem is computationally demanding, and the number of discrete variables exaggerates with the number of scenarios. Despite the large number on model statistics near optimal solutions are computed within reasonable execution times. Towards the analysis of results, the breakdown of the expected costs for the stochastic formulations is visualised in Fig. 3.

Emergency deliveries cost constitutes 4.7%. In the same context, revised minimum levels have been proposed for 16 of 56 of the examined hospitals. However, the maximum revision does not exceed 8% of the maximum capacity of a hospital. Besides, the adopted strategy for next day satisfaction of the violated levels do not allow wide and lasting violations and the stochastic problem remains feasible. Exemplary inventory profiles of hospitals at a certain scenario path are presented in Fig. 4.



**Figure 4:** Inventory profiles for hospitals in a certain scenario path.

## 5. Conclusions

The contribution of this problem is twofold: (i) to extend the PIRP planning of oxygen distribution to a stochastic context using stochastic programming, (ii) to propose a SG framework based on scenario lattice approach to discretise the demand stochastic process. As investigated problem remains computationally complex, future work within our research group focuses on the development of tailor-made solution techniques.

**Acknowledgements:** The authors would like to acknowledge financial support from the UK Engineering & Physical Sciences Research Council (under project EP/V050168/1).

## References

- A.P. Barbosa-Povoa, J.M. Pinto, 2020. Process supply chains: Perspectives from academia and industry. *Comput. Chem. Eng.* 132, 106606.
- G.L. Bounitsis, L.G. Papageorgiou, V.M. Charitopoulos, 2022. Data-driven scenario generation for two-stage stochastic programming. *Chem. Eng. Res. Des.* 187, 206-224.
- COVID-19 Scenario Analysis Tool, 2022. MRC Centre for Global Infectious Disease Analysis, Imperial College London. [www.covidsim.org](http://www.covidsim.org)
- B. Fahimnia, R.Z. Farahani, R. Marian, L. Luong, 2013. A review and critique on integrated production–distribution planning models and techniques. *J. Manuf. Syst.* 32, 1–19.
- Y. Lee, V.M. Charitopoulos, K. Thyagarajan, I. Morris, J.M. Pinto, L.G. Papageorgiou, 2022. Integrated production and inventory routing planning of oxygen supply chains. *Chem. Eng. Res. Des.* 186, 97–111.
- C. Li, I.E. Grossmann, 2021. A Review of Stochastic Programming Methods for Optimization of Process Systems Under Uncertainty. *Front. Chem. Eng.* 2, 622241.
- Z. Li, C.A. Floudas, 2014. Optimal scenario reduction framework based on distance of uncertainty distribution and output performance: I. Single reduction via mixed integer linear optimization. *Comput. Chem. Eng.* 70, 50–66.
- N. Löhndorf, D. Wozabal, 2021. Gas storage valuation in incomplete markets. *Eur. J. Oper. Res.* 288, 318–330.
- E. Nikzad, M. Bashiri, F. Oliveira, 2019. Two-stage stochastic programming approach for the medical drug inventory routing problem under uncertainty. *Comput. Ind. Eng.* 128, 358–370.
- L.G. Papageorgiou, 2009. Supply chain optimisation for the process industries: Advances and opportunities. *Comput. Chem. Eng.*, 33, 1931-1938.
- G.C. Pflug, A. Pichler, 2012. A Distance For Multistage Stochastic Optimization Models. *SIAM J. Optim.* 22, 1–23.
- UK Health Security Agency, 2022. Healthcare in UK. <https://coronavirus.data.gov.uk>
- F. You, J.M. Pinto, E. Capón, I.E. Grossmann, N. Arora, L. Megan, 2011a, Optimal distribution-inventory planning of industrial gases. I. Fast computational strategies for large-scale problems. *Ind. Eng. Chem. Res.* 50, 2910–2927.
- F. You, J.M. Pinto, I.E. Grossmann, L. Megan, 2011b, Optimal distribution-inventory planning of industrial gases. II. MINLP models and algorithms for stochastic cases. *Ind. Eng. Chem. Res.* 50, 2928–2945.

# Life-cycle assessment of marine biofuels from thermochemical liquefaction of different olive residues in Spain

Sivaramakrishnan Chandrasekaran,<sup>a</sup> Puck Wammes,<sup>a</sup> John A. Posada,<sup>a</sup>

<sup>a</sup>*Delft University of Technology, Mekelweg 5, Delft and 2628 CD, the Netherlands*

## Abstract

Advanced biofuels from thermochemical liquefaction, such as pyrolysis (PY) and Hydrothermal liquefaction (HTL), of olive residues in the Andalusian region of Spain (specifically in the province of Jaen) can potentially play a crucial role in the reduction of greenhouse gas (GHG) emissions in the maritime sector. In this study, an attributional life-cycle assessment (ALCA) was performed to estimate and compare the GHG emissions for producing marine biofuels via pyrolysis and HTL from olive pomace (COP) and pruning biomass (OTPB), to provide 1 megajoule (MJ) of marine biofuel, as a functional unit. For convenience, the different technology-feedstock combination scenarios are represented as scenario 1 (PY\_COP), scenario 2 (PY\_OTPB), scenario 3 (HTL\_COP), and scenario 4 (HTL\_OTPB). The life-cycle GHG emissions of the biofuels were 42.0, 44.1, 22.1, and 32.1 g CO<sub>2</sub>-eq/MJ for PY\_COP, PY\_OTPB, HTL\_COP and HTL\_OTPB scenarios, respectively, corresponding to 47–73% GHG emissions reduction compared with petroleum fuels. The scenarios were also evaluated based on other impact categories such as Sulphur dioxide in the air, Nitrogen oxides in the air, Particulates in the air, and Non-methane volatile organic compounds (NMVOCs) in the air. The scenarios reduced the SO<sub>2</sub> emissions, Nitrogen emissions, NMVOCs, and particulates in the air by at least 50%, 90%, 20%, and 25% respectively in comparison to fossil fuels. A contribution analysis revealed that olive cultivation and upgrading as hot spots for emission in pyrolysis-based systems. Likewise, HTL conversion and upgrading steps were emitting more emissions for an HTL-based system. Therefore, marine biofuel obtained through the thermochemical conversion of olive residues has better environmental performance on a life cycle basis, with a preference for HTL based system over pyrolysis.

**Keywords:** Marine biofuels, hydrothermal liquefaction, Pyrolysis, Life-cycle assessment, Greenhouse gas emissions.

## 1. Introduction

Ships play a central role in the global supply chain of any commodity. About 80% of the world's products are transported by the marine sector. This makes it one of the largest consumers of fossil-based fuels. The total consumption is estimated to be around 330 million metric tons of heavy fuel annually (Hsieh & Felby, 2017). In 2019, Spain consumed approximately 5 million tons of in terms of international maritime bunkering (Eurostat, 2022). Even though maritime transportation emits fewer pollutants per ton kilometer, it contributes to 15% of global carbon emissions. If left unregulated, the emissions from this sector are expected to grow by between 250% in 2050 (Rutherford & Comer, 2018). Decreasing these emissions can therefore greatly contribute to the current

world's environmental challenges. Lignocellulosic biomass-based marine biofuels are seen as a must in reaching these goals because the CO<sub>2</sub> emitted during combustion is in balance with the CO<sub>2</sub> absorbed during the cultivation of biomass. Also, the lower Sulphur and nitrogen content in this type of biomass make them suitable to attain IMO regulations. The residues from the olive sector are one such potential candidate for producing biofuel for the maritime sector. In terms of availability, Spain devotes 2.5 million hectares with more than 180 million trees to this industry. In the province of Jaen, over 91% of the land is occupied by olive trees, making it one of the highest occupants in the entire EU. With accounting for 80% production capacity of olive oil in Spain, the Andalusian region is one of the largest agricultural sectors in the olive- and olive oil industry (UNESCO, 2017). Consequently, the sector produces a large amount of lignocellulosic biomass waste such as olive tree pruning biomass (OTPB), olive pits, leaves, and crude olive pomace (COP). For example, pruning biomass alone has an estimated waste of between 1.5 and 3 tons per hectare in a year (Ruiz et al., 2017). Thereby valorizing these residues will provide an opportunity for developing a bio-economy.

Thermochemical liquefaction, such as pyrolysis, hydrothermal liquefaction (HTL) and gasification, of biomass to bio-crude along with sequential upgrading to biofuel is one of the potential conversion routes for marine biofuel production. Although various studies have been conducted to investigate the thermochemical conversion of olive residues, very few studies have been conducted for biofuel production via HTL based on residues from the olive sector (De Filippis et al., 2016; Evcil et al., 2021). Based on this literature the thermochemical liquefaction of olive residues to produce biofuel is quite promising. With the amount of residues available in the region of Jaen, almost 10% of Spain's national HFO demand can be satisfied. Therefore, apart from techno-economical assessment, analysis based on the environmental aspect is necessary to evaluate the potential of these novel alternatives. Some of the literature also go further beyond the experimental work to evaluate the environmental performance of these methods (Hani & Hailat, 2016; Parascanu et al., 2018). As result indicating better performance in terms of the potential for carbon emission reduction along their life cycle. However, a comparative attributional life cycle analysis is needed to understand and select a suitable conversion pathway for marine biofuel production.

This study's objective is to provide a deeper understanding of the environmental impacts of utilizing olive industry residues, such as COP and OTPB, to biofuels for the maritime sector. To evaluate the environmental performance by quantification of the GHG emissions of these designs by identifying key performance indicators and performing life cycle analysis (LCA). Lastly, this work will serve as a basis for future studies by providing recommendations on improving GHG emission performance. To reach the objective, this thesis is structured to answer the following question: "What are the environmental life cycle impacts of utilizing lignocellulosic residues of the olive sector in Spain to produce marine biofuels via HTL or Pyrolysis?"

## 2. Methodology

This life cycle analysis (LCA) aims to evaluate the environmental impacts and hotspots of the thermochemical conversion of olive industry residues to produce marine biofuels in Spain. In this LCA, the entire proposed supply chain from cradle to grave is considered, from olive cultivation to the combustion of marine biofuel in ships. In this study, an Attributional LCA (ALCA) is performed. The ALCA is to find information on the global

burdens that can be associated with a product life cycle, it uses cut-off rules and allocation to isolate the production system from the larger system to assess solely the route to marine biofuel. As for geographical coverage, the LCA is focused on the Jaen in the Andalusian region of Spain. The entire chain will be modeled in Spain or Europe. If necessary, global data was used. Figure 1 shows the overall system boundary encompassing all the scenarios. The system includes olive cultivation, primary Virgin olive oil (VOO) extraction, secondary extraction, thermochemical conversions, biofuel upgrading using hydrogen, and final combustion of marine biofuels.

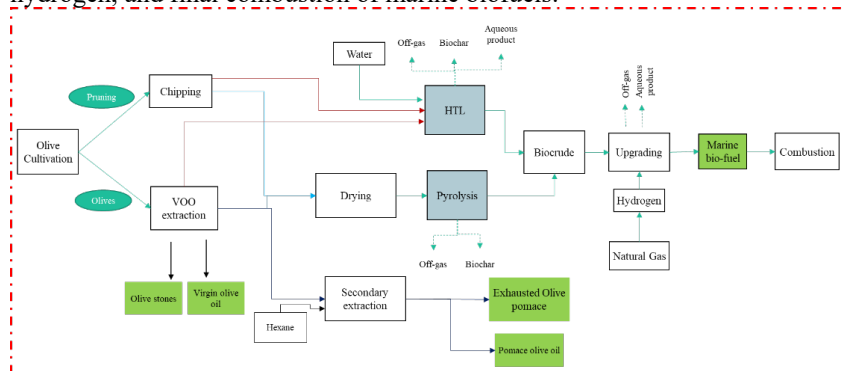


Figure 1: System boundary of scenarios to produce marine biofuels from olive residues in Spain. The green box indicates the products and co-products of the systems, and the blue-grey box indicates the thermochemical pathways. Dotted lines indicate the waste streams.

The analysis follows ISO 14040, the international standard for LCA, and the functional unit is set to be 1 megajoule (MJ) of biofuel produced. All emissions from each process and its associated upstream supply chain were accounted for in this study. But, the emission originating due to the construction of infrastructure, manufacture of equipment, wastewater treatment, biochar, and off-gas utilization were not included within the scope. Data quality and specificity are widely considered key criteria for any LCA studies. Local-specific data such as environmental uniqueness and spatial variation are one of the concerns that require special attention. Hence, the data consistency and accuracy are enhanced by utilizing local data of olive residues obtained from a field visit, e.g., the feedstock availability, the locations of biofuel supply chain nodes, the transportation, and the electricity mix. For the processes modeling and development, e.g., biofuel production and delivery including catalysts, hydro treating and hydrogen production, and nitrogen fertilizer, the GREET 2015 (Greenhouse gases, regulated emissions, and energy use in transportation) or Activity browser coupled with Eco invent v3.7 database was used to model the process emissions with modifications. When the data could not be found in the software database described above, they are collected from peer-reviewed journal articles or reports issued by the government and widely recognized scientific organizations (e.g., IPCC) or laboratories (e.g., PNNL and NREL).

The emissions from each process are obtained based on the physical allocation method. Specifically, mass and energy balances were initially calculated for each process through the Aspen Plus v12 process simulator and then multiplied by the corresponding emission factors. The aspen process model developed includes biomass pretreatment, hydrothermal reactor system, product separation and purification, and byproduct valorization. The off-gas and aqueous stream were recycled within the system. The collected raw data from the various sources described above were first compiled in Microsoft Excel to build the

lifecycle inventory of marine biofuels, and then, IPCC 2007 Global warming potential factors were used to convert CO<sub>2</sub>, CH<sub>4</sub>, and N<sub>2</sub>O into CO<sub>2</sub>-eq for a time horizon of 100 years.

### 3. Results

#### 3.1. GHG emissions from product allocation

The GHG emissions obtained as a result of producing marine biofuel to reach the reference flow of 1 MJ is presented in Figure 2 as g CO<sub>2</sub>-eq physically allocated to the fuel production system. Scenarios 1 and 2 emit 42.0 and 44.1 g CO<sub>2</sub>-eq/MJ, respectively compared to 83.6 g CO<sub>2</sub>-eq/MJ for HFO as calculated for 1 MJ by using the Eco invent database. This is approximately a 50% decrease in emissions. Likewise, Scenarios 3 and 4 even decreased the GHG emissions to approximately 25% of that of HFO with 22.1 and 32.1 g CO<sub>2</sub>-eq/MJ. A similar reduction can be observed in terms of kg CO<sub>2</sub>-eq/kg fuel where scenarios 1-4 releases 1.6, 1.2, 0.7, and 0.8 kg CO<sub>2</sub>-eq/kg fuel compared against 3.42 kg CO<sub>2</sub>-eq/kg fuel for HFO. These values are in the range of data reported in the literature (Capaz et al., 2020, 2021; Nie & Bi, 2018; Parascanu et al., 2018; Santos et al., 2018; Tanzer et al., 2019),

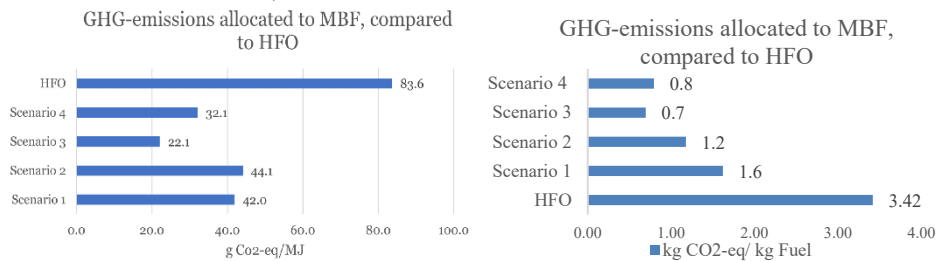


Figure 2: The physically allocated g CO<sub>2</sub>-eq/MJ fuel (left) and kg CO<sub>2</sub>-eq/kg fuel (right) per scenario compared to HFO fromecoinvent.

#### 3.2. Other impact categories

The environmental performance of the four scenarios considering the other environmental impact criteria is discussed in this section. Figure 3, shows the LCI results for emissions SO<sub>2</sub>, NO<sub>x</sub>, NMVOC, and particulates <2.5 am, in centigrams per megajoule. This is

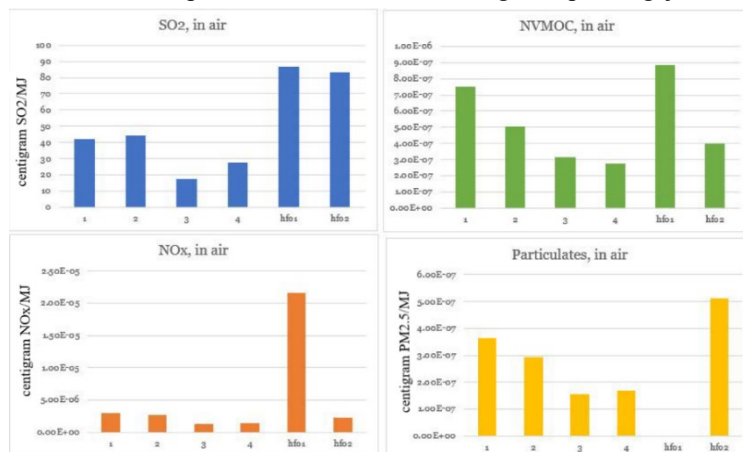


Figure 3: LCI results for emissions to air, compared to HFO from Eco invent and HFO from (Andersson et al., 2016) based on physical allocation

compared to HFO LCA results from the Eco invent database (hfo1) and the LCA results from literature (hfo2) (Andersson et al., 2016). The scenarios reduced the SO<sub>2</sub> emissions, Nitrogen emissions, NMVOCs, and particulates in the air by at least 50%, 90%, 20%, and 25% respectively in comparison to fossil fuels.

### 3.3. Contribution analysis

To determine the hotspots in the biofuel product systems, a contribution analysis is run through activity-browser, on the allocated model. The figure shows the life-cycle stage-wise GHG emissions of the four biofuel production scenarios. The contribution from main unit processes: Olive cultivation, virgin olive oil extraction, drying and crushing of biomass, HTL or pyrolysis conversion and upgrading, are analyzed.

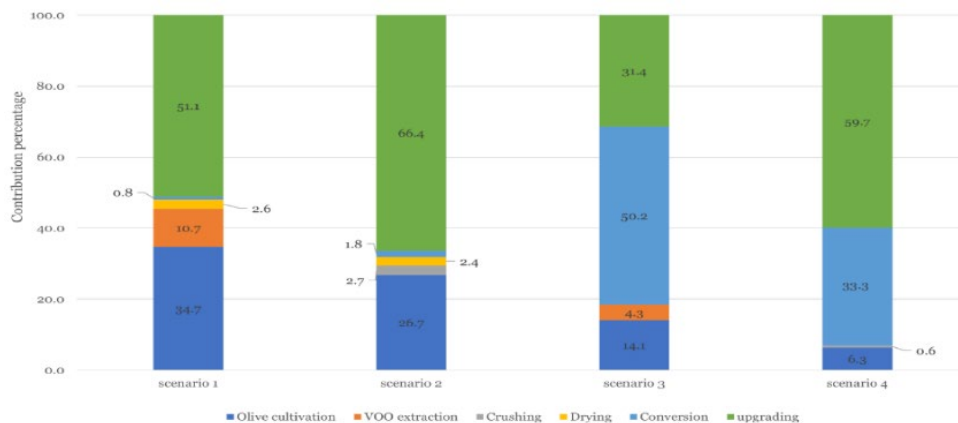


Figure 4: Contribution percentage of the unit processes that contribute to the scenarios

The most dominant contributor of Pyrolysis systems to GHG emissions is bio-crude upgrading, which makes up around 50-60% followed by olive cultivation. The most dominant for HTL\_OP is the biomass conversion process, which counts for almost 50%, followed by upgrading with 31.5%. However, an opposite trend was observed for HTL\_OTPB with upgrading causing 60% emission followed by conversion. This is might be due to the assumption of similar fuel characteristics of pyrolysis OTPB biofuel and HTL\_OTPB biofuel which was made due to data unavailability.

## 4. Conclusion

This study shows that the product system towards marine biofuel (MBF) compared to HFO has lower allocated GHG emissions. The final GHG emissions from scenarios 1-4 are modeled to be: 42, 44.1, 22.1, and 32.1 g CO<sub>2</sub>-eq/MJ, respectively. Which is all a decrease compared to 84 g CO<sub>2</sub>-eq/MJ from HFO. Also, the use of MBF reduced the amount of SO<sub>2</sub>, particulates, and PMVOC emissions into the air, specifically from 83 g SO<sub>2</sub>/MJ (HFO) to 20-40g SO<sub>2</sub>/MJ. Therefore, implementing the proposed systems could reduce two of the major environmental impacts of using HFO in marine shipping. Based on the analysis of bio-crude obtained from the thermochemical conversion pathways the nitrogen content was negligible hence corresponding NO<sub>x</sub> emissions while combusting the MBF are expected to be meager. Finally, the contribution analysis indicated that the conversion pathways were the hotspots for GHG emission in the scenarios. Thereby optimizing the process will lead to a further reduction in emissions to the environment.



However, sensitivity analysis and system expansion studies are recommended to effectively understand the impact of the scenarios.

## References

- Andersson, K., Brynolf, S., Lindgren, J. F., & Wilewska-Bien, M. (2016). Shipping and the Environment: Improving Environmental Performance in Marine Transportation. *Shipping and the Environment: Improving Environmental Performance in Marine Transportation*, 1–426. <https://doi.org/10.1007/978-3-662-49045-7>
- Capaz, R. S., de Medeiros, E. M., Falco, D. G., Seabra, J. E. A., Osseweijer, P., & Posada, J. A. (2020). Environmental trade-offs of renewable jet fuels in Brazil: Beyond the carbon footprint. *Science of the Total Environment*, 714, 136696. <https://doi.org/10.1016/j.scitotenv.2020.136696>
- Capaz, R. S., Guida, E., Seabra, J. E. A., Osseweijer, P., & Posada, J. A. (2021). Mitigating carbon emissions through sustainable aviation fuels: costs and potential. *Biofuels, Bioproducts and Biorefining*, 15(2), 502–524. <https://doi.org/10.1002/bbb.2168>
- De Filippis, P., De Caprariis, B., Scarsella, M., Petruccio, A., & Verdone, N. (2016). Biocrude production by hydrothermal liquefaction of olive residue. *International Journal of Sustainable Development and Planning*, 11(5), 700–707. <https://doi.org/10.2495/SDP-V11-N5-700-707>
- Eurostat. (2022). *Supply, transformation and consumption of oil and petroleum products*. [https://ec.europa.eu/eurostat/databrowser/view/NRG\\_CB\\_OIL\\_custom\\_1359468/bookmark/?lang=en&bookmarkId=57d36982-0e9a-4226-9634-3ee6ead50431](https://ec.europa.eu/eurostat/databrowser/view/NRG_CB_OIL_custom_1359468/bookmark/?lang=en&bookmarkId=57d36982-0e9a-4226-9634-3ee6ead50431)
- Evciil, T., Tekin, K., Ucar, S., & Karagoz, S. (2021). Hydrothermal liquefaction of olive oil residues. *Sustainable Chemistry and Pharmacy*, 22(June), 100476. <https://doi.org/10.1016/j.scp.2021.100476>
- Hani, F. F. B., & Hailat, M. M. (2016). Production of Bio-Oil from Pyrolysis of Olive Biomass with/without Catalyst. *Advances in Chemical Engineering and Science*, 06(04), 488–499. <https://doi.org/10.4236/aces.2016.64043>
- Hsieh, C.-W. C., & Felby, C. (2017). *Biofuels for the marine shipping sector*. 86. <http://task39.sites.olt.ubc.ca/files/2013/05/Marine-biofuel-report-final-Oct-2017.pdf%0Ahttps://www.ieabioenergy.com/wp-content/uploads/2018/02/Marine-biofuel-report-final-Oct-2017.pdf>
- Nie, Y., & Bi, X. (2018). Life-cycle assessment of transportation biofuels from hydrothermal liquefaction of forest residues in British Columbia. *Biotechnology for Biofuels*, 11(1), 1–14. <https://doi.org/10.1186/s13068-018-1019-x>
- Parascanu, M. M., Puig Gamero, M., Sánchez, P., Soreanu, G., Valverde, J. L., & Sanchez-Silva, L. (2018). Life cycle assessment of olive pomace valorisation through pyrolysis. *Renewable Energy*, 122, 589–601. <https://doi.org/10.1016/j.renene.2018.02.027>
- Ruiz, E., Romero-García, J. M., Romero, I., Manzanares, P., Negro, M. J., & Castro, E. (2017). Olive-derived biomass as a source of energy and chemicals. *Biofuels, Bioproducts and Biorefining*, 11, 1077–1094. <https://doi.org/10.1002/bbb.1812>
- Rutherford, D., & Comer, B. (2018). *THE INTERNATIONAL MARITIME ORGANIZATION'S INITIAL GREENHOUSE GAS STRATEGY*. [https://theicct.org/sites/default/files/publications/IMO\\_GHG\\_StrategyFinalPolicyUpdate042318.pdf](https://theicct.org/sites/default/files/publications/IMO_GHG_StrategyFinalPolicyUpdate042318.pdf)
- Santos, C. I., Silva, C. C., Mussatto, S. I., Osseweijer, P., van der Wielen, L. A. M., & Posada, J. A. (2018). Integrated 1st and 2nd generation sugarcane bio-refinery for jet fuel production in Brazil: Techno-economic and greenhouse gas emissions assessment. *Renewable Energy*, 129, 733–747. <https://doi.org/10.1016/j.renene.2017.05.011>
- Tanzer, S. E., Posada, J., Geraedts, S., & Ramirez, A. (2019). Lignocellulosic marine biofuel: Technoeconomic and environmental assessment for production in Brazil and Sweden. *Journal of Cleaner Production*, 239, 117845. <https://doi.org/10.1016/j.jclepro.2019.117845>
- UNESCO. (2017). *The Olive Grove Landscapes of Andalusia*. <https://whc.unesco.org/en/tentativelists/6169/>

# An optimisation approach for the design of distributed solar farms with shared energy storage in Pakistan

Obaid Khan<sup>a</sup>, Ishanki De Mel<sup>a</sup>, Robert A. Steven<sup>a</sup>, Michael Short<sup>a\*</sup>

<sup>a</sup>*Department of Chemical and Process Engineering, University of Surrey, Guildford, Surrey GU2 7XH, United Kingdom.*  
*m.short@surrey.ac.uk*

## Abstract

In many developing nations, the carbon intensity of the grid is high while capacity is also constrained. Local distributed energy resources are expensive to deploy and own by private households. This work presents an optimisation-based approach to optimally locate renewable technologies at both local prosumer's dwellings and at community-level locations, incorporating the addition of distribution lines in a distributed energy system (DES) model to implement a solar photovoltaic (PV) energy farm with energy storage at the community level in densely populated urban areas. A case study of three scenarios in Karachi, Pakistan is selected to test the model. Scenario 1 presents a decentralised approach where every node has the possibility to install batteries locally. Scenario 2 considers a more centralised approach where only 2 central battery nodes are connected to different residential nodes. Lastly, the scalability of the model is tested with Scenario 3 which considers a model with 4 times as many nodes. The optimisation model indicates the profitability of different centralised/decentralised approaches with the highest level of decentralisation showing both the lowest cost and highest profit made, demonstrating the trade-off between battery sizing and line costs for privatised local energy storage.

**Keywords:** Distributed energy systems, renewable energy, energy systems.

## 1. Introduction

A long-term sustainable and renewable solution to increasing emissions resulting from fossil fuel consumption may be achieved through high penetration of renewable energy technologies such as solar PV, wind turbines, battery technologies, etc. which will be essential to meet rising energy demands while reducing greenhouse gas emissions. Solar energy systems require large space requirements to install, particularly in large cities, and in many countries there are political, financial, and institutional barriers that may make large-scale, grid-integrated solutions difficult to achieve. Battery storage can help to alleviate the intermittency associated with renewable generation and mismatches between generation and load but the cost of investment and operations makes it difficult to readily employ at scale and developing countries which may already have an energy shortfall cannot implement these technologies at current costs.

Pimm et al. (2018) studied local distribution networks and how privately-owned energy storage can save on peak tariff rates by storing and reducing peak load. Cedillos Alvarado et al. (2016) presented a technology and selection optimisation model in which the author

optimised the real-time renewable energy resources with energy demand. This model is purely data-driven and encompasses whole life costing, carbon emissions and energy pricing. Most DES problems are formulated as mixed-integer linear programs (MILPs), which is a popular and effective strategy for solving large and complicated processes such as microgrid energy demand-supply planning (Silvente et al., 2015), albeit without integrating detailed power flow concerns (de Mel et al., 2022). While several studies have looked at optimal neighbourhood electrical storage, the area that this should service, along with whether electrical networks may need to be expanded, has not been studied in depth. This paper presents a DES model which seeks to find the optimal design of a distributed solar farm with energy storage, including the addition of distribution lines (Zakeri et al., 2021). In this case, both solar PV and battery storage are included and the resultant model is tested in a case study involving a neighbourhood in Karachi, Pakistan, where Khan & Arsalan (2016) determined that by utilising a neighbourhood's complete rooftop area, it is possible to create 10 times more electricity than it consumes by installing PVs.

## 2. Model formulation

The model formulation in this paper develops a DES MILP optimisation model which introduces additional distribution lines to connect the solar PV installed by end-users to battery storage distributed throughout a neighbourhood. The model for DES is based on the work of de Mel et al. (2021) and Mehleri et al. (2012). The geographical boundary of the neighbourhood is defined and the enclosed area is divided into several nodes, as shown in Figure 1, where each node contains both houses with and without solar PV installed and each house may be connected to battery storage units by a distribution line. The battery storage (in this case Li-ion batteries) is either sited at a point within each node or in separate centralised storage nodes.

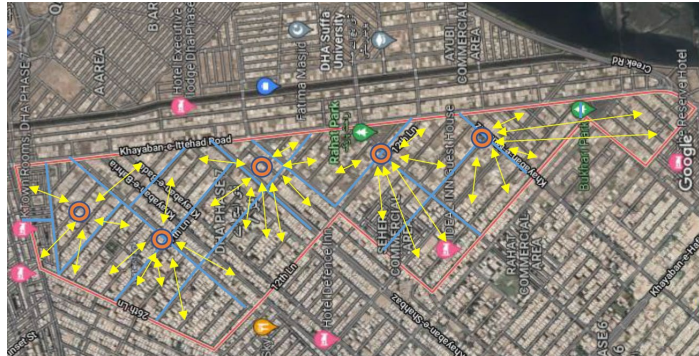


Figure 1: Sample topology with red lines indicating neighbourhood boundaries, blue lines node boundaries, red circles indicating battery storage and yellow arrow indicating distribution lines.

The model itself is formulated as an MILP problem, where the objective function (1) minimises the system total annualised cost ( $TAC$ ). This contains distribution line costs ( $DLC$ ), electricity purchased from the grid ( $C_s^{grid}$ ), distributed energy resource installation ( $AC^{INV,DER}$ ) and operational costs ( $C_s^{OM,DER}$ ) and seasonal selling costs of generated electricity ( $I_s$ ).

$$TAC = DLC + \sum_{DER} AC^{INV,DER} + \sum_{s \in S} \left( C_s^{grid} + \left( \sum_{DER} C_s^{OM,DER} \right) - I_s \right) \quad (1)$$

Here, the overall distribution line costs ( $DLC_{i,k}$ ) are defined as:

$$DLC_{i,k} = \left( (BDC) * (CC) * (ReC) * \frac{Node\ Distance * CRF}{4} \right) * N_{houses} \quad (2)$$

This contains the base distribution cost ( $BDC$ ), conductor cost ( $CC$ ), the re-conductor coefficient ( $ReC$ ), capital recovery factor ( $CRF$ ), the distance from each house to the central node (each house is only connected to a single node) and the total number of houses attached to the node. The summation subscript  $s \in S$  is used to denote a seasonal formulation whereby both the irradiance and load profile datasets contain separate values for each of the four seasons.

The model constraints include power flow between the centralised battery storage and houses, considering both charging and discharging cycles, PV generation capacity, electricity export to the grid and battery capacities, omitted in this paper for brevity.

Further constraints contain capital costs, seasonal formulations, ensuring energy is not bought and sold at the same time, operational and maintenance costs and the cost of both buying and selling electricity to the grid.

Here, the total solar PV power generated is:

$$E_{i,t}^{PV,used} + E_{i,t}^{PV,sold} + \sum_{c,k} E_{i,k,t,c}^{PV,charge} \leq N_i^{Panel} * A^{Panel} * Irr_t * \eta^{Panel} \quad (3)$$

Where the power consumed ( $E_{i,t}^{PV,used}$ ), generated ( $E_{i,t}^{PV,sold}$ ) and used to charge the battery storage ( $E_{i,k,t,c}^{PV,charge}$ ) is less than or equal to the product of the number of panels installed ( $N_i^{Panel}$ ), the area of the panels ( $A^{Panel}$ ), the irradiance received per panel ( $Irr_t$ ) and the efficiency of the panel ( $\eta^{Panel}$ ). The number of panels that can be installed is limited by the maximum roof area available,  $N_i^{Panel} * A^{Panel} \leq A_i^{roof}$ .

Battery parameters are also included, with battery energy balance equations are given as:

$$\xi_{k,t,c}^{stored} = (E_{k,t,c}^{ch} * \eta^{ch} * \Delta t) - \frac{\sum_i E_{i,k,t,c}^{disch} * \Delta t}{\eta^{disch}}, t = 1 \quad (4)$$

$$\xi_{k,t,c}^{stored} = \xi_{k,t-1,c}^{stored} + (E_{k,t,c}^{ch} * \eta^{ch} * \Delta t) - \frac{\sum_i E_{i,k,t,c}^{disch} * \Delta t}{\eta^{disch}}, t \geq 2 \quad (5)$$

Containing the charging ( $E_{k,t,c}^{ch}$ ) and discharging ( $E_{k,t,c}^{disch}$ ) battery powers, with their corresponding efficiencies ( $\eta^{ch}$ ,  $\eta^{disch}$ ).

Table 1: Model Parameters

Description	Values	Units
Electricity price in £ per kWh	0.090651	£/kWh
Capital cost of PV in £ per panel	110	£/panel
Fixed operational cost of PV	12.5	£/kWh-yr
Variable operational cost of PV	0.005	£/kWh
Tariff for exporting in £ per kWh	0.0465	£/kWh
Surface area per panel	2.1	m <sup>2</sup>
Base transmission cost	5	£/m
Conductor Cost	2.5	£/m
Re-conductor Cost	1.7	£/m

### 3. Case study

A local neighbourhood in the city of Karachi was selected due to its wide-open spaces and large unutilised rooftop areas as well as its sizeable electricity demand.

For this case study, both real-world irradiance and load profile datasets were obtained, with the load data anonymised from a subset of residents and used to generate average load profiles. Other parameters such as PV cost, battery cost and the electricity tariff were set according to the local market rates. GAMS CPLEX 20.1.10.1 (IBM, 2021) was used to solve the MILP. A subset of the model parameters are given in Table 1.

#### 3.1. Scenario 1

In the first scenario, five nodes with 12 houses per node are considered with one battery potentially installed at each node. The optimal system initially purchases energy from the grid during the night but this drops to near zero during the day as energy from solar panels is able to meet demand, with excess generation used to charge the batteries. Demand from the grid is also reduced in the evening hours as the battery storage is discharged.

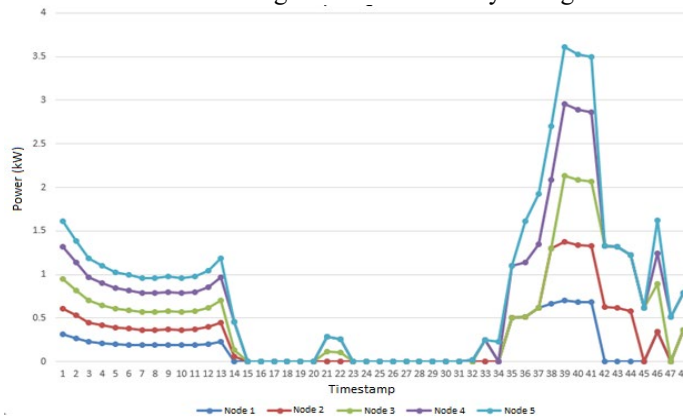


Figure 2 Scenario 1 external power purchased by nodes in representative autumn day

#### 3.2. Scenario 2

In the second scenario, the same five nodes are considered but now battery storage is present at two separate, more centralised, battery storage nodes. Here, decision variables in the model determine which distribution lines should be installed, alongside whether house nodes should be powered from the battery storage or from the grid. The model determined that four of the five nodes should be connected to battery storage but that for node 2 it was more viable to draw power from the grid. Here, the model behaves in a similar way to Scenario 1, with power drawn from the grid at night, especially during the evening when there is a surge in demand. From timestamp 39 onwards the model determined that it was more suitable to fulfil the demand from the centralised batteries.

#### 3.3. Scenario 3

Scenario 3 tests the validity of the model on a larger scale system, where the number of nodes is increased from 5 to 20 (now 3 houses per node), with 5 central battery storage nodes distributed across the neighbourhood. Unlike the previous results, here the model determined that the solar energy from the PV panels should be used solely to charge the battery storage, which is then used to provide power from timestamp 39 onwards. This is due to the higher cost of purchasing electricity from the grid after this point.

### 3.4. Cost Analysis

The cost components of the TAC objective function are shown in Table 2 and highlight the advantages and disadvantages of the various levels of decentralization within the community-level distributed PV farm with battery storage.

## 4. Discussion

The results of Scenarios 1 and 2 demonstrate the effectiveness of the model, with the power drawn from the grid dropping to nearly 0 kW during the day and the grid dependence being reduced in the evening due to the contribution from the stored battery power. The cost analysis shows the favourable values that can be obtained by placing batteries at each node (Scenario 1), shared among 12 households, in terms of all component metrics. Reduced distribution line costs come from lower distances between the individual houses and the battery storage nodes. Whilst power is still purchased from the grid, in Scenario 1 this cost is completely offset by the income made from selling power back to the grid. Scenario 3 shows the potential limits of the model, with comparatively large battery storage costs due to high fixed costs and low PV penetration. The results indicate that energy is mainly purchased from the grid during the day whilst the PV are used to charge the battery storage, subverting the original aim of the model. The high costs of purchasing this power from the grid greatly increases the TAC, making it an economically unattractive option. Note that in different locations with different price structures and regulatory conditions, the model results are likely to alter significantly. By examining the changes in the different components that make up the TAC, these results can be used to study the tradeoffs between centralisation/decentralisation of the DES, especially in terms of the relative storage and distribution line costs.

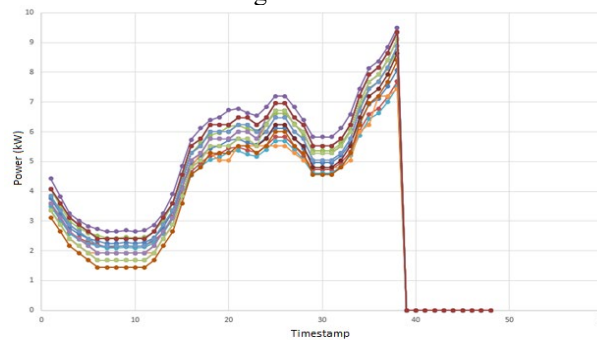


Figure 3 Scenario 3 Power Drawn from Grid

Table 2: TAC Components

Variables	Scenario 1	Scenario 2	Scenario 3
Annual grid cost	£734.89	£10,842.97	£340,677.40
Annual investment cost of PVs	£11,595.10	£47,895.23	£2,986.81
Annual operational cost of PVs	£7,441.30	£9,786.59	£1,170.78
Annual investment cost of battery	£284.36	£1,754.78	£8,625.82
Annual operational cost of battery	£240.40	£1,156.23	£5,556.47
Export income from selling electricity	£932.76	£9,776.87	£570.45
Distribution line cost	£4,124.33	£10,596.89	£26,589.87
Total annual cost	£23,487.62	£72,255.82	£385,036.70
Total annual cost per household	£391.46	£1,204.26	£6,417.28

## 5. Conclusion

In this study, an MILP-based optimisation model is proposed to minimize the total annual cost of a DES at the community level, with three different scenarios to demonstrate its applicability in a neighbourhood of Karachi, Pakistan. The suggested model uses solar PV as the DER, combined with battery storage. It also provides cost-based feasibility analysis and can determine design parameters such as the number of solar PVs required, the placement of distribution lines and the size and number of batteries to be installed. Whilst the given scenarios are for a local neighbourhood, the design parameters of a different environment can be calculated by providing the ground parameters and constraints of the specific area. The output of the model provides both investment and operational costs for the system as well as providing a tool to assess the correct scale of centralisation/decentralisation for community energy in a given neighbourhood. The three scenarios presented in this research validate the ability of the model to select the optimum energy sources based on cost efficiency and by examining the different components of the TAC, the relative benefits to prosumers and the larger community can be seen.

## References

- D. C. Alvarado, S. Acha, N. Shah & C. N. Markides, 2016, A Technology Selection and Operation (TSO) optimisation model for distributed energy systems: Mathematical formulation and case study, *Applied Energy*, 180, 491–503
- I. A. de Mel, O. V. Klymenko & M. Short, 2021, Levels of Approximation for the Optimal Design of Distributed Energy Systems, *Computer Aided Chemical Engineering*, 50, 1403–1408
- I. de Mel, O. V. Klymenko & M. Short, 2022, Balancing accuracy and complexity in optimisation models of distributed energy systems and microgrids with optimal power flow: A review, *Sustainable Energy Technologies and Assessments*, 52
- A. González-Briones, J. Prieto, F. De La Prieta, E. Herrera-Viedma & J. M. Corchado, 2018, Energy optimization using a case-based reasoning strategy, *Sensors (Basel)*, 18, 3, 865-892
- IBM, 2021, User's Manual for CPLEX 20.1.0, <https://www.ibm.com/docs/en/icos/20.1.0?topic=cplex-users-manual>
- I. Kantor, J.-L. Robineau, H. Büttin & F. Maréchal, 2020, A Mixed-Integer Linear Programming Formulation for Optimizing Multi-Scale Material and Energy Integration, *Frontiers in Energy Research*, 8
- J. Khan & M. H. Arsalan, 2016, Estimation of rooftop solar photovoltaic potential using geo-spatial techniques: A perspective from planned neighborhood of Karachi – Pakistan, *Renewable Energy*, 90, 188–203
- J. F. Marquant, R. Evins, L. A. Bollinger & J. Carmeliet, 2017, A holarchic approach for multi-scale distributed energy system optimisation, *Applied Energy*, 208, 935–953.
- E. D. Mehleri, H. Sarimveis, N. C. Markatos & L. G. Papageorgiou, 2012, A mathematical programming approach for optimal design of distributed energy systems at the neighbourhood level, *Energy*, 44, 1, 96–104
- A. J. Pimm, T. T. Cockerill & P. G. Taylor, 2018, The potential for peak shaving on low voltage distribution networks using electricity storage, *Journal of Energy Storage*, 16, 231–242
- J. Silvente, G. M. Kopanos, E. N. Pistikopoulos & A. Espuña, 2015, A rolling horizon optimization framework for the simultaneous energy supply and demand planning in microgrids, *Applied Energy*, 155, 485–501
- B. Zakeri, G. C. Gisse, P. E. Dodds & D. Subkhankulova, 2021, Centralized vs. distributed energy storage – Benefits for residential users, *Energy*, 236

# Integrating environmental sustainability in next-generation biopharmaceutical supply chains

Miriam Sarkis<sup>a</sup>, Jesslyn Fung<sup>a</sup>, Ming Hei Lee<sup>a</sup>, Andrea Bernardi<sup>a</sup>, Nilay Shah<sup>a</sup>,  
Maria M. Papathanasiou<sup>a,\*</sup>

<sup>a</sup>*Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom*

*\*[maria.papathanasiou11@imperial.ac.uk](mailto:maria.papathanasiou11@imperial.ac.uk)*

## Abstract

Maximizing product availability to the public and minimizing costs are primary objectives in the biopharmaceutical sector. Nevertheless, awareness of the environmental sustainability of supply chain operations is becoming increasingly relevant in recent years. To assist decision-makers in balancing financial and environmental sustainability we present an optimization framework which determines candidate supply chain structures network designs and operational plans. Supply chain structures are assessed with respect to total cost and environmental score, with the latter integrating environmental impacts related to climate change, water usage and energy consumption. A Pareto set of candidate solutions is found which provides insights in complex trade-offs between impact categories and cost: centralized manufacturing is selected to lower unit production cost and better use water resources, whilst decentralized manufacturing improves energy usage. Emissions from CO<sub>2</sub> are lowered through cost minimization.

**Keywords:** mathematical programming, supply chain optimization, pharmaceutical manufacturing, environmental sustainability.

## 1. Introduction

The biopharmaceutical industry is uniquely placed to create value for the society, with novel therapeutics enabling the cure of an increasing number of life-threatening diseases. The significant societal impact of the life science sector drives manufacturers to orchestrate and expand resources to maximize product availability to patients from clinical to commercial stages. Minimizing capital and operating costs translate in higher returns on upfront R&D and capacity investments. Recent innovation is brought by Advanced Therapy Medicinal Products (ATMPs), which demonstrate promising results in the cure of life-threatening diseases, including cancer and neurodegenerative diseases. Gene therapies are ATMPs that utilize vectors to deliver extra-chromosomal material into target cells. The sector is expected to grow significantly with the FDA forecasting approvals for 50-60 cell and gene therapy products by 2025 (FDA, 2019). Viral vectors are at the forefront of gene therapy manufacturing and delivery and represent a case where supply chains are transitioning from clinical to commercial stages. Therefore, increasing pressure from governmental agencies to improve sustainability of the industrial operations is expected (HM Government, 2021; EFPIA, 2022; NHS 2022). In biopharmaceutical manufacturing, process steps are often water and energy intensive, while pollutants are emitted during distribution.



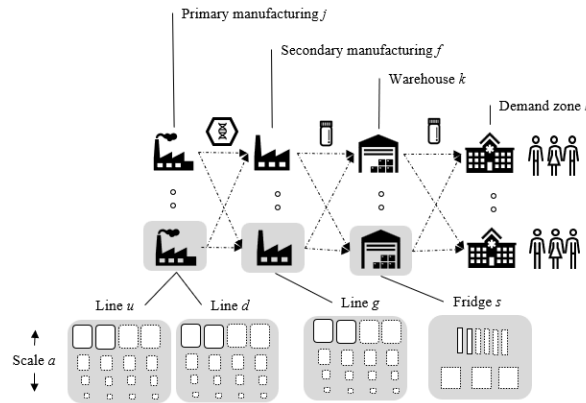


Figure 1: Schematic of supply chain superstructure

In this space, computer-aided decision-making can help assess trade-offs between financial and environmental sustainability. Integrating sustainability in supply chain planning has been explored in PSE operations research with multi-objective optimization approaches (Guillén-Gosálbez & Grossman, 2009; You *et al.* 2012; Barbosa-Póvoa *et al.*, 2018). A focus on the life science sector is presented in Duarte *et al.* (2022), where a large-scale vaccine multi-objective supply chain planning problem is integrated with life cycle analysis (LCA). In this work, we focus on a case study from viral vector-based therapeutics and consider a set of additional model decisions relevant to industrial scenarios, including selection of production scale and storage. A multi-objective optimization problem is presented, which compares minimization of cost and LCA-based supply chain environmental footprint.

## 2. Methodology

This study is based on a previously developed mixed-integer linear programming model presented in Sarkis *et al.* (2022), which consists in a multi-site capacity and distribution optimization for viral vector supply chains, where total costs are minimized and manufacturing is represented in terms of upstream (USP), downstream (DSP) and fill and finish (F&F) processes (Fig.1). Candidate locations are selected for primary manufacturing  $j$  comprising USP  $u$  and DSP  $d$  lines and available at a range of scales  $a$ , secondary manufacturing process  $f$  comprising  $g$  parallel filling lines and available in scales  $a$ , storage nodes  $k$  and demand zones  $l$ . Allowed node connections are  $j$ - $f$ ,  $f$ - $k$ ,  $k$ - $l$ ,  $f$ - $k$  and are generalized as  $n$  and  $n'$  (Table 1). Viral vector-based product  $i$  is considered.

### 2.1. Data collection

Information regarding a generalized viral vector processes was obtained from the literature. A techno-economic model for primary and secondary manufacturing platform for lentivirus-based vectors was developed in SuperPro Designer (Intelligen) to calculate batch sizes  $x_{a,i}$ , process times  $\alpha_u$ ,  $\alpha_d$ ,  $\alpha_g$ , and cycle times  $r_u$ ,  $r_d$ ,  $r_g$ , costs, resources  $w$  consumed (water, energy) or emissions ( $\text{CO}_2$ )  $EF_{a,w}^u$ ,  $EF_{a,w}^d$ ,  $EF_{a,w}^g$  by the process at different scales  $a$  of primary manufacturing, namely 50, 200, 1000 L and 2000 L bioreactor working volumes and secondary manufacturing, namely 50, 100, 400 vials  $\text{min}^{-1}$  filling rates. Viral vectors must be stored at low temperatures, therefore information regarding commercially available ultra-low storage freezers was collected (Martínez-Molina *et al.*, 2020). Specifically, costs  $C_{k,s}^{\text{CAP}}$  and capacity  $V_{\text{box}}$  for MATOS PLUS Cloud 300 UF (Cloud) and MATOS PLUS Eco 300 UF (Eco) freezers are used. Finally, information regarding

packaging  $m_{box}$  and shipping was obtained, including distances between nodes, costs, and emissions.

*2.2. Supply chain network model*

Given the inputs presented in Section 2.1, the optimization framework determines (i) supply chain structure, (ii) investment in new capacity at scale  $a$ , (iii) production levels in terms of batches at scale  $a$  of product  $i$   $B_{j,u,a,i}$ ,  $B_{j,d,a,i}$ ,  $B_{f,g,a,i}$  processed by  $u$ ,  $d$  and  $g$  respectively (iv) transport flows  $Q_{n,n',a,i}$  between nodes  $n$  and  $n'$ , (v) emissions from supply chain activities. Production is planned in campaigns, with number of batches and allocated times calculated as a function of process times for  $\alpha_u$ ,  $\alpha_d$  and  $\alpha_g$  and cycle times for  $r_u$ ,  $r_d$  and  $r_g$ . In this work, a more detailed description of storage and transport is presented compared to Sarkis *et al.* (2022). The number of freezers of type  $s$  at the warehouse  $k$ ,  $N_{k,s}$ , is found as a function of inbound flows of product  $i$  and freezer capacity (Eq.1). Capital cost of storage is estimated based on type of freezers installed and their cost  $C_{k,s}^{CAP}$  (Eq.2), whereas operating costs are calculated in terms of handling cost (Eq.3). Transport costs include contributions from the cost of freight  $TC^{freight}$  and dry ice requirements  $TC^{DI}$  for cold-chain shipping (Eq.4-5); which are a function of distances between nodes ( $D_{n,n'}$ ) and dry ice requirements per dose ( $C^{DI}$ ), respectively.

Table 1: Model equations to augment framework presented in Sarkis *et al.* (2022)

N. Description	Equation	N.
Number of freezers	$N_{k,s} = \frac{CAP_s}{DP_{vial}} \sum_{f,a,i} Q_{f,k,a,i} x_{a,i}$	$\forall k, s$ 1
Capital cost of storage	$TC_k^{CAP} = \sum_k C_{k,s}^{CAP} N_{k,s}$	2
Operating cost of storage	$TC_k^{OP} = \sum_{t,l} C^{OPki} Q_{t,k,l,a,i} x_{a,i}$	3
Cost of transport (freight)	$TC^{freight} = t_c \sum_{n,n',a,i} D_{n,n'} x_{a,i} (Q_{n,n',a,i})$	4
Cost of transport (dry ice)	$TC^{DI} = m^{DI} C^{DI} \sum_{n,n',a,i} x_{a,i} (Q_{n,n',a,i})$	5
Storage emissions	$E^k = \sum_{k,s} N_{k,s} \varphi_s$	6
Emission from DI production	$E^{DI} = m^{DI} \sigma^{DI} \sum_{n,n',a,i} x_{a,i} (Q_{n,n',a,i})$	7
Load from transport $n$ to $n'$	$L_{n,n'} = \sum_{a,i} Q_{n,n',a,i} x_{a,i} (m^{DI} + m_{box})$	$\forall n, n'$ 8-11
Emission from transport $n$ to $n'$	$E^{n,n'} = \sigma^{trans} (\sum_{n,n'} D_{n,n'} L_{n,n'})$	$\forall n, n'$ 12-14
Total emissions from transport	$E^t = E^{DI} + \sum_{n,n'} E^{n,n'}$	16
Emission from manufacturing	$E_w^j = \sum_{j,u,d,a,i} EF_{a,w}^u B_{t,v,a,u,i} + EF_{a,w}^d B_{t,j,a,d,i} + \sum_{f,g,a,i} EF_{a,w}^g B_{t,f,a,g,i}$	$\forall w$ 17
Impact factor for manufacturing	$I_w^j = \frac{E_w^j}{NF_w}$	$\forall w$ 18
Impact factor for storage	$I^k = \frac{E^k}{NF_{energy}}$	19
Impact factor for transport	$I^t = \frac{E^t}{NF_{CO2}}$	20
Environmental score	$ES = \sum_w WF_w I_w^j + WF_{CO2} I^t + WF_{energy} I^k + WF_{CO2} I_{penalty}$	21
Total cost	$TC = TC_j^{CAP} + TC_j^{OP} + TC_k^{CAP} + TC_k^{OP} + TC^{freight} + TC^{DI}$	22

Table 2: Summary of normalization (NF) and weighting (WF) factors for each environmental impact considered in this work. NF has been scaled down by a factor of  $10^8$  (Sala *et al.*, 2018).

Impact category	Units	NF <sub>w</sub>	WF <sub>w</sub>
Climate change	kg CO <sub>2</sub> -eq	$5.74 \times 10^5$	56%
Water usage	kg H <sub>2</sub> O-eq	$7.86 \times 10^8$	22%
Energy consumption	kWh	$1.24 \times 10^6$	22%

### 2.3. Environmental constraints

A set of environmental constraints is formulated to quantify an environmental score (ES) of each candidate optimization solution, with higher ES corresponding to larger environmental footprint. Emissions from manufacturing nodes  $E_w^j$  account for the resource requirements of water and electricity per batch for each process section via  $EF_{a,w}^u$ ,  $EF_{a,w}^d$ ,  $EF_{a,w}^g$  (Eq.17). CO<sub>2</sub> emissions result from manufacturing and transport, due to CO<sub>2</sub> emitted per kg of DI produced  $E_{DI}^t$  and emitted CO<sub>2</sub> during shipping  $E^{n,n'}$ . These are respectively calculated as a function of dry ice requirements  $m^{DI}$ , emissions  $\sigma^{DI}$ ,  $\sigma^{trans}$  and the loads for each transport link  $L_{n,n'}$  (Eq 7-16). Storage emissions  $E^k$  result from electricity usage  $\varphi_s$  for each storage type  $s$  (Eq. 6). The normalised impact for each of these emissions can be obtained, namely  $I_w^j$ ,  $I^k$  and  $I^t$  (Eq 18-20). The environmental score is calculated as weighted sum of the normalized impacts (Crenna *et al.*, 2019) (Eq. 21), implementing weights  $WF_w$  obtained from Sala *et al.* (2018) (Table 2).

### 2.4. Multi-objective Optimization

Multi-objective optimization provides means of considering multiple objectives and analyze trade-offs. In this study, we compare the weighted-sum method (WSM) and the  $\varepsilon$ -constrained method, both of which enable the multi-objective problem to be reformulated as a single-objective one. The WSM assigns weights  $\alpha$  and  $(1-\alpha)$  to each objective respectively, ES and TC, and systematically optimizes the sum of objective functions obtained for realizations of  $\alpha$ . TC and ES are normalized using the min-max method and limited between 0 and 1, preventing one attribute with large initial values from dominating. The  $\varepsilon$ -constrained method optimizes one objective function at the time, with an upper-level constraint placed on the other objective. The  $\varepsilon$ -constraint method is used to investigate potential non-convexity along the Pareto curve.

## 3. Results & Discussion

### 3.1. Single-objective optimization

Single-objective optimization was performed to investigate its decisions under two scenarios: (i) minimizing total cost (*min*-TC) and (ii) minimizing environmental score (*min*-ES). *min*-TC yielded the minimum TC and maximum ES, whereas *min*-ES produced the maximum TC and minimum ES. Alternative capacities of sites and network structures were selected in each scenario to meet equivalent demands across  $l$  (Figure 1). *min*-TC favors a centralized manufacturing as opposed to a more distributed configuration in *min*-ES. The selection of a single 2000 L scale bioreactor is justified by economies of scale, as unit production cost lowered at higher scales. Instead, a 2000 L is 2.3 times more energy intensive than a 1000 L, which means energy consumed per unit product is lower at lower scales. Hence in *min*-ES the optimal configuration involves 2 1000L sized USP and DSP lines. Although a larger scale would result in better usage of water per unit produced, the weighting by  $NFs$  and  $w$  in *ES* is lower than that of energy intensity. Therefore, the model prioritizes minimizing of energy usage, followed by CO<sub>2</sub> and water usage. With respect to secondary manufacturing the 50 vials min<sup>-1</sup> fill-finish plant is sufficient to process the incoming flows from the DSP process. Therefore, a higher number of batches of smaller scale are produced. In *min*-ES, only 4 batches were produced at small scale 50 vials min<sup>-1</sup> F&F size, with the rest processed by a single 100 vials min<sup>-1</sup> F&F. This is because larger scale improves water-efficiency in F&F. *min*-ES also results in the selection of Cloud as storage, as more energy-efficient; on the other hand *min*-TC favors Eco for its lower costs. With a total demand of 22500 doses, the capacity of a single freezer is sufficient.

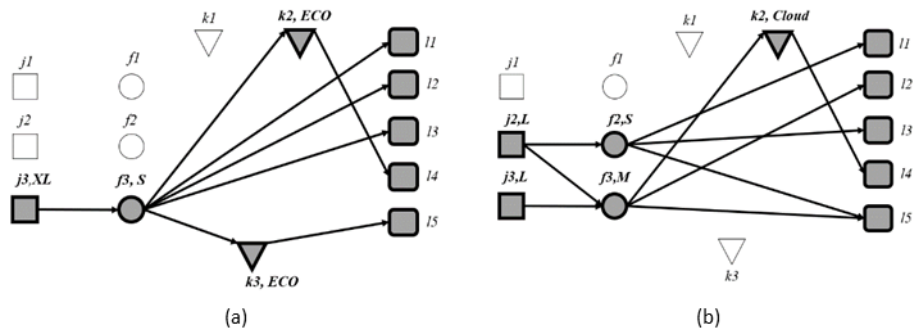


Figure 2: (a) single-objective optimisation network structures for (a) *min*-TC and (b) *min*-ES single-objective optimisation solutions. J: primary manufacturing, F: secondary manufacturing, K: storage, L: demand zones. Different scales are S (50 L), M (200 L), L(1000 L), XL(2000 L).

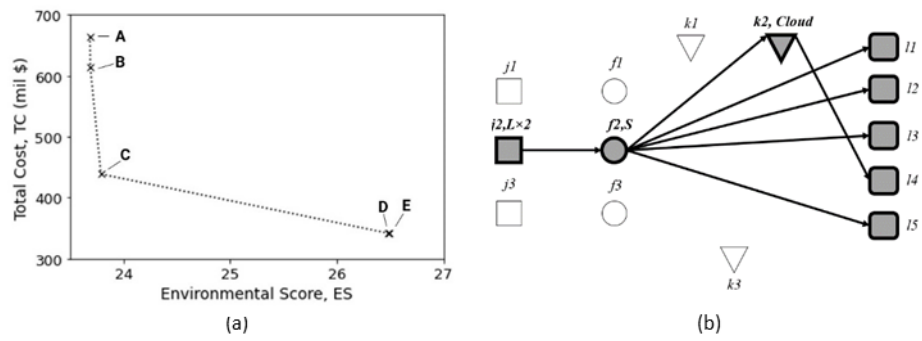


Figure 3: (a) multi-objective optimisation solutions on TC-ES Pareto frontier. Solutions correspond to different values of  $\alpha$ : A ( $\alpha=0$ ) or *min*-ES, B ( $0 < \alpha < 0.05$ ), C ( $0.05 < \alpha < 0.75$ ), D ( $0.75 < \alpha < 1$ ), E ( $\alpha=1$ ) or *min*-TC. (b) Network structure for C.

*min*-TC suggests installing  $k3$  as the shipping cost reduction over this route outweighs the cost of installing a fridge in the location. In *min*-ES, demand in  $l5$  was served directly by  $f2$  and  $f3$ , removing storage node  $k3$  with associated energy consumption.

### 3.2. Multi-objective optimization

The Pareto front constructed via WSM provides solutions which achieve a compromise between *min*-ES and *min*-TC extrema (Solution A & E, Figure 3a). At  $\alpha=0$ , solution E corresponds to the centralized SC of *min*-TC. As  $\alpha$  increases, the relative importance of ES on TC increases. The network scales-out favoring two lines of 1000 L scale at the same location (Solution C, Figure 3a & b) as opposed to one at 2000 L (Solution A). At  $\alpha=1$ , solution A is obtained, with two 1000 L lines 1 at different locations  $j2$  and  $j3$ . The methodology resulted in 5 identical solutions of that of WSM, confirming the lack of additional solutions and highlights the convexity of the Pareto front. Solution C presents lower costs to B, with lower capital and operating costs due to the absence of a second  $f1$  site and transport costs with shorter travel distances. Water-efficiency is improved with C implementing lower scales; however, this is outweighed by higher energy usage hence ES increases. By comparing C and D, solution C select two 1000 L USP lines and one DSP line. D instead opts for single USP, DSP and F&F lines at 2000 L scale.

This results in lower capital and operating costs per dose. D results 28% lower in TC and 10% higher ES and this is due again to higher net energy intensity of USP and DSP lines. C is found to be an example solution of a balanced compromise between TC and ES and covers a wide range of  $\alpha$ . Capacity allocations switch due to the relative importance of cost, water-efficiency, improved at higher scales and energy-efficiency improved at lower scale. ES benefits from lower CO<sub>2</sub> emissions for solutions where transport cost is lowered.

#### 4. Conclusion

This study presents an optimization framework to support decision-makers in biopharmaceutical sector in finding good compromises between environmental and financial sustainability. The problem of minimizing supply chain costs is integrated with life cycle analysis (LCA). Impacts brought by CO<sub>2</sub>, energy and water usage are combined in a dimensionless indicator for environmental impact. Results from minimizing total supply chain costs show how centralized manufacturing at larger scale also minimizes of transport distances and CO<sub>2</sub> emissions and selects cheaper storage options. This translates in higher emissions from storage, however better utilization of water resources. By contrast, minimizing environmental impact favors smaller scales of manufacturing as smaller batches are more energy-efficient, with a removed benefit from economies of scale. As the identified solutions are likely sensitive to the assumed weights assigned to each impact category, on-going work is assessing the impact of such inputs on optimization outputs and considering additional impact categories. These may include emissions from the selection of using plastics-based single-use versus stainless steel equipment and extending the analysis to a wider range of resources used in manufacturing and distribution of viral vector-based advanced therapeutics.

#### References

- A.P. Barbosa-Póvoa, C. da Silva, A. Carvalho, 2018, Opportunities and challenges in sustainable supply chain: an operations research perspective, *European Journal of Operations Research*, 268, 399-431.
- E. Crenna, M. Secchi, L. Benini, S. Sala, 2019, Global environmental impacts: data sources and methodological choices for calculating normalization factors in LCA, *International Journal of Life Cycle Assessment*, 24.
- I. Duarte, B. Mota, T. Pinto-Varela, A.P. Barbosa-Póvoa, 2022, Pharmaceutical industry supply chains: how to sustainably improve access to vaccines, *Chemical Engineering Research and Design*, 182, 324-341.
- G. Guillén-Gosálbez & I. Grossman, 2009, Optimal design and planning of sustainable chemical supply chains under uncertainty, *AIChE Journal*, 55(1), 99-121.
- EFPIA. 2022, *Environment, Health, Safety and Sustainability*.
- E. Martínez-Molina, C. Chocarro-Wrona, D. Martínez-Moreno, J.A. Marchal, H. Boulaiz, 2020, Large-scale production of lentiviral vectors: current perspectives and challenges, *Pharmaceutics*, 12, 1051.
- HM Government, 2021, *Life Sciences Vision*.
- S. Sala, A.K. Cerutti, R. Pant, 2018, Development of a weighting approach for the environmental footprint. Publications Office of the European Union: Luxembourg.
- M. Sarkis, K. Tak, B. Chachuat, N. Shah, M.M. Papanthasiou, 2021, Towards resilience in next-generation vaccines and therapeutics supply chains, *Computer Aided Chemical Engineering*, 51, 931-936.
- NHS England. 2022, *Delivering a 'Net Zero' National Health Service*
- F. You, L. Tao, D.J. Graziano, S.W. Snyder, 2012, Optimal design of sustainable cellulosic biofuel supply chains: multi-objective optimization coupled with life cycle assessment and input-output analysis, *AIChE Journal*, 58(4), 1157-1180.
- U.S. Food & Drug Administration, 2019, *Statement from FDA Commissioner*

# Techno-economic assessment of sustainable energy planning on renewable electricity demand-supply networks: A deep learning approach

Byeongmin Ha<sup>a</sup>, Taehyun Kim<sup>a</sup>, Jou-Hyeon Ahn<sup>a</sup>, Soonho Hwangbo<sup>a,\*</sup>

<sup>a</sup> *Department of Chemical Engineering, Gyeongsang National University, Jinjudae-ro 501, Jinju-si, Gyeongsangnam-do, 52828 Korea*  
*s.hwangbo@gnu.ac.kr*

## Abstract

Most renewable energy networks rely on wind and solar energy, known as variable renewable energy (VRE), and its generation process is heavily dependent on weather conditions, leading to fluctuations in the supply side. As a result, this study aims to: 1) develop an optimal forecasting model to predict the supply-demand balance, 2) provide different thresholds to generate potential scenarios, and 3) compare the scenarios using a techno-economic assessment. The optimal model in this study is a GRU, which has an  $R^2$  score of 0.994. The levelized cost of electricity (LCOE) ranges from 0.03 USD/kWh to 0.07 USD/kWh. The key conclusions of the study are as follows: 1) conversion factors are used to show that the processed data can be converted to match the feasible pattern of the target year's VRE data; 2) the sampling method accounts for uncertainties in future data and those caused by limited time-series data; 3) the optimal model can be identified by comparing various models using sample data as input; 4) the feasibility of scenarios consisting of a techno-economic component is validated by IRENA; and 5) the probability of LCOE can inform expected budget for energy policy.

**Keywords:** Variable renewable electricity, Forecasting model, Variational auto-encoder, Techno-economic assessment, National energy strategy

## 1. Introduction

The world is facing threats such as climate change and climate refugees caused by global warming. Since the signing of the 2015 Paris Agreement, many nations have been working to reduce carbon emissions and achieve a carbon-neutral economy. Solar and wind power generation, also known as variable renewable energy (VRE), are highly dependent on weather conditions. Despite this uncertainty, VRE accounts for a significant portion of renewable energy (RE) mix. Power generation is difficult to predict due to fluctuations and the costs of projects can vary greatly based on technological advancements and spatial and temporal regions. However, national policy-based energy sector projects must be evaluated to reach the goal of a carbon-neutral economy. To our knowledge, few studies have used deep learning models to perform techno-economic assessments based on national policy. The following steps are taken to achieve this goal: 1) Determine the generation/demand conversion factor using projected growth in RE sources and factors affecting generation and demand, 2) Use a DL-based sampling method to account for future uncertainties, 3) Calculate insufficient generation based on the scenario using a DL-based supply/demand prediction model, 4) Perform techno-economic assessment and find the optimal scenario. These steps allow for the consideration of future uncertainties and the validation of the feasibility of national policy.

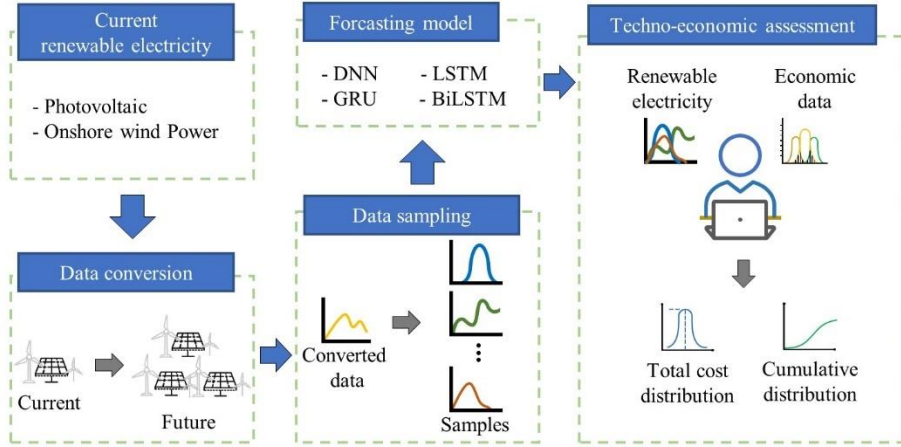


Fig. 1. Major steps in the techno-economic assessment.

## 2. Method

### 2.1. Gated recurrent unit

The GRU, a model based on recurrent neural networks (RNNs), has attractive performance in time-series data (Nam et al., 2020). The GRU revised from LSTM has only two gates and the hidden state, the GRU presents better performance for long-term forecasting compared to the LSTM. The primary mathematical equations of the GRU are as follows:

$$r_t = \sigma(W_{xr}^T \cdot x_t + W_{hr}^T \cdot h_{t-1} + b_r) \quad (1)$$

$$z_t = \sigma(W_{xz}^T \cdot x_t + W_{hz}^T \cdot h_{t-1} + b_z) \quad (2)$$

$$g_t = \tanh(W_{xg}^T \cdot x_t + W_{hg}^T \cdot (r_t \circ h_{t-1}) + b_g) \quad (3)$$

$$h_t = (1 - z_t) \circ h_{t-1} + z_t \circ g_t \quad (4)$$

where  $\tanh$  and  $\sigma$  are the hyperbolic tangent and the sigmoid function, respectively.  $W$ ,  $x$ , and  $b$  in Eq. (1) to (3) represent weights, the input data, and biases, respectively. The hidden state of the previous time ( $t-1$ ) and the input of the present time are used to calculate the reset gate ( $r_t$ ) and update gate ( $z_t$ ), which have a value between 0 and 1 by sigmoid (Eqs. (1) and (2)). Reset gate decides how much past hidden state should be forgotten (Eqs. (3) and (4)). Update gate determines information composition composed previous information and current information (Eq. (4)).

### 2.2. Variational auto-encoder

Variational auto-encoder (VAE) has a similar structure to auto-encoder, but it performs a different role. VAE, deep learning-based generative model, consists of the reconstruction term and regularization term and minimizes argument of these two terms (Qi et al., 2020). The reconstruction term explains how this model can accurately reconstruct the input data from a sampling function, using the negative loglikelihood function (Eq. (5)). Regularization term manipulates the samples to make samples rely on the prior assumption distribution (Eq. (6)).

$$\sum_i -\mathbb{E}_{q_\theta(z|x_i)} [\log(p(x_i|g_\theta(z)))] \quad (5)$$

$$KL(q_\theta(z|x_i)||p(z)) \quad (6)$$

$$\arg\min_{\theta, \phi} \sum_i -\mathbb{E}_{q_\theta(z|x_i)} [\log(p(x_i|g_\theta(z)))] + KL(q_\theta(z|x_i)||p(z)) \quad (7)$$

### 3. A case study

With the 3020 Implementation Plan, Korea plans to have 20% of its energy mix from renewable sources by 2030. To meet 3020 implement plan, Korea already meets target capacity of wind power, but solar power has to be supplemented (MOTIE, 2017).

Capital expenditure (CapEx), operating expenditure (OpEx), life span, and shortage of electricity supply are considered for components related to TEA. CapEx and OpEx estimations of PV based on IRENA are utilized for TEA-related to target year. CapEx and OpEx estimations of FPV are decreased by 5% and equal value compared to PV, respectively. Life span of PV is 20 years, and FPV ranges from 25 to 30. Since shortage of electricity related to target year is considered, supply/demand data are passed through 3 steps as follows: 1) converting the raw data to reference data based on national policy, 2) generating samples considering uncertainties, and 3) forecasting VRE supply/demand data using various models. The conversion factor consists of data that comes from the Korea Power Exchange (KPX), and generates 2030 reference data of supply/demand. Sampling method considers uncertainties of conversion and develops various scenarios. Techno-economic assessment evaluated scenarios and this operation was validated by LCOE.

### 4. Results and discussion

According to Korea's RE strategy, the conversion factor was calculated using the current capacity, OpEx, CapEx, and the capacity factor for 2030. Since components of conversion factors for solar power generations and demand do not change over period (e.g., capacity factors of solar power generations were assumed as constant), the conversion factor for each scenario was constant in the case of demand and solar power generations. However, wind power generation's capacity factor increases over period. The conversion factor of onshore ranged from 0.89 to 1.63, and offshore ranged from 276.53 to 506.97. Required capacity derived from shortage of electricity were considered for two types (i.e., PV 100% or FPV 100% additional projects). The conversion factor for offshore had a high value compared to other generation types since offshore generation plans to build lots of additional capacity.

Sample data, which was generated from reference data, had a large gap with reference data in the case of wind power generation (Fig. 2). The difference between solar power and wind power represents characteristic (i.e., solar power can function only during the day. Wind power is affected by many weather conditions.).



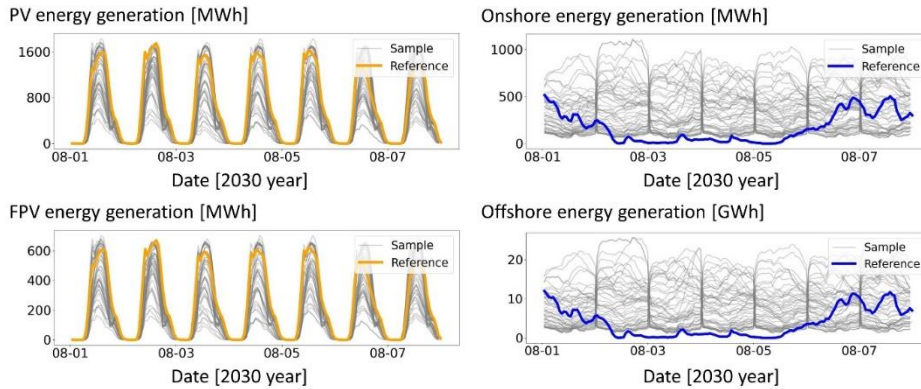


Fig. 2. Sample data and reference data of RE sources.

The  $R^2$  score of the GRU was 0.994, which was comparable to the average  $R^2$  score of 0.996 for other RNNs-based models, and the GRU had AIC/BIC values of 30,776 and 53,180, which are 6.5% and 15% lower than the LSTM for PV test data (Table. 1). The GRU was selected as the forecasting model in this study since it had comparable performances and lower values of AIC/BIC.

Table 1. A comparison of performance metrics and information criteria considering several forecasting models based on the PV test data; B, L, G, D, and A means BiLSTM, LSTM, GRU, DNN, and ARIMA.

	$R^2$	RMSE	MAE	WAPE [%]	AIC [ $\times 10^3$ ]	BIC [ $\times 10^3$ ]
B	0.997	28	17	3.7	81	284
L	0.994	40	25	5.6	32.9	62.7
G	0.994	40	26	5.7	30.8	53.1
D	0.987	60	44	9.6	-	-
A	0.948	120	74	16.1	-	-

After calculating the maximum and minimum RE demand/supply using the generated sample and GRU, the scenario was assumed as follows:

Best scenario: maximum supply, minimum demand

Worst scenario: maximum demand, minimum supply

Considering Korea's 2030 target of RE generation, RE generation was insufficient by 15,239,418 MWh/year and 71,035,093 MWh/year in the best scenario and the worst scenario, respectively. The high proportion of offshore wind farms with large fluctuations appears to be the main factor in the difference in RE generation between scenarios. Required capacity and the related costs were presented in Table. 2. The total annual cost of PV ranged from 460.9 to 4,911 million USD/year, and FPV ranged from 330.9 to 3,740 million USD/year.

Table 2. Costs according to scenarios.

	Best scenario		Worst scenario	
	PV 100%	FPV 100%	PV 100%	FPV 100%
Required Capacity [MW]	12,363	10,670	57,151	49,322
CapEx [million USD/year]	337.3	224.2	3,825	2,803
OpEx [million USD/year]	123.6	106.7	1,086	937
Total cost [million USD/year]	460.9	330.9	4,911	3,740

Since several FPV expansions were already anticipated for Korea, PV-based validation was conducted using IRENA LCOE based on the 2030 global average. The theoretical LCOE, which ranged from 0.02 to 0.08, was revised from 0.03 to 0.07 to consider Korea’s technical level of RE. In Table. 3, the theoretical total cost was calculated by multiplying the theoretical LCOE and supply shortage, and the measured total cost presented the overall cost of this study. The results of validation show 0.8-1.2 absolute error, which means the feasibility of Korea's net-zero carbon.

Table 3. Cost verification by scenario.

	Best scenario	Worst scenario
LCOE [USD/kWh]	0.03	0.07
Theoretical total cost [million USD/year]	457.2	4,972
Measured total cost [million USD/year]	460.9	4,911
Absolute error [%]	0.8	1.2

Fig. 3 shows the cumulative density function of 1,000 total cost scenarios using key variables (i.e., CapEx, OpEx, and capacity). The result of this study (blue solid line) has a similar shape to the normal distribution (green solid line), and is most likely to form a price range of around 2,000 million USD. About 490 million USD will be allocated to RE in Korea in 2022, and it appears that this budget will need to be gradually increased in order to meet the country's 2030 RE objective.

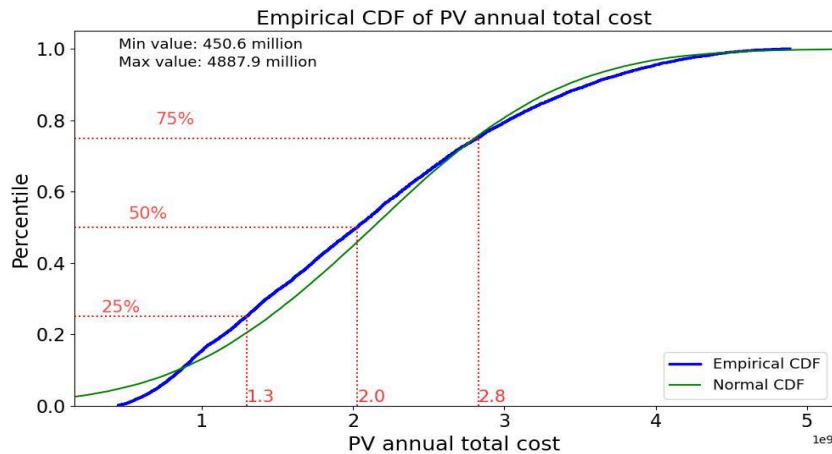


Fig. 3. Empirical cumulative density function of PV annual total cost.

## 5. Conclusions

This study carried out a techno-economic assessment based on national policy for a target year. A case study in Korea was used to calculate the LCOE for RE networks under different scenarios. The following steps were taken to perform the techno-economic assessment of the scenarios: 1) Consider the conversion factor estimated by the installed capacity and capacity factor over time, 2) Use a deep generative model-based sampling method to account for uncertainties, 3) Forecast 2030 supply and demand data using machine learning models. In this study, the GRU-based forecasting model was found to be the optimal model with the best performance and the lowest information scores. Using 2030 supply and demand data and technical data, a techno-economic assessment was carried out for a carbon-neutral economy based on national policy. The results showed that FPV are up to 28% more cost-effective compared to PV. The validity of the results was confirmed by the LCOE of each scenario, which showed an absolute error range of 0.8–1.2 compared to the theoretical total cost, indicating that a carbon-neutral economy in Korea is feasible. Further research will be conducted to consider various factors and extend the study to a global model.

## References

- MOTIE, 2017, Renewable energy 3020 implementation plan, Seoul, Korea., The Government of the Republic of Korea.
- K. J. Nam, S. Hwangbo, C. K. Yoo, 2020, A deep learning-based forecasting model for renewable energy scenarios to guide sustainable energy policy: A case study of Korea, *Renewable and Sustainable Energy Reviews*, 122(March 2019); 109725.
- Y. Qi, W. Hu, Y. Dong, Y. Fan, L. Dong, M. Xiao, 2020, Optimal configuration of concentrating solar power in multienergy power systems with an improved variational autoencoder, *Applied Energy*, 274(June); 115124.
- G. R. Timilsina, 2021, Are renewable energy technologies cost competitive for electricity generation?, *Renewable Energy*, 180; 658–672.
- H. Wang, Z. Lei, X. Zhang, B. Zhou, J. Peng, 2019, A review of deep learning for renewable energy forecasting, *Energy Conversion and Management*, 198(April); 111799.

# Evaluation of centralized/decentralized configuration schemes of CO<sub>2</sub> electrochemical reduction-based supply chains

Thijmen Wiltink<sup>a</sup>, Stijn Yska<sup>a</sup>, Andrea Ramirez<sup>a</sup>, Mar Pérez-Fortes<sup>a</sup>

<sup>a</sup> *Department of Engineering, Systems, and Services, Faculty of Technology, Policy, and Management, Delft University of Technology, Jaffalaan 5, 2628 BX Delft  
T.J.Wiltink@TUDelft.nl*

## Abstract

Electrochemical reduction of CO<sub>2</sub> (CO<sub>2</sub>ER) is an emerging technology with the potential to limit the use of fossil-based feedstocks in the petrochemical industry by converting CO<sub>2</sub> and renewable electricity into useful products such as syngas. Its successful deployment will depend not only on the technology's performance but also on its integration into the supply chain. In this work, a facility location model is used to gain insights regarding the capacity of CO<sub>2</sub>ER plants that produce syngas and the implications for the central/decentral placement of these CO<sub>2</sub>-based syngas plants. Different optimal configurations are examined in the model by changing the syngas transport costs. In this exploratory case, the results indicate that centralization is only an option when the syngas and CO<sub>2</sub> transport costs are similar. When syngas transport is more expensive, decentralizing CO<sub>2</sub>-based syngas plants in the supply chain appears more feasible.

**Keywords:** CO<sub>2</sub> electrochemical reduction; CO<sub>2</sub> utilization, supply chain modeling, optimization, supply chain configurations

## 1. Introduction

The European Green Deal includes the 2030 Climate Target Plan targeting a greenhouse gas (GHG) emission reduction of at least 55% in 2030 compared to 1990 levels and a net-zero GHG emissions target by 2050 (European Commission, 2019). This requires a drastic change in the petrochemical industry, which is challenging due to the industry's dependence on fossil fuels as its carbon feedstock. Currently, processes using CO<sub>2</sub> as an alternative feedstock are being developed. In the electrochemical reduction of CO<sub>2</sub> (CO<sub>2</sub>ER), CO<sub>2</sub>, water, and electricity are converted into a range of intermediates and final products that can be used for further chemical and fuel synthesis in multiple sectors. CO<sub>2</sub>ER is not yet a mature technology, and ultimately it will be integrated into already existing supply chains (SCs), which affects the SC configurations in terms of implementation scale and (de)centralization of the technology. In order to study the impact of the CO<sub>2</sub>ER technology in connection with its SC, a facility location optimization model was developed in the current work to understand the potential trade-offs of decentralized/centralized configurations when replacing fossil-based syngas with CO<sub>2</sub>-based syngas from a CO<sub>2</sub>ER process. Centralized/decentralized business cases for CO<sub>2</sub>ER are explored from an SC perspective in the European petrochemical context using a hypothetical case study. Syngas is chosen as the product of interest, as it is a large-scale fossil-based industrial product, commercialized as a versatile commodity that can be used as a precursor for a wide range of processes in the petrochemical industry (Choe et al., 2022; Ebbehøj, 2015). Syngas is a mixture of hydrogen and carbon monoxide (CO). The CO molecule in syngas makes the mixture toxic, which adds safety requirements to the

transportation and handling of this feedstock. Alternative feedstocks for syngas production are described in SC literature; the primary focus has so far been on the upstream part of the SC (e.g., (Ahmadvand and Sowlati, 2022; Marufuzzaman et al., 2016)). In these studies, the current business model is employed where syngas is produced on-site (i.e., without storage and transport). Transporting syngas to establish a centralized market is not considered. This work investigates how an industrial scale CO<sub>2</sub>-based syngas SC could look like for different syngas transportation costs in relation to CO<sub>2</sub> transport costs. The facility location model selects the optimal locations, the number of CO<sub>2</sub>-based syngas plants, the plant capacities, and connections between echelons in a three-stage SC, see Figure 1A.

## 2. Mathematical model formulation of CO<sub>2</sub>-based syngas supply chain

A fixed charge facility location problem (FLP) is formulated as a mixed integer linear program (MILP). This model is an uncapacitated multiple allocation FLP; the maximum size of the CO<sub>2</sub>-based syngas plant is not constrained, and individual echelons can connect to multiple other echelons. The problem is formulated as a p-median problem, the foundations of which are laid by Hakimi (1964). The model is based on the SiLCaRD (simultaneous location of central and regional distribution facilities) model formulated in the work of Götzinger (2013).

Mass balance constraints ensure that the amount of CO<sub>2</sub> captured in the system is in equilibrium with the amount of syngas at the demand locations and that mass conservation is guaranteed. Location constraints specify where to establish and open the CO<sub>2</sub>-based syngas plants, while allocation constraints select and connect the different echelons in the system. The piecewise linear transport constraints allow for dealing with non-linearities in transport. The mathematical is elaborated below and the decision variables are emphasized in bold font:

$H$	set of CO <sub>2</sub> source locations, indexed by $h$
$J$	set of potential CO <sub>2</sub> -based syngas plant locations, indexed by $j$
$K$	set of syngas demand locations, indexed by $k$
$Q$	set of transport segments of the piecewise linear cost function, indexed by $q$
$A_q$	slope [€/2020/kilometer/Mtonne transported] of segment $q$ in the piecewise linear transport cost function
$B_q$	intercept [€/2020] of segment $q$ in the piecewise linear transport cost function
$C_{hj}$	distance [kilometers] from CO <sub>2</sub> source $h$ to CO <sub>2</sub> -based syngas plant $j$
$C_{jk}$	distance [kilometers] from CO <sub>2</sub> -based syngas plant $j$ to syngas consumer $k$
$CC_h$	average carbon capture cost [€/2020/Mtonne] at the CO <sub>2</sub> source (at cluster level)
$CF_{st}$	cost factor of syngas transport [€/2020 syngas transported /€/2020 CO <sub>2</sub> transported]
$SG_{Conf}$	CO <sub>2</sub> -to-syngas conversion factor [Mtonne syngas/Mtonne CO <sub>2</sub> ]
$\mathbf{y}q_{hj}$	1 if the CO <sub>2</sub> -based syngas plant $j$ is assigned to CO <sub>2</sub> source $h$ in segment $q$ of the piecewise linear transport function, 0 if not
$\mathbf{y}q_{jk}$	1 if the syngas consumer $k$ is assigned to the CO <sub>2</sub> -based syngas plant $j$ in segment $q$ of the piecewise linear transport function, 0 if not
$\mathbf{z}_{hjq}$	size [in Mtonne of CO <sub>2</sub> transported] from CO <sub>2</sub> source $h$ to CO <sub>2</sub> -based syngas plant $j$ in segment $q$ of a piecewise linear function
$\mathbf{z}_{jkq}$	size [in Mtonne syngas] transported from CO <sub>2</sub> -based syngas plant $j$ to syngas consumer $k$ in segment $q$ of a piecewise linear function

The objective function aims to minimize the yearly capture costs at the industrial cluster (1a), the yearly pipeline transport cost of CO<sub>2</sub> from CO<sub>2</sub> sources to CO<sub>2</sub>-based syngas plants (1b), and the yearly (hypothetical) pipeline transport cost of syngas from CO<sub>2</sub>-based syngas plants to syngas consumers (1c). The amount of CO<sub>2</sub> captured ( $z_{hj q}$ ) and the size of the electrolyzer ( $z_{jk q}$ ) is determined as a result of the minimization of the objective function to fulfill the syngas demand in the SC. A piecewise linear transport cost function is applied to all the transport links between the echelons in the model. The amount of CO<sub>2</sub> (from source to CO<sub>2</sub>-based syngas plant) is multiplied with a CO<sub>2</sub>-to-syngas conversion factor to calculate the plant's syngas production; see equation (2). Objective function - minimize:

$$\begin{aligned}
 & \sum_{h \in H} \sum_{j \in J} \sum_{q \in Q} z_{hj q} * CC_h \quad (1a) \quad + \sum_{h \in H} \sum_{j \in J} \sum_{q \in Q} A_q * z_{hj q} \quad (1b) \quad + \sum_{j \in J} \sum_{k \in K} \sum_{q \in Q} (A_q * z_{jk q} * C_{jk} + B_q * y_{qjk}) \quad (1c) \\
 & * C_{hj} + B_q * y_{qhj} \quad * CF_{st} \\
 \hline
 & SG_{Conf} * \sum_{h \in H} \sum_{q \in Q} z_{hj q} = \sum_{j \in J} \sum_{q \in Q} z_{jk q} \quad \forall j \in J \quad (2)
 \end{aligned}$$

### 3. Case study – CO<sub>2</sub>-based syngas supply chains

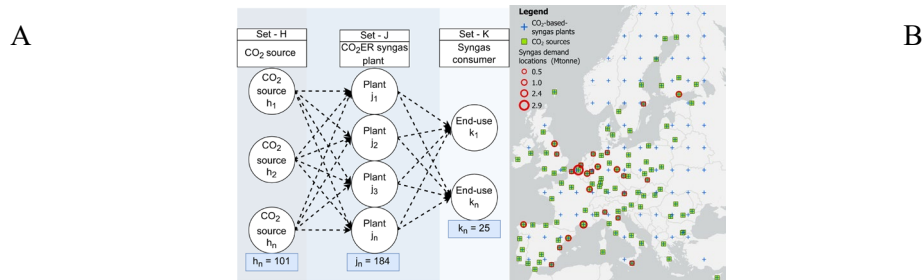


Figure 1 - A – Representation of the three-echelon supply chain model described in this case study. B – Industrial CO<sub>2</sub> sources in Europe, potential CO<sub>2</sub>-based syngas plant locations, and syngas demand locations with sizes.

The model uses geographical data from industrial clusters in Europe (EU-27+UK). It uses a discrete grid of potential CO<sub>2</sub>-based syngas plant locations. A 100 x 100 grid was placed to create potential location sites for CO<sub>2</sub>ER plants from the borders of Portugal to Finland. For the purpose of this study, plant locations over the sea were disregarded, and the industrial cluster locations were added as potential location sites, resulting in 185 potential electrolyzer locations, see Figure 1B. The current case study contains CO<sub>2</sub> emission data from 101 industrial clusters (with 944 individual plants within these clusters from 9 types of industries: ammonia, cement, lime, iron & steel, refining, petrochemical, oil & gas, power generation, and aluminum). Industry-specific CO<sub>2</sub> impurities were not considered, and it was assumed that CO<sub>2</sub> is in the same conditions irrespective of its source. The dataset is available via ArcGIS online (Boston Consulting Group, 2021) and is based on the European Pollutant Release and Transfer Register (European Environment Agency, 2020). The following considerations were taken into account:

- The worldwide syngas market was 150 Mtonne in 2018 (Jouny et al., 2018), with a CAGR of 9.5% it is estimated to reach 180 Mtonne in reference year 2020 (Inkwood Research, 2017). Europe uses approximately 22% of global syngas, with 51% in the

petrochemical industry (IMARC, 2021). The estimated syngas demand in Europe's petrochemical industry used in this work was 20.3 Mtonne/year. In the model, the demand needs to be assigned to specific locations. In order to assign different syngas demands to different (hypothetical) sites, a syngas plant size of 0.482 Mtonne/year was assumed based on the syngas demand of a standard methanol plant size of 0.4 Mtonne/year (Ebbehøj, 2015). To fulfill the syngas demand using this standardized size, 42 CO<sub>2</sub>-based syngas plants would be necessary. 42 of the 84 petrochemical plants in the European dataset were randomly selected to serve as syngas demand locations. The demand was summed when the selected plants were part of the same cluster.

- The costs of CO<sub>2</sub> transport were based on (hypothetical) existing onshore pipelines (i.e., only operating costs are considered). A piecewise linearization is used based on the values by d'Amore (2021).
- This work assumes an existing syngas transportation infrastructure (i.e., only operating costs are considered). Currently, there is limited data available on the transportation of syngas. Direct use of syngas is the dominant business case causing the lack of existence of commercial-scale syngas transportation networks. The costs of syngas transport are therefore highly uncertain and likely more expensive than CO<sub>2</sub> transport due to additional safety requirements. For the latter, Knoope (2015) has shown how CO<sub>2</sub> transport costs increase with additional safety measures. In this study, the costs of syngas transport are based on the transport costs of CO<sub>2</sub> but were multiplied with different cost factors to investigate its effect on the centralized/decentralized deployment and the SC configuration.
- The capture costs of CO<sub>2</sub> were averaged at the cluster level based on the mass of CO<sub>2</sub> emissions of individual plants in that cluster and the capture costs per industry. These capture costs [in €<sub>2020</sub>] were based on the capture cost from Global CCS Institute (2021) and CO<sub>2</sub> compression costs from IEA (2020). The CO<sub>2</sub> coming from the petrochemical, oil, and gas industries was assumed to have the same capture and compression cost as from the oil refining industry.

#### 4. Case study results

The problem formulation was formulated in GAMS (41.1) and optimized using the CPLEX 22.1.0.0 MIP solver. The system ran with an Intel®Core™ i7-1185G7 CPU 3GHz processor and 32GB RAM. The problem comprised 228,893 variables, of which 93,526 were binary and 350,303 equations. The model was solved for five different transport cost factors (CF) values (1-3) and took 11 minutes.

The SC configurations with varying cost factors for syngas transport are presented in Figure 2, while the different SC characteristics are presented in Table 1. For all the CFs, the same 25 syngas demand locations were selected. Plants were co-located when the CO<sub>2</sub> source, the CO<sub>2</sub>-based syngas plant, and the syngas consumer were within the same industrial cluster. Co-locating reduces transport dependency; however, in some cases, CO<sub>2</sub>/syngas transport was still needed to fulfill the syngas demand or desired to optimize costs. Remote CO<sub>2</sub> capture and transport can be cheaper to (partly) fulfill the demand compared to on-site capture due to capture cost differences between clusters. When the transport of syngas was more expensive than CO<sub>2</sub> co-locating the CO<sub>2</sub> source with the CO<sub>2</sub>-based syngas plant became more viable. In CF2.5, there remains a need for CO<sub>2</sub> transport to 5 clusters to fulfill syngas demand. The cost factors affect the capacity of the

Table 1 Supply chains characteristics for different syngas transport cost factors

Cost factor	Capacity CO <sub>2</sub> -based syngas plants			CO <sub>2</sub> sources	CO <sub>2</sub> -based syngas plants	Co-located plants
	Min	Max	Avg.	[#]	[#]	[#]
1	0.24	3.5	0.99	21	21	15
1.5	0.24	3.5	0.87	24	24	18
2	0.24	3.5	0.87	25	24	20
2.5	0.50	3.0	0.83	26	25	21
3	0.50	3.0	0.83	26	25	21

individual plants, as the average capacity is higher at a lower cost factor. In other words, lower transport cost increase plant capacity leading to a more centralized SC.

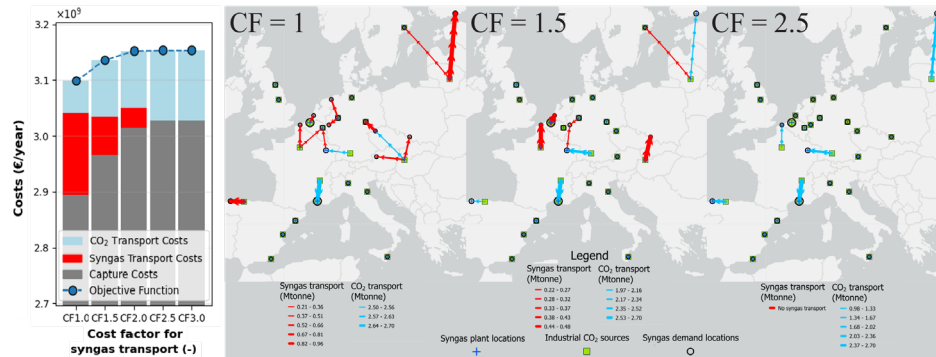


Figure 2 - Parts of the objective function for different syngas transportation costs factors and CO<sub>2</sub>-based syngas supply chains with different syngas cost factors.

The total SC cost in the objective function is hardly affected by an increase in syngas transportation cost, see Figure 2. The CF1 SC can benefit from the cheapest CO<sub>2</sub> capture sources. When syngas transport becomes more expensive, the model chooses to pay more for capturing CO<sub>2</sub> at less favorable locations to avoid transportation. At a CF of 2.5, syngas transport is eliminated.

### 5. Conclusion and future work

A simplified CO<sub>2</sub>-based syngas supply chain facility location model was developed in this work. The objective minimized the CO<sub>2</sub> capture, transportation, and syngas transport costs. In this case, when syngas transportation has the same price as CO<sub>2</sub> transport, there are options for centralization in the CO<sub>2</sub>-based syngas SC. At higher syngas transportation costs, decentral SCs are preferred. When syngas transport is as expensive as CO<sub>2</sub> transport, syngas transport is favored due to its lower transport volume and potential to capitalize on clusters with lower capture costs. CO<sub>2</sub> transport is occasionally necessary for clusters with insufficient CO<sub>2</sub> to fulfill their syngas demand after CO<sub>2</sub>ER conversion.

In the next iteration of the model, a sensitivity analysis will be performed by exploring different syngas demand locations and sizes, which will help generalize the findings of this specific case. The current model is a starting point for developing a more exhaustive CO<sub>2</sub>-to-syngas-based SC model, focusing on enhancing transportation details and refining the objective function. With this exploratory case, the first insights were gained regarding the scale of CO<sub>2</sub>-based syngas plants and the preference for a centralized/decentralized configuration of the SC.



## Acknowledgments

This research receives funding from the project "Addressing the multiscale challenge of CO<sub>2</sub> electrochemical reduction", NWO ECCM tenure track grant (project number ECCM.TT.009), and "Sustainable design of multiscale CO<sub>2</sub> electrochemical conversion", MVI grant (ECCM.TT.MVITU.006).

## 6. References

- Ahmadvand, S., Sowlati, T., 2022, A robust optimization model for tactical planning of the forest-based biomass supply chain for syngas production, *Computers and Chemical Engineering*, 159, 107693.
- Boston Consulting Group, 2021, Hubs and Industry layer, Retrieved November 18, 2022, from <https://edu.nl/g3twa>
- Choe, C., Cheon, S., Gu, J., Lim, H., 2022, Critical aspect of renewable syngas production for power-to-fuel via solid oxide electrolysis: Integrative assessment for potential renewable energy source, *Renewable and Sustainable Energy Reviews*, 161, April, 112398.
- d'Amore, F., Romano, M. C., Bezzo, F., 2021, Optimal design of European supply chains for carbon capture and storage from industrial emission sources including pipe and ship transport, *International Journal of Greenhouse Gas Control*, 109.
- Ebbehøj, S. L., 2015, Integration of CO<sub>2</sub> air capture and solid oxide electrolysis for methane production, Department of Energy Conversion and Storage, Technical University of Denmark.
- European Commission, 2019, The European Green Deal, European Commission, 53, 9.
- European Environment Agency, 2020, The European Pollutant Release and Transfer Register (E-PRTR).
- Global CCS Institute, 2021, Technology Readiness and Costs of CCS. Global CCS Institute, March, 50.
- Göttinger, M. J. W., 2013, Facility Location Planning for Distribution Networks and Infrastructure Locations, University of Freiburg, 160.
- Hakimi, S. L., 1964, Optimum Locations of Switching Centers and the Absolute Centers and Medians of a Graph, *Operations Research*, 12, 3, 450–459.
- IEA, 2020, Energy Technology Perspectives 2020 - Special Report on Carbon Capture Utilisation and Storage.
- IMARC, 2021, Syngas Market: Global Industry Trends, Share, Size, Growth, Opportunity and Forecast 2022-2027, Retrieved November 7, 2022, from <https://www.imarcgroup.com/syngas-market>
- Inkwood Research, 2017, Syngas Market - Global Trends, Size Share, Analysis Report 2017-2026, Retrieved November 9, 2022, from <https://edu.nl/77qea>
- Jouny, M., Luc, W., Jiao, F., 2018, General Techno-Economic Analysis of CO<sub>2</sub> Electrolysis Systems, *Industrial and Engineering Chemistry Research*, 57, 6, 2165–2177.
- Knoope, M. M. J., 2015, Costs, safety and uncertainties of CO<sub>2</sub> infrastructure development, 359.
- Marufuzzaman, M., Li, X., Yu, F., Zhou, F., 2016, Supply Chain Design and Management for Syngas Production, *ACS Sustainable Chemistry and Engineering*, 4, 3, 890–900.

# Optimization of large-scale energy systems to achieve carbon emissions neutrality

Sanja Potrč,<sup>a</sup> Andreja Nemet,<sup>a</sup> Lidija Čuček,<sup>a</sup> Petar Varbanov,<sup>b</sup> Zdravko Kravanja,<sup>a</sup>

<sup>a</sup>*Faculty of Chemistry and Chemical Engineering, University of Maribor, Smetanova ulica 17, 2000 Maribor, Slovenia*

<sup>b</sup>*Sustainable Process Integration Laboratory – SPIL, NETME Centre, Faculty of Mechanical Engineering, Brno University of Technology – VUT Brno, Technicka 2896/2, 616 69 Brno, Czech Republic*

## Abstract

This contribution presents a synthesis of sustainable renewable energy supply networks to achieve a carbon emission neutral energy system in the EU by 2050, considering the generation and supply of heat and electricity from renewable and nuclear sources, second- and third-generation liquid biofuels, hydrogen and by-products from different sources using different conversion and storage technologies. A multi-period, mixed-integer linear programming (MILP) model is developed that aims to maximize the Sustainability Net Present Value of the EU's future energy system, emphasizing the role of energy storage technologies in achieving carbon emissions neutrality while taking into account the Earth's great potential for self-regeneration. The results show that in 2050, 502 TWh/y of electricity demand could be covered by electricity stored in batteries, followed by 213 TWh/y covered by pumped hydro storage, while applied thermal energy storage technologies are expected to meet 389 TWh/y of heat demand. The integration of the studied electricity and thermal storage technologies will also make an important contribution to achieving a carbon-neutral energy system in the EU with a 7.6 % reduction in CO<sub>2</sub> emissions in 2050 for the identified energy mix compared to the system without energy storages.

**Keywords:** supply network optimization, energy storage, renewable energy, energy transition, carbon emissions neutrality.

## 1. Introduction

In order to accomplish the EU's climate and energy goals, a rapid transformation of the energy system and further developments toward more efficient energy generation, supply and use are required. The integration of energy storage technologies plays an important role in the transition towards carbon emissions neutrality and provides a concrete means to improve the energy efficiency of the future energy system based on renewable energy sources, while increasing energy security in the EU (European Commission, 2020). Several studies have examined the impact of integrating storage technologies in the transition to carbon emissions neutrality at different spatial and temporal scales. Child et al. (2018) examined the role of storage technologies in the gradual transition to a 100 % renewable power sector in Europe by 2050 and concluded that up to 15 % of electricity demand could be covered by battery storage, highlighting its important supporting role. In addition, the deployment of flexible renewable energy generation and energy storage could significantly reduce the levelized cost of electricity, especially when optimization

is performed at larger geographic scale (Child et al. 2019). The results of the study by Dowling et al. (2020) suggest that the introduction of long duration energy storage, such as pumped hydro storage and compressed air energy storage, could significantly reduce the cost of a future electricity system that relies on intermittent energy sources, and to a much greater extent that could be achieved by equivalent reductions in battery costs. Most of the transition pathways towards the renewable energy system have been simulated using the LUT energy system transition model or the EnergyPLAN hourly simulation model.

This study presents the optimization of renewable energy supply networks for the gradual transition to a renewable energy system in the EU to achieve carbon emission neutrality by 2050. A multi-period mixed-integer linear programming (MILP) model was developed that simultaneously considers all three aspects of sustainability, economic, environmental and social. Electricity, heat, second and third generation biofuels, and hydrogen generation and supply are taken into account to satisfy the demand in end-use sectors. Our previous work (Potrč et al., 2022) was extended to include energy storage technologies to show their potential in the EU's future energy system with a high share of renewables and their impact on achieving carbon emissions neutrality, while also considering the Earth's large potential for natural carbon sequestration.

## 2. Methodology for the synthesis of a large-scale renewable energy supply networks

The synthesis of large-scale supply networks involves a large number of highly complex interactions and by addressing these interactions simultaneously with effective solution procedures, many of the synergistic effects of the subsystems can be exploited. The simultaneous solving of these subsystems, which also guarantees the optimality of the results for a given set of conditions, is only possible with a mathematical programming approach. It is also important that the optimization criteria are sustainable, covering economic, environmental and social aspects of sustainability. For this purpose, a maximization of the composite criterion Sustainability Net Present Value (Zore et al., 2018) was used as an objective, Eq. (1).

$$\max NPV^{\text{Sustainability}} = NPV^{\text{Economic}} + NPV^{\text{Eco}} + NPV^{\text{Social}} \quad (1)$$

$NPV^{\text{Sustainability}}$  is defined as the sum of economic, environmental and social  $NPVs$ , with economic  $NPV$  calculated as the sum of cash flows over the years, where each annual cash flow is discounted to the present. For the environmental  $NPV$ , burdening and unburdening effects of energy sources, pretreatment, conversion technologies, products, transportation and land-use were considered, while for the social  $NPV$ , the creation of additional jobs was the main contributor.

A gradual transition to a renewable energy system to achieve carbon emissions neutrality by 2050 was conducted with 10-year time steps, taking into account dynamic changes through the supply network using monthly, daily, and hourly time periods. Electricity and heat generation, demand and storage, as well as availability of variable sources, were modelled on an hourly basis, while monthly time periods were used for biofuels. Future price trends and potential improvements in the efficiency of existing and emerging technologies, as well as the EU projection for the nuclear energy use, were also considered.

The mathematical model accounts for several model size reduction techniques to keep the model solvable: (i) patterns in intermittent sources availability and electricity and heat generation, demand, and storage were applied to reduce computation time, where time periods were reduced to two periods per week and four periods per day; (ii) connectivity in a renewable energy supply network was reduced by specifying the maximum allowable distances between supply, pretreatment, storage, production, and demand sites; (iii) unnecessary variables and constraints with zero-flows were eliminated (Lam et al., 2011); (iv) the number of zones was reduced from 132 zones in the original formulation (Potrč et al., 2020) to 47 zones, emphasizing the regional characteristics in each country using geographically specific data. In addition, the nonlinear investment terms were linearized using a piecewise linear approximation, transforming the MINLP model to a MILP model. The model takes into account the generation and supply of electricity and heat, liquid biofuels (second and third generation), hydrogen, and other byproducts from various renewable and clean energy sources and biowaste, using different conversion and storage technologies. All integration options are shown in Figure 1.

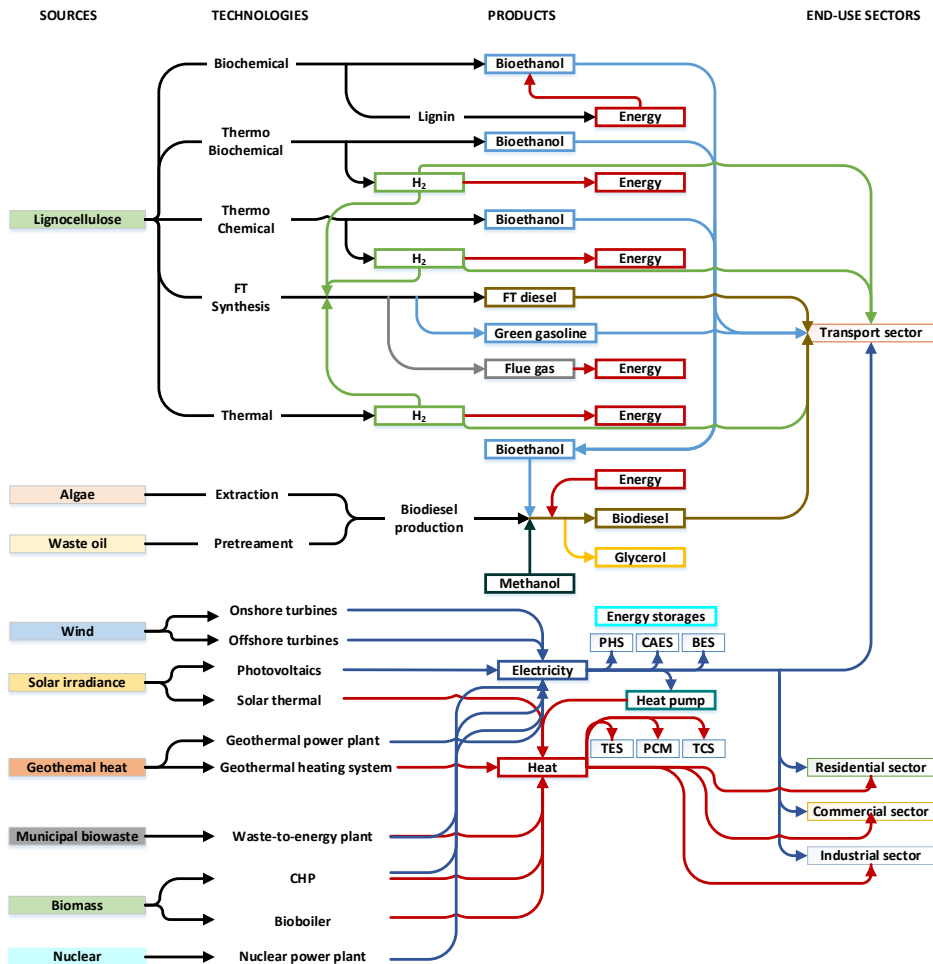


Figure 1. Integration possibilities in renewable energy supply network (modified after Potrč et al., 2022).

In addition, the possibility of converting thermal power plants currently in operation into biomass cogeneration plants is also being considered. The storage technologies considered are pumped hydro storage (PHS), utility-scale battery storage systems (BES), compressed air energy storage (CAES), thermal energy storage (TES), storage in phase change materials (PCM), and thermo-chemical energy storage (TCS). The TES technologies investigated in this study are hot water tanks for domestic hot water storage, large water tanks (seasonal), underground thermal energy storage, and molten salt storage. Data on the performance and costs of the storage technologies were taken from (IEA-ETSAP and IRENA, 2013; IRENA, 2019).

The model includes production and conversion constraints, constraints on plant capacities, pretreatment, investment, operating, storage and transportation costs, economic, environmental and social constraints, and also the area that could be used for energy generation in each zone is limited to preserve biodiversity. To predict the future concentration of CO<sub>2</sub> in the atmosphere, the Earth's CO<sub>2</sub> self-sequestration capability was considered, and the prediction model was developed based on historical data (Potrč *et al.*, 2022). The overall synthesis model was formulated in the modeling system GAMS, and the solutions were obtained in about a week using the solver CPLEX on the HPC server DL580 G9 CTO (4 processors, 32-core, Intel® Xeon® CPU E5-4627 v2 @ 3.30 GHz, 768 GB RAM).

### 3. Results

The following gradual transition to highly renewable energy system in the EU by 2050 was studied to achieve carbon emission neutrality: (i) at least 50 % share of RES in electricity and heat supply in all end-use sectors, and 30 % share of RES in road transportation by 2030, (ii) at least 75 % share of electricity and heat from RES and at least 60 % share of RES in road transport by 2040, and (iii) in 2050, electricity supply in all sectors should be met by renewables or nuclear energy and the highest possible share of RES in heat supply and road transport should be achieved to reach the goal of a carbon neutral EU.

Figure 2 shows the proposed electricity and heat generation, and energy storage output of the various technologies over the years. The optimization results indicate that in the first years of the energy transition (until 2030), most of the electricity from RES will be generated by wind farms, 798 TWh/y, followed by 497 TWh/y from biomass cogeneration plants, 352 TWh/y from hydropower plants and 301 TWh/y from solar PV. The possibility of converting thermal power plants currently in operation into biomass cogeneration plant is also proposed in most of the Member States, as the investment costs of reconstruction are lower compared to the construction of new CHP plants. After 2030, faster spreading of electricity generation from wind turbines and solar PV is proposed, while the construction of additional CHP plants is not proposed after 2030. Electricity demand in the EU is expected to increase to about 5,400 TWh/y by 2050. 13.4 % of electricity generation would be covered by nuclear power plants, 38.2 % by wind turbines (onshore and offshore), 26.7 % by solar PV, while other technologies considered will contribute another 21.7 %. In terms of heat generation and supply, air-source heat pumps are expected to meet 37.5 % of heat demand, biomass cogeneration plants (newly constructed and converted from thermal power plants) 26.2 %, and solar thermal 16.1%, while smaller shares are expected from geothermal energy, biomass boilers, and WtE using biowaste fraction of municipal solid waste as a source.

As can be seen in Figure 2, the energy output from storage technologies correlates with energy production, especially that from intermittent sources. Pumped hydro storage,

which currently accounts for 97 % of all electricity storage in the EU, is expected to meet 160 TWh/y of electricity demand by 2030. Utility-scale battery storage system is expected to cover 92 TWh in 2030, while by 2040 BES output is proposed to exceed PHS with 218 TWh/y of energy output to provide balances in shorter time periods due to the higher share of electricity generated from intermittent sources. CAES will enable an additional 115 TWh/y of electricity storage. BES is preferred over CAES due to its higher round-trip efficiency. In 2050, 1,254 TWh/y of electricity and heat demand is proposed to be met by applied storage technologies, of which 502 TWh/y corresponds to batteries, 324 TWh/y to sensible thermal energy storage, and 213 TWh/y to pumped hydro storage, while CAES, storage in PCM, and thermochemical energy storage are expected to meet 150 TWh/y, 29 TWh/y, and 36 TWh/y, respectively. Applied electricity storage technologies are expected to cover 865 TWh/y or 16 % of electricity demand across all end-use sectors and applied thermal energy storage technologies 389 TWh/y or 9 % of heat demand by 2050.

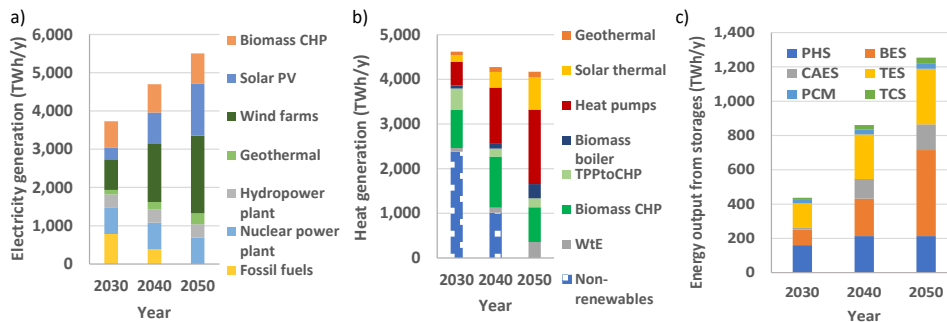


Figure 2. Production of a) electricity and b) heat, and c) energy output from storages in EU by 2050.

In the road transport sector, electricity is proposed to cover more than 55 % of the energy demand, while biodiesel and FT-diesel will account for 28 %, bioethanol and ‘green’ gasoline for 11 %, and hydrogen for 6 %. The Sustainability NPV of the proposed transition to a carbon-neutral energy system amounts to 5,865,975 M€, with the environmental NPV being the largest contributor.

It is interesting to note that the integration of storage technologies will also significantly reduce CO<sub>2</sub> emissions from generation and supply of energy. In our previous work (Potrč et al., 2022) the Earth’s great potential for self-regeneration has been assessed, indicating that more than a half of the annual anthropogenic CO<sub>2</sub> emissions are removed through natural sequestration. The results show that anthropogenic CO<sub>2</sub> emissions would have to decrease by 69 % from 1990 levels to achieve emissions neutrality, with about 1.2 Gt/y of CO<sub>2</sub> released into the atmosphere and the same amount removed from the atmosphere by the Earth. Electricity storage is expected to account for 4.4 % and thermal energy storage for 3.2 % of the reduction in CO<sub>2</sub> emissions in 2050 at the energy mix identified for that year. Compared to the current energy mix, the reduction in CO<sub>2</sub> emissions will be even higher, accounting for 15.3 % of the emissions reduction.

## Conclusions

A high share of renewables, integration of energy storage technologies, and cross-sectoral energy integration will play a very important role in achieving carbon emissions neutrality by 2050. It is also important that the Earth’s large CO<sub>2</sub> self-regeneration potential is considered when planning a transition pathway to achieve carbon emissions neutrality, as more than half of annual anthropogenic CO<sub>2</sub> emissions are removed from the atmosphere

through natural carbon sequestration. The synthesis of such large systems involves a large number of highly complex interactions. For this purpose, various model linearization, convexification and model size reduction techniques have been applied, e.g., i) piecewise linear approximation of nonlinear nonconvex terms, reducing the MINLP model to a MILP model, ii) use of surrogate models for different production plants after the detailed synthesis of their individual flowsheets, iii) reduction of the number of zones, etc. The optimization of renewable and clean energy supply networks for the gradual transition to a highly renewable energy system in the EU to achieve carbon emissions neutrality was based on a multi-period MILP model that considers the production and supply of heat and electricity from renewable and nuclear sources, second- and third-generation liquid biofuels, hydrogen, and by-products from various sources using different conversion and storage technologies to meet the demand in all end-use sectors. The results show that pumped hydro storage is expected to dominate in the first year of the transition, while grid-scale battery storage systems are expected to increase after 2030 due to the larger share of intermittent sources. By 2050, 1,254 TWh/y of energy demand is expected to be met by applied storage technologies, with batteries accounting for 40 %, thermal energy storage technologies for 26 %, and pumped hydro storage for 17%, while the shares of compressed air energy storage, thermochemical energy storage and storage in phase change materials are expected to reach 12 %, 3 % and 2 %, respectively. The integration of the studied electricity and thermal storage technologies will also contribute significantly for achieving carbon emission neutrality with a 7.6 % reduction in CO<sub>2</sub> emissions in 2050 for the identified energy mix.

## References

- M. Child, D. Bogdanov, C. Breyer, 2018, The role of storage technologies for the transition to a 100% renewable energy system in Europe, *Energy Procedia*, 155, 44-60.
- M. Child, C. Kemfert, D. Bogdanov, C. Breyer, 2019, Flexible electricity generation, grid exchange and storage for the transition to a 100% renewable energy system in Europe, 139, 80-101.
- J.A. Dowling, K.Z. Rinaldi, T.H. Ruggles, S.J. Davis, M. Yuan, F. Tong, N.S. Lewis, K. Caldeira, 2020, Role of Long-Duration Energy Storage in Variable Renewable Electricity Systems, *Joule*, 4, 9, 1907-1928.
- European Commission, 2020, Study on energy storage - Contribution to the security of the electricity supply in Europe, Publications Office, <https://data.europa.eu/doi/10.2833/077257>, accessed on 29th November 2022.
- IEA-ETSAP and IRENA, 2013, Technology Briefs, <https://www.irena.org/publications/2013/jan/irena-iea-etsap-technology-briefs>, accessed on 11th November 2022.
- IRENA, 2019, Utility-Scale Batteries Innovation Landscape Brief, [https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2019/Sep/IRENA\\_Utility-scale-batteries\\_2019.pdf](https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2019/Sep/IRENA_Utility-scale-batteries_2019.pdf), accessed on 12th November 2022.
- H. L. Lam, J.J. Klemeš, Z. Kravanja, 2011, Model-size reduction techniques for large-scale biomass production and supply networks, *Energy*, 36, 8, 4599-4608.
- S. Potrč, L. Čuček, M. Martin, Z. Kravanja, 2020, Synthesis of European Union Biorefinery Supply Networks Considering Sustainability Objectives, *Processes*, 8, 12 1588.
- S. Potrč, A. Nemet, L. Čuček, P. Varbanov, Z. Kravanja, 2022, Synthesis of a regenerative energy system – beyond carbon emissions neutrality, *Renewable and sustainable energy reviews*, 169, 112924.
- Ž. Zore, L. Čuček, D. Širovnik, Z. Novak Pintarič, Z. Kravanja, 2018, Maximizing the sustainability net present value of renewable energy supply networks, *Chemical Engineering Research and Design*, 131, 245-265.

# Multi-objective sustainable supply chain design under uncertainty in energy price

F.L. Garcia-Castro,<sup>a</sup> R. Ruiz-Femenia,<sup>a</sup> R. Salcedo-Diaz,<sup>a</sup> J. A. Caballero<sup>a</sup>

<sup>a</sup>*Institute of Chemical Process Engineering, University of Alicante, Aparatado de Correos 99, Alicante 03080, Spain*

## Abstract

In our society, sustainability and global responsibility concepts have been gaining in importance over the years. For the manufacturing companies and supply chains there is a direct impact on their benefits, not only through an increasing demand of sustainable products, but also through the regulations that aim to reduce the total emissions over the years. Moreover, due to the current political situation, the prices of energy could jeopardize many companies, whose production is no longer profitable.

For all these reasons, it is more important than ever to apply a holistic view in the design of supply chains and ensure a model that considers risk management.

Our case-study is a petrochemical three echelon supply chain in Europe, where we incorporate the cap-and-trade model from the European Union Emissions Trading System, to directly link the emissions associated to consumption of raw material, energy consumption and transportation to the total benefit of the supply chain.

We present an optimal design of a supply chain facing uncertainty in energy prices, which is modelled by fitting historical data in an ARIMA model. It involves the solution of a multi-objective multi-period mixed integer linear program that was implemented in Python using Pyomo and solved using IBM's CPLEX solver.

**Keywords:** Energy price, carbon price, multi-objective optimization, risk management.

## 1. Introduction

Regarding the design of a supply chain, a vast amount of information and metrics need to be considered during the process of decision-making. One important aspect to consider are uncertainties of the underlying parameters, such as demand, carbon prices, energy prices, etc., leading to stochastic numerical models. Solving a stochastic model yields a supply chain design that is more robust against changes in the uncertain parameters, and will have a superior expected performance over designs resulting from deterministic models. The consideration of uncertainty has been widely addressed in the literature mainly for cost and demand, c.f. Yang et al. 2015, Robles et al. 2020. In this work, we consider uncertainty of energy prices, a topic that has recently become very important in the euro zone, c.f. the article on energy prices and security of supply (2022). In order to generate possible pricing scenarios, we identify the best-suited parameters of an ARIMA model and fit it to the historical price data. We then use this model to forecast energy prices for the next 10 years.

Another important concept in the design of supply chains is evolving carbon regulations, which governments are implementing to control the available amount of carbon emissions for the industry. These regulations have been changing fast over the last years. In our



model, we considered a cap-and-trade system based on the European Union Emissions Trading System. Each company is assigned a fraction of a total quantity of emissions that is reduced over time. The companies may then produce their goods and / or trade with their assigned emissions.

When it comes to making design decisions, it is not only important to assure robustness against possible future parameter changes, but also to consider the impact of decisions and the risk associated to them. Hence, instead of only maximizing the profit, we considered a multi-objective model that also aims to minimize the downside risk (DR). The conjoint study of uncertainty in energy prices, risk management and carbon regulations will provide important information when taking design-decisions for supply chains. To the best of our knowledge, a similar study has never been performed.

## 2. Methodology

### 2.1. Problem statement

In this manuscript, we consider a three-echelon supply chain superstructure that consists of 6 plants, located in Hungary (Kazincbarcika), Germany (Leuna), Italy (Montova), Czech Republic (Neratovice), Spain (Tarragona) and Poland (Wloclaweck). All plants can produce acetone, acetaldehyde, cumene, acrylonitrile, isopropanol and phenol using one of six available technologies involving up to 18 different chemicals. For each plant, there is an associated warehouse having an initial inventory of zero tons. All plants and warehouses have a yearly expansion limit between 10 to 400 kt. The products can be sold on four different European markets, being Germany (Leuna), Czech Republic (Neratovice), Portugal (Sines) and Spain (Tarragona). For more details on the superstructure see Guillen-Gosalbez et al. (2009).

To obtain a more realistic model, no negligible quantities may be transported between plants and warehouses, respectively between warehouse and markets. A penalty term is introduced into the calculation of the net present value (NPV), that penalizes any insatisfaction of demand by 30% of the product price on each market. The initial cap for carbon dioxide emissions is set to  $2 \cdot 10^8$  kg with a yearly reduction rate of 2.2%, see the EU ETS handbook for details (2015).

### 2.2. Supply chain model

The model proposed has the objectives of maximizing the profit and minimizing the downside risk, while satisfying equations related to mass balance, capacity constraints, environmental impact assessment and carbon regulations. The last one directly links the environmental to the economic performance. In this section, we will refer the reader to the articles of Garcia-Castro (2022) and Guillen-Gosalbez (2009) for details regarding the equations, variables and parameters of the model. Here, we will only focus on the equations that introduce uncertainty into the model.

The cost of transportation between a plant  $p$  and a warehouse  $w$  at timestep  $t$  in pricing scenario  $s$ ,  $COSTTRA(p, w, t, s)$ , depends on the quantity  $TRA(p, w, t, s)$  transported, the distance  $DISTA(p, w)$ , and the price of diesel / energy,  $PRICE_{ENERGY}(t, s)$ ,

$$COSTTRA(p, w, t, s) = TRA(p, w, t, s) \cdot DISTA(p, w) \cdot PRICE_{ENERGY}(t, s) \cdot 0.014.$$

Similarly, the cost of transportation between a warehouse  $w$  and a market  $k$  at timestep  $t$  in pricing scenario  $s$ ,  $COSTTRB(w, k, t, s)$ , depends on the transported quantity  $TRB(w, k, t, s)$ , the distance  $DISTB(w, k)$ , and the price of diesel / energy,

$$COSTTRB(w, k, t, s) = TRB(w, k, t, s) \cdot DISTB(w, k) \cdot PRICE_{ENERGY}(t, s) \cdot 0.014.$$

According to the Spanish Ministry of Public Works and Transportation, an articulated vehicle with a payload of 25t that travels around 120000 km annually, will have a fuel expenditure of more than 40700 liters of diesel, c.f. the report by the Spanish Ministry of Public Works and Transport (2019). This is equivalent to approx. 0.014 liters of diesel required to transport one ton of goods over a distance of one kilometer. The distances *DISTA* and *DISTB* are given in kilometers. Since the most common type of transportation of goods in Europe is via truck, we assumed for simplicity that all transportation is done in this way.

### 2.3. Environmental impact assessment and carbon regulation

In the model, we account for emissions related to raw material consumption, energy consumption, as well as emissions due to transportation between plants, warehouses and markets. We use Life Cycle Impact Assessment (LCIA) data obtained from the ECOINVENT Database to calculate the emissions in terms of global warming potential for a timespan of 100 years (GWP100a), c.f. Frischknecht et al., 2005 for details. The cap-and-trade system incorporated by the European Union Emissions Trading System aims to help the EU become climate-neutral through the reduction of the maximum amount of emissions allowed over time. An initial emissions cap is introduced, together with a yearly reduction rate. The plants are assigned a fraction of the total emissions and they are allowed to also trade their emissions with other players.

### 2.4. Uncertainty in energy prices

In a supply chain, there are many parameters with possible uncertainties. Driven by the current political and social situation in Europe, we will in this work focus on uncertain future changes of the energy prices. To account for uncertainty in our model, various future pricing scenarios will be generated. This is done by fitting a suited autoregressive integrated moving average (ARIMA) model to the historical diesel prices of the last 25 years. ARIMA models are defined by three parameters, being the non-seasonal polynomial degree, the number of moving average terms, as well as the order of differentiation applied to the data in order to remove seasonal trends. To find the parameters that best fit the historical data, we employed the Akaike and Bayesian information criteria. It turns out, that the ARIMA(2,1,3) model is best suited to forecast future pricing scenarios. For details on time-series forecasting, see Brockwell et al. (2016).

### 2.5. Risk management

Uncertainty of parameters and increasing complexity of supply chains require risk analysis to take well-informed decisions. Different approaches for quantitative supply chain risk management are known and should be employed to predict performance. Some examples of risk metrics are the downside risk (DR), the value at risk (VaR), or the conditional value at risk (CVaR). These metrics are not redundant and all provide valuable information for the decision-makers. For details on a variety of risk metrics and their usage, see Heckmann 2014.

In this work, we will optimize the economic performance of the supply chain together with the expected downside risk, that is the expected unfavorable deviation of the profit from a target value  $\Omega$ ,

$$DR(s, \Omega) = \mathbf{E}[\delta_{\Omega}(s)] = \int_{-\infty}^{\infty} \delta_{\Omega}(s) P(s, \xi) d\xi.$$

Here,  $P$  is the probability distribution function, and

$$\delta_{\Omega}(s) = \begin{cases} \Omega - \text{Profit}(s), & \text{if } \text{Profit}(s) < \Omega \\ 0, & \text{if } \text{Profit}(s) \geq \Omega. \end{cases}$$

### 2.6. Multi-objective simulation

To solve the stochastic model, we generate 100 possible energy pricing scenarios and reduce them using the BIRCH (balanced iterative reducing and clustering using hierarchies) data mining algorithm. We use the  $S\_dbw$  index, to measure the quality of the clustering for different numbers of scenarios. For an application of the clustering, each scenario is identified as a two-dimensional point consisting of its net present value and downside risk.

In the real world, the decision-makers have more than one objective when it comes to supply chain design, such as maximization of profit, minimization of risk, emissions, cost, etc. Therefore, we implement a multi-objective approach using the  $\varepsilon$ -constraint method. This approach requires more computational effort, but it provides a set of possible solutions, from which the decision-makers obtain a better overview of the dynamics of the supply chain.

First, we optimize the stochastic problem including all of the reduced scenarios by maximizing the profit. The corresponding downside risk is denoted  $DR_{maxProfit}$ . Then, we re-solve the model by minimizing the downside risk, obtaining  $DR_{minDR}$ . The obtained solutions represent the two extreme points of the pareto set of solutions. Afterwards, we add a new constraint to the model that sets a limit  $\varepsilon$  for the expected downside risk and re-solve the model maximizing the profit. This bound is modified for each pareto point  $i_{pp}$  considered,

$$\varepsilon = DR_{minDR} + \frac{DR_{maxProfit} - DR_{minDR}}{\text{number pareto points}} \cdot i_{pp}.$$

## 3. Results and discussion

The stochastic model was implemented and solved in Python using Pyomo and IBMs CPLEX algorithm. The simulations were conducted on a MacBook Pro model 2020 with 16 GB Ram and running macOS 11.2.1.

In figure 1, the left plot shows the original 100 scenarios that were grouped into 14 clusters and the “x” indicates the centroid of each cluster. In the right subplot from figure 1, the  $S\_dbw$  index is represented for a changing number of clusters. It can be seen that the error measured in the  $S\_dbw$  index is less than 10% for numbers of clusters larger than 13. The value of this indicator drops sharply for small numbers of clusters, before showing asymptotic behavior. The  $S\_dbw$  index measures the mean dispersion of datapoints inside the clusters, as well as a between-cluster density, to compute a scoring for a given clustering (c.f. Desgraupes, 2017). Its values are positive and will be relatively small for a clustering that splits a highly connected point cloud approximately evenly. The foregoing explanation motivates our choice of 14 reduced clusters whose centers will be used as reduced set of scenarios.

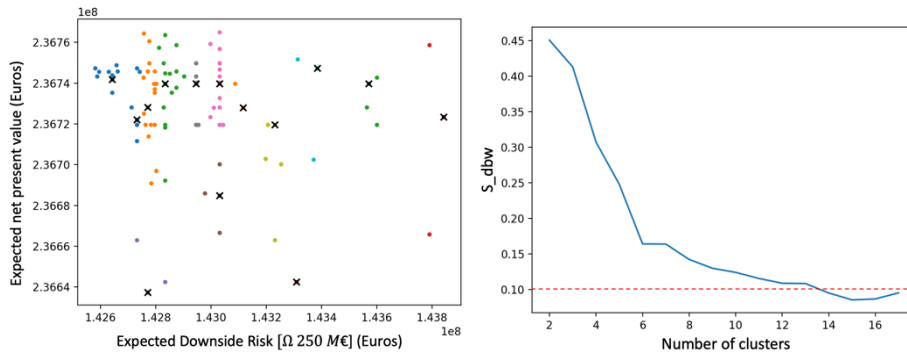


Figure 1: (Left) Economic performance and downside risk of the 100 deterministic scenarios. Each scenario is represented by a two-dimensional point, and all the points that form a cluster have the same color. The centroid of each cluster is indicated with “x”. (Right) S\_dbw index for different number of clusters.

In order to obtain a pareto curve, several steps have to be performed. First, the problems that provide the upper and lower limit for the epsilon are solved. Afterwards, the value of epsilon is varied in this interval to obtain the different pareto solutions. Figure 2 shows the cumulative distribution function for the solution maximizing the expected net present value. The obtained supply chain design was exposed to 100 possible energy pricing scenarios and the values of different risk metrics were calculated. The numerical results show, that the NPV of the worst-case scenario deviates from the expected NPV by less than 5%.

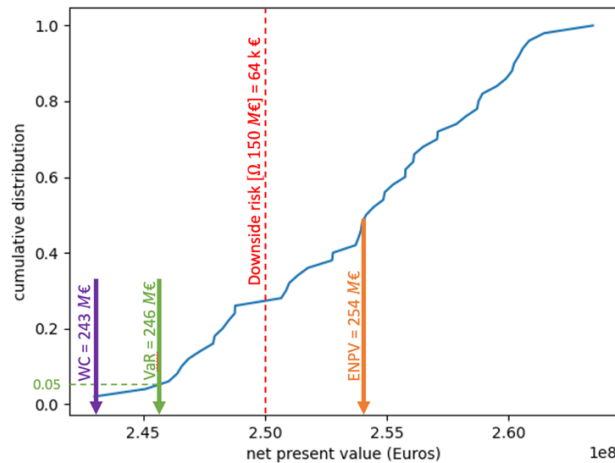


Figure 2: Cumulative distribution function for the obtained supply chain design. The expected net present value, downside risk for a target value of 250 million euros, VaR and worst-case scenario are represented.

#### 4. Outlook and further work

So far, we only considered the influence of energy prices in the cost of transportation. However, since the model under study is a petrochemical supply chain, the cost of energy also influences the cost of raw materials and the cost of production.

Once we incorporated these influences into the model, we will then move on to computing the pareto curve of possible solutions using the  $\varepsilon$ -constraint method. The results will give us a better understanding of the influence of energy prices in the overall SC design.

As stated before, different risk indices provide valuable information and so far, we only implemented the downside risk. In upcoming work, we will introduce other of these risk indices as objectives into the model.

#### Acknowledgements

The authors gratefully acknowledge financial support to the Conselleria de Innovacion, Universidades, Ciencia y Sociedad Digital of the Generalitat Valenciana, Spain under project PROMETEO/2020/064 and project PID2021-124139NB-C21.

#### References

- Spanish Ministry of Public Works and Transport, October 2019, Observatorio de Costes del Transporte de Mercancías por Carretera
- Frischknecht et al., 2005, The ecoinvent Database: Overview and Methodological Framework, The International Journal of Life Cycle Assessment
- Brockwell, P., Davis, R., 2016, Introduction to time series and forecasting, Springer Switzerland
- Heckmann, I., Comes, T., Nickel, S., 2014, A Critical Review on Supply Chain Risk – Definition, Measure and Modeling
- European Council, November 2022, Energy prices and security of supply, <https://www.consilium.europa.eu/en/policies/energy-prices-and-security-of-supply/>
- Guillen-Gosalbez, G., Grossmann, I. E., 2009, Optimal Design and Planning of Sustainable Chemical Supply Chains Under Uncertainty, AIChE Journal, 55, pp. 99 – 121
- European Union, 2015, EU ETS handbook, [https://www.sallan.org/pdf-docs/ets\\_handbook\\_en.pdf](https://www.sallan.org/pdf-docs/ets_handbook_en.pdf)
- Desgraupes, B., 2017, Clustering indices, <https://cran.r-project.org/web/packages/clusterCrit/vignettes/clusterCrit.pdf>
- Yang, G., Liu, Y., Yang, K., 2015, Multi-objective biogeography-based optimization for supply chain network design under uncertainty, Computers & Industrial Engineering, 85, pp.145-156
- Robles, J.O., Azzaro-Pantel, C., Aguilar-Lasserre, A., 2020, Optimization of a hydrogen supply chain network design under demand uncertainty by multi-objective genetic algorithms, Computers & Chemical Engineering, 140

# An Ecologically Safe and Socially Just Supply Chain Design for Li-battery

Ying Xue<sup>a</sup>, Yazeed M. Aleissa<sup>a</sup>, Bhavik R. Bakshi<sup>a</sup>

<sup>a</sup> *William G Lowrie Department of Chemical and Biomolecular Engineering, The Ohio State University, Columbus OH 43210, USA*

## Abstract

Global goals like “Net-zero”, “Nature-positive”, and “Socially Just” require human activities to reduce emissions, restore nature, and be socially equitable. This work proposes an approach that includes ecological capacity and social justice requirements to guide engineering decisions and designs. We utilize the supply and demand of ecosystem services to identify the safe and just operating space<sup>1,2</sup>. The ecologically safe space is determined by the multiscale framework of Techno-Ecological Synergy (TES). The degree of overshoot quantifies the absolute environmental sustainability (AES) at the relevant spatial scale<sup>3,4</sup>. For the socially just space, we calculate a minimum threshold of necessary goods and services to meet basic food, energy, and water<sup>5</sup> needs. We demonstrate this approach by a multiobjective supply chain design of Li-battery which minimizes the ecological and social overshoot simultaneously.

**Keywords:** Safe and Just Space, Absolute Environmental Sustainability, Nature-Based Solution, Supply Chain Design, Multiobjective Optimization

## 1. Introduction

In the past several years, the COVID-19 pandemic, wildfires, unprecedented heat waves and other extreme natural events reinforced the fact that economic and social structures are untenable without resilient nature. Meanwhile, global goals such as “Nature-positive by 2030”, “Net-zero by 2050”, have been proposed and widely accepted as guidelines for future development. However, achieving these goals must not exacerbate social issues such as poverty, food insecurity, and equity. Hence, the concept of the “Safe and Just Space”<sup>1,2</sup> was introduced, which identifies an operating space for human activities to be within the ecological capacity of the planet and the social threshold where everyone has access to natural resources to meet their basic needs. This requires industry to develop and apply methods for achieving global goals. Due to the concerns of energy and sustainability, the demand for electric vehicles (EV) is growing fast. Li-battery is widely used in EV and its impact is important to the sustainable EV development. This study focuses on designing an ecologically safe and socially just supply chain for Li-battery.

Current sustainability assessment methods such as life cycle assessment (LCA) are useful for quantifying relative sustainability which encourages the reduction of environmental impacts<sup>4</sup>. Environmental sustainability requires human impact to not exceed nature’s carrying capacity<sup>6</sup> but most methods ignore the capacity of nature. Reducing negative impact alone is not enough to get the full recovery of nature by 2050, as envisioned by the “Nature-positive” goal<sup>7</sup>. The idea of absolute environmental sustainability (AES) includes nature’s carrying capacity in sustainability metrics as a reference value. Xue and Bakshi developed a multiscale approach for absolute environmental sustainability assessment (AESA) which encourages “Nature-positive” decisions<sup>4</sup>. Through biophysical models, this techno-ecological synergy (TES) based AES metric quantifies ecosystem

services at different spatial scales to identify the ecological threshold with high geographical resolution. Furthermore, Aleissa and Bakshi<sup>3</sup> proposed a quantitative social threshold that aligns with the ecological threshold by using the flow of ecosystem goods and services. The social foundation represents the minimum human demand to meet their basic food, energy, and water needs from that ecosystem service.

This work brings in the concept of AES and social threshold into Li-battery supply chain design under the global goals of “Nature-positive” and “Socially Just”. Four main sectors are considered along the supply chain: mining, processing production and packaging. Ecological and social thresholds are quantified at country level for each process using the methods mentioned in the previous paragraph. The ecological and social transgression levels for Li-battery supply chain are minimized simultaneously as design objectives rather than merely constraints. The transgression level is expressed as the ratio of human impact and ecological/social threshold, making it a linear fractional optimization problem. It is first transformed into a linear program through the Charnes–Cooper method and then solved by  $\varepsilon$ -constraint method. We also show how this framework can contribute to identifying hotspots and future improvement opportunities towards global goals. The major novelties of this work are: 1) A multiscale approach of ecological threshold quantification in supply chain design. 2) Quantifying and incorporating aspects of social justice in supply chain design. 3) Application to global safe and just supply chain design for Li-battery.

## 2. Methodology

### 2.1. The Ecological Ceiling

One popular approach for AESA is based on the Planetary boundary (PB) framework. PB framework identifies nine important earth system processes and defines the safe operating space (SOS) for human development<sup>8</sup>. This method downscales the SOS to specific processes/systems based on a sharing principle such as population which provides holistic perspectives but ignores spatial heterogeneity of ecosystems and has high subjectiveness. These shortcomings may be overcome by a multiscale TES approach.

The TES framework is built upon the concept of ecosystem services (ESs). It has been integrated with LCA (TES-LCA) for assessment and process design<sup>9</sup>. Instead of directly downsampling SOSs, the multiscale TES approach quantifies the capacity of ecosystems with biophysical models from different spatial scales, which is illustrated in Figure 1a. Public and private ownerships are considered for ecological threshold partitioning. Private ownership implies that only landowners own the ecosystem services, while for public ownership, ESs belong to every activity inside the region. Figure 1b illustrates a two-scale system (e.g. county and state). Considering a  $n$ -scale TES, the generalized mathematical expression of the system’s ecological threshold can be expressed as:

$$S_i^{tot} = S_{i,1}^{pvt} + \sum_{j=2}^J S_{i,j}^{pub} \prod_{m=1}^{m=j} P_{m+1,m} \quad (1)$$

Here  $i$  represents process  $i$ ,  $j$  represents the  $j$ -th scale,  $m$  is a dummy variable.  $S$  is the ecological threshold,  $P$  denotes the sharing principle. In this study Equation (1) will be used to estimate the ecological safe space for each process in the supply chain. For process  $i$  the absolute environmental sustainability metric – ecological overshoot ( $EO_i$ ) is defined in Equation 2 and  $D_i$  represents environmental impacts.

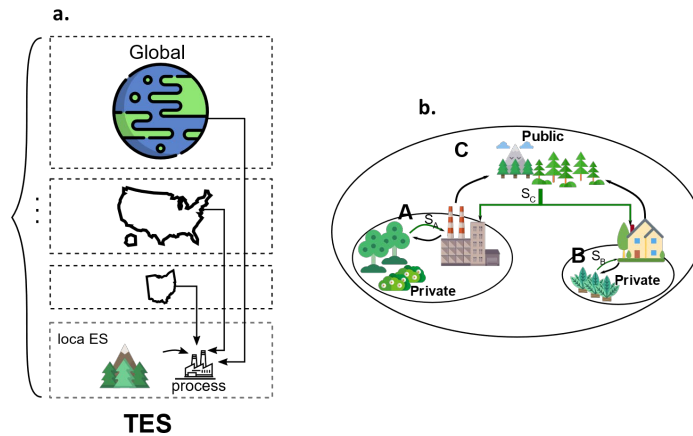


Figure 1. (a) A four level TES example (local, state, country and global). (b) Illustration of public and private ownership of ecosystem supply.

$$EO_i = \frac{D_i}{S_i} \quad (2.)$$

## 2.2. The Social Foundation

Aleissa and Bakshi<sup>3</sup> define the social threshold as the population's minimum demand ( $D^{min}$ ) to meet their basic food, energy, and water needs from ecosystem services. This aligns with the UN sustainable development goals of zero hunger, clean water, and energy for everyone, which specify global standards that define the required amount of water, caloric intake, and electricity to sustain human lives<sup>10</sup>. Then, the social threshold can be formulated as the required emissions from food production and electricity generation to meet the energy and caloric intake thresholds for the entire population. The metric of social shortfall (SS) from the social threshold is expressed as:

$$SS_i = \frac{D_i^{min}}{D_i} \quad (3.)$$

After identifying the safe and just space through the ecological and social thresholds, we can use the current level of demand to assess the operating conditions relative to the thresholds. Figure 2 shows the scenarios that arise with the relative demand levels and the other thresholds.

## 3. Development of Ecologically Safe and Socially just Supply Chains

### 3.1. General Problem Statement

The approaches described in the previous section will be integrated with LCA, multi-objective optimization for designing a safe and just supply chain. Life cycle environmental impact of each process, from cradle to gate, will be estimated. There are two objectives: the ecological objective is to minimize the ecological overshoot and the social objective is to minimize the social shortfall. Both objectives focus on carbon emissions. The major decision variables are selection of suppliers, location of warehouses, and transportation modes. The provided parameters are a set of suppliers and locations for each process in the supply chain, a set of transportation modes, production capacity, demand for the final product, etc. To simplify this problem, this study mainly focuses on selecting suppliers, in another word finding the most suitable country to produce a specific product from "Nature-positive" perspective.



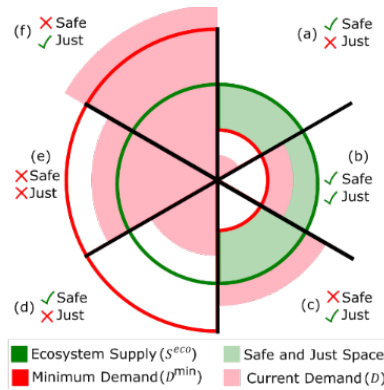


Figure 2. Possible scenarios for the SJS are determined by the current level of demand and the ecological and social thresholds.

### 3.2. Lithium-battery Supply Chain

Environmental impacts of EVs are mainly from the use phase and battery production. There are different supply chains for battery manufacturing based on the final product and the design application. However, many of these supply chains share similar stages: procurement of raw materials, processing these raw materials into valuable chemicals, production of battery components, and fabrication and packing of the battery cell. Each of these stages has its complex supply chains across the world currently. In this problem, we assume it is a linear supply chain, also products and processes has a one-to-one corresponding relation.

### 3.3. Supplier Selection

We want to find the optimal set of suppliers that minimize the ecological overshoot and the social shortfall (Equation 2, 3) along the supply chain from the perspective of CO<sub>2</sub> emission. Using these metrics as objectives results in a mixed-integer linear fractional programming (MILFP) problem that can be optimized to minimize a single objective locally at each stage of the supply chain, or to minimize the metrics in terms of the aggregate supply and demands for the entire supply chain. The set of countries ( $I$ ) and main processes ( $J$ ) are given.  $S_i$  denotes the ES supply (ecological threshold) and  $D_i^{min}$  represents the minimum impact to meet human needs in country  $i$  ( $i \in I$ ).  $x_{i,j}$  and  $c_{i,j}$  are both binary variables.  $x_{i,j} = 1$  if process  $j$  is to be established in country  $i$ ; 0 otherwise.  $c_{i,j} = 1$  when country  $i$  can have process  $j$ ; 0 otherwise.  $D_i$  is the demand of ecosystem services (environmental impact). The general form of total emission can be expressed in Equation 4 where  $g_j$  is life cycle emission factor (EF) for product  $j$ ,  $e_{j,k}$  is EF for product  $j$  in transportation mode  $k$ .  $Q_{j,k,i,i'}$  and  $l_{i,i',k}$  are flow rate and distance from  $i$  to  $i'$ .

$$D_i = \sum_i \sum_{i'} \sum_j \sum_k g_j Q_{j,k,i,i'} + e_{j,k} Q_{j,k,i,i'} l_{i,i',k} \quad (4)$$

The model formulation for the entire supply chain in this study is shown as below.

$$\min EO = \frac{\sum_i D_i}{\sum_i S_i} \quad (5)$$

$$\min SS = \frac{\sum_i D_i^{min}}{\sum_i S_i} \quad (6)$$

$$s. t. \quad x_{i,j} \in \{0,1\} \quad \forall i \in I, j \in J \quad (7.)$$

$$\sum_{i \in I} x_{i,j} = 1 \quad \forall j \in J \quad (8.)$$

$$x_{i,j} \leq c_{i,j} \quad \forall i \in I, j \in J \quad (9.)$$

$$c_{i,j} \in \{0,1\} \quad \forall i \in I, j \in J \quad (10.)$$

The constraint in equation (8) ensures that each process included in Li-battery supply chain is assigned to only one country. Equation (9) makes sure that process  $j$  is assigned to country  $i$  only when country  $i$  can produce that product. Figure 3 shows the results for the optimal supply chains that minimize the ecological overshoot and social shortfall, for both local and aggregate approaches.

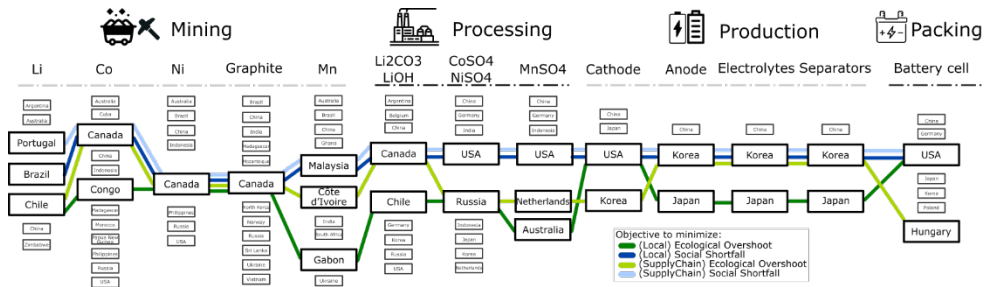


Figure 3. Supply chain results for optimal suppliers' network for the different objectives. For example, minimizing the ecological overshoot at each stage of the supply chain result gives a solution of (Chile, Congo, Canada, Canada, Gabon, Chile, Russia, Australia, South Korea, Japan, Japan, Japan, USA). More details are available in the work of Aleissa<sup>11</sup>.

These results represent the best combinations of suppliers in the current market. However, political relations and economic deals, in addition to logistic and transportation constraints, can limit the selection of supplier networks. In addition, single-objective optimization does not reveal information about the interaction with other objectives. Hence, we performed multi-objective analysis and optimization to find the best set of solutions that define the trade-offs between the objectives. We also compare the optimal global supply chains to the current supply chains of Tesla using their limited supplier network<sup>12</sup>, as shown in Figure 4.

These results can be used to identify hotspots along the supply chains that affect the performance of the social and ecological objectives. For example, the suppliers for lithium for Tesla performed worse than other suppliers in the market in terms of ecological overshoot. Tesla's recent investment in manganese mines in the United States improved their supply chain social performance but greatly downgraded the ecological performance compared to suppliers from Gabon or Ghana.

#### 4. Conclusions

In this work we demonstrate the necessity of including absolute environmental sustainability and social justice into sustainability assessment and design to achieve the 'Nature-positive' and 'Socially-Just' goals. We propose a method for sustainable supply network design which accounts for nature's capacity and social needs. This method can be tailored to activities at different scales and has been applied to a battery supply chain design. LCA coupled with multiobjective optimization guides the selection of suppliers

from ecological and social aspects, and the Pareto curve is obtained showing the trade-off between these two aspects.

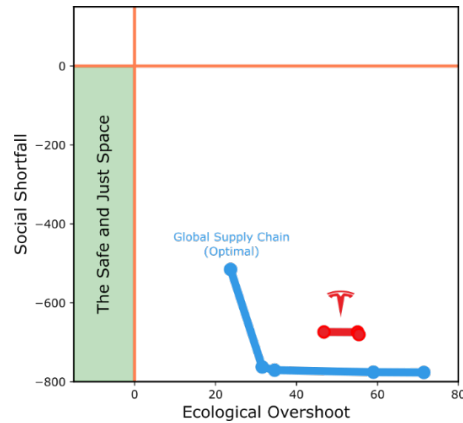


Figure 4. Comparison between Tesla's supply chains and the global optimal supply chains for Lithium batteries.

**Acknowledgements:** King Abdulaziz University, Jeddah, Saudi Arabia for support of Yazeed Aleissa and the National Science Foundation (SBE-1739909; CBET 1804943).

## References

1. J. Rockström, et al., 2009, A safe operating space for humanity, *Nature*, 461.7263, 472-475.
2. K. Raworth, 2012. A safe and just space for humanity: can we live within the doughnut. Oxfam.
3. Y. M. Aleissa, and B. R. Bakshi, 2022, Meeting National Food-Energy-Water Needs in an Environmentally Safe and Socially Just Manner, Submitted.
4. Y. Xue, and B. R. Bakshi, 2022, Metrics for a nature-positive world: A multiscale approach for absolute environmental sustainability assessment. *Science of The Total Environment*, 846, 157373.
5. B. R. Bakshi (editor), *Engineering and Ecosystems: Seeking Synergies for a Nature-Positive World*. in preparation. Springer, 2022.
6. Fiksel, Joseph. "Sustainability and resilience: toward a systems approach." *Sustainability: Science, Practice and Policy* 2.2 (2006): 14-21.
7. Locke, Harvey, et al. "A nature-positive world: the global goal for nature." (2021).
8. Steffen, Will, et al. "Planetary boundaries: Guiding human development on a changing planet." *science* 347.6223 (2015): 1259855.
9. Liu, Xinyu, and Bhavik R. Bakshi. "Ecosystem services in life cycle assessment while encouraging techno - ecological synergies." *Journal of Industrial Ecology* 23.2 (2019): 347-360.
10. United Nations, 2015, *Transforming our world: The 2030 agenda for sustainable development*. A/RES/70/1.
11. Y. M. Aleissa, 2022, *Sustainable Process Design to Meet Ecological and Social Goals Through Novel Simulation Tools and Optimization*, Doctoral dissertation, The Ohio State University.
12. Tesla, 2022, *Tesla impact report* (2021).

# Leveraging Semantics and Machine Learning to Automate Circular Economy Operations for the Scrap Metals Industry

Manolis Vasileiadis<sup>a</sup>, Konstantinos Mexis<sup>a</sup>, Nikolaos Trokanas<sup>a</sup>, Theodoros Dalamagas<sup>b,d</sup>, Thomas Papageorgiou<sup>c</sup>, and Antonis Kokossis<sup>a\*</sup>

<sup>a</sup>*School of Chemical Engineering, National Technical University of Athens, Athens, Greece;* <sup>b</sup>*Symbiolabs, SYMBIOLABS, Circular Intelligence, Artemidos 6 & Epidavrou, Marousi 6, 15125, Athens, Greece;* <sup>c</sup>*ANAMET S.A., 160 NATO Av., 19300, Aspropyrgos, Greece;* <sup>d</sup>*IMIS ATHENA Research Center, Athens, Greece*

\*Corresponding author [akokossis@mail.ntua.gr](mailto:akokossis@mail.ntua.gr)

## Abstract

Circular Economy operations and supply chains are becoming increasingly intricate. From raw material extraction, manufacturers and retailers to consumers and recyclers, the benefits stemming from an organized and transparent circular economy network are numerous. Scrap materials have the potential to contribute significantly to the reduction of energy consumption and resource consumption. One is challenged to combine data from scrap metal collection and management, shredding and separation stages, data related to the suppliers and customers as well as background data as they are available from reference sources and classifications available by the metal industry. Integration of data from dispersed and heterogenous sources with in-house ERP systems, alongside a need to automate and extract knowledge and support, call upon a semantically enabled approach as the one demonstrated in the paper. The work combines knowledge from a sizeable actual plant (ANAMET), data technologies (Symbiolabs) and modelling expertise (NTUA). The approach proposes the means to enhance circular economy operations for scrap metal materials with the use of semantic technologies and machine learning techniques presenting examples and cases as they have been used in real-life.

**Keywords:** circular economy, scrap, ontology, machine learning

## 1. Introduction

Circular Economy operations for many types of materials, including metals, are characterized by high complexity. From raw material extraction, manufacturers and retailers to consumers and recyclers, the benefits stemming from an organized and transparent circular economy network are numerous. The use of semantic web technologies, and ontologies in specific, offers many benefits towards efforts to automate low-end knowledge work, automate processes and extract insights for Circular Economy operations. The collection, processing and valorization of secondary materials (waste and/or by-products), have a key role to play in the realization and successful operation of circular economy. Scrap metals (e.g. ferrous and non-ferrous, such as copper, aluminium and iron), in particular, have the potential to contribute significantly in the reduction of energy consumption and resource consumption. Scrap metals are commonly by-products and waste of the manufacturing and consumption of vehicles, electrical and electronic

devices, machinery, wiring, tubes and many other products. This makes their valorization (recycling or otherwise) an attractive proposition for businesses and environmentalists alike. The collection, processing and valorization of by-products or waste materials are key components of a successful circular economy. More specifically, the identification/classification and valorization of scrap metals can play a key role in reducing energy consumption, as well as the waste ending up in landfills. Scrap metals are usually a by-product of manufacturing operations characterized by variability in composition and in their defining characteristics. This variability makes identifying, classifying and standardizing scrap metal burdensome and time-consuming. Circular Economy operations for scrap metals rely on tacit knowledge, embedded on the humans involved in the process. However, the lack of a systemic and standardized approach makes it difficult to leverage this knowledge. In this paper, we attempt to bridge that gap with the use of semantic technologies and machine learning techniques. Ontologies and knowledge graphs (KG) are used to formalize the knowledge in the domain of scrap (and non-scrap) materials and waste, as well as the processing technologies that participate in the valorization paths of scrap metals. Machine learning (ML) techniques are employed to build on the formalized knowledge of the ontology and facilitate further inferences that can help automate the process of identifying and classifying scrap materials, with a view to improve circular economy operations.

## 2. Semantics & Machine Learning for Scrap Metal

### 2.1. Ontologies & Knowledge Graphs

Ontologies and knowledge graphs are knowledge modelling tools that have been extensively used in Process Systems Engineering (PSE), addressing a wide range of knowledge-intensive tasks in a wide range of sectors, with applications ranging from R&D linked data facilitating data sharing, different facets of product development including smart production and data interoperability for manufacturing and product development (Hailemariam and Venkatasubramanian, 2010) as well as decision making for specific processing steps (Akkisetty et al., 2009). Ontologies have also been widely used in applications related to supply chains across a range of sectors. In process industries, ontologies have been used for data integration (Chang and Terpenney, 2009), improved product traceability, information sharing (Munoz et al., 2012), knowledge representation and decision-making (Trokanas et al., 2016). In discrete industries, ontologies have been used for decision-making tasks, such as supply chain negotiations (Wang et al., 2013), knowledge and data integration (Huang and Diao, 2008) among others.

In this work, we developed an ontology to represent the knowledge of the domain. The ontology model is simple, aiming to facilitate ML tasks. To that end, Scrap is modelled as follows (Figure 1). Scrap materials can be classified in three distinct ways, allowing the end-user to extract as much

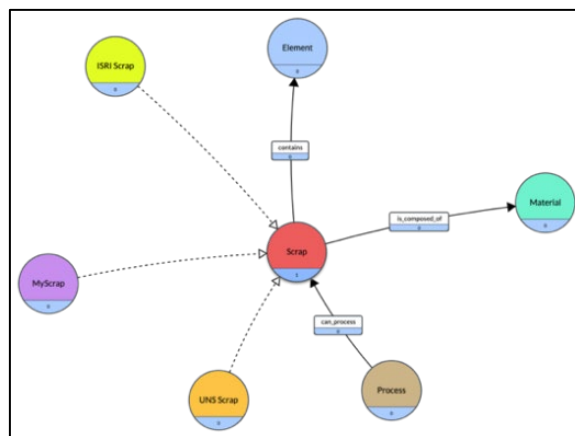


Figure 1. Top-level Scrap Ontology

information as possible when classifying a scrap metal. ISRI Scrap (ISRI, 2018), The Scrap Specifications Circular is a set of guidelines for scrap trading, UNS Scrap, Unified Numbering System is a composition-based system of commercial materials, and MyScrap: the reference company classification (name to indicate the generic approach to the problem). Each type of scrap classification is populated from the relevant documentation and relationships to other relevant entities (i.e. Elements, Materials) are defined.

Materials class represents mostly metal materials that are needed to properly define scrap metals. They form a taxonomy of metals, the top-level structure of which is depicted in (simplified) Figure 2. The ontology is populated with indicative data provided by the partner company, to create the application-level knowledge graph. This KG will be used as input for the ML models and

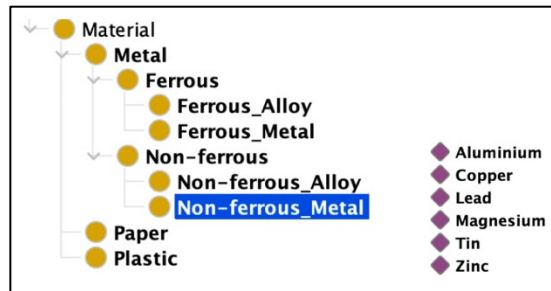


Figure 2. Materials ontology branch and sample instances

facilitate inference of new knowledge about our data. The use of an ontology provides the means to break existing information ‘silos’ and integrate data from different sources. The amalgamation of different sources allows the inference of additional knowledge about different types of scrap and offers a flexibility in describing scrap materials based on three different approaches. All the different scrap metals populate the ontology to form the ScrapKG, a knowledge graph that contains data about scrap materials, their composition, their potential to be processed by certain processing technologies as well as chemical and physical characteristics. ScrapKG is ‘fed’ into the machine learning models (described in the section below), enabling a number of circular economy applications.

## 2.2. Machine Learning

Machine Learning is a branch of Artificial Intelligence (AI) that focuses on the study of algorithms and statistical models that computer systems use to perform a specific task without being explicitly programmed. With the abundance of datasets available, the demand for machine learning is on the rise. In general, ML algorithms have been widely used in waste management, and Artificial Neural Networks (ANN) was the most commonly used method. Several studies have applied ML algorithms to predict the fill levels of waste bins and improve the transportation costs (Pereira Ramos et al., 2018). Finally, various computer vision methods have been implemented for waste classification for effective waste disposal (Ping et al., 2020). Through classification, useful materials are recycled and utilized, which is crucial for a circular economy. Most of the existing literature works use a computer vision approach to the waste management problem, using various waste related images as inputs for the machine learning algorithms. In our work, we attempt to enhance circular economy operations for scrap metal materials with the use of semantic technologies and machine learning techniques. Knowledge graphs and ontologies could formalise (codify) the existing knowledge in the waste management problem, while the use of ML techniques will assist to facilitate further inferences that can help automate the process of identifying and classifying scrap materials, with a view to improve circular economy operations. In this work, we leverage ML techniques to achieve 2 main objectives, namely to:

- (i) identify the main (missing) material of scrap by identifying missing links in the knowledge graph. This builds on processing data provided by our partners and identifying links between scrap materials and metals. The nature of knowledge graphs allows us to use these inferences to expand our KG and facilitate other ML tasks.
- (ii) develop an ML model that classifies scrap metal materials, based on the ISRI standard and the proprietary company scrap classifications of the ontology, allowing users to make inferences about the type of scrap they are dealing with.

### 3. Use Cases/Applications

#### 3.1. Identifying missing links

For this task, we leverage the structure of the graph and the data it contains to develop supervised classification models to identify additional relationships between graph entities. In the case of ScrapKG, we used classification ML methods to identify the composition of different scrap metals. The aim of this task is to improve the completeness of the data for new scrap metals when they arrive at the processing/recycling location, enabling better decision making, by predicting the ‘hasMaterial’ relationship, between scrap metals and materials. To achieve that, several supervised classification methods were trained on the description provided on the ISRI Scrap dataset to predict the material composing the referred scrap. The input of the model contained the text provided in the scrap’s description, whereas the answer set contained the material composing each scrap. Different types of scrap are represented as nodes in the graph and converted into vectors. The models were initially fitted on the training set and then the fitted models were used to predict the responses for the observations in the validation set. The validation set provides an unbiased evaluation of a model’s fit on the training set. The accuracy of all the fitted models on the validation set is presented in Table 1.

Table 1. ML model accuracy

Model	Accuracy on <i>Validation set</i>
<b>Support Vector Machines</b>	85.18%
<b>Random Forest Classifier</b>	85.19%
<b>Gaussian Process</b>	92.25%
<b>K-Nearest Neighbors Classifier</b>	81.88%

To test the validity of the model, we created a new and external test set, consisting of scrap metal descriptions (defined by the seller) from an online marketplace. The input is the description of the scrap materials. Each ML model predicts a material and these results participate in a voting mechanism, to identify the top composition material(s) (Figure 3). Some examples are presented in Table 2 (predicted material). The predicted results are correct, and the accuracy of the test set is 85%.

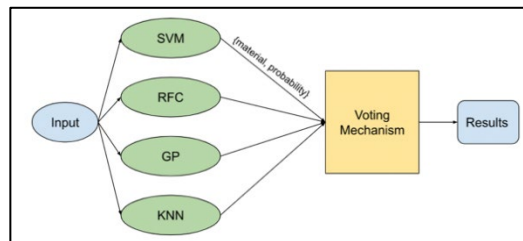
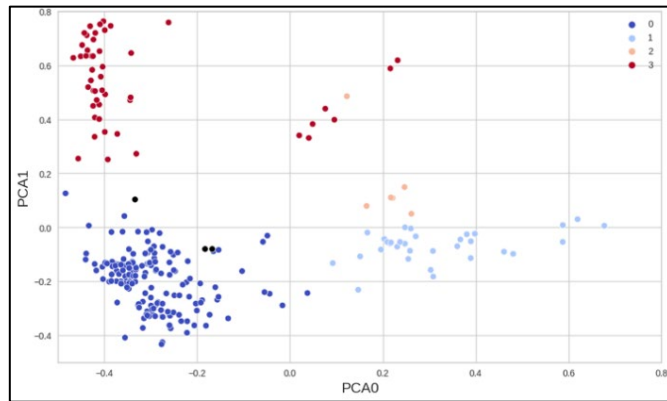


Figure 3. ML Process

#### 3.2. Classification of scrap

For this task, we leverage the structure of the knowledge graph and the data it contains to develop unsupervised learning models that allow us to classify any given scrap. Classification is based on two distinct ontology branches and entities (scrap) are classified

and compared to ISRI classification and the proprietary company classification, allowing users to evaluate the scrap by juxtaposing new types of scrap into existing scrap types. For this task, both ISRI Scrap dataset and the proprietary company scrap dataset were combined to develop unsupervised clustering models to identify clusters of similar scrap metals. The aim of this task is to allow users to make inferences about the type of scrap they are dealing with. To achieve that, we implemented the KMeans clustering algorithm, an unsupervised learning algorithm that partitions  $n$  observations into  $k$  clusters, where each observation belongs to the cluster with the nearest mean serving as a prototype of the cluster. The KMeans algorithm was trained on the description provided on both the ISRI and the proprietary company datasets. Similar to the previous task, graph nodes were converted into vectors. When a new, unknown scrap is provided, the KMeans algorithm will predict the closest cluster it belongs to. Finally, the most similar ISRI or proprietary company scrap will be defined based on the cosine similarity of the nodes of the two scraps. The selected number of clusters was  $k = 4$ . Using principal component analysis to transform the data in two-dimensional space we are able to visualise the clustering (Figure 4). The *black dots* in Figure 4 represent the new and external test data described on the previous task, defined by the seller in an online marketplace.



All of them were classified at cluster 0. The most similar, in terms of cosine similarity, ISRI scrap types are presented in Table 2. From that we can compare the new scrap types to existing categories, to inform decision making.

Table 2 Sample results for composition prediction based on scrap classification

Scrap Description	Closest Category	Predicted Material
Open Extinguisher with no heads   6061 grade	ISRI_85	Aluminium
Regularly purchase of galvalume dross, capacity 300-500mt/month	ISRI_99	Zinc
With no cable, plate or grit. CuSN in cable, plate or grit, separately. Not mixed. In the case of small packaged (30x30 cm) and palletized cable, ball and shot in BIG BAGS	ISRI_61	Copper, Aluminium

#### 4. Discussion & Conclusions

We have presented a method that combines semantic technologies such as ontologies and knowledge graphs and machine learning approaches, to design a framework aiming to automate and assist with circular economy operations for scrap metal materials. We have organized the knowledge of this domain into an ontology and a knowledge graph and used these structured data to identify the key material of a scrap and subsequently classify



it based on existing and known classifications of scrap metals. The ‘new’ scrap is then represented as a graph and is compared and clustered along with already known categories, aiming to identify a cluster and the most similar categories in our graph, allowing users to understand what they have and allow them to make informed decisions regarding the next steps of the scrap material in question.

In future work, the knowledge graph will be extended to include more types of scrap, hence reducing the bias and allowing the implementation of more advanced machine learning models, such as Artificial Neural Networks, aiming to learn and make predictions with minimal data. Additionally, processing technologies will be modelled and populated in the graph, allowing us to predict the best valorization for each scrap metal at the point of receipt.

### Acknowledgments

This research has been co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code:T1EDK-03010)

### References

- ISRI (2018) Scrap specifications circular: guidelines for nonferrous scrap, ferrous scrap, glass cullet, paper stock, plastic scrap, electronics scrap, and tire scrap. Washington
- Akkisetty, G.V. Reklaitis and V. Venkatasubramanian, 2009. Ontological informatics based decision support system for pharmaceutical product development: Milling as a case study. *Computer Aided Chemical Engineering*, 26, pp.159-164.
- Chang, and J. Terpenney, 2009. Ontology-based data integration and decision support for product e-Design. *Robotics and Computer-Integrated Manufacturing*, 25(6), pp.863-870.
- Hailemariam & V. Venkatasubramanian. *J Pharm Innov* (2010) 5: 88. doi:10.1007/s12247-010-9081-3.
- Huang, and S. Diao, 2008. Ontology-based enterprise knowledge integration. *Robotics and Computer-Integrated Manufacturing*, 24(4), pp.562-571.
- Trokanas, M. Bussemaker, and F. Cecelja, 2016. Utilising Semantics for Improved Decision Making in Bio-refinery Value Chains. *Computer-Aided Chemical Engineering*, 38, pp.2097-2102.
- Wang, T.N. Wong, and X. Wang, 2013. An ontology based approach to organize multi-agent assisted supply chain negotiations. *Computers & Industrial Engineering*, 65(1), pp.2-15.
- Xia, W., Jiang, Y., Chen, X., & Zhao, R. (2021). Application of machine learning algorithms in municipal solid waste management: A mini review. *Waste Management & Research: The Journal for a Sustainable Circular Economy*, 40(6), 609–624.
- Abbasi M, El Hanandeh A. (2016) Forecasting municipal solid waste generation using artificial intelligence modelling approaches. *Waste Management* 56: 13–22.
- Pereira Ramos TR, de Morais CS, Barbosa-Povoa AP. (2018) The smart waste collection routing problem: Alternative operational management approaches. *Expert Systems with Applications* 103: 146–158.
- Ping P, Xu G, Kumala E, et al. (2020) Smart street litter detection and classification based on faster R-CNN and edge computing. *International Journal of Software Engineering and Knowledge Engineering* 30: 537–553.

# System identification of scrap metal shredders under minimal process and feedstock information

Manolis Vasileiadis,<sup>a</sup> Thomas Papageorgiou,<sup>b</sup> Antonis Kokossis<sup>a,\*</sup>

<sup>a</sup> *National Technical University of Athens, 9 Heron Polytechniou Str., 15780, Athens, Greece*

<sup>b</sup> *ANAMET S.A., 160 NATO Av., 19300, Aspropyrgos, Greece*

\* *Corresponding author: akokossis@mail.ntua.gr*

## Abstract

Shredders are essential processes in the metal scrap industry. Their efficiency is critical to sustainable recycling as shredders account for the largest energy consumers and the most expensive units to operate in the plant. While such processes are based on relatively simple principles and in operation for a long time, their feedstock (scrap) is difficult to characterize and relate to standard control schemes; operation is carried out empirically with an apparent scope for energy savings once the system is better understood. The paper presents a mathematical modelling approach, specifically aiming at the identification of a conventional shredder with minimum information about its input and output flows. A set of hypotheses, together with a mathematical optimization approach, are combined to estimate energy consumption and the breakdown of feedstock into the major metals involved. The model can be exploited to minimize energy and to connect with available model-based control technology in applications that are not addressed in this paper.

**Keywords:** Scrap metal shredder, mathematical optimization, model identification, energy efficiency

## 1. Introduction

Metal recycling is imperative for environmental and sustainable development. Recycling technologies have attracted significant attention, getting increasingly mature (Reuter et al. 2013) and expected to assume a fundamental role in the transition towards circular economy (George et al. 2015). Rather than based on a fixed composition of virgin materials, scrap metals account for a wide range of by-products (waste) and industrial activities (also referred as end-of-life products) such as end-of-life vehicles (ELVs), waste electric and electronic equipment (WEEE), used beverage cans (UBCs), etc. The importance of metal recycling is undisputed (Hageluken, 2016) with numerous companies already in place to recycle metal scrap. Trading and treatment companies receive scrap from collector companies, recover the metal by physical processes such as dismantling, shredding, etc., and then sell it to metal refining industries. A systems approach to analyze the energy use in the shredders is still missing in the literature, due to a general lack of information and data about the process.

Metal shredders are core processes in trading and treatment companies. Process efficiency is critical to sustainable recycling as shredders account for the largest energy consumers and the most expensive units to operate on-site. While process optimization could deliver significant savings in operating costs, process feedstocks (scrap) are difficult to characterize or systematically relate to the energy input required by the process; the

shredder is eventually operated empirically. To that purpose, the paper introduces a systems approach for the estimation of input to a metal shredder using minimal information largely based on metals composition in the output. The approach fills in for missing information using material properties (especially those related to energies) while its model parameters are prepared for tuning and data embedding using online measurements. The development and validation of the methodology have been coordinated with real-life metal shredders from ANAMET S.A., the largest trading and treatment recycling company in Greece.

## 2. Shredding process: introduction and background

Shredding is a process ultimately aiming in the separation of various metals found in end of life (EOL) products. Shredding generally may be performed in various ways, including metal shredders and aims in the fragmentation of EOL products at a sufficiently fine granularity. Subsequently separation units separate the various materials composing the feedstock that has been shredded. The underlying assumption is that fragments of the original EOL products are sufficiently small so that each one consists of a single material. Energy use relates to the use of hammers (in our case 16 hammers) that break metals into fragments (shredded products) that are usually separated with magnetic separators followed by eddy-current separators. Handpicking is also used at various stages of the process. Products may vary and are typically classified as (1) “Shredded”, (2) Zorba, (3) CuFe, (4) Inox and a by-product (5) Fluff. Feeds, however, include a much wider range of materials and EOL products that include whole cars, car engines, water heaters etc. In our case, they are organized in 21 different scrap categories (ISRI Scrap Specifications Circular, 2022).

Several batch processes,  $N_d$ , are repeated throughout a working day,  $d$ . Each day,  $d$ , a prespecified subset,  $S_d \in \mathcal{P}(\{1,2, \dots, 21\})$ , of the 21 different scrap types is fed to the shredder. In each batch,  $t \in \{1,2, \dots, N_d\}$ , a mass,  $m_t$  of one of the 21 different scrap types,  $s_t \in S_d$ , is fed to the shredder. Once the shredder drum is empty a new batch is introduced. While products  $M_d^i$   $i = 1,2, \dots, 5$ , are weighted every day and they are known, the actual process input is not known precisely. Namely, every day,  $d$ , the subset  $S_d$  is known but the total mass  $m_d^k$  of scrap type  $k \in S_d$  that is fed on day  $d$  is unknown. A model is subsequently required to provide estimates  $m_d^k$  for each different day  $d$ . Reliable measurements and data mostly relate to the product produced and separated on each different day. They include mass flowrates (kg/d), the energy consumption of the metal shredder (kWh), operating hours (each day), and the downtimes (each day). Operational data were made available for 6 consecutive years (2015-2021) and accounted for almost 1200 production days. Measurements of feedstocks are limited and are used as empirical and grossly uncertain estimates.

## 3. Energy consumption at the shredder $\widetilde{E}_d$

The energy consumption of the metal shredder depends largely on the feedstock,  $(m_t, s_t)$ ,  $t = 1, \dots, N_t$ . We assume that (i) the sequence that the various batches are processed is not significant, and (ii) the mechanical properties of the scrap metals are the ones mostly contributing to the energy dissipation in the process. The mechanical properties are strong functions of the composition of each type of scrap, but we can ignore other affecting reasons (such as the scrap shape, heterogeneity etc.). Instead, we are in search for an analytical expression,  $f$ , in the form

$$\overline{E_d} \approx f(m_d^1 \cdot x^{1,1}, \dots, m_d^{21} \cdot x^{5,21}) \approx \sum_{i=1}^5 \sum_{k=1}^{21} e^i \cdot m_d^k \cdot x^{i,k} + c \quad (1)$$

The function  $f$  is unknown and, as simplification, it was assumed that it is linear. Then,  $e^i$  are parameters to be determined via regression analysis of available data. Simple algebraic manipulation of equation (1) results in:

$$\left(\frac{\overline{E_d}}{M_d}\right) \approx \sum_{i=1}^4 e^i \cdot \frac{M_d^{j_i}}{M_d} + c \quad (2)$$

where  $j_i \in \{1,2,3,4,5\} \quad \forall i = 1,2,3,4$ .  $M_d$  is the total production on day  $d$  and  $M_d^{j_i}$  is the mass of product  $j_i$  on day  $d$ . In Equation (2) the choice of index to be excluded from  $j_i$  is arbitrary and aims to achieve better results. Here we allow  $j_i = 1,2,3,4$ . Equation (2) is fitted to available measured data of energy consumption and mass of products for one month. To simplify this process and ensure that  $e^i > 0, \quad \forall i = 1, \dots, 4$  we assumed that there are correlations between parameters  $e^i$  when fitting our available data. These correlations are assumed to be of the form:

$$e^1 = a^{\kappa_1} \cdot e^2; e^3 = a^{\kappa_2} \cdot e^2; e^4 = a^{\kappa_3} \cdot e^2 \quad (3)$$

where the  $\kappa_i$  constant parameters measure the average strength of each type of scrap;  $a$  is a parameter determined by the solution of the constrained regression:

$$\min_a \sum_{d=1}^{20} \frac{100}{20} \cdot \left( \frac{(\overline{E_d/M_d}) - (E_d/M_d) \left( \overline{M_d^i}; c(a), e^2(a) \right)}{(\overline{E/M})_d} \right) \quad (4)$$

s. t.  $c(a), e^2(a) > 0$

The objective spans 20 working days of November 2011 to which Equations (2) and (3) are fitted;  $\kappa_i$  are estimated using values of the tensile strength (Calister, 2017) of the major material comprising each product.

#### 4. Estimation of feedstock

The feedstock is estimated to minimize real and expected production. This is formally expressed in Equation (5), where  $\overline{M}_d^i$  is the estimation of the mass of product,  $i$ , based on the assumed feedstock and B.O.M.  $M_d^i$  is the measured mass of product and  $i, x^{i,k}$  represent the bills of material (B.O.M.) that express fractional compositions of the scrap type,  $k \in \{1, 2, \dots, 21\}$  on product  $i$ . The first set of linear constraints express the mass balances for each product. These expressions are easily derived considering  $x_t^i = x^{i,k} \quad \forall t \in \{t': s_{t'} = k\}$ . Both values refer to date,  $d$ ;  $\widehat{m}_d^k$  is the empirical estimate of the scrap feedstock  $k$  on each day. Finally,  $D$ , is a parameter (essentially a degree of freedom) that defines the number of days, that together with a specific day,  $d'_1$ , specify a date range within which the optimization problem is solved.

$$\min_{\{m_d^k: k=1,2,\dots,21\}} \sum_{d=d'_1}^{d'_1+D} \sum_{i=1}^5 (\overline{M}_d^i - M_d^i)^2 \quad (5)$$

$$\text{s. t. } \begin{cases} \tilde{M}_d^i = \sum_{k=1}^{21} m_d^k \cdot x^{i,k} & \forall i = 1,2,3,4,5 \\ M_d = \sum_{i=1}^5 \tilde{M}_d^i & d = d_1, d_1 + 1, \dots, d_1 + D \\ \sum_{d=d_1}^{d_1+D} m_d^k = \sum_{d=d_1}^{d_1+D} \hat{m}_d^k & \forall k = 1,2, \dots, 21 \\ 0 \leq m_d^k \leq m^{k,U} & \forall k = 1,2, \dots, 19 \end{cases}$$

The optimization minimizes deviations between the estimated production and the production as measured within a specific date range. The second set of constraints enforces that, besides daily deviations between estimates and real measurements, the total production,  $M_d$ , should be equal. Similarly, the last set of equality constraints enforces that, besides deviations in empirical estimates,  $\hat{m}_d^k$ , and model estimates,  $m_d^k$ , the total flowrates should be equal within the specified date range,  $\{d'_1, \dots, d'_1 + D\}$ . The constraints essentially guarantee that the new estimation  $m_d^k$  across a prespecified date range add up to the total scrap quantity of type,  $k$ , that the company received within the date range. This constraint could change as the practice to monitor the overall process changes.

## 5. Results

### 5.1. Energy Estimation

The optimization problem of equation (4) is solved by the enumeration of the objective function on 41 evenly distributed points in the interval  $a \in [1,3]$ ;  $a$  is selected to match the lower objective while simultaneously satisfying the constraints of equation (4). Figure 1 shows the objective function of the optimization problem of equation (4) as a function of parameter  $a$ . Given these results the parameter  $a$  is selected to the value 2. The average error in the estimation of energy consumption increases with the use of Equation (2) (e.g. by imposing physical constraints and positive coefficients on  $e^i$ ) but the latter constraint is important to maintain a physically interpretable solution. The estimated parameters are  $e^2 = 0.7485 \text{ kWh/kg}$  and  $c = 29.1857 \text{ kWh}$ .

### 5.2. Feedstock Estimation

The methodology for the feedstock estimates, as presented in Section 4, has been applied starting from a reference day  $d_1$  (set to 1/4/2019) towards the end of the horizon ( $D=20$ ). The problem is solved in subsequent iterations whereby  $D$  is set to 18, 16, ..., 2 etc., essentially as many times as necessary to cover the largest subset of periods from day  $d_1$ .

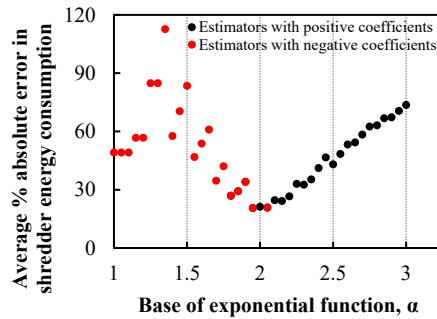


Figure 1. Average error in the estimation of energy consumption as function of  $a$  in equation (4).

The approach always produced feasible solutions with the characteristics described in the previous section. A measure of accuracy was based on an average error:

$$\Delta = \left\langle \frac{1}{5} \sum_{i=1}^5 |M_d^i - \tilde{M}_d^i| \right\rangle \quad (6)$$

Brackets  $\langle \rangle$  in Eq. (6) denote the average error over 20 dates of  $d_1$ . Figure 2 compares  $\Delta$  with empirical measurements with an apparent improvement in the overall estimation as  $\Delta$  decreases as function of  $D$  reaching an eventual plateau.

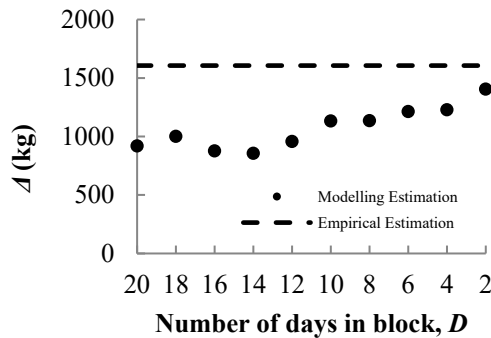


Figure 2 Average error of estimated production over a twenty-day working period starting from 1<sup>st</sup> of April 2019

Figure 3 illustrates the average production error of the proposed methodology on day  $d_1$ ; the error is consistently lower in comparison with empirical estimates for values in  $D$ . The average production error on  $d_2$  appears instead fluctuating, as a consequence of the optimization problem (5) whereby production errors are simultaneously minimized for more days.

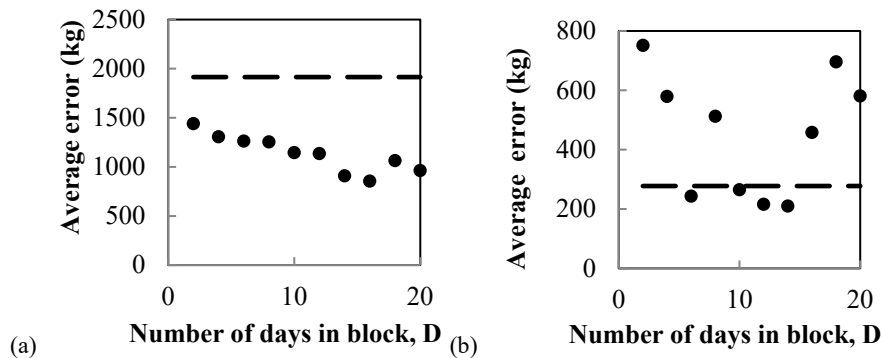


Figure 3 Average error in days 1&2Conclusion

A mathematical modelling approach is presented to identify the composition of unknown feedstocks in scrap metal industries, a problem challenged by incomplete information and process complexity that prevents proper monitoring of the operation. Available data include production flows and energy consumption; feedstock information is available by means of empirical classes of scrap types sorted at the gate. Even with the limited information available, a mathematical modelling approach is validated to deliver reliable input-output relationships with assumptions that include a linear function for energy consumption and five independent parameters assigned for the different product types. Constraints include mass balances, property constraints, classification constraints (feedstocks, products) alongside the postulation of a linear energy function with plenty of room for improvements in considering additional properties for the scrap. The approach is tested on real-life shredders against existing empirical approaches and demonstrating significant improvements for the entire time window tested. Deviations and errors remain significant (e.g. 20%) but this is acceptable given the nature of the problem and the uncertainties that characterize scrap. The work further prepares shredders for control applications by providing reliable estimates for feedstock manipulation (e.g. scrap mixtures) while most expressions in the classification constraints and the energy function are natural entries for data embedding and machine learning. Future research will certainly consider these challenges.

### Acknowledgments

This research has been co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code:T1EDK-03010)

### References

- M. Reuter et al., 2013, Metal recycling. Opportunities, Limits, Infrastructure, United Nations Environment Programme, Editor: International Resource Panel, Working Group on the Global Metal Flows
- D.A.R. George, B. C-A Lin, Y. Chen, 2015, A Circular economy model of economic growth, Environmental, Modelling & Software, 73, 60-63
- C. Hageluku et al., 2016 The EU circular economy and its relevance to metal recycling, Recycling 1, 242-253
- Institute of Scrap Recycling Industries, 2022, ISRI Specs: ISRI Scrap Specifications Circular.
- W.D. Callister, D.G. Rethwisch, 2015, Fundamentals of materials science and engineering. An integrated approach,

# Development of a user-friendly platform for binary interaction parameter estimation

Oscar D. Lara-Montaño, <sup>a</sup> Fernando I. Gómez-Castro, <sup>a\*</sup> Emilio Alba-Robles, <sup>a</sup>  
Brayan Alejandro González-Abundis, <sup>a</sup> Sergio Yovani Rodríguez-Rojó, <sup>a</sup>  
Cristian Geovanni Gutiérrez-Cano, <sup>a</sup> Patricia Isela de la Cruz-Morales, <sup>b</sup> Luis  
Fernando Corado-Castañeda <sup>c</sup>

<sup>a</sup>*Departamento de Ingeniería Química, División de Ciencias Naturales y Exactas, Campus Guanajuato, Universidad de Guanajuato, Noria Alta S/N, Guanajuato, Guanajuato 36050, Mexico*

<sup>b</sup>*División de Ingeniería Química, Campus Centla, Tecnológico Nacional de México, C. Ejido S/N, Col. Siglo XXI, Centro, Frontera, Tabasco, 86751, México*

<sup>c</sup>*Escuela de Ingeniería Química, Facultad de Ingeniería, Campus Central, Universidad de San Carlos de Guatemala, 11 Avenida Zona 12, Ciudad de Guatemala, Departamento de Guatemala, 01012, Guatemala*  
*fgomez@ugto.mx*

## Abstract

Thermodynamic modeling and phase equilibrium calculations are among the core aspects of the formation of chemical engineers. In this work, a platform has been developed to estimate binary interaction parameters for the best fitting of the NRTL and the Wilson models from data provided by the user. Moreover, the user can choose which optimization algorithm will be used. The options currently available are Differential Evolution, Grey Wolf Optimizer, and Jaya Algorithm. To test the platform, the binary mixture composed of allyl alcohol–ethylene is employed. Although the four optimization algorithms can converge to the best solution with enough iterations, the most efficient algorithm for both optimization problems is Differential Evolution. This algorithm produces the smallest deviation from experimental data.

**Keywords:** phase equilibria, binary interaction parameter, user-friendly platform.

## 1. Introduction

Many different chemical engineering processes rely on accurate phase equilibrium calculations. Cubic equations of state, such as Peng-Robinson, Redlich-Kwong, and Soave-Redlich-Kwong, are commonly used to predict the vapor-liquid equilibrium (VLE). However, these equations cannot accurately describe the behavior that occurs in the liquid phase for non-ideal mixtures (Tosun, 2021). To overcome this issue, binary activity coefficient models are employed to predict the interactions occurring in the liquid phase. Some examples of these liquid-phase predicting models are the Wilson equation and the Non-Random Two-Liquid (NRTL) model. Both representations require binary interaction parameters obtained from fitting with experimental data. Once the value of



the binary interaction parameters is obtained, they can be used and introduced into simulation environments to predict the thermodynamic behavior of complex systems.

To identify the value of the binary interaction parameters of the Wilson and NRTL models, it is necessary to have experimental data as the vapor and liquid fractions of the components when phase equilibrium occurs. An optimization procedure can be employed to find the best values of the binary interaction parameters. The objective is the minimization of the deviation between the experimental and the calculated values (Tosun, 2021).

Metaheuristic optimization algorithms are powerful techniques for global optimization. Different authors have employed various metaheuristic techniques to search interaction parameters for activity coefficient models due to the nonlinearity of the equations involved in such optimization problems. Khansary and Sani (2014) used particle swarm optimization (PSO) and genetic algorithms (GAs) to calculate the binary interaction parameters for four quaternary systems for the Wilson, NRTL, and UNIQUAC models. They conclude that the NRTL and UNIQUAC models can better predict the liquid-liquid equilibria (LLE); moreover, the PSO algorithms perform better than GAs. Vakili-Nezhaad et al. (2013) employed GAs to identify the values of binary interaction and non-randomness parameters for the NRTL model and two variants considering 20 ternary aromatic extraction systems containing 16 different ionic liquids at different temperatures. The calculated values produced a satisfactory prediction of the LLE. Jaime-Leal et al. (2015) identified the binary interaction parameters for two case studies using the cuckoo search (CS) algorithm and the e-NRTL model. They report that the CS has a success rate of approximately 90%; also, it is mentioned that the optimization algorithm may fail, especially when evaluating a local optimum solution comparable with the global optimum solution.

This study presents a user-friendly platform developed in Python to determine the binary interaction parameters for the Wilson and NRTL models. The platform is compatible with the three main operative systems (Windows, macOS, and Linux). It can be easily used for teaching purposes to help the student to understand the importance of counting with accurate binary interaction parameters. The user can select among three different metaheuristic algorithms widely used to solve engineering optimization problems; these are differential evolution (DE), grey wolf optimizer (GWO), and Jaya algorithm (JA). Thus, it allows to understand the importance on selecting an adequate optimization algorithm, and on its proper tuning. Also, the user can select among two different objective functions.

## 2. Optimization algorithms

The platform presented in this work allow selecting among three metaheuristic algorithms. The three optimization algorithms are population-based. Each algorithm is briefly described below.

### 2.1. Differential Evolution

The Differential Evolution (DE) algorithm is an evolution-based algorithm proposed by Storn, (1996). This algorithm implements mutation, crossover, and selection as operators in its structure. Different versions of DE are available. In this work, the version DE/x/y/z is employed, where DE refers to Differential Evolution,  $x$  indicates a string denoting the base vector that is being perturbed,  $y$  refers to the number of different vectors used in the perturbation of  $x$ , and  $z$  involves the type of crossover. In mutation, a candidate solution,  $V_i^t$ , is generated with the equation 1, where  $x_{r1}^t$ ,  $x_{r2}^t$ , and  $x_{r3}^t$  are three different random

solutions, and  $t$  refers to the current iteration.  $F$  is a scaling factor that for the selected version is equal to 0.5.

$$V_i^t = x_{r_1}^t + F(X_{r_2}^t - X_{r_3}^t) \tag{1}$$

The crossover operator is applied to the mutant vector  $V_i^t$  to produce a new vector  $U_i^t = u_{i,2}, u_{i,2}, \dots, u_{i,dim}$ , where  $dim$  refers to the number of dimensions. The candidate solution  $X_i^t$  and the vector  $V_i^t$  are considered in the crossover process that uses a probability criterium  $c_r$  as shown in equation 2.  $rand(0,1)$  is a random number between 0 and 1. In the selection process, the new solution vector  $U_i^t$  passes to the new iteration only if it is fitter than the candidate solution  $X_i^t$ .

$$u_{i,j} = \begin{cases} v_{i,j}^t & \text{if } rand(0,1) \leq c_r \\ x_{i,j}^t & \text{otherwise} \end{cases} \tag{2}$$

### 2.2. Grey Wolf Optimizer

The Grey Wolf Optimizer (GWO) was proposed by Mirjalili et al. (2014). It emulates the hunting behavior of grey wolves in nature, which expose a highly hierarchized society. Three main stages are identified in the hunting process: track and chase the prey, pursue, and encircle the prey, and attack the prey.

In this algorithm, the best three solutions of each iteration are identified as  $\alpha$ ,  $\beta$ , and  $\delta$ ; these candidate solutions lead the search process. The first phase of the algorithm (encircling the prey) is modeled according to equations 3 and 4.  $A_i$  and  $C_i$  are vectors of coefficients calculated as  $A_i = 2a \cdot r_1 - \alpha$  and  $C_i = 2r_2$ , where  $r_1$  and  $r_2$  are random numbers between 0 and 1.  $A$  is a parameter that linearly decreases from 2 to 0 in the iterative process.

$$\vec{D} = \left| \vec{C}_i \cdot \vec{X}_p^t - \vec{X}_i^t \right| \tag{3}$$

$$\vec{X}_i^{t+1} = \vec{X}_p^t - \vec{A} \cdot \vec{D} \tag{4}$$

According to the inspiration of this algorithm, the optimal global solution is the prey; nevertheless, in most optimization problems, the solution is unknown. The best three candidate solutions in the current iteration are used to estimate the position in the  $dim$ -dimensional space, as shown in equation 5.  $X_1$ ,  $X_2$  and  $X_3$  are the best, second-best, and third-best candidate solutions, respectively.

$$\vec{X}_p^t = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{5}$$

As the value of the parameter  $a$  decrease, the transition from exploration to exploitation occurs.

### 2.3. Jaya Algorithm

The Jaya algorithm (JA) presented by Rao (2016) is a simple parameter-free metaheuristic optimization algorithm based on the idea that the newly generated candidate solutions must be created towards the best current solution, avoiding the space close to the worst current solution. Equation 6 is employed to update the candidate solutions,  $X$  is the matrix of candidate solutions,  $j$  indicates the decision variable,  $k$  is the candidate solution, and  $i$  the current iteration. The subscript *best* and *worst* refers to the best and worst current candidate solutions, respectively.  $X'_{j,k,i}$  is only accepted if its evaluation in the objective function is lower than the evaluation of  $X_{j,k,i}$ , for a minimization problem. The parameters  $r\_1$  and  $r\_2$  are random numbers between 0 and 1.

$$X'_{j,k,i} = X_{j,k,i} + r_{1,j,i}(X_{j,best,i} - |X_{j,k,i}|) - +r_{2,j,i}(X_{i,worst,k} - |X_{j,k,i}|) \tag{6}$$

### 3. Optimization problem

The purpose of the presented platform is to find parameters for the Wilson and NRTL models, to represent the phase equilibrium from binary experimental data. For a binary system, to calculate the activity coefficients through the Wilson approach, equations 7 and 8 are employed, where  $x_1$  and  $x_2$  are the mole fractions of the liquid phase for components 1 and 2.  $\Lambda_{12}$  and  $\Lambda_{21}$  are the parameters whose value must be determined through the optimization procedure. Equations 9 and 10 are employed to calculate the activity coefficient using the NRTL model, where  $\tau_{21}$  and  $\tau_{12}$  are temperature-dependent parameters.  $G_{12}$  and  $G_{21}$  are calculated as  $G_{12} = \exp(-\alpha\tau_{12})$  and  $G_{21} = \exp(-\alpha\tau_{21})$  and  $\alpha$  is the nonrandomness parameter.

$$\ln \gamma_1 = -\ln(x_1 + \Lambda_{12}x_2) + x_2 \left( \frac{\Lambda_{12}}{x_1 + \Lambda_{12}x_2} - \frac{\Lambda_{21}}{x_2 + \Lambda_{21}x_1} \right) \quad (7)$$

$$\ln \gamma_2 = -\ln(x_2 + \Lambda_{21}x_1) + x_1 \left( \frac{\Lambda_{12}}{x_2 + \Lambda_{12}x_1} - \frac{\Lambda_{12}}{x_1 + \Lambda_{12}x_2} \right) \quad (8)$$

$$\ln \gamma_1 = x_2^2 \left[ \tau_{21} \left( \frac{G_{21}}{x_1 + G_{21}x_1} \right)^2 + \frac{\tau_{12}G_{12}}{(x_2 + G_{12}x_1)^2} \right] \quad (9)$$

$$\ln \gamma_2 = x_1^2 \left[ \tau_{12} \left( \frac{G_{12}}{x_2 + G_{12}x_1} \right)^2 + \frac{\tau_{21}G_{12}}{(x_1 + G_{21}x_2)^2} \right] \quad (10)$$

The platform allows the user to select between two objective functions. The first is shown in equation 11; this minimizes the squared error between the experimental and the calculated activity coefficients. The second, shown in equation 12, minimizes the squared error between the experimental and calculated mole fractions on the liquid phase.  $N$  is the number of experimental values (Tosun, 2021).

$$F = \sum_{i=1}^N \left[ (\gamma_{1,exp} - \gamma_{1,calc})_i^2 + (\gamma_{2,exp} - \gamma_{2,calc})_i^2 \right] \quad (11)$$

$$F = \sum_{i=1}^N \sum_{k=1}^2 \left[ (x_{i,k}^{I,exp} - x_{i,k}^{I,calc})^2 + (x_{i,k}^{II,exp} - x_{i,k}^{II,calc})^2 \right] \quad (12)$$

The decision variables involved in the optimization problem where the Wilson model is present are  $\Lambda_{12}$  and  $\Lambda_{21}$ . Lower and upper limits are given for decision variables, for both these are 0 and 10, respectively. In the optimization problem where the NRTL model is employed, the decision variables are  $\tau_{12}$  and  $\tau_{21}$ , with -5 and 5 as lower and upper boundaries.  $\alpha$  is also a decision variable, the lower boundary is 0.2 and the upper boundary is 0.5.

### 4. Results

The platform to find the binary interaction parameters is developed in Python. The interface is presented in Figure 1. To introduce the experimental data, the user must click on the button “VLE Data”. This opens a window, where it is necessary to select a .xlsx or a .csv file that contains the experimental values. If the user requires to save the obtained values of the parameters it is necessary to click on *Select record*, then the user can choose where to save the results for the optimization problem. After, the user can proceed to introduce the names of the substances that conforms the binary mixture and the pressure of the system. Also, the objective function must be chosen, and the general parameters of the optimization algorithms such as population and number of individuals must be given.

As a results, the user obtains the  $x$ - $y$  and  $T$ - $xy$  diagrams, and the best values found for the binary interaction parameters.

To test the platform, the system allyl alcohol–ethylene glycol is employed. Experimental equilibrium data has been taken from Dong et al. (2018). Table 1 shows the results obtained for both models using the three optimization algorithms. The calculated squared error (SE) is also included. A population size of 50 individuals and 100 iterations are employed. The results obtained in the Wilson model produce very small variations between experimental and calculated activity coefficients. The values obtained for the parameters required in the NRTL model produce a slightly larger SE, but it still correctly predicting the phase equilibrium. DE calculates the lowest SE for both activity coefficient models.

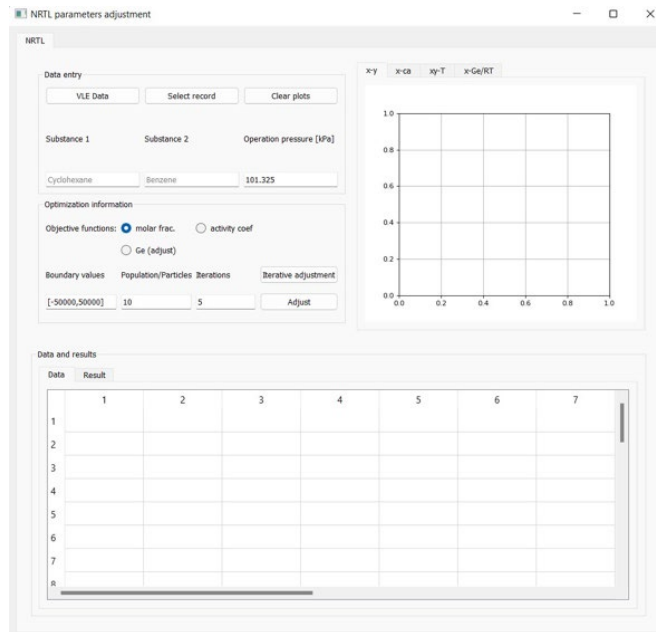


Figure 1. User interface of the platform.

Table 1. Optimal binary interaction parameters.

Optimization algorithm	Wilson model			NRTL model			
	$\Lambda_{12}$	$\Lambda_{21}$	SE	$\tau_{12}$	$\tau_{21}$	$\alpha$	SE
DE	0.5932	0.3480	0.0054	1.0355	0.5524	0.5	0.0117
GWO	0.5923	0.3486	0.0055	1.0363	0.5517	0.4735	0.0117
JA	0.5932	0.3480	0.0054	1.0372	0.5402	0.4851	0.0143

Figure 2 shows the plot for the experimental and calculated activity coefficient for allyl alcohol using the parameters generated by the DE algorithm for the Wilson model. It can be appreciated that all the points lie on the 45° line. On the other hand, Figure 3 presents the experimental and calculated activity coefficient using the best value found for the parameters of the NRTL model with the DE algorithm. Although almost all the points are very close to the 45° line, the data points of the center slightly deviate. This could be due to the higher complexity of the NRTL model. However, the deviation is small. In

general terms, it is observed that the three algorithms generate low values for the deviation, producing reliable values for the interaction parameters.

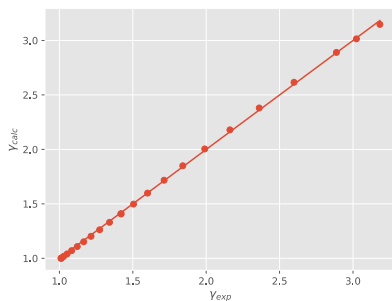


Figure 2. Experimental and calculated activity coefficient for allyl-alcohol using the Wilson model with  $\Lambda_{12} = 0.5932$  and  $\Lambda_{21} = 0.3480$ .

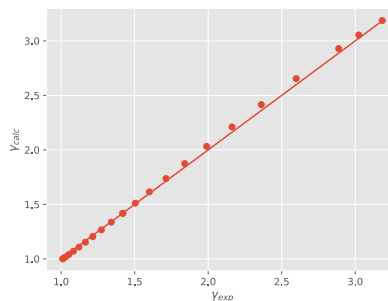


Figure 3. Experimental and calculated activity coefficient for allyl-alcohol using the NRTL model with  $\tau_{12} = 1.0355$ ,  $\tau_{21} = 0.5524$  and  $\alpha = 0.5$ .

## 5. Conclusions

The presented platform is designed to be used with educational purposes to obtain the optimal interaction parameters for binary mixtures represented by the Wilson and NRTL models. The obtained parameters allow the lowest deviation between experimental and calculated composition and activity coefficient values. The platform has a user-friendly interface that allows introducing own experimental data from .csv or .xlsx files. Three optimization algorithms are employed independently either to find the best values for the parameters, or to compare the performance of the algorithms for this application. The quality of the solution depends on the size of the population and the number of iterations. The interface deploys different plots to compare the experimental and the calculated data.

## References

- Y. Dong, R. Zhu, Y. Guo, Z. Lei, 2018, A united chemical thermodynamic model: COSMO-UNIFAC, *Industrial & Engineering Chemistry Research*, 57, 46, 15954–15958.
- J.E. Jaime-Leal, A. Bonilla-Petriciolet, V. Bhargava, S.E.K. Fateen, 2015, Nonlinear parameter estimation of e-NRTL model for quaternary ammonium ionic liquids using Cuckoo Search, *Chemical Engineering Research and Design*, 93, 464–472.
- M.A. Khansary, A.H. Sani, 2014, Using genetic algorithm (GA) and particle swarm optimization (PSO) methods for determination of interaction parameters in multicomponent systems of liquid–liquid equilibria, *Fluid Phase Equilibria*, 365, 141–145.
- S. Mirjalili, S.M. Mirjalili, A. Lewis, 2014, Grey Wolf Optimizer, *Advances in Engineering Software*, 69, 46–61.
- R.V. Rao, 2016, Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems. *International Journal of Industrial Engineering Computations*, 7, 19–34.
- R. Storn, 1996, On the usage of differential evolution for function optimization, *Proceedings of North American Fuzzy Information Processing*, 519–523.
- I. Tosun, 2021, *The Thermodynamics of Phase and Reaction Equilibria*, Elsevier Science Publishing Co Inc.
- G. Vakili-Nezhaad, M. Vatani, M. Asghari, 2013, Calculation of the binary interaction and nonrandomness parameters of NRTL, NRTL1 and NRTL2 models using genetic algorithm for ternary ionic liquid systems, *Chemical Engineering Communications*, 200, 8, 1102–1120.

# Liquid-Liquid Equilibrium Data Correlation: Predicting a robust and consistent set of initial NRTL parameters

Juan A. Labarta\*, José A. Caballero, Antonio F. Marcilla

*Department of Chemical Engineering & Institute of Chemical Process Engineering,  
University of Alicante, PO 99, Alicante E-03080, Spain.*

## Abstract

In the present work, the NRTL model has been analyzed to obtain a good representation of the different ternary liquid-liquid equilibria that this thermodynamic model can reproduce satisfactorily. To do that, the main characteristics of different possible binary subsystems, ternary binodal curves (location in the composition diagram, size, tie-lines orientation, and plait point location), and LLLE tie triangles have been evaluated. With more than two hundred systems studied, the different behaviors have been parametrized to create, as the main objective of the present work, a database and a graphical user interface associated. This resource (publicly available for teaching and research uses) allows, given a set of ternary experimental LLE data to obtain by comparison with the elements of the database, a consistent set of initial NRTL parameters ( $\tau_{i,j}$ ,  $\alpha_{i,j}$ ) to start the corresponding correlation data procedure with enough guarantees.

**Keywords:** NRTL model, phase equilibria, experimental data correlation, LLE, initial parameters.

## 1. Introduction

The correlation of liquid-liquid equilibrium data is still currently a challenging problem not fully resolved, especially for complex systems such as those involving for instance ionic liquids (Zhang et al. 2020), biodiesel (Bazooyar et al. 2021), etc. where the high nonlinearity and non-convexity of the equations involved produce convergence problems, with a strong dependence on the initialization values used of the obtained solution (Labarta et al. 2022a, Klerk and Schwarz, 2023). At this point, it is also important to remark that parameters of thermodynamic models (such as NRTL and UNIQUAC) that are inconsistent with the experimental behavior that is intended to be correlated can be found in the literature as it has been pointed out previously (Marcilla et al. 2017, Glass et al. 2017).

At present, different tools have been developed to help researchers, professionals, and students that have to deal with phase equilibrium correlation to avoid the intrinsic convergence difficulties of these calculations, reduce the search space, and/or confirm the consistency of the obtained results (interaction parameters) in the whole range of composition. For instance, the Graphical User Interface (GUI) Boundaries\_LL\_NRTL allows the visualization of the different LL and LV equilibrium regions that are possible to observe in the NRTL binary interaction parameters space. The modeling as constraints of these boundaries helps to reduce significantly the search space of the unknown parameters in the correlation process (Labarta et al. 2022b). Other GUIs such as GMcal\_TieLinesLL or GMcal\_TieLinesLV (Figure 2) allow checking the consistency of the parameter obtained in the correlation of binary and ternary LLE or (isobaric or

isothermal) LVE experimental data, respectively, by analyzing the topology of the corresponding Gibbs Energy of mixing functions (calculated with NRTL, UNIQUAC or any other alternative model) and its consistency with the equilibrium of the system under study (Labarta et al. 2022c-d). These tools include also, in the case of the LLE, the analysis of the Hessian matrix, Spinodal curve, and critical point location, and in the case of the LVE, additionally the analysis of the necessity of an additional dependence with the temperature or pressure of the binary interaction parameters, and also the possible existence of distillation boundaries.

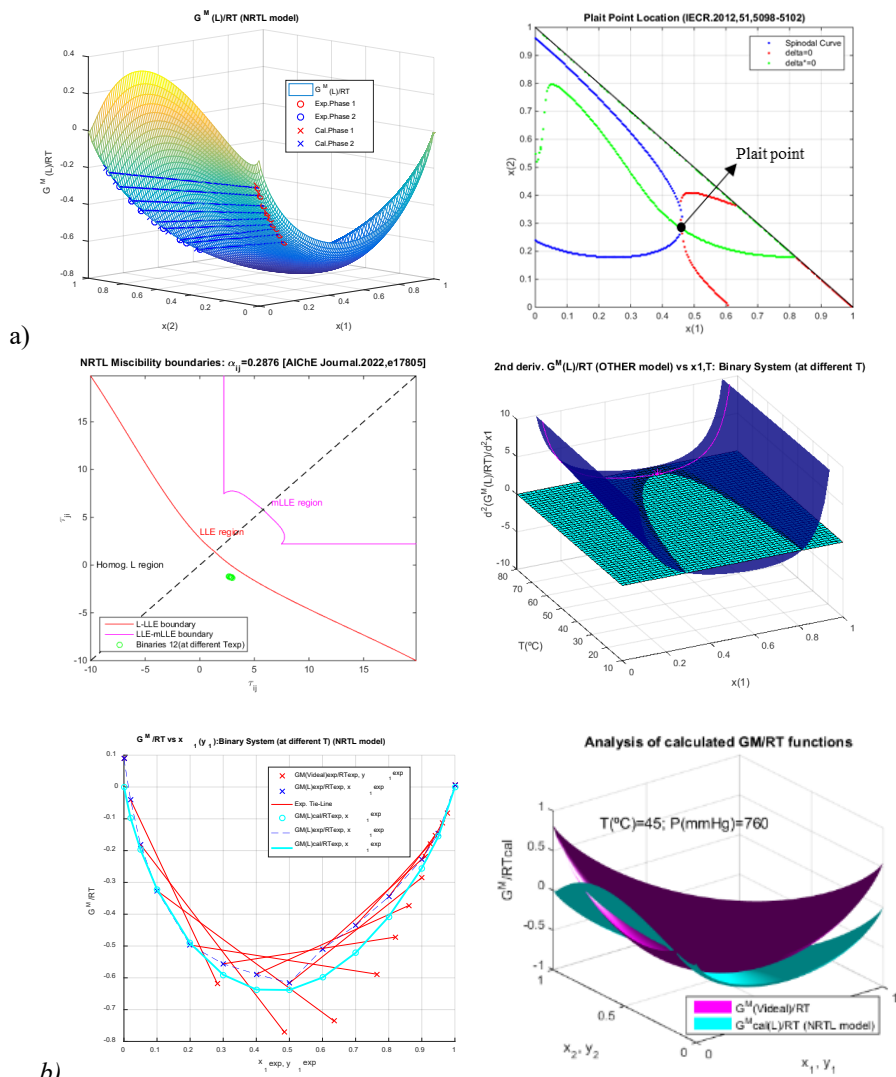


Figure 1. GUIs for the Topological Analysis of Binary and Ternary LLE or LVE Correlation Data Results. Examples of calculations: a) GMcal\_TieLinesLL. b) GMcal\_TieLinesVL

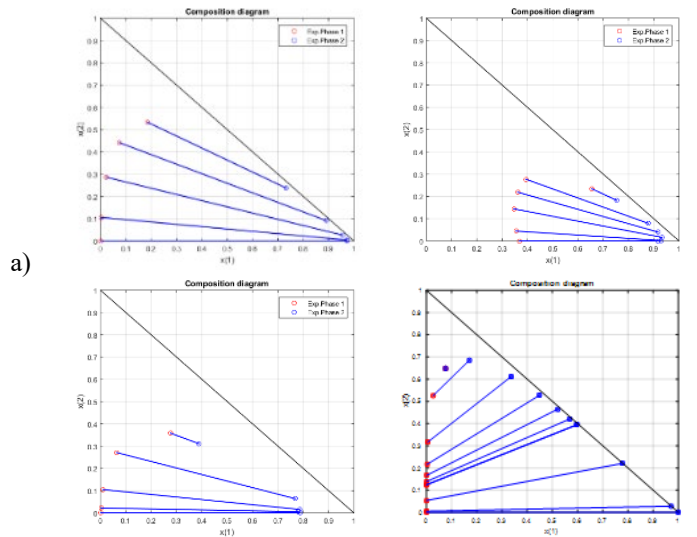
## 2. Methodology and results

In the present work, the NRTL model has been analyzed through more than 200 LLE ternary systems (including type I, II, III, and 0 or island, following the Treybal

classification) to obtain a good representation of the different possible LLE (i.e. binodal curves, tie triangles, etc.) that this model can reproduce satisfactorily, in a wide enough interval of values of the different parameters ( $\tau_{ij}$ ,  $\alpha_{ij}$ ). Figure 2 shows different examples of the results obtained in this scanning of the NRTL model, where it is possible to observe several systems with differences in the number of binary subsystems partially miscible, in the size of the LLE region, and even presenting different possibilities of LLE tie-triangles.

The equilibrium region of all these systems has been parametrized regarding the main characteristics of the corresponding LLE depending on the classification of the system:

- Systems type 1 (with only one binary subsystem partially miscible). In this case, the location and extension of the equilibrium LL regions have been characterized by 5 parameters:
  - Size and average composition of the binary partially miscible (2 parameters)
  - Composition of the plait point (2 parameters)
  - Average slope of the experimental ternary tie lines (1 parameter)
- Systems type 2 (with two binary subsystems partially miscible). 6+4 parameters:
  - Size and average composition of the two partially miscible binaries (4 parameters)
  - Maximum size and average slope of the experimental ternary tie lines (2 parameters)
  - Composition of the two plait points, if exit (4 additional parameters)
- Systems type 3 (with three binary subsystems partially miscible). 12 parameters:
  - Size and average composition of the three partially miscible binaries (6 parameters)
  - Composition of the three points that configure the tie triangle, or the composition of the different plait points, if exit (6 parameters)
- Systems type 0 (with no binary subsystems partially miscible). 8 parameters:
  - Average composition of all the tie lines (2 parameters)
  - Average and maximum size of the tie lines (2 parameters)
  - Composition of the two plait points (4 parameters)





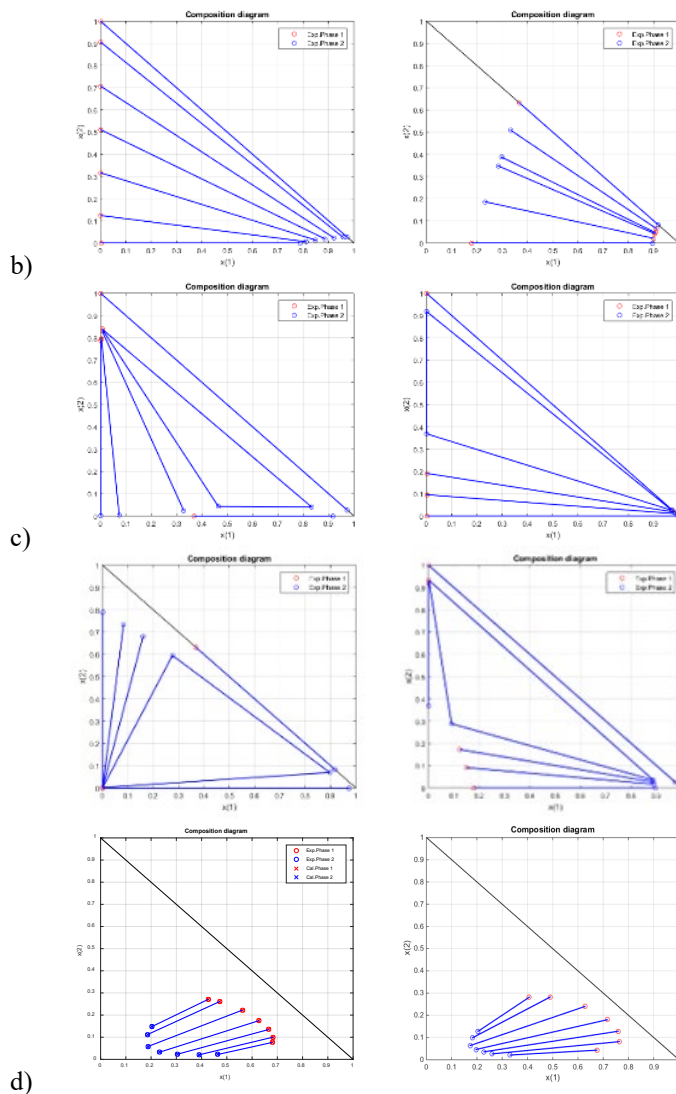


Figure 2. Examples of different LLE behaviors: a) Type I with only one binary partially miscible, b) Type II with two binaries partially miscible, c) Type III with three binaries partially miscible and a tie-triangle, d) Type 0 without binaries partially miscible but with an internal region of two liquids partially miscible.

With this parametrization, we have created a database and a new graphical user interface ParamIni\_LL\_NRTL (Labarta et al. 2022e) that allows loading a set of liquid-liquid experimental data, to obtain by comparison with the different elements of the database, a consistent set of initial NRTL parameters that predicts the parametrized LLE nearest the experimental one (Figure 3). The topology of the Gibbs Energy of mixing function regarding the behavior of the experimental data studied can be also checked, in order to evaluate the goodness of the binary interaction parameters suggested as an initial point in the corresponding correlation data algorithm.

Finally, the suggested set of the binary interaction parameters can be used now in any correlation data algorithm (using the corresponding equilibrium condition: isoactivity,

minimum of the global Gibbs energy of mixing, or the Gibbs energy of mixing minor common tangent plane) to obtain the final rigorous solution (Marcilla et al. 2011, Díaz et al. 2019).

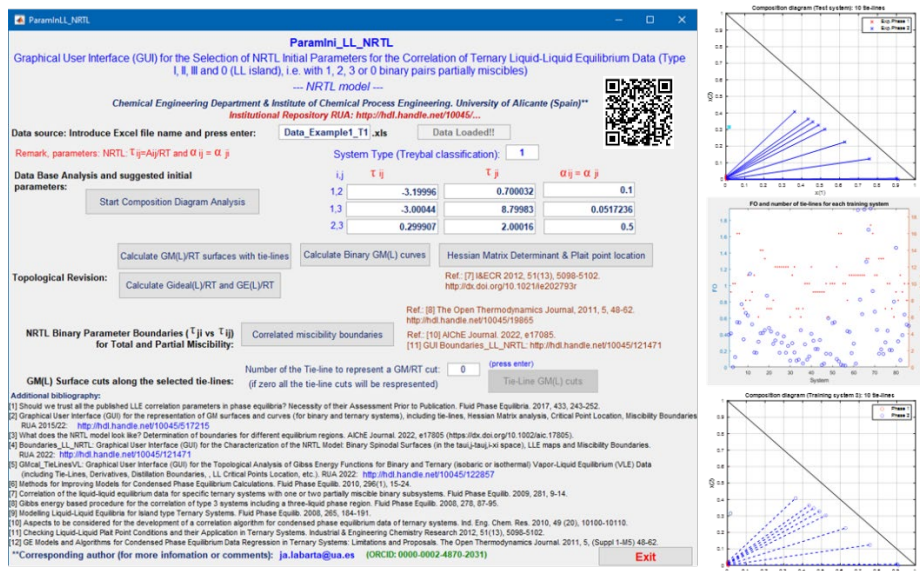


Figure 3. ParamIni\_LL\_NRTL: Graphical User Interface (GUI) for the Selection of NRTL Initial Parameters for the Correlation of Ternary Liquid-Liquid Equilibrium Data (Type I, II, III and 0 (LL island), i.e. with 1, 2, 3 or 0 binary pairs partially miscible).

### 3. Conclusions

We believe that the present work with the analysis and tool developed can help in the task of correlating experimental liquid-liquid equilibrium data, in the sense that allows obtaining a consistent set of initial NRTL parameters, regarding the main characteristics of the experimental data in comparison with the elements of the database, in order to start the corresponding correlation data process. Obviously, the intention is to periodically update the database with more systems, and also to include the UNIQUAC model, which together with the NRTL model is still currently one of the most used models to calculate fluid phase equilibria, taking into account that the correct calculation of the phase equilibrium is a key point in the optimal (efficient and sustainable) design of the corresponding processes and equipment.

### Acknowledgments

The authors gratefully acknowledge the financial support by the Ministry of Science and Innovation from Spain, under the project PID2021-124139NB-C21: SUS4Energy, 2022/00666/001 (AEI).

### References

- B. Bazooyar, F. Shaahmadi, M.A. Anbaz, A. Jomekian, 2021, Intelligent modelling and analysis of biodiesel/alcohol/glycerol liquid-liquid equilibria, *Journal of Molecular Liquids*, 322, 114972.

- A. Bazyleva, J. Abildskov, A. Anderko, O. Baudouin, Y. Chernyak, J. Hemptinne, V. Diky, R. Dohrn, J. Elliott, J. Jacquemin, J. Jaubert, K. Joback, U. Kattner, G. Kontogeorgis, H. Loria, P. Mathias, J. O'Connell, W. Schröer, G. Smith, A. Soto, S. Wang, R. Weir, 2021, Good reporting practice for thermophysical and thermochemical property measurements (IUPAC Technical Report). *Pure and Applied Chemistry*93(2): 253-272.
- I. Díaz, M- Rodríguez, E.J. González, M. González-Miquel, 2019, A simple and reliable procedure to accurately estimate NRTL interaction parameters from liquid-liquid equilibrium data. *Chem. Eng. Sci.*, 193: 370-378.
- M. Glass, M. Aigner, J. Viell, A. Jupke, A. Mitsos, 2017, Liquid-liquid equilibrium of 2-methyltetrahydrofuran/water over wide temperature range: Measurements and rigorous regression. *Fluid Phase Equilib.*, 433, 212-225.
- D.L. de Klerk, C.E. Schwarz, 2023, Simplified Approach to the Parameterization of the NRTL Model for Partially Miscible Binary Systems: Trt LLE Methodology, *Industrial & Engineering Chemistry Research*, 62, 4, 2021-2035.
- J.A. Labarta, J.A. Caballero, 2022e, ParamIni\_LL\_NRTL: Graphical User Interface (GUI) for the Selection of NRTL Initial Parameters for the Correlation of Ternary Liquid-Liquid Equilibrium Data (Type I, II, III and 0 (LL island), i.e. with 1, 2, 3 or 0 binary pairs partially miscible), Institutional Repository of the University of Alicante, publicly available at: <http://hdl.handle.net/10045/130017>.
- J.A. Labarta, M.M. Olaya, A. Marcilla, 2022a, What does the NRTL model look like? Determination of boundaries for different fluid phase equilibrium regions, *AIChE Journal*, e17805.
- J.A. Labarta, M.M. Olaya, A. Marcilla, 2022b, Boundaries\_LL\_NRTL: Graphical User Interface (GUI) for the characterization of the NRTL model: Binary Spinodal Surfaces (in the  $\tau_{i,j}-\tau_{j,i}-x_i$  space), LLE maps and Miscibility Boundaries, Institutional Repository of the University of Alicante, publicly available at: <http://hdl.handle.net/10045/121471>.
- J.A. Labarta, M.M. Olaya, A. Marcilla, 2022c, GMcal\_TieLinesLL: Graphical User Interface (GUI) for the representation of GM surfaces and curves (for binary and ternary systems), including tie-lines, Hessian Matrix analysis, Critical Point Location, Miscibility Boundaries, etc. Institutional Repository of the University of Alicante, <http://hdl.handle.net/10045/51725>.
- J.A. Labarta, M.M. Olaya, A. Marcilla, 2022d, GMcal\_TieLinesVL: Graphical User Interface (GUI) for the Topological Analysis of Experimental and Calculated GM Functions for Binary and Ternary (isobaric or isothermal) Vapor-Liquid Equilibrium (VLE) data (including Tie-Lines, Derivatives, Distillation Boundaries, LL Critical Points Location, etc.), Institutional Repository of the University of Alicante, <http://hdl.handle.net/10045/122857>.
- A. Marcilla, J.A. Labarta, M.M. Olaya, 2017, Should we trust all the published LLE correlation parameters in phase equilibria? Necessity of their Assessment Prior to Publication, *Fluid Phase Equilibria*. 433, 243-252.
- A. Marcilla, J.A. Labarta, M.D. Serrano, M.M. Olaya, 2011, GE Models and Algorithms for Condensed Phase Equilibrium Data Regression in Ternary Systems: Limitations and Proposals. *The Open Thermodynamics Journal*, 5 (Suppl 1-M5), 48-62.
- L. Zhang, X. Bing, Z. Cui, J.A. Labarta, D. Xu, J. Gao, S. Zhou, Y. Wang, 2021, Multiscale Exploration and Experimental Insights into Separating Neutral Heterocyclic Nitrogen Compounds Using [emim][NO3] as an Extractant, *ACS Sustainable Chem. Eng.* 2020, 8, 14, 5662-5673.

# Practical learning activities to increase the interest of university applicants in STEM careers in the era of Industry 4.0

Monica Tirapelle\*, Dian Ning Chia, Fanyi Duanmu, Konstantinos Katsoulas,  
Alberto Marchetto and Eva Sorensen

*<sup>a</sup>Department of Chemical Engineering, University College London, Torrington Place,  
London WC1E 7JE, UK*

\*[m.tirapelle@ucl.ac.uk](mailto:m.tirapelle@ucl.ac.uk)

## Abstract

Inspiring young students, especially young girls, about STEM disciplines is crucial to address the current shortage of engineers. Since the engineering skills that are required by graduates are evolving in line with technological progress, there is now an even stronger need for graduates with strong Process Systems Engineering skills. In this work, we describe an effective way to promote the chemical engineering curriculum, with particular emphasis on computational tools, to a group of Year 12 high school students during a one-week course in our department. The course was designed to engage students in active learning through interactive sessions and practical hands-on activities. Through the course, the students gained a better understanding of the importance of STEM subjects and, in particular, of the challenges and opportunities that engineers encounter in the era of Industry 4.0 with ever-increasing use of digitalization in process design and operation.

**Keywords:** Engineering Education, Hands-on activities, Active learning, Orientation to university, PSE computational tools

## 1. Introduction

The world is currently facing unprecedented challenges in our efforts to achieve the UN's sustainability goals, of which many are closely related to the work of chemical engineers. To achieve these goals, we need a workforce that is suitably equipped to handle the complexities of these challenges, and key to this is science, technology, engineering and maths (STEM) skills (Perkins, 2013). The recruitment into STEM subjects is, at least in the UK, still limited, particularly with respect to gender. In 2019, female students in the UK accounted for only 39% of pre-university, or A-level, examination entries in mathematics, 22% in physics, and 13% in computing. Importantly, relatively few girls opted for two or more STEM subjects at A-Level, which is a requirement to access many university STEM degrees (Department for Education, 2020). To increase the interest of secondary school students towards STEM disciplines and to encourage the enrolment in engineering courses, universities offer outreach activities. Several authors have looked at the impact of these activities and found that, generally, the participants gain a better understanding of STEM fields and enhance their STEM skills (Little and León de la Barra, 2009; Gumaelius et al., 2016). Process Systems Engineering (PSE) involves the use of not only fundamental STEM skills but also of computational tools, and these are not generally considered in schools when teaching science and maths. As a result,

potential applicants to chemical engineering degrees may not be aware of the career opportunities that PSE may offer, or of the new challenges that the profession is facing in the Industry 4.0 era.

In this work, we describe an effective way to promote PSE to Year 12 students, in particular to female students, which is considered at UCL. The students spent one week in the department with the aim of increasing their motivation toward STEM disciplines, and of making them aware of challenges and opportunities that engineers will encounter in the era of Industry 4.0. The course is taking place in the summer before the final year of their final pre-university education (between Year 12 and Year 13). The course was designed to engage students in active and practical learning (Biggs and Tang, 2011; Lucas and Hanson, 2021). To deliver basic key concepts, traditional teaching techniques were replaced with brief interactive sessions. To consolidate learning and to apply knowledge in practice, the students were assigned hands-on activities that involved the use of typical PSE tools, e.g., Excel, Python, MATLAB, and Aspen Plus. The topics covered included basic principles of fluid separations, optimisation, and machine learning. As a final task, the students were asked to present their work to peers and other teaching staff.

The following sections will present the various hands-on activities that were used, as well as the practical requirements for a department to deliver such a course. We also discuss our students' perceptions of STEM before and after attending the course. We found that their attitudes towards STEM, and their interest in pursuing STEM-related careers, increased. More interestingly, their attitudes towards the PSE-related aspects of engineering increased considerably.

## **2. Teaching methods**

The aim of the one-week course is to encourage high-school students, particularly female students, to pursue studies within STEM disciplines and to introduce them to the current challenges and opportunities of Industry 4.0 and the associated use of PSE tools. The week was therefore designed based on topics related to these areas that are not usually covered in the high school curriculum (details can be found in Section 3). To deliver these topics, which can be far from trivial to understand for a high school student, we employed a student-centred teaching approach (Catalano and Catalano, 1999) by providing ad-hoc hands-on activities to enable active learning and by creating an informal learning environment.

Wisniewski et al. (2019) introduced the easy-to-hard effect, which states that a person is more likely to learn a difficult topic if the learning process is progressed from easy to difficult levels. Therefore, the activities in the course were sorted by complexity. In the first couple of days, the students were introduced to the basic concepts and operating principles of mainstream separation processes (see Sections 3.1 to 3.3) by using either Excel spreadsheets or commercially available process simulators. The students were then introduced to the more concept- and code-based topics of optimisation and machine learning (see Sections 3.4 and 3.5). Each topic was kick-started with a simple introduction followed by everyday life examples. This approach elicited the students' curiosity and made them realise how most of the concepts within STEM are inspired by nature and are closely related to their lives.

To deliver the topics, an active learning and inquiry-based approach was adopted. The students were actively engaged in the learning process by allowing them time to ask questions, participate in peer group discussions and hands-on activities to encourage critical thinking (Felder and Brent, 2009). For the hands-on activities, students were split into pairs to work in a team (see Section 3 for more details). To consolidate their learning

and to enhance their communication skills, the students were also given time to record the different topics and their achievements by gradually preparing a PowerPoint presentation that they refined and improved throughout the week and then presented to their peers and teaching staff during the last day of the course. The overall teaching method, with details on the primary goals, is summarized in Figure 1.

Last but not least, we focused the programme development and delivery also on diversity and equality, which we believe to be key concepts underpinning teaching and learning. The teaching team consisted of two female and three male researchers coming from different cultural backgrounds, and during each session, at least one female and one male researcher were present. Such a multicultural teaching team allowed for a free and wide-reaching exchange of ideas and gave the students a taste of university student life, especially in a global university like UCL, which has had strong commitments to diversity and inclusion from its foundation almost 200 years ago. Moreover, by providing female role models, we tried to break the stereotype that engineering is mainly for males and to inspire the female students to apply for STEM university degrees.

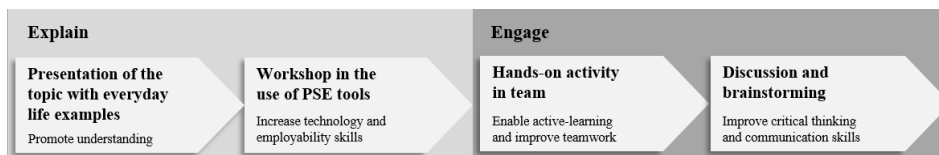


Figure 1. Pedagogical framework

### 3. Case studies

In the following, we will discuss the five main topics covered in the course. The topics cover traditional fluid separation processes such as chromatography and distillation, as well as various methods commonly employed in PSE that are of importance for attaining the UN's sustainability goals. Note that all of these topics are embedded within the chemical engineering undergraduate curriculum at UCL.

#### 3.1. Topic 1 - Liquid Chromatography

On the first day of the course, we introduced the students to the basic principles of liquid chromatography, focusing on the influence of key design and operating parameters on the degree of separation. To get the students interested, we started by showing clips of movies where chromatography is employed by forensic scientists. We proceeded with a brief overview of the topic followed by several short educational videos covering the main terms and definitions of chromatography as well as the basic principles involved. The students were then given an assignment, working in pairs, to find a combination of operational conditions to achieve the separation of an analyte whilst satisfying certain constraints such as maximum time of analysis and minimum peak resolution. To this end, students were provided with an Excel spreadsheet developed by Guillaume and Veuthey (Guillaume and Veuthey, 2016; Veuthey and Guillaume, 2016). The students could explore by themselves the impact of some of the main design parameters on the degree of separation, such as column length, temperature, flow rate, and solvent strength. Furthermore, they learned that trial-and-error approaches are time-consuming, may lead to non-optimal results and, in a real industrial setting, may lead to an increase in operating costs and wastage. With this experience and awareness, the students were ready to be introduced to optimisation techniques later on in the course (see Section 3.4).

### 3.2. Topic 2 - Adsorption

The first session of day two involved an introduction to adsorption, a physicochemical phenomenon involved in various processes studied in Chemical Engineering (i.e., chromatography, catalytic reactions and fluid separation). The students first watched YouTube videos that explained the basic principles of adsorption, and briefly covered the derivation of very simple mass balances and isotherms. We then focused on the significance of isotherm equations, while pointing out that an isotherm includes parameters that are often unknown, and thus have to be estimated, normally from fitting an isotherm equation to experimental data. The next activity involved a hands-on session where the students fitted the Langmuir isotherm to experimental data provided to them by using non-linear regression methods in Excel. The students were able to estimate the isotherm parameters by minimising the mean squared error between the experimental and calculated values of solid phase concentration. This activity allowed the students to handle some experimental data and to experience parameter estimation.

### 3.3. Topic 3 - Distillation

The students were then introduced to distillation. This topic started with the fundamental concepts of distillation, including vaporisation, condensation, and vapour-liquid equilibrium. Then, two different types of distillation processes, batch and continuous distillation, were briefly presented. A group discussion followed about the relative benefits, drawbacks, and applications of each mode of operation, not only to encourage the students' critical thinking but also to examine their understanding. Next, a commonly used process simulator, Aspen Plus (Aspen Technology Inc., 2017), was introduced with step-by-step guidance on how to set up a proper distillation simulation. Next, a hands-on activity of separating a binary mixture in a continuous distillation column was considered, followed by a discussion of the effects of the key design variables on the separation performance using sensitivity analysis. This topic ended with a discussion about distillation column sequencing for the separation of multi-component mixtures (without considering azeotropes) and of energy efficiency. By the end of the day, the students had a better understanding of how highly energy-demanding processes, such as distillation, may become more energy-efficient, smarter, and safer when digital tools are used in their process design and optimisation.

### 3.4. Topic 4 - Optimisation techniques

The third day was focused on the concept of optimisation as one of the main steps for designing a sustainable chemical process. The students were first introduced to the basic elements of optimisation, such as objective function and problem constraints, and then to the following optimisation methods: Linear Programming (LP), Genetic Algorithm (GA) and Particle Swarm Optimisation (PSO).

LP is one of the simplest methods with which to perform deterministic optimisation of linear problems. We first demonstrated how an LP problem could be formulated through equations, that is, by specifying the objective function and the constraints, and then demonstrated the solution using the graphical method. When the students became more familiar with the equation formulation, we demonstrated the solution to the same problem using the Excel optimiser. For the hands-on session, the students had to formulate the equations for four different problems and find the optimal solution using Excel. After the hands-on problems, the students were confident in their ability to express an LP problem using linear equations and in finding the optimum using Excel spreadsheets.

Unlike LP, GA and PSO are heuristic-based stochastic optimisation methods designed to solve more complex problems. We started this topic by showing the students a video on how humans have evolved following the concept of "survival of the fittest", which are

the key features of GA (parent selection, crossover, and mutation). Then, a hands-on activity based on the famous travelling salesman problem, built-in in MATLAB, was used to help the students understand the power of GA optimisation and how to tune the GA parameters. Next, a similar teaching procedure was used for PSO. The students were tasked with obtaining the global optimum by changing the key PSO settings, such as the number of swarms, cognitive parameter, and social parameter. This activity also made the students realise that the results from the optimisation may not be 100% reliable, and that it is necessary to double-check the results using different approaches (e.g., different optimisation methods combined with sound engineering judgement).

### *3.5. Topic 5 - Machine learning*

On the final day, the students were introduced to data-driven modelling and cheminformatics related activities. In particular, they were shown how to apply regression techniques for molecular aqueous solubility prediction starting from selected molecular descriptors as input (Todeschini and Consonni, 2008). Data used to build the machine-learning models were taken from the literature (Huuskonen, 2000). A fully interactive approach for this topic would be quite difficult due to its high complexity. Nevertheless, the students were actively involved when they were asked to calculate some molecular descriptors of interest from the Simplified Molecular Input Line Entry System (SMILES) representation of ten selected molecules among those available in the data set. AlvaDesc (Mauri, 2020), a commercial cheminformatics software, was adopted for this purpose. The students were also shown common preliminary data analysis and data curation techniques, used prior to data-driven model development. Modelling-wise, the students were shown how to perform a linear regression with regularisation techniques to penalise model complexity, as well as how to present results through parity and box plots.

## **4. Results and discussion**

The hands-on activities were designed to give the students the opportunity to consolidate their learning and to put their new knowledge to work. Even if the student background on the topics covered was either very limited or null, the students' performance was outstanding. They were able to complete each task in the time allocated and to attain reliable results. This implies that adopting teaching methods that enable active learning is advantageous and has a positive impact on student learning. To ascertain the students' perspective on the course, we asked the students to fill out the same survey both before and at the end of the course. We designed the survey to target both the students' motivation of pursuing STEM careers (i.e., commitment, perceived value, and awareness), as well as their knowledge and understanding of the topics covered during the week. Questions were Likert-type questions based on a 5-points scale ranging from strongly disagree to strongly agree. Due to the small number of students involved, we cannot do meaningful statistical analysis on the survey results, however, looking at the average scores and comparing the answers given before and after the placement, we can conclude that the course had not actually impacted much on their interest in STEM careers, as this was already high. The students have probably already been exposed to STEM disciplines and were already conscious that a strong STEM workforce is needed, today more than ever, to achieve a sustainable process industry. It appeared, however, that the students felt much more confident about the topics they had considered, and their interest in computational tools had increased. This reflects the fact that the two-ways teaching methods, the teaching environments and the hands-on activities have been adequately applied and had achieved our goal of introducing them to the wonderful world



of PSE. At the end of the course, we also asked the student to fill a questionnaire about their general satisfaction. All students would recommend the placement to a friend.

## 5. Conclusions

In this work, we have presented an engaging way to promote the chemical engineering curriculum to high school students during a one-week course by utilizing an interactive teaching method that involves brief interactive sessions followed by hands-on activities. Through the course, we have enhanced student awareness about the role of chemical engineers in the era of Industry 4.0, and we have promoted the role of females in STEM disciplines. From pre- and post- surveys, we were able to gather some student feedback. Even if we did not significantly impact their interest in pursuing a STEM career, we were able to improve their knowledge of the basic concepts of fluid separation, but more importantly, to increase their curiosity towards the use of computational tools commonly employed by process system engineers. This evidences that the implemented pedagogical method and the informal learning environment, as well as the contents covered in the course, represent effective tools towards getting students engaged and enthused about the subject. We believe that this course can help students in their career choices and, hopefully, get them started on a career journey within Process Systems Engineering. To improve even further the effectiveness of this course, we suggest including an informal meeting with a few engineering employers.

## References

- Aspen Technology Inc., 2017. Aspen Plus V10.  
URL <https://www.aspentech.com/en/products/engineering/aspen-plus>
- J. Biggs, C. Tang, 2011. Teaching for quality learning at university. McGraw-hill education (UK).
- G. D. Catalano, K. Catalano, 1999. Transformation: From teacher-centered to student-centered engineering education. *Journal of Engineering Education* 88 (1), 59–64.
- Department for Education, 2020. Applying behavioural insights to increase female students' uptake of stem subjects at a level.
- R. M. Felder, R. Brent, 2009. Active learning: An introduction. *ASQ higher education brief* 2 (4).
- D. Guillarme, J. L. Veuthey, 2016. HPLC teaching assistant: a new tool for learning and teaching liquid chromatography, part i. *LCGC North America* 34 (10), 804–811.
- L. Gumaelius, M. Almqvist, A. Árnadóttir, A. Axelsson, J. A. Conejero, J. P. García-Sabater, ... & M. Voss, 2016. Outreach initiatives operated by universities for increasing interest in science and technology. *European Journal of Engineering Education*, 41 (6), 589-622.
- J. Huuskonen, 2000. Estimation of aqueous solubility for a diverse set of organic compounds based on molecular topology. *Journal of Chemical Information and Computer Sciences* 40 (3), 773–777.
- A. J. Little, B. A. León de la Barra, 2009. Attracting girls to science, engineering and technology: An Australian perspective. *European Journal of Engineering Education*, 34 (5), 439-445.
- B. Lucas, J. Hanson, 2021. Reimagining practical learning in secondary schools: A review of the evidence.
- A. Mauri, 2020. *alvaDesc: A tool to calculate and analyze molecular descriptors and fingerprints*. Humana Press Inc.
- J. Perkins, 2013. Professor John Perkins' review of engineering skills. London, Department for Business Innovation and Skills.
- R. Todeschini, V. Consonni, 2008. *Handbook of molecular descriptors*. John Wiley & Sons.
- J. L. Veuthey, D. Guillarme, 2016. HPLC teaching assistant: A new tool for learning and teaching liquid chromatography, part ii. *LCGC North America* 34 (12), 906–915.
- M. G. Wisniewski, B. A. Church, E. Mercado, M. L. Radell, A. C. Zakrzewski, 2019. Easy-to-hard effects in perceptual learning depend upon the degree to which initial trials are “easy”. *Psychonomic Bulletin and Review* 26 (6), 1889–1895.

# Integrating Python in the (bio)chemical engineering curriculum: challenges and opportunities

Fiammetta Caccavale<sup>a</sup>, Carina L. Gargalo<sup>a</sup>, Krist V. Gernaey<sup>a</sup>, Ulrich Krühne<sup>a</sup>

<sup>a</sup>*PROSYS, Dept. of Chemical and Biochemical Engineering, Technical University of Denmark, Søtofts Plads, Building 228 A, 2800 Kgs. Lyngby, Denmark.*  
*ulkr@kt.dtu.dk*

## Abstract

In times when developments in digital transformation and online education are accelerating due to the push of the Fourth Industrial Revolution and the Covid-19 pandemic, Education and Digitalization are becoming increasingly intertwined. These recent events have proved the necessity for engineers to adapt by acquiring new technical skills. To facilitate the shift to digital technologies, as well as to enhance students' engagement, this study investigates the introduction of Python as a tool to solve chemical engineering problems, aiming to improve students' learning experience by providing stimulating material and gamification. The Python material, covering topics such as stoichiometry, design of chemical reactors, and mass and energy balances, was embedded into the curriculum of a BEng. Course in Chemical Reaction Engineering at the Technical University of Denmark. Preliminary results seem to confirm that digital tools generate higher engagement and participation among students, leading to better understanding and higher performance.

**Keywords:** Python, digital education, programming in engineering curriculum

## 1. Introduction

In these times of educational disruption, drastic changes have been accelerated, resulting in significant advances in adopting digitalization strategies (Gargalo *et al.*, 2020). In this transformation climate, engineers should be adequately educated to face the challenges and acquire the new skills imposed by Industry 4.0 (dos Santos *et al.*, 2018; Narayanan *et al.*, 2020), coupled with the new learnings from the Covid-19 pandemic.

Previous studies (C. de Las Heras *et al.*, 2021) highlight a clear urgency expressed by students to acquire programming skills in (bio)chemical engineering education, and shed light on the insufficient presence of programming in the curriculum and students' preference to use only one programming language (Python favored over MATLAB).

Python is considered a favorable choice for students and engineers that want to adapt to this digital transformation due to its widespread usage, straightforward syntax, readability, available libraries, broad documentation, support, and suitability for data science applications. Moreover, contrary to tools like MATLAB, it is open-source. Because of all these reasons, Python is attracting the interest of both students and professionals in the chemical engineering community.

Although some tailored courses to learn Python, both broadly and specifically for chemical engineers, are currently available (PyCRE, 2022; AIChE, 2022), they are often

not open access, while the few open-source courses available do not seem to have thorough explanations of either the theory or coding guidelines.

This work aims to evaluate the introduction of a hands-on Python programming course in the (bio)chemical engineering curriculum to (i) address the need to prepare engineers for Industry 4.0; (ii) increase students' engagement and participation in the courses; and, (iii) answer their request for more programming. The course chosen to test the introduction of Python is the BEng Chemical Reaction Engineering course (<https://kurser.dtu.dk/course/28342>) at the Technical University of Denmark (DTU). This course was considered a suitable candidate since educators observed that students generally perform significantly better in group assignments than in the final exam, attributing this to the low active participation in class, especially in the practical exercise sessions.

The main contributions of this work are the following:

- (i) We implement and release a Python course on Chemical Reaction Engineering on GitHub;
- (ii) We embed this course in the educational curriculum and assess it through both quantitative metrics and qualitative evaluations;
- (iii) We discuss the learned challenges and opportunities in embedding programming in the chemical engineering curriculum and summarize our suggestions in a structured framework.

## 2. Methods

### 2.1 Course framework

The *Chemical Reaction Engineering in Python* course was developed at DTU and it seeks to teach Python by providing real use cases and applicable tools. It covers topics such as stoichiometry, design of chemical reactors, and mass and energy balances. The course was created based upon three pillars: (i) relevance and usefulness, with a focus on chemical reaction engineering exercises; (ii) informativeness, aiming to provide thorough explanations, background, fun facts, links to bonus material; and, (iii) entertainment through stimulant exercises and gamification, with embedded reward through three difficulty levels. Practically, the material for each week is divided into two sections: (i) a theory folder, including Jupyter Notebook tutorials, explanations, and coding examples; and, (ii) a practice folder, where students are given exercises.

This material was introduced in the Chemical Reaction Engineering course at DTU (fall 2022). The programming material was designed to run in parallel to the conventional chemical reaction engineering lectures. The general course consisted of four hours per week (plus 5 hours of self-study), where two hours were allocated for lectures and two for exercises. In the last two hours of the class, the Python tutorials were explained and students could work on the exercises and get help from the teachers. The exercises generally followed the theory explained in class, however, some more advanced exercises covered examples solved manually by the students the previous week. This was designed to grant students the time to understand the theory and group exercises explained in class and, once they were confident about these, then they only needed to code the solution in Python. In this manner, the students also had the opportunity to revise the topics from the previous week.

## 2.2 Metrics

To evaluate the success of the course, it was necessary to establish rigorous and concrete metrics. Since the focus of this work is on education, it is essential to measure the learnings. This was done by comparing students' performance in the first given assignment with the results of the same assignment given before the introduction of Python. Additionally, both students' and teachers' perspectives were taken into consideration. Two surveys were conducted to measure the students' opinions regarding learning Python, therefore both quantitative and qualitative analysis can be performed.

## 3. Results and Discussion

### 3.1 Students' perspective

The updated course description, to reflect the introduction of coding in the course, was made available to students before they had to register and start the course. However, the initial survey, conducted at the beginning of the course to quantify the students' interest in learning Python, highlights that only 11.5% of students (among the 26/40 who responded to the survey) were aware that Python would be used in the course. This could be considered as a conclusion by itself since it seems that students did not read the course description in detail and were perhaps not prepared to learn to program.

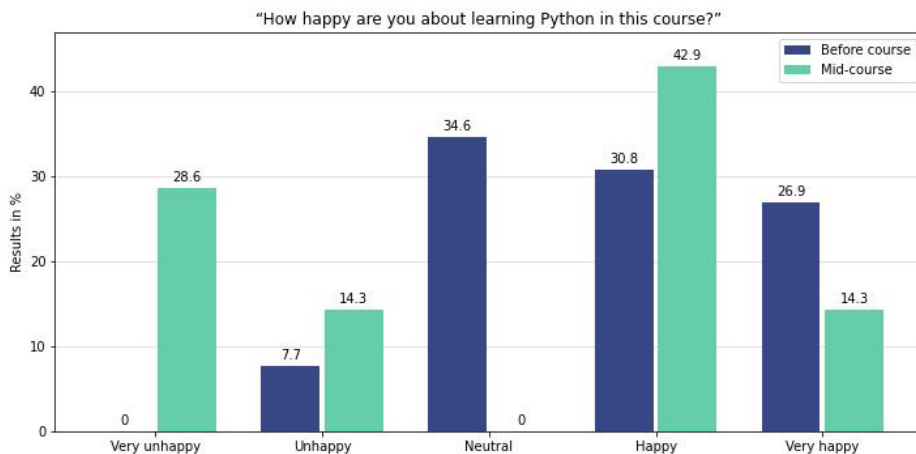


Figure 1: “How happy are you about learning Python in this course?” - Comparison between students' responses before the course started and midway through the course.

Among the various questions, students were asked how happy they were with learning Python. The results in Figure 1 show that 57.7% of the students were either happy or very happy to learn Python, the remaining being mostly neutral (34.6%). We administered another survey to evaluate the Python material in the middle of the course. In this second survey, 57.2% of the students were happy or very happy, while the rest were either unhappy (14.3%) or really unhappy (28.6%). The interesting trend shown is that all the students that had a neutral opinion regarding Python at the beginning, express more polarity in the second survey, turning mostly to unhappy.



### *3.3 Performance in the assignments*

To help students learn additional skills and material, in the 2022 fall edition of the course it was decided to only administer two assignments instead of five. Python exercises, allowing students to practice their developing knowledge were however given every week, but these were not graded. The first assignment was submitted by 72.5% of the students (29/40) enrolled in the course, while in the fall 2021 edition, 72% of students (36/50) submitted an answer to the assignment released in the same period. The same assignment was given to students in the fall 2021 edition earlier in the semester, and 42/50 students submitted it. Worth mentioning is that some of the students registered do not actively participate in the course, but are only interested in the re-exam.

In 2021 students gained, on average, a score of 7.86/10, while the students taking the course in the fall of 2022 have an average score of 8.17/10. Although the difference is not significant, it could indicate that programming helps students to be more engaged. The course is still ongoing, therefore we do not know the performance of the students in the final exams; however, we expect them to perform better than in the last years, given that many students seemed to be engaged in the lectures and the exercise sessions, asking questions and actively participating.

## **4. Learnings and Educational Framework**

Among the learnings from this experiment, is that students consider the amount of work required by the course to be too high since they also need to learn and understand the regular material designed for the specific course. However, it is also important to reflect upon and weigh students' opinions regarding the amount of material, since the course needs to be challenging and stimulate their critical thinking. Therefore, it is necessary to balance the amount of work required while making the course intellectually stimulating and educational. This could be solved in the future by dividing the material into mandatory and optional modules. This division would allow the students that do not feel too comfortable with programming to still learn the basic syntax and how to solve easier exercises, while the students that are more motivated to learn Python could challenge themselves with more advanced exercises.

An issue that students have pointed out is the variety of programming languages and tools required by different courses during their education. An effort should be made at institutional level to align on a common tool, or at least to give the students the option to choose among a few selected software/programming languages for each course. This would also reduce the added stress of having to learn to program during another course because it would be a skill that could be taught independently as a stand-alone course.

The course was designed to be self-sufficient, enabling students to go through the material and learn by themselves. However, as also suggested by the students in the survey, probably more guided and in-class programming would be helpful to students to better understand how to read, write and think in code.

Although the subject of the course (chemical reaction engineering) might seem very specific, the content covered (i.e., general syntax, data structures, data visualization, stoichiometry, solving differential equations and integrals) can be used in many more topics within the chemical engineering field. Its large applicability and relevance are the motivations behind the open-source nature of the course. The Python course is publicly available on GitHub at <https://github.com/FiammettaC/Chemical-Reaction-Engineering-in-Python> and in the near future, it will be integrated into BioVL ([www.biovl.com](http://www.biovl.com)), an online educational platform for (bio)chemical processes developed by the authors.

This project is part of a series of courses designed to teach Python to Chemical Engineers. In the future, we plan to compare the results of the course in two different settings: (i) embedded into the curriculum of a selected Bachelor's course at DTU, and (ii) open-source on GitHub, therefore available to whoever wants to take the course. The aim of this comparison is to investigate various approaches to teach chemical engineering students how to code in Python and share our learnings to improve the educational methods needed to prepare future engineers for Industry 4.0.

## 5. Conclusion

This work establishes and evaluates a pedagogical framework to teach chemical reaction engineering in Python. The course has been embedded in the educational system at DTU. It aims to educate future engineers for Industry 4.0 and improve the quality of education of (bio)chemical engineering processes and increase students' engagement.

Even though generating changes in the educational system and promoting digitalization initiatives is rather challenging, we evaluate the introduction of a new tool in the engineering curriculum and discuss the learnings from this experience. Among these, students considered having the Python material introduced in parallel to the normal course as time-consuming and quite difficult, since they had to develop two sets of skills at the same time. A better approach could be to have a stand-alone programming course at the beginning of their education, which the other courses build upon. This would also solve another challenge discovered in this work, that there is no institutional coherence in the tools chosen by every single course, resulting in students having to learn a variety of programs for different courses. Overall, students found the course to be useful. The course, which has large applicability and relevance across various chemical engineering areas, is publicly available on GitHub.

## References

- de Las Heras, S.C., Gargalo, C.L., Gernaey, K.V. and Krühne, U., 2021. Programming skills across the (bio) engineering curriculum—a students' perspective. In *Computer Aided Chemical Engineering* (Vol. 50, pp. 2039-2044). Elsevier.
- dos Santos, M.T., Vianna Jr, A.S. and Le Roux, G.A., 2018. Programming skills in the industry 4.0: are chemical engineering students able to face new problems?. *Education for Chemical Engineers*, 22, pp.69-76.
- Gargalo, C.L., Heras, S.C.D.L., Jones, M.N., Udugama, I., Mansouri, S.S., Krühne, U. and Gernaey, K.V., 2020. Towards the development of digital twins for the bio-manufacturing industry. *Digital Twins*, pp.1-34.
- Narayanan, H., Luna, M.F., von Stosch, M., Cruz Bournazou, M.N., Polotti, G., Morbidelli, M., Butté, A. and Sokolov, M., 2020. Bioprocessing in the digital age: the role of process models. *Biotechnology journal*, 15(1), p.1900172.
- S. C. de Las Heras, et al, 2021, Programming skills across the (bio) engineering curriculum—a students' perspective, in *Computer Aided Chemical Engineering*, vol. 50, pp. 2039–2044, Elsevier.
- PyCRE 2022, Python for Chemical Reaction Engineering (PyCRE), 2022, Available at: <https://www.lct.ugent.be/PyCRE> (Accessed 30/11/2022).
- AIChE 2022, Introduction to Python for Chemical Engineers, Available at: <https://www.aiche.org/academy/courses/ela270/introduction-python-chemical-engineers> (Accessed 30/11/2022).

# An Online Course for Teaching Process Simulation

Daniel R. Lewin,<sup>a</sup> Assaf Simon,<sup>a</sup> Sapir Lifshiz Simon,<sup>a</sup> Asia Matatyaho  
Ya'akobi,<sup>a</sup> and Abigail Barzilai<sup>b</sup>

<sup>a</sup>*Department of Chemical Engineering, Technion. I. I. T., Haifa 32000, Israel*

<sup>b</sup>*Centre for the Promotion of Learning and Teaching, Technion, Haifa 32000, Israel*

*Email for correspondance: [dlewin@technion.ac.il](mailto:dlewin@technion.ac.il)*

## Abstract

A course in process simulation prepares students to take on a capstone design project, in that it teaches them to use a commercial process flowsheet simulator effectively. This paper describes an online process simulation course with the following features: (a) Self-paced instruction using online multimedia modules; (b) Regular practice using UniSim<sup>®</sup> to ensure efficient and effective modeling capability; (c) Online examination of the students' mastery including programming assignments using UniSim. The paper shows that not only do the students accept the importance of acquiring practical capabilities in process simulation, but analysis also shows that the learning outcomes have improved since the move to online practical examination as part of the course.

**Keywords:** Process simulation, process design, online education, commercial simulators.

## 1. Introduction

Most undergraduate programs in chemical engineering incorporate a capstone design project completed in the last year of study, which usually requires the student participants to use a commercial process flowsheet simulator efficiently, intelligently, and critically. It is impractical to expect students to be able to tackle large-scale design projects at the same time that they are learning how to acquire adequate capabilities in process simulation. Belton (2016) advocates teaching students to use the simulator in the first year of studies. However, most chemical engineering programs teach the principal building-blocks of the curricula over time. In the first year, students would not have the background and engineering sophistication to use the simulator, with the danger that students would resolve problems technically, and without understanding the physical phenomena underlying each unit operation simulated. Alternatively, as suggested by Dahm et al. (2002), this training could be integrated throughout the curriculum, by introducing elements of simulation applications in all of the core courses (e.g., thermodynamics, heat transfer, separation processes, and reactor design). However, summative assessment of the skills acquired by students is more difficult for this alternative (Lewin et al., 2006). The solution that has been implemented at Technion since 2005 is to require students to take a course in process simulation one semester in advance of the capstone design course (Lewin and Barzilai, 2021). This is similar to the learning approach offered in the book "Learn Aspen Plus in 24 Hours," (Adams, 2022). The main advantage of these "just in time" solutions is that students receive training in the usage of a process simulator in parallel to their formal instruction in core subjects such as multi-component thermodynamics, reactor engineering, and multiple-stage separations, and are thus better equipped to integrate their formal knowledge of these subjects with their practical implementation.



Before the COVID-19 pandemic, the course was run as a two-hour/week lab, in which students worked at their own pace, guided by multimedia modules (Lewin et al., 2003). At Technion, we use UniSim<sup>®</sup> as the process simulator. To exhibit mastery, students were required to pass a closed-book exam comprising of largely theoretical, multiple-choice questions. It has become clear that successfully passing such an exam is not a guarantee that students have actually mastered the practical usage of UniSim, a fact that was observed in the average student's performance on the capstone design project in the semester following the course. The COVID-19 pandemic required the students to work 100% online, meaning that they had to install UniSim on their own PCs, and the weekly sessions were moved to Zoom Breakout Rooms. Furthermore, the exam was run online also, making it feasible to test the students' practical capabilities on typical design tasks using UniSim as part of the exam. This change was the most influential in significantly improving the students' capabilities. Since the pandemic, the 100% online format of the course and its exam has been retained, thus maintaining the level of mastery required from the students. This paper presents the course syllabus and its learning outcomes. We also compare course outcomes before and after the changes described above. Finally, we present salient conclusions from a poll conducted on the students in the class of 2022, the year after the COVID-19 pandemic largely abated, which was the third time that we ran the course 100% online.

## 2. Course Learning Objectives, Syllabus and Schedule

The course runs over 12 weeks, according to the schedule shown in Table 1. The main learning objective of the course is to instruct students to effectively use UniSim to model process flowsheets. Each week, students proceed at their own pace, guided by self-paced multimedia modules (Lewin et al, 2003). The multimedia modules covering UniSim are freely available at the link: <https://www.seas.upenn.edu/~cbe400/hysys-unisim/>.

Table 1. Simulation Laboratory Weekly Schedule.

Week	Topic	Content
1	Getting started in UniSim	User interface, simulation basis, stream definition, and unit operations. Definition of degrees of freedom.
2	M&E balances and recycle	Recycle convergence. Dealing with multiple recycles.
3	Thermo-package selection	Package selection rules. Adjust and set operations.
4	Heat exchangers	Coolers, heaters, and heat exchanger operations.
5	Plantwide simulation	Modeling the reaction section in the Hydrodealkylation of toluene (HDA) process.
6	Reactors (Part 1)	Defining reactions: conversion, equilibrium and PFR operations. Simulation of a cold shot NH <sub>3</sub> converter.
7	Reactors (Part 2)	CSTR operation. Multiplicity.
8	Separations	Separator, splitter, shortcut column, and distillation column operations.
9	Distillation column workshop	Separation of a ternary mixture of hydrocarbons into three products of specified purity using two columns.
10	Optimization	Two examples using the UniSim optimizer.
11	Project	Developing a separation sequence to produce ethanol from beer (involving flash, distillation and absorbing columns).
12	Recap	Preparation for multiple-choice exam questions.

Weekly quizzes and several extended programming assignments enable students to formatively gauge their acquired capabilities, with a final exam intended for summative assessment. The course grading is distributed as follows: Final exam – 75%, Three assignments – 13%, Seven weekly quizzes – 7%, Weekly attendance – 5%. Note that 25% of the course grade is awarded for attending and carrying out the in-course tasks, with the lion’s share being awarded to successfully demonstrating mastery in the exam. Before the COVID-19 pandemic, the weekly meetings were held in our computer laboratory, in 2-hr sessions of 20 students in each session, with work carried out using UniSim installed on the lab’s computers. The final exam involved theoretical multiple-choice questions, which did not adequately test practical abilities on UniSim. Since the annual enrolment in the course is usually in excess of 50 students, the computer lab’s capacity of 20 stations effectively eliminated even having a portion of the exam to include practical testing of UniSim mastery.

The COVID-19 pandemic forced all teaching to move to 100% online, and this meant moving the simulation course to be run on Zoom, requiring all students to also transfer their UniSim licenses to their own laptops, which is permitted by the company. Consequently, the weekly exercise sessions were run in Zoom Breakout Rooms, enabling self-paced learning by the students as before, but in addition permitting staff to assist individual students on demand. In each session, sufficient breakout rooms were opened, allowing students to work alone or in teams of their choosing, and supporting peer instruction and teaching them to work in groups. The summative exams also had to be run on Zoom, with appropriate proctoring. In the first year of the pandemic, the previous exam format relying on theoretical multiple-choice questions was retained, but in the following year, the format was changed to include two extended UniSim assignments, intended to test the students’ capabilities in detail. Typical UniSim questions would be:

- Given a fully defined feed stream consisting of three components, design a distillation train to produce three product streams of 99% purity, while ensuring reboiler duties do not exceed a specific amount, while maintaining the total number of trays used below a certain amount.
- Set up a refrigeration cycle using ammonia as the refrigerant so that the heat duty in the evaporator meets a given specification.

### 3. Outcomes Assessment

Table 2. Exam grades and fitted bimodal model parameters, 2017-2022.

Year	$N$	$N_{\text{fail}}$	$\mu$	$\sigma$	$\mu_1$	$\mu_2$	$\sigma_1$	$\sigma_2$	$p$
2017	81	7 (9%)	0.72	0.15	0.77	0.73	0.08	0.18	0.33
2018	61	17 (28%)	0.68	0.18	0.80	0.42	0.10	0.07	0.78
2019	54	14 (26%)	0.67	0.16	0.77	0.45	0.08	0.06	0.74
2020	40	5 (13%)	0.68	0.15	0.75	0.70	0.04	0.18	0.34
2021	46	3 (7%)	0.78	0.26	0.88	0.74	0.15	0.03	0.90
2022	58	13 (22%)	0.73	0.22	0.94	0.38	0.29	0.08	0.91

Table 2 presents summative data from the last six years, that is, exam grade averages, standard deviations, and failure rates, as well as computed parameters for a bimodal distribution model fitted to the grade distributions. Figure 1(a) shows the evolution of course outcomes, as brought by the final exam grade distribution, as approximated following the approach of Lewin (2021); each exam grade distribution is approximated by a bimodal distribution model comprising a weighted sum of two normal distributions, yielding estimates for averages and standard deviations of high- and low-performing

subpopulations ( $\mu_1$ ,  $\sigma_1$ ,  $\mu_2$ , and  $\sigma_2$ ), as well as the fraction of high-performers ( $p$ ). The fitted model for the grade distribution is therefore:

$$f(x) = \frac{p}{\sigma_1\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2} + \frac{1-p}{\sigma_2\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2}. \quad (1)$$

Figure 1(b) shows a typical fit, for the exceptionally good exam grade distribution of 2021, the first year that the final exam included practical assignments using UniSim.

In Figure 1(a), each year's result is represented by a disk centered on the average grade for the high- and low-performing students,  $\mu_1$  and  $\mu_2$ , respectively, with a disk diameter proportional to  $p$ . Evidently, a good outcome is either characterized by high values of  $\mu_1$  and  $\mu_2$ , in which case the magnitude of  $p$  is not important, or high values of  $\mu_1$  and  $p$ , in which case the magnitude of  $\mu_2$  is less important. Thus, disks of arbitrary diameter positioned on the top right of the bubble plot, or disks of large diameter positioned anywhere on the right of the diagram constitute desirable outcomes. In Figure 1(a), the red bubble shows the outcome in 2020, when the course was taught online for the first time, using an exam format as in previous years. It is noted that the outcomes are in the range seen in previous years, when the course was taught face-to-face in the computer laboratory. From 2021, the online exam has included practical exercises involving the usage of UniSim, and it is noted that the outcomes have significantly improved: Both the blue disk, for 2021 and the green disk, for 2022, are to the right of the plot, indicating significantly improved outcomes over those achieved before the switch to requiring mastery in practical UniSim assignments. The exam format change did not affect its level of difficulty. The difficulty level of the multiple choice questions was the same as in previous years, as were the practical UniSim questions.

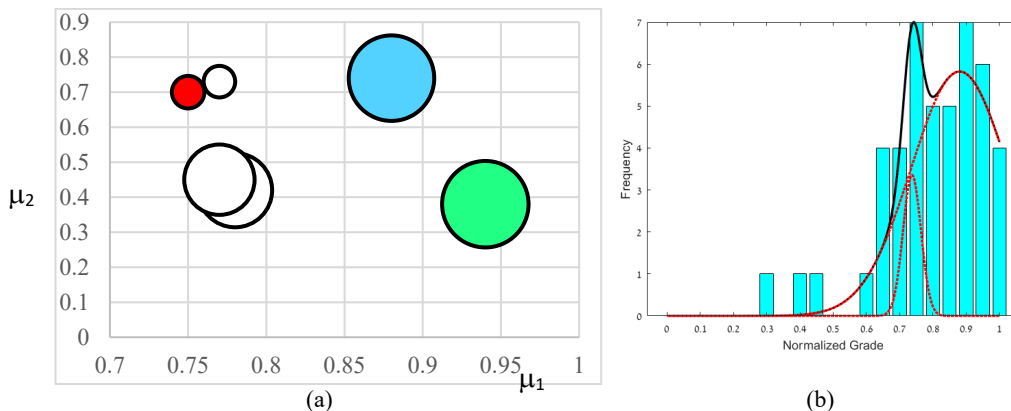


Figure 1. (a) Bubble plot showing the evolution of course outcomes from 2017-2022. The white disks indicate the outcomes for 2017-2019, when the course was taught face-to-face and examined on paper. From 2020, the course was taught and examined online, and from 2021, the online exam has included UniSim assignments. The 2020, 2021 and 2022 outcomes are shown in red, blue, and green, respectively. (b) A typical bimodal distribution fitted to the exceptionally good exam grades for 2021, where the normal distribution for the high- and low-performing students are shown in red-dotted lines, the combined bimodal distribution is shown in black, and the histogram of actual grades is shown in blue.

#### 4. Feedback from Students

The students enrolled in the Simulations Lab in the Spring of 2022 were asked to respond to an online questionnaire, involving 11 questions, mostly requiring a response on a Likert scale from 1, indicating the lowest level of confidence or highest level of disagreement to a statement, through 3, indicating ambivalence, to 5, indicating the highest level of confidence or highest level of agreement to a statement. The number of responses (out of a total of 65 enrolled students), average scores and standard deviations for each question are presented in Table 3.

The first 10 questions all returned positive, non-ambivalent responses (scored averages greater than 3), indicating positive feedback overall. Q1 and Q4 scored the lowest, most likely because many students may have been uncertain about the meaning of “critical thinking” and “troubleshooting.” Q6 – Q9 scored the highest: The highest score was returned on the exam as a motivator for learning to use UniSim, with a slightly lower score also on the subsequent semester’s design project as a motivator, though the difference is not statistically significant. Most students agreed that the online format is the most suitable way of teaching software skills.

Table 3. 2022 Questionnaire and Responses

Question	<i>n</i>	$\mu$	$\sigma$
To what extent did working with the simulator help you acquire/improve the following skills:			
Q1. Critical thinking	50	3.06	1.17
Q2. Solving engineering problems	51	3.67	0.91
Q3. Solving complex engineering problems	50	3.54	0.95
Q4. Troubleshooting	50	3.34	1.12
Q5. Application of knowledge from previous courses	51	3.71	1.10
Q6. Design skills	49	3.82	0.86
Q7. To what extent do you recommend learning software in a format similar to the one studied in the current course?	50	3.94	1.08
Q8. To what extent did the knowledge that you would be required to solve problems with the simulator during the exam affect your motivation to learn to work with the simulator during the course?	50	4.02	1.08
Q9. To what extent did the knowledge that you would be required to solve problems with the simulator in next semester’s design project affect your motivation to learn the simulator during the course?	51	3.84	1.16
Q10. To what extent do you recommend continuing with this type of exam format?	49	3.65	1.01
Q11. Any other comments?	See typical responses below		

The last question prompted verbal responses from many students, and we report here some example responses, indicating mostly a favorable impression of the course’s pedagogical approach:

- The format of the course is very efficient in terms of independent work and the presence of the teaching assistants when needed.
- I really like learning in this style. It was a perfect course and excellent practitioners!!
- In this course, distance learning works very well because there is only self-learning. And when the moment of practice comes, working online works very well because all the operations are only on the computer.

- All in all, I enjoyed learning to apply the knowledge acquired in the degree with the help of UniSim.
- Overall, the course was fine, sometimes some of the videos were not the most understandable and we got a little stuck.
- Sometimes watching the videos and doing what they do feels pointless.
- In my opinion, it was good that the course was via Zoom. I missed that we almost didn't see live demonstrations of you doing things and why.
- In my opinion, this course is the most fun and among the most useful in the degree. I would have preferred if had there been more time to learn the principles of the software and its capabilities in depth instead of focusing on the basics.

## 5. Conclusions

It is self-evident that students in the capstone design course would benefit from having acquired the necessary skills in operating a commercial process simulator in advance. The simulation lab offered at the Technion one semester ahead of the design course is one way of achieving this objective. This paper has described the current online format used in the course, its syllabus and schedule, and the online assessment methods used to verify mastery. The main conclusions are: (1) The level of support that can be tendered to students by course staff in an online setting using Zoom Breakout Rooms is at least as good if not better than that provided in regular face-to-face lab sessions; (2) Having students work on their own PCs with UniSim improved the level of individual student competence, as shown by the summative assessment evolution as well as the observed performance of the students in the subsequent capstone design course; (3) Online exams including practical usage of UniSim made a huge impact on the resulting competency of the average student: Students opined that the knowledge that their practical UniSim skills would be tested in the final exam and be required in the capstone design project were significant motivation for them to learn these skills during the semester; (4) We therefore recommend moving examinations online to allow for practical examination of competence in all courses that merit this feature (for example, in practical design of control systems using Matlab and Simulink for the course on process control).

## References

- T. A. Adams III, 2022, "Learn Aspen Plus in 24 Hours," 2<sup>nd</sup> Edition, McGraw-Hill, New York.
- D. J. Belton, 2016, "Teaching process simulation using video-enhanced and discovery/inquiry-based learning: Methodology and analysis within a theoretical framework for skill acquisition," *Education for Chemical Engineers*, 17, 54-64.
- K. D. Dahm, R. P. Hesketh, and M. J. Savelski, 2002, "Is process simulation used effectively in ChE courses?" *Chem. Eng. Educ.*, 191-198.
- D. R. Lewin, W. D. Seider, J. D. Seader, E. Dassau, J. Golbert, D. Goldberg, M. J. Fucci, E. Filiba, and R. B. Nathanson, 2003, "Using Process Simulators in the Chemical Engineering Curriculum – A Multimedia Guide for the Core Curriculum," Version 2.3, Multimedia CD-ROM, John Wiley, New York.
- D. R. Lewin, E. Dassau, and A. Goldis, 2006, "Effective Process Design Instruction: From Simulation to Plant Design," *16th European Symposium on Computer Aided Process Engineering*, 719-724.
- D. R. Lewin and A. Barzilai, 2021, "Teaching Process Design to Chemical Engineering Undergraduates – an Evolution," *Chem. Eng. Educ.*, 55(3), 157-172.
- D. R. Lewin, 2021, "What Can We Learn from Exam Grade Distributions?" *International Journal for the Scholarship of Teaching and Learning*, 15(2), Article 7.

# Designing an interactive environment to share educational materials and resources. Application to the Geomatics Hub at UniLaSalle Beauvais

Réjanne LE BIVIC<sup>a</sup>, Sébastien OTTAVI<sup>a</sup>, Pierre SAULET<sup>a</sup>, Pauline LOUIS<sup>a</sup>,  
Arnaud COUTU<sup>a</sup>

<sup>a</sup>*Institut Polytechnique UniLaSalle, 19 rue Pierre Waguet, 60000 BEAUVAIS, France  
arnaud.coutu@unilasalle.fr*

## Abstract

Teaching is a process requiring a logical and progressive frame, which is formalized in the pedagogic curriculum. Selecting, arranging, and framing the content is necessary to enhance the learning experience (Deng, 2011). The making of a curriculum implies the creation of materials that can be numerous. These pedagogic resources can be of different kinds: course materials, exercise instructions, or help tools. Using technology to provide easier access to these resources to students through an online platform is a great way to perpetuate these materials and ease the work of teachers (Nachmias et al., 2008). Furthermore, the development of such online media permits a valorization of the pedagogic resources and their innovations (Mahajan, 2010). It is a step towards an open science approach to teaching geomatics which goes further than actual textbooks (Bell et al., 2018). The focus of this paper is on the online platform created by the Institut Polytechnique UniLaSalle Beauvais for Geographic Information System (GIS) courses, known as the Geomatics Hub. The general public has access to over 75% of the content on this platform for free, while the remaining content is only available to students. 54 answers were collected from a survey to collect the experience of students on this online tool. 87% declared that this platform allows better access to pedagogic resources while 89% indicated that its organization is clear and logical. The Geomatics hub is also a way to display the results of projects led by either teachers or students. The use of an innovative solution with digital open access has a reach that is not limited to the students. The display of the education tools produced by UniLaSalle Beauvais for GIS can be noticed by companies, potential partners, or former students.

**Keywords:** Innovation diffusion; open science; e-Learning; Educational innovation

## 1. Introduction: teaching geomatics at UniLaSalle – Campus of Beauvais

Geomatics is an umbrella term for all disciplines related to collecting, analyzing, and storing geographic data. It gathers Geographic Information Systems (GIS), Remote Sensing, and Photogrammetry. The sensors and platforms are extremely varied, from ground-based GNSS to imaging satellites, or airborne LiDAR systems. The full Data Life Cycle is included, from the acquisition to the processing and communication of the results (Levitin and Redman, 1993). At UniLaSalle, the topics studied are mostly centered around Earth Observation, and more precisely Agriculture and Geology.

UniLaSalle Beauvais is an engineering school, that teaches about 3000 students in 5 years.

Geomatics is an important part of the curriculum, with an average of 10 % of the time spent learning GIS, Remote Sensing, and Photogrammetry, with an emphasis on the former (REN, 2022). Since the students have different career ambitions in Earth and Life sciences, being able to offer a wide range of real-life examples, especially for lab sessions, is a major way to keep them interested and therefore more engaged. One goal is to bring the students to work autonomously on labs that they choose from the online repository. Taking ownership of their learning experience by choosing exercises could increase the efficiency of the course (Kolb and Kolb, 2005).

Teaching involves creating a significant amount of materials, and although some teachers opt to make them accessible on their professional or personal website, some of these resources can still be lost. Making all materials available for the teachers of an organization could provide more time for others to improve their curriculum, without having to reproduce materials that already exist. It is also in line with the Open Access movement claims, which is that everyone would gain from making materials public, by avoiding duplicate work (Hylén, 2021).

Creating interactive online course materials was selected as the way to evolve teaching Geomatics at UniLaSalle. The Geomatics Hub was invented as a platform to gather all those materials and bring them beyond the classroom, in the light of the new paradigm of Education 4.0 which emphasizes experimentation as a way to learn, especially by taking advantage of emerging technologies (Almeida and Simoes, 2019).

## **2. Development of the Geomatics Hub**

### *2.1. The Geomatics Hub: a repository for course material*

The Geomatics hub was developed in 2019. It is freely accessible online at <https://hub-geomatique-unilasalle-fr.hub.arcgis.com/>. The main strategy is to provide students, as future technicians and engineers, with real-world examples and problem-based exercises to work on.

The original idea is to give more autonomy to students by making them work on a workflow-based approach, where they were provided data, some context, a general workflow, and access to a user manual created by the teacher. This allows the teacher to focus on students with more difficulties. Students can work in complete autonomy, by choosing a library of labs with varying levels of difficulty, and also check their skills autonomously using lists of technical skills.

The three main disciplines are easily available: GIS, Remote Sensing, and Photogrammetry. For GIS, the main parts are the lecture slides, about different topics related to GIS, the exercises as lab Sessions, and online help on three software: ArcGIS Pro, ArcGIS Online, and qGIS. Online skills trees also guide students to different parts of the user manual depending on the skill they want to practice.

For the Photogrammetry part, lecture slides are available as well as online help on the software Agisoft Metashape (proprietary). There is a similar structure for Remote Sensing: lecture slides, labs, and online help on ENVI (proprietary). In terms of web hosting, the user manuals are written and hosted on Scenarii, the lecture slides are in Genially, and the labs are on ESRI's servers in the form of Story Maps.

## *2.2. The Hub as a pedagogic tool*

The use of StoryMaps allows us to set levels for the lab sessions. For the GIS part, all lab sessions are taught as Problem-based-Learning, where a situation is described to the student, who then has to follow the steps of the processing tasks to complete the scenario. While in the first labs, the workflow is often given to the student, it is progressively less and less detailed, to let the student, improve and decide on their own which tools to use, with the end goal of writing and developing their workflows.

In Photogrammetry, the students are provided with a generic workflow and a detailed user manual, developed by the teaching team. The student then has to apply the processing to given sets of data, which can be changed year after year depending on the current need, for example when we want to student to work on new data.

The display tools chosen as support (StoryMaps, ESRI hub, Scenari) are easy to edit if the leading teacher is not in charge of the Hub anymore, as has been the case in this study. As for now, the platform is in French only.

## *2.3. The Geomatics hub is a showcase platform to enhance relations between companies and education*

The Geomatics Hub is in close relation with the GéoLab, an innovation center at UniLaSalle Beauvais that provide an access to students to work on innovative numeric tools for 3D modeling and 3D printing. This includes high-quality PCs, software licenses, technical support, and more. The Hub is a way to present current and past student projects. Beyond being a platform for students, it is also an opportunity to build bridges between the academic world and industry. By showcasing the researcher's and students' work, companies are more willing to offer internships or job offers or rent the services of the GéoLab for one-time projects, such as fossils and minerals 3D modeling, and bio-mimicry research work. By interacting both with students and professionals, the benefits are coming back to students because the teachers are on the front line and can estimate the needs of the companies, by being in direct communication with them (see Figure 1).



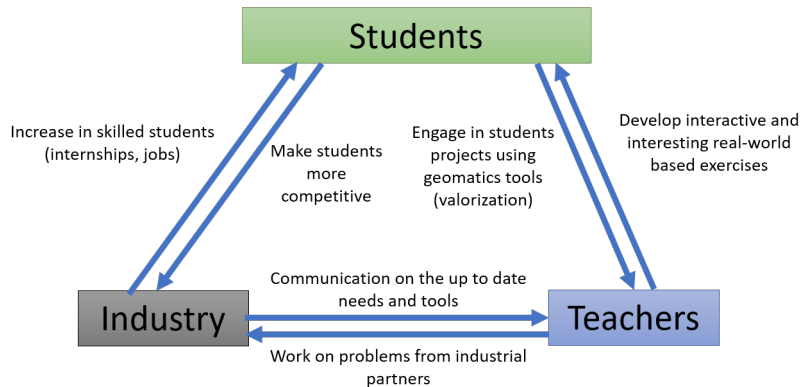


Figure 1 - Multi-parts benefits of Open Educational Resources with an emphasis on the communication between all three actors of learning: students, teachers, and the industry

### 3. Three years in evaluating the use of the platform

The main objectives of the Geomatics Hub are (1) to be an efficient tool for current students at UniLaSalle Beauvais, (2) to be useful for former students (3) to be used as a showcasing website for external users, whether they are individuals or companies. Two metrics are here studied to quantify the success of the platform. Firstly, we will evaluate the impact on a student through surveys. Secondly, a study of the hub engagement metrics will be used to quantify the impact of the website.

#### 3.1. Students survey

A study was conducted in September 2022, and 66 students answered out of the 406 students surveyed. 48.5% of the students who replied are in their third year at UniLaSalle. The questions were divided into three sections: The Geomatics Hub in general, the chronology of the lab Sessions, and the skills tree.

To the first question (“To which occasions are you using the Geomatics Hub?”), and besides a use during the class, 67% say they use the Hub when studying for the exam, 17% say they use it to know more about geomatics in general, 56% say they use it for projects outside of the class, and 15% say they are using it when working for a company (either during an apprenticeship, internships, or in their new position for the recently graduated students). When asked which resources the students were looking at the most, the first place comes for the course materials (73%), then the chronology of Exercises in ArcGIS (59%), the user manuals (55%), the library of exercises/examples (41%). About the use of the Hub outside of the classroom, a third of the student said they were using it regularly (33%), a third replied they were using it outside of the classroom but not regularly (35%) and another third replied that they were not using it outside of the classroom (31%). Three following questions are about the display and ease of use of the platform. A large majority, of students, find that the Hub is easy to use (90%), and would recommend it to someone who is a total beginner in GIS (94%). However, those results are put in perspective with some written comments by students who say that the interface is sometimes too busy and that there are too many steps to get to the course content.

The second part of the survey deals with the chronology of the exercises. Almost all the students say that the chronology is easy to understand and use (98%) and that it allows them to understand better the course in terms of skills to learn, by putting all the exercises in perspective (94%). The third part of the survey is about the skills list. While only half of the students had seen the skills list before (53%), a large majority of them think that it can be useful for them to learn GIS (85%) and that they would use it autonomously, without being expressively told to during a lecture (75%).

This first experience in surveying the students provided insights on the use of the platform but will also be used to improve the next surveys, especially in terms of focusing questions and getting the students more engaged.

### *3.2. Website engagement overview*

In the past year (Nov.2021 to Nov. 2022), over 15,000 visits were recorded on the Hub website, with an average of 1,400 visits per month over the past three months. Those metrics are limited but give a good quantitative indicator of the website's relevance and use.

## **4. Conclusion**

After three years of use of the platform, the feedback from students and teachers alike is in the majority positive. The students can work autonomously, during labs the teacher is in a support position. Students can follow their own pace and see the big picture of the skills expected from them at any time of their education. During the lifetime of the platform, teachers in Geomatics have partially changed. The existence of the platform made those changes easier and lessened the disruption for students. The school's industrial partners also gave positive feedback. ESRI France, the French branch of the leader in GIS and Remote Sensing software, broadcasted the Geomatics Hub in their magazine "SIGMag" in December 2020.

With the improvement in topics such as small objects photogrammetry, there is a need to update the tutorials. New topics related to Spatial Data Science should be included, like Machine Learning techniques, now broadly used in Spatial Data Science, including Geomatics. A part on data mining and data downloading should also be added, as well as reflecting on how to include the data acquisition part. A translation into English would also be a plus. Following the survey with students, it seems important to modify the architecture of the website to streamline access to the different parts of the materials. A focus should also be made on accessibility (option for contrast, image alternative text for example). One main risk taken with the architecture of the hub is that it relies heavily on external storage. The StoryMaps and main interface are hosted on ESRI's servers, the data is hosted internally on the school's platform, Claroline. The software manuals are hosted on Scenarii, and the slides are hosted on Genially and also archived on the teacher's computers. A way to bypass this would be to develop a server internal to the school, like an ESRI ArcGIS Enterprise solution.

In terms of pedagogy, more didactic tools should be tested, such as glossaries, quizzes, and serious games. The pandemic led to a chance for teachers to reexamine their teaching practices. A study from the University of Genoa showed that techniques such as Team-based learning, Problem-based learning, and tools such as Glossaries and quizzes are effective (Botto *et al.*, 2022). Our lab uses Problem-based learning and plans to build on the study's positive results. The use of free open-source software was also beneficial, as it's more accessible and flexible (Quinn, 2021). This is part of our future plans at the Geomatics Hub. While universities widely use Learning Management Systems, recent studies show a low use of the interactive tools (Al-Sharhan *et al.*, 2020). The Geomatics Hub will aim to utilize these past findings to overcome limitations and offer students a more engaging environment.

## References

- Almeida, F., Simoes, J., 2019. The role of serious games, gamification and industry 4.0 tools in the education 4.0 paradigm. *Contemp. Educ. Technol.* 10, 120–136.
- Al-Sharhan, S., Al-Hunaiyyan, A., Alhajri, R., Al-Huwail, N., 2020. Utilization of Learning Management System (LMS) Among Instructors and Students, in: Zakaria, Z., Ahmad, R. (Eds.), *Advances in Electronics Engineering, Lecture Notes in Electrical Engineering*. Springer, Singapore, pp. 15–23. [https://doi.org/10.1007/978-981-15-1289-6\\_2](https://doi.org/10.1007/978-981-15-1289-6_2)
- Bell, S., Ross, H.M., Epp, J., Bates, K., 2018. *Geomatics and Open Textbooks* (No. 2516–2314). EasyChair.
- Botto, M., Federici, B., Ferrando, I., Gagliolo, S., Sguerso, D., 2022. Innovations in geomatics teaching during the COVID-19 emergency. *Appl. Geomat.* 1–14.
- Deng, Z., 2011. Deng, Zongyi, "Revisiting Curriculum Potential," *Curriculum Inquiry*, 41 (December, 2011), 538-559.
- Hylén, J., 2021. Open educational resources: Opportunities and challenges.
- Kolb, A.Y., Kolb, D.A., 2005. Learning styles and learning spaces: Enhancing experiential learning in higher education. *Acad. Manag. Learn. Educ.* 4, 193–212.
- Levitin, A.V., Redman, T.C., 1993. A model of the data (life) cycles with application to quality. *Inf. Softw. Technol.* 35, 217–223. [https://doi.org/10.1016/0950-5849\(93\)90069-F](https://doi.org/10.1016/0950-5849(93)90069-F)
- Mahajan, V., 2010. *Innovation diffusion*. Wiley Int. Encycl. Mark.
- Nachmias, R., Mioduser, D., Forkosh-Baruch, A., 2008. Innovative pedagogical practices using technology: The curriculum perspective, in: *International Handbook of Information Technology in Primary and Secondary Education*. Springer, pp. 163–179.
- Quinn, S., 2021. Using Free and Open Source Software to Teach University GIS Courses Online: Lessons Learned During a Pandemic. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*
- REN, F., 2022. *L'enseignement des Systèmes d'Informations Géographiques à UniLaSalle Beauvais*. UniLaSalle Beauvais.

# Immersive learning through simulation: implementing twin screw extrusion in Unity

Pedro Santos Bartolomé<sup>a</sup>, Daniel Just<sup>b</sup>, Ariana Bampouli<sup>a</sup>, Simon Kemmerling<sup>b</sup>,  
Aleksandra Buczko<sup>b</sup>, Tom Van Gerven<sup>a</sup>

<sup>a</sup>*Process Engineering for Sustainable Systems (ProcESS), KU Leuven. Celestijnenlaan  
200F, 3001, Leuven, Belgium*

<sup>b</sup>*Fraunhofer-Institut für Chemische Technologie ICT, Polymer Engineering. Joseph-  
von-Fraunhofer-Str. 7, 76327 Pfinztal, Germany*

## Abstract

The use of immersive environments for education is a topic that attracts growing interest, particularly within process engineering. The engines for the design of immersive environments are ill-suited for process simulation (as is the case with Unity), leading to one of three main options: either discarding the use of these tools, with the consequent discarding as well of the easy integration in many platforms and graphical and physics engines; setting an interconnection with external simulation engines, with the consequent limitations; or focusing on a simulation-free implementation altogether.

In this work, we demonstrate instead the implementation of a twin-screw extrusion simulation within Unity by benefitting from the available C# optimization library *alglib* and using complementarity constraints, in contrast with the previous literature. We study the computational time demanded and compare the results with simulations carried out with *LUDOVIC*, a state of the art simulator for twin-screw extrusion. The obtained results show that this alternative provides solid performance while simplifying the deployment of the final tool for its use by students.

**Keywords:** Unity, Twin Screw Extrusion, Educational Simulation, Complementarity Constraints.

## 1. Introduction

Twin screw extrusion is a popular process in the polymer, chemical and food industries, where an intermeshing extruder (see Fig.1) is used for mixing highly viscous materials, favored for its good mixing performance and short residence times, amongst other properties (Eitzlmayr et al., 2014). However, it is also a complex process: the operating conditions mean that most properties, such as the filling ratio or the pressure, have little identifiability in practice; the screw must be designed specifically in different sections, taking into account the geometry, length and rotating speed, a design which is meant to compensate for the complex physical phenomena of mass and heat transport, as well as the temperature rise due to viscous dissipation; all of these effects come into play in the operation of the process. Finally, it is also complex to formulate and solve, as the mass balance results in a constrained Ordinary Differential Equation (ODE).

In view of the complexity of the process, there is a clear incentive to develop training tools for both operators and engineers which would allow them to familiarize themselves with the different elements and mechanics of the process. There are many alternatives for this: the first is simulation-free environments, where the interaction of each element is

hand-coded by the designer: this has the advantage of requiring little time for simple cases, but has the disadvantage of exponential complexity that arises as a new element added has to interact with all the other previously designed ones. Alternatives include connecting with available simulators, such as the authors previously discussed in (Santos Bartolome and Van Gerven, 2022), but this alternative was not available in the field of extrusion simulation. Finally it can be opted to develop the tool in a different software, not intended for environment design, such as a user interface in python: the caveat for our case is that this offers limited functionality and realism.

Instead, we have opted to implement a simulation of a twin extruder directly in Unity, by benefiting from available libraries in C#, particularly `alglib` (ALGLIB, 2022). This allows for great flexibility in the design of the environment, at a cost of rigidity in the simulation software itself. The prototype tool has been developed to demonstrate the feasibility of this approach, and in it we have developed the model presented in Figure 1, which consists of 49 compartments, and for which data has also been generated with a rigorous polymer extrusion software (Vergnes et al., 1998).

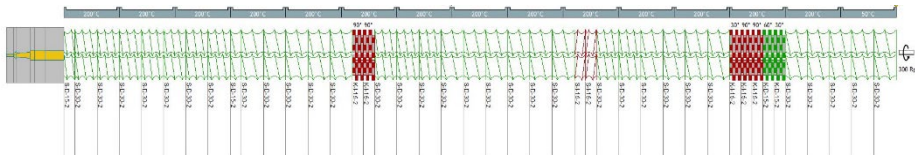


Figure 1 : Twin extruder design used in the simulation.

We will introduce the model in the following section, after which we expand on the technical implementation into Unity. Section 4 shows the results obtained, followed by conclusions and proposals of future work.

## 2. Twin screw extruder modelling

The twin screw extrusion system can be rigorously modeled with CFD simulation, however, this results in a great computational demand, which makes its use unadvisable for an educational tool (Solmaz and Van Gerven, 2022). Instead, we will use the 1D model developed by (Eitzlmayr et al., 2014), which provides an extensive study of the effects of each physical phenomenon, based on the properties of the geometry and the chosen polymer. This model subdivides the screw elements into sections which are considered perfectly mixed, with homogeneous properties.

Due to its complexity we will not extend on every term of the model here, and instead will focus on the main elements of the mass and energy balance. We take  $f_m^{(i)}(u, x)$ , to provide the derivative of the mass of the  $i^{th}$  screw element, and  $f_Q^{(i)}(u, x)$ , to provide the derivative of its heat. Both are functions of  $u$ , the process inputs, and  $x$ , the internal states of the process, resulting in:

$$f_m^{(i)}(u, x) = \sum_{j=1}^k \dot{m}_{conv}^{(k \rightarrow i)} + \dot{m}_p^{(k \rightarrow i)}$$

Where  $k$  is the index of the last screw element,  $\dot{m}_{conv}^{(k \rightarrow i)}$  is the flow by convection from the  $k^{th}$  element to the  $i^{th}$ , and  $\dot{m}_p^{(k \rightarrow i)}$  is the pressure flow from the same.

The heat balance is instead composed of:

$$f_Q^{(i)}(u, x) = \sum_{j=1}^k C_p^{(j)} T^{(j)} (\dot{m}_{conv}^{(k \rightarrow i)} + \dot{m}_p^{(k \rightarrow i)}) + \alpha^{(i)} A_b^{(i)} (T_b^{(i)} - T^{(i)}) + Q_{diss}^{(i)}$$

Where  $C_p^{(i)}$  is the heat capacity at the  $i^{th}$  element,  $T^{(i)}$  it's temperature,  $\alpha^{(i)}$  the heat transfer coefficient with the barrel, which is assumed to be constant,  $A_b^{(i)}$  the surface area of the barrel segment,  $T_b^{(i)}$  the temperature of the barrel, and  $Q_{diss}^{(i)}$  the heat generated by viscous dissipation, which, due to the complexity of full viscous dissipation terms, we simplify significantly to:

$$Q_{diss}^{(i)} = (K_{diss}^{(i)} n)^2 \eta^{(i)}$$

Once the model has been chosen, we have designed an optimization problem to solve the mass balance, and a linear equation for the energy balance, which are solved sequentially to obtain the mass balance and energy balance. This sequential approach was chosen due to the already significant time required to solve the mass balance by the chosen solver, which we preferred to not prolong by increasing the number of variables. The balances are connected due to the changes in property values, but we consider this changes not significant on sufficiently small step sizes (when calculating steady state results, five iterations are done to arrive to convergence).

### 2.1. Mass (and Pressure) Balance with Complementarity Constraints

The greatest challenge when calculating the mass balance of the twin extruder is the condition that the pressure cannot rise higher than the atmospheric pressure as long as the compartment is not filled completely. This can be expressed as a constrained ODE in the form:

$$\begin{aligned} \frac{dm^{(i)}}{dt} &= f_m^{(i)}(u, x) \quad \forall i \\ 0 \leq m^{(i)} &\leq m_{max} \quad \forall i ; P^{(i)} \geq P_{atm} \quad \forall i \\ P^{(i)} &= P_{atm} \quad \forall i \in \{i | m^{(i)} \neq m_{max}\} \end{aligned}$$

In the literature we have found approaches to find the solution to these balances by using iterative calculation (Vergnes et al., 1998), by iteratively classifying compartments as filled or not filled (Choulak et al., 2004), and by introducing an additional differential equation for the pressure (Cegla and Engell, 2021). Some of the limitations of these approaches are, respectively: not using the information available due to ignoring the gradients; required small time steps or use of more complex integration schemes; and need of tuning of an additional parameter.

Instead, in this work we propose to tackle this process based complementarity constraints, as described in (Biegler, 2010). These are paired constraints for which one of them must always be active, which is the case in our system. This turns the equations to:

$$\begin{aligned} \frac{dm^{(i)}}{dt} &= f_m^{(i)}(u, x) \quad \forall i \\ 0 &\leq m^{(i)} \quad \forall i \\ m^{(i)} \leq m_{max} &\perp P^{(i)} \geq P_{atm} \quad \forall i \end{aligned}$$

This equation can then be solved as an NLP with the approach listed in (Biegler, 2010) as RegComp:

$$\begin{aligned} & \min_m \varepsilon \\ & \frac{dm^{(i)}}{dt} = f_m^{(i)}(u, x) \quad \forall i \\ & 0 \leq m^{(i)} \leq m_{max} \quad \forall i \\ & P^{(i)} \geq P_{atm} \quad \forall i \\ & (-m^{(i)} + m_{max})(P^{(i)} - P_{atm}) \leq \varepsilon \quad \forall i \end{aligned}$$

We can observe that both terms on the last constraint will be positive in the feasible region of the rest of the problem, and  $\varepsilon$  reaches its minimum value of 0 when at least one of the constraints is active for all compartments, meaning that they are each either full or are at atmospheric pressure.

### 2.2. Heat Balance

Once the mass balance is solved, the heat balance is comparatively simpler, as there are no restrictions necessary and instead is solved only as an ODE:

$$\frac{dQ^{(i)}}{dt} = f_Q^{(i)}(u, x) \quad \forall i$$

Where some of the states in  $x$  have been obtained in the previous step.

## 3. Implementation in Unity

After the design of the model we implement the simulation into the game engine Unity, which allows for the creation of an interactive educational environment. We give a summary in this section:

### 3.1. Optimization

The optimization is conducted with the SQL solver included in (ALGLIB, 2022) within their augmented Lagrangian (minnlc) C# package, which was found to be the best performant of all available algorithms. Restraints are added as boxed, linear or non-linear, in that order of preference. The heat balance is solved with the matrix solver `matrixsolve`. All derivatives are implemented with the backwards Euler method.

The library can be added directly in the assets folder of the project with no further modifications. Although multiple warnings are displayed in the editor regarding duplicated namespaces, we found no issues either in the editor nor when compiling or playing a compiled version.

### 3.2. Game elements

Each compartment is integrated as a game object, whose properties such as length  $\Delta x$ , diameter  $D$  or barrel temperature  $T_b$  can be changed directly, whereas the simulation is centralized into an empty object which contains a list of all active compartments, as well as all central variables such as the time step  $\Delta t$  and the rotational speed  $n$ .

Plots are added to represent the evolution of the variables either with regards to the position in the extruder or with time. These are added from the Chart and Graph utility (Graph And Chart, 2022), and edited by the main simulation object upon each new result.

#### 4. Results

We have collected both the computational time needed to solve the system and the accuracy of the steady-state results, compared with results obtained in the state of the art twin extrusion simulator Ludovic (Vergnes et al., 1998). Since the parameters dependent on screw geometry were not available, they have been fitted by comparing with the Ludovic results. All compartments not part of the extruder die have been considered to have the same geometry to avoid overfitting. Due to the results only needing to be generally correct for learning, the parameters were tuned by visual comparison instead of using more sophisticated methods such as data reconciliation (Pitarch et al., 2019).

All data are collected within the Unity 2021.3.4f1 editor in Windows 10 Enterprise, with a 11th Gen Intel(R) Core(TM) i7-11850H at 2.50 GHz and 16 GB of RAM.

##### 4.1. Computational time demanded

We have observed that compartments where the pressure goes above atmospheric are the most computationally consuming, which led to the use as a comparison a series of mixing elements with a single forward conveying element in the first location. Figure 2 shows the evolution of the average computational time demanded with the number of elements, which shows that the time demand is prohibitive for real-time simulation of the extruder design proposed (which contains 49 compartments). However, it is still possible to use the system to show students the basic effects in smaller, less representative designs. Further work on the solver settings might lead to a significant improvement.

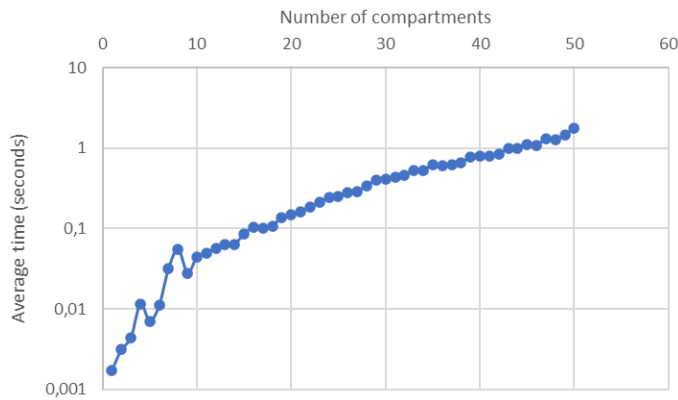


Figure 2. Average time consumed by solving the mass balance in Unity (logarithmic scale) against number of compartments.

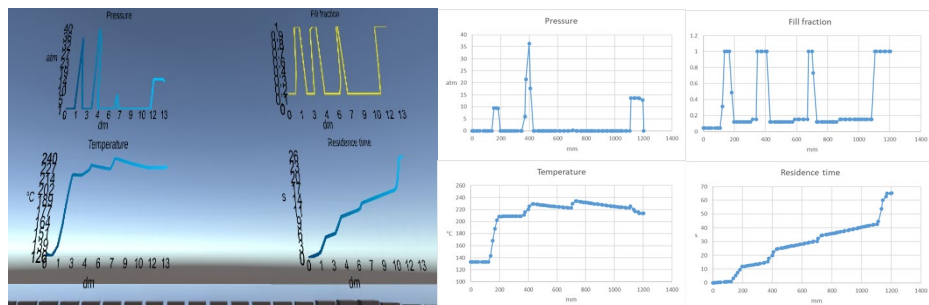


Figure 3. Comparison of results in Unity (left) and Ludovic (right). The game objects representing the extruder compartments can be seen in the bottom left as default black boxes.



#### 4.2. Comparison of results

Figure 3 shows the comparison of the results obtained in Ludovic and in Unity for temperature, pressure, filled fraction and residence time. Although there are differences of magnitude, which we attribute to the limited ability to tune the parameters, it appears that the profiles are sufficiently similar for the students to learn the effects and interactions of different screw designs and process parameters.

### 5. Conclusions

The work has satisfactorily dealt with the simulation of a twin extruder in an educational environment designed in Unity, showing that the available tools are sufficient to consider direct development of the model in the engine as a valid alternative for design. Due to the computational load, the prototype designed is of limited use for dynamic simulation of systems with a large number of compartments, but is otherwise useful for steady state simulation, and for dynamic simulation of simpler systems, as the profiles obtained have been shown to be similar to those obtained by state of the art tools. We believe that these educational environments can be used as tools for the benefit of the learning of students and trainees of all kinds.

Besides this main goal, we have also introduced the formulation of the twin extruder as an ODE with complementarity constraints, as opposed to the previous approaches present in the literature. It remains necessary to study how this approach compares in practice with the already existing ones, both in stability and time demand, both for pure modelling and for other objectives such as control.

### Acknowledgements

This project has received funding from the European Union's EU Framework Programme for Research and Innovation Horizon 2020 under Grant Agreement 812716 (SIMPLIFY). Project website: [www.spire2030.eu/simplify](http://www.spire2030.eu/simplify).

### References

- ALGLIB, 2022. ALGLIB - C++/C# numerical analysis library. <https://www.alglib.net>
- Biegler, L.T., 2010. Nonlinear Programming: Concepts, Algorithms, and Applications to Chemical Processes. Chapter 11.
- Cegla, M., Engell, S., 2021. Application of Model Predictive Control to the reactive extrusion of e-Caprolactone in a twin-screw extruder. IFAC-Pap., 16th IFAC Symposium on Advanced Control of Chemical Processes 2021 54, 225–230.
- Choulak, S. et al., 2004. Generic Dynamic Model for Simulation and Control of Reactive Extrusion. *Ind. Eng. Chem. Res.* 43,
- Eitzlmayr et al., 2014. Mechanistic modeling of modular co-rotating twin-screw extruders. *Int. J. Pharm.* 474, 157–176.
- Graph And Chart, 2022. Graph And Chart | GUI Tools <https://assetstore.unity.com/packages/tools/gui/graph-and-chart-78488>.
- Pitarch, J.L., Sala, A., de Prada, C., 2019b. A Systematic Grey-Box Modeling Methodology via Data Reconciliation and SOS Constrained Regression. *Processes* 7, 170.
- Santos, P., Van Gerven, T., 2020. Aspen Hysys – Unity Interconnection. An Approach for Rigorous Computer- Based Chemical Engineering Training, CACE, 30. pp. 2053–2058.
- Solmaz, S., Van Gerven, T., 2022. Automated integration of extract-based CFD results with AR/VR in engineering education for practitioners. *Multimed. Tools Appl.* 81, 14869–14891.
- Vergnes, B., Valle, G.D., Delamare, L., 1998. A global computer software for polymer flows in corotating twin screw extruders. *Polym. Eng. Sci.* 38, 1781–1792.

# An Educational Workshop for Effective PSE Course Development

Daniel R. Lewin,<sup>a</sup> Edwin Zondervan,<sup>b</sup> Meik Franke,<sup>b</sup> Anton A. Kiss,<sup>c</sup> Stefan Krämer,<sup>d</sup> Mar Pérez-Fortes,<sup>e</sup> Artur M. Schweidtmann,<sup>c</sup> Petronella M. (Ellen) Slegers,<sup>f</sup> Ana Somoza-Tornos,<sup>c</sup> Pieter L.J. Swinkels,<sup>c</sup> and Bart Wentink<sup>g</sup>

<sup>a</sup> *Department of Chemical Engineering, Technion. I. I. T., Haifa 32000, Israel*

<sup>b</sup> *Department of Chemical Engineering, University of Twente, the Netherlands*

<sup>c</sup> *Department of Chemical Engineering, TU Delft, the Netherlands*

<sup>d</sup> *Bayer AG, Engineering & Technology, Leverkusen, Germany*

<sup>e</sup> *Engineering, Systems and Services, TU Delft, the Netherlands*

<sup>f</sup> *Operations Research and Logistics, Wageningen University & Research, the Netherlands*

<sup>g</sup> *BASF SE, Corporate Technology, Ludwigshafen, Germany*

Email for correspondence: [dlewin@technion.ac.il](mailto:dlewin@technion.ac.il)

## Abstract

An educational workshop for developing Process Systems Engineering (PSE) courses will be held during ESCAPE-33, following the model workshop that was run during the CAPE Forum 2022 held at the University of Twente, in the Netherlands. This 3-hour workshop distributes the participants into four teams working together to develop the outline of a course on a novel application area in PSE motivated by a selected plenary or keynote talk at the conference, with each team led by authors of this contribution. This paper provides an overview of the approach used in the workshop for the effective development of a PSE course.

**Keywords:** PSE education, course development, learning objectives.

## 1. Introduction

From 14-16 September 2022 a workshop on developing Process Systems Engineering (PSE) courses was held at the CAPE Forum 2022 at the University of Twente in the Netherlands ([www.pse-nl.com/index.php/cape-forum-2022](http://www.pse-nl.com/index.php/cape-forum-2022)). Working in teams, participants of the forum with different backgrounds developed education plans for a complete course on one of four selected subjects: Carbon-neutral PSE, Sustainable bio-based PSE, Energy efficient PSE, and Artificial Intelligence (AI) in PSE. These are representative novel application areas described by Lewin et al. (2022) as suitable emerging fields that offer new horizons to students from the basic areas of process design, process control and numerical methods, and, of course, rely heavily on them as foundations. These topics are also fully in line with the recently described PSE developments in industry and academia (Kiss and Grievink, 2020). The specific topics selected in CAPE Forum 2022 were inspired by the four keynote talks delivered at the meeting and the team leaders for the course development were the four keynote speakers from industry and academia (Bart Wentink, Ellen Slegers, Stefan Krämer, and Artur Schweidtmann).

The workshop was held in two sessions. In the first session, each team worked on developing learning objectives and a syllabus for the entire course. They also proposed more detailed learning objectives and content for at least one week of activity of the course. In the second session, each team developed suitable activities for students for each of the weeks of activity specified in the first session. Each of the courses developed relied on the *flipped classroom* paradigm (Lewin and Barzilai, 2022; Lewin, 2022), in which every week of class activity is divided into three sequential phases: (1) *preparation*, in which students prepare for the week's activities by completing online lessons consisting of a series of short video clips and associated quiz questions that test comprehension; (2) *class meeting*, in which students improve their understanding, comprehension, and capabilities by tackling open-ended exercises and discussion mentored by the teacher, and (3) *active tutorial*, in which students solve exercises on their own or in small groups with their peers, mentored by the course staff. For more details, the reader is referred to the referenced papers, and the YouTube video clips mentioned in Section 3.

## 2. Course Development Steps

A systematic framework for course development follows the following top-down approach:

**Step 1:** At the highest level, one needs to answer the question: “*What do I want my students to learn and why?*” In other words, to specify what skills and attributes students need be able to demonstrate on successfully completing the proposed course. These are best defined in terms of learning objectives. For example, for a typical 13-week course on *Process Dynamics and Control* (equivalent to 5 ECTS), the course learning objectives would be divided into those concerning dynamic modelling and those concerning control design, defining what a student needs to exhibit on successfully completing the course:

Process modelling (Weeks 1-6):

L01: Capability to model dynamic response of processing systems using material and energy balances

L02: Ability to linearize models and generate transfer function approximations

L03: Ability to generate overall transfer functions from block diagrams

L04: Ability to analyse the transient response of linear systems

Process control synthesis (Weeks 7-13):

L05: Frequency domain analysis capabilities, enabling transition between Laplace, frequency, and time domains

L06: Stability analysis capability

L07: Knowing how to tune PID controllers effectively

L08: Capability to synthesize feedback control systems to meet closed-loop performance specifications

L09: Capability to design effective cascade and feedforward control systems

**Step 2:** Identify the best teaching approaches for the type of learning you want. For efficient learning, we recommend active learning strategies. This implies freeing some or all of the contact student-staff time to allow students to get involved in staff-assisted problem-solving. An established way to do this is by *flipping the classroom* (Lewin, 2022), which moves some or all of the lecture

material to student online self-study activities. An alternative approach would be to reduce the volume of lecture material, and indeed, any approach that frees some student-staff contact time for practice is appropriate.

**Step 3:** Select the course content, organized as a weekly schedule, in which each week is driven by its learning objectives. For example, for the process control course, the weekly schedule could be as shown in Figure 1. Therein, specific learning objectives are defined for each week. For example, in Week 10, the learning objectives would concern the Root Locus method:

On completing week 10 of the process control course, the student should be able to:

- Sketch the root locus, given transfer functions for a process and its controller,
- Sketch the output response of a controlled process, given locations of closed-loop poles on the root locus, and
- Use the root locus to design a controller, given the transfer function of the process to be controlled, to meet the desired closed-loop response specifications (e.g., overshoot and settling time).

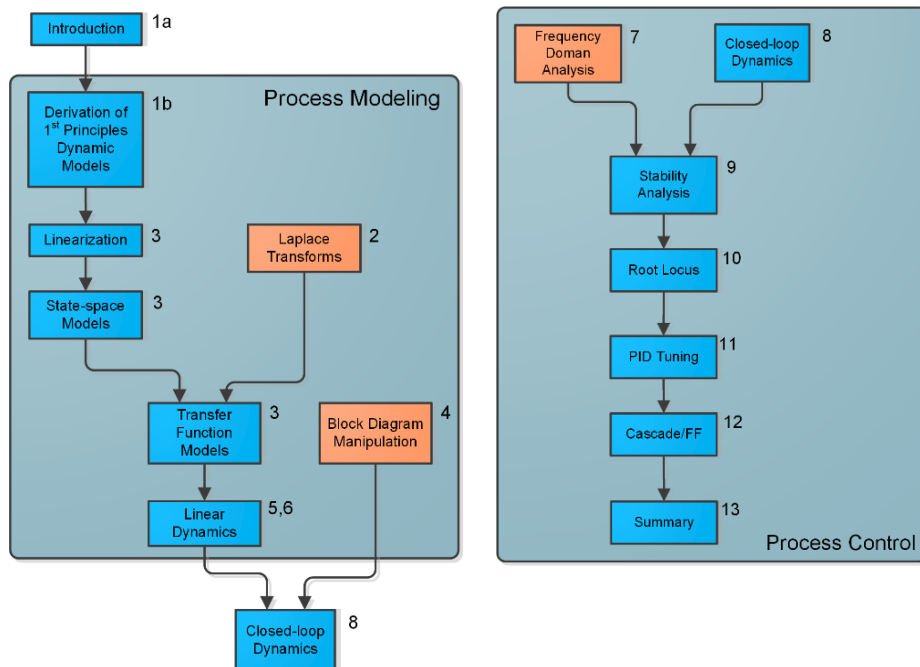


Figure 1. Weekly schedule for a course on process control; numbers in the diagram refer to weeks in the course.

**Step 4:** For each week, develop student assignments, projects, and class activities so that all the learning objectives are practiced by the students, with repeats as necessary. This is by far the most important content of the course, as it defines the mechanisms by which students will “learn by doing.” Some of the activities should be designed to be guided activities (e.g., open-ended problem solving in class meetings, guided by the instructor), which other activities are designed to be for students working alone or in groups. An example exercise for Week 10

of the process control course that tests accomplishment of all three learning objectives would be the following:

$$\text{Given the open loop transfer function: } PC(s) = \frac{K_C (s+2)(s+4)}{s(s+1)(s+5)(s+10)}$$

- (a) Draw the root locus for  $PC(s)$ .
- (b) Determine the value of  $K_C$  that will give the fastest possible closed-loop response with no oscillations.
- (c) Sketch (qualitatively – there is no need to compute the response) the closed-loop response to a unit step change in the setpoint, with the value of  $K_C$  as determined in part (b).

**Step 5:** No work on course development is ever final. One should always make an effort to improve the course plan with class experience and feedback.

### 3. Outline and outcomes of the CAPE Forum 2022 Workshop

#### 3.1. Workshop activity

In the spirit of “flipping” the course, the Forum participants were invited to review the short videos describing how the flipped class is used to teach PSE subjects before attending the CAPE Forum:

Webinar talk (14 min): [www.youtube.com/watch?v=O3hoSIYaGo4](http://www.youtube.com/watch?v=O3hoSIYaGo4)

ESCAPE-31 Keynote (28 min): [www.youtube.com/watch?v=b6w6mqSPxp0](http://www.youtube.com/watch?v=b6w6mqSPxp0)

The workshop was organized with three points of contact during CAPE Forum 2022. Splitting the workshop into three points of contact gave the opportunity for group members to get together informally and discuss course content during the entire forum. The three points of contact were:

1. 15<sup>th</sup> September 2022, morning – Participants were distributed into one of the four topics by a suitable color code on their badges. A 5-min announcement introduced the workshop. Each of the four teams was asked to prepare an outline for a new course in the following subjects:  
 Topic 1: Carbon-neutral PSE, led by Bart Wentink and Leyla Özkan  
 Topic 2: Sustainable biobased PSE, led by Ellen Slegers and Meik Franke  
 Topic 3: Energy efficient PSE, led by Stefan Krämer and Pieter Swinkels  
 Topic 4: Artificial intelligence (AI) for PSE, led by Artur Schweidtmann and Ana Somoza-Tornos
2. 15<sup>th</sup> September 2022, afternoon – First 90-minute session. The workshop presenter introduced the activity and initiated a hands-on session in which the four groups worked on developing each of the courses (**Steps 1 and 3**), led by one of the keynote speakers. This session was positioned after the last of the four keynotes.
3. 16<sup>th</sup> September 2022, afternoon – Second 90-minute session. The workshop presenter introduced a hands-on session in which the four groups worked on **Step 4**, developing activities for students appropriate for at least one of the detailed weeks planned in **Step 3** for the first hour, each group being led by one of the keynote speakers. In the last 30 minutes, the groups presented their course

plan outlines to all participants and received immediate feedback from the instructor and the group.

### 3.2. Example course layout

As an example of the outcomes that can be expected from the course development workshop, we present here in detail the course outline for Topic 4, on “AI in PSE,” as completed by a team of eight participants, headed by the keynote speaker, Artur Schweidtmann. Similar outcomes were achieved by the other three teams. This section details the course learning objective, and syllabus and expands on the planned activities for Week 2 of the course.

#### 3.2.1. Topic 4: Course learning outcomes and schedule

The team working on Topic 4 developed the outline for a master’s level elective course of 4-5 ECTS on “AI for PSE.” Students of the course need to have chemical engineering basics, matrix algebra, calculus, and an introductory course in statistics as prerequisites. It would be advantageous to also have programming experience in Python and have taken an introductory course in optimization. At the end of the course, the student will be able to:

- LO1: Explain the concepts of AI and data analysis
- LO2: Discuss the potential and limitations of different AI methods/tools for given chemical engineering applications
- LO3: Design an AI system to solve a given chemical engineering problem
- LO4: Implement a given or designed AI system using Python
- LO5: Validate an AI model based on data
- LO6: Comprehend state-of-the-art AI research

The course schedule as devised by the team was divided into two approximately equal parts, the first covering the foundations and the second covering applications and project work:

- Week 1: Introduction to Python
- Weeks 2-6: AI foundations
  - Week 2: Data analysis
  - Weeks 3-4: Regression
  - Weeks 5-6: Classification
- Weeks 7-13: Applications of AI in PSE
  - Week 7: Molecular property prediction
  - Week 8: Hybrid modelling
  - Week 9: Demand forecasting
  - ...
  - Week 13: Project presentations and assessment

#### 3.2.2. Topic 4: Course plan for Week 2

Team 4 also produced a detailed plan for Week 2, covering data analysis methods. The week’s activity is driven by four learning objectives that indicate that after successfully completing week 2, the student will be able to:

- LO1: Select an appropriate regression method for a given problem
- LO2: Train a regression model on data using common packages
- LO3: Validate a regression model based on data

- LO4: Discuss advantages and disadvantages of common regression tools for a given application

These LOs are achieved by a combination of two phases of self-learning supported by online videos and quiz questions, each followed by in-person sessions involving online quizzes, class discussions, and programming activities in the lab. As shown in Figure 2, the activities have multiple connections to all four learning objectives.

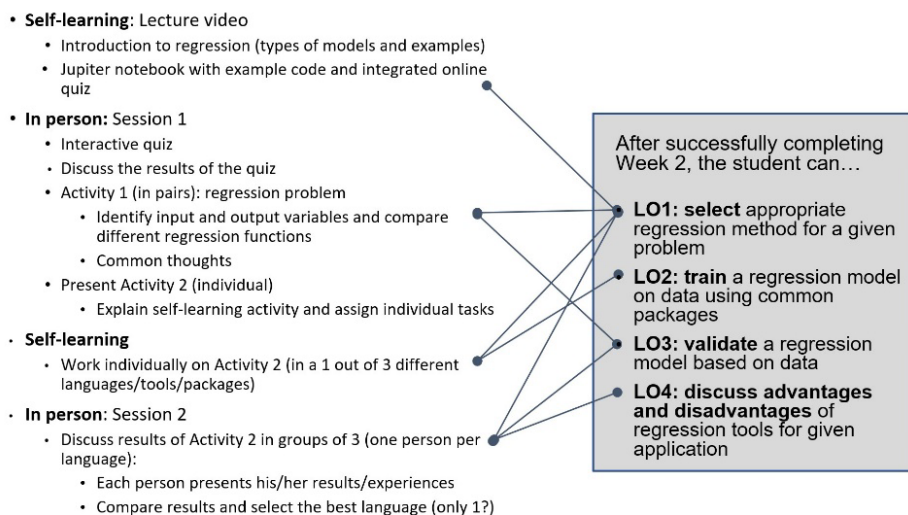


Figure 2. Mapping activities in Week 2 to learning objectives of “AI for PSE”.

## 4. Conclusions

The outcomes of the CAPE Forum 2022 workshop were four master plans, one for each of the topics discussed in the keynote talks (i.e. Carbon-neutral PSE, Sustainable biobased PSE, Energy efficient PSE, and AI for PSE). Each master plan was characterized by a syllabus and a week-by-week schedule, of which most were linked to the learning objectives. In addition, the teams devised exercises, assignments, and projects to be carried out by students of the courses, which address the same learning objectives. The workshop gave a useful course development experience to junior faculty, PhD and EngD students. Following these successful outcomes, a similar educational workshop will be offered as part of the ESCAPE-33 conference program.

## References

- A. A. Kiss, J. Grievink, 2020, “Process Systems Engineering developments in Europe from an industrial and academic perspective”, *Computers & Chemical Engineering*, 138, 106823.
- D. R. Lewin and A. Barzilai, 2022, “The Flip Side of Teaching Process Design and Process Control to Chemical Engineering Undergraduates – and Completely Online to Boot,” *Education for Chemical Engineers*, 39, 44-57.
- D. R. Lewin, 2022, “Teaching PSE Mastery During, and After, the COVID-19 Pandemic,” *Comput. Chem. Eng.*, 160, 107741.
- D. R. Lewin, E. M. Kondili, I. T. Cameron, G. Léonard, S. Soheil Mansouri, F. G. Martins, L. Ricardez-Sandoval, H. Sugiyama, and E. Zondervan, 2022, “Agile Process Systems Engineering Education: What to teach, and how to teach it,” submitted to *Computers and Chemical Engineering*.

# Teaching strategies for the effective use of computational tools within the Chemical Engineering curriculum

A. Tsatse, E. Sorensen

*Department of Chemical Engineering, University College London, Torrington Place,  
WC1E 7JE London, UK*

## Abstract

This work discusses the rationale behind, and strategies followed, when delivering chemical engineering modules which have a significant requirement of computational work, and how to coordinate this delivery across an entire curriculum. It will be shown how the complicated aspects of PSE-related subjects can be considered and introduced to students through several modules spread across UCL Chemical Engineering Year 2 programme, and how student learning is supported, whilst providing them with the best opportunity for applying their new computational knowledge and skills and for innovation. It will also be discussed how this procedure prepares students for their capstone design project in Year 3, indicating that this approach has several benefits, including but not limited to, students' understanding of PSE tools and the development of their critical engineering thinking.

**Keywords:** Chemical Engineering, Process Systems Engineering, problem-based learning, computational tools

## 1. Introduction

UCL Engineering's Integrated Engineering Programme (IEP) was introduced in 2014/15 and is based on innovative teaching practices which include the use of scenario- and problem-based learning (Graham, 2018) across Years 1 and 2 of the programme. Within the Department of Chemical Engineering, several IEP teaching activities have strong computational elements (Tsatse and Sorensen, 2021). Most of the computational tools are introduced in the Year 2 curriculum, which is the focus of this work, as this year includes various modules that make use of computational tools (see Figure 1). The Department employs a wide range of techniques to ensure that the Year 2 students know how to use a wide range of computational tools; that they have the required knowledge of how to select the most suitable software based on the given problem; and for them to reflect on their work before proceeding to the capstone design project in Year 3 of the programme. In addition, teaching staff needs to regularly evaluate the material delivered, to ensure that the concepts and topics discussed align with current research directions and the constantly evolving industrial needs (e.g. integration of multiple computational tools).

In all modules that will be discussed below, assessed projects and team reports evaluate whether the learning outcomes are met, whilst tutorials for the relevant computational tool(s) take place before the release of the material for which the software is required (e.g. project, coursework etc.). In addition, computational help desk sessions are scheduled before the respective assessment deadline so that support by staff is provided to students



who may be facing difficulties with the specific software used. As this work focuses on Year 2 modules, it is useful to note that, in Year 1, students are made familiar with the fundamentals of GAMS (GAMS, 2022) and gPROMS (Siemens Process Systems Engineering, 2022), and are thus able to solve simple computational problems such as how to solve a mass balance around a buffer tank and calculate the respective liquid height. Year 1 does not include any teaching material on AspenPlus (Aspentech 2022), so this is an entirely new computational tool for the students when starting Year 2.

During their studies, students have continuous access to online learning material for all the computational tools used across the entire curriculum (i.e. AspenPlus, gPROMS, GAMS and MATLAB). The material for each of the tools includes how-to videos and step-by-step presentations that explain various concepts and processes, and how these can be implemented and solved using the respective tool. The material is updated regularly to make sure it aligns with the theory taught, or any new elements added, as well as updated versions of the software. For instance, material on how to consider distributed models in gPROMS, and relevant examples thereof, is available for students to download, study and apply in their own time. Relevant material is available online also for aspects that are not formally considered or assessed as part of any of the taught modules, such as solids processing, so that students can refer to this if they need to consider these aspects as part of an open-end project, for instance in a Scenario (Sorensen, 2016). This supporting material provides additional resources which students appreciate, especially when they proceed to more challenging projects in their 3<sup>rd</sup> or 4<sup>th</sup> year, when they need to consider processes not previously covered in individual taught modules.

Each of the following sections discusses the rationale and teaching strategy for the use of computational tools within our programme. The case studies discussed are presented in order of increasing difficulty for the students, and the discussion includes the strategy for how the students learn how theory and experimental data can be combined, as well as how a proposed process synthesis can be verified using computational tools. In addition, the strategy for how to include computational work in open-ended problem-based learning in the form of a week-long IEP project (i.e. Scenarios) is also presented, along with a discussion of how students are prepared for their 3<sup>rd</sup> Year capstone design project through all these teaching activities. A diagram of the PSE-related modules in Year 2 and their associated, increasing, difficulty is presented in Figure 1.

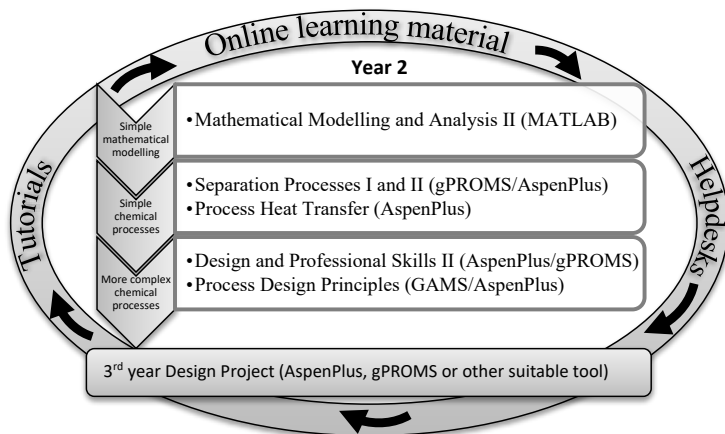


Figure 1: Distribution of PSE related modules and available support over the Year 2 curriculum

## **2. Computational tools for mathematical modelling and analysis**

One of the core modules in Year 2 is Mathematical Modelling and Analysis II (the first part of the topic is taught in Year 1) (UCL Module Catalogue, 2022). For each of the topics considered (e.g. Taylor and Fourier series, Transforms, Partial Differential Equations etc.), students participate in a tutorial where they have the opportunity to solve real-life examples, relevant to their engineering discipline, for which the specific mathematical concept is required. They mainly use MATLAB (Mathworks, 2022) in order to solve these problems.

As an example, students are asked to find the tangential velocity of a space capsule when it re-enters Earth's atmosphere, using Euler's method. Students first provide their answer symbolically using the relevant theory taught, and then proceed to solve the problem and find the exact solution using MATLAB. This exercise helps students to familiarise themselves with MATLAB and to learn, in general, how to use a suitable computational tool to solve challenging mathematical problems that are difficult, if not impossible, to solve by hand or analytically. Furthermore, part of the aim is for the students to realise that they can use MATLAB to solve similar problems later in their studies.

## **3. Combination of computational and experimental work**

A particular aspect of our programme is the combination of computational and experimental work within most modules, which is taught in an integrated manner alongside the theoretical concepts. This is to ensure that the students not only understand the fundamentals, but also know how to solve relevant problems using computational tools, as well as understand how the various unit operations function practically through lab experiments and how experimental input may be required when modelling these units.

### *3.1 Separation Processes I and II*

Within the Separation Processes I module, students have two pieces of coursework worth 20% each to the overall mark, each of them focusing on a different computational tool. For the first piece of coursework, when students have been made familiar with the concepts of flash separations and binary and multi-component distillation and their relevant equation sets, students are asked to model a dynamic equilibrium stage within gPROMS and observe its performance under different conditions (e.g. varying relative volatility and stage holdup). As mentioned above, gPROMS is also taught in Year 1 so students are expected to be familiar with how this software works, however, this is the first time they are given a more complex problem involving multiple components.

For the second piece of coursework, students are asked to simulate heat transfer and various separation processes (heat exchangers followed by distillation, absorption and/or liquid-liquid extraction) in AspenPlus with data that they themselves have obtained in the laboratory during a relevant experiment within the Engineering Experimentation module, and to discuss and justify any differences observed between the experimental and computational results. This coursework is due after the completion of the module to ensure that all students have obtained the relevant experimental data as the experiments are conducted throughout several weeks due to the large cohort size. The aim of this exercise is to encourage students to familiarise themselves with core industrial processes;

to understand how these work from both a practical and theoretical perspective; to think about the validity of any assumptions made when modelling each unit operation; as well as to reflect on the underlying assumptions and equations and their implications which are often 'hidden' within the computational tool.

As an example, in a distillation experiment students calculate the stage efficiency of the small-scale column they have operated. Then, they are asked to simulate the corresponding distillation process in AspenPlus using a rigorous column model such as RADFRAC, and to compare the cases when stages are ideal (i.e. stage efficiency is considered 100%) and when stages have the efficiency they found in the experiment. Discussion of the impact of stage efficiency on process performance is required, to encourage students to think about the importance of this parameter when designing a distillation unit. An additional aspect of this coursework includes the selection of the thermo-physical property method to be used in AspenPlus and for students to properly justify their choice, which usually requires them to look into literature to find experimental data that confirm the suitability of the specific method chosen. Other examples include students modelling a packed column with data obtained in the lab to observe and understand how column hydraulics work, as well as to simulate a continuous liquid-liquid-extraction process and discuss how the findings differ from the equivalent batch process they considered in the lab. Both of these exercises are also making use of AspenPlus.

For the Separation Processes II module, students are asked to simulate a crystallisation process using data from the relevant experiment. gPROMS is most suitable for this and therefore used in this case, and students are able to simulate the process, analyse, plot and discuss the generated results.

### *3.2 Process Heat Transfer*

Within the Process Heat Transfer module, students are asked to use their data, obtained in the relevant experiment, to model a shell and tube heat exchanger in AspenPlus as part of their 20% coursework, after having made a number of theoretical calculations. Students are asked to design the heat exchanger using a relevant heat exchanger model, state any assumptions made and provide their results such as the overall heat transfer coefficient and stream outlets.

The learning outcome for the students, in addition to learning how to model heat exchanger units using a higher level of detail, is to understand and critically discuss the underlying equations and assumptions when modelling such units, their design characteristics (e.g. in terms of equipment geometry) as well as the results generated (e.g. temperature profiles). This knowledge is expected to help them later on in their studies, especially when they come across the concept of heat integration and how heat exchangers are among the units which can help achieve this.

## **4. Modelling and simulation of more complex PSE concepts**

### *4.1 Process Design Principles*

The Process Design Principles module focuses on using theoretical approaches for the synthesis and design of more complex processes (i.e. processes that include several units). It is crucial that students learn not only how to consider process synthesis but also how to consider their design using computational tools (GAMS and AspenPlus in this module).

In the first part of the module, GAMS is used to solve the synthesis model of interest. Students should be able to identify all the input based on the problem description (e.g. feed conditions and unit specifications), and structure the code properly so that they are able to generate the desired results, for instance product compositions.

Later in the term, students are given a project on a specific design problem (e.g. the production of formaldehyde etc.) and are initially asked to synthesise and design a suitable process using the Douglas approach. They are then asked to simulate the conceptual process using AspenPlus, which typically requires the steady-state design of core units including, but not limited to, heat exchangers, plug-flow or continuous-stirred reactors and separation units. Students are also asked to find the heating and cooling requirements based on the generated results, as the overall aim of the project is to evaluate whether the design is economically attractive or not, making use of the relevant Economics Evaluation tool in AspenPlus.

In this project, students typically struggle with the existence of one or more recycle streams and how to properly include them in the flowsheet, and must consider the addition of tear streams and the selection of suitable initialisation procedures and initial conditions so that the simulation converges. This module clearly builds upon the previously discussed modules in order to expand the students' knowledge and capabilities on how to use computational tools by introducing them to more complex flowsheets and demanding procedures such as the inclusion of recycle streams.

#### *4.2 Design and Professional Skills II (Scenarios)*

In addition to the traditional taught modules discussed above, the chemical engineering programme at UCL includes a number of problem-based activities typically running over a whole week each, whereby the students work in teams to solve real-world problems, the so-called Scenarios (Graham, 2018). The aim of these problem- and scenario-based activities is to produce independent and self-directed engineers (Sorensen, 2016). Scenarios 3 and 4 (out of a total 6 Scenarios taught over Year 1 and Year 2) are typically related to Process Systems Engineering. This is an excellent opportunity for students to apply their knowledge from the taught modules, including the link between computational and experimental work mentioned above, but also apply their own ideas for the Scenario deliverables.

For instance, one Scenario 4 problem focused on Process Intensification and Reactive Distillation, where students considered fundamental process modelling and alternative process improvements using relevant simulation tools (in this case AspenPlus). The deliverables included, but were not limited to, column internals consideration; critical discussion of temperature and composition profiles; evaluation of process economics and environmental impact; as well as process safety. For each aspect they were required to justify their decisions based on their simulation findings from a well-rounded perspective. This particular Scenario also aimed at reminding students of the relevant computational knowledge obtained within Process Heat Transfer, to train and encourage students to make connections between different modules and apply knowledge obtained in one module for the purposes, and within the context of, a different module. For Scenario 4, students were asked to consider various types of distillation column condensers and reboilers, their suitability based on the specific process considered, as well as report the heating and cooling demands of the process, along with the required utilities.

## 5. Transition to 3<sup>rd</sup> year capstone design project

Once students have completed their Year 2 modules, they are familiar with the majority of the capabilities of the various computational tools and are therefore ready and well-equipped for the 3<sup>rd</sup> year's Process Plant Design Project module. The design project requires the modelling of a complete chemical engineering process using several computational tools (most often gPROMS and Aspen Plus depending on the particular units on which they are focusing) and the consideration of various aspects of those designs. These aspects, which are all considered using these tools and the critical engineering thinking of the students, include the optimal design of the process with respect to product quality, control performance, economics, environmental impact and sustainability, as well as process safety. This project works in practice as evaluation and confirmation of whether the learning outcomes of the Y2 modules and strategies discussed were successfully met, and how students employed the knowledge previously acquired for their open-ended project in Y3.

## 6. Conclusions

UCL's chemical engineering programme offers a number of opportunities for students to familiarise themselves with various computational tools through taught modules and problem-based learning. Students gradually become regular and confident users of PSE-related computational tools, learn how each of these works, understand their underlying equations and assumptions, as well as appreciate where experimental data is required. At the end of Year 2, students start becoming digitally literate as they are capable of simulating a wide range of simple and more complex chemical processes, including more challenging concepts such as recycle streams and heat integration, and evaluate these processes from various perspectives such as economical, environmental etc. The most important learning outcome, however, is that they perform these tasks making at the same time use of their own critical engineering thinking, and therefore make full use of all the potential benefits that PSE has to offer chemical engineering graduates.

## References

- Aspentech (2022). Aspen Plus USA, Aspen Technology. Inc., <https://www.aspentech.com/en/products/engineering/aspen-plus> [Accessed 27/10/2022]
- GAMS (2022). <https://www.gams.com> [Accessed 27/10/2022]
- Graham, R. H. (2018). The global state of the art in engineering education. Cambridge, MA: Massachusetts Institute of Technology.
- Mathworks (2022). <https://uk.mathworks.com/products/matlab.html> [Accessed 27/10/2022]
- Siemens Process Systems Engineering (2022). gPROMS, <https://www.psenterprise.com/products/gproms/process>, 1997-2022.
- Sorensen, E. (2016). Changing the World. The Chemical Engineer, 27-30.
- Tsatse, A. and Sorensen, E. (2021) Reflections on the development of scenario and problem-based chemical engineering projects. Computer Aided Chemical Engineering 50, 2033-2038.
- UCL Module Catalogue, <https://www.ucl.ac.uk/module-catalogue/> [Accessed 27/10/2022].

# Teaching courses to STEM students – Lessons learned from the Pandemics

Nelson Chibeles-Martins,<sup>a,b</sup> Lourdes B. Afonso,<sup>a,b</sup>

<sup>a</sup>*Departamento de Matemática, FCT-NOVA , 2829-516 Caparica, Portugal*

<sup>b</sup>*NOVAMATH Centro de Matemática e Aplicações, FCT-NOVA , 2829-516 Caparica, Portugal*

## Abstract

The current pandemic has been having a huge effect in teaching at all levels. Particularly, in Portugal, for STEM (Science, Technology, Engineering and Mathematics) students attending Universities and taking high learning degrees, it was particularly harsh. Nevertheless, the teaching crisis triggered by the SARS-CoV-2 pandemic was fertile terrain for pedagogical experimenting, learning and growing. In a previous work (Chibeles-Martins and Afonso, 2022) the authors described their experience during three eventful semesters teaching Applied Mathematics courses to STEM students.

In this work the authors present several teaching practices that were firstly introduced during the aforementioned period as an emergency patch-up but revealed to be so curiously thought provoking that caused them rethink how courses were taught before the Pandemic versus how they should be effectively taught. The authors explore how these teaching practices had to be readapted to a face-to-face teaching situation with the Lockdowns' end and what were their effect on students' learning experiences in a post-pandemic environment. In class and home-work assignments; working and discussion groups; lectures by students; more dynamical MOODLE pages, filled with on-line asynchronous activities; and more diverse evaluation moments are some of the practices that will be described and explored though out this presentation.

**Keywords:** Teaching Practices, Pandemics, Active Learning, Peer Instruction.

## 1. Introduction

FCT NOVA (NOVA SST - School of Sciences and Technology) is one of the most renowned Sciences and Engineering schools in Portugal. It offers doctorate, master and bachelor's degrees in mathematics, Computer Sciences, Physics, Chemistry, Industrial Management, Geology, Environmental Sciences, Material Sciences, among others. All offered degree programs are accredited by the national agency for assessment and accreditation of higher education and all Engineering programs have the EUR-ACE certification. Around 8,000 students, 400 professors, 1,000 researchers and 200 staff (NOVA School of Sciences and Technology, 2022) study, work and live inside the campus.

The current CoViD19 pandemic had a tremendous impact in this community slowing down research and teaching activities, leading to students dropping out of their hard-earned higher learning opportunities. Teachers had to reinvent their pedagogical practices

in order to keep their lectures, tutoring classes and laboratorial sessions interesting during the successive lockdowns. And amid the crisis opportunities bloomed.

## 2. Lessons Learned

During the Pandemics that has been affecting us since the end of 2019 Academia had to face terrible challenges like every institution. But by facing them teachers also had the opportunity to become more deeply aware of some issues that were mainly dormant. In the following the authors will address some of the situations they experienced and the lessons they learned from them.

### 2.1. Recorded Lectures

At NOVA SST the majority of professors would feel uncomfortable when confronted with the notion of having to record their lectures. This occurs due to several factors:

- Discomfort from potential over exposition;
- Fear of students skipping classes because they have the lecture's recording available;
- Apprehension from the technology needed to create a good quality recording.

The pandemic forced us to be creative on the ways to reach our students and a lot of teachers accepted the challenge of recording their lectures. A considerable number of students were missing remote synchronous lectures due to illness, technical issues or simply because all their families was disputing the only available computer at their homes.

And despite all the previous reservations we found that:

- Non expensive movie maker software is easily available online;
- The recordings' access can be monitored. It is impossible to ensure it will not end up on YouTube but you can have some control on who can initially view your movies;
- Students rather attend to lectures. They usually only resort to recorded lectures if they cannot attend to the real deal. Or when they do not remember a particular detail and want to review it.

On the other hand, keeping a collection of lecture recordings online updated can be particularly time consuming and the learning curve for the necessary software and equipment can be off putting.

Nevertheless, recorded lectures are a very useful pedagogical tools and teachers should not be afraid to use them.

### 2.2. e-learning platforms

During the 2000s, at NOVA SST, some members of the teaching staff became aware of the MOODLE<sup>1</sup> project and an e-learning office was created to help teachers who wanted

---

<sup>1</sup> MOODLE is the acronym to Modular Object-Oriented Dynamic Learning Environment. It is a digital learning platform designed to provide an integrated, secure

to learn how to create and manage MOODLE pages. In spite of its usefulness the use of MOODLE pages was considerably limited to a small number of more gimmick inclined teachers.

During the harsher initial Lockdowns midterm tests, presential lectures and face-to-face tutoring sessions were cancelled. Suddenly, most teachers had to fall back to digital online tools to complement on-line lectures. And realized how the use of e-learning platforms, namely MOODLE pages, would help tremendously with organizing and providing studying material, creating on-line evaluations, informing students, monitoring students' interaction with the proposed tasks, ...

At NOVA SST the number of Courses being supported by MOODLE increased immensely overnight. And did not stop after the Pandemics lost steam. Currently the number of active MOODLE pages per semester is so high that the server where they are lodged is getting overworked with an alarming frequency and it is on the process of being replaced/upgraded.

It is undeniable that online support provided by e-learning tools greatly improves the dynamics of any academic course. It does not really matter the preferred environment. MOODLE pages, Virtual Classrooms, simple online home pages. Circumstances that would be hard to remedy, like students not attending classes due to illness, athletic competitions or because they accumulate their academic life with a full-time job, are easily solved with the support of an e-learning platform.

Still, e-learning platforms need to be lodged somewhere and the generalized use of these tools imply a considerable investment in stable servers and internet connections. That can be an obstacle for learning institutions with tighter budgets.

But, despite its technical demands, currently, there are no motives to keep lectures solely to the physical classrooms.

### *2.3. Diversified Evaluation*

Midterm and final tests, and exams are one of the most common kind of evaluations available to teachers. And in spite of its more complicated and cumbersome organizational logistics it is a very comfortable way of individually and quantitatively assessing students' performance in a course. Nevertheless, during the Lockdowns tests and exams were heavily restricted and teachers had to opt between postponing them *sine die* or adopt an alternative way of evaluating their students.

Group assignments, on-line presentations, essays, short on-line quizzes, small workshops, individual reports, and exercises solved and submitted on a short period of time were different types of evaluations that were explored as alternatives to midterm and final tests. And that imposed diversity in evaluation brought to mind an issue that was usually ignored: a more diversified evaluation enables assessing skills, abilities, competences and

---

and highly flexible system for teachers, educators and students. (The Moodle Project, 2020). Often used by schools and universities as an e-learning platform.



knowledges that are harder to evaluate through a standard test. Additionally, they alleviate tests and exams' psychological load and their relative importance from students' minds.

It is also easier to schedule a single test to a convenient date if courses do not have their evaluation centered solely on midterm and final tests. There will be a lesser number of tests to schedule and therefore the overlapping risk is much lower.

Obviously, this diversity can generate an all-new batch of evaluation problems the teaching staff is not used to address. What is the adequate balance between traditional tests and oral presentation evaluations? How to minimize personal biases during the more personal and potentially more subjective evaluation moments like group assignment discussions or lectures done by students? Additionally, paradigm shifts like these often have unexpected consequences... What will be the characteristics of top grades students under a generalized evaluation focus shift?

In conclusion, the constraints imposed by the Pandemics revealed that it is possible to evaluate students more accurately, more effectively and with less burden on the students during the critical periods of midterm and semesters' end.

#### 2.4. Videoconferencing

Not many years ago using videoconference systems was not exactly an everyday event for most academics. Nevertheless, the Pandemics forced teachers and researchers to resort to these communication platforms daily, sometimes *ad nauseam*. Currently with Life in Academia going back to some kind of normalcy one would expect to go back to those times when videoconferencing was seldom used.

Actually, nowadays it became part of our routine. As most companies already knew it is an incredibly useful tool for team meetings. Training courses for teachers can have both videoconference components and presential sessions, making them more accessible to non-local teachers. Tutoring times can be scheduled to more convenient hours for both students and tutors when held on a videoconference.

With the investment already made in videoconference technology, hybrid lectures are the new normal, allowing some students to follow the lecture remotely. Additionally, for public learning institutions it is currently more economically viable to invite guest professors for occasional seminars. Travelling is becoming more expensive and having an invited speaker lecturing by videoconference may be less appellative for the attendants, but it certainly is a lesser expense for the inviting school.

#### 2.5. Versatility, adaptability and inclusion

Change is usually faced with some apprehension in the Academia. Teaching is more comfortable when grounded on known routines. When teachers know what to expect and everything goes according to what was planned their time can be more efficiently managed and they can be more productive. Nevertheless, almost every other semester some particular circumstances may challenge those routines and submit the course's planning to a stress test. An unexpected foreigner student with visa problems is not able to attend the course presentially, a sudden change on the regular class format due to space constraints, a student that has special learning needs enrolls on the course... All these circumstances can be stressful and anxiety inducing to even a well-prepared teacher.

But the Pandemic brought the unexpected and the exceptional to our Academia in unprecedented wave after wave. Teachers had to adapt quickly and ruthlessly. Every week something worse was happening and demanding another adjustment. We had to invent, create, become more tolerant and versatile. Way beyond we ever thought we were able to do.

It was tremendously hard, stressful and time consuming. But we learned we could get more efficient, more effective, and abler to deal with accidental and unforeseen events and circumstances affecting our teaching activity.

In the aftermath of the hardest part of the current Pandemic we became aware that, by comparison, the previous stressful happenstances that came knocking at our classroom's doors were mere nuisances. And that teachers are much more versatile, adaptable and resilient they ever considered being before.

### **3. Conclusions**

The pedagogical tools mentioned previously were already known but seldom used by a large number of Academia members. The current CoViD19 pandemics put us all to a hard and long stress test. Teachers had to become more inventive, creative and resourceful resorting to those already available tools and use them at their full potential in order to keep students motivated, involved and obtaining high performance levels.

After the end of the Lockdowns, we noticed we kept using this tools even when lectures were fully presential. And that their contribution for a better learning experience were still considerable. Actually, in certain circumstances their effect was even better during on site classes.

We believe we will not go back to a time when teaching and learning did not include these tools and mind frames. Some of them can be time consuming and expensive. But a considerable investment in equipment and training was made during the Pandemics. The social distancing forced the Academia to pay a terrible mental and emotional price that is yet to be fully understood. It would be a huge waste of resources and acquired knowledge if we do not exploit to their fullest all the hard lessons, we learned during this trouble times.

### **Acknowledgments**

This works was partially funded by national funds through the FCT – Fundação para a Ciência e a Tecnologia, I.P. (National Foundation for Science and Technology), under the scope of the project UIDB/00297/2020.

This work was partially funded by Magentakconcept, lda.

### **References**

- N. Chibeles-Martins and L. B. Afonso, 2022, Teaching courses heavily dependent on computational resources to STEM students during Pandemics, *Computer Aided Chemical Engineering*, 51, 1687-1692

Nova School of Sciences and Technology, 2022, *Overview*. <https://www.fct.unl.pt/en/about-fct/overview>

The Moodle Project, 2020, *About Moodle*. [https://docs.moodle.org/311/en/About\\_Moodle](https://docs.moodle.org/311/en/About_Moodle)

# Enhancing Human Machine Interface design using cognitive metrics of process operators

Mohammed Aatif Shahab,<sup>a</sup> Babji Srinivasan,<sup>a,c</sup> Rajagopalan Srinivasan<sup>b,c</sup>

<sup>a</sup>*Department of Applied Mechanics, Indian Institute of Technology Madras, Chennai, 600036, India*

<sup>b</sup>*Department of Chemical Engineering, Indian Institute of Technology Madras, Chennai, 600036, India*

<sup>c</sup>*American Express Lab for Data Analytics, Risk and Technology, Indian Institute of Technology Madras, Chennai, 600036, India*

[babji.srinivasan@iitm.ac.in](mailto:babji.srinivasan@iitm.ac.in), [raj@iitm.ac.in](mailto:raj@iitm.ac.in)

## Abstract

In a typical process industry, Human Machine Interfaces (HMIs) are essential for all aspects of communication and interaction between operators and processes, vital for process safety, quality, and efficiency. Good interface design enables operators to accomplish their duties efficiently and effectively with minimal errors. Consequently, it is crucial to design HMIs in a way that facilitates collaboration between automation and operators. Current HMI design practices have a lot of subjectivity and often ignore operator's cognitive behavior. In this work, we propose a quantitative approach that uses cognitive metrics – association and salience metrics – obtained using eye-tracking to improve HMI design for process operators. The association metric helps in identifying information sources which are often used together while salience metric informs about information sources which are extensively used from the HMI. Based on insights from these cognitive metrics, we designed a new HMI. To evaluate the relative efficacy of each HMI, human subject studies were conducted. The results indicate that the new HMI developed with the use of cognitive metrics improved the operators' efficiency. Thus, the proposed approach can help enhance usability of industrial HMIs.

**Keywords:** Human Machine Interface, Cognitive Behavior, Eye-tracking, Correspondence Analysis, Safety

## 1. Introduction

The Human Machine Interfaces (HMIs) used in modern industries depict complex systems, massive amounts of data, high degrees of automation, and safety management systems. These developments place more cognitive workload on operators increasing their chances of errors. Good interface design enables operators to accomplish their duties efficiently and effectively with minimal errors. Poor design results in an increase in reaction times and likelihood of error in perception and understanding. A poor HMI information display reduces the likelihood of properly anticipating an abnormal event by five times (Krajewski, 2014). Studies reveal that human error accounts for 70% of industrial accidents (Mannan, 2013). Therefore, it is crucial to develop HMIs driven by operators' requirements to enhance how they utilize, manage, and retain the process information.

The development of industrial HMIs has traditionally been guided by standards issued by organizations such as the International Society of Automation (ISA). Common practices

in this area include the use of P&ID images (Lee et al., 2017) as a display for Distributed Control Systems (DCS) and the depiction of real-time process data. Despite significant advancements in the display capabilities of DCS, a significant proportion of operator graphics still follow this paradigm, with only limited improvements in their practical effectiveness and capability. Information on HMIs is typically organized according to hierarchies, with data pertaining to the physical components that support key functions displayed together (Jamieson et al., 2007). To assess the usability of these HMIs, various measures have been proposed, including task-based measures, operator actions, and expert opinion. An expert-based approach to evaluate HMI usability involves using a well-defined process with specific evaluation items to pinpoint areas for improvement (Aranburu et al., 2019). This method has been tested in experiments and found to enable participants to identify numerous aspects of the evaluated interface that could be improved. In yet another study, task-based measures such as success rate, total time on task, response time and subjective rating scale were used to evaluate usability of a power plants HMI (Lee et al., 2017).

In summary, current HMI design practices rely heavily on the expertise of subject-matter experts and lack objective evaluations of the cognitive abilities of operators (Ikuma et al., 2014). These abilities, such as information processing, reasoning with process states, and decision-making processes, play a crucial role in reducing the likelihood of human error. To address this issue, this study proposes the use of eye tracking to develop cognitive metrics for enhancing usability of industrial HMIs. The validity of this approach is demonstrated through human subject studies conducted on an in-house chemical process simulator.

## **2. Literature Studies**

With the digitalization in industries, HMI systems are becoming pivotal in interaction between the operator and the process. It is, therefore, necessary to keep track of operator cognitive behavior while designing HMI. An efficient HMI reduces operators' cognitive workload and assists them in accomplishing their duties with minimal errors. The performance of operators improves when they are provided with physical, functional, and task information in addition to the traditional physical and functional data on HMI (Jamieson et al., 2007).

In recent years, researchers have used eye-tracking to evaluate and improve HMIs in high-risk industries using the cognitive behavior of operators. An eye-tracking based methodology was developed to enhance the usability of a nuclear power plants HMI (Ha et al., 2018). Eye-tracking measures such as fixation-to-importance ratio (FIR) and selective attention effectiveness (SAE) were developed to identify the effective attention of operators on the HMI. The FIR indicates the usefulness of a single information source, whereas the SAE is a measure of the usefulness of all available information sources taken together. FIR and SAE values were found to be much lower on steam generator level suggesting less attention to the information source. The participants reported that they had a difficulty understanding the pattern of level change of the variable. Accordingly, the HMI design was enhanced.

The current studies to enhance HMI design using operator cognitive behavior are restricted in their ability to be generalizable to other industries and do not give an objective foundation to improve HMI design. Previously, we have used eye-tracking to get insights into operator's cognitive behavior in process industries (Shahab et al., 2021a; Iqbal et al., 2018). Shahab et al. (2022) developed Hidden Markov Model based

methodology to capture operators' mental models using eye-tracking. The study of mental models is very relevant to HMI design as it reveals how operators perceive the information from the HMI. In this work, we propose a methodology that uses cognitive metrics to improve HMI design for process control applications.

### 3. Experimental Studies and Methodology

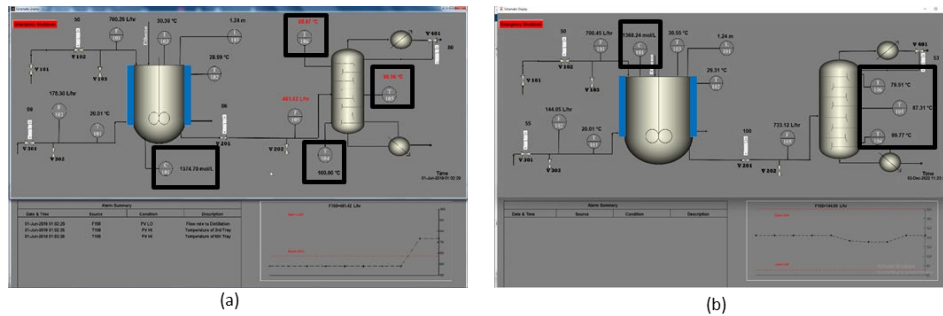


Fig. 1 (a) Old HMI of the experimental testbed (b) New HMI of the experimental testbed

The experimental testbed consists of a simulated ethanol process plant. A typical distributed control system is used to control the process. The operators need to monitor the process and intervene during abnormalities. Ten participants interacted with the process through an HMI (Fig. 1 (a)). The study consisted of six different scenarios and each participant repeated all the scenarios multiple times, henceforth called as trials. This ensures that participants develop enough expertise in monitoring and dealing with process abnormalities. We recorded process, data, operator actions and eye gaze data during each trial.

The proposed methodology is based on the percept that the operators pay attention only to those information sources that they think have a relationship with the current abnormal situation. We used correspondence analysis (CA), a multivariate tool, to obtain an asymmetric plot showing operator's attention allocation on various information sources of HMI, henceforth called as Areas of Interest (AOIs), during the abnormal scenarios. CA shows the relationship between two categorical variables, in the current study, areas of interest (AOIs) on the HMI and various abnormalities relating to the disturbance in the plant. The operator's attention on various AOIs is crucial for determining their preferences during abnormal situations and can be exploited to place information sources on the HMI to help them achieve the task efficiently.

We develop two quantitative cognitive metrics – association and salience metric – based on the angle between a scenario and an AOI and the distance of an AOI from the origin on the asymmetric plot. The association metric captures the operator's attention to key variables that they believe to have a causal relation with the abnormal situation. Association metric reaches unity if a variable(s) has a causal relationship with an abnormal scenario. Therefore, such variables should be placed together on the HMI to increase its efficiency. The salience metric captures the prevalence of AOIs, which are extensively used on the HMI. Such AOIs should be easily accessible during abnormal situations. Interested readers are referred to Shahab et al. (2021b) for a detailed description of the calculation procedure. Using the insights from CA of eye gaze data, the HMI is redesigned. For the sake of simplicity, we will call the redesigned HMI as the new

HMI and the previous HMI as the old HMI. We conducted similar human subject studies to compare the performance of operators in both the HMIs (old HMI and new HMI).

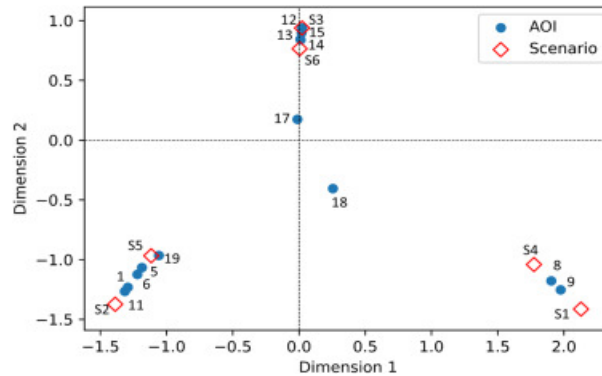


Fig. 2 Asymmetric plot obtained from correspondence analysis of eye gaze data from an expert operator

#### 4. Results and discussion

In this section, we illustrate the application of the proposed methodology in enhancing the HMI design.

As discussed in the previous section, we obtained the asymmetric plot from the correspondence analysis of the operator's eye gaze data. Fig. 2 shows the asymmetric plot for a typical expert operator. The numbers 1-19 represent the areas of interest (AOIs) on the human-machine interface (HMI), while S1-S6 indicate six disturbance scenarios. The scenarios represent disturbances in one or more process variables, including low coolant flow rate and low reflux ratio. As shown in Fig. 2, AOIs 12-15 are closer to scenarios S3 and S6 (low reflux ratio disturbance), indicating that the operator primarily focuses their attention on AOIs 12 (temperature of the 8th tray of the distillation column), 13 (reflux flow control valve V401), 14 (temperature of the 4th tray), and 15 (temperature of the 3rd tray). The association metric between the scenarios and AOIs is approximately one.

The operator's focus on only a few information sources (AOIs) on the HMI indicates the significance of these variables in relation to the abnormal situation at hand. Hence, these variables should be kept together on the HMI. In the current HMI design, T104, T105, and T106 are in different places and separate from the control valve V401 (see Fig. 1(a)). Based on the insights from the asymmetric plot, these variables should be kept together. Therefore, the HMI was redesigned and these variables were kept together, as shown by rectangular boxes in Fig.1(b).

The proposed methodology also has the potential to identify incorrect information placement on the HMI. For example, as shown in Fig. 2, for scenarios S2 and S5, the operators' preferences are AOIs 1 (feed flow to the reactor tag), 5 (feed flow to distillation column control valve), 6 (feed flow to the reactor control valve), 1 (flow rate to distillation column tag), and 19 (concentration of ethanol in the reactor). These scenarios are related to a disturbance in the feed flow to the reactor, which triggers alarms at F101 (AOI 1), F105 (AOI 11) and C101 (AOI 19). The correct action is to manipulate valve V102 (AOI 6). Although the operator successfully completes the task using V102, Fig. 2 shows that AOI 5 is close to scenarios S2 and S5 indicating they have also thought about V201 which making decision to control the abnormality. This confusion is due to the proximity of

C101 and V201 on the HMI (Fig. 1(a)). To avoid such confusion, in the new HMI design (Fig. 1(b)), the location of C101 was changed to be near the feed flow to the reactor line, helping operators diagnose the root cause of alarm C101 as a disturbance in the feed flow to the reactor. Additionally, Fig. 2 shows that AOI 17 (trend panel) is close to the origin on the asymmetric plot (saliency metric approximately equal to 0), indicating that the trend panel was extensively used during all subnormal scenarios and should be easily accessible during abnormalities. Therefore, the position of the trend panel was left unchanged. Next, we compare the operator's performance in both HMIs.

Table 1: Operator performance measures obtained from both the HMIs

<b>Operator performance measures</b>	<b>Old HMI</b>	<b>New HMI</b>
Time taken to first correct control action (seconds)	9.9	5.7
Time to first fixation on correct action (seconds)	7.5	3.7
Scanpath length from first alarm to correct action (pixels)	2056	1206
Task completion time (seconds)	38.7	38.2
Fixation count	35	20

#### 4.1. Operator performance evaluation

As discussed previously, similar experiments were carried out using the new HMI. To evaluate the efficiency of HMIs, we used operator performance measures that describe the speed and accuracy of the task performed by the operator.

Consider two participants, operator A, and operator B, who deal with an abnormal situation involving a disturbance in feed flow to the reactor resulting in alarm F101, alarm F105, and alarm C101. Operator A interacted with the process using the old HMI, while operator B interacted using the new HMI. Both operators have a similar level of expertise. For the proposed methodology to be effective, it is expected that operator performance in the new HMI should be better. The performance measures of operators A and B are listed in Table 1. The time to the first fixation on correct control action and time taken to correct action in response to the abnormal situation is lower in the new HMI compared to the old HMI. The lesser time taken with the new HMI signifies that the new HMI design helps the operator for early identification of the corrective action in response to the abnormality.

Further, the scanpath length from alarm occurrence to the application of the first control action is lower in the new HMI (Table 1). Scanpath length is a productivity measure that can be used for defining optimal visual search (Goldberg and Kotval, 1999). The lower values of scanpath length signify that the new HMI aids the operator during visual searches, such as locating the correct control action. The task completion time shows no difference, probably because the operators have similar expertise. The number of fixations is related to the number of components the user must process. When searching for a target, a large number of fixations indicates the user explored many other information sources prior to selecting the target, as if distracted or hindered from isolating the target (Goldberg and Kotval, 1999). As reported in Table 1, the number of fixations after the occurrence of the alarm till the correct control action is lower using the new HMI, indicating an enhancement in operator performance. The critical measure of an HMI is how well it performs when the demands are greatest – in an abnormal situation. As such, the new HMI design based on developed cognitive metrics outperforms the old HMI in all aspects of performance. The results revealed that the new HMI aids the operator in dealing with process abnormality.



## 5. Conclusions

We put forward the idea of incorporating operator cognitive behavior into the design of industrial HMIs to improve their usability. Our approach is based on the percept that operators tend to attend only relevant information sources during abnormal situations. Therefore, such information sources should be placed together on the HMI to enhance its usability. We used correspondence analysis of operator eye gaze data to develop quantitative cognitive metrics - association and salience metrics - that identify the most relevant information sources during various abnormal scenarios. Based on these findings, we redesigned the HMI and kept the relevant information sources together, thereby reducing confusion and improving the operator's ability to diagnose and resolve disturbances efficiently. The efficacy of the new HMI was validated through human subject studies, which showed enhanced operator performance compared to the original HMI. In future studies, we plan to conduct statistical analysis with large scale human subject experiments. Additionally, we will utilize electroencephalography (EEG) to measure the cognitive workload of operators using both HMIs.

## References

- Aranburu, E., Lasa, G., & Gerrikagoitia, J. K. (2019, June). HEMEI: new user experience evaluation tool for Human-Machine Interfaces. In Proceedings of the XX International Conference on Human Computer Interaction (pp. 1-2).
- Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: methods and constructs. *International journal of industrial ergonomics*, 24(6), 631-645.
- Ha, J. S., Byon, Y. J., Cho, C. S., & Seong, P. H. (2018). Eye-tracking studies based on attentional-resource effectiveness and insights into future research. *Nuclear Technology*, 202(2-3)
- Ikuma, L. H., Harvey, C., Taylor, C. F., & Handal, C. (2014). A guide for assessing control room operator performance using speed and accuracy, perceived workload, situation awareness, and eye tracking. *Journal of loss prevention in the process industries*, 32, 454-465.
- Iqbal, M. U., Srinivasan, B., & Srinivasan, R. (2018). Towards obviating human errors in real-time through eye tracking. In *Computer Aided Chemical Engineering* (Vol. 43, pp. 1189-1194). Elsevier.
- Jamieson, G. A. (2007). Ecological interface design for petrochemical process control: An empirical assessment. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans*, 37(6), 906-920.
- Krajewski, J. (2014). Situational awareness—the next leap in industrial human machine interface design. White paper, Invensys Systems, Houston, USA.
- Lee, S. T., Kim, S. Y., & Gilmore, D. (2017, October). Human-in-the-loop evaluation of human-machine interface for power plant operators. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 34-39). IEEE.
- Mannan, S. (2013). *Lees' Process Safety Essentials: Hazard Identification, Assessment and Control*. Butterworth-Heinemann.
- Shahab, M. A., Iqbal, M. U., Srinivasan, B., & Srinivasan, R. (2022). HMM-based models of control room operator's cognition during process abnormalities. 1. formalism and model identification. *Journal of Loss Prevention in the Process Industries*, 76, 104748.
- Shahab, M. A., Iqbal, M. U., Srinivasan, B., & Srinivasan, R. (2021b). Metrics for objectively assessing operator training using eye gaze patterns. *Process Safety and Environmental Protection*, 156, 508-520.
- Shahab, M. A., Srinivasan, B., & Srinivasan, R. (2021a). Evaluating control room operator training outcomes through eye gaze augmented multi-scale data. In *Computer Aided Chemical Engineering* (Vol. 50, pp. 1307-1312). Elsevier.

# A Cloud-based Collaborative Interactive Platform for Education and Research in Dynamic Process Modelling

Vinay Gautam, Alberto Rodríguez-Fernández, and Heinz A. Preisig

*Department of Chemical Engineering, Norwegian University of Science and Technology, Trondheim, Norway, 7491, Norway.*  
*heinz.preisig@ntnu.no*

## Abstract

Process modelling is used in many disciplines for various purposes like simulation, process design, optimisation and control. Although there is an increasing demand to build large, complex process models involving multiple disciplines, a systematic approach to developing such models is largely missing from engineering education and research. The ontology-based methodology and ProcessModeller (ProMo) tool suite developed over the years (Preisig, 2021, 2020; Elve and Preisig, 2018) help users to build multi-disciplinary and multi-scale models systematically. This paper presents a cloud-based platform ProMo-Remote that uses Symphony-Remote, a free and open-source web service. In this platform, the graphical user interfaces of the ProMo tool suite are accessible using a regular web browser, freeing the user from installing the desktop tool locally. Users can save, download/upload modelling data and use it in computational workflows utilising cloud-based interactive computing tools like Jupyter notebooks. The user can also collaborate with others by screen sharing in real time. The platform can easily be configured and deployed with additional applications to facilitate further education, training and research on the systematic development of complex process models.

**Keywords:** Process modelling; computer-aided modelling; Symphony-remote; Jupyterhub.

## 1. Introduction

Models play a central role in many engineering disciplines and process systems engineering is particularly driven by models. Model design moves into a more dominating position as a research and development domain since the processes involved with domain models are the key bottleneck. The formulation of the model and its effective handling all the way through the generation of the target

code are just two of the numerous difficulties that must be overcome. The basic models are traditionally written by hand, which means that equations from published or developed models are converted into a target language and then made accessible as a module in a library. These library modules of unit operations or equivalent are then combined using interactive tools to create models.

The modelling methodology implemented in the ProMo tool (Preisig, 2021) captures multidisciplinary aspect in the model generation process by using reductionism, where model generation is based on identifying the smallest underlying basic entities in the context of the intended application. A network of these interacting basic entities then describes the behaviour of the modelled system.

Modern processes and materials modelling workflows are becoming increasingly dependent on cloud infrastructure, e.g. AiiDAlab (Yakutovich et al., 2021) and Materials Cloud (Talirz et al., 2020). The sheer volume of data generated by models and tools used in a workflow simulation would be hard to analyse without high-performance computing infrastructure. In the cloud, a task that could have taken days or weeks on a desktop computer might be finished in a matter of hours. This gap is frequently great enough to execute previously deemed impossible calculations.

We developed a cloud-based deployment of the ProMo tool to give remote access to users on the web. The platform uses Symphony-Remote (SR), a web service created as part of the Horizon2020 Symphony and MarketPlace projects and is free and open-source. The ProMo-Remote platform has the following highlighting features. First, ProMo is a modelling suite that consists of a set of modules built using Python. Several external dependencies are required and installing them is rather demanding for users without previous knowledge of Python. ProMo-Remote does not require users to install anything on their machine because the tool can be accessed using any regular browser. Second, the ProMo-Remote runs remotely in a container, making it easier to administer and configure the overall platform. Third, onboard apps can share a dataspace while running in their own Docker containers, making it simple to create and share computational workflows utilising cloud-based interactive computing tools like Jupyter notebooks.

## 2. The ProMo Workflow

Figure 1 shows the basic ProMo workflow used when designing a model. It all starts by building an ontology (Gruber, 1993) tailored to the application domain.

The ontology provides a common namespace and the fundamental structures for the behaviour description and represents the construction blueprint for complex models (Elve and Preisig, 2018). Next, all the equations necessary to describe the application processes are inserted. Following this, entities are constructed, and groups of equations representing particular behaviours are linked. Up to this point, a high level of expertise is required to ensure that all the information is included correctly.

The regular user will interact more with the modules to the right of the dashed line in Figure 1. In the Model Composer, the previously constructed entities can be used as basic blocks to graphically represent the application processes' physical topology. A high level of re-usability is promoted by allowing to save and reuse of parts of this topology. Finally, in the Task Generator module, the user may choose to generate the code necessary to perform different tasks using the model designed.



Figure 1: The ProMo workflow consists of several modules used in sequence. There is a clear division between modules based on the level of expertise required to use them.

### 3. **Simphony-Remote Web Service**

Simphony-Remote is an open-source project that resulted from the development efforts of several European projects, starting from the SimPhoNy project and the most recent the H2020 FORCE and Marketplace projects. The SR provides a JupyterHub<sup>1</sup> based web interface for running desktop and web applications in Docker containers. Figure 2 provides a graphical overview of the main components in the SR. It uses JupyterHub to configure and control the cloud

---

<sup>1</sup> JupyterHub provides a multi-user Hub that spawns the Jupyter Notebook server per user. The Hub consists of a configurable proxy to serve requests from the client's browser and an authentication class to manage how users interact with the system. It provides several authentication options, e.g., Lightweight Directory Access Protocol (LDAP) and OAuth.

infrastructure. Three types of users can access the service: admin, the non-admin and remote viewer. The admin user has roles in building and configuring the platform and extends roles to authenticated non-admin users to access containerised onboard apps. The third type of users are remote users who can access the application via a shared link from the authenticated users and do not need to authenticate directly to get access.

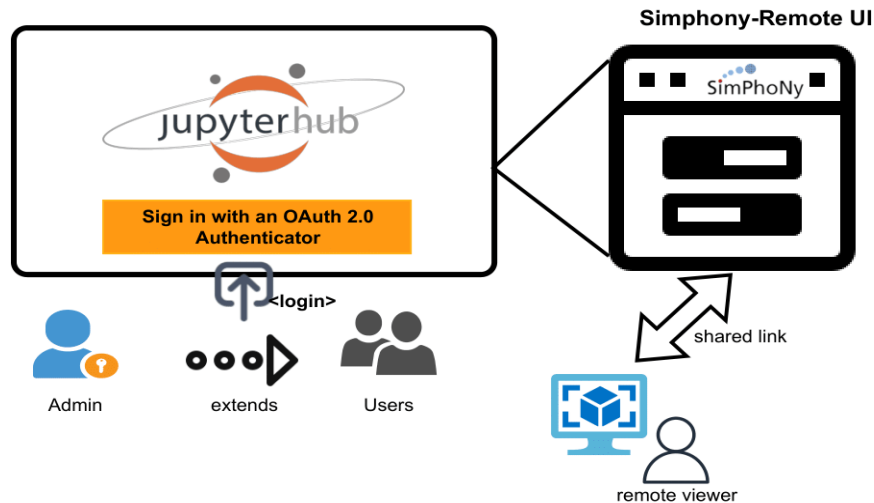


Figure 2: An overview of the Simphony-Remote web service.

#### 4. The ProMo-Remote Platform

The ProMo-Remote platform is preconfigured with three applications: the ProMo app, a file manager app and a Jupyter notebooks app, as shown in Figure 3. These containerised apps are first built using the Docker build process and then deployed on cloud infrastructure, as discussed below. The Docker images encapsulate applications as stand-alone software entities that can be run independently as software processes while sharing a common dataspace among the apps.

##### 4.1. Building application containers

Containerised apps are built in two steps. First, a base image of the host operating system with some add-ons, such as virtual network control (VNC) application, is built. Second, the application's Docker image is built on top of the base image.

## *4.2. Cloud Deployment*

Cloud deployment of the ProMo-Remote platform involves two steps. First, an instance of SR is configured and launched as a web service that uses nginx and reverse proxy. Second, the prebuilt containerised images of the apps are uploaded to the SR web service from its user interface.

## *4.3. Running the Applications*

The users can log in to the platform and will see a list of applications that they have access to, which is enabled by admin. As shown in Figure 3, the user 'vinay.gautam' has access to all three currently deployed apps in the ProMo-Remote. The user can select an application and run it by simply clicking the target application's start icon, which opens the GUI of the application. The user can interact with it and switch between different apps. The user can also share an app's running session with other remote users for collaborating in real time. Running apps are affected neither by the web browser's accidental closing nor by temporal network connectivity loss. It should be noted that closing an app would erase all its data, and the app will reset to its initial state, except the data saved on the shared dataspace directory named 'workspace'.

## **5. Conclusion**

The ProMo-Remote provides a cloud-based platform to systematically design multidisciplinary process models using the ontology-based methodology and ProMo tool. The platform is easy to manage, configure and deploy on cloud infrastructure, and its web interface is accessible on any regular web browser. The platform has a shared dataspace, making it easy to create and share computational workflows, for example, by utilizing cloud-based interactive computing tools like Jupyter notebooks. The platform can be used for education, research and knowledge transfer in the dynamic processes modelling community by enabling higher collaboration between expert and non-expert users. We provide a set of model-building exercises to help users familiarise themselves with the ProMo tool's usage to create process models systematically. We aim to create more training and pedagogical content to assist others in using this infrastructure to educate and train modellers in systematically building multidisciplinary, complex process model workflows.

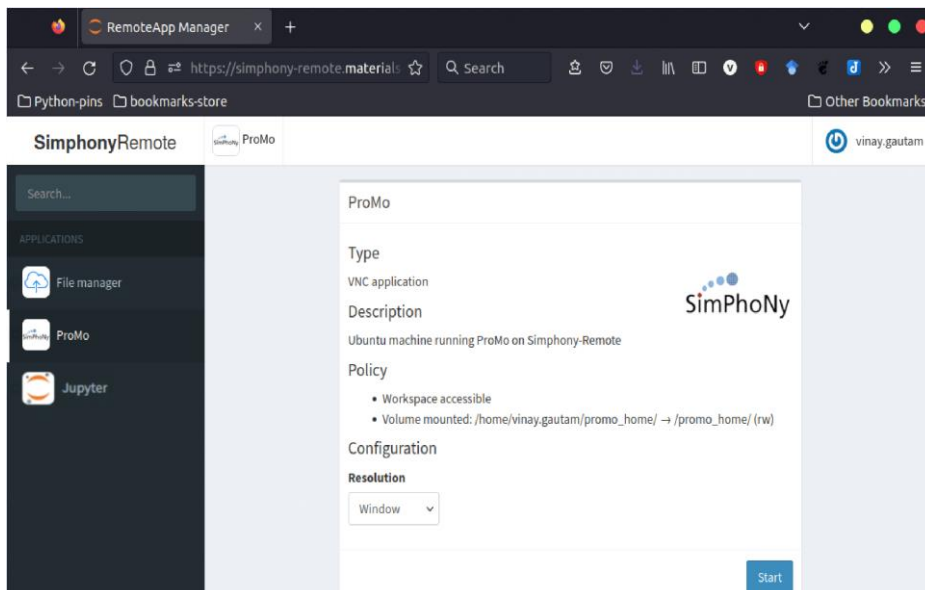


Figure 3: A snapshot of the ProMo-Remote platform running on a web browser.

### Acknowledgements

Bio4Fuels RCN: 257622, MarketPlace H2020-NMBP-25-2017:760173, VIPCOAT H2020-NMBP-TO-IND-2020:952903, NanoLodge NFR 299363.

### References

- Elve, A. T., & Preisig, H. A., 2018. A framework for multi-network modelling. *Computer Aided Chemical Engineering*, 44, 2215–2220.
- Gruber, T. R., 1993. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2), 199–220.
- Preisig, H. A., 2020. Promo - a Multi-disciplinary Process Modelling Suite. *Computer Aided Chemical Engineering*, 48, 571–576.
- Preisig, H. A., 2021. Ontology-Based Process Modelling-with Examples of Physical Topologies. *Processes 2021*, Vol. 9, Page 592, 9(4), 592.
- Talirz et al., 2020 . Materials Cloud, a platform for open computational science. *Scientific Data 2020 7:1*, 7(1), 1–12.
- Yakutovich et al., 2021. AiiDALab – an ecosystem for developing, executing, and sharing scientific workflows. *Computational Materials Science*, 188, 110165.

# INDEX

## A

- Aas, Vegard 613  
Abarca, Jose Antonio 3059  
Abbasi, Hamid Reza 1037  
Abbas, Safdar 619  
Abdallah, Mohamad 3269  
Abdullah, Fahim 1511  
Abdullatif, Yasser M. 1407  
Abildskov, Jens 575  
Abonyi, János 1377  
Abou El Qassim, M. 1703  
Abraham, Elizabeth J. 2329  
Acevedo, Andres R. 331  
Aceves-Lara, César A. 583  
Acha, Salvador 697  
Acosta-Pavas, Juan C. 583  
Adelgren, Nathan 433  
Adjiman, Claire S. 973, 1153  
Afonso, Lourdes B. 3505  
Aghayev, Zahir 2629  
Ahmadian, Ali 533, 2907  
Ahmed, A. 2499  
Ahmednooh, Mahmoud A. 1781  
Ahmed, Usama 2233  
Ahn, Byeongchan 2239  
Ahn, Jou-Hyeon 3409  
Aizpuru, Judit 2611  
Ajagekar, Akshay 1563  
Akintola, Ayooluwa 3157  
Akkermans, Simen 2653  
Akoh, Benjamin 3157  
Akturk, Ilke 1437  
Akundi, Sahithi Srijana 1227  
Alabid, Maytham 3071, 3217  
Alakent, Burak 607  
Al-Ansari, Tareq 89, 273, 569, 1407, 1425, 1473, 1583, 1589, 1613, 2153, 2159, 2177, 2191, 2295  
Al-Banna, Adnan 1111  
Alba-Robles, Emilio 3451  
Albo, Jonathan 3059  
Alcántara-Avila, J. Rafael 889, 897  
Alcivar-Espinoza, K. 519  
Aleissa, Yazeed M. 3433  
Alexandridis, Alex 1787  
Alhajaj, Ahmed 1007, 1299, 3163  
Alhajeri, Mohammed S. 1511  
Alherbawi, Mohammad 89, 273, 1425, 1473, 2153, 2159, 2177, 2191, 2295, 2301  
Alimirzaei, Fatemeh 2689  
Ali, Moustafa 1839, 2029, 2629  
Ali, Rehannah 2877  
Aliyev, Mammad 1871  
Allam, A. 2795  
Almansoori, Ali 533, 2907  
Almeida, José L. G. 2921, 2945  
Al-Mohannadi, Dhabia M. 2329  
Alnajdi, Aisha 1511  
Alnouri, Sabla Y. 7  
AlNouss, Ahmed 1613, 2159  
Al-Obaidi, Mudhar A. 45  
Al-Qadri, Ali A. 2233  
Alrebei, Odi Fawwaz 1407  
Alsarayreh, Alanood A. 45  
Al-Sobhi, Saad 2117  
Alva-Argaez, Alberto 3249  
Alvarado-Rodriguez, C. E. 411, 825, 1345  
Amaran, Satyajith 3195  
Amaro, Ana 3367  
Amhamed, Abdulkarem I. 1407  
Amiri, Amirpiran 2349, 3169  
Ammari, Bashar L. 1329  
Amorim, Ana Paula Alves 3017  
Anandan, Paul Danny 1665  
Anand, Shikhar 1901  
Anantharaman, Rahul 267, 3011  
Anastasakis, Konstantinos 2227  
Anaya-Ruiz, Guillermo Adolfo 2265  
Andersson, Leif Erik 267, 831, 1543, 3011  
Angeli, Panagiota 1099  
Anglou, Elisavet 2403  
An, Nahyeon 2397  
Anne, Sarath Babu 673, 1013, 2975  
Anpalagam, Sivanesh Kumar 101  
Antos, Dorota 1025



- Aouichaoui, Adem R. N. 575  
 Appiani, Federico 2641  
 Araújo, Cáo C. S. 1821, 1827  
 Araujo, Hamilton 1133  
 Arellano-Garcia, Harvey 2963, 2969,  
 3017  
 Arges, C. G. 1449  
 Armaou, Antonios 801  
 Arnaud, Manon 1931  
 Arnese-Feffin, Elia 163  
 Aroniada, Magdalini 2099, 2105  
 Arora, Pratham 2445, 3373  
 Arshad, Nahyan 1299  
 Arvanitidis, Achilleas L. 77  
 Asensio, Anne 1931  
 Ashraf, Waqar Muhammad 2939  
 Asprión, Norbert 463  
 Asrav, Tuse 227  
 Assawajaruwan, Supasuda 2709  
 Asuni, Oludayo 2895  
 Attarakih, Menwer 2623  
 Augusto, Ana S. 2421  
 Avalos, José L. 2599  
 Aviso, Kathleen B. 3263  
 Avraamidou, Styliani 861  
 Awasthi, Utsav 1557  
 Awwad, Mohammed 2493  
 Aydin, Duygu 227  
 Aydin, Erdal 227  
 Azapagic, Adisa 2993  
 Azzaro-Pantel, Catherine 3209
- B**
- Baader, Florian Joseph 2745  
 Babaei, Masoud 1037  
 Babi, Deenesh K. 399  
 Baboolal, Anastasia 2877  
 Baccouche, Ahmed 2851  
 Back, Seoin 691  
 Bada, Samson 2525  
 Badejo, Oluwadare 3343  
 Badr, Sara 985  
 Bahamon, Daniel 3163  
 Bahamon, Pablo Rodriguez 2611  
 Bailey, Phoebe 813  
 Bai, Yiming 1715  
 Bakas, Nikos 279  
 Bakshi, Bhavik R. 2481, 3433  
 Balhorn, Schulze Lukas 2009  
 Baltazar-Álvarez, José Luis 2265, 2373  
 Bamberg, A. 2055, 2061  
 Bampouli, Ariana 3487  
 Bañares-Alcántara, René 2757  
 Bano, Gabriele 39, 2099  
 Baratsas, Stefanos G. 539  
 Baratti, Roberto 751  
 Barbati, Daniele 915  
 Barbieri, Camilla 2507  
 Barbosa-Povoa, Ana Paula 1283, 3349,  
 3367  
 Barbosa, Vinicius P. 1351  
 Barcia-Quimi, A. F. 519, 2321  
 Bareschino, Piero 1197  
 Barhate, Yash 1499  
 Barolo, Massimiliano 39, 163  
 Barrera-Reyes, Ernesto 2265  
 Barreto, Amaro G. 1833  
 Bart, Hans-Jörg 2623  
 Barton, Ryan 2635  
 Barzilai, Abigail 3475  
 Basha, Nausheen 511  
 Bassani, Gabriel Soares 2067  
 Batet, Lluís 1801  
 Batista, Clyde-Theodore Nguimbous  
 3169  
 Batista, Gustavo 2271  
 Bausa-Ortiz, Irina 721  
 Bawa, Solomon Gajere 1  
 Bayer, Christoph 2379  
 Bechara, Christelle 7  
 Becker, Per Julian 1051  
 Beenfeldt, Christian 2665  
 Beisswenger, Lucien 1057  
 Beleli, Yuri Souza 3231  
 Benfenati, Emilio 2715  
 Benjelloun, Saad 367, 643  
 Benyahia, Brahim 255, 349, 439, 667,  
 1665, 2259  
 Berglihn, Olaf T. 757, 769  
 Bermak, Amine 1589  
 Bernabé-Martínez, Rubén 3299  
 Bernardi, Andrea 1653, 2343, 2433,  
 3403  
 Bernardino, Lucas F. 1697  
 Berrow, James 197  
 Bestwick, Tate 1607

- Bettenbrock, Katja 2599  
Beykal, Burcu 2629  
Bezzo, Fabrizio 39, 163, 2123, 2135  
Bhagaloo, Keeara 2877  
Bhati, Naveen 183  
Bhonsale, Satyajeet S. 1069, 1493, 1769, 2653  
Bianchini, Cosimo 2349  
Biegler, Lorenz T. 1845  
Binns, Michael 1045  
Birgen, Cansu 2141, 2147  
Birkelbach, Felix 113  
Bisgaard, Thomas 2665  
Bisotti, Filippo 757, 769  
Bispo, Artur S. 2921, 2945  
Biundo, Antonino 2561  
Bi, Xiaolin 763  
Bizon, Katarzyna 1197  
Blackburn, Daniel 3269  
Blagov, Sergej 463  
Blakseth, Sindre Stenen 831  
Blanco, Elena C. 1967  
Blank, Lars M. 2683  
Blekos, Kostas 1251  
Blom, Richard 313  
Boix, Marianne 2883  
Bollas, George M. 733, 1557, 3041  
Boly, Vincent 1985  
Bommineni, Praveen Kumar 673, 2981  
Bonetti, Darlan 1743  
Bonilla-Petriciolet, Adrian 1257  
Bonino, Sergio G. 3317  
Bontron, Camille 3379  
Bornemann, Luka 2865  
Bortz, Michael 463  
Bouallou, Chakib 367  
Bouchkira, Ilias 349, 439, 643  
Boukharfane, Radouan 139, 175, 367  
Boukouvala, Fani 1311, 2403  
Bounaceur, Roda 625  
Bounitsis, Georgios L. 739, 3047, 3385  
Bowskill, David H. 1153  
Boys, Ian 601  
Bozek, Ewa 1191  
Braeken, Leen 457  
Brandner, Dean 595  
Braz, Catarina G. 2421, 3379  
Bregy, Michael 2487  
Bremer, Jens 1487  
Bresciani, Antonio Esio 2067  
Bretado-Aragón, Luis Alberto 2265  
Briceno-Mena, L. A. 1449  
Brilman, Wim 2307  
Brown, Solomon 1075, 1461, 2889, 2895, 3139  
Brown, Solomon F. 1431, 2647, 3083  
Bublitz, Saskia 119  
Buczko, Aleksandra 3487  
Bugryniec, Peter 1075  
Bui, Mai 2087  
Burtea, Radu-Alexandru 1641  
Busse, Corinna 1191  
Bussemaker, Madeleine 2677, 3361
- C**  
Caballero, José A. 1925, 3427, 3457  
Cabezas, Heriberto 2469  
Cabrera-Ruiz, Julián 889, 897  
Caccavale, Fiammetta 3469  
Cafaro, Diego C. 3189  
Caçada, Luís Américo 1351  
Calixto, Ewerton E. S. 2921, 2945  
Calvo-Serrano, Raul 2487  
Camarda, Kyle V. 1607  
Cammann, Lucas 3145  
Canchola-López, Perla G. 1345  
Cao, Yankai 1337  
Carabineiro, Hugo M. D. 949  
Cardenas, Cristian 505  
Cardenas, Ilya Orson Sandoval 83  
Cárdenas, Milton D. 1147  
Carrasco, Francisco 267  
Carvalho, Ana 3349  
Casán, Daniel Bagan 2343  
Casas-Orozco, Daniel 1167, 1437  
Casella, Francesco 915  
Castro-Amoedo, Rafael 929  
Castro-Hernández, Salvador 897  
Cattaneo, Simone 3243  
Cecelja, Franjo 2677, 3361  
Cecílio, Duarte M. 2421  
CELSE, Benoit 1051  
Cenci, Francesca 1  
Cezareth, Pedro 19  
Chabaud, Valentin 1543  
Chachuat, Benoit 2215, 2343, 2433

- Chaitanya, Vuppanapalli 3121  
 Chakrabarti, Dhurjati 2901  
 Chakraborty, Arijit 1087  
 Chakraborty, Souvik 1895  
 Chandra-Ambhorn, Walairat 1907  
 Chandrasekaran, Sivaramakrishnan  
     3053, 3391  
 Chang, Jo-Shu 2579  
 Chang, Yuchen 2403  
 Charalambous, Margarita A. 2927  
 Charitopoulos, Vassilis M. 739, 1159,  
     1601, 3023, 3047, 3385  
 Chaton, Nicolas 1931  
 Chaturvedi, Nitin Dutt 3287  
 Chatziasteriou, Christos C. 2769  
 Chatzidoukas, Christos 1467  
 Cházaro-Ruiz, Luis Felipe 2315  
 Chebeir, Jorge 751  
 Chen, Bingzhen 247  
 Chen, Chang 131  
 Cheng, Xi 1979  
 Chen, Tao 57, 209, 2165  
 Chen, Wei-Han 2845  
 Chen, Xi 801, 1735  
 Chen, Yaoxi 3101  
 Chia, Dian Ning 235, 1919, 3463  
 Chianese, Simeone 1961  
 Chibeles-Martins, Nelson 3505  
 Chiesa, Paolo 3095  
 Chin, Matthew 2703  
 Cho, Hyungtae 1395, 1413, 2397, 2539  
 Choi, Soo Hyoung 1081  
 Chonghyo, J. 881  
 Choubert, Jean-Marc 703  
 Chrisandina, Natasha J. 2951  
 Christodoulou, Charalampos 39, 2099,  
     2105  
 Christodoulou, Christos 279  
 Christofides, Panagiotis D. 71, 1511,  
     1517  
 Chuang, Yao-Chen 297  
 Cieno, Fabio 2123  
 Citmaci, Berkay 1517  
 Clements, Chris 3269  
 Cloete, Schalk 313  
 Cockx, Arnaud 583  
 Cogliati, Alessandro 575  
 Collas, Alain 1493  
 Çölmekci, Ceylan 3089  
 Colombo, Gabriele 2283  
 Colombo, Giovanni 2641  
 Continillo, Gaetano 1197  
 Contreras-Zarazúa, Gabriel 145, 1317,  
     1529  
 Coppens, Marc-Olivier 2703  
 Cordiner, Joan 1431  
 Cormos, Ana-Maria 1365, 1691, 2775  
 Cormos, Calin-Cristian 1365, 1691,  
     2775  
 Corradini, Leonardo L. 2361  
 Corrales, David Camilo 583  
 Corriou, Jean-Pierre 589  
 Cortez-González, Jazmín 825, 1345  
 Cosgun, Ahmet 227  
 Cos, Ralph 2751  
 Cots, Olivier 709  
 Couenne, Françoise 703  
 Coupan, Romuald 2851  
 Coutu, Arnaud 1931, 3481  
 Cowley, Thomas 1461  
 Cremaschi, Selen 475, 1203, 1323,  
     1481, 1505, 2721  
 Cristea, Smaranda P. 721  
 Cristea, Vasile Mircea 1877  
 Cristiu, Daniel 2123  
 Cruz Bournazou, Mariano Nicolas  
     2611  
 Čuček, Lidija 979  
 Cutillo, Enrico A. 1197
- D**
- da Cruz, Antonio José Gonçalves  
     2271  
 da Cruz Meleiro, Luiz Augusto 1351  
 Dai, Jindong 1549  
 Dalamagas, Theodoros 3439  
 d'Amore, Federico 2123, 2739  
 D'Angelo, Sebastiano C. 2487  
 Daoutidis, Prodromos 655  
 da Paz, Priscila Marques 2671  
 da Silva, Cátia 3349  
 da Silva, Luciana A. 1815  
 da Silva Osório, Tiago 1815  
 Da Soghe, Riccardo 2349  
 de Angelis, Alberto R. 3243  
 De Blasio, Cataldo 241, 1961

- de Brito Alves, Rita Maria 2067  
 Decardi-Nelson, Benjamin 1659  
 De Carfort, Johan Le Nepvou 775  
 de Carvalho Servia, Miguel Ángel 33,  
 795  
 De Clercq, Dries 457  
 De Guido, Giorgia 1443  
 De-León Almaraz, Sofia 3209  
 Deliismail, Ozgun 679  
 del Rio Chanona, Antonio 511, 1647,  
 1801  
 del Rio-Chanona, Ehecatl Antonio 33,  
 83, 469, 649, 2659  
 Demandt, Sören 2745  
 De Mel, Ishanki 3397  
 de Moraes, Marcellus G. F. 1833  
 Deng, Yushi 1323  
 de Paiva, José Luis 3231  
 de Prada, Cesar 499, 721, 1271  
 Deshpande, Parag A. 2015  
 de Souza Jr., Maurício B. 1351, 1833  
 Deussen, Philipp 789  
 Dhanraj, Rashi 3237  
 Diab, Samir 39, 2099  
 Diakoumi, Eleonora 95  
 Diangelakis, Nikolaos A. 1709  
 Dias, Ricardo N. 1117  
 Diaz-Damacillo, Lamberto 411  
 Diaz, Maria Soledad 355, 2073  
 Díaz-Sainz, Guillermo 3059  
 Di Caprio, Ulderico 457, 3281  
 Dikaiakos, Antonios 95  
 Di Legge, Giampiero 2349  
 Di Martino, Marcello 51  
 Dimitriadis, Christos N. 2763  
 Dinca, Cristian 2775, 3071, 3217  
 Dinuccio, Elio 3133  
 di Pretoro, Alessandro 2439, 3329  
 Dirza, Risvan 613  
 Diwekar, Urmila 2469  
 Doe, Symon 279, 285  
 Doganis, Philip 1595  
 Dokl, Monika 979  
 Dologlu, Pelin 685  
 Domingos, Meire Ellen Gorete Ribeiro  
 1233  
 Domingos, Meire Ribeiro 3175  
 Domínguez-Ramos, Antonio 2277  
 Donthula, Kiran 1013  
 Dorneanu, Bogdan 2963, 2969, 3017  
 Dorne, Jean-Lou C. M. 2715  
 dos Santos, Gustavo de S. 2945  
 dos Santos, Moisés Teles 1233  
 Doyle, Stephen J. 1031  
 Dragan, Mihaela 1365  
 Dragan, Simion 1365, 1691, 2775  
 Drouven, Markus G. 1859  
 Duanmu, Fanyi 235, 1919, 3463  
 Duarte, Belmiro P. M. 1019  
 Duarte, Inês 3367  
 Dua, Vivek 2939  
 Dube, Samukeliso 2801  
 Du, Binglin 425  
 Dumas, Claire 583  
 Duran, Jairo 3249  
 Dürra, Robert 1025  
 Dürr, Robert 2221  
 Duvigneau, Stefanie 2221  
 Du, Wenli 1635  
 Duyar, Melis 197
- E**
- Eden, Mario R. 475, 1203, 1323  
 Efthymiadou, Margarita E. 3023  
 Eijssink, Vincent G. H. 757, 769  
 Ejeh, Jude O. 1431, 3083  
 El-Halwagi, Mahmoud M. 2951  
 Eljack, Fadwa 2117  
 Elkamel, Ali 533, 2907  
 Elkarii, Marwane 367  
 Elmaz, Furkan 457  
 Elmisaoui, Safae 175  
 Elmisaoui, Sanae 643  
 Elmocayd, Nabil 367  
 Elomri, Adel 89  
 Elsheikh, Mohamed 1751  
 Engell, Sebastian 1127, 1751  
 Enriquez-Posligua, J. 2321  
 Errington, Ethan 2253  
 Erturk, Emrullah 679  
 Esche, Erik 119, 595  
 Esenboga, Elif Ecem 227  
 Espinel-Ríos, Sebastián 2599  
 Espinoza-Vázquez, Yulissa M. 2733  
 Espuña, Antonio 961, 1703, 2289,  
 2533

Esquivel-Barajas, Gabriela Guadalupe  
2265, 2373  
Estrada, Vanina 355, 2073  
Eterigho-Ikelegbe, Orevaoghene 2525  
Èuèèk, Lidija 3421  
Evripidou, Nikola 1099  
Ez-zahra El Hamra, Fatima 139

**F**

Faaij, André P. C. 3077  
Facco, Pierantonio 1, 163  
Fakhraddinfakhriazar, Salar 1179  
Fakid, Sadman 2519  
Falk, Véronique 1985  
Fan, Xiaolei 1955, 1979  
Farinas, Cristiane Sanchez 2271  
Farnsworth, Amanda 2819  
Farrusseng, David 505  
Fasse, Sylvia 2463  
Fechtner, Marcus 343  
Fedeli, Matteo 2439  
Felices-Rando, Beatriz 2135  
Feller, Rolf 1173  
Femat, Ricardo 2315  
Fermiglia, Maurizio 1949  
Fernández-González, Javier 2277  
Ferrari, Adrián 291  
Ferreira, Jimena 1105  
Ferreira, Newton Libanio 2067  
Figueiredo, Maurício M. F. 1821,  
1827  
Figueredo, Manuel 1147  
Fikar, Miroslav 1001  
Filbig, Melanie 2683  
Fileti, Ana M. F. 1743, 1821, 1827  
Filipe, Rui M. 1117, 1257  
Findeisen, Rolf 2599  
Finn, Adrian J. 1031  
Fiorini, M. Josefina 3189  
Firstauthorm, Anne 525  
Fischer, Martin 215  
Fleiger, Kristina 267  
Flores-Mendoza, Omar 2367  
Flórez-Orrego, Daniel 1233, 3175  
Foglia, J. 2499  
Fontoura, Tahyna B. 1815  
Foo, Dominic C. Y. 3263  
Forina, Mattia 2547

Forster, Tim 1239  
Foulkes, Lesley 197  
Fraga, Eric 3269  
Fragkopoulos, I. S. 553  
Franke, Meik 1851, 3493  
Frinklea, Ferdous 2721  
Friso, Andrea 13  
Fritz, Laura J. 355  
Fröhlich, Jochen 325  
Fuchs, Uwe 125  
Fujii, Shoma 2475, 3035  
Fung, Jesslyn 3403

**G**

Gadekar, A. 1985, 2499  
Galeazzi, Andrea 745, 1277, 2533  
Galindo, Amparo 973  
Gallage Dona H. K. 1449  
Gallo, Francesco 1277  
Gallucci, Fausto 2457  
Galvan-Alvarez, Nicolas 1305  
Galvanin, Federico 1, 13, 631, 789,  
1099, 3269  
Gando-Ferreira, Licínio M. 1889  
Ganesan, Arvind 2403  
Gani, R. 221, 2499  
Gänsch, Jonathan 1389  
Gao, Qinghe 2003  
Gao, Shang 2259  
Garcia-Castro F. L. 3427  
Garcia Garcia-Verdier, Tomas 1271  
García-Trejo, Juan Fernando 337  
Gargalo, Carina 2635  
Gargalo, Carina L. 3469  
Ga, Seongbin 2397  
Gautam, Vinay 3517  
Gavriliadis, Asterios I  
Gawande, Gayatri 3373  
Gazzani, Matteo 3077  
Gedraite, Rubens 1351  
Geissler, Caleb H. 3311  
Gençer, Emre 1139, 2819, 2831  
Georgiadis, Georgios P. 95, 107  
Georgiadis, Michael C. 77, 95, 2763,  
2769  
Gerbaud, Vincent 1937  
Geremia, Margherita 39  
Gergaud, Joseph 709

- Gernaey, Krist V. 2635, 3469  
 Ghasemi Naraghi, Saba 3335  
 Ghavami, Niloufar 1961  
 Ghiat, Ikhlal 1583  
 Ghidaglia, Jean-Michel 175  
 Ghilardi, Lavinia Marina Paola 915  
 Ghosh, Snigdha 715  
 Gietmann, Delia 1191  
 Gilardi, Matteo 757, 769  
 Gilleskie, Gary 2635  
 Gkizas, George 2813  
 Glaude, Pierre-Alexandre 625  
 Goldhahn, Ruben 2391  
 Golla, Christian 325  
 Gomez, Camilo 1859  
 Gómez-Castro, Fernando Israel 169,  
     337, 1913, 2733, 3293, 3451  
 Gomez, Jorge M. 1859  
 Gonçalves, Bruno S. 1283  
 González-De-La-Cruz, Sérgio 1257  
 Gonzalez, Marcial 813, 1577  
 Gonzalez, Paola Ibarra 2451  
 Goswami, Umang 1883, 1895  
 Gott, David 2715  
 Gottschalk, Axel 2463  
 Govindan, Rajesh 569, 1583, 1589  
 Graells, Moisés 961, 2289  
 Graf, Rafael 3203  
 Granacher, Julia 929, 1233  
 Grénman, Henrik 715  
 Griffin, Sholeem 2653  
 Grizeau, Alexandre 589  
 Grossi, Caroline Dias 1351  
 Grossmann, Ignacio E. 247, 1371,  
     1859, 3317  
 Grytli, Tuva 2141, 2147  
 Gu, Boram 2397  
 Gu, Chenxi 1549  
 Guillén-Gosálbez, Gonzalo 1239,  
     2487, 2927  
 Guo, Miao 2253, 2715  
 Gupta, Nikita 1901  
 Gupta, Rishabh 1357  
 Gutiérrez-Antonio, Claudia 337, 3299  
 Gutierrez, Gloria 1271  
 Gutiérrez-Guerra, Roberto 1345  
 Gutiérrez-Limón, Miguel A. 3065  
 Gutierrez, Maria F. 2933  
 Gutiérrez, Soledad 291  
 Gu, Xuerong 943
- H**
- Ha, Byeongmin 3409  
 Hachhach, Mouad 405  
 Hagedorn, Dörthe Franzisca 2745  
 Haider, Md Alquma 3287  
 Haji, Maryam 273, 1425, 2153  
 Halba, Ankush 2445, 3373  
 Hall, Keegan K. 2379  
 Hamedí, Nazanin 1685  
 Han, Chengyu 1549  
 Han, Dongho 3005  
 Han, Yueheng 921  
 Hanzelik, Pál P. 1377  
 Haralambos, Sarimveis 1595  
 Harjunkoski, Iiro 3005  
 Harris, Zoe 2165  
 Hashemi, Mohammadjafar  
     2721  
 Hassna, Bashar 89  
 Hayashi, Yusuke 65  
 Hayat, Khizar 3163  
 Hazrat, Bilal 569  
 Hecht, Felix 1751  
 Heide, Matti 325  
 Heiduk, Caroline 837  
 Heiho, Aya 2361, 2475  
 Heim, Dennis Manuel 463  
 Held, Christoph 955  
 Hellgardt, Klaus 33  
 Hellinckx, Peter 457, 3281  
 Hemrle, Jaroslav 3203  
 Heng, Jerry 2253  
 Henrichfreise, Lars 1173  
 Hernández-Camacho, Nereyda Vanessa  
     3293  
 Hernández, Salvador 337, 1913, 3299  
 Hernández-Vargas, Esteban Abelardo  
     1529  
 Herrera-Velázquez, Josué J. 889, 897  
 Herwig, Christoph 481, 487, 909, 2605  
 Hii, King Kuok (Mimi), 33  
 Hiltunen, Eero 319  
 Hirtreiter, Edwin 2009  
 Hoch, Patricia M. 2415  
 Hodnett, Neil 2099

Hoffmann, Christian 393  
 Hofmann, René 113, 215  
 Hogstrand, Christer 2715  
 Hong, Fugui 131  
 Hong, Xiaodong 2957  
 Horler, Mark 2165  
 Horsch, Anna S. 331  
 Hørsholt, Steen 1775  
 Horst, Nora 2605  
 Hua, Kaixun 1337  
 Huang, Yan-Shu 813  
 Huber, David 113  
 Huerta-Rosas, Brenda 1317  
 Hu, Guoqing 2203  
 Hui, Wai Fung 2573  
 Hur, Inyoung 1167  
 Huskova, Nadiia 1389  
 Hutton, Timothy 1461, 3139  
 Huusom, Jakob Kjøbsted 361, 1087,  
 1775, 2573, 2635  
 Hwangbo, Soonho 3115, 3409  
 Hyungtae, C. 881

**I**

Iakovou, Eleftherios 2951  
 Iapteff, Loic 1051  
 Ibarra-Gonzalez, Paola 1209  
 Ierapetritou, Marianthi 2183, 3343  
 Igwe, Chika Linda 481  
 Ikonen, Teemu J. 3005  
 Ilea, Flavia-Maria 1691  
 Illner, Markus 393  
 Interlenghi, Stefano Ferrari 2067  
 Ioannidou, Sofia-Maria 2513  
 Ipkovich, Ádám 1377  
 Iqbal, Mohd Umair 1729  
 Irabien, Angel 2277, 3059  
 Iraola, Eduardo 1801  
 Irhamna, Adrian R. 3041  
 Isella, Andrea 2283, 2547  
 Iseri, Funda 539, 2629  
 Iten, Muriel 2783

**J**

Jackson, George 973  
 Jaewon, L. 881  
 Jafary, Nur Fitriyanni 101  
 Jafari, Soheil 2349

Jagana, Jnana Sai 3195  
 Jallut, Christian 703  
 Jang, Joon Baek 1517  
 Jang, Shi-Shang 451  
 Jankoviæ, Tamara 2111  
 Jäschke, Johannes 1723, 1763, 3145  
 Jayaraman, Suresh Kumar 1185  
 Jena, Sankalp 2709  
 Jensen, Anker D. 1775  
 Jiang, Zhen-Feng 297  
 Jiang, Zheyu 819, 1621, 3335  
 Ji, Cheng 131, 1549, 1627  
 Jimbo, Roberto Xavier Jimenez 649  
 Jiménez, Laureano 1925  
 Jiménez-Romero, Julia 2993  
 Jobson, Megan 2021  
 John, Yakubu M. 783  
 Jones, Marie 1233  
 Jonghun, L. 881  
 Joo, C. 1395, 2539  
 Jooss, Yannick 3011  
 Jørgensen, John Bagterp 1775  
 Jul-Rasmussen, Peter 361, 1087  
 Junghwan, K. 881  
 Jurado, Nelia 3269  
 Just, Daniel 3487

**K**

Kabak, Özgür 727  
 Kado, Shigeru 1973  
 Kafarov, Viatcheslav 1305  
 Kager, Julian 487, 2605  
 Kakaodkar, Rahul 1839  
 Kakunuri, Manohar 673, 1013  
 Kaldellis, J. K. 2129  
 Kalian, Alexander D. 2715  
 Kalogeropoulos, Ioannis 1787  
 Kanematsu, Yuichiro 2361, 2475, 3035  
 Kang, Jia-Lin 297, 451  
 Kang, Yung-Shun 1499  
 Kano, Manabu 151  
 Kaps, Lothar 561  
 Kardamaki, Argyri 1595  
 Karimi, I. A. 1291  
 Karpov, I. 2499  
 Kassianides, Christoforos 279  
 Kassianides, Symeon 279, 285  
 Kato, Shota 151

- Katsini, Lydia 2653  
 Katsoulas, Konstantinos 3463  
 Kavakiotis, Ioannis 2677  
 Kaven, Luise F. 807  
 Kavousanakis, Mihalis 1595  
 Kaya, Gizem Kuşoğlu 727, 3089  
 Kayahan, Emine 3281  
 Kayar, Gizem 685, 1871  
 Kay, Harry 1523, 2593  
 Kay, Sam 1523, 2593  
 Kazepidis, Panagiotis 2807, 3109  
 Kazi, Monzure-Khoda 2117  
 Keates, Adam 1153  
 Keestra, Hilbert 2307  
 Kelly, Jeffrey D. 1111, 1781  
 Kemmerling, Simon 3487  
 Kenefake, Dustin 1227, 1839, 2029  
 Keßler, Tobias 343, 561  
 Khaitan, Pratham 2445  
 Khalafalla, Siddig S. 1991  
 Khaled, Houda 2105  
 Khalil, Roger 3225  
 Khamar, Lhachmi 175  
 Khan, Obaid 3397  
 Khaydarov, Valentin 2709  
 Khumalo, Nomadlozi L. 1419  
 Kienle, Achim 343, 561, 1025, 2221  
 Kieslich, Chris A. 2689, 2695  
 Kikkinides, Eustathios S. 2769  
 Kikuchi, Yasunori 2361, 2475, 3035  
 Kilari, Hemalatha 1499  
 Kim, J. 2539, 2987  
 Kim, Jong Woo 2611  
 Kim, Junghwan 1395, 1413, 2397  
 Kim, Minsu 849  
 Kim, So-mang 1179  
 Kim, Taehyun 3409  
 Kim, Y. 1413, 2539  
 Kiparissides, Costas 637  
 Kiss, Anton A. 1031, 2111, 3493  
 Klamt, Steffen 2599  
 Kleingesinds, Eduardo Krebs 2665  
 Klein, Peter 2093  
 Knuutila, Hanna K. 3257  
 Koch, A. 2055, 2061  
 Kodamana, Hariprasad 1455, 1883, 1895, 1901  
 Köhler, Hannes 325
- Ko, Jaerak 3115  
 Kokossis, Antonis 1653, 2561, 2567, 2727, 3439, 3445  
 Koksal, Ece Serenat 227  
 Komodromos, Charis 285  
 Konchakova, Natalia 2093  
 Kondili, Emilia M. 2129, 3323  
 Kontoravdi, Cleo 375, 381, 469, 1997, 2659  
 Kookos, Ioannis K. 553, 2513  
 Kordatos, Ioannis 1787  
 Korichi, Mourad 661  
 Kostoglou, Margaritis 77  
 Koulouris, Alexandros 107  
 Koutinas, Apostolis 2513  
 Kramer, Gert Jan 3077  
 Krämer, Stefan 1751, 3493  
 Kravanja, Zdravko 979, 3421  
 Krishnamoorthy, Dinesh 613  
 Krishnamurthy, Shreenath 313  
 Krog, Halvor A. 1723  
 Kroll, Mariann 955  
 Kruber, Kai Fabian 463, 955  
 Krühne, Ulrich 775, 3469  
 Kumar, Avan 1455  
 Kumar, Deepak 1883, 1901  
 Kumar, R. 1449  
 Kummer, Alex 1377  
 Kunde, Christian 561  
 Kurban, Sena 727  
 Kushwah, Anamika 203  
 Kusoglu, Gizem 227  
 Kwon, Hyukwon 1395  
 Kyriakides, Alexios-Spyridon 2807
- L**  
 Labarta, Juan A. 1925, 3457  
 Ladakis, Dimitrios 2513  
 Lagare, Rexonni B. 813, 1577  
 Lagoutière, Frédéric 703  
 Lahlou, Fatima-Zahra 569  
 Lainez, Jose 3317  
 Laird, Carl D. 1329, 1859  
 Lakelin, Matthew 1653  
 Lal, Apoorv 2837  
 Lampe, Matthias 1851  
 Lane, Amanda 1523  
 Lange, Hannes 1265



- Lapiedra Carrasquer, Pau 1769  
 Lara-Montaño, Oscar D. 3451  
 Latifi, Abderrazak 625  
 Latifi, Abderrazak M. 439, 589, 643  
 Lazaro-Molina, Brenda I. 2337  
 Le bivic, Réjanne 3481  
 Leblebici, M. Enis 457, 3281  
 Lechtenberg, Fabian 961  
 Lee, Byungjun 861  
 Lee, Chul-Jin 2385, 3127  
 Lee, Donggyun 849  
 Lee, Jaewon 1395, 1413  
 Lee, Jay H. 3005  
 Lee, Ming Hei 3403  
 Lee, Moonyong 875  
 Lee, Tim B. Jaewook 997  
 Lee, Wonbo 493  
 Lee, Yena 3385  
 Le Jemtel, Antoine 1985  
 Lemire, Mélissa 1793  
 Léonard, Grégoire 1179  
 Leong, Yoong Kit 2579  
 Leonhard, Kai 807  
 Leonzio, Grazia 2215  
 Le Roux, Galo A. C. 1723, 2079, 2671,  
 3231  
 Lesage, François 589  
 Lewin, Daniel R. 3475, 3493  
 Liang, Runze 921  
 Lian, Guoping 2165  
 Liang, Xiaodong 361, 1087  
 Liao, Yangyanbing 1979  
 Liao, Zuwei 2957  
 Li, Chenxi 25  
 Li, Dan 261  
 Liebmann, Vera 325  
 Li, Hanxiu 935  
 Li, Haoran 305  
 Li, Jie 261, 1955, 1979  
 Li, Liang 1769  
 Lima, Fernando Arrais R. D. 1833  
 Lima, Rodrigo M. 1815  
 Lim, J. 1413, 2539  
 Lindahl, Simon B. 399  
 Linke, Patrick 51, 2329  
 Lin, Xiaowen 801  
 Lipke, Elizabeth 2721  
 Li, Qing 1031  
 Liski, Antti 319  
 Lista, Alberto 2283  
 Liu, a Pei 425  
 Liu Chia-Hsiang 451  
 Liu, Chong 1549  
 Liu, Guilian 2859  
 Liu, Jay J. 493  
 Liu, Jialin 2579  
 Liu, Jiaxu 667  
 Liu, J. Jay 2239  
 Liu, Pei 25, 189, 417, 547, 3275  
 Liu, Rui 801  
 Liu, Xinying 2801, 2825  
 Liu, Yu-Ting 57  
 Li, Xiang 1671  
 Li, Zheng 25, 189, 417, 425, 547, 3275  
 Llewelyn, Philip 313  
 Logist, Filip 1069  
 Lopez-Arenas, Teresa 2337, 2367  
 Lorenz, Heike 1389, 2933  
 Louis, Pauline 3481  
 Louloudi, Sofia 285  
 Ľubušký, Karol 1001  
 Luca, Alexandra-Veronica 1877  
 Luceño Sánchez, José A. 3029  
 Lucia, Sergio 595  
 Luderer, Lynn 2009  
 Lugo-Leyte, Raúl 3065  
 Luna, Martin F. 2611  
 Luo, Junwei 1517  
 Luo, Yuqing 2183  
 Lyons, Ben 2433  
  
**M**  
 Macchietto, Sandro 3029  
 Mace, Lucie 1931  
 Mackey, Hamish R. 569  
 Mackey, Jaron 1437  
 Madayag, Jan Vincent 1571  
 Madbhavi, Rahul 2041  
 Mädler, J. 1215, 1265, 2055, 2061  
 Maduako, Adaeze C. 733  
 Ma, Fangyuan 131  
 Magnússon, Atli F. 1221  
 Mahmood, Farhat 1589  
 Manca, Davide 2283, 2547, 2641  
 Mancini, Michael A. 2629  
 Mancusi, Erasmo 1197

- Mandis, Marta 751  
Manenti, Flavio 745, 1277, 2289, 2409,  
2439, 2533  
Mangold, M. 1389, 2795  
Mann, Vipul 221  
Mansouri, Seyed Soheil 2665  
Mansour, Khaled 661  
Maravelias, Christos T. 433, 445, 3311  
Marchetto, Alberto 3463  
Marcilla, Antonio F. 3457  
Marcoulaki, Effie 1251  
Maréchal, François 183, 929, 1063,  
1233, 3175, 3203  
Marias, Frédéric 3133  
Marinangeli, Pablo 2415  
Marinho, Chrislaine B. 2921, 2945  
Marousi, Asimina 1601  
Martelli, Emanuele 915, 1179, 3095  
Martensen, Carl Julius 561  
Martinez, Ernesto C. 2611  
Martínez-Guido, Sergio Iván 3299  
Martínez-Herrera, G. 825  
Martínez-Lomovskoi, Adrián 241  
Martín, Mariano 1967, 2197, 2871,  
3029, 3293  
Martin, Philip 1523  
Martins, F. G. 1227  
Martin, Írem 727  
Martynov, Sergey B. 3083  
Mashi, Rocky 3209  
Matamoros, Fatima 625  
Matar, Omar K. 511  
Mateus, Maria M. 2421  
Mathias, Marcelo 3249  
Matias, Jose 1865  
Matichecchia, Angelo 3243  
Matos, Henrique A. 949, 1117, 2421,  
2783, 3379  
Mauricio-Iglesias, Miguel 2247  
Mavukwana, Athi-enkosi 1401  
Mazzei, Lorenzo 2349  
Mazzetti, Marit Jagtøyen 831  
McKay, Gordon 1473, 2153, 2159,  
2191, 2295  
Mechleri, Evgenia 2969  
Medrano-García, Juan D. 2927  
Mehling, Tanja 2047  
Mei, Chentao 1735  
Mendez, Carlos 1271  
Méndez, Carlos A. 3317  
Méndez-Cruz, Ladislao E. 3065  
Mendoza, Cesar 1523  
Menezes, Brenno C. 1111, 1757, 1781  
Meng, Qingbo 1665  
Mercelis, Siegfried 457  
Merino-Garcia, Iván 3059  
Mero-Benavides, M. 2321  
Mexis, Konstantinos 2727, 3439  
Mhamdi, Adel 807, 1173  
Miah, Sayeef 2969  
Michel, Justine 1931  
Micunovica, Sanja 285  
Migo-Sumagang, Maria Victoria 3263  
Miguel Ángel 33  
Mihály, Norbert-Botond 1877  
Milton, Robert 2647  
Milton, Robert A. 1431  
Minor, Ann-Joelle 2391  
Mio, Andrea 1949  
Mione, Federico M. 2611  
Misra, Swayam P. 2015  
Mitrai, Ilias 655  
Mitsos, Alexander 807, 1173, 2683  
Mock, Michael 1265  
Mohammed, Sadah 2117  
Moioli, Stefania 2507, 3257  
Mojto, Martin 1001  
Mok, Dong Hyeon 691  
Mondino, Giorgia 313  
Monnier, Hubert 625  
Montañés, Rubén Mocholí 831  
Montastruc, Ludovic 2439, 2883, 3305,  
3329  
Monteiro, Mariana 375  
Montesa, Daniel 499  
Moolya, Akash 3317  
Moon, Il 849, 2987  
Morales-Guio, Carlos 1517  
Morales-Gutierrez, Rosa Angelica  
1317  
Morales-Rodriguez, Ricardo 843, 1093  
Morchain, Jérôme 583  
Moreira, Maria Teresa 2513  
Moreno, J. 2499  
Mores, Wannes 1069  
Morgan, Joshua C. 1329

Morud, John 313  
 Moshammer, Kai 2087  
 Mostafaei, Hossein 319  
 Mostafa, Mahmoud 2915  
 Mota, Bruna 3367  
 Moulines, E. 1703  
 Mowbray, Max 795, 1523  
 Muir, Ryan 2751  
 Mujtaba, Iqbal M. 45, 783  
 Mukta, Chinmoy B. 1203  
 Müller, Don Fabian 487  
 Mulopo, Jean 2525  
 Munia, Amanda 2301  
 Muñoz, Hernán D. 393  
 Muñoz, Raul 721  
 Murakami, Sei 985  
 Murillo-Alvarado, Pascual Eduardo  
     2265, 2373  
 Murillo-Andrade, Edgar I. 1093  
 Murrieta-Dueñas, R. 825, 1345  
 Murzin, Dmitry Yu. 405  
 Musmarra, Dino 1961  
 Mutlu, Ilhan 1685  
 Mwengula, Gentil Kahilu 2525

**N**

Nabavi, Seyed Ali 2349  
 Naşcu, Ioan 1635  
 Nagy, Zoltan K. 813, 1167, 1437,  
     1499, 1577, 1943  
 Naik, Sakshi 1845  
 Naim, Nashat 2301  
 Nair, Purusothmn Nair S Bhasker 1383  
 Nair, Sankar 2403  
 Na, Jonggeol 493, 691  
 Nakama, Caroline Satye Martins 2671  
 Nakamura, Takuma 2475  
 Namany, Sarah 89, 273, 1425, 2153  
 Nam, Seolji 3115  
 Naquash, Ahmad 875  
 Narasimhan, S. 3121  
 Narciso, Diogo A. C. 1227  
 Nazeeruddin, Mohammad Khaja 183  
 Nazemzadah, Nima 2665  
 Ncube, Selusiwe 2825  
 Negny, Stéphane 3305, 3329  
 Negri, Francesco 1277, 2409  
 Negri, Valentina 2927

Nehmeh, Mohamad 1937  
 Nemet, Andreja 3421  
 Nevolianis, Thomas 807  
 Ng, Kok Siew 1383  
 Nikolaidis, Theoklis 2349  
 Nikolakopoulos, Athanassios 1595  
 Nikolic, Igor 1209  
 Nisal, Apoorva 2469  
 Nomura, Yuki 2475  
 Nougues, José M. 1801  
 Nova, Anna 1277, 2409

**O**

Obar, Farah 2191  
 Oberleitner, Thomas 909  
 Odemis, Sadik 1871  
 O'Dwyer, Edward 697  
 Ohara, Satoshi 2361  
 Oishi, Kota 985  
 Okamura, Kozue 985  
 Okibe, Maureen Chiebonam 3361  
 Oktay, Emirhan 3089  
 Olayiwola, T. 1449  
 Oliveira, Miguel Castro 2783  
 Oliveira, Nuno M. C. 949, 1019  
 Oliveira, Rafael D. 1723  
 Onaizi, Sagheer A. 2171  
 O'Neill, Eric G. 3311  
 Ordóñez, Diego Freire 2087  
 Orjuela, Álvaro 393  
 Orkoulas, Gerassimos 71  
 Ortmanns, Yak 1751  
 Osborne, Olivia J. 2715  
 Ostuni, Raffaele 2283  
 Othman, Mohamad Rizza 101  
 Ottavi, Sébastien 3481  
 Otto, Eric 1025  
 Ou, Feiyang 71  
 Ouseleye, Abdifatah 1133  
 Ovalle, Daniel 1859  
 Owens, James 2831  
 Özdenkçi, Karhan 1961

**P**

Pably, Philipp 487  
 Pacheco-López, Adrián 2289, 2533  
 Padma, Prarana 387  
 Padole, Manjusha C. 2015

- Pajander, Jari P. 1221  
Palazzo, Roberto 915  
Pandhal, Jagroop 2647  
Pankajakshan, Arun 1  
Pan, Qiong 1955  
Pantelides, Constantinos C. 1153  
Papadokostantakis, Stavros 619  
Papadopoulos, Athanasios I. 2807,  
2813, 3109  
Papageorgiou, Lazaros G. 739, 1601,  
3023, 3385  
Papageorgiou, Thomas 3439, 3445  
Papaiakovou, Stavros 1653  
Papapostolou, Christiana M. 2129,  
3323  
Papastratos, Stelios 2751  
Papathanasiou, Maria M. 381, 1653,  
1997, 2789, 2901, 3403  
Pappas, Iosif 1709  
Paredes-Ortiz, Jose Alfredo 169  
Parris, Vasileios 1467  
Parker, Robert 1845  
Park, J. 1413, 2539  
Parschau, Johannes 2683  
Parthasarathy, Prakash 1473, 2295  
Patel, Bilal 1419  
Patel, Raj 783  
Patnaikuni, Venkata Suresh 2981, 2999  
Patrocínio, Francisco J. G. 949  
Pauk, Jan Niklas 481  
Paulen, Radoslav 1001  
Paulo, Helena 1283  
Pedemonte, Martín 1105  
Pedroza-Medina, Uriel Roberto 2315  
Pellegrini, Laura A. 2507, 3243, 3257  
Peña-Torres, Daniel 2883  
Peng, Yuxing 3275  
Pennington, Oliver 2587  
Perdomo, Felipe A. 973  
Pereira, Cecile 313  
Perez-Cisneros, Eduardo S. 1093  
Pérez-Fortes, Mar 2427, 2451, 3151,  
3415, 3493  
Perez, Hector D. 1371  
Pessoa, Fernando L. P. 2921, 2945  
Peterson, Luisa 1487  
Petrescu, Letitia 2775  
Pfeiffer, Bernd-Markus 1191  
Phani Kumar, V. Sai 2015  
Phoochahan, Yuttana 387  
Pich, Andrij 807  
Pilario, Karl Ezra 1571  
Pillai, B. Basant Kumar 2975  
Pinelo, Manuel 2573  
Pinheiro, Carla C. 1257  
Pini, Ronny 2789  
Pinto, Jose M. 1601, 3317, 3355, 3385  
Pinto, Tiago 775  
Pinto-Varela, Tânia 1283, 3367  
Pishkari, Rojjar 343  
Pistikopoulos, Efstratios N. 51, 539,  
1227, 1709, 1839, 2029, 2629,  
2951  
Pitarch, José Luis 499  
Plantard-Wahl, Valentin 1985  
Plate, Christoph 561  
Plaza, Eric 411  
Polanska, Monika 2653  
Ponce-Ortega, José María 3293  
Ponce-Rocha, Jaime D. 843, 1093  
Pongboriboon, Nattapat 1907  
Pontes, Karen Valverde 3017  
Pornchuti, Bawornpong 387  
Posada, John 2457  
Posada, John A. 3053, 3391  
Potrè, Sanja 3421  
Potter, Claire 2715  
Pradhan, Snigdhendubala 2295  
Prado-Rubio, Oscar A. 2573, 2635  
Preisig, Heinz A. 2093, 2555, 3517  
Presser, Demian J. 3189  
Prieto, Carlos 2871  
Prifti, Kristiano 745, 2289, 2533  
Protoulis, Teo 1787  
Prousalis, Thomas 2813  
Przywara, Mateusz 1025  
Psycha, Melina 2561  
Pugnet, Veronique 313  
Pujan, Robert 2555  
Pulsipher, Joshua L. 1859  
Pyrgakis, Kostas 2561, 2567
- Q**  
Qian, Yu 3101  
Qi, Meng 2987  
Qin, Ruoshi 867

Qiu, Kang 1543  
 Qiu, Sizhe 2617  
 Qiu, Tong 305, 855  
 Quilló, Gustavo L. 1493  
 Quina, Margarida J. 1889  
 Quiroz-Ramírez, Juan José 145, 1317,  
 1529

**R**

Rabbani, Arash 1037  
 Rahma, Fadilla Noor 3225  
 Rajan Gopa, Kumar 1057, 1215, 1265,  
 2709  
 Rajendiran, Shenbageshwaran 2721  
 Ramirez, Andrea 1209, 3151, 3415  
 Ramirez, Andrea Ramirez 2427, 2451  
 Ramírez-Márquez, César 843  
  
 Ramos, Fernando 2073  
 Ramos, Matías 2073  
 Ramteke, Manojkumar 1883, 1901  
 Rani, Fatima 1685  
 Rani, Jyoti 1883, 1895  
 Rathi, Tushar 3355  
 Rato, Tiago J. 1889  
 Raval, Sanketkumar 2463  
 Ravutla, Suryateja 1311  
 Rawajfeh, Khaled 2623  
 Rebolledo-Leiva, Ricardo 2513  
 Recker, Sebastian 1173  
 Regueira, Alberte 2247  
 Reimschüssel, Sophie 125  
 Reina, Tomas Ramirez 203  
 Reinert, Christiane 2745, 2865  
 Reis, Marco S. 1889  
 Reis, Martina Costa 19  
 Reklaitis, Gintaras 1167, 1943  
 Reklaitis, Gintaras V. 813, 1437, 1577  
 Relvas, Susana 1283  
 Rengifo, Camilo 1147  
 Ren, Jiayang 1337  
 Repke, Jens-Uwe 119, 393, 595  
 Ressia, Jorge A. 2415  
 Reyes-Lúa, Adriana 1543  
 Riboldi, Luca 3011  
 Richard, Marc-André 1793  
 Rielly, Chris D. 1665  
 Rihko-Struckmann, Liisa 2391

Risco-Bravo, A. 519  
 Ritschel, Tobias K. S. 1775  
 Rittig, Jan G. 1173  
 Riyad, Mahmudul Hassan 2519  
 Roberts, Diarmid 2889  
 Robinson, Ada Josefina 619  
 Robles-Rodríguez, Carlos. E. 583  
 Roccaro, Ernesto 3243  
 Rodrigues, Diogo 3379  
 Rodriguez-Donis, Ivonne 1937  
 Rodríguez-Fernández, Alberto 3517  
 Rodríguez Garrido, Gabriel 2415  
 Rodriguez-Gomez, Divanery 1093  
 Rodriguez, José 1133  
 Rodríguez-López, Verónica 1093  
 Rodríguez-Zúñiga, Úrsula Fabiola 169  
 Roger, Kevin 709  
 Roger, Paul-Emeric 1931  
 Rogers, Alexander W. 83, 2593  
 Rojas-Casadieago, David 1305  
 Rolland, Thulasi 463  
 Romagnoli, J. A. 751, 1449, 1807  
 Román-Godínez, José Joel 2265  
 Román-Martínez, Alicia 2315  
 Romano, Matteo 3095  
 Romano, Matteo C. 2739  
 Romero-García, Ana-Gabriela 241  
 Romero-Izquierdo, Araceli G. 337,  
 2733  
 Romo-Bucheli, David 1305  
 Rosbo, Joachim Weel 1775  
 Roßmann, Volker 1751  
 Roudaut, Adrien 1931  
 Roufou, Styliani 2653  
 Roussanaly, Simon 3011  
 Rüdiger, Frank 325  
 Ruengrit, Suchada 387  
 Ruiz-Femenia, R. 3427  
 Rumayor, Marta 2277  
 Russo, Vincenzo 405, 715  
 Ryu, Yubin 493

**S**

Sabal Bermúdez, Jaime 1647  
 Saccardo, Alberto 2135  
 Sachio, Steven 1997, 2789  
 Sadeek, Sherard 2901  
 Sager, Sebastian 561

- Salah, Nora B. 3053  
Sala, Lorenzo 3095  
Salcedo-Diaz, R. 3427  
Salem, Rania 2301  
Sales-Cruz, Mauricio 3065  
Salmi, Tapio 405  
Salmon, Nicholas 2757  
Salucci, Emiliano 715  
Samadi, Amin 433  
Samouilidou, Maria E. 95  
Sánchez, Antonio 2871  
Sánchez-Gómez, Jahaziel Alberto 1913  
Sanchez Medina, Edgar Ivan 2035  
Sánchez-Ramírez, Eduardo 145, 241, 1317, 1529, 2733  
Sand, Guido 125  
Sandoval, Ilya Orson 33  
Sandru, Marius 2775, 3071, 3217  
Sansana, Joel 1889  
Santana, Leonardo O. S. 2921, 2945  
Santhamoorthy, Pooja Zen 1505  
Santos Bartolomé, Pedro 3487  
Santos, Gustavo S. 2921  
Santos, Magno Fonseca 2067  
Sanzida, Nahid 2519  
Saraiva, Samuel V. 1743  
Sarimveis, Haralambos 1787  
Sarkis, Miriam 3403  
Sarraf, Sahin 679  
Sarris, Iraklis 2567  
Satyakam, Anupam 3237  
Saulet, Pierre 3481  
Savage, Tom 511  
Sawaly, Razan 2177  
Saxén, Henrik 715  
Schäfer, Thomas 2227  
Scheid, Cláudia Mirian 1351  
Schenkendorf, René 837  
Schenzel, Karl-Wilhelm 215  
Schneider, Adrian 3203  
Scholl, Stephan 837  
Scholz, Benedikt X. 65  
Schuck, Carolin Eva 2227  
Schueler, Michael 2047  
Schulze, Jan C. 1173  
Schulze, Peter 2933  
Schweidtmann, Artur M. 2003, 2009, 3493  
Seanner, Florian 1851  
Secchi, Argimiro R. 1833  
Seckler, Marcelo Martins 3231  
Secondauthor, Tim B. 525  
Sedano, Luis 1801  
Seferlis, Panos 2807, 2813, 3109  
Seghers, Estelle E. 1807  
Segovia-Hernández, Juan Gabriel 145, 241, 1317, 1529  
Seidel-Morgenstern, Andreas 561, 1389, 2933  
Seidel, Tobias 463  
Selalame, Thabang W. 783  
Selejan, Alessandra-Diana 1365  
Selle, Kurt 2635  
Selvarajan, Subiksha 837  
Sempuga, Baraka C. 2801, 2825  
Sempuga, Celestin Baraka 1401  
Semrau, Robin 1127  
Sen, Amrita 2481  
Sener, Aysegul 1871  
Seongbin, G. 881  
Sergent, Lionel 589  
Serna, Juliana 1985  
Sforza, Eleonora 2135  
Shahab, Mohammed Aatif 2041, 3511  
Shahbaz, Muhammad 1473  
Shah, Nilay 381, 649, 697, 1653, 2215, 2433, 3403  
Shajahan, Thasnimol Valuthottiyil 2041  
Shamaki, Patience B. 2079  
Shanbhag, Shachi M. 2003  
Shao, Zhijiang 1671  
Sharma, Shivom 1063, 3203  
Shastri, Yogendra 2469, 3237  
Shcherbakova, Nataliya 709  
Shehab, Moaaz 2087  
Shihada, Rawan 2301  
Shijaz, Haneef 2457  
Shi, Kenian 967  
Shimono, Ryoko 2361  
Shin, Daeun 691  
Shin, Dongil 997  
Shin, Sunkyu 493  
Shirazi, Ahad Sarraf 3249  
Shoji, Yasuhiro 1481  
Shokry, A. 1703

- Short, Michael 197, 203, 209, 3361, 3397  
 Sievers, Carsten 2403  
 Siirola, John D. 1329  
 Sikarwar, Shailesh Singh 2999  
 Sildir, Hasan 679  
 Silva, Ana Carolina Borges 2271  
 Silva, Carlos A. M. 1743  
 Silva, e Flávio V. 1743  
 Simon, Assaf 3475  
 Simon, Sapir Lifshiz 3475  
 Singh, Suraj Prakash 1291  
 Singtothong, Pravitt 387  
 Sin, Gürkan 291, 399, 575, 1221  
 Siniscalchi, Amira 355  
 Sixon, David 813  
 Sizaire, Paul 1139  
 Skiadas, Ioannis V. 2605  
 Skiborowski, Mirko 331, 463, 955  
 Skogestad, Sigurd 613, 1697  
 Slavu, Nela 3071, 3217  
 Slegers, Petronella M. (Ellen) 3493  
 Smith, Chloe 2751  
 Smith, Robin 763, 1031, 2993  
 Sochard, Sabine 3133  
 Sodiq, Ahmed 1407  
 Soh, Qiao Yan 697  
 Somoza-Tornos, Ana 2289, 3493  
 Song, Hyeju 2695  
 Song, Zeyuan 819  
 Son, Jinkyung 997  
 Sorensen, Eva 235, 1919, 3463, 3499  
 Souza, Tiago F. 1821, 1827  
 Sow, Mame Cheikh 1931  
 Spagna, Mario 2547  
 Spatolisano, Elvira 2507, 3243  
 Srinivasan, Babji 1729, 2041, 3511  
 Srinivasan, Prahalad 2183  
 Srinivasan, Rajagopalan 1291, 1729, 2041, 3511  
 Steimel, Jochen 601  
 Steinberger-Wilckens, Robert 3169  
 Steiner, Philipp 2487  
 Stepchuk, Inna 2427  
 Stephanopoulos, George 2481  
 Steven, Robert A. 3397  
 Stevnsborg, Mads 2573, 2635  
 Stinchfield, Georgia 1329  
 Stocks, Stuart M. 1221  
 Stossi, Fabio 2629  
 Straathof, Adrie 2457  
 Straathof, Adrie J. J. 2111  
 Straub, Adrian 2963  
 Strelet, Eugeniu 1889  
 Stylianopoulou, K. G. 2129  
 Suarez Castaneda, José Luis 1793  
 Subramanian, Avinash 267  
 Subraveti, Sai Gokul 3011  
 Sugiyama, Hirokazu 65, 985  
 Su, Jie-Ying 451  
 Sumankumar, Agarwal Shruti 2981  
 Sundarkumar, Varun 1943  
 Sundmacher, Kai 991, 1245, 1487, 2035, 2391  
 Sunghyun, C. 881  
 Sun, Wei 131, 1549, 1627  
 Supankanok, Rasa 1907  
 Suryavanshi, Atharva V. 1511  
 Swartz, Christopher L. E. 1865  
 Swinkels, Pieter L. J. 2499, 3493  
 Symes, Andrew 2343  
 Syu, Cian-Shan 451  
 Szafran, Adam T. 2629  
 Szymanski, Pierre 2851
- T**
- Tabak, Alihan 685  
 Taşçı Çilak, Tutku 2463  
 Taguchi, Tomoyui 1973  
 Taifan, Garry S. P. 445  
 Taifouris, Manuel 2197  
 Talis, Torben 595  
 Tan, Benjamin I. 1153  
 Tan, Michael D. 1209  
 Tan, Raymond R. 3263  
 Tao, Min 1431  
 Tappe, Aike Aline 837  
 Tauqir, Wajeha 733  
 Tenneson, Albane 1931  
 Tesfamichael, Brook 3305  
 Testa, Romina Lasry 2073  
 Thakker, Vyom 2481  
 Theodoropoulos, Constantinos 1037, 1045  
 Theodoropoulou, Artemis 285  
 Thiebaud-Roux, Sophie 1937

- Thirdauthora, James Q. 525  
 Thomsen, Mette H. 2463  
 Thota, Naresh 1013  
 Thyagarajan, Karthik 3385  
 Tian, Yuhe 2029, 3157  
 Tiggeloven, Julia L. 3077  
 Tillmann, Maerthe Theresa 631  
 Tinoco-Caicedo, D. L. 519, 2321  
 Tirapelle, Monica 3463  
 Tiso, Till 2683  
 Todd, Lucy 2703  
 Tom, Matthew 71  
 Topac, Emir 679  
 Torres, Ana I. 247, 1105  
 Torres-Magaña, María Fernanda  
     2373  
 Trajano, Heather 1337  
 Tran, Khanh-Quang 3225  
 Triantafyllou, Niki 381, 1653  
 Tripura, Tapas 1895  
 Trokanas, Nikolaos 3439  
 Trokanas, Nikos 2677  
 Tronci, Stefania 751  
 Tsatse, A. 3499  
 Tsay, Calvin 1641, 1647  
 Tschentscher, Roman 757, 769  
 Tsimopoulos, Evangelos G. 2763  
 Tula, Anjan K. 475  
 Turan, Evren Mert 1763  
 Turati, Daniele 163
- U**
- Udugama, Isuru A. 2379  
 Ukawa, Chinatsu 1679  
 Urbas, Leon 1057, 1215, 1265, 1685,  
     2055, 2061, 2709  
 Usai, Alessandro 1045  
 Uzunlar, Erdal 3089
- V**
- Vaes, Julien 1159  
 Valappil, Jaleel 3181  
 Valdramidis, Vasilis 2653  
 Valentin, Claire 703  
 van de Berg, Damien 649  
 Van der Hauwaert, Lucas 2247  
 Van Gerven, Tom 3281, 3487  
 Van Herle, Jan 1063  
 Van Impe, Jan F. M. 1069, 1493, 1769,  
     2653  
 Varbanov, Petar 3421  
 Varela, Christopher 2915  
 Várnai, Anikó 757, 769  
 Vasilaki, Michaela 553  
 Vasilakis, Pantelis 2561  
 Vasileiadis, Manolis 3439, 3445  
 Vasileiadou, Athina 637  
 Vázquez, Daniel 1239  
 Vázquez-Hernández, Ariadna E. 1345  
 Vedant, Shivam 2951  
 Vedin, A. Eren 685, 1871  
 Vega, Lourdes F. 3163  
 Vega, Pastora 1967  
 Vega-Ramon, Fernando 157  
 Venkatarathnam, G. 3121  
 Venkatasubramanian, Venkat 221,  
     1087  
 Verheyleweghen, Adriaen 1543  
 Vermeire, Florence 457  
 Vernuccio, Sergio 1075  
 Vertis, Carolina S. 1019  
 Viedt, Isabell 1057, 1215, 1265  
 Vieira, Miguel 1283  
 Villarreal-de-Aquino, Maria de los A.  
     1093  
 Vincotte, Yohan 3209  
 Vlachos, Dionisios G. 2183  
 Voldsund, Mari 267  
 Volpe, Maria A. 2415  
 von der Aßen, Niklas 2745, 2865  
 Vooradi, Ramsagar 2975, 2999  
 Vos, Josephine 3151  
 Vrana, Til Kristian 1543
- W**
- Waldherr, Steffen 3281  
 Walker, George F. 2629  
 Wallrath, Roderich 1851  
 Walmsley, Timothy G. 2379  
 Wammes, Puck 3391  
 Wanderley, Ricardo R. 3257  
 Wang, Bei 2617  
 Wang, Haiting 469  
 Wang, Henrik 71  
 Wang, Jingdai 2957  
 Wang, Jingde 1549, 1627



- Wang, Kexin 1671  
 Wang, Liang 2165  
 Wang, Wanning 1943  
 Wang, Wei 157  
 Wang, Yang 2165  
 Wang, Yifan 2865  
 Wang, Zhaoxing 2183  
 Wang, Zihao 1245  
 Ward, Adam 2789  
 Ward, Keeran 2877, 2901  
 Ward, Oliver 3269  
 Watanabe, Toshiyuki 1973  
 Weber, Mirko 267  
 Wehbe, Nader 1931  
 Wei, Xinyi 1063  
 Wentink, Bart 3493  
 Wenzel, Tom 2709  
 Wibbing, Daniel 487  
 Wilisch-Neumann, Annette 2221  
 Wilkes, Mathew Dennis 2889, 2895  
 Wiltink, Thijmen 3415  
 Wittgens, Bernd 757, 769  
 Wolter, Nadja 807  
 Wong, David Shan-Hill 297  
 Won, Wangyun 2239  
 Wu, Chuan-Yu 57  
 Wu, Deyang 1535  
 Wu, Guoquan 795  
 Wu, H. 2499  
 Wu, Min 457, 3281  
 Wu, Wei 157, 1907
- X**
- Xenios, Stefanos 2727  
 Xie, Jiahan 1563  
 Xiouras, Christos 1493  
 Xue, Ying 3433  
 Xu, Hao Yang 3011  
 Xu, Shuang 475
- Y**
- Ya'akobi, Asia Matatyaho 3475  
 Yamashita, Yoshiyuki 1679, 1973  
 Yang, Aidong 1383, 2617  
 Yang, David 1589  
 Yang, Haoyu 2003  
 Yang, Nan-Hua Nadja 1383  
 Yang, Siyu 943, 967, 3101  
 Yang, Xi 733  
 Yang, Yao 2957  
 Yang, Yongrong 2957  
 Yang, Zhijie 2617  
 Yao, Yuan 57, 297  
 Yaqot, Mohammed 1111, 1757, 1781  
 Yasmal, Asli 3089  
 Yoon, Ha-Jun 3127  
 Yoon, Hayoung 2385  
 Yoon, Sungho 2385  
 Yoon, Taeksang 2385  
 Yoo, Yup 1413  
 You, Fengqi 1563, 1659, 2203, 2209,  
 2837, 2845  
 Young, Brent R. 2379  
 Young, Nathan 2721  
 Yousefi, Negin 2665  
 Yska, Stijn 3415  
 Yuan, Xuming 255  
 Yuan, Zhihong 247, 921  
 Yuki Junior, Giberto M. 3133  
 Yu, Luxi 2659  
 Yun, Sungil 71  
 Yup, Y. 881  
 Yusuf, Noor 1613
- Z**
- Zaabout, Abdelghafour 313  
 Zafar, Umair 553  
 Zahel, Thomas 909  
 Zahid, Umer 1991, 2493  
 Zamarripa, Miguel 1329  
 Zarei, Mohammadamin 3127  
 Zaryab, Syed Ali 1179, 3095  
 Zeballos, Luis J. 3317  
 Zeng, Hong 2617  
 Zhai, Jianyuan 1311  
 Zhang, Chen 855  
 Zhang, Dongda 33, 83, 157, 1523,  
 2587, 2593  
 Zhang, Duo 1671  
 Zhang, Jiaqi 547  
 Zhang, Lifeng 247  
 Zhang, Nan 763, 2021  
 Zhang, Nuobei 417  
 Zhang, P. 2795  
 Zhang, Qi 1357, 3195, 3355  
 Zhang, Ruiyu 189

- Zhang, Ruosi 197, 209  
Zhang, Wei 1383  
Zhang, Xiang 991  
Zhang, Xiangping 361  
Zhao, Fei 1735  
Zhao, Jinsong 867, 1535, 1715  
Zhao, Liang 935  
Zhao, Liwen 2859  
Zhao, Xiang 2209  
Zhao, Yi 929
- Zheng, Jingyuan 967  
Zheng, Taicheng 261  
Zhou, Teng 1245  
Zhu, Chenguang 2021  
Zhu, Lingyu 1735  
Ziegler, Anita L. 2683  
Zomer, Simeone 2099  
Zondervan, Edwin 2087, 2307, 2915,  
3493  
Zotică, Cristina 1543

